IRis

Northeastern University

Economics Dissertations

Department of Economics

January 01, 2011

Three essays in empirical industrial organization

Daniel Greenfield
Northeastern University

Recommended Citation

Greenfield, Daniel, "Three essays in empirical industrial organization" (2011). *Economics Dissertations.* Paper 5. http://hdl.handle.net/2047/d20001086

This work is available open access, hosted by Northeastern University.

THREE ESSAYS IN EMPIRICAL INDUSTRIAL ORGANIZATION

A dissertation presented

by

Daniel Greenfield

to
The Department of Economics

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in the field of

Economics

Northeastern University
Boston, Massachusetts
June, 2011

THREE ESSAYS IN EMPIRICAL INDUSTRIAL ORGANIZATION

by

Daniel Greenfield

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Graduate School of Arts and Sciences of
Northeastern University, June, 2011

Abstract

The first chapter is titled "Competition and Service Quality: New Evidence from the Airline Industry." This paper examines the relationship between competition and service quality in the airline industry. Using panel data and two different instrumental variable (IV) strategies, it directly addresses potential biases in previous studies. One of the IV strategies exploits changes in local market structure generated by a merger between two airlines that operate on a global scale. While this variation in market structure does not occur in a controlled environment, its source, if not random, differs from the variation in market structure otherwise observed in the airline industry. The findings indicate the effect of competition on airline delays may be three times stronger than previous studies suggest. The validity of the merger as a source of exogenous variation is examined along several lines.

The second chapter is titled "Evidence on the Product Market Effects of the Merger between Delta and Northwest." This paper employs a difference-in-differences model to estimate the price effects of the merger between Delta Air Lines and Northwest Airlines on October 29, 2008. The transaction combined two of the largest air carriers in the U.S. The analysis focuses on markets where the change in horizontal market structure was substantial enough to raise competitive concerns under the horizontal merger guidelines of the U.S. Department of Justice and the Federal Trade Commission. While the focused nature of the analysis does not offer a complete assessment of the merger's competitive effects, the results are noteworthy in that they contrast earlier studies of airline mergers that have found anticompetitive price increases. The empirical analysis highlights the importance of not only accounting for shocks to supply and demand, but also accounting for differences in the way those shocks are

transmitted to prices across treatment and control markets.

The third chapter is titled "The Cost Structure of Regional Transmission Organizations," a paper co-authored with John Kwoka. RTOs now cover well over one-half of customers and sales of electricity in the U.S. As they have expanded in geographic coverage and functional scope, controversies have arisen about their rapidly growing costs and overall effectiveness. We model their costs, exploiting the fact that the seven existing RTOs initiated their various functions at different points in time over their roughly ten-year history. Specifically, we investigate the costs of each of the market functions administered by RTOs, the relative costliness of different RTOs in performing the same functions, the possibility of learning economies from either their individual or industry-wide experience, and economies of scale over the sampled range of RTO "output." Our results confirm the importance of some of these factors but raise doubts about others, in all cases based on systematic modeling and data analysis.

Acknowledgements

I am deeply indebted to my dissertation committee co-chairs, James Dana and John Kwoka, for their constant support and guidance. I would also like to thank committee members Steven Morrison and Zhongmin Wang, whom were instrumental in my development as a researcher. Gratitude is also owed to Kamran Dadkhah, Bill Dickens, Jordi Jaumandreu, Kai-Uwe Kuhn, Maria Luengo-Prado and Jackson Osborne for helpful comments and discussion on various parts of this dissertation. I would also like to express sincere thanks to Neil Alper, Kathy Downey, Cheryl Fonville, and Gregory Wassall for their help while navigating graduate school. Finally, my deepest appreciation goes to my family—particularly my parents, for their encouragement and advice, and my wife, for being the consummate teammate and partner in all aspects of life.

TABLE OF CONTENTS

Abstract 2

Acknowledgements 5

Table of Contents 6

Chapter 1 - Competition and Service Quality: New Evidence from the Airline Industry
- Introduction 8
- Airline Delays 11
- Data 12
- Empirical Specification 18
- Results 22
- Discussion 28
- Conclusions 29

Chapter 2 - The Product Market Effects of the Merger between Delta and Northwest
- Introduction and Literature Review 42
- Data and Model 46
- Results 54
- Conclusions 59

Chapter 3 - The Cost Structure of Regional Transmission Organizations
- Introduction 75
- Background on Regional Transmission Organizations 76
- Empirical Model 82
- Data 85
- Results and Discussion 88
- Conclusions 97

Chapter 1

Competition and Service Quality: New Evidence from the Airline Industry

1. Introduction

Economists traditionally assess market power by measuring the difference between price and marginal cost. Standard economic theory explains that in the absence of competition firms will increase price above cost, reducing social welfare and transferring rents from consumers to producers. Yet the relationship between competition and product quality is unclear. Economic theory predicts competition can cause quality to either increase or decrease.[1] Given the uncertainty, empirical research can help inform theory, particularly in specific industry settings.

To be sure, the relationship between competition and product quality has important policy implications as well. For services such as health care, legal counsel, and education, quality may rival price as a determinant of demand and consumer welfare. Despite this importance, empirical research on the relationship between competition and product quality is still relatively sparse.[2] Among the exceptions are papers by Mazzeo (2003) and Rupp, Owens, and Plumly (2006), each of which finds that the incidence and duration of airline flight delays are greater on routes with fewer competitors. Flight delays are generally considered an important measure of service quality in the airline industry. A recent study commissioned by the FAA reported the total cost of US air transportation delays in 2007 was $32.9 billion, including $3.9 billion of lost consumer surplus incurred by consumers who avoided flying because of

[1] See for example, Swan (1970), Spence (1975), Mussa and Rosen (1978), Klein and Leffler (1981), Hörner (2002), Kranton (2003), Bar-Isaac (2005), Aghion et al. (2005) and Dana and Fong (2010). Gaynor (2006) provides a comprehensive survey.

[2] See Gaynor (2006) for a survey on healthcare markets, Domberger and Sherr (1989) on legal services, Hoxby (2000) on public schools, and Matsa (2010) on the supermarket industry.

delays and $16.7 billion directly borne by airline passengers.[3]

This study extends the literature on competition and airline delays with an emphasis on establishing a causal relationship. Using panel data and two IV strategies, it directly addresses potential biases in previous studies. The hypothesis is that the relative incidence and duration of flight delays decreases as market structure becomes relatively more competitive. Competition makes poor service quality costly in terms of reduced future demand, thus, carriers exert effort to minimize delays on more competitive routes.

An OLS regression of flight delays on market structure is problematic, as there are two specific reasons to expect that market structure is endogenous. First, an airline's level of service quality affects demand for its own and competing carriers' service. Thus, market structure and quality are codetermined. Second, an airline's profit maximizing choice of service quality is likely a function of demand and cost factors which also affect market structure. Any such factors that are unobserved or poorly measured by the researcher will lead to correlation between market structure and the error term. Additionally, given market structure is not randomly assigned, there is an ever-present concern that some exogenous determinant of delays may confound the relationship.

To address this final concern, prior studies have relied on variation in market structure across routes at the same airport. This approach effectively controls for airport effects such as congestion and weather. However, as Mazzeo (2003) mentions, it leaves open the possibility "that there is some unobserved factor explaining delays on particular routes that is correlated

[3] The estimated $16.7 billion borne by passengers was calculated based on lost time due to flight delays, cancellations and missed connections, plus expenses incurred while being away from home for additional time. See, http://its.berkeley.edu/sites/default/files/NEXTOR_TDI_Report_Final_October_2010.pdf

with the propensity for airlines to offer service on that route." This study uses variation in competition and on-time performance within routes over time. Carrier-route fixed effects control for potentially confounding differences across routes. Then, two distinct instruments are employed to address the other aforementioned endogeneity concerns.

The first instrument strategy is to use lagged market structure as an instrument for current market structure. This approach is analogous to that taken by Evans, Froeb, and Werden (1993) and Davis (2004) while studying the effect of market structure on prices. Using lagged values of market structure as an instrument is a logical way to address simultaneity bias and also softens the threat of omitted variable bias. The second IV strategy relies on variation in competition caused by the 2008 merger between Delta Air Lines (DL) and Northwest Airlines (NW). In the spirit of Aghion et al. (2005) and closely following Dafny, Duggan, and Ramanarayanan (2010), the merger is used as a natural experiment. In the first stage regression, post-merger market structure is predicted as a function of the premerger market shares of DL and NW. While merger partners are not randomly assigned, the resulting variation in local market structure certainly arises from a different mechanism than would otherwise be observed in the data, and is not likely subject to the same endogeneity concerns.

The OLS estimates are similar in magnitude to those found by Mazzeo (2003) and support the hypothesis that competition positively effects on-time performance (Rupp, Owens, and Plumy (2006), use slightly different measures of competition to yield qualitatively similar results). More interestingly, the IV estimates indicate that the impact of competition on flight delays is roughly three times stronger than predicted by the OLS models.

The remainder of the paper is organized as follows: Section 2 provides some background

on airline delays and discusses the related literature. Section 3 describes the data set. Section 4 details the empirical specification. Section 5 presents the results, while 6 discusses their implications. Section 7 concludes.

2. Airline Delays

Airlines can exert effort to prevent delays. Employing extra labor and aircraft can mitigate delays arising from unscheduled maintenance, crew problems, fueling, aircraft cleaning, and baggage handling. Moreover, when an imminent delay arises, airlines can influence which routes are affected. For example, airplanes, as well as flight and maintenance personnel, can be redeployed to routes where the costs of delay are highest. One of the reasons delays are costly to an airline is because they can reduce future demand. In a more competitive environment, where customers can choose among additional carriers, the cost of a flight delay is likely greater. Conversely, the adverse reputational effects of delays are presumably least costly in a monopoly market.

It is also true that many flight delays arise due to exogenous events outside an airline's control, such as congestion and weather. Nonetheless, airlines can again influence which flights are most affected by such events. When faced with schedule disruptions, airlines may be forced to choose which flights are delayed. All else equal, the cost of a delay is likely greater on more competitive routes, where repeat customers have greater opportunities to switch carriers.

In both of the aforementioned examples, factors of production are employed across routes in the short-run, after route networks and flight schedules have been set. The correlation between on-time performance and competition documented by Mazzeo (2003) and Rupp,

Owens, and Plumly (2006) is consistent with this hypothesis. If the relationship between quality and future demand is driving the association between competition and on-time performance, delays should have a negative effect on demand. Prior research indicates that flight delays do reduce airline demand and market fares. Morrison and Winston (1989) estimate a logit demand model and find that a one percentage point increase in the share of flights delayed more than 15 minutes reduces passengers' willingness-to-pay by $2.60 per round-trip fare.[4] Consistent with those findings, Forbes (2008a) finds that fares fall by $2.43 with each percentage point increase in the share of flights delayed.[5] Forbes exploits a legislative change at LaGuardia Airport which allowed new entry on a subset of routes, creating an exogenous change in congestion on all routes at LaGuardia.[6] Suzuki (2000) also offers support for the underlying theory. He constructs a model in which a traveler's choice of airline at time t depends on whether they experienced a flight delay at $t - 1$. Using data from a specific city pair market, he finds that fluctuations in market share can be explained by past on-time performance.

3. Data

3.1. Data sources

[4] Forbes (2008b) finds the number of consumer complaints increases with the percentage of flight delays. Using an ARIMA model, Foreman and Shea (1999) find that publication of on-time performance data for U.S. domestic airline operations was associated with increased quality distribution, enhanced demand, and reduced price-quality rigidity.

[5] Dollar values from earlier studies have been inflated to 2008 dollars using the CPI. Forbes uses one-way fares which are multiplied by two.

[6] Other papers using airline on-time performance data include Mayer and Sinai (2003) and Rupp (2009) who examine whether airlines internalize congestion at hub airports. Rupp and Holmes (2006) investigate the determinants of flight cancelations. Lee and Rupp (2007) utilize on-time performance as a proxy for employee effort to assess how airline pilots respond to wage reductions. Ciliberto and Schenone (2010) investigate the relationship between bankruptcy and several measures of airline quality, including on-time performance. Finally, Lederman and Forbes (2010) measure the effects of vertical integration on firm performance using several measures of on-time performance.

The two primary sources of airline data are the U.S. Department of Transportation's (DOT) Airline On-Time Performance dataset and Airline Origin and Destination Survey (DB1B) dataset. The DB1B is a ten percent sample of airline tickets sold in the U.S. These ticket data provide detailed itinerary information including the fare, total passengers, origin and destination airports, and all intermediate stops on each directional trip. The quarterly ticket data are aggregated to provide prices and quantities for each carrier in each market.

The Airline On-Time Performance Data covers all non-stop domestic flights by carriers with one percent or more of total domestic revenues, plus two additional carriers that report voluntarily.[7] These data are used to construct variables for the duration and incidence of delays and the total number of flights at each airport in the sample. The DOT data are supplemented with weather data from National Oceanic and Atmospheric Administration's Climatic Data Center (NCDC).[8] Demographic data from the Bureau of Economic Analysis (BEA) and the U.S. Census Bureau are also used to control for changes in local-market personal income and population.

3.2. Data and variable definitions

Carrier. The sample consists of major legacy and low-cost carriers.[9] All regional and commuter airlines are excluded, as many of their flights are operated under codeshare agreements. As explained by Lederman and Forbes (2010) the incentive and ability of regional

[7] Carriers are required to report on operations to and from 29 large airports, however, all reporting carriers have voluntarily provided data for their entire domestic systems.

[8] Data on airplane characteristics from the Federal Aviation Administration's (FAA) aircraft registry were also used in several unreported regressions.

[9] The carriers in the sample are American, Alaska, JetBlue, Continental, Delta, Frontier, AirTran, Northwest, United, US Airways, and Southwest.

operating carriers to respond to schedule disruptions is likely different than traditional carriers, particularly when the partners are not commonly owned. While the principal results are not sensitive to this specification choice, it also facilitates direct comparisons with prior research. Carriers are indexed by *i*.

Route. A route is an ordered pair of airports on which one or more carriers operate nonstop flights. Thus, on a particular route, carriers fly directional non-stop flights between two specific airports. For example a flight from Boston Logan (BOS) to Los Angeles International (LAX) is flown on a different route than a flight from LAX to BOS. Routes are indexed by *r*.

Market. A market is an unordered pair of airports where carriers offer airline service, either on a direct or connecting basis.[10] Both one-way and round-trip tickets are included in the market. While markets are sometimes defined more narrowly, it is common practice to include non-stop and connecting service in the same market. A passenger that experiences poor on-time performance on a non-stop flight—or expects to based on an airline's reputation—may choose a competing carrier that offers connecting service. Markets are indexed by *m*.[11]

Another reason markets are defined to include connecting traffic in the baseline specifications is because the merger between DL and NW had limited effect on markets that are defined to include only direct passengers. Nonetheless, in secondary regressions, the measure of market structure is redefined using only direct traffic.

Unit of observation. The unit of observation in the final estimation data is indexed by

[10] Connecting service refers to airline service that involves an intermediate stop, typically at a hub airport, before reaching the final destination of a one-way or directional trip. Tickets with more than one stop are eliminated from the data set.

[11] The results are not sensitive to defining markets as airport-pairs.

irmt. The dependent variable in the regressions represents the aggregate on-time performance of flights operated by carrier *i* on a route *r* during year-quarter *t*. The explanatory variable of interest is the level of competition in market *m* during period *t*, where market *m* corresponds to both routes *r* that have the same end-point airports as market *m*. The indexing subscript *m* is included whenever a variable varies over *r* even though all routes map to only one market.

On-time Performance. Two measures of on-time performance are employed: one to capture the prevalence of delays and the other to capture the average duration of delays. $MINS_{irmt}$ is the mean discrepancy, in minutes, between the scheduled arrival time and actual arrival time of flights operated by carrier *i* on route *r* during quarter *t*. Note this measure counts early arrivals as negative delays. The second measure of delays, $ODDS_{irmt}$, is the log odds of being delayed. Let P_{irmt} be the fraction of flights operated by carrier *i* on route *r* during period *t* that are delayed upon arrival by 15 minutes or more, plus .01 (to ensure the logarithm is defined). Then, $ODDS_{irmt} = \ln\left((P_{irmt})/(1 - P_{irmt})\right)$.

Market Structure. The primary measure of market structure is the Herfindahl-Hirschman Index (HHI). The variable HHI_{mt} is the sum of the squared market shares in market *m* during quarter *t*. It is constructed from the DB1B data using both direct and connecting passengers that make a single change of plane.[12] Market shares are calculated using revenue, not passengers. As an alternative, the variable $HHID_{mt}$ is defined analogously using only direct traffic.

Control variables. All specifications include demographic variables to control for

[12] Tickets with more than one connection are dropped from the dataset.

changes in demand. These include the sum of the origin and destination personal income, $INCOME_{mt}$, and the sum of the origin and destination population, POP_{mt}, both in log form. $INCOME_{mt}$ comes from the BEA Quarterly State Personal Income series and POP_{mt} comes from the U.S. Census Bureau. These variables are measured at the state level. The sign of the coefficients on these variables is a priori ambiguous. Absent a specific theoretical model, the effect of a ceteris paribus change in demand on a carrier's choice of service quality is unclear. On the other hand, increased demand, holding flight schedules constant, may result in greater congestion related delays.

In all regressions, the expectation of on-time performance is also conditioned on observable weather characteristics at the route endpoints. $RAINORG_{rmt}$ and $RAINDST_{rmt}$ equal the number of days with greater than or equal to 1 inches of precipitation at the origin and destination airports of route r during period t. The impact of snowfall is captured with $SNOWORG_{rmt}$ and $SNOWDST_{rmt}$, which represents total snow fall at the route endpoints measured in feet. All of the weather controls are expected to be positively correlated with delays.

There are also variables that control for differences in airport traffic and flight schedules. $TIME_{irmt}$ is the average scheduled arrival time of the flights operated by carrier i on route r during time period t.[13] Prior research indicates, as one would expect, that flights scheduled later in the day have worse on-time performance. $ORGFGHT_{rmt}$ and $DSTFGHT_{rmt}$ are respectively the total number of departures and arrivals at the origin and destination airports of route r during

[13] $TIME_{irt}$ is scaled from 0 to 1 with 0 = 12:01 AM and 1 = midnight.

period *t* in natural logs. These variables are specified to capture changes in airport traffic. The coefficients on these variables are expected to be positive, as greater airport traffic leads to additional congestion delays.

3.3. Description of estimation sample

The estimation dataset spans 2006:Q1 through 2010:Q3. This timeframe provides more than 2 years of data prior to the announcement of the merger, and 2 years of data after the transaction was finalized.[14] After constructing the variable $ORGFGHT_{rmt}$ and $DSTFGHT_{rmt}$, the on-time performance data is restricted to flights between the top-100 airports in terms of total flights, over the entire sample. Then carrier-route-year-quarter combinations with fewer than 24 flights (approximately 2 flights per week) are eliminated. The on-time performance data is then matched to the market data constructed from the DB1B survey. The appendix describes how the DB1B data is cleaned prior to matching with the on-time data. The final sample includes 1369 airport-pair markets and 3981 carrier-route combinations.

In addition to standard summary statistics, table 1 includes the minimum, maximum, and standard deviation of each variable after time-demeaning by carrier-route. These statistics describe the variation that occurs over time, holding the carrier-route effects constant. Section 4.1, which outlines the merger instrument strategy, provides a more detailed description of the consolidation caused by the merger between DL and NW.

[14] When the sample is restricted to 2007 through 2009, the results are qualitatively similar. The estimated relationship between competition and on-time performance is in fact stronger.

4. Empirical Specification

On-time performance is modeled as a function of competition, controlling for scheduled arrival times, airport characteristics, endpoint weather conditions, and exogenous demand and cost shifters. Rather than specifying an explicit theoretical model, the regression equation is assumed to arise from a broad conceptual framework where airlines simultaneously choose price and quality after entering markets and setting flight schedules. Instead of including price in the regression, which would introduce another endogenous regressor, demand and cost factors are included to control for exogenous determinants of both price and quality.[15] The regression equation takes the form:

$$Y_{irmt} = \alpha + \beta\, HHI_{mt} + X_{irmt}\gamma + \delta_t + \varphi_{irm} + \varepsilon_{irmt},$$

where Y_{irmt} is either $MINS_{irmt}$ or $ODDS_{irmt}$. X_{irmt} is the vector of control variables, δ_t represents the year-quarter fixed effects, and φ_{irm} represents the carrier-route fixed effects. The year-quarter effects capture industry-wide changes in demand and costs, such as fluctuations in the business cycle or fuel prices. The carrier-route effects hold constant time-invariant factors including those particular to a carrier's flights on a given route.[16] These include carrier-specific hub effects and other airport characteristics, which may be difficult to measure, and are not of particular interest here. It is also important to note that the directional nature of the route effects allows for unique origin and destination airport effects, as well as differences in geography or

[15] Gaynor (2006) suggests this approach. Following a similar logic, carrier output, market shares, and load factors are excluded from the regression equation.

