THREE ESSAYS IN APPLIED ECONOMICS:

Topics in Transportation, Industrial Organization and Health Economics

A dissertation presented

By

Pukar KC

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
March, 2018

THREE ESSAYS IN APPLIED ECONOMICS:

Topics in Transportation, Industrial Organization and Health Economics

A dissertation presented

By

Pukar KC

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the College of Social Sciences and Humanities of
Northeastern University
March, 2018

Abstract

My dissertation is comprised of three papers that empirically explore the impact of changes in public policy and firm dynamics on markets. Two of my chapters are focused in transportation and industrial organization, and the third is a topic in health economics. The three chapters of my dissertation share the spirit of being socially important questions that have not been addressed satisfactorily in the past empirical literature.

The first chapter explores the impact of exogenous changes in entry barriers on competition and market structure. This empirical study uses the case of a policy change at Dallas Love Field (DAL) airport as a natural experiment: On October 13, 2014, regulators repealed a perimeter rule (The Wright Amendment), but simultaneously introduced gate restrictions for airlines serving this airport. The relaxation of the perimeter rule allowed DAL based Southwest Airlines to enter long-haul non-stop markets from Dallas, but their capacity was constrained due to the new gate restrictions. This study finds that the policy changes at Love Field led to reduction in airfares on routes between Dallas and cities beyond the neighboring states of Texas, but increase in airfares on routes between Dallas and destinations in Texas and its surrounding states (collectively called the "Wright Perimeter"). Southwest's entry in markets where they were previously denied entry due to the perimeter rule contributes to the drop in fares. The fare increase in the short-haul Wright Perimeter markets indicates the impact of binding gate constraints. A heuristic capacity-constrained entry model is used to explain the opposite effects in different markets.

The second chapter conducts a retrospective analysis of a unique merger between two low cost carriers: Southwest Airlines and AirTran Airways. The paper begins with a detailed study of price effects for a variety of routes affected by the merger such as overlapping markets, markets where either carrier exerted potential competition, and markets where AirTran ceased service following the merger. The significant magnitude of price increase indicates that a merger between two carriers that had an industry reputation of disciplining fares of other carriers has wide ramifications on welfare. The price analysis is followed by a structural model of airline competition that is used to quantify the impact of the merger on welfare in overlapping markets. The finding shows that following the merger, consumer welfare decreased and airlines' profits increased.

The third chapter (co-authored with Ngoc Ngo), uses the dependent coverage extension component of the Affordable Care Act (ACA) as a natural experiment to study the causal impact of health insurance provision on the consumption of preventive health care services. Using a fuzzy regression discontinuity design, the study reveals that the policy change had no significant impact on the forms of preventive care services studied, although it significantly increased coverage among young adults. Alternative empirical analyses conducted using difference-in-differences and propensity scores show that the general findings are insensitive to the choice of econometric methodology. Since the analysis is based on the preventive health care usage of young adults affected by the policy change, it could be indicative of moral hazard behavior specific to this age group. A theoretical framework is also devised to gauge the relation between moral hazard, insurance provision, and the usage of preventive care.

Overall, my dissertation contributes to the applied economics and public policy literature by empirically examining the consequences of policy changes that affect large parts of society. It is hoped that the research will be useful to inform some ongoing debates in industrial policy, antitrust economics and healthcare.

Acknowledgements

I would like to thank my committee members: Professors Steve Morrison, James Dana and John Kwoka for their invaluable support to me throughout graduate school. I am also grateful for the constructive feedback from Professor Imke Reimers on the first two essays. The third essay would not have been possible without the generous help from Professor Mindy Marks. I would also like to thank my friend and co-author Ngoc Ngo for being a such a supportive individual throughout the research process. Lastly, I want to thank my father, whose contributions to the Nepalese aviation industry continue to inspire and foster my deep interest in the airline industry.

TABLE OF CONTENTS

Chapter 1

Non-Stop Love: A Study of Entry Barriers in the Airline Industry Using Policy Changes at Dallas Love Field

1.1 Introduction

Studies of firm entry, and its impact on market structure are at the heart of the industrial organization literature. The consensus in neoclassical models is that in markets other than natural monopolies, firm entry is usually desirable since the increase in competitiveness drives down prices, and enables markets to achieve a more allocative efficient equilibrium. However, barriers to entry exist in several industries that inhibit firms to operate in markets. In some instances, incumbent firms have been found to erect entry barriers to preserve their economic profit. Strategies deployed to achieve such ends include controlling key resources, predatory pricing, and collusive agreements. Sometimes, entry barriers are intrinsic outcomes of a firm's operation. Such is the case when a firm experiences economies of scale, or develops superior technology that gives it a clear advantage over potential competitors.

Setting aside firms' sphere of influence in the market, entry barriers are sometimes upshots of laws instituted by regulators. Some such regulations, like patent protection laws, incentivize innovation at the cost of anticompetitive outcomes. Others are instituted due to pure political motivations, and may lack a clear economic justification, or have one that appears vestigial. In the context of the airline industry, the Wright Amendment and gate reductions at Dallas Love Field are paradigms of the latter.

Gate restrictions and perimeter rules are popular forms of regulations used to suppress airline competition. Other such regulations include slot controls (usually instituted to control congestion at airports), and air traffic rights (instituted to allow or deny carriers from foreign countries). Airport gates are used by operating carriers to board/disembark passengers, and is hence a crucial determinant of an airline's capacity at an airport. If availability of gates is reduced at an airport, the airline could recover some capacity by improving flight scheduling and reducing the lead time between flights, but there is a limit to the extent of such tools.

A perimeter rule at an airport restricts carriers serving that airport from offering flights outside an indicated perimeter. Perimeter rules are currently in effect at New York LaGuardia and Reagan National, DC. The motivation often cited for instituting a perimeter rule is to shift traffic from centrally-located airports to newly-built regional airports. Implementing a perimeter rule at an airport leads to airlines diverting long-haul flights to a non-perimeter restricted substitute airport, which would then spur infrastructure development at the substitute airport.

The Wright Amendment (WA) was a perimeter rule imposed on Dallas Love Field (DAL). When it was fully repealed on October 13, 2014, gate restrictions were simultaneously introduced at DAL by reducing the number of gates from 32 to 20. Southwest Airlines was mostly affected by these changes since during that time, over 95 percent of passengers enplaned at DAL were Southwest's customers.[1]

The policy changes at Love Field had opposing effects: on one hand entry barriers were relaxed by repealing the perimeter rule, but on the other, new barriers were introduced in the form of gate restrictions. These events present a valuable opportunity to study the impact of entry barriers in the airline industry, primarily due to the exogenous nature of the policy changes. To recognize that the policy changes at Love Field were exogenous with regards to the state of airline markets during the time frame that this study uses, we need to understand the history and politics of the Wright Amendment.

[1]BTS data, 2013-14.

The Wright Amendment: History and Politics

In the context of airports in the Dallas region, there were once four airports in operation around the area: Dallas Love Field (DAL), Greater Southwest Airport, Red Bird Airport and Meacham Field. Federal officials drafted a proposal to build a single airport and decommission the smaller competing airports. This new airport began operations in 1974 as Dallas Fort Worth International Airport (DFW). With the completion of the construction of DFW, all airlines except Southwest Airlines agreed to relocate their operations to DFW from surrounding airports.

Southwest Airlines initially started off as a low-cost intrastate carrier in Texas, and was hence initially exempt from Civil Aeronautics Board (CAB) regulations. Their operations were based at DAL, and only included destinations within Texas. Southwest's refusal to abandon services at DAL and relocate to the newly built DFW was challenged by the cities of Dallas and Fort Worth in a court case. Although Southwest won the court case and could base operations at DAL, airline deregulation in 1978 reignited the discussion. Airline deregulation allowed Southwest to expand beyond Texas, and since its services were based out of DAL, officials at DFW worried that the amount of air traffic Southwest would now generate could seriously challenge DFW.

Supporters of DFW lobbied the then US House of Representatives Speaker Jim Wright (D-Texas) to institute a law to protect the newly built DFW from competing with DAL. This led to the "Wright Amendment" being implemented in 1980. The implications of the Amendment were the following:[2]

1. It became illegal for any airline at DAL to offer flights to destinations beyond Texas and its four contiguous states: Louisiana, Arkansas, Oklahoma and New Mexico. These states were collectively demarcated as the "Wright Perimeter"

2. Airlines were prohibited to offer or advertise the availability of any connecting flights between DAL and any city outside the Wright Perimeter.

3. Airlines at DAL may not use aircraft with more than 56 seats for commercial purposes to destinations outside the Wright Perimeter.[3]

[2] Love Terminal Partners, et al. Plaintiffs

[3] Note that provision (3) allows airlines to operate flights anywhere from DAL, but only in extremely small

Over the years, Southwest Airlines launched several campaigns claiming that perimeter restrictions at DAL lead to a decrease in consumer welfare due to higher fares. Such campaigning led to a series of relaxations of the perimeter rule, and several states were added to the Wright Perimeter:

1. Shelby Amendment, 1997: Sponsored by Senator Richard Shelby of Alabama, the Amendment allowed carriers to operate non-stop flights from DAL to Alabama, Mississippi and Kansas.

2. Bond Amendment, 2005: Sponsored by Senator Christopher 'Kit' Bond of Missouri, the Amendment allowed flights to Missouri from DAL.

3. Wright Amendment Reform Act, 2006: The agreement laid the groundwork for the complete repeal of the Wright Amendment. Conditions were:

    - From October 2006 to October 2014, airlines having flights from DAL could sell tickets to any destination in the country as long as the flights made a stopover within the Wright Perimeter.

    - From October 2014, airlines would be allowed to operate nonstop flights to anywhere in the country from DAL but the number of gates at the airport would be reduced from 32 to 20.

The final repeal of the Wright Amendment took effect on October 13, 2014. This policy change was enthusiastically welcomed by Southwest; the airline immediately launched non-stop flights to seven new destinations in October, and eight more destinations from DAL were added in November.[4] Virgin America also welcomed the decision, switching its Dallas flight operations from DFW to DAL. However, as a political compromise between the cities of Dallas and Fort Worth, gate restrictions were imposed, and these continue to act as a major source of stress for airlines operating from DAL. [5]

Awareness of the history of the WA reveals that the reasons behind the passage of the perimeter rule and its eventual repeal (and accompanying gate reductions) are deeply rooted in politics, and

capacities. Therefore, throughout the Wright Amendment literature, it is stated that airlines out of DAL were simply not allowed to operate flights anywhere from DAL.

[4] Airlines for America (A4A, 2014)

[5] A Dallas news article discusses how gate limitations have led to a conflict between Southwest and Delta: https://www.dallasnews.com/business/business/2016/07/11/southwest-delta-share-dallas-love-field-gate-working

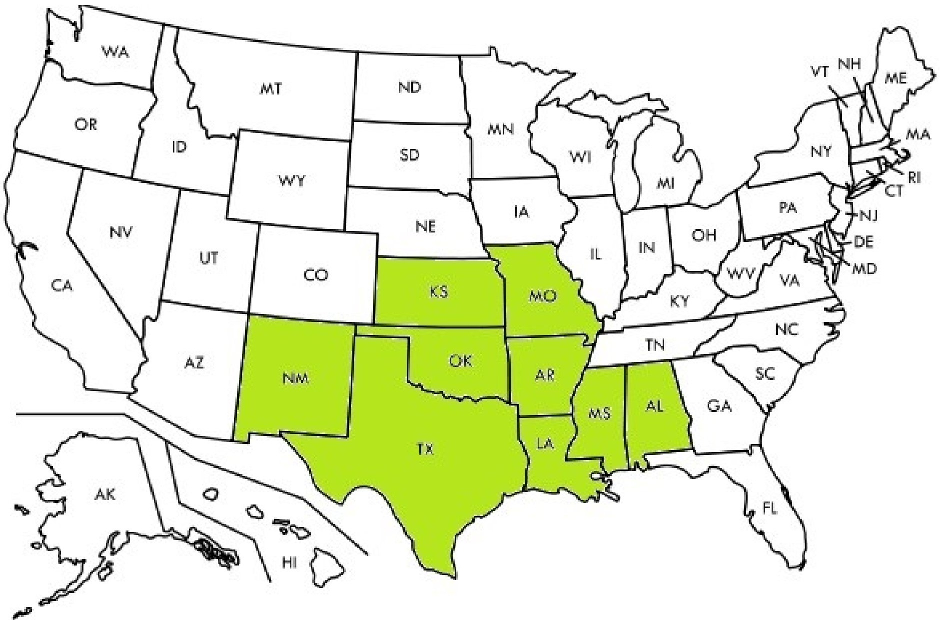


Figure 1.1: The Wright Perimeter (WP) as of October 2014 (shaded in green)

are not quite related to the market structure of the airline industry in the affected markets. One could nevertheless argue that Southwest's meteoric rise as a successful low-cost carrier could have cajoled policymakers to repeal the WA. Southwest's business success is quite reasonably correlated to unobserved (to the econometrician) parameters of airline markets, thereby casting doubt on the exogeneity of the WA repeal. However, it must be noted that the policy changes at Love Field were set to be implemented in 2014 by a reform act that was announced in 2006. Policy makers in 2006 could not have had an accurate bearing about how the industry would be eight years later, thus making the perimeter rule relaxation and gate reductions exogenous to the state of airline markets at the time these policy changes were implemented.

The empirical investigation in this paper begins with a difference-in-differences approach to quantify the fare impact of the October 13, 2014 policy changes. Analysis is presented separately for the impact on fares at DAL and DFW airports, and for routes that connect Dallas to destinations within the Wright Perimeter (WP), and outside the WP. The motivation to categorize markets in this way is to shed light on disentangling the effect of the perimeter rule relaxation from the gate cuts. The fare regressions are supplemented with regressions with passenger quantity as the dependent variable.

Fare movements are the results of changes in competition and costs. The WA repeal allows airlines operating at DAL, primarily Southwest Airlines, to enter non-stop routes between DAL and destinations outside the WP. Southwest's entry causes fares in these markets to drop. The econometrician should exercise care while attempting to quantify the magnitude of fare change from Southwest's entry since entry is non-random and endogenous to market conditions. It is reasonable to believe that Southwest would enter non-stop routes where it has unobserved advantages over rival carriers. Therefore, an OLS coefficient obtained from a difference-in-differences regression would yield a biased large magnitude. This endogeneity is controlled for by using several measures of airport presence from the end-point airports as instruments for Southwest's entry.

The rest of the paper is divided into five parts: The Related Literature section summarizes papers in empirical Industrial Organization that are related to my work. The section that follows discusses a heuristic capacity-constrained entry model that motivates the analysis of fares to gauge the strength of gate constraints. The Data section discusses the details of the dataset used, and how the relevant data were filtered using standard methods used in airline studies. The Empirical Analysis discusses the setup of the regression analyses, and their results. The Conclusion presents a summary, and discusses the welfare implications of the policy changes.

1.2 Related Literature

Many papers in empirical IO have explored the impact of firm entry on markets. These papers use both reduced-form as well as structural approaches. One widely cited paper is Bresnahan and Reiss (1991) in which the authors work with a dataset lacking detailed firm level data on prices, costs and quantities. The authors show that firm entry leads to increased competitiveness, which decreases margins. Another notable work includes Mazzeo (2002), in which the author endogenizes a firm's product type decision, and shows that the decrease in margins arising from firm entry also depends on the relative product space location of competitors.

With regards to the airline entry literature, Berry (1992) considers the role of airport presence in determining an airline's profitability in a given market. As in Bresnahan and Reiss (1991), Berry begins his analysis by noting that entry by firms indicates the potential of profit in the market.

Using a maximum likelihood estimator, he finds that airport presence[6] at either end-point of a city pair is strongly correlated with entry decision to serve the city pair market. I use Berry's findings in devising instruments for Southwest's entry (discussed later). Berry concludes that while increasing airport access may make it possible to decrease market concentration, the equilibrium number of firms entering the market will also depend on competition within the market. Strong within market competition can limit the number of entering firms, even when policies that effectively increase market access are implemented.

Morrison (2001) estimates consumer savings resulting from Southwest Airlines directly operating a route (airport pair), a route adjacent to the airport pair, or simply exerting potential competition (by operating other routes from both airports, but not operating the route connecting the airports). The monumental estimate of savings resulting from Southwest's influence- $12.9 billion leads him to the policy recommendation that policies targeted towards easing entry can tremendously increase consumer welfare. Another work in similar vein is Goolsbee and Syverson (2008), in which the authors show that simply a threat of entry by Southwest airlines is enough to lower fares by incumbent carriers in a route.

Despite being an interesting policy change case, the Wright Amendment has received relatively little attention among economists. In a consultant report in 2006, Morrison estimated that total consumer savings from the Wright Amendment repeal would be at least $726 million for full repeal versus $117 million to $281 million for through ticketing.[7] Boguslaski et. al. (2004) present an analysis of how Southwest's entry and network strategies have evolved over time. They find that the Wright Amendment has imposed a binding constraint on Southwest's operations, and has led to large foregone savings for passengers.

The effect of the relaxations of the perimeter rule via the Bond, Shelby Amendments, and the Reform Act on pricing and consumer welfare was studied by Bold (2013) in one of his dissertation essays. Bold's research shows that the previous relaxations of the WA have led to quite substantial fare decreases in affected routes, and also contributed to the increase in market share of Southwest

[6] Airport presence in Berry (1992) is either the mean value of passenger miles across the two cities in the pair, or the mean number of destinations served out of either end points.

[7] Through ticketing is the arrangement that allows airlines operating out of DAL to market tickets to destinations beyond the WP, although a stopover needs to be made within the perimeter.

airlines.

While Bold's work focused on the effects of the previous relaxations of the Wright Amendment on market fares and consumer welfare, this is the first paper in the literature that discusses the most recent relaxation and final appeal of the Wright Amendment. This paper also contributes to the airline entry literature by devising instruments to account for endogenous entry. Furthermore, this paper provides a more holistic analysis of the repeal of the WA by tying together the fare and quantity changes using a simple framework.

1.3 A Model of Capacity-Constrained Entry

There exist two markets:

M_1: represents Dallas to Texas and its neighboring states (Wright Perimeter). This market has two competing firms: A (representing Southwest) and B (representing American Airlines, the dominant carrier based at DFW).

M_2: represents Dallas to outside Wright Perimeter. B is the monopolist in this market before repeal and after repeal, A enters this market.

Assume a linear demand function with unit mass of consumers. Marginal costs are zero. The products of firms A and B are assumed to be homogenous. Firms compete in quantities.

Before repeal:

In M_1, $P_1 = 1 - Q_1 = 1 - q_{A1} - q_{B1}$. Solving FOCs yields $P_1 = 0.33$, $q_{A1} = q_{B1} = 0.33$. There exists a constant capacity "K" for A. Assume K>0.33, so capacity is non-binding.

In M_2, B is the monopolist and $P_2 = 1 - q_{B2}$. This yields $P_2 = 0.5$ and $q_{B2} = 0.5$.

Post repeal, A enters M_2.

For A,

$$\sum\pi = \pi_{M1} + \pi_{M2} = (1 - q_{A1} - q_{B1})q_{A1} + (1 - q_{A2} - q_{B2})q_{A2}, \text{ such that } q_{A1} + q_{A2} \leq \text{K}$$

For B,

$$\sum\pi = \pi_{M1} + \pi_{M2} = (1 - q_{A1} - q_{B1})q_{B1} + (1 - q_{A2} - q_{B2})q_{B2}$$

Note that B is not capacity constrained, but capacity is limited for A due to gate reductions imposed post repeal.

Solving FOCs reveals $q_{A1} = q_{A2} = \frac{K}{2}$, $q_{B1} = q_{B2} = \frac{2-K}{4}$, and $P_1 = P_2 = \frac{2-K}{4}$

The counterfactual where no binding capacity constraint is imposed on A can be checked by solving the respective unconstrained FOCs. This reveals that P_1=P_2=$\frac{1}{3}$, i.e., prices would fall in the markets where entry would take place (M_2), and not change in M_1. Also, q_A would be $\frac{2}{3}$. Thus capacity binds when K<$\frac{2}{3}$. With the presence of the capacity constraint parameter "K", a variety of movements in P_1 and P_2 are possible. Figure 1.2 plots K against P_1 and P_2 (both of which are equal to $\frac{2-K}{4}$ after policy change, per the calculation above).

The horizontal red and blue lines are the prices in M_1 and M_2 before repeal (equaling $\frac{1}{3}$ and $\frac{1}{2}$) respectively. These have been plotted to compare ex-ante prices in the two markets with the ex-post prices, which depend on K. The dark black line shows the relation between K and P after repeal. We see that when capacity is small, price after repeal falls in M_2 but rises in M_1.

This simple model illustrates how we can use price movements in the two markets to gauge how strict or lax the gate constraint was. Following the discussion above, at very restrictive gate constraints, prices will fall in M_2 but rise in M_1. On the other hand, when gate constraints are lax and non-binding, prices will fall in the market where entry takes place, but does not change in the other market. In the empirical section, price and output movements in the two markets will be studied, which will help infer the restrictiveness of the gate constraint.

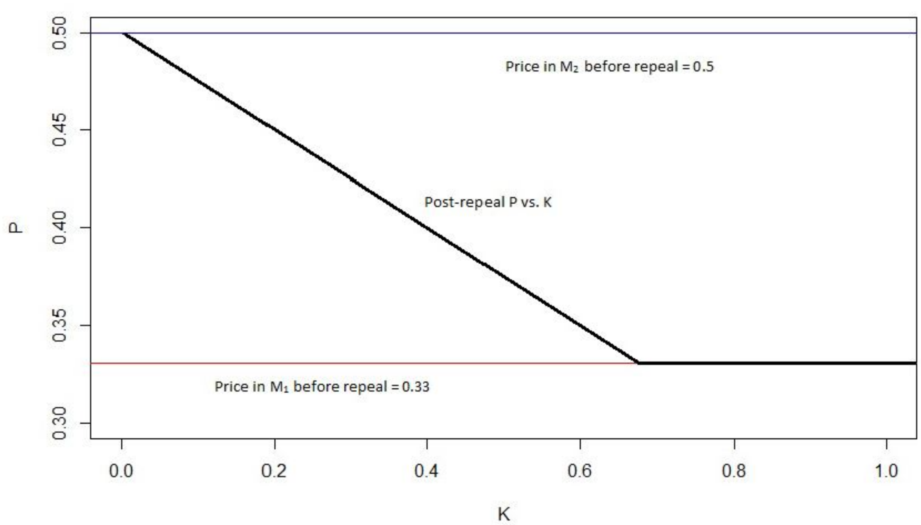


Figure 1.2: Capacity Constraint (K) versus Price (P)

1.4 Data

Air fare data were obtained using the Airline Origin and Destination Survey (DB1B). Published quarterly by the Bureau of Transportation Statistics (BTS) of the Department of Transportation (DOT), the DB1B is a 10% ticket sample of airline tickets from reporting carriers.[8] This useful data source contains detailed information about the ticket such as market fare, origin and destination, number of passengers with the same flight, etc.

A market in this paper refers to a non-directional airport pair. The DB1B gives individual ticket level data but for this research, data were coalesced to market-carrier-year quarter level, i.e., an observation refers to an airline-specific market in a given quarter of a year. For example, Boston Logan-Dallas Love field on Southwest in Q1 2014 is one observation.[9]

The relevant data were all quarters of 2013, quarters one through three of 2014, all quarters of 2015, and quarters one through three of 2016 (i.e., seven quarters before, and seven after the policy change).[10] The fourth quarter of 2014, i.e., the quarter in which the change in regulation took place has been omitted due to the impossibility of breaking the quarterly data accurately into pre and post October 13 sets.

All observations with market coupons[11] greater than three were dropped as these tend to be open jaw tickets. Bulk fares were dropped as well. All tickets with market fares less than $30 and greater than $5000 were also dropped. The abnormal fares could be the result of coding errors, or frequent flier miles. City pairs with less than 30 sample passengers in an entire quarter were dropped for that quarter.[12] To control for airline network evolution, only markets that are present in both the pre (2013-14) and post (2015-16) Wright Amendment eras are considered.

The T-100 DS (Domestic Segment) database was used to obtain information on non-stop flight segments. Published monthly, this data table contains flight-specific information as reported by

[8] Reporting carriers in the DB1B include all the major airlines operating domestic routes in the US.

[9] Robustness checks were carried out with market year quarter as an observation. The findings are reported in the appendix.

[10] Robustness checks were conducted with different time periods before and after the policy change: 2013 q4 - 2014 q3 as pre, 2015 q1 - 4 as post (four quarters before and after), and 2012 q2 - 2014 q3 as pre, 2015 q1 - 2017 q2 as post (ten quarters before and after). The results are reported in the Appendix.

[11] A coupon in the DB1B represents a boarding pass.

[12] Follows Kwoka (2010).

participating carriers. It provides flight-level data such as the origin and destination, routing of the flight, passengers enplaned, frequency, etc.

City demographics data were obtained from the Census Bureau.

1.5 Empirical Analysis

1.5.1 Baseline Fare Regressions

The baseline fare regressions use a difference-in-differences approach to investigate the causality of change in air fares resulting from policy change. Two different control groups are used:

CONTROL-A: The set of markets both of whose end-point airports are outside the Wright Perimeter, and do not have Southwest Airlines operating in any city-market pair originating from the end-point airports.

Since the policy changes primarily affected the operations of Southwest Airlines, all Southwest markets are excluded from this control group as they could be affected when the airline restructures its operations. For instance, Cincinnati/Northern Kentucky International (OH) Jackson Hole (WY) is one of the markets in the control group. The end point airports are in Ohio and Wyoming, both states are outside WP. Southwest Airlines does not operate the city market pair Cincinnati Jackson, nor does it operate any flights to other destinations out of Cincinnati and Jackson.