[16] Throughout 2009, after its merger with Delta, Northwest's planes continued to operate under the airline code NW. For routes served by both DL and NW, once market structure variables are appropriately defined, DL and NW flights are combined into a single observation both before and after the merger.

weather patterns.

In all regressions the standard errors are clustered by market. While the carrier-route effects capture all time-invariant group error components, the clustered variance estimator accounts for serial correlation in the remaining error term. Clustering at the market level also accounts for correlation due to time-varying error components that are common across carrier-routes within a specific market—e.g. unobserved weather shocks or measurement error in the market structure variable. Finally, in all regressions, observations are weighted by the number of flights.

4.1. Instruments

Two different instruments are employed to address the potential endogeneity of HHI. The first instrument is lagged HHI. Using lagged values of HHI is an obvious approach to circumvent simultaneity bias and also reduces the likelihood of omitted variable bias. While unobserved determinants of delays may be persistent over time, current unobserved cost or demand shocks are certainly less likely to be correlated with lagged values of HHI than with current HHI. Following this reasoning, a six-quarter lag of HHI is employed to balance relevance and exogeneity. Exploratory IV regressions were run using lags ranging from one to eight quarters as the sole excluded instrument. While all of these estimators produce a coefficient on HHI larger than the OLS estimator, the coefficient on HHI gradually increases as higher order lags are employed. Lags of less than four quarters appear to be endogenous.[17]

[17] While the coefficient on HHI in the IV estimator is larger than the OLS estimator, when using a lag of less than four quarters, we cannot reject statistical equivalence between the just identified IV estimator and the OLS

However, lags of an order greater than six provide weak correlation in the first stage regression (F-statistic less than 10, see Staiger and Stock (1997)).

The second instrument predicts changes in market structure using the merger between DL and NW. The carriers announced their intention to merge in April of 2008, and finalized the transaction after receiving antitrust approval at the end of October of 2008. While DL and NW offered overlapping non-stop service in a limited number of airport-pair markets, the merger generated a change in horizontal market structure in hundreds of markets where both carriers provided either nonstop or connecting service.

During ex ante merger analysis, the predicted change in HHI is often approximated as the change which would occur if the transaction combined the market shares of the merging firms, leaving rivals' market shares unchanged. Following Dafny, Duggan, and Ramanarayanan (2010), this notion of an all else equal predicted change in HHI is used to construct the excluded instrument. For each market *m*, let DL_m and NW_m, be the market shares of DL and NW during 2007. Then, the predicted change in HHI is:

$$\Delta_m = (DL_m + NW_m)^2 - DL_m^2 - NW_m^2 = 2 * DL_m * NW_m.$$

Figure 1 illustrates the distribution of Δ_m.[18] The excluded instrument, $CHGHHI_{mt}$, is the

---

estimator. Additionally, tests of the overidentifying restrictions, following 2SLS estimations that combine lagged HHI with the second instrument described below, suggest that lags of an order less than four are endogenous.

[18] $\Delta_m > .01$ in 213 of the 1369 airport-pair markets in the estimation sample. 191 of these markets meet the criteria of having "the potential to raise significant competitive concerns" under 2010 U.S. Department of Justice/Federal Trade Commission horizontal merger guidelines ($\Delta_m > .01$ and post-merger HHI > .25). Another 667 markets in the sample were affected by the merger; however, the expected change in HHI was less than .01. $\Delta_m = 0$ in 489 markets.

interaction of Δ_m and a dummy variable that equals zero prior to 2008:Q4 and equals one in 2008:Q4 and all later periods.[19]

Given its structure, it seems reasonable to assume $CHGHHI_{mt}$ is uncorrelated with unobserved demand and cost shocks. Moreover, there is no reason to believe it is simultaneously determined with on-time performance. Thus, it does not appear subject to the same endogeneity concerns as HHI_{mt}. However, the excluded instrument is valid only if it has no effect on airline delays other than through its relationship with post-merger market structure. Thus, an implicit identifying assumption is that absent the merger routes served by both DL and NW would not have experienced a relative change in on-time performance, conditional on the explanatory variables. While merger partners are not randomly assigned, a survey of potential motivations for the transaction does not present an obvious reason to discount the exclusion restriction. The U.S. Government Accountability Office (GAO) has reported that experts generally believe airlines seek to merge to increase profitability. The GAO explains airlines can lower costs by combining complementary assets and eliminating redundant activities, and also increase demand by "offering more seamless travel to more destinations." Moreover, airlines may reduce capacity to exercise greater market power and generate "increased revenues from higher fares in existing markets."[20] Although the transaction survived antitrust review by the DOJ, it was criticized as

[19] Although the merger was officially consummated during the fourth quarter of 2008, it is possible that the two firms changed their objective functions to reflect the expected merger prior to the official closing of the deal. Due to this uncertainty in when measured market structure and on-time performance should reflect the combination of the two firms, the robustness of this specification choice was tested by estimating a model with 2008 excluded from the sample. The results were qualitatively similar.

[20] http://news.delta.com/index.php?s=43&item=216

anticompetitive by some observers.[21] Alternatively, upon consummating the merger, DL claimed the deal brought "together two of the industry's most financially secure airlines to produce a best-in-class cost structure and an industry-leading balance sheet." The company further stated the transaction was expected to generate $2 billion in annual revenue and cost synergies.

While none of these opinions seem to explicitly cast doubt on the exogeneity of the excluded instrument, one key threat does arise. The merger may have caused changes in on-time performance, which are causally unrelated to the changes in local market competition, yet correlated with the excluded instrument. For example, the merger may have caused a change in operational performance as the two firms began integrating their factors of production into a single organization.[22] After presenting the main results, several alternative specifications are estimated to assess the possibility that unobserved "merger effects" are confounding the results.

5. Results

Two regressions are estimated for each specification; one with $MINS_{irmt}$ as the dependent variable and another with $ODDS_{irmt}$. The presentation of the results is organized in three sections. Section 5.1 presents the benchmark OLS regressions and discusses the control variables. Section 5.2 discusses the corresponding two-stage estimators and 5.3 discusses the validity of the instruments. Sections 5.4 and 5.5 present two sets of ancillary regression results.

[21] See, Bailey, Jeff. "Delta's Chief Faces Skeptics as He Tries to Sell a Deal With Northwest." The New York Times 16 Apr. 2008. Web. 23 Oct. 2010.

[22] See Kole and Lehn (2000) for a detailed discussion of the potential challenges of combining two airlines. It should be noted that the popular press has generally characterized the integration of Delta-Northwest as a smooth transition.

5.1. Benchmark results

The first two columns of table 2 present a parsimonious specification in which $MINS_{irmt}$ and $ODDS_{irmt}$ are regressed on market structure without any control variables. Only the carrier-market and year-quarter fixed effects are included. The coefficient on HHI_{mt} is positive and statistically significant in both regressions, consistent with the hypothesis that competition improves on-time performance. The third and fourth columns add in the control variables. With the control variables included, a 0.25 increase in HHI_{mt} is associated with a 1.2 minute increase in the expected arrival delay and a 0.011 increase in the probability of being delayed (when the initial probability of a delay is .22, the mean). A 0.25 increase in HHI_{mt} would occur if two equal-sized competitors with market shares of approximately 0.35 merged.[23]

Columns 5 through 10 present the instrumental variable estimates. Before comparing these results with the OLS estimates, the control variables warrant some discussion. Scanning across columns 3 through 8, all the weather control variables are associated with longer and more numerous delays. Income and population are also positively correlated with delays. All else equal, greater local market demand appears to decrease on-time performance. Changes in total traffic are also associated with reduced on-time performance. As hypothesized, additional flights at the origin and destination airport lead to longer and more frequent delays.

5.2. Instrumental variable estimates

In columns 5 and 6, when lagged HHI is used as the sole excluded instrument, there is a

[23] Note the Delta-Northwest merger generated an expected change in HHI greater than 0.25 in only one market, the Atlanta and Detroit airport-pair market.

two- to three-fold increase in the magnitude of the coefficients on HHI_{mt} in comparison to the OLS estimates. As expected the IV estimates have larger standard errors. Nonetheless, with $MINS_{irmt}$ as the dependent variable we can reject the statistical equivalence OLS and IV estimators—equivalence both in terms of the coefficients on HHI (via a z-test) as well as the estimators as a whole (via a regression-based Haussmann test with robust standard errors). This indicates the OLS estimator is biased. With $ODDS_{irmt}$ as the dependent variable, the imprecision of the two-stage estimator prevents us from rejecting statistical equivalence, however, the larger point estimate on HHI_{mt} again suggests the OLS estimator is biased toward zero.

In columns 7 and 8, $CNGHHI_{mt}$ is employed as the sole excluded instrument. With this approach, both of the regressions are statistically different than corresponding IV estimates. Again, we see a two- to three-fold increase in the magnitude of the coefficients on HHI_{mt}. Finally, in columns 9 and 10, the 2SLS estimators are presented. Once more, the results of the two-stage estimates indicate the OLS estimator is biased toward zero. While the similarity of these three approaches suggests both instruments are effectively remedying correlation between HHI_{mt} and the error term, we can perform a formal test of the overidentifying restriction. For both dependent variables, we are unable the reject the null hypothesis that the instruments are uncorrelated with the error.[24]

The coefficient of 14.03 on HHI_{mt} in column 9 indicates if HHI_{mt} increases by 0.25 then

[24] The Sargan-Hansen test is used to test the overidentifying restriction. The joint null hypothesis is that the instruments are uncorrelated with the error term. With *MINS* as the dependent variable the Chi-square(1) P-value = 0.699 and with the *ODDS* as the dependent variable the Chi-square(1) P-value = 0.738.

the expected duration of an arrival delay increases by approximately 3.5 minutes. The results in column 10 indicate a similar change in HHI_{mt} is associated with a 0.030 change in the probability of being delayed. To attach some economic significance to these results, Morrison and Winston (1989) estimate passengers are willing to pay $7.80 for a 3 percentage point change in carriers' on-time performance records.[25]

The lower portion of table 2 includes the coefficients on the excluded instruments and an F-statistic for the excluded instruments. That the estimated coefficients on $CHNGHHI_{mt}$ are less than one indicates, as standard oligopoly theory would suggest, that DL and NW's market shares contracted slightly as they combined and those of their rivals expanded.[26]

5.3. Instrument validity

The validity of the instruments rests on the assumption that they are: (a) relevant and (b) exogenous. The first stage results presented in table 2 show that both instruments are relevant. In all the first stage regressions the coefficients on HHI_{mt-6} and $CHNGHHI_{mt}$ are significant at the one-percent level. Moreover, all the F-statistics on the excluded instruments are well above ten, the generally accepted threshold (see, Staiger and Stock (1997) and Stock and Yogo (2005)).

On the other hand, exogeneity cannot be *directly* tested. However, the fact that we cannot reject the null hypothesis that both instruments are exogenous, lends considerable credence to the choice of instruments. It should also be noted that the instruments are seemingly correlated with HHI_{mt} for very different reasons. Thus, it is unlikely the two instruments are

[25] Per round-trip in 2008 dollars.

[26] Dafny, Duggan, and Ramanarayanan (2010) find that rivals' market shares increased substantially following the merger of two health insurance providers.

endogenous for the same reason, and hence, preventing us from rejecting the null in a test of the over identifying restrictions.

Despite the aforementioned evidence, it is worth further exploring the possibility that $CHNGHHI_{mt}$ may be correlated with delays for reasons other than the merger-induced changes in market structure. A chief concern is the possibility that DL and NW's on-time performance declined following the merger for reasons unrelated to changes in local market competition. Declines in operational performance may have occurred as the flight schedules, labor forces, and information technology systems of the two merging firms were integrated.

To examine the possibility that the merger altered merging carriers' operational performance, and that such a change is confounding the results, the regressions in columns 9 and 10 of table 2 are re-estimated excluding DL and NW from the sample. The first two columns of table 3 present these regressions. As one can see, while they are statistically equivalent to the full sample results, the coefficients on HHI_{mt} actually increase in magnitude, suggesting the full-sample results are not being driven by a decline in the merging carriers' on-time performance.

A slightly different type of threat to the exogeneity of $CHNGHHI_{mt}$ is the possibility of a uniform decline in on-time performance in all markets served by both DL and NW. Operational performance may decline as the airlines integrate their gates, ticket counters, and other airport-specific capital and labor. Furthermore, the merger may cause on-time performance to suffer in all markets served by both carriers, as the airlines integrate market-specific assets and re-optimize their schedules to reflect the change in ownership. To simultaneously address these concerns, the third and fourth columns of table 3 restrict the sample to only those markets served by both DL and NW prior to the merger. By restricting the sample this way, we eliminate the

possibility of a confounding effect common across all markets served by both DL and NW. It is encouraging to see the coefficients on HHI_{mt} are statistically equivalent to the full sample estimates.

5.4. Scheduled flight time and competition

One way carriers can achieve better on-time performance is by "padding" their schedules. That is, carriers can schedule additional flight time making it easier to arrive on time. Mazzeo (2003) finds that carriers actually schedule shorter elapsed flight times, all else equal, on more competitive routes—suggesting his results may underestimate the total effect of competition. As Mazzeo does, we can use *scheduled elapsed flight time* as the dependent variable in a model similar to that used to test the relationship between on-time performance and competition. Such a regression answers an interesting question in its own right. Scheduled elapsed flight time is a component of product quality distinct from on-time performance. Carriers likely compete along both dimensions. When searching for an airline ticket passengers may consider both the total scheduled flight time, as well as a given carrier's on-time performance record.

Table 4 presents two estimations that use the average scheduled elapsed time of flights offered by carrier *i* on route *r* during period *t*, $SCHDTIME_{irmt}$, as the dependent variable. The first column presents the OLS estimator and the second presents the 2SLS estimator. As the table indicates, there is no evidence that competition influences scheduled flight times.

5.5. Alternative measure of competition

As a final robustness check, the measure of local market structure, HHI_{mt}, is redefined to include only passengers flying on direct itineraries, that is excluding the connecting

passengers. Table 5 presents regressions that employ this new variable, $HHID_{mt}$. Columns 1 and 2 present estimates using the same excluded instruments as the initial results, while columns 3 and 4 use excluded instruments that are redefined using only direct traffic. It is reassuring to see both approaches yield estimates that are qualitatively similar to the baseline regressions.

6. Discussion and policy implications

As the airline industry continues to consolidate, an understanding of the likely non-price effects is of interest. What policy implications can be taken away from this study? Should antitrust agencies be concerned that consolidation in the airline industry will decrease consumer surplus via reduced service quality?

It is important to note that the results herein do not imply that industry-wide consolidation will lead to changes in the overall, industry-wide, level of delays. With carrier-route and time fixed effects, the coefficient on HHI implies that, all else equal, routes that experienced a relative increase (decrease) in local market concentration experienced a relative increase (decrease) in delays. The time fixed effects capture changes in the overall level of delays.[27] The overall level of delays is affected by exogenous factors such as congestion, weather, and long-run infrastructure investments, as well as carriers' scheduling decisions at the network and airport level. This study attempts to hold these effects constant to ascertain how local market competition affects delays.

The results suggest that consumers in markets that experience relative consolidation will experience a relative decline in service quality. However, a complete consumer welfare analysis

[27] Examining the year-quarter fixed effects indicates the overall level of delays declined after the merger, which one might speculate is due to a reduction in airport congestion associated with the coincident recession.

would require an understanding of the distribution, across markets, of consumers' willingness to pay for on-time performance.[28] Policymakers should also recognize the possibility that industry consolidation may impact on-time performance via changes in competition outside local markets, not captured by this study. For example, Prince and Simon (2009) find that multimarket contact increases delays. Also, Rupp (2009) elucidates the effects of airport level concentration with a special emphasis on how delays are measured. His findings, which reverse earlier results by Mayer and Sinai (2003), explain that airport consolidation leads to greater delays, particularly when delays are measured as deviations from scheduled departure/arrival times, rather than excess travel time.[29]

7. Conclusions

Empirical evidence on the relationship between competition and product quality is growing, yet still relatively sparse. The gap in the literature exists despite the fact that in many industries quality is a crucial determinant of demand and consumer welfare. Moreover, economic theory yields largely ambiguous predictions on the sign of the relationship. While lower price-cost margins erode the benefits of a strong reputation, heightened market discipline may spur competition in non-price attributes. This paper's results indicate the latter effect dominates the relationship between local market competition and on-time performance in the airline industry.

[28] Mazzeo (2003) emphasizes this point. He notes a total welfare analysis requires detailed knowledge of the costs of providing quality. Demand estimation that captures heterogeneity in consumers' willingness to pay for on-time performance would afford more decisive welfare statements.

[29] While carrier-route effects control for time invariant airport effects, it should be noted that the relationship between competition and service quality documented in this paper *is not conditioned* on changes in airport concentration that occur over time.

This study offers two key contributions to the literature. First, the within market estimator mitigates concerns that unobserved (time-invariant) route characteristics cause the correlation between competition and on-time performance. Second, using two different instruments, the study finds that OLS estimates are likely biased toward zero. The results indicate the effect of competition on airline delays is three times stronger than an OLS estimator suggests. Finally, while decisive statements regarding consumer and producer surplus are beyond the scope of this study, the results represent a first step in such an analysis. Further examination of the distributional and total welfare implications of the results is a potentially interesting area for future research.

References

Aghion, Philippe et al. "Competition and Innovation: An Inverted-U Relationship*." *Quarterly Journal of Economics* 120.2 (2005): 701-728.

Bar-Isaac, Heski. "Imperfect competition and reputational commitment." *Economics Letters* 89.2 (2005): 167-173.

Bar-Isaac, Heski, and Steven Tadelis. "Seller Reputation." *Foundations and Trends in Microeconomics* 4.4 (2008): 273-351.

Ciliberto, Federico, and Carola Schenone. "Are the Bankrupt Skies the Friendliest?" (2010): Available at SSRN: http://ssrn.com/abstract=1496519.

Dafny, Leemore, Mark Duggan, and Subramaniam Ramanarayanan. "Paying a Premium on Your Premium? Consolidation in the U.S. Health Insurance Industry." *Forthcoming, The American Economic Review* (2011).

Dana, James, and Yuk-Fai Fong. "Product Quality, Reputation, and Market Structure." (2010). Working paper.

Davis, Peter. "The Effect of Local Competition on Admission Prices in the U.S. Motion Picture Exhibition Market." *The Journal of Law and Economics* 48.2 (2005): 677-707

Domberger, Simon, and Avrom Sherr. "The Impact of Competition on Pricing and Quality of Legal Services." *International Review of Law & Economics* 9.1 (1989): 41-56.

Dranove, D., and W. D. White. "Recent Theory and Evidence on Competition in Hospital Markets." *Journal of Economics & Management Strategy* 3.1 (1994): 169-209.

Evans, William N., Luke M. Froeb, and Gregory J. Werden. "Endogeneity in the Concentration—Price Relationship: Causes, Consequences, and Cures." *The Journal of Industrial Economics* 41.4 (1993): 431-438.

Forbes, Silke J. "The Effect of Air Traffic Delays on Airline Prices." *International Journal of Industrial Organization* 26.5 (2008a): 1218-1232.

Forbes, Silke J. "The Effect of Service Quality and Expectations on Customer Complaints." *Journal of Industrial Economics* 56.1 (2008b): 190-213.

Foreman, Stephen Earl. "Publication of Information and Market Response: The Case of Airline on Time Performance Reports." *Review of Industrial Organization* 14.2 (1999): 147-162.

Gaynor, Martin. “What Do We Know About Competition and Quality in Health Care Markets?” *National Bureau of Economic Research Working Paper Series* No. 12301 (2006).

Hörner, Johannes. “Reputation and Competition.” *The American Economic Review* 92.3 (2002): 644-663.

Hoxby, Caroline M. “Does Competition among Public Schools Benefit Students and Taxpayers?” *The American Economic Review* 90.5 (2000): 1209-1238.