CONTROL-B: The subset of CONTROL-A containing only markets where at least one low-cost carrier[13] *is present in the respective city-market pair.*

Since the markets being studied in this study are strongly impacted by the operation of Southwest Airlines, a low-cost carrier, elements in the control group would ideally comprise of markets that are also impacted by low-cost carriers. It is widely accepted that markets operated by LCCs are substantially different than those operated by legacy carriers. LCC markets have a point-to-point rather than a hub-and-spoke typical to legacy airline markets. Furthermore, LCCs cater more to leisure passengers whereas legacy airlines target business passengers. In a difference-in-differences framework, the treatment and control groups should only differ on the grounds of being subject to

[13] Following Kwoka (2016), a low-cost carrier is any of the following: AirTran, JetBlue, Frontier, Allegiant, Spirit and Southwest. Note that markets operated by Southwest are excluded from CONTROL-B since they are not present in CONTROL-A.

the treatment. Hence, the exclusion of non-LCC markets could be appropriate.

However, note that factors that contribute to LCC markets being different from legacy markets may not be directly observed in the dataset, and thus could be correlated with the regression error. This would imply that there could exist a selection bias with the way CONTROL-B is constructed. In apropos of this argument, CONTROL-A might be a more appropriate control group. However, CONTROL-A is more dissimilar to the treatment markets with regards to the presence of LCCs. From this discussion, a clear trade-off of using either control groups becomes apparent: one of them is more similar to the treatment and potentially endogenous, whereas the other is more dissimilar to the treatment but less endogenous. Therefore, analysis is presented in this study using both control groups.

The primary specification for the OLS regression is as follows:

$$ln(Fare_{ikt}) = \alpha_0 + \alpha_1 * Treatment_{ij} + \alpha_2 * Post_t + \alpha_3 * Post_t * Treatment_{ij} + \alpha * X_{ikt} + \epsilon_{ikt}$$

Here, "j" is the treatment dummy (equal to 1 if "i" is in treatment group "j"), "i" is the market (non-directional airport pair), "k" is the carrier and "t" represents the year-quarter. The dependent variable, $\ln(Fare_{ikt})$, is the logarithm of the average airline specific market fare. Depending on the nature of the market area being studied, several treatment groups (j) are defined as follows:

Treatment 1: Dallas Love to/from outside Wright Perimeter. This can be sub categorized as follows (relevant in the entry discussion):

Treatment 1a: Markets entered by Southwest following repeal.

Treatment 1b: Markets not entered by Southwest following repeal.

Treatment 2: Dallas Fort Worth to/from outside Wright Perimeter.

Treatment 3: Dallas Love to/from Wright Perimeter.

Treatment 4: Dallas Fort Worth to/from Wright Perimeter.

Categorizing treatment groups in this manner also allows to disentangle the impact of the gate constraints from the perimeter rule relaxation. The perimeter rule was binding in treatment 1 markets, but not in treatment 3. However, both treatment 1 and 3 were impacted by gate constraints since these are Love Field markets. Therefore, the fare movements in treatment 1 are due to the perimeter rule repeal and gate reduction, but the movements in treatment 3 are only due to the new gate restraints.

The "post" variable takes value "1" if the data are in the 2015-16 (post WA) period. The interaction of the post and treatment dummies is the primary independent variable of interest. An advantage of the difference-in-differences model is that influences on the dependent variable that are common to both the treatment and control group drop out while running the regression. This is useful because variables such as inflation (which would affect all routes) do not need to be explicitly controlled for. Nevertheless, a number of control variables are included in the regression since they may not be constant across the treatment and control groups over time. These are denoted as X_{ikt} and described as follows:

1. Distance: The influence of distance between the end-point airports is accounted for by including the logarithm of average market miles flown between end-point airports as a covariate. Since some flights with stopovers may have been converted to non-stop flights in different time periods, distance between the end-point airports for the cross-section of markets may also be slightly different over time.

2. Population: To control for population, logarithm of the population product at the end-point metropolitan areas is included as a regressor.[14]

3. Effective Competitors (EC): This is the reciprocal of the HHI of the market. I use this specification instead of the HHI due to the more intuitive appeal of understanding competition using the number of firms in the market. However, market structure is endogenous in the regression equation. Some specifications include the EC variable and others do not.

4. Quarter dummies: These are included to control for seasonal variation in air fares.

5. Other dummies to indicate whether

    - the airport is slot controlled.[15]

[14]This follows the logic of gravity models used in urban geography literature.

[15]Some airports in the United States are slot controlled, i.e., restrictions are imposed on airlines operating at that airport from making more than a given number of take-offs and landings.

- the airport is a hub for any of the carriers.[16]
- the city where the airport is located is a tourist destination.[17]

These dummies are included to check the validity of the regression setup rather than to control for their influence on the coefficients of interest.

Table 1.1 presents quick summary statistics on the treatment and control groups. Note that an observation is a market-airline-year-quarter. The general price movements before and after the policy changes of October 2014 can be gauged by studying the table: prices increased in Treatment groups 3 and 4, but fell in 1 and 2. Regression results will more carefully describe the causality of the policy change, and the resulting price movements.

Table 1.1: Summary Statistics of Treatment and Control Groups

Group	Observations	Average sample fare before repeal (USD)[18]	Average sample fare after repeal (USD)
Control-A	39,119	350	351
Control-B	3,024	289	278
Treat 1: DAL - outside WP	5,105	263	251
Treat 1a: Markets entered by Southwest	3,464	253	226
Treat 1b: Markets not entered by Southwest	1,641	288	292
Treat 2: DFW - outside WP	22,446	294	280
Treat 3: DAL - WP	895	197	209
Treat 4: DFW - WP	3,436	188	212

The results of the baseline regressions using Controls A and B are reported in Tables 1.4 and 1.5 respectively. In both tables, specifications (1) and (2) are OLS with robust standard errors, (3) includes market fixed effects, (4) includes market, year and airline fixed effects. (1) excludes the endogenous EC variable, whereas (2), (3) and (4) include it. The coefficients of the Effective Competitors (EC) variable is positive in (2). This may be due to omitted variables bias: more airlines may be attracted to long haul markets that have numerous connections and topological network advantages.[19] Such long-haul markets have higher fares since they involve connecting two

[16] The definition of an airline's hub follows the information given on their websites.

[17] A tourist destination was defined using http://travel.usnews.com/rankings/best-usa-vacations/

[19] As an example, a market with end-points that have more connecting markets, or more airlines operating (presenting opportunities of codesharing) may have more airlines.

cities that are far apart, hence making other forms of transportation connecting them (road or train) unlikely. The EC variable is also correlated with the policy change, as the repeal allowed Southwest to enter several markets. Such issues with the EC variable may cast doubt about its inclusion in the regression specification. Including the variable may help to isolate the effect of the policy change from other reasons airlines may be entering or exiting markets. I therefore run specifications with the EC variable and without, observing little change in the coefficients of interest.

Including fixed effects reverses the sign of the EC coefficient, making it consistent to economic theory: *cet. par* competitive markets would have lower fares. One possible conjecture for the sign reversal on the coefficient could be that the tendency of some markets to have higher fares and higher number of participating firms could be due to the idiosyncratic nature of the market. For instance, a long-haul market could have higher fares and higher number of competitors due to the discussion above. "Long-haul-ness" could be a market specific fixed effect, which is controlled for in (3) and (4). Including year fixed effects helps control for the impact of the changes in jet fuel prices. Airline fixed effects help control for firm-specific cost shocks.

Specification (4)'s coefficients will be used to infer the results since it is the most well-specified model that accounts for most of the unobservable factors. We see that the primary variables of interest (the ex post treatment interactions) follow predictions from theory. Markets between Dallas Love (DAL) and outside WP destinations experience a drop in fares following the repeal of the WA (7.7%: with CONTROL A, 4.6%: with CONTROL B). Since Dallas Fort Worth (DFW) is a substitute for DAL, it is understandable that DFW to outside WP markets too experienced a drop in fares (5.5% with A, 2.8% with B). The drop in fares is due to entry by DAL-based Southwest airlines (discussed more in the next section), but could also be due to cost improvements. Although including airline fixed effects in (4) helps account for airline specific cost changes, due to the lack of market-specific cost data, I am unable to investigate this part empirically.

Markets between DAL and within Wright Perimeter destinations show an increase in fares (6.3% with A, 10.2% with B). The capacity constraint faced by airlines operating at DAL, attested by the fact that number of gates were decreased in DAL along with the repeal of the perimeter rule, lead to the fare increase. Airlines at DAL face a tradeoff of operating either DAL-outside WP cities, or

DAL-inside WP. The fact that they enter more long-haul markets at the cost of short-haul markets indicates that long-haul markets are more profitable. As they redeploy their resources to the long-haul markets, they serve smaller capacity in the short-haul DAL-inside WP markets (discussed under quantity regressions).

It could be argued that the fare movements in Treatment 3 are partly due to the repeal of the perimeter rule, and thus are not purely indicative of the impact of gate constraints. This could be true since following the policy change, an airline from DAL does not have to make a redundant stopover within the Wright Perimeter to connect to destinations beyond, and in essence would be exiting the respective DAL - WP segment of the overall long-haul route not due to gate limitations, but due to the absence of perimeter restrictions. To identify markets that were only directly affected by the gate constraints, a subset of Treatment 3 was constructed, which comprise of markets that were not a segment of any multi-stop DAL to outside WP markets in the ex-ante period. One constituent of this subset is the DAL - Houston Hobby market. Southwest did not use Houston Hobby as a stopover to connect to destinations beyond the WP from DAL. Fare regressions on this subset are presented in Tables 1.6 and 1.7 using the two control groups. The magnitude of the fare increase in these markets (8.3% with A, 10.3% with B) is similar to the fare increase in Treatment 3 markets.[20] This indicates that gate constraints were indeed binding in the post period.

A market clearing higher fare in DAL-inside WP leads to the increase in fares of the substitute DFW-inside WP markets (15.3% with A, 18.9% with B). The fare increase in DFW based routes is much higher than the adjacent DAL routes. A possible conjecture is that following the exit of Southwest from DAL based short-haul markets[21], leisure passengers, who are the primary users of services offered by LCCs, could have switched to other modes of transportation for inter-city passenger travel. Such a scenario is reasonable since short-haul airline travel faces substantial competition with other modes of transportation. Furthermore, most leisure passengers would use DAL since it is an airport dominated by Southwest Airlines, a low-cost carrier. Hence, only business

[20] An analysis of the change in output and capacity in this subset of Treatment 3 validates the claim that the fare rise results from Southwest exiting these markets following the policy changes: these markets had 300 thousand seats, 4,589 departures in pre, and 29 thousand seats and 530 departures in post.

[21] DAL-WP and DFW-WP are referred as "short-haul" markets since they involve end-point cities that are smaller in distance than DAL-outside WP and DFW-outside WP, which are referred as "long-haul" markets.

passengers would be left in these airline markets, and legacy carriers serving them from DFW might find it profit-maximizing to charge much higher fares.[22]

The other covariates show their expected signs in relation to other works in the airline literature: slot controlled airports are more expensive, hub airports are more expensive, tourist destinations are cheaper, and perhaps due to traffic densities, highly populated airport pairs are cheaper.

1.5.2 Non-stop versus multiple-stop flights

The policy changes of October 13, 2014 allowed airlines operating at DAL (primarily Southwest Airlines) to operate non-stop flights between DAL and any destination in the United States. Before the full repeal of the WA in October, Southwest still operated flights connecting DAL to destinations all over the United States, but these flights made a stopover somewhere within the Wright Perimeter. Following the policy change, many multiple-stop flights from DAL were converted into non-stop flights.

It is reasonable to expect that the impact of the policy change on non-stop and multiple-stop flights would be different in magnitude. To investigate, we can break down Treatment 2 (DFW-outside WP) into two categories: non-stop and multiple-stop flights. Note that we cannot similarly analyze Treatment 1 (DAL-outside WP), since the ex-ante observations are only multiple-stop flights.

In this section, DB1B data were coalesced to market-carrier-year quarter-market coupon level. The number of market coupons helps identify if the flight is non-stop, where market coupons equals one, or multiple stop, where market coupons are greater than one. The dependent variable is the coupon-specific average market-airline fare across year-quarters. The same X_{ikt} and control groups as in the baseline regression are used. The results are reported in Tables 1.8 and 1.9 using controls A and B respectively.

As in the baseline regressions, specification (3) was run with market fixed effects, and (4) includes market, year and airline fixed effects. The results reveal that the fare impact was much

[22]It could be argued that airlines could have used various forms of price discrimination to charge different fares to business and leisure passengers even in the ex-ante period. While this is true, the exit of leisure passengers from airline markets in the ex-post period in Dallas-based short-haul markets makes it easier for carriers to devise policies catered only for business passengers.

stronger on non-stop flights (decrease by 4.8% with A, 4.7% with B)[23] than multi-stop flights (1.8% decrease with A, result with B not statistically significant). This is reasonable since the repeal of the Wright Amendment allowed DAL based Southwest to enter non-stop routes from DAL, which impose competitive pressure on non-stop flights out of the adjacent DFW airport. It is also reasonable to expect some effect of the policy change on multi-stop flights as well since a non-stop flight between cities X and Y also competes with multi-stop flights between the same two cities. However, the results show a small impact. The coefficients of other covariates are similar to the baseline regression results, which have been discussed in the preceding section.

1.5.3 A deeper look at entry using Instruments

The goal in this section is to measure the effect of Southwest's entry on market prices following the repeal of the perimeter rule. As mentioned earlier, the markets entered by Southwest in this context are a subcategory of Treatment 1 (DAL – outside WP). Since Southwest's entry decision is non-random, and quite possibly correlated to market characteristics, a two-staged least squares approach is the appropriate econometric tool.

Following the entry literature, a firm enters a market only if it expects to earn a profit in the Nash equilibrium. Profitability is a function of both demand and cost conditions; Southwest would enter a market where it can get sufficient demand to fill up its planes, and where it has a cost advantage over incumbent carriers. The IV needs to be correlated with the entry decision, but uncorrelated with the regression error. Since data on population at end-point cities is included as a regressor, the demand side condition motivating entry is addressed. However, since route-specific airline cost data are unavailable, the regression error comprises of such unobserved costs. Therefore, the IV should predict entry, but not be correlated with Southwest's cost advantage.

According to Berry (1992), airport presence at either end-point of a city pair is strongly correlated with entry decision to serve the city pair market. This makes intuitive sense and may be best understood using an example from the dataset. Southwest introduced non-stop flights in the DAL-Boston Logan market after WA repeal, but did not introduce non-stop flights in the DAL-Asheville

[23]The percent figures used to discuss the magnitude of effects use specification (4) since it includes all fixed effects and hence accounts for most unobservables compared to other specifications.

market. In both these markets, Southwest was operating in the ex-ante period but with a stopover in the Wright Perimeter. Consider the fact that Southwest operates flights to many more cities out of Boston than Asheville. Introducing nonstop flights between DAL and Boston would enable Southwest to design flights for passengers traveling from Dallas to any of the numerous cities connected through Boston. In a way, this would be like simultaneously entering a multiple stop market like DAL – X, where X is a city Southwest flies to from Boston Logan. In other words, entering the DAL – Boston market non-stop would also feed more traffic into Boston, thereby increasing the demand of Southwest's other flights out of Boston.

A smaller fixed cost of entry would also motivate Southwest to enter some markets and not others. Markets where Southwest has large airport presence at end-points already have the fixed infrastructure (gate space, baggage handling, ticketing kiosks, etc.) in place to accommodate entry. In this way, airport presence may identify fixed costs of entry.

Using airport presence as a strong predictor of entry, three instruments are devised:

1. Connected markets count: average (across year quarters) of the total number of markets connected by Southwest from the end-point[24] during the ex-ante period.

2. Passengers Connected: average of the total passengers connected by Southwest from end-point during ex-ante period.

3. RPM Connected: average of total revenue passenger miles[25] from passengers connected by Southwest from end-point during ex-ante period.

The underlying assumption for the IVs mentioned to be used in a fare regression equation is that fares changes on say, DAL-Boston Logan are not directly affected by the fact that Southwest had bigger airport presence in Boston Logan in the ex-ante period. In other words, the entry decision is correlated to ex-ante airport presence, but the change in market fares in the markets entered is exogenous to ex-ante airport presence.

Since route specific marginal costs of the firm and rival firms are unobserved in the dataset,

[24] E.G. for DAL – VEGAS, the end-point is Vegas, and the average number of markets Southwest operated flights through Vegas during ex-ante year-quarters (i.e., the instrument's value) was 85.1. Symmetric definitions apply for other two instruments based on passenger count and RPM.

[25] RPM = number of passengers*distance travelled. RPM is widely used as a measure of traffic in airline markets.

and these influence the fare change, the validity of the instrument is questionable if the instrument is correlated with such unobserved marginal costs. It is reasonable to assume that the instruments discussed above are uncorrelated to rival firms' marginal costs, because the instruments are based on Southwest's operation.

However, the instrument could be correlated with Southwest's marginal costs in few possible ways. For instance, one possibility could be if strong economies of scale exist in airport services such as baggage handling. The baggage handling costs of serving one more passenger at an airport where the airline already handles a large volume of baggage may be smaller than in the case where the airline had a smaller volume of baggage handling.

Another possibility could be that the airline allocates more efficient manpower and machinery in airports where it serves many more markets and passengers. Marginal costs of operation at an airport where more efficient resources are used would be smaller. Such possibilities lead to an upward bias (larger magnitude) on the size of fare decrease resulting from entry.

If an airport does not have excess capacity, it is also quite possible that markets where the airline has bigger presence is congested (due to multitude of operations), and thus, the marginal cost of serving more passengers rises with increase in passengers. This factor, if more pronounced that the aforementioned influences on marginal cost, would lead to a downward bias (smaller magnitude) on the size of fare decrease resulting from entry.

The results for the IV specification are presented in Tables 1.10 and 1.11. Note that Treatment 1a refers to non-stop markets entered by Southwest following repeal (DAL – outside WP). All specifications were run with market, year and airline fixed effects. The control group and all other covariates are the same as in the baseline regressions.

We see the value of the coefficient of interest (ex-post * treatment 1a) is negative, and significant at the one percent significance level (implying 11% fare decrease with CONTROL A, and 7.3% decrease with B). The magnitude of the coefficient is slightly smaller when instruments are used in a two-stage least squares regression (implying 10.4% fare decrease with A, and 6% with B)[26]. Not surprisingly, fares decrease in the markets Southwest enters, but perhaps the decrease in fares

[26]These magnitudes use specification (4) since it employs the strongest instrument.

would not have been so large had they made a pure random entry decision. The IV is introduced to dampen the selection bias; in other words, the coefficient we obtain using the IVs are closer to the true coefficient of entry on fares if Southwest had entered routes on a pure random basis.

We find that Southwest's entry has a reasonably large effect on fares, even after controlling for the endogeneity of the price decision. This suggests that many US airline markets are not very competitive. Of course, the results are specific to Southwest, so the price effect might be smaller if the entrant was some higher cost airline.

Naturally, one might be interested in quantifying fare changes in markets that Southwest did not enter after WA repeal (Treatment 1b with reference to the baseline regressions). Regressions run using market, year and airline fixed effects is reported in Table 1.12. Specifications (1) and (2) use control groups A and B respectively. The results show that fare change was not very pronounced in these markets.

1.5.4 Investigating Output and Capacity changes

In the price analyses, we observed that following the Oct. 13, 2014 policy changes at Love Field, fares decreased in DAL–outside WP and DFW–outside WP, whereas fares increased in DAL–inside WP and DFW–inside WP. These results indicate that output and capacity in DAL–outside WP could have increased at the cost of DAL–inside WP. In other words, Southwest Airlines redeployed resources from DAL–inside WP markets to DAL–outside WP markets.

Analysis begins by constructing totals of seats, flights and passengers in the different markets using the same definitions of ex-ante and ex-post periods of WA repeal. The results are reported in Table 1.2. Total seat and flight data is only available on the segment level i.e., only for non-stop markets. Seats and flights represent capacity in the airline industry, whereas passenger count is a measure of output. As seen from Table 1.2, output and capacity in DAL-outside WP markets increased, whereas it decreased in the DAL-inside WP markets[27]. The magnitudes for the corresponding DFW markets are smaller.

[27]Note that even before the WA was fully repealed, airlines could serve DAL-outside WP markets in aircrafts with not more than 56 seats (discussed under Introduction). This is responsible for the small number of non-stop total seats and flights in the ex-ante period.

The output change can also be investigated using regression analysis. Using OLS regressions

Table 1.2: Output and capacity during ex-ante and ex-post time periods

	NON-STOP				ALL STOPS	
Markets and Time Period	Total Seats	% Change	Total Flights	% Change	Total Passengers	% Change
DAL-outside WP						
Ex-ante	165,346	4993.5%	3,018	1826.5%	3,597,060	190.1%
Ex-post	8,421,970		58,143		10,434,350	
DFW-outside WP						
Ex-ante	41,593,612	4.2%	338,471	-2.3%	30,744,010	7.7%
Ex-post	43,355,095		330,529		33,109,100	
DAL-WP						
Ex-ante	10,233,436	-22.4%	76,867	-22.8%	5,784,550	-9.7%
Ex-post	7,938,252		59,379		5,224,900	
DFW-WP						
Ex-ante	15,143,729	0.2%	182,008	0.3%	4,221,420	2.5%
Ex-post	15,179,761		182,487		4,325,490	

with market, year and airline fixed effects, and the logarithm of market-level passenger totals as the dependent variable, results are reported in Table 1.13. The two specifications differ in the choice of control groups as in regressions in preceeding sections. The coefficients of the variables of interest (ex-post*treatment) reveal that output increased in DAL – outside WP markets (117-119 %) but decreased in DAL – inside WP markets (19%). Put together with the observation that fares decreased in DAL–outside WP, and increased in DAL–inside WP, the findings indicate that Southwest faced a binding capacity constraint due to gate restrictions that were simultaneously introduced with the repeal of the Wright Amendment.

The results show that output decreased in DFW–outside WP markets (4-6%). This is reasonable since the introduction of non-stop long-haul flights from the Dallas region to outside WP destinations from the neighboring DAL airport might have led to consumers switching to flying out from DAL rather than DFW.

Output decreased in DFW-inside WP markets (8-9%). It is likely that this is the consequence of monopoly power being exercised by DFW-based airlines since competing DAL-based Southwest exits the adjacent short-haul routes. Another likely scenario is that following the exit of Southwest in the adjacent DAL-WP markets, price-sensitive leisure passengers could have switched to other forms of transportation (road, train, etc.) for travelling within these short-haul markets. As legacy carriers serve DFW-inside WP sectors, it is likely that DFW is used mostly by business passengers who are less price but more time sensitive with regard to getting to their destination city. Following the loss of leisure passengers, and the existence of primarily business passengers in these markets, legacy carriers decreased output within DFW-inside WP to sharply increase fares.

The coefficients of other variables also seem reasonable: logarithm of distance squared is negative (but not significant), suggesting that demand for air travel grows with distance, but air travel between origin destination pairs that have extremely large distances generates disutility. The logarithm of the distance variable is omitted in the regression since it is multicollinear with the log of distance squared. The population variable is positive with a large magnitude, the reason being straightforward: bigger population means higher passengers.

1.6 Conclusion

Using the case study of the Wright Amendment repeal and gate restrictions, this paper has assessed the impact of relaxing some, and simultaneously introducing other entry barriers in the airline industry. The findings show that the policy changes of Oct.13, 2014 led to a decrease in fares in airline markets that connect Dallas to destinations outside the Wright Perimeter, but increase in fares in markets that connect Dallas to Wright Perimeter destinations. Quantity regressions showed that output decreased in DAL–WP, but increased in DAL–outside WP markets. The opposite price and quantity changes in different markets are tied together with the framework of a simple capacity constrained entry model. The findings indicate that the gate restrictions imposed on airlines operating out of DAL were binding, which leads to them trading off operating flights in short-haul markets by increasing output in long-haul markets that connect Dallas to destinations beyond the neighboring states of Texas.

It is interesting to note that the magnitude of fare changes observed in this research is quite small compared with the magnitude of fare change observed by Bold (2013) when he studied the impact of the Wright Amendment Reform act of 2006 that allowed airlines to operate flights anywhere from DAL if they made a stopover within the Wright Perimeter (he finds that fares dropped

17 percent). A possible reason is that full repeal allows airlines to fly non-stop from DAL to anywhere in the USA, whereas previously they were already operating the same markets with at least one stopover. A non-stop flight is a higher quality product for which consumers would be willing to pay a higher amount. Therefore, adding non-stop service to a market that had a multi-stop service may not yield a large fare difference.

Another contribution of this paper to the airline entry literature is the introduction of an instrumental variables technique to address endogenous entry by airlines following a policy change. The two-staged least squares with fixed effects regressions show that markets (non-directional airport pairs) where Southwest introduced non-stop services experienced much higher fare changes than those where they did not. Using instrumental variables decreases the magnitude of the coefficient relating the impact of Southwest's entry on fares, but by a small about. It could be quite likely that the possible correlation of Southwest's unobserved marginal costs with the instruments are leading to an upward bias of the coefficient on entry. It is also likely that endogeneity is not creating a large bias because of which the coefficients do not move by much. In the lack of detailed dataset on costs, these speculations are difficult to empirically investigate.

A full analysis of the impact of the October 13, 2014 policy changes at Love Field on welfare is difficult since some markets benefitted while others lost. The losers were the markets for which Southwest reduced service because of the capacity constraint and the DFW markets which were cannibalized by Southwest's entry. It is quite likely that the policy changes led to an increase in Southwest's profits since they moved services to more profitable markets. Other rival carriers might have experienced a decrease in profits due to the increase in competition arising from Southwest's entry. The overall impact on producer surplus is hard to investigate in the absence of route-specific cost data.

A detailed analysis of the impact of the policy changes on consumer surplus would require a structural approach. However, structural analysis using publicly available DB1B data requires some strong assumptions. An attempt to gauge the change in consumer surplus is made in this study by assuming a constant elasticity demand curve. Following the estimates produced by IATA[28], a value

[28] IATA estimates that short-haul and long-haul Intra North American air travel routes have price elasticity of demand equalling 1.5 and 1.4 respectively.

of 1.5 is used as the price elasticity of demand for DAL - WP and DFW - WP markets, and a value of 1.4 is used for DAL - outside WP and DFW - outside WP. The estimates of consumer welfare changes for the different markets using the two control groups are presented in Table 1.3. As seen from the table, the overall change in consumer surplus is estimated to be from 130 to 330 million USD. The large positive increase in consumer surplus in DAL - outside WP markets is enough to offset the negative changes in the other markets.