Klein, Benjamin, and Keith B. Leffler. “The Role of Market Forces in Assuring Contractual Performance.” *The Journal of Political Economy* 89.4 (1981): 615-641.

Kranton, Rachel E. “Competition and the Incentive to Produce High Quality.” *Economica* 70.279 (2003): 385-404.

Lee, Darin, and Nicholas G. Rupp. “Retracting a Gift: How Does Employee Effort Respond to Wage Reductions?” *Journal of Labor Economics* 25.4 (2007): 725-761.

Matsa, D. A. “Competition and product quality in the supermarket industry.” *Unpublished Manuscript*.

Mayer, Christopher, and Todd Sinai. “Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Evil.” *The American Economic Review* 93.4 (2003): 1194-1215.

Mazzeo, Michael J. “Competition and Service Quality in the U.S. Airline Industry.” *Review of Industrial Organization* 22.4 (2003): 275-296.

Morrison, Steven A., and Clifford Winston. “Enhancing the Performance of the Deregulated Air Transportation System.” *Brookings Papers on Economic Activity.* (1989): 61-123.

Prince, J., & Simon, D.H. “Multimarket Contact and Service Quality: Evidence from On-Time Performance in the US Airline Industry.” *Academy of Management Journal*. (2009) 52: 336-35

Rupp, N., D. Owens, and L. Plumly. “Does Competition Influence Airline On-Time Performance?” *Advances in airline economics: Competition policy and antitrust* (2006): 251–272.

Rupp, N. G., and G. M. Holmes. “An Investigation into the Determinants of Flight Cancellations.” *Economica* 73.292 (2006): 749-783.

Rupp, Nicholas G. “Do carriers internalize congestion costs? Empirical evidence on the

internalization question." *Journal of Urban Economics* 65.1 (2009): 24-37.

Spence, A. Michael. "Monopoly, Quality, and Regulation." *The Bell Journal of Economics* 6.2 (1975): 417-429.

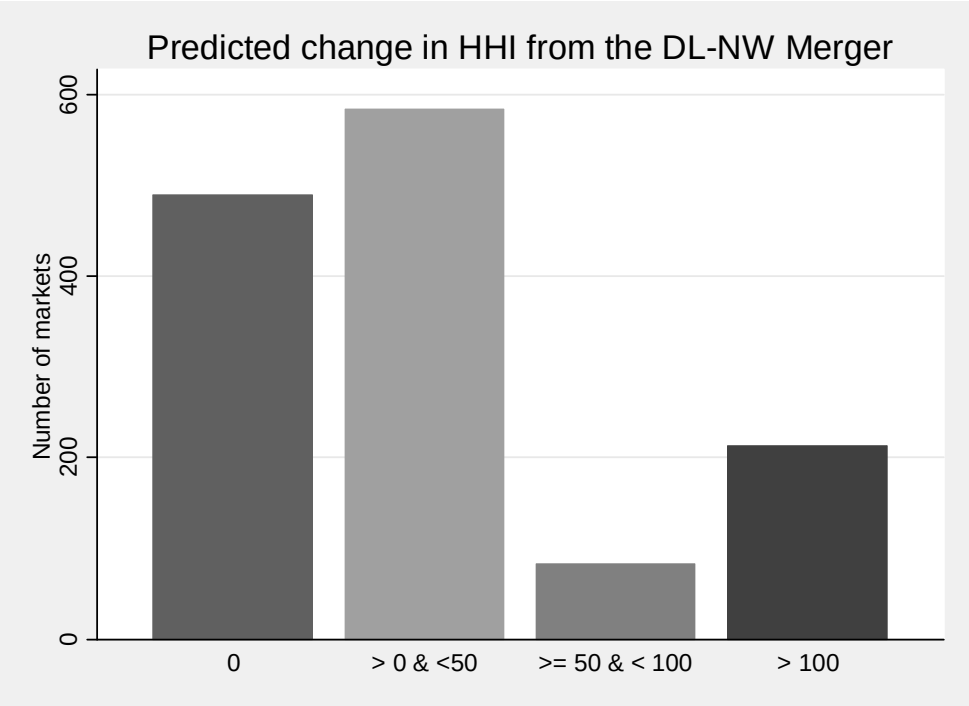
Staiger, D. and J.H. Stock. "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65 (1997): 557 – 586.

Stock, James H. and Motohiro Yogo. "Testing for Weak Instruments in IV Regression," in Identification and Inference for Econometric Models. (2005).

Suzuki, Yoshinori. "The relationship between on-time performance and airline market share: a new approach." *Transportation Research Part E: Logistics and Transportation Review* 36.2 (2000): 139-154.

Swan, Peter L. "Durability of Consumption Goods." *The American Economic Review* 60.5 (1970): 884-894.

Figure 1



HHI scaled from 0 – 10,000.

Table 1: Summary Statistics

					Time-demeaned by carrier-route		
Variable	Mean	Std. Dev.	Min	Max	Std. Dev.	Min	Max
MINS	6.412	7.319	-33.618	78.094	5.804	-41.276	54.263
ODDS	-1.335	0.511	-4.595	1.501	0.384	-3.262	2.326
HHI	0.615	0.225	0.155	1.000	0.056	-0.432	0.441
TIME	0.639	0.074	0.003	0.999	0.042	-0.815	0.606
ORGDPT	10.852	0.870	7.589	12.537	0.109	-0.827	0.649
DSTDPT	10.853	0.870	7.589	12.537	0.109	-0.827	0.649
INCOME	12.931	0.664	10.510	14.299	0.040	-0.136	0.099
POP	16.148	0.665	13.575	17.433	0.015	-0.045	0.043
RAINORG	2.238	2.212	0.000	13.000	1.721	-5.600	9.467
SNOWORG	0.320	0.802	0.000	9.775	0.694	-3.992	7.367
RAINDST	2.236	2.212	0.000	13.000	1.720	-5.600	9.467
SNOWDST	0.320	0.802	0.000	9.775	0.693	-3.992	7.367
CHNGHHI	0.002	0.018	0.000	0.472	0.014	-0.199	0.273

1369 markets, 3981 carrier-route combinations, N = 60217 (carrier-route-year-quarter)
Means and standard deviations are weighted by the number of flights.

Table 2

VARIABLES	1 *MINS*	2 *ODDS*	3 *MINS*	4 *ODDS*	5 *MINS*	6 *ODDS*	7 *MINS*	8 *ODDS*	9 *MINS*	10 *ODDS*
HHI	4.699***	0.236***	4.931***	0.262***	16.500**	0.530	13.500***	0.720***	14.030***	0.687***
	[0.922]	[0.060]	[0.875]	[0.056]	[7.821]	[0.460]	[3.589]	[0.222]	[3.349]	[0.200]
TIME			13.31***	1.185***	13.07***	1.180***	13.12***	1.175***	13.12***	1.177***
			[0.894]	[0.061]	[0.902]	[0.061]	[0.897]	[0.061]	[0.896]	[0.061]
ORGDPT			3.737***	0.319***	4.252***	0.332***	4.104***	0.339***	4.137***	0.338***
			[0.532]	[0.0301]	[0.584]	[0.036]	[0.560]	[0.033]	[0.548]	[0.032]
DSTDPT			4.206***	0.287***	4.718***	0.299***	4.575***	0.307***	4.602***	0.305***
			[0.580]	[0.035]	[0.656]	[0.039]	[0.609]	[0.036]	[0.604]	[0.036]
INCOME			10.60**	1.215***	15.54***	1.339***	14.16***	1.409***	14.43***	1.397***
			[4.353]	[0.261]	[5.943]	[0.345]	[4.672]	[0.285]	[4.728]	[0.286]
POP			46.32***	0.638	43.24***	0.556	44.17***	0.521	43.85***	0.523
			[10.500]	[0.553]	[10.870]	[0.577]	[10.510]	[0.567]	[10.590]	[0.568]
RAINORG			0.214***	0.0138***	0.202***	0.0134***	0.206***	0.0133***	0.204***	0.0133***
			[0.018]	[0.001]	[0.021]	[0.001]	[0.018]	[0.001]	[0.018]	[0.001]
SNOWORG			0.327***	0.0346***	0.328***	0.0346***	0.328***	0.0346***	0.328***	0.0346***
			[0.050]	[0.003]	[0.051]	[0.003]	[0.050]	[0.003]	[0.050]	[0.003]
RAINDST			0.297***	0.0199***	0.285***	0.0196***	0.288***	0.0194***	0.287***	0.0195***
			[0.018]	[0.001]	[0.020]	[0.001]	[0.018]	[0.001]	[0.018]	[0.001]
SNOWDST			0.226***	0.00410	0.228***	0.00421	0.226***	0.00414	0.227***	0.00422
			[0.052]	[0.003]	[0.053]	[0.003]	[0.052]	[0.003]	[0.052]	[0.003]
EXCLUDED INSTRUMENTS										
HHI_{t-6}					0.096***				0.091***	
					[0.019]				[0.019]	
CHNGHHI							0.926***		0.914***	
							[0.051]		[0.051]	
F-test for excluded instruments					25.42		328.69		167.52	
Observations	60,217	60,217	60,217	60,217	60,086	60,086	60,217	60,217	60,086	60,086
	0.234	0.204	0.265	0.243	0.252	0.241	0.258	0.239	0.257	0.239

Cluster robust standard errors in brackets. All regressions include carrier-route and year-quarter effects. First-stage coefficients on the instruments included in the second stage are suppressed.

*** p<0.01, ** p<0.05, * p<0.1

Table 3

	1	2	3	4
	Rivals Only		Overlap Routes	
VARIABLES	*MINS*	*ODDS*	*MINS*	*ODDS*
HHI	20.320***	0.998***	18.620***	0.927***
	[4.431]	[0.268]	[4.048]	[0.235]
TIME	13.49***	1.215***	13.33***	1.116***
	[1.016]	[0.071]	[1.030]	[0.068]
ORGDPT	5.486***	0.436***	4.752***	0.323***
	[0.686]	[0.040]	[0.711]	[0.039]
DSTDPT	5.100***	0.332***	6.053***	0.351***
	[0.762]	[0.044]	[0.783]	[0.044]
INCOME	17.23***	1.637***	29.53***	2.081***
	[5.313]	[0.321]	[8.407]	[0.481]
POP	40.70***	-0.155	36.43**	-0.177
	[12.81]	[0.662]	[15.27]	[0.859]
RAINORG	0.230***	0.0144***	0.172***	0.010***
	[0.020]	[0.001]	[0.024]	[0.001]
SNOWORG	0.268***	0.028***	0.384***	0.035***
	[0.0567]	[0.004]	[0.061]	[0.004]
RAINDST	0.270***	0.018***	0.271***	0.017***
	[0.019]	[0.001]	[0.023]	[0.001]
SNOWDST	0.214***	1.69e-05	0.383***	0.009**
	[0.060]	[0.004]	[0.062]	[0.004]
EXCLUDED INSTRUMENTS				
HHI_{t-6}	0.114***		0.050***	
	[0.020]		[0.017]	
CHNGHHI	0.954***		0.911***	
	[0.079]		[0.054]	
F-test excluded instruments	84.14		144.20	
Observations	48,786	48,786	43,198	43,198
R-squared	0.266	0.257	0.270	0.243

Cluster robust standard errors in brackets. All regressions include carrier-route and year-quarter effects. First-stage coefficients on the instruments included in the second stage are suppressed.
*** p<0.01, ** p<0.05, * p<0.1

Table 4

VARIABLES	1 *SCHDTIME* (OLS)	2 *SCHDTIME* (2SLS)
HHI	1.561	-0.045
	[2.613]	[0.629]
TIME	3.703***	3.655***
	[0.614]	[0.618]
ORGDPT	3.010***	3.089***
	[0.383]	[0.399]
DSTDPT	0.558	0.620*
	[0.347]	[0.365]
INCOME	-7.992***	-7.297**
	[2.862]	[3.051]
POP	-29.25***	-29.80***
	[5.759]	[5.794]
Observations	60,217	60,086
R-squared	0.193	0.193

Cluster robust standard errors in brackets. All regressions include carrier-route and year-quarter effects.
First-stage regression suppressed in column 2.
*** p<0.01, ** p<0.05, * p<0.1

Table 5

VARIABLES	1 *MINS*	2 *ODDS*	3 *MINS*	4 *ODDS*
HHID	13.530***	0.657***	10.820***	0.549***
	[3.120]	[0.187]	[3.049]	[0.189]
TIME	13.000***	1.167***	13.030***	1.162***
	[0.904]	[0.061]	[0.919]	[0.062]
ORGDPT	4.478***	0.354***	4.288***	0.347***
	[0.578]	[0.034]	[0.560]	[0.033]
DSTDPT	4.945***	0.321***	4.729***	0.309***
	[0.641]	[0.038]	[0.623]	[0.037]
INCOME	15.43***	1.441***	14.00***	1.381***
	[4.852]	[0.293]	[4.846]	[0.292]
POP	42.71***	0.472	43.70***	0.497
	[10.640]	[0.576]	[10.65]	[0.571]
RAINORG	0.210***	0.0136***	0.211***	0.0136***
	[0.018]	[0.001]	[0.018]	[0.001]
SNOWORG	0.331***	0.0346***	0.325***	0.0342***
	[0.051]	[0.003]	[0.050]	[0.003]
RAINDST	0.294***	0.0198***	0.296***	0.0199***
	[0.018]	[0.001]	[0.018]	[0.001]
SNOWDST	0.232***	0.00448	0.232***	0.00442
	[0.052]	[0.003]	[0.0525]	[0.003]
EXCLUDED INSTRUMENTS^				
HHI_{t-6}	0.111***		0.101***	
	[0.058]		[0.019]	
CHNGHHI	0.919***		0.866***	
	[0.058]		[0.041]	
F-test excluded instruments	134.12		230.48	
Observations	59,644	59,644	58,556	58,556
R-squared	0.251	0.235	0.259	0.240

Cluster robust standard errors in brackets. All regressions include carrier-route and year-quarter effects. First-stage coefficients on the instruments included in the second stage are suppressed.

^Excluded instruments in columns 3 and 4 are constructed using direct traffic only.

*** p<0.01, ** p<0.05, * p<0.1

Appendix: DB1B data construction

All three tables from the Origin and Destination Survey (DB1B) are downloaded from the U.S. Department of Transportation website for 2004:Q1 through 2010:Q3. These tables are then merged to create a table that includes a record for each coupon of each domestic itinerary. Data cleaning proceeds as follows. First, itineraries with more than 4 coupons in total or more than 2 coupons on a directional leg are excluded. Second, itineraries involving a domestic portion of an international O&D, or a foreign carrier flying between two U.S. points are excluded. Third, itineraries with a fare in the top-1 percentile of the distribution for the given quarter, less than 20 dollars, or deemed non-credible by the DOT are excluded. Fourth, itineraries involving an airport outside the contiguous 48 states are excluded. Fifth, interline itineraries (those with a change in ticketing carrier) are excluded. Sixth, itineraries that are neither one-way nor round-trip are excluded. Seventh, using the DOT T100 data, itineraries involving a coupon that does not correspond to a flight provided 12 or more times in a given quarter are eliminated.

Finally, the data are aggregated by year, quarter, ticketing carrier, and non-directional airport-pair market. Each observation contains a variable for the passenger weighted average directional fare and total passengers for carrier i in market m during year-quarter t. All observations with fewer than 5 passengers are dropped. Also, all observations associated with non-directional airport-pairs that generate fewer than 50 passengers in the DB1B sample for a particular quarter are excluded. The resulting data are utilized to calculate the revenue weighted HHI for each market m during year-quarter t, as well as the 2007 revenue weighted market shares for DL and NW. These year-quarter-market data are then matched to the on-time performance data. Each route r in the on-time performance data is matched to the market m with the same endpoint airports.

Chapter 2

The Product Market Effects of the Merger between Delta and Northwest

1. Introduction and literature review

A number of economists and practitioners have recently noted the importance of retrospective analyses of horizontal mergers (see, Ashenfelter et al. 2009, Farrell et al. 2009, Weinberg 2008, Agirst and Priske 2010, and Kwoka and Greenfield 2011). Such studies can shed light on the efficacy of prospective merger analysis and antitrust enforcement, as well as enhance the field's understanding of the effects of horizontal mergers. This paper analyses the price effects of the 2008 merger between Delta Air Lines (DL) and Northwest Airlines (NW), a transaction that was subject to careful scrutiny by the U.S. Department of Justice (DOJ).

DL and NW announced plans to merge in April of 2008. By September of the same year, the firms had received unconditional antitrust clearance from the European Commission and the overwhelming consent of their voting shareholders. The transaction was consummated immediately after receiving antitrust approval from the DOJ on October 29, 2008. Upon closing its review, the DOJ stated, "After a thorough, six-month investigation . . . the [Antitrust] Division has determined that the proposed merger between Delta and Northwest is likely to produce substantial and credible efficiencies that will benefit U.S. consumers and is not likely to substantially lessen competition." Nonetheless, the transaction was criticized as anticompetitive by many observers. Commonly cited concerns were increased exercise of market power, hub closures, diminished service to smaller airports, and the possibility that the merger would spur further consolidation among large U.S. airlines.[1] Moreover, concerns from the public appear to

[1] See for eample, Bailey, Jeff. "Delta's Chief Faces Skeptics as He Tries to Sell a Deal with Northwest." The New York Times 16 Apr. 2008.

have triggered at least three congressional hearings.[2]

Several studies have examined the price effects of airline mergers that occurred during the 1980s. Studying the 1986 merger of Northwest Airlines and Republic Airlines, Borenstein (1990) found that from 1985 to 1987 fares increased by 9.5 percent relative to the industry average on routes involving a shared hub. Alternatively, he found little evidence of a change in fares associated with the merger between Trans World Airlines (TWA) and Ozark Airlines. On routes where both of the merging carriers had a market share greater than ten percent, Morrison (1996) finds that fares decreased by 1.8 percent from the year before to the year after the Northwest-Republic merger and that fares increased by 4.4 percent over a similar time period for the TWA-Ozark merger. Morrison also finds the 1989 merger between USAir and Piedmont resulted in notable increases in fares, a result confirmed by Kwoka and Shumilkina (2010), who also find fares increased on routes where the merger eliminated a potential competitor. Kwoka and Shumilkina find that fares increase by 10 to 12 percent on routes served by both of the merging carriers and by 5 to 6 percent on routes where one of the merging carriers was incumbent and the other operated at the endpoint airports. Peters (2006) estimates the price effect of each of the aforementioned mergers as well as the merger between Continental Airlines and People Express and the merger between Delta and Western. In all five cases Peters finds a price increase in the year following the merger ranging from 7.2 to 29.3 percent. Finally, Kim

[2] (1) "Impact Of Consolidation On The Aviation Industry, With A Focus On The Proposed Merger Between Delta Air Lines And Northwest Airlines," House Transportation and Infrastructure Committee Subcommittee on Aviation, May 14, 2008; (2) "An Examination of the Delta-Northwest Merger" Senate Judiciary Committee, Subcommittee on Antitrust, Competition Policy and Consumer Rights , April 24, 2008; (3) "Competition in the Airline Industry." Hearing Before the Task Force on Competition Policy and Antitrust Laws of the Committee on the Judiciary, House of Representatives, April 24, 2008.

and Singal (1993) estimate the average effect of 14 mergers, finding fares increased by 9.44 percent on routes served by both firms prior to the merger.

All of the aforementioned papers use some form of difference-in-differences (DID) estimator to infer the price effects of the mergers. While this method is also standard in the broader merger retrospective literature,[3] it rests crucially on selecting an appropriate control group. DID estimation hinges on the assumption that the difference between outcomes in markets potentially affected by the merger (the treatment group) and outcomes in markets not affected by the merger (the control group), would have remained constant absent the transaction. The outcomes in the control group represent the assumed counterfactual, that is, the trend of prices but for the merger. Simpson and Schmidt (2007) emphasize that even if the treatment and control groups experience the same supply and demand shocks, changes in the control markets may not accurately control for supply and demand shocks because the factors that determine how shocks affect market outcomes may differ across the two groups. Differences in rivalry among producers or differences in market-specific cost structure may alter how supply and demand shocks affect market outcomes.

This study uses a DID regression model to assess the price effects of the merger between and DL and NW. An important consideration is the fact that the timing of the merger coincided with a period of volatile changes in the level of airline fares. In the same year as the merger, the airline industry saw a drastic spike in jet fuel prices and a collapse of aggregate demand due to

[3] An incomplete list of example includes: Ashenfelter and Hosken (2008), Taylor and Hosken (2007); Simpson and Taylor (2008); Tompson (2011); and Romano and Balan (2011).

the onset of the 2008-2009 recession.[4] If these shocks to supply and demand were transmitted to fares differently in treatment and control markets, they could potentially bias the estimated difference-in-differences.

To refine the treatment-control comparison, and control for the potential confounders alluded to by Simpson and Schmidt, this study includes regression specifications where the time trend is allowed to vary across predetermined groups—groups of markets that likely differ in competitive rivalry and market-specific costs. Markets are grouped based on pre-merger market structure and by route distance. Then group-time fixed effects control for the possibility of a unique time trend in each subset of markets. Treatment markets are compared to control markets of a similar distance and with a similar level of pre-merger competition.

The results offer no evidence of an anti-competitive price increase. While some specifications indicate a small decrease in relative fares, specifications with a more flexible time trend indicate there was no change in fares associated with the merger. The general conclusions are robust with regard to changes in the choice of aggregate control group, the window of time surrounding the merger, and the geographic market definition. Several subsample analyses are also performed to test for different merger effects in markets without competition from Southwest Airlines and where the merger resulted in a particularly large change in market structure.

While these results are interesting in that they contrast several earlier studies of airline

[4] See, Commercial Aviation: Airline Industry Contraction Due to Volatile Fuel Prices and Falling Demand Affects Airports, Passengers, and Federal Government Revenues, GAO-09-393 April 21, 2009, http://www.gao.gov/new.items/d09393.pdf

mergers, they should not be considered a comprehensive assessment of the merger's consequences. The relatively short window of time surrounding the merger potentially conceals long run changes in operating efficiency and market power. Additionally, the analysis focuses solely on markets where an incumbent competitor is eliminated. Changes in competition may also occur in markets where a potential competitor is eliminated (as documented by Kwoka and Shumilkina (2010)), or where the merger creates a new product as the two carriers' networks are combined. Moreover, changes in multimarket contact and the overall structure of the industry may alter coordinated pricing decisions among carriers.

2. Data and Model

Quarterly data from 2006:Q4 to 2010:Q3 are drawn from the U.S. Department of Transportation's (DOT's) Origin and Destination Databank 1B, a ten percent quarterly sample of passengers traveling on U.S. certified carriers, commonly referred to as the DB1B. This sample provides two years of data before and two years of data after the merger's consummation. After imposing the data to a standard set of filters elaborated in appendix A, the ticket data are collapsed so that each observation indicates the passenger weighted average one-way fare and total number of passengers traveling on an itinerary sold by carrier *i*, in market *m*, in a particular quarter *t*.[5]

The baseline definition of a market includes all non-stop and one-stop connecting service between two airports, otherwise known as a non-directional airport-pair market. It is also

[5] In markets where DL and NW both offered service before the merger, in each period *t* the data are collapsed into a single observation for the two carriers. This maintains consistency after the merger combined to two carriers.

common practice to define markets as city-pairs, grouping airports in the same metropolitan area. Accordingly, all of the regression analysis is replicated using a city-pair definition. The city-pair definition utilized is that suggested by Brueckner, Lee, and Singer (2010).[6]

The basic empirical framework is a DID regression model. The regression equation takes the form:

$$\ln(FARE_{imt}) = \alpha_{im} + \beta_1 \, DLPOST_{mt} + \beta_2 \, RIVPOST_{mt} + X_{mt} \, \gamma + \delta_t + e_{irt}, \qquad (1)$$

where $FARE_{imt}$ is the passenger weighted average fare of tickets sold by carrier i in market m during period t. The vector X_{mt} contains control variables for the endpoint population and personal income.[7] The α_{im} are carrier-market fixed effects that capture time-invariant differences in prices across carrier-markets and the δ_t are time fixed effects that control for the sample-wide time trend of fares. $DLPOST_{mt}$ and $RIVPOST_{mt}$ are dummy variables designed to capture the difference-in-differences associated with the merger. Two different post-merger dummies are employed to allow for a different effect on the fares of the merged carrier and those of its rivals. $DLPOST_{mt}$ equals one after the merger if carrier i is DL (the combined carrier) and market m is a market potentially affected by the merger, otherwise it equals zero. $RIVPOST_{mt}$ equals one after the merger if carrier i is not the merged carrier and market m is one of the markets potentially affected by the merger, otherwise it equals zero.

[6] The definition of a market can also differ by type of service, i.e. non-stop service only versus non-stop and connecting together. Replicating the analysis in this paper with markets defined to include only direct service yields little evidence of anti-competitive price increase.

[7] These data are drawn from the U.S. Bureau of Economic Analysis and U.S. Census Bureau. There are two variables: one representing the geometric mean of the endpoint populations and the other representing the geometric mean of the endpoints personal incomes.

The markets potentially affected by the merger, i.e. the treatment markets, are defined with the 2010 U.S. Department of Justice (DOJ) and Federal Trade Commission (FTC) horizontal merger guidelines in mind. The guidelines state that "mergers resulting in highly concentrated markets [HHI above 2500] that involve an increase in the HHI of more than 200 points will be presumed to be likely to enhance market power." Thus all markets where the anticipated change in HHI is greater than 200 points and the anticipated post-merger level of HHI is greater than 2500 are defined as treatment markets.[8] Markets that were served by DL and NW prior to the merger, which do not meet this threshold, are excluded from the analysis. Ancillary specifications are also included to test for a different price effect in markets where the merger resulted in a post-merger HHI greater than 4000 and a change in HHI greater than 400 points.[9]

In all regressions the standard errors are clustered by city-pair market. While the carrier-route effects capture time-invariant group error components, the clustered variance estimator accounts for serial correlation in the remaining error term. Clustering at the city-pair level also accounts for correlation due to time-varying error components that are common across carrier-routes within a market (either city-pair or airport-pair).

2.1. Control groups

As mentioned in the introduction, selection of a comparison group is paramount in a difference-in-differences analysis. Accordingly, this study employs three different control

[8] The anticipated change in HHI is equal to twice the product of the of the market shares of the merging firms.

[9] This is an arbitrary cut-off that roughly splits the sample of treatment routes in half.

groups. Group one, the preferred control, consists of all fares in markets not served by DL or NW before the merger. Markets served by only one of the merging carriers are excluded as they may have experienced a change in price if the non-incumbent merging carrier was acting as a potential competitor or if the merged carrier realized efficiency gains after the merger. Nevertheless, a potential problem with this preferred control group is that the underlying markets may differ systematically from the treatment markets. After all, these are the markets that DL and NW chose not to serve before the merger. Referring to table 1, it is worth noting that control group one differs from the treatment group along several observable characteristics. Specifically, fares in the treatment group were higher before the merger, market structure was less concentrated, there was less competition from Southwest, and the average nonstop distance is longer.

While some of the regression specifications allow the underlying time trend of fares to vary by the observable characteristics described in table 1, there is still a concern—due to potentially unobserved differences—that the time trend of fares in the control group may be different from what would have occurred in the treatment group absent the merger. Any differences in the way shocks to supply and demand are passed through to fares will bias the estimated difference-in-differences.

Control group two consists of the non-merging carriers' fares in markets served by only one of the merging carriers with a market share less than 15%. As shown in table 1, the average distance, pre-merger fare, and pre-merger level of HHI are all closer to the respective averages in the treatment group than was the case with the control group one. Nonetheless, it should be noted that markets in control group two are intentionally excluded from control group one, as

they are potentially affected by the merger. However, that the merging carrier had a market share less than 15 percent limits the concern that the merging carrier had a strong competitive influence on other incumbents.

Finally, control group three includes the fares of non-merging carriers in all markets not served by both of the merging carriers. Thus, this group contains control groups one and two, as well as fares from markets served by one of the merging carriers with a market share greater than 15 percent. While it is more difficult to argue this set of markets was unaffected by the merger, the third control group is most similar to the treatment group in terms of the average level of fares and is also more representative of the industry as a whole.[10]

2.2. Group-time effects

In the baseline model the time effects, δ_t, are captured with a set of year-quarter dummy variables. They control for sample-wide changes in fares, arising from macroeconomic fluctuations, changes in input prices, and inflation. However, as explained in the introduction, it is often unrealistic to assume industry-wide cost and demand shocks will affect outcomes in all markets identically. Market-specific cost structures and the level of competition may affect the extent to which cost and demand shocks are transmitted to changes in price.

For example, long distance markets may be affected differently by changes in demand, as there are fewer feasible substitute modes of transportation, e.g. car or rail. Also the effect of

[10] In the DID analysis, to the extent that fares in the control group experienced a post-merger increase because of a reduction in competition, the estimated effect of the merger will have a negative bias. Alternatively, if fares in the control group are lower that they would have been but for the merger, the estimated merger effect will have a positive bias.

changes in input prices (e.g. fuel) may also vary by route distance. Accordingly, in some specifications, markets are grouped into quartiles based on the nonstop distance between market endpoints. Then, the time dummies are interacted with a set of dummy variables for the distance quartiles (the quartile of shortest distance routes is the base group). The resulting distance-time effects allow for a unique time trend of fares in each distance quartile.[11]

It may also be the case that changes in demand and costs are transmitted to fares differently depending on the extent of route competition. For example, changes in cost may be passed through to fares more readily in highly competitive markets. Accordingly, markets are grouped into quartiles based on the pre-merger level of HHI.[12] The time period dummies are interacted with a set of three indicators for the HHI quartiles, allowing a unique time trend for each group (the quartile with the lowest values of HHI serves as the base group). Table 2 provides a matrix that lists the number of markets in the treatment and control groups that respectively fall into each quartile.

Finally, the time fixed effects are allowed to vary based on the presence of Southwest Airlines. Several studies such as Morrison (2001), Goolsbee and Syverson (2008), and Breuckner, Lee, and Singer (2010) have shown that competition from Southwest provides a unique form of downward pressure on fares. This special form of competition may also alter the

[11] The distance-time effects confer the type of comparison achieved by Borenstien (1990) and Kim and Singal (1993), who compare the change in average fares for each route affected by a merger to the change in average fares in a group of control markets of a similar distance. However, in the regression framework markets potentially affected by the merger are not only compared to markets of similar distance, they are also compared to routes with a similar level of competition.

[12] Formally, the estimation data set is collapsed so there is a single observation for each market *m* that indicates the passenger-share HHI for the entire year in 2007. This market level data on pre-merger HHI is then divided into quartiles.

way demand and cost shocks are passed through to fares. It is also worth noting that Southwest's aggressive fuel hedging strategy, which has been widely covered in the popular press, may have altered the time trend of fares in markets served by Southwest.[13] While hedging strategies do not alter the opportunity cost of using jet fuel in the production process, they certainly can have consequences for an airline's financial condition, which may in turn alter pricing strategies (see, Morrison and Winston 1995; Busse 2002; Ciliberto and Schenone 2010).[14] To control for a potentially different time trend in markets served by Southwest, we define a dummy equal to one in all time periods if market *m* was served by Southwest before the merger.[15] This dummy is then interacted with the time period dummies, yielding a set of Southwest-time effects.

In the regressions that employ a full set of group-time effects, the identifying assumptions are weaker. Markets in the treatment group are compared to markets in the control group with similar pre-merger characteristics. The assumption is that absent the merger treatment markets would have experienced the same change in fares over time as markets in the control group with a similar distance and a similar level of premerger competition.

2.3. Merger Window

Another important specification choice is the window of data surrounding the merger. While the timing of the change in ownership structure is a discrete event, the effects of a merger

---

[13] See for example, Justin Baer, "Southwest Airlines' fuel hedging boosts profits." The Financial Times, July 25 2008.

[14] In fact the GAO reported that "by the third quarter of 2008, jet fuel prices increased 60 percent over 2007 levels. Seven smaller airlines ceased operations during the first half of the year and others entered bankruptcy, in large part because of high fuel prices . . . 2008, Air Midwest, Aloha Airlines, ATA Airlines, Champion Air, EOS Airlines, Big Sky Air, and Skybus Airlines ceased operation, while Frontier and Sun Country filed for Chapter 11 bankruptcy but are still operating."

[15] The dummy is formally defined to equal one if Southwest had a passenger market share greater than 10 percent in 2007.

may materialize gradually over time. Nonetheless, as Ashenfelter and Hosken (2008) point out, it is preferable to employ an event window that is long enough to capture changes in price associated with the merger, however, short enough to limit the possibility of confounding time-varying effects from other changes in the industry. Focusing on this tradeoff, three different windows are employed.

The preferred window includes 5 quarters of pre-merger data, 2007:Q1 to 2008:Q1, and 5 quarters of post-merger data, 2009:Q2 to 2010:Q2. Eliminating the data surrounding the merger avoids the contentious issue of determining when the merging firms began coordinating their pricing decisions.[16] Ending the post-merger period in 2010:Q2 limits the possibility that the merger between Continental Airlines and United Airlines on October 1, 2010 will contaminate the control group. The second window spans 2006:Q4 to 2008:Q2 and 2009:Q1 to 2010:Q3. The third window does not drop any data, spanning 2006:Q4 to 2010:Q3 with the post-merger period beginning in 2009:Q4.

Again, one should note that a potential disadvantage of these relatively short event windows is that some merger effects may not occur for several years. While the literature has generally found that efficiency gains are slower to materialize than the exercise of market power (e.g. Kim and Signal (1994) and Focarelli and Panneta (2003)), the bias may go in either direction, as capacity adjustments and coordinated pricing effects may take time to materialize. However, even if additional data were available, it may not be desirable to use. Morison (1996) uses more than 7 years of post-merger data and notes that as the time period after the merger

[16] Kim and Singal (1994) and Borenstien (1990) both fare increases prior to official merger consummation dates.

lengthens, "it is increasingly likely that the observed changes are due to other factors in addition to the merger."

3. Results

It is helpful to preface the formal regression analysis with a graphical comparison of fares in the treatment and control groups. Figure 1 offers two such illustrations. The upper panel plots the mean of $FARE_{imt}$ by quarter for the treatment and control groups. The lower panel plots the difference between the mean of ln($FARE_{imt}$) in the treatment group and the mean of ln($FARE_{imt}$) in each control group. This lower plot is directly analogous to the regression models. Movements along the vertical axis represent log point changes in the mean fare of the treatment group relative to the control group. The interpretation is equivalent to the coefficients on $DLPOST_{mt}$ and $RIVPOST_{mt}$—though the fares of DL and its market rivals are averaged together in these plots.

As the upper plot shows, fares are highest in the treatment group. Control group one has the lowest level of fares, followed by groups two and three respectively. However, the same general time path emerges in all groups. Fares initially increase and then fall dramatically following third quarter of 2008. Referring to the lower panel we can see that fares in the treatment group relative to the control groups reach a minimum in 2008:Q4 as the general level of fares is falling dramatically. While there is certainly some evidence that relative fares in the treatment group are on average lower after the merger, the change does not seem dramatic.

3.1. Regression analysis

Discussion of the regression results focuses on table 3 which presents estimates from the regressions that employ the airport-pair market definition and the preferred merger window. The

first column presents the baseline model as specified in equation (1). Then in columns 2 through 4 the group-time effects are sequentially added to refine the treatment-control comparison. Column 4 provides the most flexible time trend as the temporal fixed effects vary based on quartile of market distance, quartile of pre-merger HHI, and the presence of Southwest airlines. Columns 5 and 6 present the subsample analyses. In addition to having separate rows for each regression covariate, as is standard in a regression table, table 3 is also divided horizontally into three larger panels. Each panel represents a separate set of regressions that employ a particular control group. While the discussion focuses on table 3, the basic analysis is replicated using the alternative merger windows and the city-pair market definition. Tables 4 through 8 are organized in the same manner, however, they present the estimates using these alternative specification choices.

In column 1 we can see the merger is associated with a three to five percent reduction in relative fares. While the magnitude of the coefficients on $DLPOST_{mt}$ and $RIVPOST_{mt}$ differ, the discrepancies are not statistically significant. This general pattern of a similar effect among DL and its rivals persists throughout most of the estimations. In columns 2 and 3, when the time dummies are interacted with indicator variables for the distance and HHI quartiles, the estimated merger effect becomes smaller in absolute terms, indicating roughly a two to three percent decline in relative fares. A notable change in the results occurs when we control for the presence of Southwest. After allowing for a different time trend in markets served by Southwest, there is minimal evidence of a decline in relative fares following the merger.

Notable among these findings is the lack of any evidence of a price increase following the merger. Potentially, the treatment group is too broad. The merger guidelines are not

industry-specific; hence, the concentration thresholds may be too low for the airline industry. In fact 94 percent of the airport-pair markets in the sample are considered "highly concentrated" under the guideline's threshold of 2500. To test if the threshold for the treatment group in the baseline specifications is too low, column 5 includes a regression where the treatment group has been redefined to include only markets with an anticipated post-merger HHI greater than 4000 and an anticipated change in HHI greater than 400. While this cutoff is arbitrary, it splits the original treatment group roughly in half leaving, 1,345 markets.[17] Because the regressions presented in columns 5 include only the distance-time effects, it should be compared to the similarly specified regressions in column 2. As the table indicates, the results in column 5 are statically equivalent to those in column 2, signally that there is no anti-competitive effect even when the criteria for the treatment group are more stringent.

It is also plausible that merger-induced consolidation had a stronger anticompetitive effect in markets not served by Southwest. As discussed previously, numerous studies have shown that Southwest has a particularly strong constraining effect on fares. Potentially, the merger led to a relative fare increase only in markets where Southwest was not present, and hence, grouping all treatment markets together is obscuring this effect. Accordingly, the regressions presented in column 6 are estimated after restricting the sample to markets not served by Southwest. However, yet again, there is no evidence of a price increase. Similarly specified regressions that exclude all low cost carriers yield similar results. It should be noted that all markets served by Southwest—both in the control or treatment group—are excluded from the

[17] More granular subsample analyses along these lines provide no evidence of systematic difference in the merger's effect.

regression in column 6. In addition to testing for a different merger effect, this sample restriction also effectively controls the confounding effect that is captured by the Southwest-time effects in column 4.

3.2. Importance of a flexible time trend

While there is some evidence that relative fares actually fell after the merger, that result is not robust. In particular, after allowing the time trend of fares to vary based on market distance and the premerger level of competition, the merger is no longer associated with a decline in relative fares. Controlling for a unique time trend in markets served by Southwest appears to be particularly important.

Why is it important to control for these potentially confounding factors? Plotting the time fixed effects offers some insight.[18] Figure 2 plots the Southwest-time fixed effects. The magnitude of the coefficients on the Southwest-time dummies, which is represented on the vertical axis, denotes the log-point change in fares *relative* to the markets not served by Southwest.[19] In markets served by Southwest, relative fares declined slightly as the overall level of fares in the industry grew to a peak in 2008:Q3. Then, during the post-2008:Q3 collapse and rebound of overall fares levels, relative fares in market served by Southwest grew steadily. Southwest's pricing behavior during this volatile time period was apparently different than other carriers.

That markets served by Southwest experienced an increase in relative fares and that they

[18] All of the fixed effects plotted in figures 2-4 arise from the regression equation presented in column (4) of table 8.

[19] It is worth keeping in mind the overall trend in fare levels illustrated in the upper panel of figure 1 while interpreting the Southwest-time fixed effect.

are overrepresented in each of the control groups, imparts a negative bias on the estimated difference-in-differences (when the regression model does not include the Southwest-time fixed effects). Markets served by Southwest cause the difference over time in the control groups to be greater, all else equal—making the difference over time in the treatment group appear smaller.

A similar investigation of the estimated distance-time effects is also helpful in explaining their importance. As illustrated in figure 3, the distance-time effects for longer markets tend to decline in the latter periods, particularly the fixed effects for the longest distance markets. Returning to table 2, one can see that including the distance time-effects eliminates much of the negative bias found in column one for control group one. Moreover, including the distance-time effects appears to be less important when using control groups two and three. The likely reason for this is that the distribution of market distance in control groups two and three is similar to the distribution of market distance in the treatment group. On the other hand, in control group one there is a preponderance of markets in the bottom (shortest distance) quartile. Because longer distance routes—which experienced a relative decline in fares over time—are underrepresented in control group one there is a negative bias on the estimated difference-in-differences (in regressions without distance-time effects). When we don't condition on the distance-time effects the change over time in the control group is likely greater than what would have occurred in the treatment markets *but for* the merger.