Overall, this study showcases the cost of excessive regulation in the airline industry. In the absence of the Wright Amendment, Dallas-based passengers could have experienced lower fares and better service in earlier years. It is also reasonable to speculate that the fare drop in Dallas-outside WP markets could have been much larger, and fare rise in Dallas-inside WP markets could have been absent in the absence of gate restrictions imposed at Love Field. As shown in models of neoclassical economic theory, excessive regulation and the resulting market distortions bring about deadweight losses that curb societal welfare. The findings of this paper are in line with such theoretical predictions. Perhaps the crux of this paper is best put in Alfred Kahn's words: "Whenever competition is feasible it is, for all its imperfections, superior to regulation as a means of serving the public interest."

Table 1.3: Consumer welfare calculations with constant elasticity demand curve

Using Control A	
Market	**Δ CS (in millions of USD)**
DAL - outside WP	2,416
DFW - outside WP	(2,097)
DAL - WP	(74)
DFW - WP	(113)

Using Control B	
Market	**Δ CS (in millions of USD)**
DAL - outside WP	2,531
DFW - outside WP	(1,951)
DAL - WP	(110)
DFW - WP	(139)

1.7 References

Baumol, William J., and John C. Panzar. *"Contestable Markets and the Theory of Industry Structure"* New York: Harcourt Brace Jovanovich, 1982. Print.

Berry, Steven T. "Estimation of a Model of Entry in the Airline Industry." *Econometrica*, vol. 60, no. 4, 1992, pp. 889–917.

Boguslaski, C., Ito, H., and Lee, D. "Entry Patterns in the Southwest Airlines Route System." *Review of Industrial Organization*, vol. 25, no.3, 2004, pp. 317–350.

Bold, Tuvshintulga, "Three essays on competition and productivity in the U.S. airline industry" (2014). *Economics Dissertations.* Paper 15.

Borenstein, Severin. "Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry." *The RAND Journal of Economics*, vol. 20, no. 3, 1989, pp. 344–365.

Borenstein, Severin, and Nancy L. Rose. "Competition and Price Dispersion in the U.S. Airline Industry." *Journal of Political Economy*, vol. 102, no. 4, 1994, pp. 653–683.

Bresnahan, Timothy F., and Peter C. Reiss. "Entry in Monopoly Markets." *The Review of Economic Studies*, vol. 57, no. 4, 1990, pp. 531–553.

Bresnahan, Timothy F., and Peter C. Reiss. "Entry and Competition in Concentrated Markets." *Journal of Political Economy*, vol. 99, no. 5, 1991, pp. 977–1009.

Ciliberto, Federico, and Elie Tamer. "Market Structure and Multiple Equilibria in Airline Markets." *Econometrica*, vol. 77, no. 6, 2009, pp. 1791–1828.

Dana, James D. Jr., and Eugene Orlov."Internet Penetration and Capacity Utilization in the US Airline Industry." *American Economic Journal: Microeconomics*, 2014, 6(4): 106-37.

Daniel, Joseph I., "Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues," *Econometrica*, March 1995, 63(2), pp. 327-70.

Goolsbee, Austan, and Chad Syverson. "How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines." *The Quarterly Journal of Economics*, vol. 123, no. 4, 2008, pp. 1611–1633.

Government Accountability Office, "Slot Controlled Airports," Report to the Committee on Commerce, Science, and Transportation, U.S. Senate, September 2012.

Jan K. Brueckner, “Airport Congestion when Carriers have Market Power,” *American Economic Review*, December, vol. 92, 2002, pp. 1357-1375.

Kwoka, John, and Evgenia Shumilkina. “The Price Effect of Eliminating Potential Competition: Evidence from an Airline Merger” *The Journal of Industrial Economics*, vol. 58, no. 4, 2010, pp. 767–793.

Kwoka, J., Hearle, K., & Alepin, P. “From the Fringe to the Forefront: Low Cost Carriers and Airline Price Determination.” *Review of Industrial Organization,* 2016, *48*, 247-268.

Love Terminal Partners, et al., Plaintiffs, v. The United States, Defendant. (2011) United States Court of Federal Claims, No. 08-536 L.

Mazzeo, Michael J. “Product Choice and Oligopoly Market Structure.” *The RAND Journal of Economics*, vol. 33, no. 2, 2002, pp. 221–242.

Pearce, Brian, and Smyth, Mark, “Air Travel Demand,” *IATA Economics Briefing*, no. 9, April 2008.

Seim, Katja. “An Empirical Model of Firm Entry with Endogenous Product-Type Choices.” *The RAND Journal of Economics*, vol. 37, no. 3, 2006, pp. 619–640.

Steven A. Morrison and Clifford Winston, “Empirical Implications and Tests of the Contestability Hypothesis,” *Journal of Law and Economics*, 30, April 1987, pp. 53-66.

Steven A. Morrison, “Actual, Adjacent, and Potential Competition: Estimating the Full Effect of Southwest Airlines,” *Journal of Transport Economics and Policy*, Volume 35, Part 2, May 2001, pp. 239-256.

Steven A. Morrison and Clifford Winston. “*The Economic Effects of Airline Deregulation*,” Washington, D.C.: Brookings Institution, 1986. Print.

1.8 Regression Tables

Table 1.4: Baseline regression results using CONTROL-A. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
Post*treat 1 (DAL-outside WP)	-0.0799***	-0.0897***	-0.101***	-0.0804***
	(0.0105)	(0.0106)	(0.00965)	(0.00951)
Post*treat 2 (DFW-outside WP)	-0.0517***	-0.0552***	-0.0590***	-0.0570***
	(0.00601)	(0.00600)	(0.00551)	(0.00545)
Post*treat 3 (DAL-inside WP)	0.0358	0.0190	0.0332	0.0613***
	(0.0229)	(0.0228)	(0.0232)	(0.0223)
Post*treat 4 (DFW-inside WP)	0.164***	0.148***	0.154***	0.142***
	(0.0153)	(0.0153)	(0.0112)	(0.0108)
Distance	0.363***	0.353***	0.281***	0.282***
	(0.00331)	(0.00328)	(0.00823)	(0.00812)
Population	-0.0165***	-0.0229***	0.249***	-0.00225
	(0.00107)	(0.00109)	(0.0964)	(0.118)
Slot	0.0460***	0.0494***	0	0
	(0.0104)	(0.0104)	(.)	(.)
Hub	0.114***	0.119***	0	0
	(0.00476)	(0.00471)	(.)	(.)
Tourist	-0.0948***	-0.0820***	0	0
	(0.00433)	(0.00433)	(.)	(.)
Effective Competitors		0.0324***	-0.00689***	-0.00786***
		(0.00122)	(0.00184)	(0.00177)
Observations	71001	71001	71001	71001
Adjusted R^2	0.372	0.378	0.559	0.597
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	No	No	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[28]Some coefficients have been omitted due to space constraints. The appendix presents regression results using alternative specifications. Standard errors are in parentheses.

Table 1.5: Baseline regression results using CONTROL-B. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
Post*treat 1 (DAL-outside WP)	-0.0173	-0.0277	-0.0593***	-0.0469***
	(0.0211)	(0.0207)	(0.0156)	(0.0146)
Post*treat 2 (DFW-outside WP)	0.0188	0.0121	-0.0178	-0.0282**
	(0.0193)	(0.0189)	(0.0131)	(0.0121)
Post*treat 3 (DAL-inside WP)	0.103***	0.0883***	0.0781***	0.0969***
	(0.0297)	(0.0293)	(0.0275)	(0.0256)
Post*treat 4 (DFW-inside WP)	0.234***	0.219***	0.197***	0.173***
	(0.0239)	(0.0236)	(0.0168)	(0.0156)
Distance	0.412***	0.400***	0.380***	0.453***
	(0.00645)	(0.00651)	(0.0115)	(0.0114)
Population	-0.0577***	-0.0602***	-0.0972	-0.429**
	(0.00182)	(0.00182)	(0.135)	(0.185)
Slot	0.144***	0.141***	0	0
	(0.0110)	(0.0110)	(.)	(.)
Hub	0.504***	0.480***	0	0
	(0.0196)	(0.0193)	(.)	(.)
Tourist	-0.0603***	-0.0507***	0	0
	(0.00667)	(0.00672)	(.)	(.)
Effective Competitors		0.0206***	-0.00822***	-0.0108***
		(0.00157)	(0.00297)	(0.00277)
Observations	34906	34906	34906	34906
Adjusted R^2	0.335	0.338	0.474	0.552
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	No	No	Yes	Yes

Table 1.6: Regression results for non-stopover markets using CONTROL-A. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
Post*treat (DAL-WP, only gate)	0.0698*	0.0440	0.0543	0.0794**
	(0.0369)	(0.0369)	(0.0441)	(0.0389)
Distance	0.367***	0.354***	0.161***	0.131***
	(0.00384)	(0.00376)	(0.0226)	(0.0209)
Population	0.00917***	-0.00309**	0.578***	0.384*
	(0.00139)	(0.00146)	(0.184)	(0.214)
Slot	0	0	0	0
	(.)	(.)	(.)	(.)
Hub	0.0802***	0.0917***	0	0
	(0.00501)	(0.00497)	(.)	(.)
Tourist	-0.0428***	-0.0368***	0	0
	(0.00595)	(0.00584)	(.)	(.)
Effective Competitors		0.0512***	-0.00453	-0.00579**
		(0.00178)	(0.00286)	(0.00283)
Observations	39461	39461	39461	39461
Adjusted R^2	0.387	0.403	0.641	0.660
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	No	No	Yes	Yes

Table 1.7: Regression results for non-stopover markets using CONTROL-B. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
Post*treat (DAL-WP, only gate)	0.0221	-0.00654	0.0703	0.0984**
	(0.0441)	(0.0417)	(0.0440)	(0.0436)
Distance	0.642***	0.476***	0.548***	0.215**
	(0.0188)	(0.0194)	(0.138)	(0.0832)
Population	-0.0305***	-0.0317***	-0.479	-0.107
	(0.00621)	(0.00611)	(0.566)	(0.679)
Slot	0	0	0	0
	(.)	(.)	(.)	(.)
Hub	0.256***	0.222***	0	0
	(0.0289)	(0.0272)	(.)	(.)
Tourist	-0.131***	-0.117***	0	0
	(0.0236)	(0.0224)	(.)	(.)
Effective Competitors		0.124***	-0.00379	-0.0152
		(0.00461)	(0.00808)	(0.00942)
Observations	3366	3366	3366	3366
Adjusted R^2	0.539	0.604	0.809	0.870
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	No	No	Yes	Yes

Table 1.8: Investigating heterogenous effects on non-stop and multiple-stop flights using CONTROL-A. Dependent variable is logarithm of coupon-specific average market-airline fare.

	(1)	(2)	(3)	(4)
Post*Non-stop	-0.0427***	-0.0536***	-0.0491***	-0.0496***
	(0.00848)	(0.00852)	(0.00776)	(0.00749)
Post*Multi-stop	-0.0137*	-0.00951	-0.0219***	-0.0178**
	(0.00773)	(0.00775)	(0.00742)	(0.00725)
Distance	0.340***	0.332***	0.0603***	0.0878***
	(0.00283)	(0.00283)	(0.00876)	(0.00867)
Population	-0.0129***	-0.0186***	0.403***	0.0881
	(0.000908)	(0.000940)	(0.0937)	(0.115)
Slot	0.0452***	0.0479***	0	0
	(0.00892)	(0.00891)	(.)	(.)
Hub	0.119***	0.125***	0	0
	(0.00410)	(0.00409)	(.)	(.)
Tourist	-0.0586***	-0.0487***	0	0
	(0.00389)	(0.00389)	(.)	(.)
Effective Competitors		0.0260***	-0.00345**	-0.00443***
		(0.00109)	(0.00176)	(0.00171)
Observations	94057	94057	94057	94057
Adjusted R^2	0.296	0.301	0.455	0.495
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	No	No	Yes	Yes

Table 1.9: Investigating heterogenous effects on non-stop and multiple-stop flights using CONTROL-B. Dependent variable is logarithm of coupon-specific average market-airline fare.

	(1)	(2)	(3)	(4)
Post*Non-stop	-0.0412***	-0.0507***	-0.0494***	-0.0477***
	(0.00840)	(0.00845)	(0.00859)	(0.00806)
Post*Multi-stop	0.0496***	0.0467***	0.0143	0.00945
	(0.0177)	(0.0173)	(0.0141)	(0.0132)
Distance	0.343***	0.328***	0.0163	0.154***
	(0.00639)	(0.00645)	(0.0132)	(0.0134)
Population	-0.0462***	-0.0496***	0.0893	-0.333*
	(0.00177)	(0.00178)	(0.138)	(0.190)
Slot	0.126***	0.123***	0	0
	(0.00959)	(0.00959)	(.)	(.)
Hub	0.513***	0.494***	0	0
	(0.0173)	(0.0170)	(.)	(.)
Tourist	-0.0395***	-0.0268***	0	0
	(0.00652)	(0.00660)	(.)	(.)
Effective Competitors		0.0217***	-0.00135	-0.00464*
		(0.00152)	(0.00291)	(0.00276)
Observations	44090	44090	44090	44090
Adjusted R^2	0.196	0.199	0.317	0.404
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	No	No	Yes	Yes

Table 1.10: IV regression results using CONTROL-A for markets entered by Southwest Airlines. Dependent variable is logarithm of average market-airline fare.

	(1)	(2)	(3)	(4)
Post*treat 1(a)	-0.117***	-0.110***	-0.109***	-0.110***
	(0.0111)	(0.0176)	(0.0168)	(0.0113)
Distance	0.185***	0.185***	0.185***	0.185***
	(0.0111)	(0.0111)	(0.0111)	(0.0111)
Population	0.289**	0.264*	0.262*	0.266*
	(0.145)	(0.152)	(0.151)	(0.145)
Effective Competitors	-0.00649***	-0.00659***	-0.00660***	-0.00658***
	(0.00211)	(0.00212)	(0.00211)	(0.00211)
Observations	42583	42583	42583	42583
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Instruments	None	RPM	Population	Markets
First stage F	-	1401.8	1595.9	43361.2

Table 1.11: IV regression results using CONTROL-B for markets entered by Southwest Airlines. Dependent variable is logarithm of average market-airline fare.

	(1)	(2)	(3)	(4)
Post*treat 1(a)	-0.0763***	-0.0541*	-0.0502	-0.0624***
	(0.0163)	(0.0328)	(0.0311)	(0.0169)
Distance	0.544***	0.546***	0.546***	0.545***
	(0.0307)	(0.0307)	(0.0307)	(0.0307)
Population	-0.506	-0.647	-0.672	-0.594
	(0.393)	(0.433)	(0.428)	(0.394)
Effective Competitors	-0.0140**	-0.0152**	-0.0154***	-0.0148**
	(0.00573)	(0.00595)	(0.00592)	(0.00574)
Observations	6488	6488	6488	6488
Quarter dummies	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Instruments	None	RPM	Population	Markets
First stage F	-	464.1	490.1	7115.3

Table 1.12: Fixed effects regression results for markets not entered by Southwest Airlines. Dependent variable is logarithm of average market-airline fare.

	(1)	(2)
Post*treat 1(b)	-0.0277*	-0.00570
	(0.0149)	(0.0178)
Distance	0.133***	0.233***
	(0.0115)	(0.0480)
Population	0.390***	0.0606
	(0.145)	(0.388)
Effective Competitors	-0.00616***	-0.0141***
	(0.00208)	(0.00545)
Observations	40760	4665
Adjusted R^2	0.653	0.822
Quarter dummies	Yes	Yes
Fixed effects	Yes	Yes
Control group	A	B

Table 1.13: Fixed effects regression results to investigate changes in output in related markets. Dependent variable is the logarithm of market-level passenger totals.

	(1)	(2)
Post*treat 1 (DAL-outside WP)	0.789***	0.774***
	(0.00939)	(0.0146)
Post*treat 2 (DFW-outside WP)	-0.0459***	-0.0663***
	(0.00540)	(0.0121)
Post*treat 3 (DAL-inside WP)	-0.205***	-0.213***
	(0.0220)	(0.0255)
Post*treat 4 (DFW-inside WP)	-0.0783***	-0.0886***
	(0.0107)	(0.0156)
Ln(distance)	0	0
	(.)	(.)
$Ln(distance)^2$	-0.0101**	-0.0220***
	(0.00402)	(0.00570)
Population	1.206***	1.187***
	(0.117)	(0.184)
Observations	71001	34906
Adjusted R^2	0.975	0.974
Quarter dummies	Yes	Yes
Fixed effects	Yes	Yes
Control group	A	B

1.9 Appendix

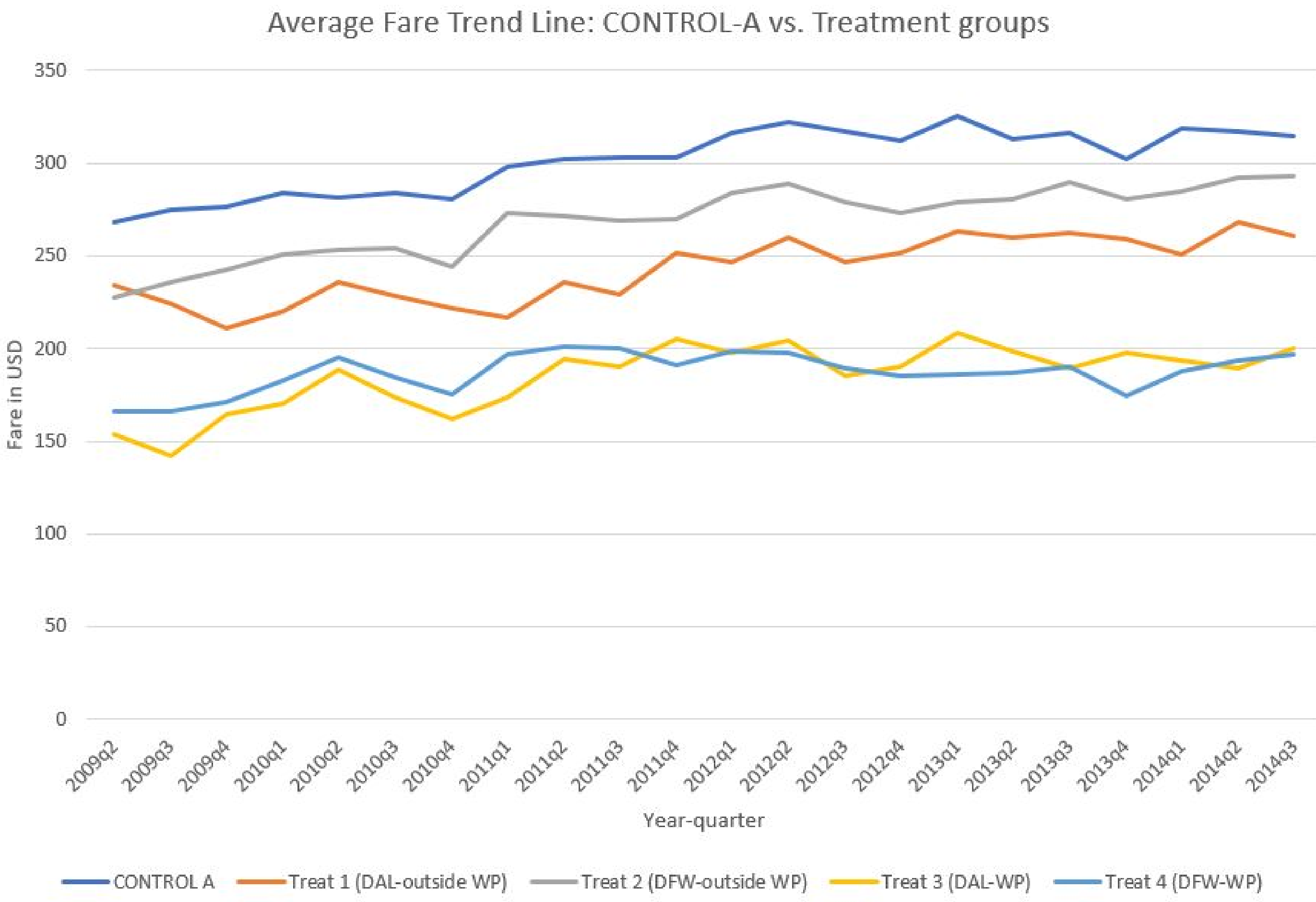


Figure 1.3: AVERAGE FARE TREND LINE (CONTROL A)

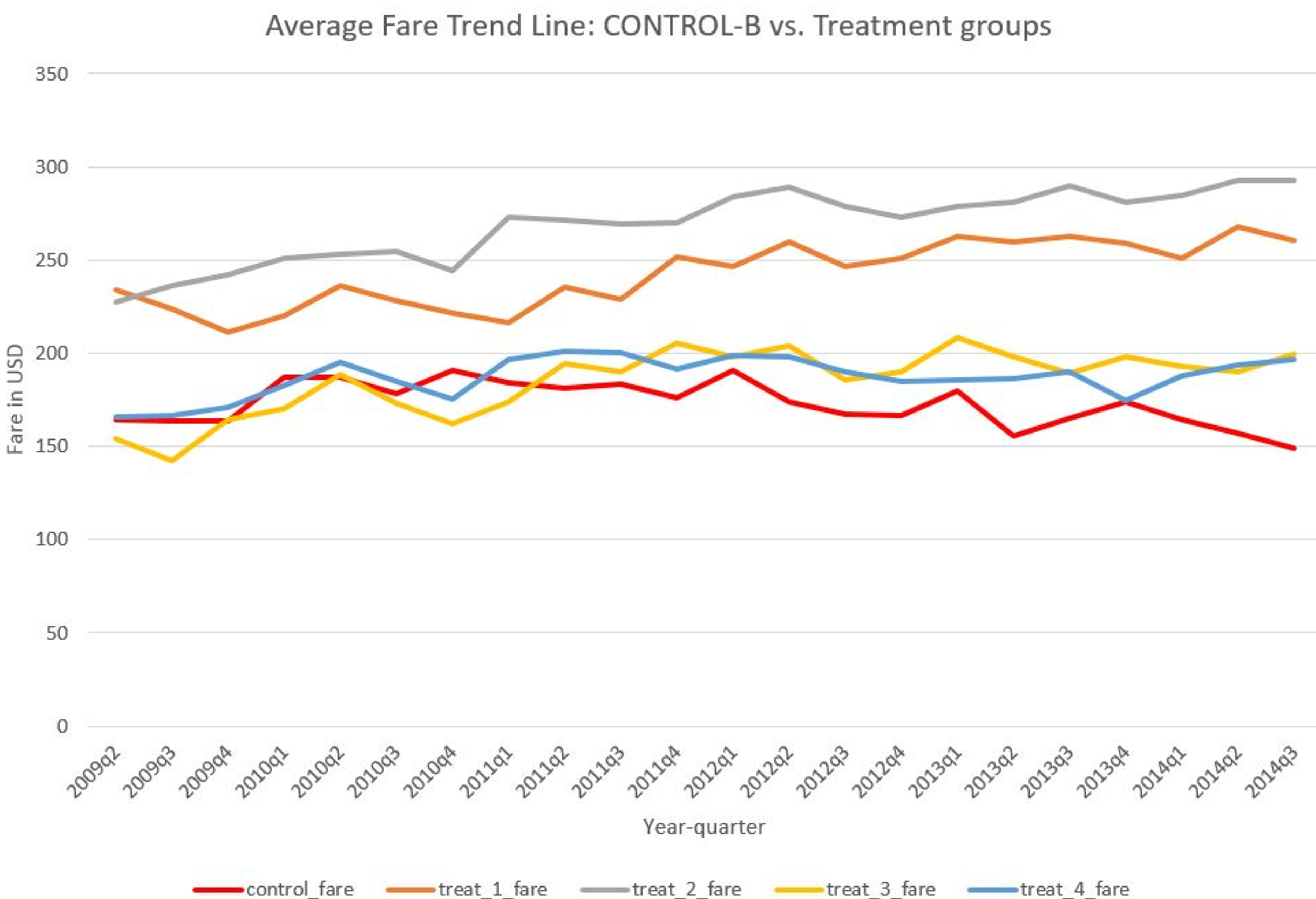


Figure 1.4: Average Fare Trend Line (control B)

Dallas, TX: Dallas Love Field (DAL)

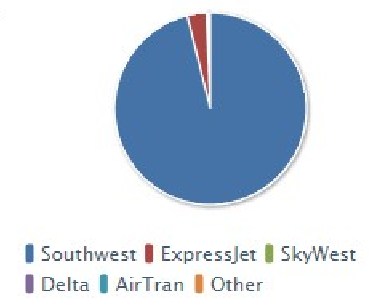
Summary Data (U.S. Flights Only)

Passengers*	2013**	2014**	%Chg	Rank***
Arrival	3,982k	4,190k	5.22%	42
Departure	3,975k	4,196k	5.57%	42
Scheduled Flights				
Departures	47,234	45,879	-2.87%	47
Freight/Mail (lb.) (Scheduled and Non-Scheduled)				
Total	22m	21m	-3.64%	132
Carriers				
Scheduled	7	9	28.57%	

* Scheduled enplaned revenue passengers.
** 12 months ending October of each year.
*** Among 819 U.S. airports, 12 months ending October 2014

Carrier Shares for November 2013 - October 2014

Carrier	Passengers	Share
Southwest	8,060	96.11%
ExpressJet	269	3.21%
SkyWest	16.15	0.19%
Delta	5.30	0.06%
AirTran	0.18	0.00%
Other	34.86	0.42%

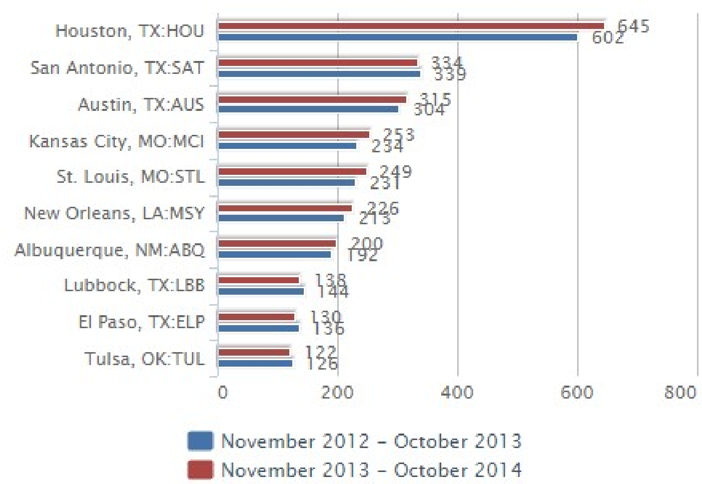


Based on enplaned passengers(000) both arriving and departing.

Total Passengers (U.S. Flights, in thousands)

* Before October 2002, only carriers operating aircraft with more than 60 seats or 18,000 pounds in payload reported traffic data.
** 2014 represents data for November 2013 - October 2014.

Top 10 Destination Airports (U.S. Only, Passengers (000))



Source: T-100 Domestic Market (US Carriers).

Figure 1.5: Airport Snapshot: Dallas Love Field (source: BTS)

Dallas/Fort Worth, TX: Dallas/Fort Worth International (DFW)

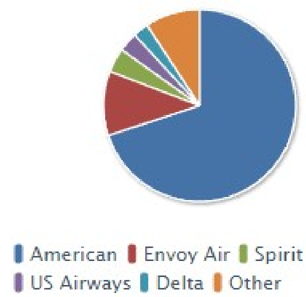
Summary Data (U.S. Flights Only)

Passengers*	2013**	2014**	%Chg	Rank***
Arrival	25,898k	27,068k	4.52%	3
Departure	25,873k	27,027k	4.46%	3
Scheduled Flights				
Departures	299,028	293,943	-1.70%	3
Freight/Mail (lb.) (Scheduled and Non-Scheduled)				
Total	677m	713m	5.43%	12
Carriers				
Scheduled	24	26	8.33%	

* Scheduled enplaned revenue passengers.
** 12 months ending October of each year.
*** Among 819 U.S. airports, 12 months ending October 2014

Carrier Shares for November 2013 - October 2014

Carrier	Passengers	Share
American	37,836	69.94%
Envoy Air	5,864	10.84%
Spirit	2,223	4.11%
US Airways	1,761	3.26%
Delta	1,412	2.61%
Other	4,998	9.24%

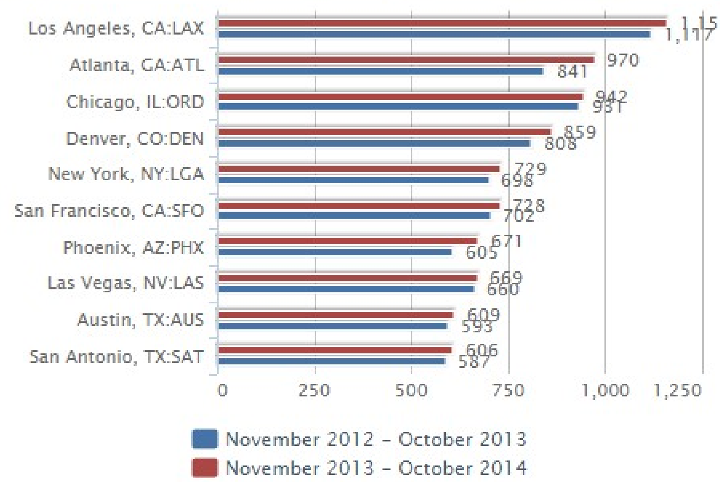


Based on enplaned passengers(000) both arriving and departing.

Total Passengers (U.S. Flights, in millions)

* Before October 2002, only carriers operating aircraft with more than 60 seats or 18,000 pounds in payload reported traffic data.
** 2014 represents data for November 2013 - October 2014.

Top 10 Destination Airports (U.S. Only, Passengers (000))



Source: T-100 Domestic Market (US Carriers).

Figure 1.6: Airport Snapshot: Dallas Fort Worth (source: BTS)

Table 1.14: Baseline regression results using outside wp - outside wp as control. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
	ln_fare	ln_fare	ln_fare	ln_fare
Post*treat 1 (DAL-outside WP)	-0.0895***	-0.0940***	-0.105***	-0.0869***
	(0.00988)	(0.00988)	(0.0163)	(0.0169)
Post*treat 2 (DFW-outside WP)	-0.0595***	-0.0608***	-0.0631***	-0.0612***
	(0.00466)	(0.00467)	(0.00964)	(0.00927)
Post*treat 3 (DAL-inside WP)	0.0366*	0.0291	0.0286	0.0619*
	(0.0219)	(0.0219)	(0.0354)	(0.0364)
Post*treat 4 (DFW-inside WP)	0.154***	0.147***	0.149***	0.139***
	(0.0148)	(0.0148)	(0.0254)	(0.0241)
Distance	0.306***	0.301***	0.210***	0.193***
	(0.000755)	(0.000755)	(0.00654)	(0.00590)
Population	-0.0414***	-0.0453***	0.286***	-0.0892
	(0.000253)	(0.000263)	(0.0451)	(0.0570)
Slot	0.104***	0.103***	0	0
	(0.00127)	(0.00127)	(.)	(.)
Hub	0.0521***	0.0561***	0	0
	(0.00102)	(0.00102)	(.)	(.)
Tourist	-0.0510***	-0.0445***	0	0
	(0.000993)	(0.00100)	(.)	(.)
Effective Competitors		0.0153***	-0.00690***	-0.00792***
		(0.000308)	(0.000831)	(0.000804)
constant	4.582***	4.668***	-3.941***	6.804***
	(0.00932)	(0.00938)	(1.271)	(1.608)
Observations	793688	793688	793688	793688
R^2	0.294	0.296	0.013	0.078
Adjusted R^2	0.294	0.296	0.013	0.078
F	13982.0	13363.0	194.4	427.3
rmse	0.345	0.344	0.294	0.284

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.15: Investigating heterogenous effects on non-stop and multiple-stop flights using CONTROL-A. Dependent variable is logarithm of coupon-specific average market-airline fare.

	(1) ln_fare	(2) ln_fare	(3) ln_fare	(4) ln_fare
Non-stop	-0.0594***	-0.0443***	0	0
	(0.00590)	(0.00595)	(.)	(.)
Multi-stop	-0.0975***	-0.0902***	0	0
	(0.00620)	(0.00622)	(.)	(.)
post	0.00434	0.0108***	-0.00643*	0.0117**
	(0.00326)	(0.00324)	(0.00388)	(0.00537)
Post*Non-stop	-0.0427***	-0.0536***	-0.0491***	-0.0496***
	(0.00848)	(0.00852)	(0.00776)	(0.00749)
Post*Multi-stop	-0.0137*	-0.00951	-0.0219***	-0.0178**
	(0.00773)	(0.00775)	(0.00742)	(0.00725)
Distance	0.340***	0.332***	0.0603***	0.0878***
	(0.00283)	(0.00283)	(0.00876)	(0.00867)
Population	-0.0129***	-0.0186***	0.403***	0.0881
	(0.000908)	(0.000940)	(0.0937)	(0.115)
Slot	0.0452***	0.0479***	0	0
	(0.00892)	(0.00891)	(.)	(.)
Hub	0.119***	0.125***	0	0
	(0.00410)	(0.00409)	(.)	(.)
Tourist	-0.0586***	-0.0487***	0	0
	(0.00389)	(0.00389)	(.)	(.)
Effective Competitors		0.0260***	-0.00345**	-0.00443***
		(0.00109)	(0.00176)	(0.00171)
constant	3.592***	3.716***	-5.664**	2.796
	(0.0348)	(0.0350)	(2.541)	(3.114)
Observations	94057	94057	94057	94057
R^2	0.297	0.301	0.463	0.502
Adjusted R^2	0.296	0.301	0.455	0.495
F	3135.3	2976.4	38.22	229.8
rmse	0.386	0.384	0.339	0.327

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.16: Investigating heterogenous effects on non-stop and multiple-stop flights using CONTROL-B. Dependent variable is logarithm of coupon-specific average market-airline fare.

	(1) ln_fare	(2) ln_fare	(3) ln_fare	(4) ln_fare
Non-stop	-0.0108*	-0.00175	0	0
	(0.00609)	(0.00614)	(.)	(.)
Multi-stop	-0.0311*	-0.00484	0	0
	(0.0162)	(0.0162)	(.)	(.)
post	-0.0600***	-0.0480***	-0.0285**	-0.0283*
	(0.0163)	(0.0159)	(0.0129)	(0.0148)
Post*Non-stop	-0.0412***	-0.0507***	-0.0494***	-0.0477***
	(0.00840)	(0.00845)	(0.00859)	(0.00806)
Post*Multi-stop	0.0496***	0.0467***	0.0143	0.00945
	(0.0177)	(0.0173)	(0.0141)	(0.0132)
Distance	0.343***	0.328***	0.0163	0.154***
	(0.00639)	(0.00645)	(0.0132)	(0.0134)
Population	-0.0462***	-0.0496***	0.0893	-0.333*
	(0.00177)	(0.00178)	(0.138)	(0.190)
Slot	0.126***	0.123***	0	0
	(0.00959)	(0.00959)	(.)	(.)
Hub	0.513***	0.494***	0	0
	(0.0173)	(0.0170)	(.)	(.)
Tourist	-0.0395***	-0.0268***	0	0
	(0.00652)	(0.00660)	(.)	(.)
Effective Competitors		0.0217***	-0.00135	-0.00464*
		(0.00152)	(0.00291)	(0.00276)
constant	4.036***	4.166***	2.885	14.39***
	(0.0734)	(0.0737)	(4.004)	(5.536)
Observations	44090	44090	44090	44090
R^2	0.196	0.200	0.322	0.409
Adjusted R^2	0.196	0.199	0.317	0.404
F	732.9	712.1	25.77	203.0
rmse	0.405	0.404	0.373	0.349

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.17: First Stage IV regression results using CONTROL-A for markets entered by Southwest Airlines.

	(1) expost_treat3	(2) expost_treat3	(3) expost_treat3
expost_inrpm	5.65e-10***		
	(3.40e-12)		
expost	-0.0249***	-0.0204***	0.00304***
	(0.00214)	(0.00207)	(0.000524)
Distance	0.00176	0.00274	-0.000371
	(0.00379)	(0.00366)	(0.000926)
quarter== 1.0000	0.00241	0.00198	-0.000238
	(0.00170)	(0.00164)	(0.000414)
quarter== 2.0000	-0.00174	-0.00170	-0.000143
	(0.00165)	(0.00159)	(0.000402)
quarter== 3.0000	-0.000415	-0.000632	-0.0000898
	(0.00166)	(0.00160)	(0.000405)
Population	2.059***	1.834***	-0.0119
	(0.0484)	(0.0468)	(0.0121)
Effective Competitors	0.00746***	0.00694***	0.00109***
	(0.000719)	(0.000694)	(0.000175)
expost_inpop		0.000000838***	
		(4.64e-09)	
expost_ins			0.0139***
			(0.0000132)
Constant	-53.26***	-47.44***	0.308
	(1.250)	(1.211)	(0.312)
Observations	42583	42583	42583
R^2	0.528	0.560	0.972
Adjusted R^2	0.514	0.547	0.971
F	1401.8	1595.9	43361.2
rmse	0.0962	0.0928	0.0235

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.18: IV regression results using CONTROL-A for markets entered by Southwest Airlines. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
	ln_fare	ln_fare	ln_fare	ln_fare
treat_3	0	0	0	0
	(.)	(.)	(.)	(.)
expost	-0.00197	-0.0206***	-0.0206***	-0.0207***
	(0.00628)	(0.00661)	(0.00660)	(0.00658)
Post*treat 1(a)	-0.117***	-0.110***	-0.109***	-0.110***
	(0.0111)	(0.0176)	(0.0168)	(0.0113)
Distance	0.185***	0.185***	0.185***	0.185***
	(0.0111)	(0.0111)	(0.0111)	(0.0111)
Population	0.289**	0.264*	0.262*	0.266*
	(0.145)	(0.152)	(0.151)	(0.145)
Effective Competitors	-0.00649***	-0.00659***	-0.00660***	-0.00658***
	(0.00211)	(0.00212)	(0.00211)	(0.00211)
Constant	-3.010	-2.373	-2.319	-2.416
	(3.740)	(3.930)	(3.901)	(3.745)
Observations	42583	42583	42583	42583
F	88.18			
rmse	0.282			

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.19: First Stage IV regression results using CONTROL-B for markets entered by Southwest Airlines.

	(1) expost_treat3	(2) expost_treat3	(3) expost_treat3
expost_inrpm	3.55e-10***		
	(7.79e-12)		
expost	0.133***	0	0.0583***
	(0.0150)	(.)	(0.00468)
Distance	-0.0119	-0.00628	-0.00315
	(0.0205)	(0.0201)	(0.00616)
quarter== 1.0000	0.00705	0.00563	-0.00246
	(0.00838)	(0.00822)	(0.00252)
quarter== 2.0000	-0.00866	-0.00853	-0.00109
	(0.00837)	(0.00821)	(0.00252)
quarter== 3.0000	-0.000162	-0.00107	-0.000166
	(0.00846)	(0.00830)	(0.00254)
Population	4.853***	4.295***	-0.489***
	(0.256)	(0.252)	(0.0798)
Effective Competitors	0.0381***	0.0373***	0.00927***
	(0.00380)	(0.00373)	(0.00115)
expost_inpop		0.000000538***	
		(1.10e-08)	
expost_ins			0.0136***
			(0.0000462)
Constant	-137.2***	-121.5***	13.80***
	(7.229)	(7.131)	(2.253)
Observations	6488	6488	6488
R^2	0.702	0.714	0.973
Adjusted R^2	0.693	0.705	0.972
F	464.1	490.1	7115.3
rmse	0.196	0.192	0.0589

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.20: IV regression results using CONTROL-B for markets entered by Southwest Airlines. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
	ln_fare	ln_fare	ln_fare	ln_fare
treat_3	0	0	0	0
	(.)	(.)	(.)	(.)
expost	-0.00404	-0.0230	-0.0236	-0.0216
	(0.0225)	(0.0238)	(0.0237)	(0.0234)
Post*treat 1(a)	-0.0763***	-0.0541*	-0.0502	-0.0624***
	(0.0163)	(0.0328)	(0.0311)	(0.0169)
Distance	0.544***	0.546***	0.546***	0.545***
	(0.0307)	(0.0307)	(0.0307)	(0.0307)
Population	-0.506	-0.647	-0.672	-0.594
	(0.393)	(0.433)	(0.428)	(0.394)
Effective Competitors	-0.0140**	-0.0152**	-0.0154***	-0.0148**
	(0.00573)	(0.00595)	(0.00592)	(0.00574)
Constant	16.16	20.15*	20.85*	18.65*
	(11.12)	(12.23)	(12.09)	(11.15)
Observations	6488	6488	6488	6488
F	61.77			
rmse	0.293			

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.21: Fixed effects regression results for markets not entered by Southwest Airlines (using CONTROL A). Dependent variable is logarithm of average market-airline fare

	(1)
	ln_fare
treat_1	0
	(.)
expost	-0.0236***
	(0.00653)
Post*treat 1(b)	-0.0277*
	(0.0149)
Distance	0.133***
	(0.0115)
Population	0.390***
	(0.145)
Effective Competitors	-0.00616***
	(0.00208)
Constant	-5.111
	(3.708)
Observations	40760
R^2	0.664
Adjusted R^2	0.653
F	76.83
rmse	0.278

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.22: Fixed effects regression results for markets not entered by Southwest Airlines (using CONTROL B). Dependent variable is logarithm of average market-airline fare

	(1)
	ln_fare
treat_1	0
	(.)
expost	-0.0194
	(0.0217)
Post*treat 1(b)	-0.00570
	(0.0178)
Distance	0.233***
	(0.0480)
Population	0.0606
	(0.388)
Effective Competitors	-0.0141***
	(0.00545)
Constant	2.549
	(10.36)
Observations	4665
R^2	0.830
Adjusted R^2	0.822
F	50.32
rmse	0.272

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.23: Fixed effects regression results to investigate changes in output in related markets (CONTROL A). Dependent variable is the logarithm of market-level passenger totals.

	(1)
	ln_pax
Post*treat 1 (DAL-outside WP)	0.789***
	(0.00939)
Post*treat 2 (DFW-outside WP)	-0.0459***
	(0.00540)
Post*treat 3 (DAL-inside WP)	-0.205***
	(0.0220)
Post*treat 4 (DFW-inside WP)	-0.0783***
	(0.0107)
Ln(distance)	0
	(.)
Ln(distance)2	-0.0101**
	(0.00402)
Population	1.206***
	(0.117)
Slot	0
	(.)
Hub	0
	(.)
Tourist	0
	(.)
eff_cptrs_2	-0.00319*
	(0.00177)
constant	-27.41***
	(3.168)
Observations	71001
R^2	0.976
Adjusted R^2	0.975
F	405.2
rmse	0.291

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.24: Fixed effects regression results to investigate changes in output in related markets (CONTROL B). Dependent variable is the logarithm of market-level passenger totals.

	(1)
	ln_pax
Post*treat 1 (DAL-outside WP)	0.774***
	(0.0146)
Post*treat 2 (DFW-outside WP)	-0.0663***
	(0.0121)
Post*treat 3 (DAL-inside WP)	-0.213***
	(0.0255)
Post*treat 4 (DFW-inside WP)	-0.0886***
	(0.0156)
Ln(distance)	0
	(.)
Ln(distance)2	-0.0220***
	(0.00570)
Population	1.187***
	(0.184)
Slot	0
	(.)
Hub	0
	(.)
Tourist	0
	(.)
eff_cptrs_2	-0.0164***
	(0.00282)
constant	-27.68***
	(5.340)
Observations	34906
R^2	0.974
Adjusted R^2	0.974
F	354.8
rmse	0.305

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.25: Market-time baseline regression results using CONTROL-A. Dependent variable is logarithm of average market fare

	(1)	(2)	(3)	(4)
	ln_fare	ln_fare	ln_fare	ln_fare
Post*treat 1 (DAL-outside WP)	-0.0677***	-0.0729***	-0.0653***	-0.0611***
	(0.0141)	(0.0147)	(0.00738)	(0.00744)
Post*treat 2 (DFW-outside WP)	0.0000800	-0.00117	-0.00306	0.00111
	(0.0108)	(0.0112)	(0.00612)	(0.00619)
Post*treat 3 (DAL-inside WP)	0.0244	-0.0222	0.0273*	0.0320**
	(0.0227)	(0.0228)	(0.0148)	(0.0149)
Post*treat 4 (DFW-inside WP)	0.229***	0.191***	0.227***	0.232***
	(0.0175)	(0.0176)	(0.0110)	(0.0110)
Distance	0.366***	0.337***	0.659***	0.646***
	(0.00417)	(0.00403)	(0.0314)	(0.0314)
Population	-0.0276***	-0.0448***	0.209**	-0.0330
	(0.00151)	(0.00158)	(0.0918)	(0.111)
Slot	0.0157	0.0189	0	0
	(0.0159)	(0.0163)	(.)	(.)
Hub	0.190***	0.199***	0	0
	(0.00749)	(0.00706)	(.)	(.)
Tourist	-0.0783***	-0.0543***	0	0
	(0.00721)	(0.00708)	(.)	(.)
Effective Competitors		0.0913***	-0.00370*	-0.00478**
		(0.00273)	(0.00199)	(0.00199)
constant	3.698***	4.103***	-4.522*	1.963
	(0.0434)	(0.0440)	(2.445)	(2.942)
Observations	15791	15791	15791	15791
R^2	0.427	0.481	0.909	0.910
Adjusted R^2	0.427	0.480	0.899	0.899
F	1060.4	1085.6	102.8	91.92
rmse	0.332	0.316	0.139	0.139

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.26: Baseline regression results using CONTROL-A (five quarters before, and after). Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)	(4)
	ln_fare	ln_fare	ln_fare	ln_fare
Post*treat 1 (DAL-outside WP)	-0.0936***	-0.104***	-0.106***	-0.0934***
	(0.0134)	(0.0134)	(0.0196)	(0.0205)
Post*treat 2 (DFW-outside WP)	-0.0449***	-0.0499***	-0.0423***	-0.0448***
	(0.00779)	(0.00778)	(0.0103)	(0.0103)
Post*treat 3 (DAL-inside WP)	0.0108	-0.00248	0.0185	0.0386
	(0.0297)	(0.0296)	(0.0417)	(0.0406)
Post*treat 4 (DFW-inside WP)	0.118***	0.105***	0.119***	0.113***
	(0.0201)	(0.0202)	(0.0273)	(0.0259)
Distance	0.358***	0.349***	0.290***	0.302***
	(0.00437)	(0.00434)	(0.0231)	(0.0223)
Population	-0.0176***	-0.0238***	-0.256	-0.282
	(0.00140)	(0.00142)	(0.206)	(0.240)
Slot	0.0672***	0.0726***	0	0
	(0.0136)	(0.0136)	(.)	(.)
Hub	0.133***	0.137***	0	0
	(0.00627)	(0.00621)	(.)	(.)
Tourist	-0.0902***	-0.0769***	0	0
	(0.00557)	(0.00559)	(.)	(.)
Effective Competitors		0.0313***	-0.00744**	-0.00689**
		(0.00153)	(0.00305)	(0.00304)
constant	3.586***	3.715***	10.56*	11.30*
	(0.0542)	(0.0540)	(5.586)	(6.516)
Observations	41430	41430	41430	41430
R^2	0.378	0.384	0.025	0.110
Adjusted R^2	0.377	0.384	0.025	0.109
F	1373.8	1341.8	27.10	44.52
rmse	0.366	0.364	0.300	0.287

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.27: Baseline regression (ten quarters before, and after) results using CONTROL-A. Dependent variable is logarithm of average market-airline fare

	(1) ln_fare	(2) ln_fare	(3) ln_fare	(4) ln_fare
Post*treat 1 (DAL-outside WP)	-0.0362***	-0.0443***	-0.0619***	-0.0403**
	(0.00890)	(0.00894)	(0.0158)	(0.0163)
Post*treat 2 (DFW-outside WP)	-0.0328***	-0.0359***	-0.0435***	-0.0435***
	(0.00520)	(0.00519)	(0.0114)	(0.0113)
Post*treat 3 (DAL-inside WP)	0.0508**	0.0324	0.0578	0.0863**
	(0.0202)	(0.0201)	(0.0360)	(0.0355)
Post*treat 4 (DFW-inside WP)	0.165***	0.147***	0.153***	0.136***
	(0.0130)	(0.0130)	(0.0265)	(0.0247)
Distance	0.382***	0.372***	0.294***	0.301***
	(0.00319)	(0.00318)	(0.0199)	(0.0185)
Population	-0.0196***	-0.0250***	0.468***	0.310*
	(0.000905)	(0.000925)	(0.134)	(0.172)
Slot	0.0463***	0.0492***	0	0
	(0.00861)	(0.00863)	(.)	(.)
Hub	0.0685***	0.0773***	0	0
	(0.00521)	(0.00519)	(.)	(.)
Tourist	-0.0874***	-0.0770***	0	0
	(0.00367)	(0.00368)	(.)	(.)
Effective Competitors		0.0293***	-0.0107***	-0.0100***
		(0.00106)	(0.00303)	(0.00297)
constant	3.486***	3.605***	-9.138**	-4.877
	(0.0357)	(0.0357)	(3.642)	(4.677)
Observations	94876	94876	94876	94876
R^2	0.376	0.381	0.024	0.108
Adjusted R^2	0.376	0.381	0.024	0.107
F	3196.3	3118.4	37.27	50.49
rmse	0.369	0.368	0.308	0.295

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 2

Higher Together: Price and Welfare Effects of a Merger between two Low Cost Carriers

2.1 Introduction

Analysis of the effect of mergers and market consolidation on competition and welfare is a topic that has received much attention in the empirical industrial organization literature. Within this topic, much has been written about airline mergers. As the US airline industry has undergone substantial consolidation in recent times, researchers have found mergers in this industry to be interesting regime change instances to use to study the impact of market consolidation. In general, the studies have not been able to agree on the direction of the effect of mergers on prices: some have shown that the unilateral fare increasing effects of the merger dominate efficiency effects, leading to an overall increase in price levels, whereas others have shown the opposite effect. One narrative that becomes apparent from the literature is that the impact of mergers tends to differ on a case by case basis.

One factor that is responsible for the observed differences in merger effects is heterogeneity across merging carriers. There are generally two types of carriers: legacy and low cost carriers (LCCs). Legacy carriers are characterized by their large network presence, which usually takes the form of a hub-and-spoke network, whereas as LCCs rely more on a point-to-point system. The cost structures and pricing strategies of the two carrier types are also significantly different, and past research have shown that LCCs have a fare decreasing effect in the markets they serve. Almost all of the past research on airline mergers has focused on mergers between legacy carriers. This is due to the fact that mergers between LCCs are quite rare. However, the recent merger between Southwest Airlines and AirTran Airways, both of which are LCCs, gives a unique opportunity to study the effects of LCC mergers.

The year 2010 witnessed what many airline industry experts believed was the first proposal of a consolidation between two major LCCs (Moss, 2010). The merger between Southwest and AirTran was completed four years later in 2014. In this paper, a detailed analysis of the price effects of this merger is presented using a reduced-form framework. A structural model is also presented that is used to quantify the welfare effects of the merger in markets that were most impacted by the merger, i.e., the markets where both carriers were present in the pre-merger period (overlapping markets).