Finally, plotting the HHI-time fixed effects also helps explain their importance. Comparing column (2) and (3) one can see that controlling for the HHI-time effects makes the estimated difference-in-differences greater. Looking at figure 4, one can see that relative fares were increasing in the more concentrated quartiles. Because all of the control groups tended to

have more concentrated markets than treatment group, when we don't control for this difference in the underlying time trends, the estimated merger effect is again biased downward.

4. Conclusions

This study examines the price effects of the merger between DL and NW, a transaction that was subject to a lengthy investigation by the DOJ and seemingly approved based on the expectation that it would not be anticompetitive. Specifically, this study tests for a change in relative fares in markets where the merger was "likely to enhance market power" under the DOJ/FTC merger guidelines. While there is some evidence of a decline in relative fares after the merger, this result is not robust. Once the empirical model is specified to control for differences in the time trend of fares across markets of differing distance and pre-merger levels of competition, the merger's effect appears benign. In this context, accounting for a unique time trend in market served by Southwest appears to be particularly important.

These results are interesting for a number of reasons. First, the merger was carefully assessed by the DOJ, thus the ultimate effects of the transaction can potentially shed light on the efficacy of a policy decision. Second, numerous earlier studies of airline mergers have found anti-competitive price increases—this transaction does not appear to fit that mold. Finally, the empirical analysis highlights the importance of not only accounting for shocks to supply and demand, but also accounting for differences in the way those shocks are transmitted to prices across treatment and control markets.

References

Ashenfelter, Orley, and Daniel Hosken. “The Effect of Mergers on Consumer Prices: Evidence from Five Selected Case Studies.” *National Bureau of Economic Research Working Paper Series* No. 13859 (2008) : n. pag.

Borenstein, Severin. “Airline Mergers, Airport Dominance, and Market Power.” *The American Economic Review* 80.2 (1990) : 400-404. Print.

Brueckner, Jan K., Lee, D., and Singer, E., “Airline Competition and Domestic U.S. Airfares: A Comprehensive Reappraisal,” (2010).

Busse, M. “Firm Financial Condition and Airline Price Wars.” *The RAND Journal of Economics* 33.2 (2002) : 298–318. Print.

Ciliberto, F., and C. Schenone. “Bankruptcy and Product-Market Competition: Evidence from the Airline Industry.” *MPRA Paper* (2010) : n. pag. Print.

Goolsbee, Austan, and Chad Syverson, “How do Incumbents Respond to the Threat of Entry? Evidence from Major Airlines,” *Quarterly Journal of Economics* 123 (2008), 1611-1633.

Focarelli, Dario, and Fabio Panetta. “Are Mergers Beneficial to Consumers? Evidence from the Market for Bank Deposits.” *The American Economic Review* 93.4 (2003) : 1152-1172. Print.

Kwoka, John, and Evgenia Shumilkina. “The Price Effect of Eliminating Potential Competition: Evidence from an Airline Merger*.” *The Journal of Industrial Economics* 58.4 (2010) : 767-793. Web. 28 Jan 2011.

Morrison, Steven A. “Airline Mergers: A Longer View.” *Journal of Transport Economics and Policy* 30.3 (1996) : 237-250. Print.

Morrison, Steven A., “Actual, Adjacent, and Potential Competition: Estimating the Full Effect of Southwest Airlines,” *Journal of Transport Economics and Policy* 32 (2001), 239-256.

Morrison, S., and C. Winston. *The Evolution of the Airline Industry*. Brookings Institution Press, 1995. Print.

Peters, Craig. “Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry.” *The Journal of Law and Economics* 49.2 (2006) : 627-649. Web. 14 Sept 2010.

Romano, P. S, and D. J Balan. “A Retrospective Analysis of the Clinical Quality Effects of the Acquisition of Highland Park Hospital by Evanston Northwestern Healthcare.” *International Journal of the Economics of Business* 18.1 (2011) : 45–64. Print.

Simpson, John, and David Schmidt. “Difference in Differences Analysis in Antitrust: What Does It Really Measure?” *SSRN eLibrary* (2007) : n. pag.

Simpson, John, and Christopher Taylor. “Do Gasoline Mergers Affect Consumer Prices? The Marathon Ashland Petroleum and Ultramar Diamond Shamrock Transaction.” *Journal of Law and Economics* 51.1 (2008) : 135-152. Print.

Taylor, C. T, and L. Silvia. “Petroleum Mergers and Competition in the Northeast United States.” *FTC Working Paper* (2010) : n. pag. Print.

Taylor, Christopher T., and Daniel S. Hosken. “The Economic Effects Of The Marathon-Ashland Joint Venture: The Importance Of Industry Supply Shocks And Vertical Market Structure —super-*.” *Journal of Industrial Economics* 55.3 (2007) : 419-451. Print.

Thompson, Aileen. “The Effect of Hospital Mergers on Inpatient Prices.” *International Journal of the Economics of Business* 18 (2011) : 91-101(11).

Appendix: DB1B data construction

All three tables from the Origin and Destination Survey (DB1B) are downloaded from the U.S. Department of Transportation website for 2006:Q4 through 2010:Q3. These tables are then merged to create a table that includes a record for each coupon of each domestic itinerary. Data cleaning proceeds as follows. First, itineraries with more than 4 coupons in total or more than 2 coupons on a directional leg are excluded. Second, itineraries involving a domestic portion of an international O&D, or a foreign carrier flying between two U.S. points are excluded. Third, itineraries with a fare in the top-1 percentile of the distribution for the given quarter, less than 20 dollars, or deemed non-credible by the DOT are excluded. Fourth, itineraries involving an airport outside the contiguous 48 states are excluded. Fifth, interline itineraries (those with a change in ticketing carrier) are excluded. Sixth, itineraries that are neither one-way nor round-trip are excluded. Seventh, using the DOT T100 data, itineraries involving a coupon that does not correspond to a flight provided 12 or more times in a given quarter are eliminated. Finally, the data are aggregated by year, quarter, ticketing carrier, and non-directional airport-pair market. Each observation contains a variable for the passenger weighted average directional fare and total passengers for carrier i in market m during year-quarter t. All observations with less than 5 total passengers are dropped. Then all remaining observations associated with non-directional airport-pairs that generate less than 50 passengers in the DB1B sample for a particular quarter are excluded.

Figure 1

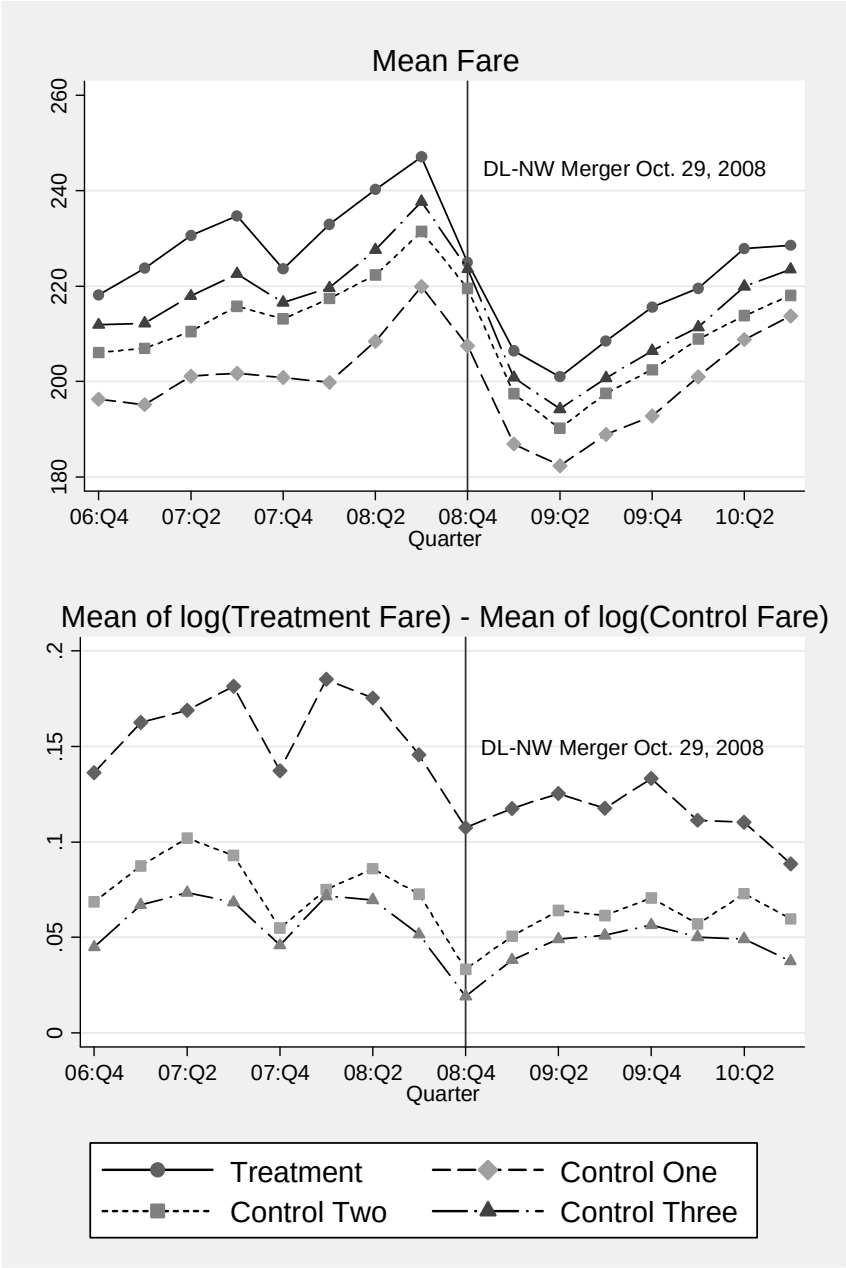

Mean Fare
260
240
220
200
180
DL-NW Merger Oct. 29, 2008
06:Q4
07:Q2
07:Q4
08:Q2
08:Q4
09:Q2
09:Q4
10:Q2
Quarter
Mean of log(Treatment Fare) - Mean of log(Control Fare)
.2
.15
.1
.05
0
DL-NW Merger Oct. 29, 2008
06:Q4
07:Q2
07:Q4
08:Q2
08:Q4
09:Q2
09:Q4
10:Q2
Quarter
Treatment
Control One
Control Two
Control Three

Figure 2: Southwest-time fixed effects

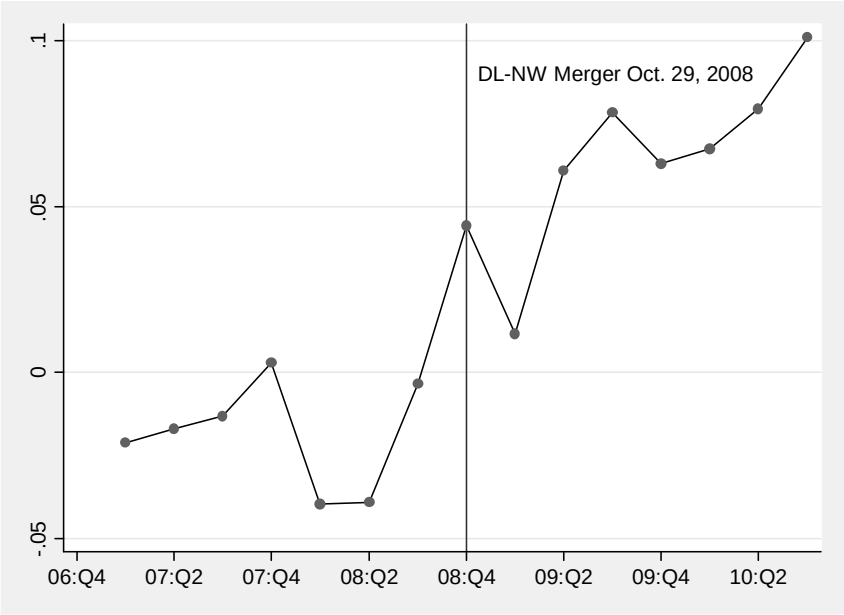
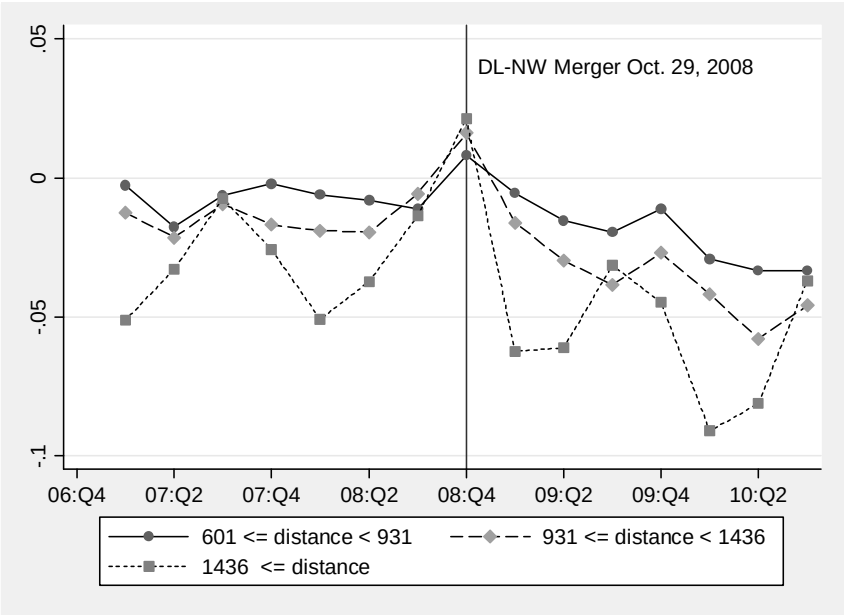
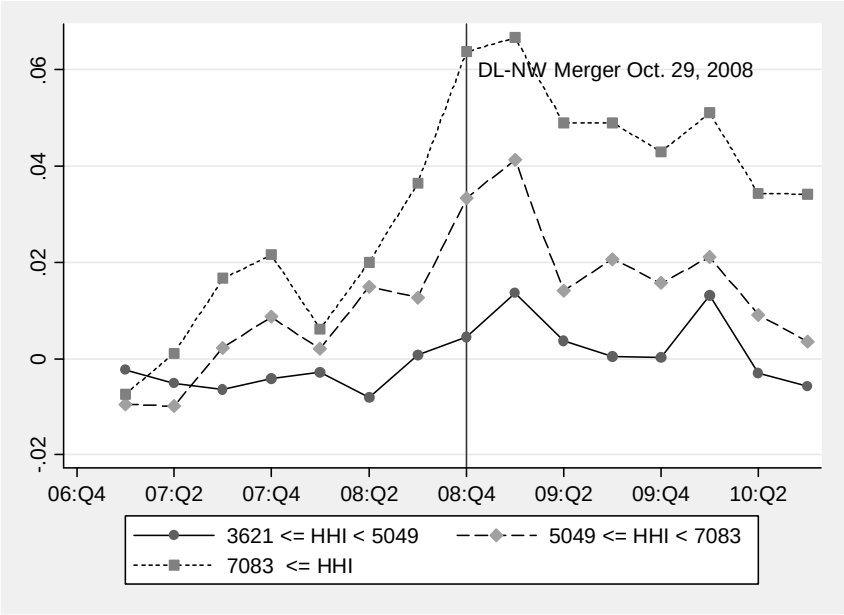


Figure 3: Distance-time fixed effects



Markets less than 601 miles serve as the base case.

Figure 4: HHI-time fixed effects



Markets with a pre-merger HHI less than 3621 serve as the base case.

Table 1: Treatment and Control Group Characteristics

	Treatment group	Control group one	Control group two	Control group three
Number of airport-pair markets	2322	1499	1236	5047
Number of carrier-market combinations	7920	2731	3735	10818
Mean non-stop distance	1107	836	1189	1102
Mean pre-merger fare (nominal US Dollars)	$231	$207	$214	$222
Mean pre-merger HHI	3708	7244	5318	5586
Percentage of markets served by Southwest	9%	27%	33%	22%

Pre-merger fare and HHI are calculated using data from 2007:Q1-2007:Q4. Markets "served" by Southwest are those markets where Southwest had a market share greater than 10 percent in 2007.

Table 2: Number of Airport-Pair Markets by Group, by Quartile

Quartiles	Treatment group	Control group one	Control group two	Control group three
HHI < 3621	1086	21	208	547
3621 ≤ HHI < 5049	809	149	276	1062
5049 ≤ HHI < 7083	332	390	288	1704
7083 ≤ HHI	95	939	464	1836
Distance < 601	456	672	319	1581
601 ≤ distance < 931	725	276	274	1188
931 ≤ distance < 1436	632	324	287	1174
1436 ≤ distance	509	227	356	1206

Table 3: Airport-pair markets, 2007:Q1 — 2008:Q1 and 2009:Q2 — 2010:Q2

	(1)	(2)	(3)	(4)	(5)	(6)
			Control group one			
DLPOST	-0.052***	-0.038***	-0.037***	0.002	-0.039***	-0.011
	[0.007]	[0.007]	[0.010]	[0.010]	[0.009]	[0.009]
RIVPOST	-0.046***	-0.030***	-0.029***	0.007	-0.025***	-0.005
	[0.007]	[0.007]	[0.010]	[0.010]	[0.009]	[0.009]
Observations	77,787	77,787	77,787	77,787	43,754	57,623
Within R-squared	0.076	0.082	0.083	0.089	0.072	0.089
			Control group two			
DLPOST	-0.042***	-0.040***	-0.036***	-0.006	-0.044***	-0.025***
	[0.006]	[0.006]	[0.007]	[0.008]	[0.008]	[0.009]
RIVPOST	-0.036***	-0.032***	-0.028***	0.001	-0.031***	-0.019**
	[0.006]	[0.006]	[0.007]	[0.008]	[0.008]	[0.008]
Observations	87,635	87,635	87,635	87,635	53,602	61,447
Within R-squared	0.073	0.079	0.080	0.085	0.069	0.087
			Control group three			
DLPOST	-0.031***	-0.027***	-0.022***	0.000	-0.030***	-0.014**
	[0.005]	[0.005]	[0.006]	[0.006]	[0.007]	[0.006]
RIVPOST	-0.025***	-0.020***	-0.015***	0.006	-0.017**	-0.008
	[0.005]	[0.005]	[0.005]	[0.006]	[0.007]	[0.006]
Observations	131,311	131,311	131,311	131,311	97,278	91,857
Within R-squared	0.069	0.074	0.075	0.082	0.066	0.084
Distance-time effects	No	Yes	Yes	Yes	Yes	Yes
HHI-time effects	No	No	Yes	Yes	No	No
WN-time effects	No	No	No	Yes	No	No

Cluster robust standard errors in brackets. All regressions include demographic controls, carrier-route fixed effects and year-quarter fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Airport-pair markets, 2006:Q4 - 2008:Q2 and 2009:Q1 - 2010:Q3

	(1)	(2)	(3)	(4)	(5)	(6)
			Control group one			
DLPOST	-0.051***	-0.037***	-0.033***	0.003	-0.038***	-0.010
	[0.007]	[0.007]	[0.009]	[0.009]	[0.008]	[0.008]
RIVPOST	-0.050***	-0.034***	-0.030***	0.003	-0.030***	-0.008
	[0.006]	[0.006]	[0.009]	[0.009]	[0.008]	[0.008]
Observations	110,728	110,728	110,728	110,728	62,531	82,235
Within R-squared	0.083	0.089	0.090	0.102	0.078	0.093
			Control group two			
DLPOST	-0.039***	-0.037***	-0.033***	-0.005	-0.041***	-0.022***
	[0.006]	[0.006]	[0.007]	[0.007]	[0.007]	[0.008]
RIVPOST	-0.038***	-0.034***	-0.031***	-0.004	-0.033***	-0.020***
	[0.006]	[0.006]	[0.007]	[0.007]	[0.007]	[0.008]
Observations	124,643	124,643	124,643	124,643	76,446	87,677
Within R-squared	0.081	0.087	0.088	0.099	0.077	0.092
			Control group three			
DLPOST	-0.028***	-0.025***	-0.020***	0.001	-0.027***	-0.010*
	[0.005]	[0.005]	[0.005]	[0.005]	[0.006]	[0.006]
RIVPOST	-0.027***	-0.022***	-0.018***	0.002	-0.019***	-0.009
	[0.004]	[0.004]	[0.005]	[0.005]	[0.006]	[0.005]
Observations	187,200	187,200	187,200	187,200	139,003	131,462
Within R-squared	0.076	0.081	0.083	0.095	0.074	0.088
Distance-time effects	No	Yes	Yes	Yes	Yes	Yes
HHI-time effects	No	No	Yes	Yes	No	No
WN-time effects	No	No	No	Yes	No	No