The rest of the paper is organized as follows: section 2.2 presents a background of the merger as it relates to the empirical exercise in this paper, 2.3 presents a synopsis of the related literature, and 2.4 presents a heuristic model that gives an insight about how the impact of an LCC merger is more distinguishable than a legacy carrier merger. Section 2.5 discusses the data used in the analysis, 2.6 presents the reduced form price analysis and 2.7 presents the structural model used to study welfare. Finally, the conclusion in section 2.8 summarizes the key findings, and discusses the public policy consequences of the merger.

2.2 Merger Background

On September 27, 2010 Southwest Airlines announced that they would acquire AirTran Airways for a total cost of $1.5 billion (Dallasnews, 2014). The deal closed on May 2, 2011. The airlines began integrating their frequent flier programs, and moved significant management infrastructure from AirTran's headquarters to Southwest's headquarters in 2011. On March 1, 2012, the airlines received the Single Operating Certificate from the Federal Aviation Administration (FAA), which cleared regulatory impediments to the merger, and allowed the carriers to move forward with the integration process. Southwest had already started services out of AirTran's hub in Atlanta on February 12, 2012 (Southwestonereport, 2011).

Southwest and AirTran began connecting their networks in 2013. This effort started on January 26, 2013 when the airlines first offered a number of connected itineraries in five markets. On March 18, 2013, the airlines announced that they had successfully completed connecting their networks (Southwest press release, 2013). The merger was finally announced as complete in 2014, and AirTran's last flight departed Atlanta for Tampa on December 28, 2014.

Ever since the merger announcement, Southwest has claimed that the integration with AirTran will allow them to efficiently expand their network, and bring about lower fares to passengers. Most notably, the merger has allowed Southwest to enter Atlanta, which was a hub for AirTran, and did not have any Southwest presence before the merger. The merger has also facilitated Southwest's entry into international markets that connect cities in Central and northern parts of South America to the continental US (Forbes, 2014).

It must be noted that three other airline mergers occurred during a similar time frame as the Southwest-AirTran merger, which further complicates the empirical challenge of disentangling the impact of this particular merger from other market phenomena. These other mergers were Delta-Northwest (2008-10), United-Continental (2010-12), and American-US (2013-15). For the purpose of this paper, the pre-merger period will be the period before Southwest and AirTran obtained the Single Operating Certificate from the FAA, and post-merger period is the time after it. Following receipt of the Single Operating Certificate, the merger effort gained momentum as the airlines began consolidating their networks. Therefore, it is reasonable to expect price and welfare impacts of the merger to follow after this event. More discussion on time period selection is presented in section 2.5.

2.3 Related Literature

The unique nature of competitive pressure exerted by LCCs has been widely studied in the airline literature. Morrison (2001) estimates consumer savings resulting from Southwest Airlines directly operating a route (airport pair), a route adjacent to the airport pair, or simply exerting potential competition (by operating routes from both airports, but not operating the route connecting the airports). Morrison's study concludes that when Southwest operates a route, fares of airlines operating not just the same route, but also adjacent and connected via sharing common airports, fall. A more recent work that extensively discusses the heterogeneity between the fare impact of LCCs and legacy carriers in Kwoka et. al. (2016). In this paper, the authors empirically show that LCCs have a much larger impact on the overall market by affecting fares of both other LCCs and legacy carriers, whereas legacy carriers' operations only affect the fares of other legacy carriers.

In terms of the literature on retrospective airline merger analysis, my paper uses similar methodology in the reduced form fare analysis as in Kwoka and Shumilkina (2010). In this paper, the authors use fare regressions to examine the market impact of the merger between USAir and Piedmont airlines. The focus of the paper is on market power attained by USAir upon eliminating Piedmont, a potential competitor to many of USAir's markets. The price regressions in my paper also bear resemblance to Morrison (1996), in which the author analyzes the long run effect of the Northwest-Republic, TWA-Ozark Airlines and USAir-Piedmont mergers.

Other notable papers in the airline literature include Borenstein (1990), which shows using the analysis of the Northwest-Republic and TWA-Ozark mergers that the gain in market power achieved by merging carriers manifests not only in the form of increased prices where the carriers overlapped, but also other routes originating from a common hub. Kim and Singal (1993) also find that prices increase in routes served by merging carriers, implying that market power effects could be offsetting the efficiency effects. Other papers that demonstrate that airline mergers lead to price increases include Werden et. al. (1991) and Peters (2006).

On the other hand, literature that shows that mergers lead to price decreases due to efficiency effects cannot be ignored. A notable work is Ashenfelter et. al. (2015), in which the authors show that efficiencies arising from the merger between Miller and Coors lead to a substantial offsetting of price increase from market power effects. In the context of the airline industry, Carlton et. al. (2016) show using the analysis of Delta-Northwest, United-Continental and American-US mergers that consolidation lead to efficiency effects, which decreased fares.

Some papers in the airline literature employ a structural approach to investigate industry level changes. A notable work is Berry and Jia (2010), in which the authors use a random coefficients framework to compare the state of the domestic US airline industry between 1999 and 2006. The research finds that during 2006, air-travel was more price sensitive, with passengers showing stronger preference for non-stop flights. The paper also finds that one major reason behind the fall in legacy carriers' profits has been the expansion of LCCs. The structural approach in my paper employs a nested-logit demand similar to framework in Cardell (1991). The nested logit model is used to quantify the change in surplus arising in overlapping markets due to the merger.

Although the airline merger literature is rich, this study aims to fill a gap in the analysis of LCC mergers. By using both a reduced form as well as structural approaches, this paper also contributes to the literature on retrospective merger analysis by using a variety of empirical methods.

2.4 Framework

A simple model is used to distinguish the price effects resulting from an LCC merger. Assume that there exist two types of firms: A: legacy carrier and B: LCC. Both types of firms face a linear demand. Firms within each category are symmetrical, and compete in quantities. Let n_k denote the number of k type firms, where k = A or B.

Here, $P_A = k_A$ - Q_A - eQ_B and $P_B = k_B$ - Q_B - fQ_A. "e" is the sensitivity of a legacy carrier's fares to the output of an LCC in the market. Similarly, "f" is the sensitivity of an LCC's fare to the output of a legacy carrier in the market. Following Kwoka et. al. (2016), we will assume e >0 and f ≈ 0. I.e., a legacy carrier's fare is sensitive to the output of an LCC, but an LCC's fare is not sensitive to the presence of a legacy carrier. Intuitively, since the LCCs have a fare suppressing effect on all carriers due to their high level of competitiveness, their presence can affect a legacy's fares but the reverse is not true.

Solving the first order conditions for B (LCC type) reveals the following:

An LCC's fare, i.e., $P_B = c_B + \frac{k_B - c_B}{1+n_B}$, where c_B is the marginal cost of the LCC.

A legacy's fare, i.e., $P_A = c_A + \frac{k_A - c_A}{1+n_A}$ - $\frac{e(k_B - c_B)}{1+n_A}$ (1 - $\frac{1}{1+n_B}$), where c_A is the marginal cost of the legacy carrier.

It becomes apparent from this exercise that the equilibrium LCC fare is inversely related to the number of LCC carriers, whereas the equilibrium legacy fare is inversely related to the number of legacy carriers as well as the number of LCCs. Hence, a merger of LCCs that reduces n_B will affect both P_A and P_B, and have a stronger overall fare increasing impact on the market than a legacy carrier merger that reduces n_A and increases only P_A.

Furthermore, note that $\frac{\partial P_A}{\partial n_B} = \frac{-e(k_B - c_B)}{(1+n_A)(1+n_B)^2}$, and $\frac{\partial P_B}{\partial n_B} = \frac{-(k_B - c_B)}{(1+n_B)^2}$. The expression for $\frac{\partial P_A}{\partial n_B}$ tells us that if e is high, the sensitivity of the increase in P_A following decrease in n_B is higher. One factor determining the value of e is the market share of type B firms (i.e., LCCs). The impact of the presence of a LCC on a legacy's fares is increasing in the share of LCCs in the market (Kwoka et. al (2016)). Thus, we would expect a larger price effect in markets where the merging carriers have a larger market share.

This model is used to emphasize the heterogeneity in the price effects of mergers between different sets of firms. A model with only one firm type would show that price effects resulting from the merger of any two firms would be symmetrical. Clearly, by differentiating the firm type parameters, we see different price effects. As shown in the model, the unilateral price effect resulting from an LCC merger (assuming efficiency gains are absent or are not passed through to the consumer) should be distinguishable by observing an overall fare increase in market fares, and that this effect increases with in the market share of merging carriers. These predictions will be empirically investigated in this paper.

2.5 Data

Air fare data were obtained using the Airline Origin and Destination Survey (DB1B). Published quarterly by the Bureau of Transportation Statistics (BTS) of the Department of Transportation (DOT), the DB1B is a 10% ticket sample of airline tickets from reporting carriers.[1] This useful data source contains detailed information about the ticket such as market fare, origin and destination, number of passengers with the same flight, etc.

A market in this paper refers to a non-directional city pair. Multiple airports within the same city or metropolitan area tend to be very good substitutes. Therefore, market activity in an airport would directly affect its adjacent airports. It is also likely that prices in nearby adjacent airports are highly correlated. Using city pair as a market instead of airport pair makes it easier to analyze

[1] Reporting carriers in the DB1B include all the major airlines operating domestic routes in the US.

and interpret the overall effects of the merger.

The DB1B gives individual ticket level data but for this research, data were coalesced to market-carrier-year quarter level, i.e., an observation refers to an airline-specific market in a given quarter of a year. For example, Boston - Atlanta on Southwest in Q1 2014 is one row in the data table.

The relevant data were 2009 quarter two through 2010 quarter two (ex-ante data), and 2012 quarter two through 2016 quarter two (ex-post data). One empirical challenge with this topic is the fact that the time since the merger announcement and completion was four years, and several changes happened in the airline industry during these intermediary years. To shorten the intermediary time span, I assume the post merger period to be the time after Southwest and AirTran received a single operating certificate (2012 quarter one).[2] Following the receipt of the SOC, the two airlines rapidly started connecting their networks and extensively started marketing joint itineraries. Therefore, it is reasonable to expect unilateral price and welfare effects following the receipt of the SOC.

All observations with market coupons[3] greater than three were dropped as these tend to be open jaw tickets. Bulk fares were dropped as well. All tickets with market fares less than $30 and greater than $5000 were also dropped. The abnormal fares could be the result of coding errors, or frequent flier miles. City pairs with less than 30 sample passengers in an entire quarter were dropped for that quarter.[4] To control for airline network evolution, only markets that are present in both the pre (2009 q2 - 2010 q2) and post (2012 q2 - 2016 q2) merger periods are considered.

The T-100 DS (Domestic Segment) database was used to obtain information on non-stop flight segments. Published monthly, this data table contains flight-specific information as reported by participating carriers. It provides flight-level data such as the origin and destination, routing of the flight, passengers enplaned, frequency, etc.

City demographics data were obtained from the Census Bureau.

[2]Robustness checks performed using alternate time periods of 2009 q2 - 2010 q2 as ex-ante, and 2015 q1 - 2016 q2 as ex-post have been reported in the appendix. As discussed earlier, these results are doubtful due to the long intermediary time gap.

[3]A coupon in the DB1B represents a boarding pass.

[4]Follows Kwoka (2010).

2.6 Price analysis using reduced-form regressions

The reduced-form price regressions use difference-in-differences approach to investigate the causality of change in air fares resulting from the merger. The control group is identified as the set of markets where neither Southwest nor AirTran is present in any of the two market end-points in both the pre as well as post merger periods. Since the merger will primarily affect markets where the merging carriers operate, it is important to define the control group as the markets where neither of these carriers are present. To clarify the construction of the control group, consider the following example. One constituent of the control group is the Alexandria, LA - Greensboro, NC market. Neither Southwest not AirTran operated this market, nor did they operate any other cities out of these end-points in both the pre and post merger periods. It is likely that this market is not directly affected by the Southwest-AirTran merger.[5]

The general specification for the OLS regression is as follows:

$$ln(Fare_{ikt}) = \alpha_0 + \alpha_1 * Treatment_{ij} + \alpha_2 * Post_t + \alpha_3 * Post_t * Treatment_{ij} + \alpha * X_{ikt} + \epsilon_{ikt}$$

Here, "j" is the treatment dummy (equal to 1 if "i" is in treatment group "j"), "i" is the market (non-directional city pair), "k" is the carrier and "t" represents the year-quarter. The dependent variable, $\ln(Fare_{ikt})$, is the logarithm of the average airline specific market fare. The definition of the treatment variable differs according to the type of markets studied, and are described in the later sub-sections. The "post" variable takes value "1" if the data are in the 2012-16 (post merger) period. The interaction of the post and treatment dummies is the primary independent variable of interest. An advantage of the difference-in-differences model is that influences on the dependent variable that are common to both the treatment and control group drop out while running the regression. This is useful because variables such as inflation (which would affect all routes) do not need to be explicitly controlled for. Nevertheless, a number of control variables are included in the regression since they may not be constant across the treatment and control groups over time. These are denoted as X_{ikt} and described as follows:

[5] A test of parallel trends is presented in the appendix to show the comparability between treatment and control groups.

1. Distance: The influence of distance between the end-point airports is accounted for by including the logarithm of average market miles flown between end-point airports as a covariate. Since some flights with stopovers may have been converted to non-stop flights in different time periods, distance between the end-point airports for the cross-section of markets may also be slightly different over time.

2. Population: To control for population, logarithm of the population product at the end-point metropolitan areas is included as a regressor.[6]

3. HHI of non-merging carriers: This variable is included to control for other influences to the market structure that might mask price effects of the merger. Note that market structure in a price regression is an endogenous variable. It is apparent from the IO literature that researchers have struggled to define instruments for HHI. An attempt is made in this paper by including the number of aircraft types as an instrument for HHI in some specifications. Number of aircraft types is correlated with the number of airlines in the market since it is usually the case that different airlines have different aircraft models. Southwest maintains a fleet of B737-700, 737-800 and 737 Max 8,[7] whereas Delta has a wider variety of models in its fleet such as B737-800, B737-900ER, B757-200, B767-300 ER, A319-100, A320-200, etc.[8] Therefore, a market with only Southwest is likelier to see less variation in aircraft models deployed than one with both Southwest and Delta. Furthermore, there is no clear economic reason to expect a direct correlation between the variety of aircraft types and the market fare. However, aircraft type information is only available for non-stop flights, so this instrument can only be used for markets that are non-stop.

4. Number of potential competitors: Previous airline literature indicates strong influence of potential competitors (Goolsbee and Syverson (2008), Morrison (2001)) in setting fares in a market. Sum of the number of potential competitors at end-point cities is included as a covariate. A potential competitor is defined as an airline that operates flights to other

[6]This follows the logic of gravity models used in urban geography literature.
[7]https://www.swamedia.com/pages/corporate-fact-sheet
[8]https://www.delta.com

markets out of a city, but not the city-pair whose fare is the dependent variable of interest. Southwest and AirTran are excluded from the potential competitor count since their effect will be captured by the causal variable of interest.

5. Dummy for the UN-CO merger: The United-Continental merger was initiated and concluded (2010-12) during the time frame that separates the pre and post periods of the Southwest-AirTran merger. A dummy for any market where United or Continental operate is included as a regressor to control for the influence of this merger.

6. Quarter dummies: These are included to control for seasonal variation in air fares.

7. Other dummies to indicate whether

   - the airport is slot controlled.[9]
   - the airport is a hub for any of the carriers.[10]
   - the city where the airport is located is a tourist destination.[11]

These dummies are included to check the validity of the regression setup rather than to control for their influence on the coefficients of interest. Some specifications also include market, airline and year fixed effects.

2.6.1 Differences across market share of merging carriers

The theoretical framework implied that the impact of the merger would be stronger in markets where the merging carriers have larger market share. To investigate empirically, joint ex-ante market shares of the merging carriers were constructed. Treatment groups were constructed by dividing overlapping markets into four groups according to the quartiles of joint Southwest-AirTran share. A paper that utilizes a similar approach is Hosken et. al. (2012). Table 2.1 summarizes the dataset.

[9]Some airports in the United States are slot controlled, i.e., restrictions are imposed on airlines operating at that airport from making more than a given number of take-offs and landings.

[10]The definition of an airline's hub follows the information given on their websites.

[11]A tourist destination was defined using http://travel.usnews.com/rankings/best-usa-vacations/

Table 2.1: Control group and overlapping markets by market share

Group	Market count	Observations
Control	890	35,007
Upto 25th percentile SW-AT share (1.3-12.8%)	12	29,218
25th - 50th percentile SW-AT share (13.0-24.4%)	117	28,984
50th - 75th percentile SW-AT share (24.5-38.4%)	115	29,182
75th - 100th percentile SW-AT share (38.5-96.1%)	126	28,809

Regression tables 2.6 through 2.10 present the results for this section. Specification (1) does not include fixed effects, (2) includes market, airline and year fixed effects, and (3) includes these fixed effects and uses number of aircraft models as an instrument for non-merging carriers' HHI. In general, specification (1) looks problematic since it does not yield the expected signs for the coefficients for HHI and potential competitors. Including fixed effects appears to fix this issue in majority of the cases. One possible conjecture for the sign reversal on the coefficient could be that the tendency of some markets to have higher fares and higher number of participating firms could be due to the idiosyncratic nature of the market. For instance, a long-haul market could have higher fares and higher number of competitors since it is likelier that these routes are more profitable as they face less competition from other modes of transportation. "Long-haul-ness" could be a market specific fixed effect, which is controlled for in (2) and (3). Including year fixed effects helps control for the impact of the changes in jet fuel prices. Airline fixed effects help control for firm-specific cost shocks.

We will use average values of the coefficient of interest (post*treat) across the three specifications to examine results. For markets where the merging carriers had upto 12.8% share, the causal effect of the merger on fares appears to be approximately four percent, followed by eight percent for 13-24.4% share, ten percent for 24.5-38.4% share, and ten percent for 38.5-96.1% share. The average effect on all overlapping markets is shown in Table 2.10, and appears to be approximately seven percent. These findings go along with the inference of the model presented earlier, which claims that the fare effect of the merger must be stronger in markets where the merging carriers have a larger market share.

2.6.2 Impact of potential competition

The merger allowed Southwest Airlines to eliminate AirTran as a potential competitor from several markets. To investigate the price effect resulting from the loss of potential competition, data were categorized according to whether Southwest operated in a given market before the merger, and AirTran existed in either or both end-points, or whether AirTran operated in a market, and Southwest existed in either or both end-points. The same control group as in the preceding section was used. Table 2.2 summarizes the dataset.

Table 2.2: Control group and markets impacted by potential competition

Group	Market Count	Observations
Control	890	35,007
SW present, AT potential entrant at both ends	106	24,972
AT present, SW potential entrant at both ends	3	380
SW present, AT potential entrant at one end	840	134,409
AT present, SW potential entrant at one end	477	82,277

One observation that becomes apparent from Table 2.2 is that AirTran was a potential competitor for Southwest in many more markets than Southwest was for AirTran. Table 2.11 presents the regression results for this section. As before, specification (1) does not include fixed effects, (2) includes market, airline and year fixed effects, and (3) includes these fixed effects and uses number of aircraft models as an instrument for non-merging carriers' HHI. The value of the coefficient of the treatment post interaction variable clearly shows that the elimination of potential competition led to a fare increase. The value of the coefficient on Post*(AT present, SW potential entrant at both ends) is unusually high for all specifications. Note that this treatment only comprises of three markets as evidenced in Table 2.2. Therefore, the extremely high coefficient could be the result of market specific idiosyncrasy, and hence we will not put a lot of weight on the result of this particular regressor. Other treatment post interaction coefficients show that modest price increases occurred in markets impacted by the loss in potential competition due to the merger. Of particular interest are the first and third coefficients where we see the fare effect of eliminating AirTran as a potential competitor. Past research that has used a similar approach to characterize potential competition (E.g. Kwoka and Shumilkina (2010)), have found that markets where a merger eliminates a potential entrant at only one end-point do not see a significant price increase. However, the significant positive values of the coefficient on Post*(SW present, AT potential entrant at one end) shows underscores the specialty of this LCC merger. We find that AirTran's elimination has decreased the fare discipline imposed on other carriers arising from potential competition, and has thus led to fare increase.

2.6.3 Impact on markets where AirTran ceased service following the merger

Although previous literature has extensively analyzed the impact of entry of a low cost carrier, little work has been done to quantify the impact of an LCC's exit from the market. Following the network integration of Southwest and AirTran, several routes where AirTran used to operate were eliminated by Southwest. In the dataset, there were 223 AirTran markets where Southwest (which used to be AirTran) ceased service following the merger. In other words, these are markets with AirTran present in the pre-merger period, but no Southwest present in the post-merger period. There were a total of 25,494 observations for these 223 markets in the dataset. Table 2.12 shows the fare regression results for these markets. The same control group as in earlier sections was used. As before, specification (1) does not include fixed effects, (2) includes market, airline and year fixed effects, and (3) includes these fixed effects and uses number of aircraft models as an instrument for non-merging carriers' HHI. It is apparent from the results that markets where Southwest eliminated AirTran's service see a substantial increase in fares. Averaging the values of the post-treat interaction variable, we observe approximately 14% price increase in these markets. The fare increase is directly attributable to the exit of a competitive player in the market. Out of all the markets studied in the price analysis, these markets appear to have the greatest magnitude of fare increase.

2.7 Welfare analysis using a structural approach

The price regressions in Section Six provide evidence that the Southwest-AirTran merger led to substantial fare increases, which eventually could translate into significant welfare effects. If consumers' preferences and willingness to pay for air travel in the affected markets do not change over the time frame studied, then the rise in prices is likely to have decreased consumer surplus. However, consumer preference could be altered if the product quality changes over time, perhaps due to the merger itself. When the merged firm provides better service quality (better on-time performance, better frequency, etc.), consumers are likely to increase their willingness to pay for the new product. The reverse is true if the new entity faces diseconomies in product quality resulting from the merger. Therefore, price movements in isolation cannot be used to gauge consumer welfare changes.

The increase in prices also signals that firm profits may have been altered by the merger. If costs remain constant, price increases can be directly translated as profit increases. In the absence of market specific cost data, it is difficult to estimate the change in profits. For the merging carriers themselves, change in profits could also come from efficiency gains.

The motivation of using a structural framework in this paper is two-fold. First, it is used to quantify the change in welfare. The nested logit model used in this paper, similar to Berry (1994), yields estimates of consumer surplus movements, and the calculated elasticities can be used to recover marginal cost estimates to gauge the change in profits as well. The second motivation is to perform a merger simulation exercise to measure how the standard model would have predicted the market to behave in the post merger period. It is likely that economists investigating the potential effects of this merger before it happened used similar models to gauge market behavior. Since this is a retrospective study, it benefits from the access of post merger market data. The comparison of actual behavior to simulation can be used to judge the accuracies of the standard methods used by a competition economist.

It must be noted however that the air-fare data being used in this paper come from the quarterly DB1B database. The discrete choice structural framework is based on the assumption that at any given time, the consumer chooses one product out of the elements available in her choice set. The exact choice set of the consumer is impossible to infer using a quarterly dataset that pools together tickets bought by passengers in different days of the same year quarter. However, since the dataset is comprised of ten percent of all tickets sold in a year-quarter, it is quite rich in construction.

Therefore, it is possible that in approximation, the DB1B captures most of the elements of the consumer's choice set during the specific time of the ticket purchase.

2.7.1 Demand Model

Consider inter-city passenger travel between two cities: A and B. Consumers decide to travel between A and B by air, or they choose the outside option that includes the choice of not traveling, or traveling by other means of transportation. Following a decision to travel by air, I will assume that the passenger first chooses the number of stop-overs, and then chooses the specific airline. Let 'g' represent a group, where g = 0, 1, 2, 3. The group g = 0 represents the outside option of not traveling or using other means of transport. g = 1 represents travel using a single market coupon, which represents non-stop flights. Similarly, g = 2 represents two coupons hence one stop-over, and g = 3 represents two stop-overs. Within each non-zero g, the passenger chooses the product of an airline 'j'. Figure 2.1 illustrates the construction of nests.

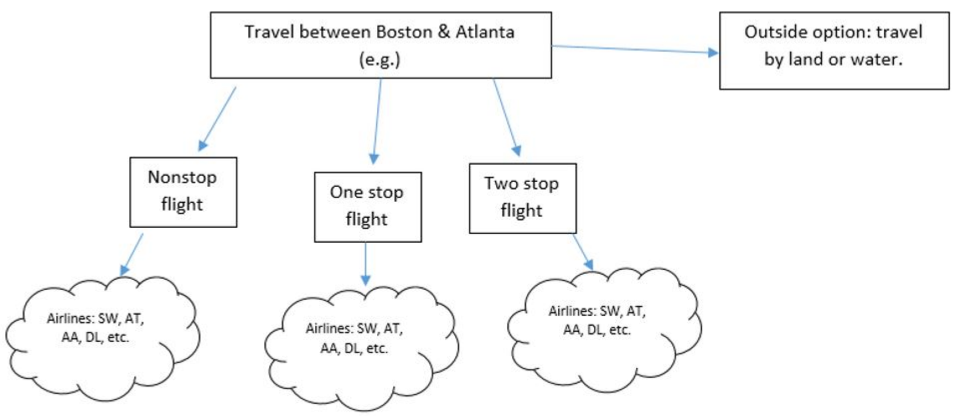


Figure 2.1: THE NESTED LOGIT DEMAND MODEL

The indirect utility of a consumer i for selecting a product j in J_g is given by:[12]

$$u_{ij} = X_j\beta - \alpha p_j + \xi_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij} = \delta_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij}$$

[12]Note that the time subscript is suppressed

X_j includes distance traveled by the aircraft, number of market coupons, geometric mean of the population at end-point MSAs, quarter, year, tourist, hub, slot and airline dummies. δ_j is the mean utility from consuming product j. ξ_j are market characteristics of product j that are unobserved to the econometrician. ϵ_{ij} are unobserved idiosyncratic individual preferences for j, and ζ_{ig} are unobserved idiosyncratic individual preferences for all airline products within group g. We will assume that the terms ϵ_{ij} and ζ_{ig} follow a type 1 extreme value distribution, and consequently so will ζ_{ig} + (1 - σ)ϵ_{ij}. σ gives us the degree of within group correlation.