Cluster robust standard errors in brackets. All regressions include demographic controls, carrier-route fixed effects and year-quarter fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Airport-pair markets, 2006:Q4 - 2010:Q3

	(1)	(2)	(3)	(4)	(5)	(6)
			Control group one			
DLPOST	-0.052***	-0.041***	-0.036***	-0.004	-0.042***	-0.015**
	[0.006]	[0.006]	[0.008]	[0.008]	[0.007]	[0.007]
RIVPOST	-0.047***	-0.034***	-0.030***	0.001	-0.031***	-0.010
	[0.006]	[0.006]	[0.008]	[0.008]	[0.007]	[0.007]
Observations	126,987	126,987	126,987	126,987	71,792	94,409
Within R-squared	0.091	0.098	0.099	0.109	0.091	0.102
			Control group two			
DLPOST	-0.042***	-0.040***	-0.037***	-0.011*	-0.044***	-0.026***
	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]	[0.007]
RIVPOST	-0.037***	-0.034***	-0.031***	-0.007	-0.033***	-0.020***
	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]	[0.007]
Observations	142,872	142,872	142,872	142,872	87,677	100,627
Within R-squared	0.090	0.097	0.098	0.107	0.090	0.102
			Control group three			
DLPOST	-0.032***	-0.029***	-0.025***	-0.005	-0.032***	-0.015***
	[0.004]	[0.004]	[0.005]	[0.005]	[0.005]	[0.005]
RIVPOST	-0.027***	-0.023***	-0.019***	-0.001	-0.021***	-0.010**
	[0.004]	[0.004]	[0.004]	[0.004]	[0.005]	[0.005]
Observations	214,892	214,892	214,892	214,892	159,697	151,164
Within R-squared	0.088	0.093	0.095	0.105	0.089	0.099
Distance-time effects	No	Yes	Yes	Yes	Yes	Yes
HHI-time effects	No	No	Yes	Yes	No	No
WN-time effects	No	No	No	Yes	No	No

Cluster robust standard errors in brackets. All regressions include demographic controls, carrier-route fixed effects and year-quarter fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 6: City-pair markets, 2007:Q1 - 2008:Q1 and 2009:Q2 - 2010:Q2

	(1)	(2)	(3)	(4)	(5)	(6)
			Control group one			
DLPOST	-0.039***	-0.029***	-0.024**	0.002	-0.040***	-0.009
	[0.008]	[0.008]	[0.011]	[0.011]	[0.009]	[0.010]
RIVPOST	-0.033***	-0.020**	-0.015	0.008	-0.024***	-0.002
	[0.008]	[0.008]	[0.011]	[0.011]	[0.009]	[0.011]
Observations	74,679	74,679	74,679	74,679	39,849	56,013
Within R-squared	0.075	0.082	0.083	0.094	0.074	0.090
			Control group two			
DLPOST	-0.034***	-0.030***	-0.027***	-0.004	-0.040***	-0.019**
	[0.007]	[0.007]	[0.008]	[0.008]	[0.008]	[0.009]
RIVPOST	-0.029***	-0.022***	-0.019**	0.002	-0.025***	-0.012
	[0.007]	[0.007]	[0.008]	[0.008]	[0.009]	[0.009]
Observations	82,874	82,874	82,874	82,874	46,063	59,817
Within R-squared	0.073	0.080	0.080	0.090	0.070	0.089
			Control group three			
DLPOST	-0.028***	-0.023***	-0.019***	-0.000	-0.029***	-0.013**
	[0.005]	[0.005]	[0.006]	[0.006]	[0.007]	[0.006]
RIVPOST	-0.022***	-0.015***	-0.011*	0.005	-0.014*	-0.007
	[0.005]	[0.005]	[0.006]	[0.006]	[0.007]	[0.006]
Observations	131,311	131,311	131,311	131,311	97,278	91,857
Within R-squared	0.069	0.074	0.075	0.082	0.066	0.084
Distance-time effects	No	Yes	Yes	Yes	Yes	Yes
HHI-time effects	No	No	Yes	Yes	No	No
WN-time effects	No	No	No	Yes	No	No

Cluster robust standard errors in brackets. All regressions include demographic controls, carrier-route fixed effects and year-quarter fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 7: City-pair markets, 2006:Q4 - 2008:Q2 and 2009:Q1 - 2010:Q3

	(1)	(2)	(3)	(4)	(5)	(6)
			Control group one			
DLPOST	-0.038***	-0.028***	-0.021**	0.003	-0.038***	-0.007
	[0.008]	[0.007]	[0.009]	[0.010]	[0.008]	[0.009]
RIVPOST	-0.037***	-0.024***	-0.017*	0.005	-0.028***	-0.004
	[0.007]	[0.007]	[0.010]	[0.010]	[0.008]	[0.009]
Observations	106,204	106,204	106,204	106,204	56,921	79,858
Within R-squared	0.082	0.089	0.090	0.101	0.078	0.093
			Control group two			
DLPOST	-0.033***	-0.029***	-0.026***	-0.004	-0.037***	-0.017**
	[0.006]	[0.006]	[0.007]	[0.007]	[0.008]	[0.008]
RIVPOST	-0.032***	-0.026***	-0.023***	-0.004	-0.028***	-0.014*
	[0.006]	[0.006]	[0.007]	[0.007]	[0.008]	[0.008]
Observations	117,884	117,884	117,884	117,884	65,729	85,361
Within R-squared	0.080	0.086	0.088	0.098	0.075	0.093
			Control group three			
DLPOST	-0.026***	-0.021***	-0.017***	0.001	-0.026***	-0.010*
	[0.005]	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]
RIVPOST	-0.025***	-0.018***	-0.014***	0.001	-0.016***	-0.008
	[0.005]	[0.005]	[0.005]	[0.005]	[0.006]	[0.006]
Observations	170,182	170,182	170,182	170,182	122,259	123,624
Within R-squared	0.075	0.080	0.082	0.093	0.072	0.087
Distance-time effects	No	Yes	Yes	Yes	Yes	Yes
HHI-time effects	No	No	Yes	Yes	No	No
WN-time effects	No	No	No	Yes	No	No

Cluster robust standard errors in brackets. All regressions include demographic controls, carrier-route fixed effects and year-quarter fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 8: City-pair markets, 2006:Q4 - 2010:Q3

	(1)	(2)	(3)	(4)	(5)	(6)
			Control group one			
DLPOST	-0.039***	-0.031***	-0.024***	-0.001	-0.042***	-0.011
	[0.007]	[0.007]	[0.008]	[0.009]	[0.007]	[0.008]
RIVPOST	-0.034***	-0.024***	-0.016*	0.004	-0.030***	-0.005
	[0.006]	[0.007]	[0.009]	[0.009]	[0.007]	[0.008]
Observations	121,829	121,829	121,829	121,829	65,379	91,680
Within R-squared	0.090	0.097	0.099	0.108	0.089	0.102
			Control group two			
DLPOST	-0.036***	-0.033***	-0.030***	-0.009	-0.041***	-0.021***
	[0.006]	[0.006]	[0.007]	[0.007]	[0.007]	[0.007]
RIVPOST	-0.031***	-0.026***	-0.023***	-0.005	-0.029***	-0.014**
	[0.006]	[0.006]	[0.007]	[0.007]	[0.007]	[0.007]
Observations	135,171	135,171	135,171	135,171	75,454	97,982
Within R-squared	0.089	0.095	0.097	0.106	0.087	0.103
			Control group three			
DLPOST	-0.029***	-0.025***	-0.022***	-0.005	-0.031***	-0.014***
	[0.004]	[0.004]	[0.005]	[0.005]	[0.006]	[0.005]
RIVPOST	-0.024***	-0.019***	-0.016***	-0.001	-0.019***	-0.009*
	[0.004]	[0.004]	[0.005]	[0.005]	[0.006]	[0.005]
Observations	195,457	195,457	195,457	195,457	140,572	142,194
Within R-squared	0.086	0.091	0.093	0.103	0.085	0.099
Distance-time effects	No	Yes	Yes	Yes	Yes	Yes
HHI-time effects	No	No	Yes	Yes	No	No
WN-time effects	No	No	No	Yes	No	No

Cluster robust standard errors in brackets. All regressions include demographic controls, carrier-route fixed effects and year-quarter fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Chapter 3

The Cost Structure of Regional Transmission Organizations

An essay co-authored with John Kwoka

1. Introduction

A key feature of restructured electricity markets in the U.S. has been the deintegration of generation, transmission, and distribution. While that change has been motivated by the quest for greater efficiency at each stage, deintegration has also created the need for some alternative method of performing the coordination functions that were previously internalized by vertically integrated utilities. In the U.S. that method has focused on regional transmission organizations, or RTOs. RTOs are independent non-profit corporations that have functional control of transmission assets in a region, assets that nonetheless continue to be owned by the utilities that distribute electricity to end users. Most fundamentally, RTOs are intended to ensure equal access to the transmission network as they centrally dispatch electricity over member utilities' transmission systems, but RTOs have also come to organize competitive wholesale markets, coordinate the complex operating decisions of member utilities, monitor wholesale market performance, and coordinate long term investment planning.

Seven RTOs now cover well over one-half of customers and sales of electricity in the U.S.[20] As they have expanded in geographic coverage and functional scope, controversies have arisen about their rapidly growing costs as well as their overall effectiveness. This paper examines the first of these issues, namely, RTO costs. We exploit the fact that existing RTOs began operation at different times and initiated various functions at different times. We estimate the relationship between these functions and several RTO characteristics to differences in their

[20] According to the ISO/RTO Council (IRC), an industry organization consisting of representatives of North American ISO/RTOs, RTOs serve more than 50 percent of Canada's population as well. See, http://www.isorto.org/.

costs and thereby cast light on several questions—the relative costs of each of the market functions administered by RTOs, the relative efficiency of different RTOs in performing the same functions, the possibility of learning economies by RTOs in performing these functions, and economies of scale over the sampled range of RTO "output."

Our statistical results establish the unsurprising fact that each RTO function—administering markets for energy, running capacity markets, developing financial transmission rights, etc.—imposes somewhat different costs on an RTO. Less expected is the finding of cost decline with RTO age, casting doubt on whether there are economies of learning by an RTO, although there is some evidence that total experience of all RTOs collectively may improve efficiency. Further, we document economies of scale with respect to certain aspects of RTO operations, and we can offer some qualified statements about the relative costs of different RTOs.

The plan of this paper is as follows. The next section provides some background on wholesale electric power markets and the origin of RTOs. Section 3 further describes the evolution of existing RTOs and their specific functions. Section 4 sets out the econometric model, hypotheses to be tested, and the data. Section 5 presents the results and section 6 concludes.

2. Background on regional transmission organizations

2.1. *History*

For most of its history the U.S. electricity sector has been dominated by vertically integrated and heavily regulated utilities that bundled and sold generation, transmission, and

distribution services. Beginning in 1978, however, reforms made inroads on this traditional structure, and under the direction of Congress, FERC began creating opportunities for utilities and independent power producers to trade power at market based rates.[21] Its major efforts—first in Order 888 and later in Order 2000—met with only mixed success. Adopted in 1996, Order 888 required utilities to "functionally unbundle," that is, to state separate rates for wholesale generation, transmission, and ancillary services, as well as to take transmission service under the same tariff as offered to its transmission customers. Integrated utilities nonetheless found methods of handicapping or foreclosing independent power producers and generation-dependent utilities from accessing their sections of the transmission network.

In December 1999, FERC issued Order 2000 with the goal that all transmission owners would place their transmission facilities under the control of an RTO. FERC set minimum characteristics and functions for RTOs, namely, (1) independence from market participants, (2) possession of operational authority for all transmission facilities under the RTO's control, and (3) exclusive authority to maintain short-term reliability. FERC argued that these institutions would bring about increased efficiency through regional transmission pricing and regional investment and expansion planning; improved congestion management; and more effective management of physical flows of electricity.[22]

Characteristics of the seven existing RTOs are listed in Table 1**.** All came into operation in the period from 1998 through 2002, but they otherwise vary greatly in terms of the load, miles,

[21] The initial reforms of the industry spawned by the Public Utilities Regulatory Policies Act of 1978 (PURPA) and the Energy Policy Act of 1992 (EPACT 1992).

[22] It is also worth noting that in 2002 FERC more aggressively sought to establish RTOs through a further rulemaking. The plan for a so-called Standardized Market Design (SMD), however, encountered opposition and was abandoned.

number of states, and population served. Despite this diversity in characteristics, some broad trends emerge in the evolution of functions by RTOs. Those RTOs in the Northeast—the New England ISO (ISONE), the New York ISO (NYISO) and PJM[23]—each evolved from power pools and now perform a group of wholesale market functions that closely correspond to FERC's Standard Market Design proposal. Along with the Northeast RTOs, the California ISO (CAISO) and ERCOT in Texas were also early adopters of the RTO model. The Midwest ISO's (MISO) development was slowed by lengthy negotiations among potential stakeholders. Initial operations began in 2002, but MISO did not begin administering market functions until 2005. The Southwest Power Pool (SPP) is the most recently constituted RTO and performs the most limited range of functions.

2.2. *Functions of RTOs*

RTOs are unusual economic institutions. They are not-for-profit corporations that assume control and management of the bulk power transport systems of their member utilities, while the latter continue to own all of those assets. RTOs are self-governed companies directed by boards that are both independent of market participants but in many cases intended to reflect all stakeholders. RTOs raise capital and receive revenues for transmission services they provide. Their most basic role is to administer the terms and conditions of transmission services and maintain the short-term reliability of the network, but as will be described below, they also perform a number of other functions. Notably, however, RTOs are not responsible for the

[23] PJM serves Pennsylvania, New Jersey, and Maryland, as well as Delaware, Illinois, Indiana, Kentucky, Michigan, North Carolina, Ohio, Tennessee, Virginia, West Virginia and the District of Columbia.

maintenance and repair, or the fixed investment costs,[24] of the transmission facilities over which they direct the flow of power. This division of responsibilities, as well as the separation of ownership and control, makes RTOs quite unique.

RTOs perform various specific functions, each of which can be expected to give rise to incremental costs. These functions are as follows:

2.2.1. Energy markets

All RTOs organize markets for electrical energy. The most basic approach involves a real-time energy market that resolves discrepancies between real-time demand from load-serving entities and supply commitments from producers. We will refer to this approach as the single-settlement energy market. Some RTOs, including ISONE, NYISO, PJM, and MISO, have implemented two-settlement energy markets, where the RTO operates a day-ahead forward market in addition to the real-time market. The day-ahead market is intended to facilitate reliability and efficiency by allowing the RTO to secure adequate supplies to meet forecasted demand ahead of real-time, and to provide buyers and sellers the opportunity to hedge real-time price volatility.

2.2.2. Ancillary service markets

Since electricity cannot be effectively stored, system operators keep an inventory of capacity—that is, generators--available to call on when system conditions change. Examples of these so-called ancillary services include spinning reserve, non-spinning reserve, reactive power,

[24] The role of RTOs in investment planning has emerged as a major point of contention. See Wu, Zheng, and Wen (2006) for a discussion of the challenges of transmission investment and expansion planning in a restructured electricity market.

and regulation, each of which represents somewhat different degrees of contingent-available supply to the RTO. Ancillary services can be procured by the RTO via cost-based payments, or the RTO may organize a market where generators bid to supply one or more of such services. For present purposes all of these variations will be considered as a single "ancillary services" market function.

2.2.3. Transmission Rights

To address possible network congestion, RTOs must have a system for allocating scarce network capacity. Some RTOs, including the Northeastern RTOs and MISO, employ "nodal pricing" in which participants submit bids to buy or sell electricity at numerous specific network locations or "nodes." Other RTOs use a broader "zonal" pricing system within which a uniform price applies in a specified region. To address the inevitable uncertainty with respect to price, RTOs have created various forms of transmission rights which are auctioned monthly and yearly. These so-called financial transmission rights (FTRs) permit all market participants to hedge the price uncertainty associated with market access.

2.2.4. Resource adequacy and capacity markets

Although settlement markets were originally believed sufficient to guide market participants with respect to capacity investment, in practice this has not occurred. As a result RTOs have devised various supplementary incentives to foster investment. Notably, both of the Northeastern RTOs have implemented markets specifically for capacity in which demand-side participants are required to procure a sufficient portfolio of generating capacity to meet a resource adequacy requirement. Since other RTOs have no such formal institutional

mechanism,[25] this variation in approaches facilitates isolating the incremental costs of capacity markets.

2.3 Issues in RTO markets.

In its original rulemaking on RTOs, FERC asserted that RTOs would be a cost effective method of achieving their important purposes. FERC staff estimated that formation of RTOs would result in cost savings of $2.4 billion annually.[26] Moreover, while acknowledging nontrivial costs, FERC predicted that later forming RTOs would experience learning economies from observing the initial RTOs. FERC explained that having the opportunity to observe "what works and what does not with respect to regional transmission entities . . . should make it somewhat easier, and more cost efficient, to create new RTOs." [27] Based on this assessment, it claimed that RTOs would confer considerably greater benefits than their costs.

Concerns about escalating RTO costs soon emerged, however. These prompted FERC to issue a report on the costs of a minimalist "Day One RTO," that is, one that ran only a real-time market and did not administer wholesale markets. It estimated such an RTO would necessitate a 35 to 75 million dollar annual revenue requirement. This report came under criticism from industry participants for its narrow focus on real-time markets—all RTOs performed additional functions—and for its overly favorable conclusions. Studies by Lutzenhiser (2004) and the Public Power Council (2004), among others, suggested that RTOs were incurring significantly

---

[25] Other RTO approaches range from energy-only (that is, no capacity mechanisms, as illustrated by ERCOT) to capacity requirements but bilateral procurement (CAISO and SPP).

[26] Order No. 2000, *Regional Transmission Organizations*, F.E.R.C. STATS. & REGS, p. 94.

[27] *Id.* at pp. 70-89.

higher costs, quite possibly exceeding even FERC's own estimates of RTO benefits.[28]

In May of 2007, the U.S. Government Accountability Office (GAO) undertook its own review of RTOs. The GAO reported concerns over swelling RTO costs, a second-best governance structure, and an apparent inability of RTO to run wholesale markets to stimulate adequate investment in transmission and generation capacity.[29] The GAO recommended that FERC systematically review RTO budgets and annual financial reports, and develop standardized measures to track RTO performance.[30] Simultaneously, many observers were concluding that regardless of costs RTOs were not performing their intended functions well, citing evidence of continuing congestion problems and inadequate investment.[31]

3. Empirical model

The purpose of this study is to advance our understanding of RTO costs beyond the relatively simplistic evaluations heretofore conducted. We do so by bringing a standard economic framework to bear on the relationship between the costs of and functions performed by RTOs. That framework is the empirical model of a Cobb Douglas cost function, which assumes total costs are log-linear in input prices and output. For present purposes we allow the intercept term to shift as the RTO adds market functions or achieves economies of learning. This permits inferences about the incremental costs of each function.

---

[28] Two consultancies, Christensen Associates (2007) and GDS Associates (2007), have also documented the significant operating expenditures of RTOs, noting that costs, in dollars per megawatt terms, vary widely across RTOs.

[29] Joskow (2005a) for a discussion of the governance challenges surrounding RTOs and Joskow (2006) for a further discussion on the challenges in regulating transmission and distribution networks.

[30] See, Drom (2007).

[31] Other analyses of the benefits and cost of RTOs can be found in Blumsack (2007), Christensen Associates Energy Consulting (2007), and Lenard and McGonegal (2008). Criticisms of RTOs have slowed their expansion as policymakers search for an alternative model for performing the necessary functions.