The share of j within g is given as $s_{j|g}(\delta, \sigma) = \frac{exp(\delta_j/(1-\sigma))}{D_g}$,
where $D_g = \sum_{j \in J_g} exp(\delta_j/(1-\sigma))$. The probability of choosing a product of type g is $s_g(\delta, \sigma) = \frac{D_g^{1-\sigma}}{\sum_{g'} D_{g'}^{1-\sigma}}$. It follows that $s_j(\delta, \sigma) = \frac{exp(\delta_j/(1-\sigma))}{D_g^{\sigma}[\sum_{g'} D_{g'}^{1-\sigma}]}$. Inverting δ_j using observed market shares yields the following equation that can be estimated using linear regression:

$$ln(s_j) - ln(s_0) = \delta_j + \sigma ln(s_{j|g})$$

Here, the mean utility of the outside good is set to zero. I assume that for any given city-pair in a year quarter, one percent of the population at an end-point considers traveling to other end-point and back. Hence, the market size for a city pair is one percent of the sum of population at the two end-point cities in a given year quarter. s_0 is the share of the outside option, and $s_{j|g}$ represents the share of product j within group g.

2.7.2 Estimation

The linear regression estimation equation above includes two endogenous right hand side variables: price of the airline product (p_j), and share of the product within a group (ln($s_{j|g}$)). Some papers utilizing a similar method in the past literature have used cost shifters as an instrument for price. Here, quadratic polynomial in distance is used as an instrument for price. The number of products within a group is used as an instrument for the within group share. Table 2.13 summarizes the regression output of the nested logit model. The estimation was run for only overlapping markets as these markets were directly affected by the merger. As shown by Berry and Jia (2010), it is likely that airline demand is different in different time periods. Hence, the estimation is conducted separately for the pre and post merger time periods.

Specification (1) is for the pre-merger period, and (2) is for post-merger. In both specifications, the coefficient of the within group share variable is positive and significant, and this rejects a simpler model such as multinomial logit to address the demand side estimation. The coefficient on price is negative and significant, showing the disutility consumers associate with an increase in price.

The table also reports the mean own price product elasticities for the two time periods. The absolute values of the elasticity magnitudes are large suggesting that airline products are highly elastic. A reason for the high estimates could be due to the nature of the quarterly DB1B data - a consumer's choice set in the quarterly data is the entire number of airline products available during the year quarter. It is not necessarily the case that all airlines that operated flights in the year-quarter were offering flights during the specific date and time the passenger wanted to travel. Therefore in reality, the elasticity estimates are likely to be much smaller. The large values are observed due to the appearance of a large number of choices to the consumer. While the magnitude of the elasticity estimates could be misleading, the difference in their magnitudes shows that elasticity fell post-merger, reflecting the loss in the number of flight options following the elimination of AirTran.

2.7.3 Consumer surplus

Although prices have risen, the product offerings of airlines have changed after the merger. AirTran's itineraries have now been transferred to Southwest, and some consumers may have found the switch to be quality enhancing. For instance, there might have existed a travel itinerary where a consumer traveling from A to B had to connect via C, and had to take AirTran from A to C, and Southwest from C to B. After the merger, such an itinerary could have been converted into a direct Southwest flight. The loss of the inconvenience factor of switching planes at stop-overs may increase the consumer's willingness to pay, and offset the price increase to enhance consumer surplus.

The consumer surplus in the nested logit model is given by the following:

$$CS = \frac{M}{\alpha} ln(1 + \sum_{g=1}^{G} [\sum_{j \in J_g} exp[\delta_j/(1-\sigma)]^{1-\sigma})$$

Here M is the market size, and α and σ are parameters obtained from the nested logit model. Since the markets (origin - destination city pairs) studied are numerous across several year-quarters, and since the number of year-quarters in the pre and post merger periods are different, I construct average consumer surplus estimates across markets and year-quarters to compare the change in consumer surplus resulting from the merger. Table 2.3 presents the estimates:

Table 2.3: Consumer surplus estimates in overlapping markets

Time period	Per consumer CS	Market size	Total CS
Pre Merger	87	110,818	9,672,195
Post Merger	79	114,773	9,022,306
Net change	-9	3,955	-649,990

Consumer surplus per average consumer has dropped approximately by \$9 due to the merger. The increase in the average market size is attributable to population growth. Total consumer surplus in the average market-year-quarter appears to have clearly decreased by approximately \$650 thousand.

2.7.4 Supply side and profit

Assuming airline competition can be explained by the Bertrand model, the relation between price, marginal costs and own price elasticities are given by the Lerner index:

$$\frac{p_j - c_j}{p_j} = -\frac{1}{e_j}$$

I make the simplifying assumption that only single product firms exist in any given market. The price elasticity, $e_j = (\frac{\alpha p_j}{1-\sigma})$*[1 - σ $s_{j|g}$ - (1 - σ)s_j]. Using the elasticity estimates and the observed airline product fares, marginal cost estimates can be recovered. Table 2.4 presents average (across markets and year-quarters) values for p_j and c_j for the two time merger specific time periods. Note that the recovered marginal costs follow the direction of price changes due to their relationship from the Lerner index. If large efficiency gains are present from the merger, then it is likely that the actual marginal costs will be lower than the recovered estimates.

Table 2.4: Average prices and recovered marginal costs in overlapping markets

Time period	p_j	c_j
Pre Merger	207	193
Post Merger	264	240

The marginal cost estimates of \$190 and \$240 may appear unintuitive in the context of the airline industry. For a flight, the marginal cost of an additional passenger is extremely low (perhaps even close to zero), since transporting an additional passenger does not really impose additional costs to an already scheduled flight. However, these estimates can be thought of as the marginal cost of one more flight, spread across the passenger total to get marginal cost per passenger.

Profits can now be calculated using the following:

$$\pi_j = (p_j - c_j) * Ms_j$$

Profit estimates from Table 2.5 show that both per product profit margin and the average (across markets and flight types (non-stop, 1-stop or 2-stop)) profit for an airline product in a year-quarter have increased following the merger. As these markets lost a fare disciplining LCC, incumbents were successful in raising profits by increasing prices.

Table 2.5: Profit estimates in overlapping markets

Time period	Per product margin	Average profit
Pre Merger	14	69,860
Post Merger	24	150,563
Net change	10	80,703

2.7.5 Merger simulation with the nested logit model

Next, a merger simulation exercise is performed to gauge how the model would have predicted post-merger prices. First, pre-merger demand is estimated, and using the Lerner Index and the estimated elasticities, marginal costs are recovered. Using the estimated demand side parameters, and the recovered costs, new prices are constructed using the following relationship:

$$p_j = \frac{\sigma - 1}{\alpha(1 - \sigma s_{j|g} - (1-\sigma)s_j)} + c_j$$

The new $s_{j|g}$ and s_j are found by recoding all AirTran observations in the pre-merger period to Southwest. It is also assumed that they set a common price in the post-merger period. The average simulated post-merger price in overlapping markets is estimated to be \$207.1, whereas the pre-merger price in the same markets was \$206.9. The small price difference is significant at the five percent level. Looking only at Southwest's fares, the pre-merger fare average was \$192.2, and the simulated average comes to \$193.1. Here too, the difference is significant at the five percent level.

According to the price regressions in section 2.6, fares in overlapping markets on average rose by six percent due to the merger. Such wide discrepancies between the merger simulation prices and the actual price change suggest that the standard tools of merger analysis may not be adequate to predict the consequence of a merger as unique as the one between Southwest and AirTran. The simulated prices are directly attributable to the small change in s_j and $s_{j|g}$ that results from combining the market shares of the merging carriers. However, real data seems to suggest that airline markets were much more sensitive to the exit of AirTran, leading to a bigger increase in fares due to a substantial lessening of competitive constraint from the loss of a major low cost carrier. An econometrician using a similar model to predict prices could conclude that the merger is not anti-competitive. However, the sharp distinction with reality once again suggests that a merger between two low cost carriers is quite atypical.

These results also seem to suggest that significant cost inefficiencies may have occurred from the merger. While this remains possible, especially in the first few years following the merger (merging carriers may initially face communication and coordination issues, as their reservation systems may not be fully integrated), the true source of cost changes would be from the change in fixed costs. Without route specific cost data, any hypothesis about the movement of costs is only as good as speculation.

2.8 Conclusion

This paper has analyzed in detail the price and welfare effects of the Southwest-AirTran merger. Reduced form price regressions show that fares in overlapping markets have substantially increased following the merger. The price increase is more pronounced in markets where the merging carriers had larger market shares. Price increase was also observed in markets impacted by the reduction of potential competition following the elimination of AirTran after the merger. Fares rose even in markets where AirTran was a potential entrant at only one end-point city. The largest magnitude of price rise was observed in markets where AirTran operated flights in the pre-merger period, but did not have Southwest post-merger. These results underscore the fare constraining behavior AirTran used to exert in markets when it existed as an independent entity. Furthermore, the price rises are also a result of the change in competitive conduct of Southwest Airlines following the merger. AirTran's elimination eased the competitive constraint on Southwest (as well as other carriers), and in turn, Southwest's change of competitive behavior enabled many other carriers to raise fares too. The compounding effect may have led to strong overall effects.

Welfare effects in overlapping markets were analyzed using a structural framework. The estimates obtained show that while consumer surplus decreased, profits of incumbents increased in the post-merger period. Aggregating the movements in consumer and producer surplus show that there was a loss of approximately USD 570 thousand in the average overlapping market in a year-quarter. However, since the model does not accurately capture possible efficiency gains that may have resulted from the merger, it is likely that this figure exaggerates the welfare loss.

Overall, this study highlights the importance of taking into account heterogeneities across firm types while evaluating the possible effects of their consolidation on competition and welfare. A merger of two LCCs may create bigger price effects due to the compounding mechanism described earlier. It is hoped that this retrospective study will be of use to antitrust authorities in deciding the future direction of competition policy in the airline industry.

2.9 References

Berry, Steven, and Panle Jia. "Tracing the Woes: An Empirical Analysis of the Airline Industry." *American Economic Journal: Microeconomics*, vol. 2, no. 3, 2010, pp. 143.

Berry, Steven T. "Estimation of a Model of Entry in the Airline Industry." *Econometrica*, vol. 60, no. 4, 1992, pp. 889–917.

Boguslaski, C., Ito, H., and Lee, D. "Entry Patterns in the Southwest Airlines Route System." *Review of Industrial Organization*, vol. 25, no.3, 2004, pp. 317–350.

Borenstein, Severin. "Airline Mergers, Airport Dominance, and Market Power." *American Economic Review*, vol. 80, no. 2, 1990, pp. 400–04.

Borenstein, Severin. "Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry." *The RAND Journal of Economics*, vol. 20, no. 3, 1989, pp. 344–365.

Borenstein, Severin, and Nancy L. Rose. "Competition and Price Dispersion in the U.S. Airline Industry." *Journal of Political Economy*, vol. 102, no. 4, 1994, pp. 653–683.

Carlton, D. W., Israel, M. A., MacSwain, I. and Orlov, E.,"Are Legacy Airline MErgers Pro- or Anti- Competitive? Evidence from Recent U.S. Airline Mergers" *SSRN*, 2017.

Ciliberto, Federico, and Elie Tamer. "Market Structure and Multiple Equilibria in Airline Markets." *Econometrica*, vol. 77, no. 6, 2009, pp. 1791–1828.

Dana, James D. Jr., and Eugene Orlov."Internet Penetration and Capacity Utilization in the US Airline Industry." *American Economic Journal: Microeconomics*, 2014, 6(4): 106-37.

Goolsbee, Austan, and Chad Syverson. "How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines." *The Quarterly Journal of Economics*, vol. 123, no. 4, 2008, pp. 1611–1633.

Hosken, D., Olson, L.M., and Smith, L.K., "Do Retail Mergers Affect Competition? Evidence from Grocery Retailing," *FTC Working Paper No. 31.*

Kim, E. Han, and Vijay Singal. "Mergers and Market Power: Evidence from the Airline Industry." *The American Economic Review*, vol. 83, no.3, 1993, pp. 549–569.

Kwoka, John, and Evgenia Shumilkina. "The Price Effect of Eliminating Potential Competition: Evidence from an Airline Merger" *The Journal of Industrial Economics*, vol. 58, no. 4, 2010, pp. 767–793.

Kwoka, J., Hearle, K., & Alepin, P. "From the Fringe to the Forefront: Low Cost Carriers and Airline Price Determination." *Review of Industrial Organization,* 2016, *48*, 247-268.

Peters, C. "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry." *Journal of Law and Economics*, vol. 49, no. 2, pp. 627–649.

Seim, Katja. "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." *The RAND Journal of Economics*, vol. 37, no. 3, 2006, pp. 619–640.

Small, Kenneth A., and Harvey Rosen, "Applied Welfare Economics with Discrete Choice Models," *Econometrica* 49, January 1981, pp. 105-30.

Steven A. Morrison and Clifford Winston, "Another Look at Airport Congestion Pricing," *American Economic Review*, December 2007, pp. 1970-77.

Steven A. Morrison and Clifford Winston, "Empirical Implications and Tests of the Contestability Hypothesis," *Journal of Law and Economics*, vol. 30, April 1987, pp. 53-66.

Steven A. Morrison, "Actual, Adjacent, and Potential Competition: Estimating the Full Effect of Southwest Airlines," *Journal of Transport Economics and Policy*, vol. 35, Part 2, May 2001, pp. 239-256.

Steven A. Morrison and Clifford Winston. "*The Economic Effects of Airline Deregulation*," Washington, D.C.: Brookings Institution, 1986. Print.

Werden, G. J., A. S. Joskow, and R. L. Johnson (1991), "The Effects of Mergers on Price and Output: Two Case Studies from the Airline Industry," *Managerial and Decision Economics*, vol. 12, no. 5, pp. 341–352.

2.10 Regression Tables

Table 2.6: Fare regression for overlapping markets where SW-AT have 1.3-12.8% share in the pre-merger period. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*treat(upto 25th pctile share)	0.0552***	0.0295***	0.00994
	(0.00635)	(0.00565)	(0.0113)
Distance	0.318***	0.262***	0.237***
	(0.00255)	(0.00763)	(0.0103)
UN-CO dummy	0.0133	0.0132	0.0305**
	(0.0126)	(0.0112)	(0.0138)
Non-merging HHI	-0.0520***	0.00243	0.195***
	(0.00262)	(0.00251)	(0.0373)
Potential competitors	0.00561***	-0.00146**	-0.00777***
	(0.000466)	(0.000622)	(0.00135)
Population	-0.0143***	-0.208***	-0.644***
	(0.00115)	(0.0372)	(0.0884)
Hub	0.0725***	0	0
	(0.00370)	(.)	(.)
Tourist	-0.0425***	0	0
	(0.00416)	(.)	(.)
Slot	0.0298***	0	0
	(0.00487)	(.)	(.)
Observations	64225	64225	35290
R^2	0.360	0.631	
Adjusted R^2	0.360	0.622	
F	1891.9	500.9	
rmse	0.342	0.287	
First stage F			26.69

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[12] Some coefficients have been omitted due to space constraints. The appendix presents regression results using alternative specifications. Standard errors are in parentheses.

Table 2.7: Fare regression for overlapping markets where SW-AT have 13.0-24.4% share in the pre-merger period. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*treat(25th-50th pctile share)	0.0946***	0.0765***	0.0654***
	(0.00622)	(0.00545)	(0.0107)
Distance	0.333***	0.320***	0.334***
	(0.00253)	(0.00760)	(0.0107)
UN-CO dummy	-0.0390***	-0.0374***	-0.0243*
	(0.0117)	(0.0103)	(0.0136)
Non-merging HHI	-0.0495***	0.00552**	0.235***
	(0.00293)	(0.00280)	(0.0649)
Potential competitiors	0.00622***	0.0000628	-0.00666***
	(0.000458)	(0.000591)	(0.00198)
Population	-0.0120***	-0.122***	-0.410***
	(0.00118)	(0.0359)	(0.0877)
Hub	0.0300***	0	0
	(0.00309)	(.)	(.)
Tourist	-0.0491***	0	0
	(0.00392)	(.)	(.)
Slot	-0.0239***	0	0
	(0.00616)	(.)	(.)
Observations	63991	63991	32327
R^2	0.397	0.545	
Adjusted R^2	0.397	0.538	
F	2221.3	674.8	
rmse	0.332	0.276	
First stage F			26.79

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.8: Fare regression for overlapping markets where SW-AT have 24.5-38.4% share in the pre-merger period. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*treat(50th-75th pctile share)	0.0984***	0.0929***	0.0888***
	(0.00633)	(0.00556)	(0.0104)
Distance	0.336***	0.356***	0.379***
	(0.00247)	(0.00725)	(0.0106)
UN-CO dummy	-0.0292**	-0.0299***	-0.0171
	(0.0128)	(0.0106)	(0.0136)
Non-merging HHI	-0.0641***	0.0130***	0.277**
	(0.00331)	(0.00356)	(0.116)
Potential competitiors	0.00558***	0.00198***	-0.00269
	(0.000459)	(0.000595)	(0.00230)
Population	-0.00797***	-0.0615*	0.149*
	(0.00116)	(0.0356)	(0.0788)
Hub	0.0510***	0	0
	(0.00308)	(.)	(.)
Tourist	-0.0544***	0	0
	(0.00386)	(.)	(.)
Slot	0.0251***	0	0
	(0.00690)	(.)	(.)
Observations	64189	64189	31797
R^2	0.402	0.577	
Adjusted R^2	0.402	0.571	
F	2342.1	725.5	
rmse	0.337	0.281	
First stage F			27.72

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.9: Fare regression for overlapping markets where SW-AT have 38.5-96.1% share in the pre-merger period. Dependent variable is logarithm of average market-airline fare

	(1) ln_fare	(2) ln_fare	(3) ln_fare
Post*treat(75th-100th pctile share)	0.101***	0.0865***	0.101***
	(0.00634)	(0.00560)	(0.00962)
Distance	0.322***	0.322***	0.352***
	(0.00250)	(0.00664)	(0.00896)
UN-CO dummy	-0.0429***	-0.0404***	-0.0327***
	(0.0120)	(0.0102)	(0.0122)
Non-merging HHI	-0.0398***	0.0235***	-0.114
	(0.00316)	(0.00298)	(0.0747)
Potential competitiors	0.00557***	0.00117**	0.00414**
	(0.000465)	(0.000583)	(0.00163)
Population	-0.00754***	-0.172***	-0.353***
	(0.00122)	(0.0349)	(0.0818)
Hub	0.0623***	0	0
	(0.00308)	(.)	(.)
Tourist	-0.0326***	0	0
	(0.00356)	(.)	(.)
Slot	0.125***	0	0
	(0.00624)	(.)	(.)
Observations	63816	63816	32509
R^2	0.455	0.620	
Adjusted R^2	0.455	0.614	
F	3025.5	733.2	
rmse	0.334	0.274	
First stage F			27.00

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.10: Fare regression for all overlapping markets. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*treat(all overlapping)	0.0608***	0.0760***	0.0637***
	(0.00530)	(0.00418)	(0.00938)
Distance	0.284***	0.308***	0.297***
	(0.00185)	(0.00435)	(0.00552)
UN-CO dummy	-0.0263***	-0.0173***	-0.00280
	(0.00585)	(0.00561)	(0.00688)
Non-merging HHI	-0.0338***	0.00402**	0.226***
	(0.00159)	(0.00165)	(0.0421)
Potential competitiors	-0.000255	0.00174***	-0.00403***
	(0.000309)	(0.000384)	(0.00113)
Population	0.00198**	-0.147***	-0.178***
	(0.000874)	(0.0259)	(0.0380)
Hub	0.0308***	0	0
	(0.00206)	(.)	(.)
Tourist	-0.0284***	0	0
	(0.00228)	(.)	(.)
Slot	0.0425***	0	0
	(0.00298)	(.)	(.)
Observations	151200	151200	107884
R^2	0.501	0.501	
Adjusted R^2	0.497	0.497	
F	3831.4	1905.4	
rmse	0.317	0.277	
First stage F			31.47

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.11: Fare regressions to investigate the impact of potential competition. Dependent variable is logarithm of average market-airline fare

	(1) ln_fare	(2) ln_fare	(3) ln_fare
Post*(SW present, AT potential entrant at both ends)	0.0878***	0.0662***	0.0678***
	(0.00699)	(0.00516)	(0.0132)
Post*(AT present, SW potential entrant at both ends)	0.286***	0.295***	0.912***
	(0.0522)	(0.0303)	(0.158)
Post*(SW present, AT potential entrant at one end)	0.0743***	0.0711***	0.0571***
	(0.00516)	(0.00369)	(0.0124)
Post*(AT present, SW potential entrant at one end)	0.0616***	0.0544***	-0.0317
	(0.00546)	(0.00391)	(0.0194)
Distance	0.267***	0.247***	0.225***
	(0.00118)	(0.00285)	(0.00601)
UN-CO dummy	-0.0652***	-0.0604***	-0.0500***
	(0.00484)	(0.00398)	(0.00835)
Non-merging HHI	-0.00381***	-0.00167*	0.437***
	(0.000944)	(0.000956)	(0.114)
Potential competitiors	0.00757***	0.00275***	-0.00878***
	(0.000200)	(0.000264)	(0.00305)
Population	-0.0277***	-0.107***	-0.0257
	(0.000662)	(0.0146)	(0.0305)
Hub	0.0256***	0	0
	(0.00139)	(.)	(.)
Tourist	-0.0532***	0	0
	(0.00133)	(.)	(.)
Slot	0.0919***	0	0
	(0.00253)	(.)	(.)
Observations	276874	276874	155056
R^2	0.334	0.497	
Adjusted R^2	0.334	0.493	
F	4835.9	3060.4	
rmse	0.301	0.263	
First stage F			42.54

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.12: Fare regressions to investigate the impact on AirTran markets where Southwest did not resume service after the merger. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*(AT in pre, no SW in post)	0.146***	0.136***	0.117***
	(0.00662)	(0.00551)	(0.0118)
Distance	0.323***	0.259***	0.271***
	(0.00273)	(0.00838)	(0.0137)
UN-CO dummy	-0.0893***	-0.0784***	-0.113***
	(0.0145)	(0.0115)	(0.0213)
Non-merging HHI	-0.0878***	0.0192***	0.194***
	(0.00408)	(0.00401)	(0.0611)
Potential competitiors	0.00847***	0.00213***	0.00136
	(0.000447)	(0.000608)	(0.00139)
Population	-0.0144***	-0.103***	-0.245***
	(0.00118)	(0.0332)	(0.0731)
Hub	0.0734***	0	0
	(0.00327)	(.)	(.)
Tourist	-0.0448***	0	0
	(0.00376)	(.)	(.)
Slot	0.0802***	0	0
	(0.00989)	(.)	(.)
Observations	60494	60494	19028
R^2	0.362	0.577	
Adjusted R^2	0.362	0.569	
F	1688.0	692.4	
rmse	0.348	0.287	
First stage F			30.53

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.13: Nested logit demand estimation for overlapping markets

	(1) Pre-merger	(2) Post-merger
Price	-0.0175***	-0.0173***
	(0.000231)	(0.000125)
σ	0.811***	0.703***
	(0.00807)	(0.00604)
Distance	0.000441***	0.000467***
	(0.0000220)	(0.0000161)
Slot	-0.620***	-0.579***
	(0.0434)	(0.0304)
Tourist	0.0540**	0.113***
	(0.0259)	(0.0192)
Hub	0.526***	0.460***
	(0.0253)	(0.0185)
Non-stop	1.398***	1.209***
	(0.0314)	(0.0212)
One-stop	1.726***	1.269***
	(0.0266)	(0.0176)
Population	3.22e-08***	3.37e-08***
	(5.49e-09)	(3.95e-09)
Mean elasticity	-16.84	-13.39
Observations	37106	84849
R^2	0.392	0.265
Adjusted R^2	0.392	0.264
Quarter fixed effects	Yes	Yes
Airline fixed effects	Yes	Yes
Coupon fixed effects	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.11 Appendix

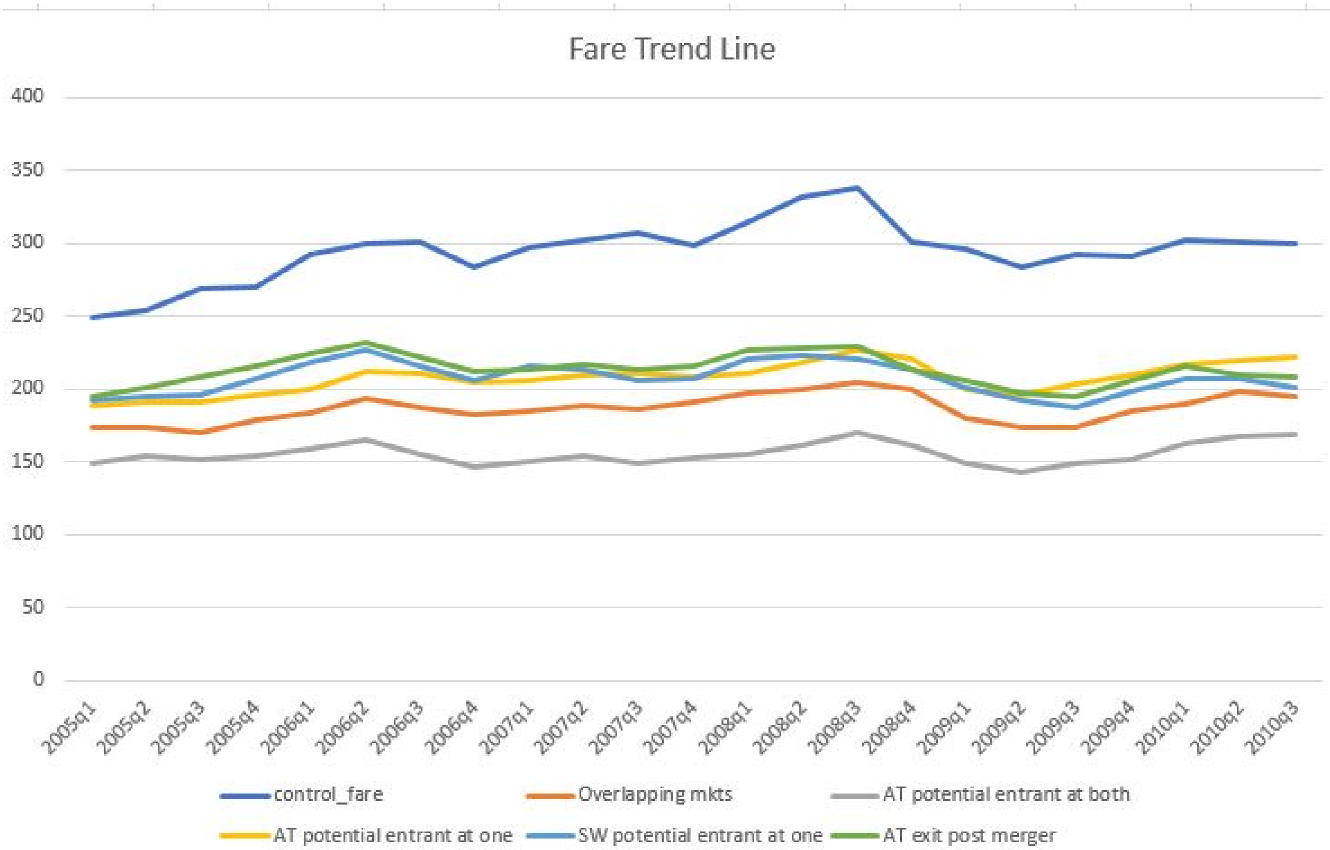


Figure 2.2: Average Fare Trend Line

Table 2.14: Fare regression for all overlapping markets with 2009 q2-2010 q2 as pre and 2015 q1-2016q2 as post. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*treat(all overlapping)	0.0409***	0.0669***	0.0589***
	(0.00635)	(0.00553)	(0.0128)
Distance	0.287***	0.367***	0.361***
	(0.00257)	(0.00630)	(0.00795)
UN-CO dummy	-0.0263***	-0.0171***	-0.00444
	(0.00584)	(0.00579)	(0.00725)
Non-merging HHI	-0.0317***	-0.00107	0.251***
	(0.00206)	(0.00223)	(0.0316)
Potential competitiors	0.0000587	0.000577	-0.00812***
	(0.000428)	(0.000532)	(0.00120)
Population	0.000329	-0.159***	-0.275***
	(0.00123)	(0.0296)	(0.0440)
Hub	0.0342***	0	0
	(0.00290)	(.)	(.)
Tourist	-0.0245***	0	0
	(0.00322)	(.)	(.)
Slot	0.0518***	0	0
	(0.00419)	(.)	(.)
constant	3.817***	7.439***	9.152***
	(0.0371)	(0.860)	(1.305)
Observations	79970	79970	57417
R^2	0.352	0.509	
Adjusted R^2	0.352	0.501	
F	2662.3	1348.9	
rmse	0.325	0.285	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.15: Fare regressions to investigate the impact of potential competition with 2009 q2-2010 q2 as pre and 2015 q1-2016q2 as post.

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*(SW present, AT potential entrant at both ends)	0.0786***	0.0732***	0.0608*
	(0.00844)	(0.00671)	(0.0352)
Post*(AT present, SW potential entrant at both ends)	0.284***	0.354***	2.076***
	(0.0545)	(0.0410)	(0.683)
Post*(SW present, AT potential entrant at one end)	0.0612***	0.0737***	0.0457
	(0.00602)	(0.00477)	(0.0373)
Post*(AT present, SW potential entrant at one end)	0.0577***	0.0644***	-0.159*
	(0.00641)	(0.00508)	(0.0867)
Distance	0.260***	0.259***	0.212***
	(0.00164)	(0.00408)	(0.0206)
UN-CO dummy	-0.0643***	-0.0585***	-0.0463***
	(0.00483)	(0.00413)	(0.0150)
Non-merging HHI	-0.00716***	-0.00518***	1.042**
	(0.00133)	(0.00142)	(0.431)
Potential competitiors	0.00779***	0.00347***	-0.0191**
	(0.000279)	(0.000365)	(0.00948)
Population	-0.0265***	-0.115***	0.256*
	(0.000939)	(0.0165)	(0.147)
Hub	0.0231***	0	0
	(0.00195)	(.)	(.)
Tourist	-0.0416***	0	0
	(0.00187)	(.)	(.)
Slot	0.0937***	0	0
	(0.00354)	(.)	(.)
constant	4.311***	6.889***	-11.42
	(0.0300)	(0.471)	(7.200)
Observations	147042	147042	82954
R^2	0.335	0.492	
Adjusted R^2	0.335	0.483	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.16: Fare regressions to investigate the impact on AirTran markets where Southwest did not resume service after the merger (using alternate time period). Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*(AT in pre, no SW in post)	0.152***	0.144***	0.140***
	(0.00803)	(0.00726)	(0.0183)
Distance	0.323***	0.254***	0.285***
	(0.00378)	(0.0115)	(0.0217)
quarter== 1.0000	-0.0263***	-0.0236***	0.0173
	(0.00714)	(0.00609)	(0.0140)
quarter== 2.0000	-0.0212***	-0.0230***	0.00336
	(0.00640)	(0.00537)	(0.0125)
quarter== 3.0000	-0.0143**	-0.0248***	-0.00233
	(0.00656)	(0.00542)	(0.0129)
UN-CO dummy	-0.0888***	-0.0776***	-0.110***
	(0.0144)	(0.0118)	(0.0250)
Non-merging HHI	-0.112***	0.00750	0.503***
	(0.00515)	(0.00587)	(0.165)
Potential competitiors	0.00704***	-0.00144*	-0.00509**
	(0.000610)	(0.000827)	(0.00226)
Population	-0.0163***	-0.0377	0.127
	(0.00166)	(0.0373)	(0.109)
Hub	0.0856***	0	0
	(0.00457)	(.)	(.)
Tourist	-0.0359***	0	0
	(0.00520)	(.)	(.)
Slot	0.0923***	0	0
	(0.0137)	(.)	(.)
constant	4.713***	4.793***	-4.167
	(0.0645)	(0.993)	(3.776)
Observations	32575	32575	10347
R^2	0.380	0.595	
Adjusted R^2	0.380	0.581	
F	1266.5	546.7	
rmse	0.354	0.291	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.17: Fare regression for all overlapping markets using benchmark control group: all markets not in treatment. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*treat(all overlapping)	0.0255***	0.0308***	0.0352***
	(0.00207)	(0.00194)	(0.00250)
Distance	0.254***	0.264***	0.262***
	(0.000649)	(0.00176)	(0.00254)
quarter== 1.0000	-0.0125***	-0.00814***	-0.00399***
	(0.00105)	(0.000921)	(0.00154)
quarter== 2.0000	0.00514***	0.00396***	0.0112***
	(0.000913)	(0.000798)	(0.00134)
quarter== 3.0000	0.0125***	0.00925***	0.0113***
	(0.000924)	(0.000812)	(0.00138)
UN-CO dummy	-0.0755***	-0.0605***	-0.0398***
	(0.00304)	(0.00245)	(0.00390)
Non-merging HHI	0.000124	0.00634***	0.149***
	(0.000610)	(0.000602)	(0.0188)
Potential competitiors	0.00246***	0.00300***	0.0000838
	(0.0000958)	(0.000143)	(0.000487)
Population	-0.0433***	-0.0975***	-0.0830***
	(0.000272)	(0.00721)	(0.0126)
Hub	0.0467***	0	0
	(0.000772)	(.)	(.)
Tourist	-0.0429***	0	0
	(0.000744)	(.)	(.)
Slot	0.0982***	0	0
	(0.00143)	(.)	(.)
constant	4.765***	6.345***	4.705***
	(0.0110)	(0.202)	(0.383)
Observations	1051716	1051716	473985
R^2	0.296	0.483	
Adjusted R^2	0.296	0.477	
F	19875.8	11154.6	
rmse	0.325	0.281	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.18: Fare regressions to investigate the impact of potential competition using benchmark control group. Dependent variable is logarithm of average market-airline fare

	(1)	(2)	(3)
	ln_fare	ln_fare	ln_fare
Post*(SW present, AT potential entrant at both ends)	0.0352***	0.0285***	0.0348***
	(0.00508)	(0.00412)	(0.00526)
Post*(AT present, SW potential entrant at both ends)	0.229***	0.269***	0.541***
	(0.0520)	(0.0323)	(0.0521)
Post*(SW present, AT potential entrant at one end)	0.0263***	0.0313***	0.0406***
	(0.00190)	(0.00186)	(0.00313)
Post*(AT present, SW potential entrant at one end)	0.0141***	0.0145***	-0.0263***
	(0.00260)	(0.00230)	(0.00438)
Distance	0.259***	0.257***	0.255***
	(0.000676)	(0.00190)	(0.00293)
UN-CO dummy	-0.0782***	-0.0681***	-0.0556***
	(0.00345)	(0.00272)	(0.00477)
Non-merging HHI	-0.00973***	0.00745***	0.160***
	(0.000686)	(0.000643)	(0.0238)
Potential competitiors	0.00592***	0.00323***	0.000484
	(0.000102)	(0.000153)	(0.000610)
Population	-0.0344***	-0.0908***	-0.0525***
	(0.000298)	(0.00741)	(0.0137)
Hub	0.0379***	0	0
	(0.000814)	(.)	(.)
Tourist	-0.0410***	0	0
	(0.000779)	(.)	(.)
Slot	0.0735***	0	0
	(0.00164)	(.)	(.)
constant	4.430***	6.188***	3.788***
	(0.0118)	(0.205)	(0.421)
Observations	935523	935523	374114
R^2	0.305	0.485	
Adjusted R^2	0.305	0.479	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.19: Fare regressions to investigate the impact on AirTran markets where Southwest did not resume service after the merger using the benchmark control group. Dependent variable is logarithm of average market-airline fare

	(1) ln_fare	(2) ln_fare	(3) ln_fare
Post*(AT in pre, no SW in post)	0.0945***	0.0936***	0.0757***
	(0.00450)	(0.00383)	(0.00695)
Distance	0.253***	0.257***	0.256***
	(0.000678)	(0.00190)	(0.00289)
quarter== 1.0000	-0.0130***	-0.00791***	-0.00453**
	(0.00112)	(0.000980)	(0.00182)
quarter== 2.0000	0.00427***	0.00350***	0.01000***
	(0.000976)	(0.000848)	(0.00153)
quarter== 3.0000	0.0140***	0.0106***	0.0147***
	(0.000986)	(0.000863)	(0.00158)
UN-CO dummy	-0.0826***	-0.0695***	-0.0567***
	(0.00346)	(0.00271)	(0.00470)
Non-merging HHI	0.00398***	0.00728***	0.139***
	(0.000647)	(0.000643)	(0.0240)
Potential competitiors	0.00390***	0.00307***	0.000603
	(0.000101)	(0.000151)	(0.000603)
Population	-0.0472***	-0.0885***	-0.0624***
	(0.000284)	(0.00741)	(0.0135)
Hub	0.0479***	0	0
	(0.000819)	(.)	(.)
Tourist	-0.0454***	0	0
	(0.000783)	(.)	(.)
Slot	0.0933***	0	0
	(0.00164)	(.)	(.)
constant	4.729***	6.132***	4.229***
	(0.0116)	(0.205)	(0.412)
Observations	935523	935523	374114
R^2	0.293	0.485	
Adjusted R^2	0.293	0.479	
F	16270.2	9574.2	
rmse	0.328	0.281	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3

The Impact of Health Insurance Provision on the Usage of Preventive Care: Evidence from the ACA (co-authored with Ngoc Ngo)

3.1 Introduction

The causal link between insurance coverage and health care usage has remained a widely discussed topic in health economics. Part of the motivation to study this issue stems from the need of designing public policies to provide high quality health care at an affordable cost to individuals and families in the United States. As the ultimate goal of health policy should focus on improving the health of the general population, it is important to consider the role insurance provision plays in the health status of an individual. Enrolling in health insurance plans can significantly reduce the financial burden faced by an individual in the case of a catastrophic event. The lessening of financial constraints may mean individuals will have more income leftover to make necessary arrangements for faster recovery. Furthermore, it may also improve mental health status as one would not need to constantly worry about pooling in large sums of money to fund health expenses. Hence, via the channel of lessening financial burden, it is reasonable to expect a positive relation between health insurance coverage and the health status of the individual.

However, the role health insurance coverage plays in improving the health of a person by preventing disease is not perfectly clear. Disease prevention can be enhanced by consuming various forms of preventive care such as immunization and routine check ups. Health insurance coverage can reduce the marginal cost faced by an individual in consuming an additional unit of preventive care service, and since preventive care service is a normal good, this may lead to an increase in consumption, which may play a critical role in improving the individual's health in future years. However, insurance coverage may also induce individuals to take up risky behavior. An individual may reason that since her insurance will cover the health bill if she falls ill, she would not need to invest time and effort in seeking preventive care services in the immediate future, thereby increasing the chance of disease.

Such possible opposing effects underscore the need of empirical analysis to fully understand the direction of the causal impact of health insurance provision on the usage of preventive care services. The ideal setting to examine this issue would be a controlled experiment in which preventive care usage of random groups of insured and uninsured populations is compared. This was the spirit of the RAND and Oregon health insurance experiments that were conducted several years ago. Our study uses more recent data by utilizing the implementation of the Patient Protection and Affordable Care Act (ACA) as a natural experiment.

The Obama Administration introduced the ACA in 2010 in an attempt to reform the health care sector. The policy change had several components, one of which was the extension of dependent coverage for young adults. Starting September 23, 2010, any individual up to the age of 26 can remain under their parents' health insurance plan as a dependent. Previously, one could be claimed as a dependent on their parent's plan only till the age of 19. It must be noted that several states already had similar dependent coverage mandates in place prior to the ACA. However, past literature has acknowledged that such state laws were much weaker and possibly ineffective (Barbaresco et. al. 2015; Monheit et al. 2011; Levine et al. 2011). The effectiveness of the ACA's dependent coverage mandate in increasing coverage rates of affected young adults has already been established by several papers that analyzed health care data from early years of the policy's implementation (Sommers et al. 2012; Cantor et al. 2012). Fewer studies have looked at the impact of the policy on the usage of health care services by young adults, and even fewer have tried to measure the impact on the usage of preventive health care services in particular. This presents us a valuable opportunity to address the gap in literature.

The rest of the paper is divided into the following parts: Section 3.2 discusses past literature related to our work. 3.3 presents a simple model that ties together the relation between insurance, moral hazard and preventive care usage. 3.4 discusses the details of the data used, and the relevant data definitions of the forms of preventive care studied. 3.5 discusses the setup of the regression discontinuity design (RDD), which is the primary empirical method used in this paper. 3.6 presents the results of the RD analysis, and 3.7 briefly discusses alternative econometric methods used to check the central findings.The conclusion in section 3.8 presents a summary, and discusses the public policy implications of the study.

3.2 Related Literature

The experimental ideal of examining the causal relation between health insurance provision and usage of health care services comes from the RAND and Oregon Health Insurance Experiments (HIE). The RAND HIE was a large-scale, multi-year controlled study conducted between 1971 and 1982 in which participants were randomly assigned various health insurance plans differing in the level of cost-sharing. Findings showed that participants that were assigned to insurance plans with cost sharing visited doctors less frequently, had fewer hospitalizations, and used less dental care and mental health treatment compared to participants that were assigned to full coverage insurance plans (RAND Corp. 2006). The Oregon study uses more recent data from a randomized lottery that took place in 2008 in Oregon to select low-income uninsured individuals who would be given a chance to apply for Medicaid. The differences in the rates of health care utilization among the treatment and control groups in the Oregon study were apparent beginning the first year of insurance coverage: individuals in the treatment group had higher rates of primary and preventive care usage as well as hospitalizations than the control group (Finkelstein et. al, 2012).

Several non-experimental research works have investigated this causal question. One such relevant work includes a paper by Card et. al (2008), in which the authors investigate the impact of Medicare on health care utilization. Using a regression discontinuity design by utilizing the fact that an individual becomes eligible for Medicare at age 65, the authors find that eligibility causes a sharp increase in the usage of health care services. We use a similar regression discontinuity design as our primary econometric tool. Some asymmetries across the type of health care service consumed are noted in their research: increases in the consumption of relatively cheaper services such as routine checks occur more among individuals who had low rates of insurance coverage before turning 65, whereas increases in more sophisticated and expensive services like hip replacements and bypass surgery occur among individuals who have supplementary insurance coverage after turning 65.

There are quite a few papers that look into the impact of the dependent coverage mandate of the ACA on the insurance coverage rate of young adults. These include Antwi et. al. (2015), Sommers and Kronick (2012) and Cantor et al. (2012). All these papers have found an unambiguous increase in insurance coverage of young adults due to the ACA. However, the study of preventive health care usage behavior of young adults is a relatively less studied topic in the literature. Barbaresco et. al. (2015) make a serious effort to study the impact of ACA on preventive health care usage of young adults. They use a difference-in-differences model with narrow age groups (ages 23-25 as treatment and 27-29 as control). Their results show that the policy did not lead to any significant increases in preventive care utilization, but led to an increase in ex-ante moral hazard behavior such as risky drinking. We aim to supplement their analysis by using few more econometric methods such as regression discontinuity and propensity scores.

3.3 Framework

We assume that the amount of preventive care service (a) consumed by an individual is a function of the price per unit of the service (p), and the risk taking tendency of the individual (r). 'r' is a function of the amount of loss faced by the individual when she falls ill (l), and a catch-all variable that represents the idiosyncratic preference towards risk for the individual (α).

The variables p, r and l are functions of an exogenous factor 't' representing time. We will assume that an increase in 't' represents a movement from pre-ACA period to post.

Thus, a = f (p(t), r(l(t),α)). The total derivative of a with respect to the exogenous t gives the following:

$$\frac{da}{dt} = \frac{\partial a}{\partial r}\frac{dr}{dt} + \frac{\partial a}{\partial p}\frac{dp}{dt} = \frac{\partial a}{\partial r}\frac{\partial r}{\partial l}\frac{dl}{dt} + \frac{\partial a}{\partial p}\frac{dp}{dt} = A + B \tag{3.1}$$

The amount of preventive care consumed has a negative relation with the risk taking tendency of the individual. Risky behavior entails not seeking preventive care. Furthermore, the law of demand implies that preventive care consumption increases with the fall in its price. Thus, $\frac{\partial a}{\partial r}$ <0, and $\frac{\partial a}{\partial p}$ <0. It is reasonable to assume that the individual would take more risks if they know that the loss they would face in the case of illness or accident is small. Hence, $\frac{\partial r}{\partial l}$ <0 as well.

Following the implementation of the ACA, p and l decrease. This indicates that $\frac{dl}{dt}$ and $\frac{dp}{dt}$ are both negative. Putting together the signs of the various components of the total derivative $\frac{da}{dt}$ reveals that A is negative whereas B is positive. Note that A represents the impact of ex-ante moral hazard on the consumption of preventive care, whereas B represents the demand effect. This simple analysis formalizes the two opposing forces at work. By empirically measuring the change in preventive care consumption following the implementation of the ACA, one can gauge the relative strengths of these opposing forces.

3.4 Data

We use the Medical Expenditure Panel Survey Household Component (MEPS-HC) for analysis. Published by the Agency for Healthcare Research and Quality (AHRQ), the MEPS-HC comprises of health care data from a nationally representative subsample of households across the United States. The dataset provides information about the amount of specific kinds of health care usage, insurance coverage, health status and socio economic status of individuals participating in the survey.

The MEPS-HC has an overlapping panel design. Every year, a new panel of households is selected, and the selected households are interviewed for a period of two years by using five rounds of interviews. Therefore, in a given year, the MEPS will comprise of data from two panels, one of which was recently selected that year, and the other being the panel that was selected the year before.

It must also be noted that the survey is conducted year round, and different households respond in different days of the year. For this reason, the MEPS uses the concept of a "reference period". For example, for panel 19 (which runs from 2014 to 2015), the reference period for the first round begins on the first day of the year of 2014 (January 1), and ends on the date when the survey respondent fills out the survey. For subsequent rounds, the reference period covers the time between last round's survey and the current round's survey for a given individual. The reference period for the final round (round 5), ends on the last calendar day of the second survey year for the selected panel.

The regression discontinuity analysis uses data from 2011-2015, whereas the robustness checks carried out using difference-in-differences and propensity scores utilize 2009 as pre-ACA and 2011 as post-ACA data.[1] There are primarily two types of dependent variables that are studied in this analysis: insurance coverage and preventive care consumption parameters. The insurance coverage indicator variable takes a value of "1" if the surveyed individual reported as having coverage throughout each interview round in the previous year.

The MEPS-HC has a number of variables classified as under preventive care usage indicators. For this study, we were able to isolate a group of usage variables that are most relevant to individuals in the young adult age group (19-34). These include annual flu shots, routine check ups, blood pressure checks, pap smear tests and breast exams.[2] Other preventive care usage variables in the MEPS such as colonoscopy, sigmoidoscopy, mammogram and prostrate checks were not considered since they are recommended for older age groups. Dental check ups were excluded from the analysis since it is not affected by the ACA coverage mandate, although it is a relevant form of preventive care usage indicator for the age group studied.

Questions about the consumption of preventive care were asked in round five of the first survey year of the panel, and round three of the second survey year. Table 3.1 summarizes the usage variables considered, and their binary classification as used in analysis.

[1]the post ACA data for pap checks and breast exams is the 2013 MEPS-HC. The reason for a separate year's data for these variables is discussed in section 3.6.2.

[2]Pap smear tests and breast exams are only relevant for females, so these variables were used for only that subset of the population.

Table 3.1: Relevant preventive care variables used in the study.

Preventive Care	Description in MEPS-HC	Method of coding[a]
Blood pressure check	Time since last check	"1" if within the last year, "0" otherwise
Pap smear test	Time since last check	"1" if within the last three years, "0" otherwise
Breast exam	Time since last check	"1" if within the last three years, "0" otherwise
Routine check	Time since last check	"1" if within the last year, "0" otherwise
Flu shot	Time since last shot	"1" if within the last year, "0" otherwise

[a]Our method of coding is consistent with recommended usage of preventive care services as outlined by the US preventive services taskforce (www.uspreventiveservicestaskforce.org)

3.5 Empirical Approach

3.5.1 Identification

Our research question is appropriate for an RD analysis since the rules determining treatment is precisely defined: post ACA implementation, young adults would lose their parents' insurance coverage after reaching age 26. However, there are some employer plans that provide coverage to a dependent throughout the 26th year. Furthermore, programs like COBRA[3] enable individuals to stay on coverage for a temporary period when they are switched off from the policy. For this reason, the cutoff is of a 'fuzzy' nature; while age would to an extent predict coverage, it is not a perfect predictor. Hence, examining preventive care usage in a sharp RDD by using age as the running variable would not be an appropriate identification approach. Instead, we use a fuzzy regression discontinuity design. Age is used to predict insurance coverage, and the fitted values of coverage are regressed on preventive care usage to determine the causal relation. More details on the fuzzy RDD setup is discussed in 3.6.2.

The causal impact in an RD design is identified assuming that each individual does not have a precise control over the assignment variable; and that all unobserved determinants of the outcome are continuously related to the running variable. In our case, the second assumption means that

[3]http://www.cobrainsurance.com

as long as two people born one month apart (of the cutoff age of 26) do not exhibit any systematic differences other than their eligibility to the ACA dependent coverage, the RD approach is as good as random. These assumptions cannot be directly tested. However, there are tests that can give us suggestive evidence whether the assumption is satisfied. We will discuss the tests in the next sub-section. Our specification follows a basic RD form:

$$y_i = \alpha + \beta t_i + f(x_i) + \epsilon_i$$

$$\forall x_i \in (c - h, c + h)$$

where y_i is the dependent variable, t_i is the treatment status, x_i is the running variable, h is the bandwidth and c is the cutoff point. $f(x_i)$ is a continuous n-order polynomial function of the running variable. We allow the slope of this function to vary on each side of the cutoff. We use local polynomial regression discontinuity point estimators, and conduct optimal bandwidth selection using MSE-optimal bandwidth algorithm. Standard error is clustered at the month-year of birth level to adjust for the specification error in the running variable.

3.5.2 Preliminary Checks

We present two standard validity checks for the RD design. First, we test the density of the forcing variable (age at time of survey) around the cutoff, to investigate if there is a discontinuity in the distribution of the forcing variable at the threshold. This is an indirect test of the identifying assumption that each individual does not have a precise control over the assignment of whether they belong to the treatment group. A discontinuity at the threshold would suggest that people can manipulate the forcing variable, which would violate our first assumption. In our context, this would imply that in the MEPS-HC, we have a measurable difference between the proportion of individuals who are just over 26, and just below. We perform a McCrary density test for this first check. As shown in Figure 3.1, there is no significant break in the density of the forcing variable (age) around the threshold (26 years). Each dot represents the proportion of individuals of a specific age in year-months. We see that the 95 percent confidence bands at either side of the cutoff overlaps to a great extent.

The second test examines the discontinuity of the covariates at the cutoff. There should be no jump at the threshold in any covariates. This presents an indirect test for the second assumption. Although the unobserved variables cannot be tested, if we find continuity for the observed variables, we have more reasons to believe that it would be the same for the unobserved variables. On the other hand, if the observed covariates jump at the cutoff, one should doubt the identifying assumption. Each graph in Figure 3.2 presents the local average of the outcomes plotted against the running variable, with overlaid smoothed linear regression lines using raw data on each side of the cutoff. Each dot in the graph represents the value of the covariate being examined for individuals grouped by their age in year-months. The two gray lines show the 95 percent confidence intervals. The predetermined covariates we include are: employment, education and marital status, individual income, and region.[4] As seen in these figures, there does not exist a significant jump at the threshold for these covariates. Next, we proceed with the RDD to investigate the effects of the ACA coverage expansion on insurance and preventive care indicators.