The empirical cost function takes the following form:

$$\ln(TC_{it}) = \mu_t + \gamma' E_{it} + \delta' M_{it} + \beta_1 ln(Q_{it}) + \beta_2 \ln(w_{it}) + \eta_i + \varepsilon_{it}.$$

In this formulation, μ_t is a constant which is allowed to vary over time as technology changes, E_{it} measures experience both within and across RTOs, M_{it} is a group of indicator variables representing market functions, Q_{it} and w_{it} are output and wage, respectively, and η_i represents RTO fixed effects. In this model, the intercept, μ_t, may vary over time due to changes in national RTO policy, such as FERC rulemakings and policy statements that potentially alter RTO behavior or production technology. The RTO fixed effects measure time-invariant differences in efficiency among the various RTOs.[32]

Before discussing the parameters of primary interest, we note that the estimated elasticity of cost with respect to the wage is captured by β_2, and returns to scale can be measured as $1/\beta_1$. There is no variable for the price of capital since the majority of an RTO's capital investments are in the development of proprietary software systems used to operate wholesale markets. As a result, it is not possible to construct a relevant variable for the price of capital. However, using the generic property that a cost function is linearly homogenous in factor prices, the implied elasticity of total costs with respect to the price of capital is $1 - \beta_2$.[33]

One of the primary goals of this study is to estimate the incremental impact of market functions on RTO costs. This impact is captured by δ, the coefficient vector on M_{it}. The diverse

[32] Of course, to the extent that there are other unobserved or difficult-to-quantify differences across RTOs, the fixed effects terms will also reflect those factors.

[33] A reviewer has raised a possible concern about endogeneity of RTO formation. Casual inspection of trade publications and FERC rulemakings indicates that RTOs' initial designs arose from negotiations among stakeholders, each region's historical industry structure, and some attention to performance. See, for example, Joskow (2005b). Accordingly, we believe that changes in market design in our sample period have been largely the product of FERC rulemaking and other mandates exogenous to any particular RTO.

groupings and implementation timelines of the various market functions across RTOs provide sufficient variation to identify the relative costs of each market function. In our most comprehensive specification, δ captures the marginal effect of market functions on costs, after controlling for RTO-specific effects, learning economies, and industry-wide changes in technology.

A further purpose of this study is to evaluate possible economies of learning, that is, reductions in costs due to experience with the production process.[34] The coefficient vector γ captures two facets of the phenomenon of "experience." The first is simply the count of the number of years of operation by an RTO, i.e., its age, which should measure its "internal" experience. One potentially confounding influence is that an RTO's age will also capture any other age-related cost forces, including some (like updates, corrections, and revisions of RTO institutions and methods) that may cause costs to rise rather than fall. To that extent, the estimated effect of age would be the net effect of learning and such other factors. "External" experience is given by the total number of years of operation across all RTOs, that is, the sum of all RTOs' ages in a particular year. To further examine learning, we later employ net energy transmitted to load, rather than years of operation, as a measure for both internal learning and external learning.

Finally for RTO output or scale, we initially employ a measure of system load, namely, the annual net energy transmitted to load. This measure is most useful since it is homogenous

[34] Recall that FERC presumed that learning economies would result in declining costs over time. Relevant studies of learning economies include Thornton and Thompson (2001), Gruber (1998), and Irwin and Klenow (1994).

across RTOs and has a straightforward interpretation in the present context. To further capture variation in the geographic scale of RTOs, we add a measure of the miles of transmission line overseen by the RTO. We view this as the best proxy variable to capture differences in costs associated with expansion of an RTO's geographic footprint, as was the case with MISO and PJM, versus simple load growth over a fixed transmission network.

First-order autocorrelation in all specifications has been corrected using a feasible GLS estimator. The estimated variance matrix is constructed from the residuals of an OLS regression, allowing for a unique variance and autocorrelation parameter in each panel.

4. Data

The analysis in this study is based on an unbalanced panel of annual data for the RTOs operating within the United States from 1998 to 2008.[35] The data set is a unique combination of data from FERC and official documents available on the RTOs' websites. The variable $OPEXP_{it}$ represents the natural log of total operating expenses. Total operating expenses are the sum of operations, maintenance, and depreciation expenses.[36] While RTOs act as an intermediary in the settlement process of members' transactions for energy, ancillary services, and related transmission products, and the financial results of these transactions are not included in the RTO financial statements. The primary categories of expenses are: salaries and related benefits;

[35] We also experimented with inclusion of data from the RTOs operating in Canada. Unfortunately we found these institutions were likely described by meaningfully different production technologies, potentially due their composition of fewer transmission owners or different geography of their bulk transmission grids.

[36] Except for ERCOT, which operates solely within the borders of Texas, all of the RTOs are subject to FERC jurisdiction and must submit annual financial statistics via the Form 1 electric utility annual report. These Form 1 data serve as our primary source for annual operating expenses. Data for ERCOT and for Southwest Power Pool (prior to being officially recognized by FERC in 2004) were collected from publicly available annual reports.

depreciation and amortization; consulting and other outside services; facility and maintenance; regulatory dues and assessments; and interest expenses. Figure 1 illustrates the general trend of $OPEXP_{it}$ by RTO, Apart from the case of California, these expenses have risen considerably for all RTOs over the ten-year period. Expenses at PJM and MISO show the most dramatic increases, generally in the years 2002 through 2004 coinciding with their geographic expansions. Most RTOs report some leveling off of expenditures after 2006.

The dummy variables that describe the grouping of markets operated by RTO i in year t were constructed based on information from RTO websites. These variables assume a value of unity for any year t in which the RTO performed the specified market function. When an RTO begins operation during month k of year t, the dummy variable takes on a value of $(12 - [k-1])/12$, that is, the fraction of the year in which it did operate.[37] Moreover, whether early or late in its first calendar year, any RTO is likely to incur costs for some time prior to the inauguration of operations. Accordingly, we define a variable $MKTPREP_{it}$ for the year of and the year before the RTO prepared for market operation. This variable equals the fraction of each of those years represented by 12 total months in advance of first operation. Thus, if an RTO implements a market during month k of year t, then $MKTPREP_{it} = (k-1)/12$ in year t and $MKTPREP_{it} = (12 - [k-1])/12$ in year t-1.

We also note that in a number of cases the initial design and implementation of a function may have been modified after inauguration in order to improve its performance (see, e.g., Chandley and Hogan (2009)). Reliable measures of these redesign costs, however, are not

[37] PJM implemented a regulation market in June of 2000 and a market for reserves in December of 2002. As a compromise, we set the *ANCILL* equal to one for all years after 2000 in PJM.

available, but their implications for interpretation of the results should be borne in mind.

With these caveats, Table 3 describes the implementation timeline of market functions for each RTO. The first variable, $REALTIME_{it}$, indicates operation of a real-time energy market. $DAYAHEAD_FTR_{it}$, equals one if RTO *i* administered a day-ahead energy market and nodal financial transmission rights in year *t*.[38] We have had to combine the day-ahead energy market and nodal transmission rights functions into a single variable because these functions were initiated simultaneously in all cases.[39] $ZONAL_FTR_{it}$, equals one in year *t* for RTO *i* if it administered zonal financial transmission rights. Finally, $ANCILL_{it}$ and $CAPACITY_{it}$ indicate operation of ancillary service and capacity market functions, respectively.

The variable $LOAD_{it}$ represents scale of output, measured as the log of annual net energy for load (GWh).[40] $WIRES_{it}$ is the natural log of the total miles of transmission line overseen by the RTO.[41] The trends for these measures of scale are respectively illustrated in figures 2 and 3. Both figures clearly illustrate the expansion of PJM and MISO from 2002 to 2005. The variable $WAGE_{it}$, the log of wages, was constructed from Bureau of Labor Statistics data on the average annual pay of workers in the electric power generation, transmission, and distribution sector (NAICS 2211) for the state where the RTO's headquarters is located. Both wage and operating expense data are deflated to year-2005 dollars with the GDP deflator.

---

[38] All RTOs that implemented a day-ahead energy market have utilized nodal pricing.

[39] PJM implemented a day-ahead market for energy 13 months after implementing financial transmission rights.

[40] This variable was constructed using annual state of the markets reports and load forecast data available on the RTO websites. For MISO we also relied on FERC's market oversight web page and the 2006, 2007, and 2008 Form 1 filings.

[41] This variable was constructed using annual reports and other documents available on the RTO websites. SPP provided us with data for 2006-2010, however, that organization did not record miles of transmission line in prior years. Accordingly, the miles of transmission line from 2001-2005 were imputed based on the average annual change from 2006-2010.

We seek to capture two dimensions of RTO experience, and each is itself measured in two alternative ways. The two dimensions of an RTO's cumulative years (that is, its age) and also the cumulative years of all RTOs up until that time. Accordingly, $RTOAGE_{it}$ is defined as the log of one plus the age of RTO *i* as of year *t*, and measures its own experience. A second variable, $TOTALYEARS_t$, is defined as the natural log of the sum over all *i* of $RTOAGE_{it}$ during year *t*. This measures the collective experience or "external" learning of all RTOs. The alternative measures are based on load rather than years, and represent a different possible metric for any possible learning. Thus, $CUMLOAD_{it}$ is defined as the cumulative sum of $LOAD_{it}$ over the current and all prior years for RTO *i*, while $TOTALLOAD_t$ is the natural log of the sum over all *i* of $CUMLOAD_{it}$ during year *t*. These measure internal and external learning, respectively.

Summary statistics for all of the non-binary variables are reported in Table 3. As this table makes clear, cost and operating data do not differ greatly among RTOs, but the timing of implementation does differ, affording the opportunity to disentangle total costs.

5. Results and Discussion

Because of the limited size of our dataset, we begin with a relatively parsimonious specification, which serves to validate the basic model. We then add covariates sequentially in order both to determine their impact on costs but also to ensure they do not fundamentally alter the core results involving market functions and costs. Finally, we explore an alternative specification of RTO "output," in part because of its intrinsic interest but also in a further effort to avoid any distortion involving the estimated effects of market functions. This approach produces a set of results that addresses the key questions posed at the outset while identifying potential sources of bias and appropriately tempering the conclusions.

5.1 Basic Results

The basic results are reported in Table 4. Column (a) contains the simplest specification, with a series of variables for market functions. Column (b) adds dummy variables for each individual RTO. Then columns (c) and (d) add variables for learning/experience economies. All four models contain the same control variables, which we discuss first.

As noted previously, the most straightforward measure of cost-causative output is load. The variable $LOAD_{it}$ is positively and highly significantly related to RTO costs in all of these regressions until other load-related variables are introduced and "steal" statistical significance from $LOAD_{it}$ itself. The magnitudes of the estimated coefficients suggest substantial economies of scale, an issue to which we shall return. The factor price $WAGE_{it}$ is positively related to costs, as expected, but it is statistically significant in only one specification. Unmeasured time-varying effects such as technology or regulatory changes are captured in the linear time trend $YEAR_t$, which is positive and highly significant in the first two models. Again, in the presence of other time-related variables in models (c) and (d), $YEAR_t$ loses its significance.[42]

The other variable of interest is $MKTPREP_{it}$, which (as discussed above) measures the cost impact of preparing to run a new RTO. In addition to substantial software development costs, RTOs typically set up and test markets for a period before going "live." The empirical results confirm that there are notable costs in the year prior to officially settling transactions in RTO markets, as the estimated coefficient on $MKTPREP_{it}$ is positive and statistically significant

[42] The use of the simple trend variable YEAR is preferable to a series of individual year dummies, both because it conserves on degrees of freedom and also since use of the latter would not permit estimation of time-dependent learning economies at all. Results on other variables of interest are not fundamentally different in either specification.

throughout.

The variables of primary interest concern the market functions and their relationship to RTO costs. The results in Column (a) show that $REALTIME_{it}$, $DAYAHEAD_{it}$, and $ANCILL_{it}$ all have the expected signs and pass the test of statistical significance at the five percent level or better. While $ZONAL_FTR_{it}$ is statistically insignificant in this model, all subsequent results suggest this is due to some omitted variable bias since (as we shall see) it becomes significant upon the addition of further explanators. The statistically insignificant coefficient on $CAPACITY_{it}$ here and throughout is unexpected. While taken literally this would seem to suggest that capacity markets have a trivial cost impact, the more likely explanation is collinearity. Capacity markets have only been implemented by the Northeastern RTOs, and $CAPACITY_{it}$ is highly correlated with other market function dummies.

The model in Column (b) adds to the previous specification a series of dummy variables for the individual RTOs. These variables control for time-invariant RTO effects and shed light on the relative efficiency of each RTO.[43] PJM is the omitted RTO, so all the remaining RTO effects are estimated differences relative to PJM.

The results suggest that MISO may be the most costly of the RTOs, with a large, positive, and statistically significant difference in its costs compared to PJM and, by implication, the other RTOs. There is a considerably weaker indication that ERCOT may be less costly than other RTOs, with its dummy variable negative but not rising to standard levels of statistical significance. The California, New England, New York, and Southwestern RTOs all have

[43] Of course, to the extent that other unobserved RTO-specific factors are important, this set of dummy variables cannot separate the effect of those factors relative to their efficiency.

indicator variables that are positive, implying higher costs than PJM but lacking statistical significance. We conclude that PJM and ERCOT seem clearly the most efficient, with MISO the least so, and all others possibly a bit less efficient than the PJM-ERCOT standard.

We further note that in this and all subsequent models, the variable $ZONAL_FTR_{it}$ is positive and statistically significant. Its emergence in models with RTO dummies suggests that the Column (a) result was the result of the effect being obscured in the "average" RTO cost experience. One policy implication from the present result is that the single settlement zonal energy markets used in Texas and California appear to be more costly than the bi-settlement nodal markets employed in the Northeast and Midwest, all else equal.[44] The estimated coefficients in column (b) imply that implementing a two-settlement energy market with nodal prices increases costs 21 percent over a single real-time energy market, while the zonal market approach increased costs by 39 percent[45]. Interestingly, CAISO implemented a two-settlement nodal energy market on April 1, 2009, and ERCOT and SPP both have plans to add nodal day-ahead energy and ancillary service market functions.

Finally, we explore possible learning economies. Column (c) adds the variables $RTOAGE_{it}$ which captures any efficiency gains associated with an RTO's own experience—but

[44] Due to the ensuing energy crisis of 2000-2001, on August 23, 2000 FERC initiated a formal investigation into the electricity rates in California. This investigation resulted in a series of orders from FERC which prompted the redesign of the California's wholesale electricity markets. In our empirical analysis, to control for costs associated with CAISO's market redesign, we have treated each implementation phase as if it were a new market function. Thus, for CAISO, *MKTPREP* takes a value of 10/12 during 2002 and a value of 2/12 during 2001 to capture Phase IA, implemented on November 1, 2002. On October 1, 2004 CAISO implemented Phase 1B. Accordingly, MKTPREP equals 9/12 during 2004 and 3/12 during 2003 for CAISO. For further information see, "Addressing the 2000-2001 Western Energy Crisis" http://www.ferc.gov/industries/electric/indus-act/wec.asp), as well as, the 2002 and 2004 versions of the Annual Report on Market Issues and Performance, Department of Market Analysis, California ISO.

[45] The marginal effect is approximated using the formula $\exp(\hat{\delta}) - 1$, see Kennedy (2003) p. 254.

as noted above, any other possible time-related effects as well. The specification also includes $TOTALYEARS_t$, number of years of operation and experience by all RTOs collectively, which measures external learning arising from aggregate industry experience. We might expect costs to fall as either or both of these variables increase, since they represent added learning and experience.

These expectations are only partly borne out. The variable $TOTALYEARS_t$ has a negative and statistically significant sign, indicating that greater cumulative experience by all RTOs collectively results in significantly lower costs, cet. par. By contrast, $RTOAGE_{it}$ appears with a positive and statistically significant sign, indicating that after controlling for industry-wide temporal changes, each RTO's costs actually increase for every additional year of operation. As noted before, other unobserved RTO functions correlated with years of operation are likely confounding these estimates, but it can be said that there is simply no good evidence of internal learning in these data.

5.2 Economies of scale

In the first three models in Table 4, the estimated coefficients on $LOAD_{it}$ range from 0.44 to 0.55. All of these estimates provide evidence of increasing returns to scale, as a value of one falls well outside the parameter's ninety-five percent confidence interval in each case. While increasing returns are not surprising, the very large returns implied by a value of 0.5 are unexpected, even unlikely. Taken literally, that value implies that a doubling of output as measured by $LOAD_{it}$ results in only a fifty percent increase in costs. That implication prompts further attention to this issue.

The RTOs in our data base have in fact undergone growth in two quite different dimensions. The first and most obvious is simply that due to increased demand and more intensive usage of a given network. But the second source of growth is illustrated by PJM and MISO, both of which expanded their geographic footprints during the sample period. It is not at all clear that these two types of growth have the same cost implications. We therefore distinguish the two by adding a variable for a characteristic of output that differs between geographic expansion and pure demand expansion. That variable is miles of transmission line operated by the RTO. The variable $WIRES_{it}$ is largely unchanged in the case of pure demand expansion, but obviously increases for geographic expansion.

Table 5 reports a set of regressions that tests for the possibly different costs implications of $LOAD_{it}$ and $WIRES_{it}$. In each case it proves useful to include the interaction term $LOAD_{it}$•$WIRES_{it}$ to capture cost complementarities from simultaneously having both greater load and more line miles.[46] We note preliminarily that the limited number of observations and high correlation between the measures of scale, yields imprecise estimates. The bivariate correlations between, $LOAD_{it}$, $WIRES_{it}$, and $LOAD_{it}$•$WIRES_{it}$, are no less than 0.82. Nonetheless, this is a reasonable approach to attempt to control for difference between the cost effects of load growth over a fixed network and load growth due to geographic expansion.

It is reassuring that our fundamental results concerning the costs of market functions and learning economies are unchanged. The statistical significance of the function variables actually increases in the presence of this more complete specification of output. With the exception of

[46] *LOAD•WIRES* is akin to economies of scope from the joint effect of *LOAD* and *WIRES*.

$CAPACITY_{it}$, all market function variables are now statistically significant. The coefficients on the year-based measures of experience, just as before, both are statistically significant. Greater $RTOAGE_{it}$ causes costs to rise and $TOTALYEARS_t$ of all RTO operation is associated with a decline in costs. Similarly, both $CUMLOAD_{it}$ and $TOTALLOAD_t$ are highly significant, with signs again indicating the RTO load is capturing some non-learning effects of an RTO's cumulative operation that cause its costs to rise, while $TOTALLOAD_t$ continues to imply that costs fall with the total experience (now measured by load) of all RTOs collectively. Additionally, the coefficients on the RTO dummies continue to indicate that MISO is less efficient than the benchmark and there is also some additional evidence the NYISO may be more costly and ERCOT less costly.

While in most cases the individual coefficients on $LOAD_{it}$, $WIRES_{it}$, and $LOAD_{it} \bullet WIRES_{it}$ are statistically insignificant, the joint effect of the three variables is highly significant in both regressions. With a chi-squared statistics of 69.07 and 52.53, in columns (a) and (b) respectively, in both cases we can soundly reject the null hypotheses that all three coefficients equal zero at the one percent level of significance. Clearly, the output variables collectively add explanatory power to both specifications and are important predictors of costs.

With two measures of output we can adopt a multiple-output measure of scale economies—ray economies of scale—which implies the change in cost associated with a proportionate change in both outputs.[47] The degree of ray economies of scale is equal to one divided by the sum of the output-specific elasticities, thus, a value of one indicates costs increase

[47] See Bailey and Friedlaender. (1982).

proportionately when both outputs increase by that same proportion. Values greater than one indicate increasing returns to scale, and values of less than one indicate decreasing returns to scale. Bearing in mind the large confidence intervals, when evaluated at the mean values of $LOAD_{it}$ and $WIRES_{it}$ the point estimates in columns (a) and (b) respectively indicate ray economies of degree 1.26 and 1.53 (cost elasticities with respect output of 0.79 and 0.65). The point estimates also indicate evidence of cost complementarity in the sign of the interaction term.[48]

An understanding of returns to scale is of practical importance since RTOs change, or have been proposed to change, in size. In 2001 FERC directed the three Northeast RTOs to enter negotiation to form a single Northeast RTO. The proposal faced opposition from stakeholders and state regulators, and eventually faded along with FERC's SMD initiative that had encouraged geographical expansion. Other attempts have met with greater success. MISO gradually increased its membership from 22 transmission owners in 2002 to 29 transmission owners in 2005, adding approximately 20,000 miles of the transmission line to its footprint. In 2004 PJM expanded westward from its original footprint into Ohio, Kentucky, Virginia, West Virginia, Illinois, Indiana and Michigan. PJM and others have argued that this expansion served to lower operating costs and created new opportunities for gains from trade.[49] Our results suggest that there may well be merit in these efforts.

---

[48] While the implied elasticity of cost with respect to load is negative over much of the range of $WIRE_{it}$ we interpret theses values as zero given the large standard errors and the logical inconsistency of negative values.

[49] While the gains from trade seem indisputable, improvements in operating efficiency remain unproven.

5.3 Policy simulations

Having used the model to address several questions surrounding the past cost experience of RTOs, we can also use the results to estimate the likely costs of various hypothetical RTOs, defined in terms of sets of market functions. In particular, we employ the parameter estimates from column (a) in Table 5, together with mean values of the control variables, as the basis for predicting the total costs of groups of functions that might be performed by RTOs in 2008.

As shown in table 6, a base-case RTO that is simply an open-access transmission provider has total costs of approximately $55 million dollars. Costs increase by $12 million per year, when that RTO adds a real-time energy market and by another $66 million, with an ancillary services market. The remaining cost increment results from the nature of the settlement process. Based on the experience of CAISO and ERCOT, the model predicts a single settlement process with zonal pricing and transmission rights costs $77 million per year. While nodal settlement costs cannot be disentangled from the costs of a day-ahead market, since in all cases both were adopted simultaneously, their combined cost is approximately $38 million. The total simulated costs of the average full-service RTO is therefore in the range of $130 million to $210 million per year.

A further use of this model is to predict the cost effects of hypothetical mergers of RTOs. As noted above, FERC sought a merger of the New England and New York ISOs, and indeed in 2002 those entities filed plans for a proposed merger with FERC. The first specification in Table 5 can again be used to predict the costs of a combined ISONE and NYISO. Based on 2008 operations, that number is $214 million. This should be compared to actual costs in 2008 of $117 million and $134 million for New England and New York, respectively, for combined total

operating expenditures of $251 million**.** Part of the reason for the relatively modest economies, is the higher concentration of load per mile of transmission line found in the Northeast relative to the rest of the sample. Because of this, the interaction term $LOAD_{it} \bullet WIRES_{it}$ is relatively small, increasing the elasticity of total costs with respect to both measures of output. This exercise underscores the uniqueness of each RTO's experience.

6. Conclusions

This analysis has achieved two important objectives. First, it moves the debate over the costs of RTOs beyond casual empiricism, graphical comparisons, and conjecture. We have imposed a theoretically appropriate framework on the cost question and tested it rigorously with the best data available. To the best of our knowledge, this has not previously been done. Secondly, despite limitations of the data, we obtain plausible statistical results, in some cases confirming beliefs about RTO costs, while elsewhere finding less expected outcomes.

Least surprisingly, we find that RTO costs are directly related to the number of market functions performed. All specifications establish that market functions for real-time energy, congestion management (either zonal or nodal transmission rights), and ancillary services each increase costs. In sum, we find convincing evidence that moving beyond their fundamental role as open-access transmission service provider causes RTOs to incur notable added costs.

Among the less expected results, we find no evidence of "internal" learning economies, that is, cost benefits associated with more years of operation of an RTO. We note that other time-dependent factors may be obscuring any learning economies. The evidence with respect to "external" learning is more compelling: Individual RTO costs appear to decline with greater cumulative experience among all RTOs. These results persist when RTO experience is measured

by cumulative load as well as by cumulative time (i.e., age).

We also find that RTOs exhibit economies of scale when output is measured as total net energy transmitted to load. Employing a more flexible cost function that includes both load and miles of transmission line as measures of scale, the parameter estimates are statistically insignificant, yet nonetheless consistent with economies of scale in a multiple output context, i.e. ray economies of scale.

Finally, examination of relative efficiency across RTOs suggests that MISO incurs significantly greater cost for its size and set of market functions. There is also some evidence that NYISO incurs greater costs, all else equal. Only ERCOT records lower costs than PJM, however its coefficient is not statistically significant in all specifications. All the other RTOs have costs insignificantly different from PJM, suggesting the apparent differences in their costs of operation are due to the functions each perform, their longevity of operation, and exogenous factors such as input costs.

Regional transmission organizations are critically important to the restructured electric power sector, but they have drawn increasing scrutiny of their costs, governance, and performance. That scrutiny has not been accompanied by much in the way of systematic evidence. This paper has developed a framework for analysis of RTO costs and has sought to cast light on the major questions raised about their cost experience. While some conclusions remain tentative, our systematic approach can hopefully serve as a foundation for further analysis of regional transmission organization and their role in liberalized electricity markets.

References

Bailey, Elizabeth E., and Ann F. Friedlaender. (1982). “Market Structure and Multiproduct Industries.” *Journal of Economic Literature* 20.3: 1024-1048

Barrios, S. and E. Strobl (2004). “Learning by Doing and Spillovers: Evidence from Firm-Level Panel Data from Firm-Level Panel Data.” *Review of Industrial Organization*, 25:175-203.

Blumsack, S. (2007). “Measuring the Benefits and Costs of Regional Electric Grid Integration.” *Energy Law Journal*, 28:147-184.

Chandley, John and Hogan, William (2009). “Electricity Market Reform: APPA's Journey Down the Wrong Path.”

Christensen Associated Energy Consulting, LLC. (2007). *The Regional Transmission Organization Report Card: Wholesale Electricity Markets and RTO Performance Evaluation Second Edition*, prepared for the National Rural Electric Cooperation Association. August 20, 2007.

Drom R.A. (2007). “New Metrics for Measuring the Success of a Non-Profit RTO." *Energy Law Journal*, 28:603-630.

GDS Associates, Inc. (2007). *Analysis of Operational and Administrative Costs of RTOs*, prepared for the American Public Power Association.

Gruber, H. (1998). “Learning by Doing and Spillovers, Further Evidence from the Semi Conductor Industry.” *Review of Industrial Organization*, 13:697–711.

Irwin, D. A. and P. J. Klenow (1994). “Learning-by-Doing Spillovers in the Semiconductor Industry.” *Journal of Political Economy*, 102:1200–1227.

Joskow, P.L. (2006). “Incentive Regulation in Theory and Practice: Electric Transmission and Distribution Networks.” *National Bureau of Economic Research Economic Regulation Project*

Joskow, P.L. (2005a). “Transmission Policy in the United States.” *Utilities Policy*, 13:95-115.

Joskow, P.L. (2005b). “The Difficult Transition to Competitive Electricity Markets in the United States.” *Electricity Deregulation*. J. Griffin and S. Puller, eds., Chicago, University of Chicago Press.

Kennedy, Peter (2003). A Guide to Econometrics, 5th edition, Cambridge, MA: MIT Press. Lenard, T. M. and S. McGonegal (2008). “Evaluating the Effects of Wholesale Electricity Restructuring.”

Lutzenhiser, M. (2004). “Dollars and Sense: Financial Data Raises Doubts about Whether Deregulation Benefits Outweigh Costs.” *Public Utilities Fortnightly*, 142:38.

Public Power Council (2004). *Public Power Council’s Response to FERC’s “Staff Report on Cost Ranges for the Development and Operation of a Day One Regional Transmission Organization”*

Thornton, R. A. and P. Thompson (2001). “Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding.” *American Economic Review*, 91, 1350–1368.

Federal Energy Regulatory Commission (2004) “Staff Report on Cost Ranges for the Development and Operation of a Day One regional Transmission Organization,” FERC Docket No. PL04-16-000.

U.S. Government Accountability Office (2008). *Electricity Restructuring: FERC Could Take Additional Steps to Analyze Regional Transmissions Organizations' Benefits and Performance*. GAO-08 987.

Wu, F.F, F.L. Zheng, F.S. Wen (2006). “Transmission investment and expansion planning in a restructured electricity market.” *Energy*, 31:6-7, 954-966.

Table 1: RTO Summary Information

	CAISO	ERCOT	ISONE	MISO	NYISO	PJM	SPP
First year of data*:	1998	1999	1998	2002	2000	1998	2001
Mean Annual Load (GWh):	230,923	291,431	128,787	546,243	162,701	447,385	196,856
Mean Miles of Transmission Line:	25,514	37,700	7,344	90,771	10,807	33,780	47,931
States Covered:	California	Texas	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont	North Dakota, South Dakota, Nebraska, Minnesota, Iowa, Wisconsin, Illinois, Indiana, Michigan and parts of Montana, Missouri, Kentucky, and Ohio.	New York	Pennsylvania, New Jersey, Maryland, Delaware, District of Columbia, Virginia, West Virginia and Ohio. Parts of Illinois, Michigan, Indiana, Kentucky, North Carolina and Tennessee.	Kansas, Oklahoma and parts of New Mexico, Texas, Louisiana, Missouri, Mississippi and Arkansas
Headquarters:	Folsom, California	Austin, Texas	Holyoke, Massachusetts	Carmel, Indiana	Rensselaer, New York	Valley Forge, Pennsylvania	Little Rock, Arkansas
Approximate Population Served (millions)**:	30	20	14	40	19	51	5

*The first annual data record for each RTO corresponds to the year which it began providing regional transmission service under a FERC open access transmission tariff. Two exceptions follow: (1) ERCOT is not under FERC jurisdiction, and hence, does not file a tariff with them. On August 21, 1996, the Public Utility Commission of Texas required ERCOT to form an ISO with the responsibilities of grid security, market facilitation, and coordination of transmission planning. On May 27, 1999, the Texas legislature passed S.B. No. 7, which provides a framework for open access transmission. In addition to operating a control area and its traditional function of preserving reliability, ERCOT was charged with ensuring access to transmission services and settling accounts. (2) SPP commenced open access transmission operations for service across 11 transmission owners in 1998; however, 2001 was the first year of publically available financial statements.

**According to the ISO/RTO Council (IRC), http://www.isorto.org.

Table 2: Implementation Dates of Market Functions

	CAISO	ERCOT	ISONE	MISO	NYISO	PJM	SPP
Real-time energy	4/98	8/01	5/99	4/05	12/99	4/97	2/07
Day-ahead energy			3/03	4/05	12/99	6/00	
FTR	2/00	2/02	3/03	4/05	12/99	5/99	
Ancillary Services	4/98	8/01	5/99		12/99	Regulation 6/00 Reserve 12/02	
Capacity Markets			5/99 – 8/00 Reopened 4/03		12/99	1/99	

Table 3: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
OPEXP	18.470	0.695	16.097	19.417
LOAD	12.401	0.500	11.669	13.493
WIRES	10.109	0.816	8.874	11.499
WAGE	11.312	0.119	11.072	11.495
YEAR	6.597	2.955	1.000	11.000
RTOAGE	1.486	0.707	0.000	2.398
TOTALYEARS	3.373	0.800	1.099	4.205
CUMLOAD	13.798	0.856	11.669	15.409
TOTALLOAD	15.732	0.917	13.191	16.742

Table 4: Regression Parameter Estimates

	a	b	c	d
REALTIME	0.407***	0.501***	0.400***	0.392***
	[0.0901]	[0.0939]	[0.0864]	[0.0869]
DAYAHEAD_FTR	0.433***	0.192*	0.107	0.0639
	[0.130]	[0.114]	[0.115]	[0.119]
ZONE_FTR	0.161	0.329*	0.299*	0.301*
	[0.175]	[0.186]	[0.157]	[0.154]
ANCILL	0.264**	0.512***	0.738***	0.787***
	[0.127]	[0.134]	[0.188]	[0.187]
CAPACITY	-0.164	0.052	0.126	0.194
	[0.132]	[0.141]	[0.165]	[0.173]
MKTPREP	0.165**	0.263***	0.169**	0.187**
	[0.0692]	[0.0764]	[0.0743]	[0.0752]
LOAD	0.499***	0.547***	0.435**	0.156
	[0.0607]	[0.148]	[0.177]	[0.193]
WAGE	0.741*	0.891	0.6	0.584
	[0.440]	[0.783]	[0.679]	[0.679]
CAISO		0.241	0.157	0.225
		[0.242]	[0.222]	[0.234]
ERCOT		-0.228	-0.159	-0.167
		[0.212]	[0.223]	[0.232]
ISONE		0.0562	-0.124	0.0787
		[0.200]	[0.237]	[0.255]
MISO		0.833***	1.497***	1.417***
		[0.241]	[0.371]	[0.364]
NYISO		0.126	0.241	0.358
		[0.172]	[0.205]	[0.229]
SPP		0.383	0.718*	0.843**
		[0.273]	[0.369]	[0.394]
YEAR	0.0583***	0.0424**	-0.00044	0.0303
	[0.0101]	[0.0175]	[0.0313]	[0.0369]
RTOAGE			0.592***	
			[0.123]	
TOTALYEARS			-0.407**	
			[0.196]	
CUMLOAD				0.568***
				[0.122]
TOTALLOAD				-0.437**
				[0.196]
Constant	2.736	0.18	5.512	7.397
	[4.847]	[8.477]	[7.421]	[7.239]
Observations	67	67	67	67
Chi-squared	660.36	1076.26	1157.91	1097.20

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Regression Parameter Estimates

	a	b
REALTIME	0.522***	0.546***
	[0.0845]	[0.0850]
DAYAHEAD_FTR	0.251**	0.201*
	[0.108]	[0.115]
ZONE_FTR	0.455***	0.481***
	[0.148]	[0.144]
ANCILL	0.686***	0.749***
	[0.144]	[0.139]
CAPACITY	0.0920	0.214
	[0.155]	[0.166]
MKTPREP	0.325***	0.353***
	[0.0795]	[0.0809]
LOAD	2.771	4.584
	[3.710]	[3.871]
WIRE	5.040	7.813*
	[4.117]	[4.358]
LOADWIRE	-0.312	-0.522
	[0.329]	[0.347]
WAGE	1.232**	1.151*
	[0.616]	[0.609]
CAISO	-0.361	-0.243
	[0.228]	[0.231]
ERCOT	-0.662***	-0.705***
	[0.226]	[0.230]
ISONE	0.716	1.485**
	[0.643]	[0.702]
MISO	1.041***	0.948**
	[0.385]	[0.371]
NYISO	0.830*	1.318***
	[0.429]	[0.470]
SPP	-0.0964	0.0898
	[0.428]	[0.442]
YEAR	-0.0160	0.0200
	[0.0273]	[0.0319]
RTOAGE	0.819***	
	[0.129]	
TOTALYEARS	-0.668***	
	[0.198]	
CUMLOAD		0.868***
		[0.134]
TOTALLOAD		-0.783***
		[0.197]
Constant	-41.82	-66.27
	[46.33]	[47.83]
Observations	67	67
R-squared	2122.55	2063.77

*** p<0.01, ** p<0.05, * p<0.1
Standard errors in brackets

Table 6: Simulation of Total Costs

Hypothetical RTO Total Costs Change in costs	Total costs	Change in total costs			
Case (A) = Open-access transmission service*	$55,213,000				
Case (B) = Case (A) + Real-time energy	$67,192,000	$11,979,000	or	22%	> case (A).
Case (C) = Case (B) + Ancillary services	$133,404,000	$66,213,000	or	99%	> case (B).
Case (D) = Case (C) + Zonal transmission rights	$210,211,000	$76,807,000	or	58%	> case (C).
Case (E) = Case (C) + Day-ahead and nodal FTRs	$171,447,000	$38,043,000	or	29%	> case (C).

*MKTPREP is set equal to one for the hypothetical open-access transmission provider.

Figure 1

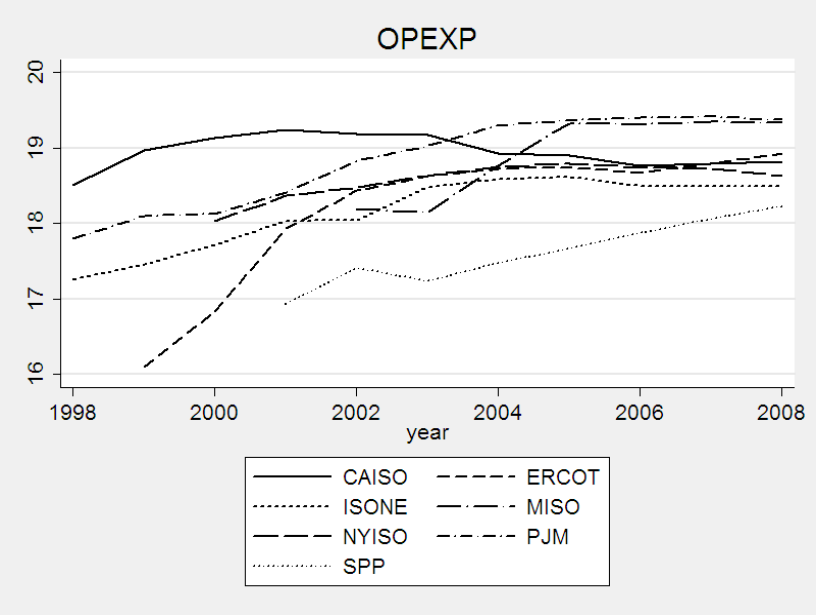

OPEXP
20
19
18
17
16
1998
2000
2002
2004
2006
2008
year
CAISO
ERCOT
ISONE
MISO
NYISO
PJM
SPP

Figure 2

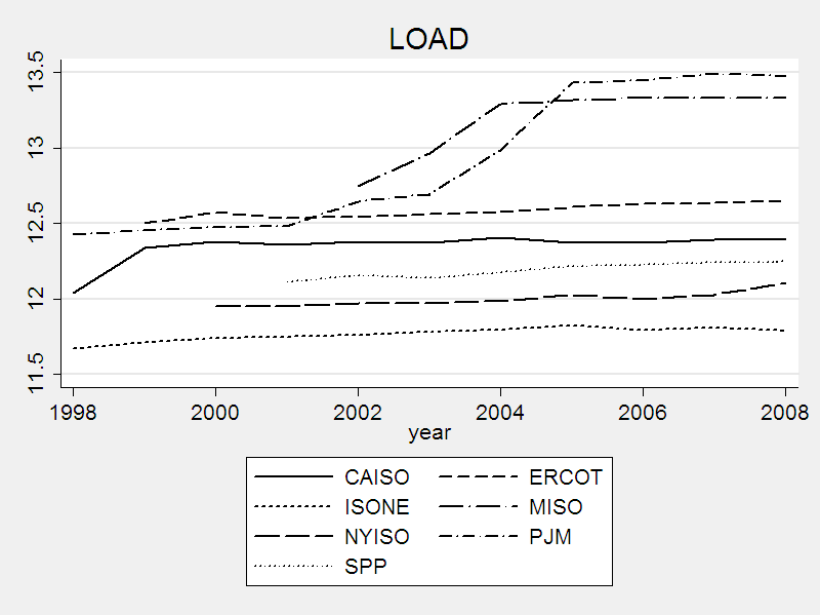
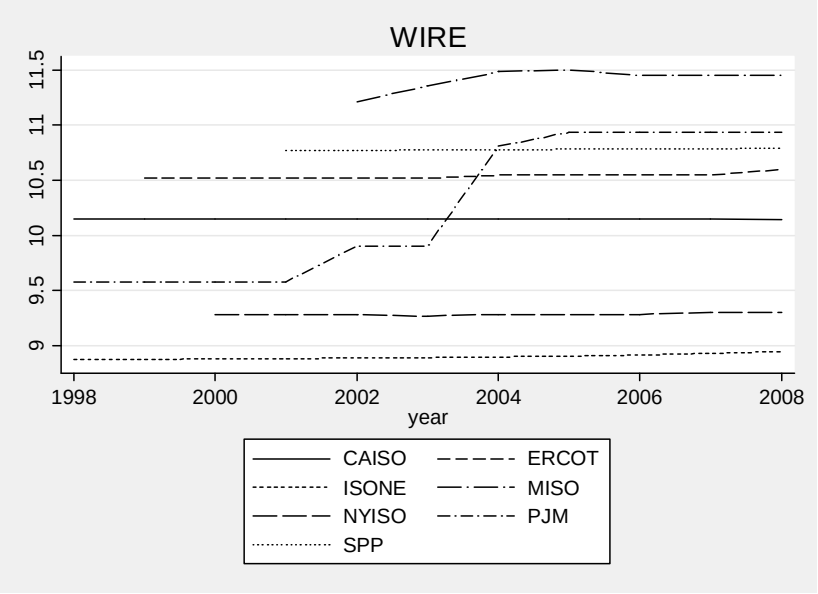

LOAD
13.5
13
12.5
12
11.5
1998
2000
2002
2004
2006
2008
year
CAISO
ERCOT
ISONE
MISO
NYISO
PJM
SPP

Figure 3


WIRE
11.5
11
10.5
10
9.5
9
1998
2000
2002
2004
2006
2008
year
CAISO
ERCOT
ISONE
MISO
NYISO
PJM
SPP