3.6 Effects of the ACA Dependent Coverage Expansion

3.6.1 Insurance Coverage

We use data from the 2011-15 MEPS-HC in this section. An individual's age (in year-months) at the time of survey is used as the running variable. Analysis begins with a visual inspection to see if there is a discrete break in insurance coverage for survey respondents turning 26 at the time of survey (the cutoff). Regarding insurance coverage, an individual is coded as being covered if they had insurance throughout the last year since the interview date.

Figure 3.3 displays the plots for insurance coverage using linear and quadratic age.[5] As before, each dot represents the average coverage rate of individuals in the same year-month bin. As seen from the two plots, there is a break at the cutoff of age 26. Individuals just above the age of 26

[4] It must be noted that several states already had dependent coverage mandates in effect before ACA implementation. The ability to identify the states where respondents belonged to could have allowed for an easier difference in difference analysis to identify the policy's effects. The MEPS-HC only has the region variable to identify geographical location.

[5] Usage of linear and quadratic functional forms of age follows past literature. See Card et. al (2008)

have lower insurance coverage rates than those below. This gives visual evidence of the change in an individual's health insurance status when they drop out of ACA eligible coverage past age 26. It is likely that the break would have been larger had programs like COBRA, that give temporary insurance to individuals right after they fall out of eligibility, not existed. This is consistent with the findings of previous literature (see Antwi et. al. (2015), Sommers and Kronick (2012) and Cantor et al. (2012)).

In addition to visual plots, we also run an RD model with age as the running variable, and insurance coverage as the outcome. Table 3.3 presents the RD treatment effects on the probability that an individual has insurance coverage. As described in the tables, the different specifications correspond to different functional forms and the inclusion of covariates. Panel A includes the whole sample in the bandwidth range. Averaging the values of the coefficients on coverage we find that aging out of ACA dependent care eligibility leads approximately to four percent point decrease in coverage.

In Panel B, we exclude individuals of age 26 as their eligibility status is ambiguous for reasons discussed earlier. There exists a trade-off in including this ambiguous year, if we drop it, the continuity requirement of regression discontinuity is violated, but including it is not ideal either since we have individuals between 26 and 27 that may still be receiving dependent coverage. We perform the exercise of dropping age 26 to compare our results to the case when we include it in the sample, as done in panel A. According to Panel B, aging out of ACA eligibility leads approximately to a 6.6 percent point decrease in coverage. Panel C shows a placebo test, in which we use 2009 data, when ACA has not been implemented. There is no significant change in the probability of having insurance coverage in 2009 around the cutoff.

We also perform some heterogeneous tests on insurance coverage by looking at population subsamples in the relevant age group that were either high school dropouts, or active in the military. The results, presented in Table 3.5 show that high school dropouts face a seven percent point decrease in coverage after crossing the threshold of 26 years of age. It is likely that high school dropouts get coverage through their parents, whereas those who complete high school may have started work or college, and have other forms of coverage, as a result of which their insurance coverage is not significantly affected by the dependent coverage mandate. Similarly, individuals in the military get coverage through other programs such as Tricare. Respondents who were not active in the military face approximately a four percent point decrease in coverage after reaching age 26.

3.6.2 Preventive Care

Next, we use a fuzzy RD setup to investigate the jump in preventive care usage caused due to the dependent coverage mandate of the ACA. A fuzzy design works by exploiting the jumps in the probability of treatment status. Fuzzy RD is appropriate in our analysis because the age cutoff may not deterministically describe the treatment status of individuals in the survey due to reasons discussed earlier.

Fuzzy RD has a setup similar to a two-staged least squares, in which the discontinuity is used as an instrument for the treatment status. Specifically, in the first stage, the cutoff of age (at 26) is used as the instrument to predict the individual insurance coverage status. In the second stage, the predicted value of coverage status, along with other control variables are regressed on preventive care indicators. The results for the second stage are reported in Table 3.4.

Before interpreting the results, we would like to acknowledge a feature of the dataset that makes accurate identification difficult to make. As described in Table one, the preventive care variables take a value of "1" if the service was consumed within the last year. Since the cut-off is defined as age 26, we would potentially have individuals on the right hand side of the cut-off in their 26th year who would have been covered by the ACA at some point within the last year, and hence would have consumed the preventive care service. Thus, there are some individuals on the right hand side of the cut-off that were still covered by the ACA during the time they used the service. Note that all individuals on the left hand side of the cut-off can be clearly identified as ACA eligible. For this reason, any break in continuity seen for preventive care usage at the cut-off will be a conservative estimate of the policy's true impact. Hence, due to the limitation of the dataset, our RDD framework's results cannot be directly translated as being the policy's true impact. However, any observed discontinuous jump at the cut-off would clearly reject the null hypothesis that the policy did not have any impact.

For blood pressure checks, routine checks and flu shots, we use MEPS-HC data from 2011 to 2015. For breast cancer checks and pap tests, we only use a sub-sample of female respondents, and use data for 2013 through 2015. The data for 2011-12 for these two tests are not considered since the recommended dosage for these services is once in three years. If we use data for 2012, an individual responding that she had a breast exam could have had it in 2009, the year before ACA was implemented. Therefore, using data for 2011-12 for these two check-ups would not provide meaningful information about the policy's effects.

As seen in Table 3.4, the only significant coefficients are in column (1). This specification uses coverage as the independent variable (along with some control variables as described in the table notes). It is likely that insurance coverage is endogenous: health conscious individuals and hypochondriacs may be more likely to have insurance coverage, and at the same time consume more preventive care. This yields the positive coefficient. Due to the possible misspecification that arises due to endogeneity, we will not put a lot of weight on the results with (1).

In column (2), the age cut-off (defined as the dummy variable 'post' which takes value of '1' if the individual is greater than age 26) is the primary independent variable. (3) and (4) are specifications for the fuzzy RD where we use the age cut-off as an instrument for coverage, and regresses the fitted values of coverage on preventive care variables. All three specifications (2-4) show statistically insignificant change in preventive care usage.

In addition, we run the model using other functional forms and bandwidth selection algorithms. The results, reported in appendix table 3.8 and 3.9 also show that there is no statistically significant change in preventive care usage in any of the specifications other than the endogenous OLS regression. Past literature suggests that the impact of health insurance coverage may vary among groups. We test for heterogeneous effects for the following categories: sex (Table 3.10), race (Table 3.11) and education (Table 3.12). All regressions fail to show a statistically significant effect of dropping out of dependent care eligibility on preventive care usage. We therefore conclude in this section that the policy does not affect the usage of preventive care. Next, we check this central finding using other econometric methods.

3.7 Alternative methods of empirical analysis

The results of the regression discontinuity design were checked by using difference-in-differences, and propensity score matching combined with difference-in-differences. Both methods use data from the 2009 MEPS-HC as pre-ACA, and 2011 MEPS-HC as post for blood pressure checks, routine checks and flu shots, and 2009 MEPS-HC as pre and 2013 as post for breast cancer checks and pap tests.

In the analysis that uses difference-in-differences only, survey respondents between the ages of 19 and 25 are considered the treatment, and 27-34 year are the control. Following literature, individuals aged 26 are not considered in our analysis since their treatment status is ambiguous as some of them could or could not be covered by the ACA mandate depending on their birth date (Akosa et. al 2013, Barbaresco et al. 2015). The control covariates include demographic variables such as age, sex, race, region, education, student status and wage income. The same dependent variables as in the RD analysis are used. The results are presented in Table 3.6. Note that since the dependent variables can only take values of "1" or "0", the regression can be interpreted as linear probability models. As seen from Table 5, none of the preventive care variables exhibit a statistically significant increase.

One crucial assumption of a difference in difference approach is the parallel trend in the ex-ante period. If this assumption does not hold, the estimates from a difference-in-differences model will be biased. We introduce propensity scores to the diff-in-diff analysis as a means of relaxing the parallel trends assumption. The identifying assumption using this approach is that, conditioned on propensity scores, the control group would have similar change in preventive care usage compared with the treatment over the two time periods if they had also affected by the policy. The covariates used to perform matching in the first stage are marital status, education level, employment status, income and family income in dollars and family income as a percentage of the poverty line. Nearest neighbor matching was performed with replacement to match 19-25 age group in 2009 with 27-34 year group in 2009. The same was done for 2011 (2013 for pap checks and breast exams). The matching for the two time periods were done separately because individuals in 2009 are not the same as individuals in 2011 in the MEPS-HC. In the second stage, the average effect on the treated is estimated using a diff-in-diff approach on the matched sample. Each observation is weighted by the number of times it is used to match. The same control covariates as in the aforementioned simple diff-in-diff model were used. The results presented in Table 3.7 once again reaffirm the central finding of our analysis: there is no statistically significant effect of the policy change on the usage of the forms of preventive care usage covered by the ACA coverage extension for young adults (at the five percent significance level, at least).

3.8 Conclusion

Empirical analysis conducted using regression discontinuity design in this paper has shown that while the dependent coverage mandate of the ACA enabled many young adults to gain health insurance coverage, it did not lead to a significant increase in the usage of preventive care services. Once again, we acknowledge that our identification approach only establishes a conservative estimate of the policy's true effects (as discussed in section 6). However, when we investigated the same question using alternative econometric methods, we reached the same conclusion that preventive care usage is not significantly affected by the policy. These robust findings are also consistent with the results of Barbaresco et. al (2015), who investigated the same topic using a different data source.

As formalized in the theoretical framework section of the paper, usage of preventive care is influenced by two opposing forces: risk taking behavior and the price effect. It could be the case that individuals start indulging in more risk taking behavior after receiving health insurance. Upon receipt of health insurance, young adults may become incentivized to be careless about preventing disease, reasoning that if they fall ill, the medical bills will be covered by insurance. Past literature on this topic provides support for this conjecture: Barbaresco et. al (2015) found that risky drinking increases following the ACA dependent coverage extension.

It must be noted however, that our results are specific to the behavior of young adults. Anecdotal evidence suggests that young adults take more health risks than older age groups. Only by conducting a counterfactual analysis for a similar policy that is applicable to individuals of other ages would allow us to generalize these findings.

Apart from the shortcomings already discussed, our work also suffers from the possibility that the data come from self-reported surveys, and are subject to the respondent's recall error. It is also likely that respondents could over or under report health care consumption due to idiosyncratic reasons. Nevertheless, this issue should not be a major problem since the sample size we use is quite large, and hence may not be systematically biased.

If the results are to be taken at face value, then it has some implications for public policy. The issue of reducing overall health care costs has dominated health policy debate for many years. Increasing preventive care usage may lead to increased health care utilization, which can raise costs for the system in the short run, but eventually it may reduce health care costs if individuals are prevented from falling ill to serious disease, and hence do not seek more expensive forms of treatment. If one of the goals of the ACA is to encourage the consumption of preventive care, then it seems to be ineffective. Along with coverage expansion, awareness programs that educate individuals about the importance of preventive care should also be implemented.

3.9 References

Antwi, Yaa Akosa, Asako S. Moriya, and Kosali Simon. "Effects of Federal Policy to Insure Young Adults: Evidence from the 2010 Affordable Care Act's Dependent-Coverage Mandate." *American Economic Journal: Economic Policy* 5.4 (2013): 1–28. Print.

Barbaresco, Silvia, Charles J. Courtemanche, and Yanling Qi. "Impacts of the Affordable Care Act Dependent Coverage Provision on Health-Related Outcomes of Young Adults." *Journal of Health Economics* 40 (2015): 54–68. *onesearch.northeastern.edu*. Web.

Cantor, Joel C. et al. "Early Impact of the Affordable Care Act on Health Insurance Coverage of Young Adults." *Health Services Research* 47.5 (2012): 1773–1790. *onesearch.northeastern.edu*. Web.

Card, David, Carlos Dobkin, and Nicole Maestas. "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare.(Report)." *American Economic Review* 98.5 (2008): 2242. Print.

Depew, Briggs, and James Bailey. "Did the Affordable Care Act's Dependent Coverage Mandate Increase Premiums?" *Journal of Health Economics* 41 (2015): 1. Print.

Gaynor, Martin, Carol Propper, and Stephan Seiler. "Free to Choose? Reform and Demand Response in the English National Health Service." 2012. *onesearch.northeastern.edu*. Web. 6 Apr. 2017.

Girma, Sourafel, and Holger Görg. "Evaluating the Foreign Ownership Wage Premium Using a Difference-in-Differences Matching Approach." *Journal of International Economics* 72.1 (2007): 97–112. *onesearch.northeastern.edu*. Web.

Han, Xuesong et al. "Has Recommended Preventive Service Use Increased after Elimination of Cost-Sharing as Part of the Affordable Care Act in the United States?" *Preventive Medicine* 78 (2015): 85. Print.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64.221 (1997): 605. Print.

Hofer, Adam N., Jean Marie Abraham, and Ira Moscovice. "Expansion of Coverage under the

Patient Protection and Affordable Care Act and Primary Care Utilization." *The Milbank Quarterly* 89.1 (2011): 69–89. Print.

Kenkel, Donald S. "The Demand for Preventive Medical Care." *Applied Economics* 26.4 (1994): 313–325. *onesearch.northeastern.edu*. Web.

Krueger, Alan B., and Ilyana Kuziemko. "The Demand for Health Insurance among Uninsured Americans: Results of a Survey Experiment and Implications for Policy." *Journal of Health Economics* 32.5 (2013): 780–793. *onesearch.northeastern.edu*. Web.

Lau, Josephine S. et al. "Improvement in Preventive Care of Young Adults after the Affordable Care Act: The Affordable Care Act Is Helping." *JAMA pediatrics* 168.12 (2014): 1101–6. *onesearch.northeastern.edu*. Web.

Lee, David, and Thomas Lemieux. "Regression Discontinuity Designs in Economics." 2009. *onesearch.northeastern.edu*. Web. 6 Apr. 2017.

Lyk-Jensen, Stéphanie Vincent, Cecilie Dohlmann Weatherall, and Peter Winning Jepsen. "The Effect of Military Deployment on Mental Health." *Economics and Human Biology* 23 (2016): 193–208. *onesearch.northeastern.edu*. Web.

Maciosek, Michael V. et al. "Greater Use of Preventive Services in U.S. Health Care Could Save Lives at Little or No Cost." *Health affairs (Project Hope)* 29.9 (2010): 1656–60. *onesearch.northeastern.edu*. Web.

Monheit, Alan C., and Jessica Primoff Vistnes. "Health Insurance Enrollment Decisions: Preferences for Coverage, Worker Sorting, and Insurance Take-Up." *Inquiry* 45.2 (2008): 153–167. Print.

Qi, Yanling. "Three Essays on the Impact of the Affordable Care Act Expansion of Dependent Coverage for Young Adults." ProQuest Dissertations Publishing, 2015. Print.

Sommers, Benjamin D. et al. "The Affordable Care Act Has Led to Significant Gains in Health Insurance and Access to Care for Young Adults." *Health affairs (Project Hope)* 32.1 (2013): 165–74. *onesearch.northeastern.edu*. Web.

3.10 Regression and Summary Tables

Table 3.2: SUMMARY STATISTICS

	Mean	S.D.	Min	Max
Insurance Coverage	0.84	0.37	0	1
Blood Pressure Check	0.54	0.50	0	1
Routine Check	0.27	0.44	0	1
Flu Shot	0.38	0.48	0	1
Breast Cancer Check	0.20	0.40	0	1
Pap Test	0.27	0.44	0	1
N	181,529			
	19 - 25 Years Old		27 - 34 Years Old	
	Mean	S.D.	Mean	S.D.
Insurance Coverage	0.73	0.45	0.72	0.45
Blood Pressure Check	0.60	0.49	0.66	0.48
Routine Check	0.46	0.50	0.49	0.50
Flu Shot	0.23	0.42	0.28	0.45
Breast Cancer Check	0.21	0.41	0.28	0.45
Pap Test	0.33	0.47	0.44	0.50
N	17,882		20,158	
Covariates				
Married	0.46	0.50	0	1
Education Level	0.89	0.31	0	1
Without College Degree	0.11	0.31	0	1
With at least College Degree				
Student Status	0.06	0.24	0	1
Employed	0.49	0.50	0	1
Income	21,422	31,548	0	409,118
Region				
Northeast	0.16	0.36	0	1
Midwest	0.19	0.39	0	1
South	0.38	0.48	0	1
West	0.28	0.45	0	1

Table 3.3: RD Treatment Effects on Insurance Coverage

	Quadratic RD $\hat{h}$ bandwidth (1)	Quadratic RD $\hat{h}$ bandwidth (2)	Quadratic RD $2\hat{h}$ bandwidth (3)	Linear RD $\hat{h}$ bandwidth (4)
Panel A: Whole Sample				
Insurance Coverage	-0.0400**	-0.0454***	-0.0460***	-0.0300**
	(0.016)	(0.017)	(0.012)	(0.014)
Bandwidth	78	78	156	41
N	33,295	33,558	58,283	17,540
Panel B: Sample excluding 26 years old				
Insurance Coverage	-0.0741***	-0.0637***	-0.0671***	-0.0571***
	(0.017)	(0.017)	(0.013)	(0.015)
Bandwidth	78	78	156	42
N	33,251	33,515	57,986	17,973
Panel C: Placebo Test (Using 2009 data)				
Insurance Coverage	-0.0605	-0.0521	-0.0307	-0.0379
	(0.048)	(0.047)	(0.036)	(0.045)
Bandwidth	86	86	172	83
N	4,199	4,237	7,430	2,009

Notes: Columns 1 and 2 report local RD regressions with quadratic polynomials in months old using the optimal bandwidth $\hat{h}$. Column 1 includes control variables, column 2 does not. Column 4 reports the local RD regressions with linear polynomial in months old using the optimal bandwidth $\hat{h}$. Column 3 reports local RD regressions with quadratic polynomials in months old using twice of optimal bandwidth $2^*\hat{h}$. The optimal bandwidth is estimated using the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. Panel A reports the results for the full sample. Panel B reports the results for the sample that excludes individuals of 26 years old. Panel C reports the results for the placebo test that use survey in 2009. The dependent variable in the first row in each panel is a dummy variable that value 1 if the individual has insurance coverage in the survey year. All specifications control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.4: Insurance Effects on Preventive Care Usage

Outcomes	OLS	RF	IV	IV
	(1)	(2)	(3)	(4)
Blood Pressure Check	0.296***	-0.00979	0.258	0.287
	(0.008)	(0.018)	(0.491)	(0.505)
Bandwidth	53	53	53	53
N	22,629	22,629	22,665	22,629
Routine Check	0.263***	-0.027	0.79	0.806
	(0.008)	(0.019)	(0.631)	(0.603)
Bandwidth	51	51	51	51
N	21,791	21,791	21,826	21,791
Flu shot	0.140***	0.0163	-0.391	-0.484
	(0.006)	(0.019)	(0.568)	(0.643)
Bandwidth	61	61	61	61
N	25,951	25,951	25,989	25,951
Breast Cancer Check	0.189***	0.021	-1.301	-1.64
	(0.012)	(0.029)	(3.146)	(4.658)
Bandwidth	79	79	79	79
N	10,183	10,183	10,196	10,183
Pap Test	0.162***	0.00731	-0.325	-0.572
	(0.013)	(0.025)	(1.685)	(2.511)
Bandwidth	80	80	80	80
N	10,183	10,183	10,196	10,183

Notes: Table reports local RD regressions with quadratic polynomials in months old using the optimal bandwidth $\hat{h}$ estimated by the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. Column 1 reports the OLS results using insurance coverage as the independent variable for an optimal bandwidth $\hat{h}$. Column 2 reports the reduced-form (RF) RD treatment effects of individuals age below 26 with a quadratic control function in the months old on each side of the discontinuity. Columns 3 and 4 report the two-stage least-squares (IV) RD treatment effects (by using treatment as an instrument for insurance coverage). Columns 1, 2 and 4 control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Column 3 does not include control variable. Breast cancer exam and Pap test specifications only include sample of female respondents. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.5: RD Treatment Effects on Insurance Coverage by High School Dropout and Being in Military

	High School Dropout		Military Active	
	Yes	No	Yes	No
	(1)	(2)	(3)	(4)
Insurance Coverage	-0.0711**	-0.0247	-0.0611	-0.0394**
	(0.030)	(0.020)	(0.065)	(0.016)
Bandwidth	78	78	78	78
N	9570	23725	134	33161

Notes: Columns 1 and 2 report the RD treatment effects on insurance coverage by whether individuals are high school dropout. Column 3 and 4 report the RD treatment effect on insurance coverage by whether individuals are currently in military. The optimal bandwidth h estimated by the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. All specifications control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.6: Difference in Difference Results

	Blood Pressure Check	Routine Check	Flu Shot	Breast Cancer Check	Pap Test
Treatment x Post	0.000856	-0.0131	-0.011	-0.0371	-0.017
	(0.017)	(0.018)	(0.015)	(0.025)	(0.021)
Treatment	0.0133	0.0243	0.0106	-0.0236	0.0255
	(0.021)	(0.022)	(0.019)	(0.031)	(0.024)
Post	0.00347	0.0340***	0.00454	-0.0444**	-0.0320**
	(0.012)	(0.012)	(0.011)	(0.019)	(0.014)
N	12,021	12,021	12,021	6,456	6,456
adj. R-sq	0.103	0.084	0.042	0.061	0.131
rmse	0.461	0.475	0.41	0.484	0.397

Notes: All specifications include control variables age, dummies for whether individuals speak English, sex, marital status, race, region of living, education levels, student status, employment status, and income. Pap test specification only includes sample of female respondents. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.7: Difference in Difference with PSM

	Blood Pressure Check	Routine Check	Flu Shot	Breast Cancer Check	Pap Test
Treatment x Post	-0.0242	0.0163	-0.0282*	0.026	0.0171
	(0.019)	(0.019)	(0.016)	(0.027)	(0.023)
Treatment	0.0384***	0.0234	-0.0292**	-0.0795***	-0.123***
	(0.015)	(0.015)	(0.012)	(0.022)	(0.017)
Post	0.0301**	0.00742	0.0210*	-0.107***	-0.0635***
	(0.013)	(0.013)	(0.012)	(0.020)	(0.015)
N	11,394	11,394	11,394	5,779	5,779
adj. R-sq	0.025	0.016	0.017	0.046	0.103
rmse	0.488	0.486	0.412	0.486	0.424

Notes: All specifications include control variables age, dummies for whether individuals speak English, sex, marital status, race, region of living, education levels, student status, employment status, and income. Pap test specification only includes sample of female respondents. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

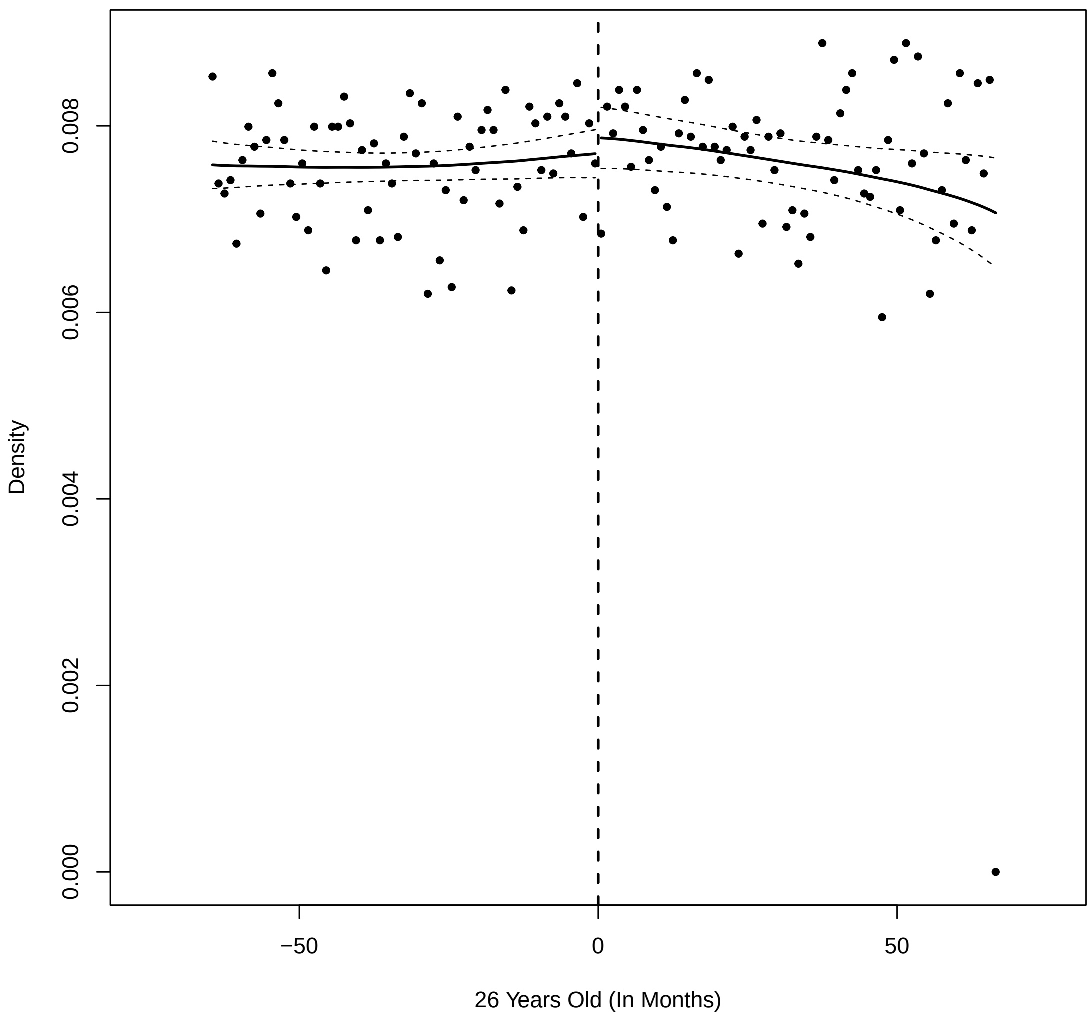


Figure 3.1: McCrary Density Test

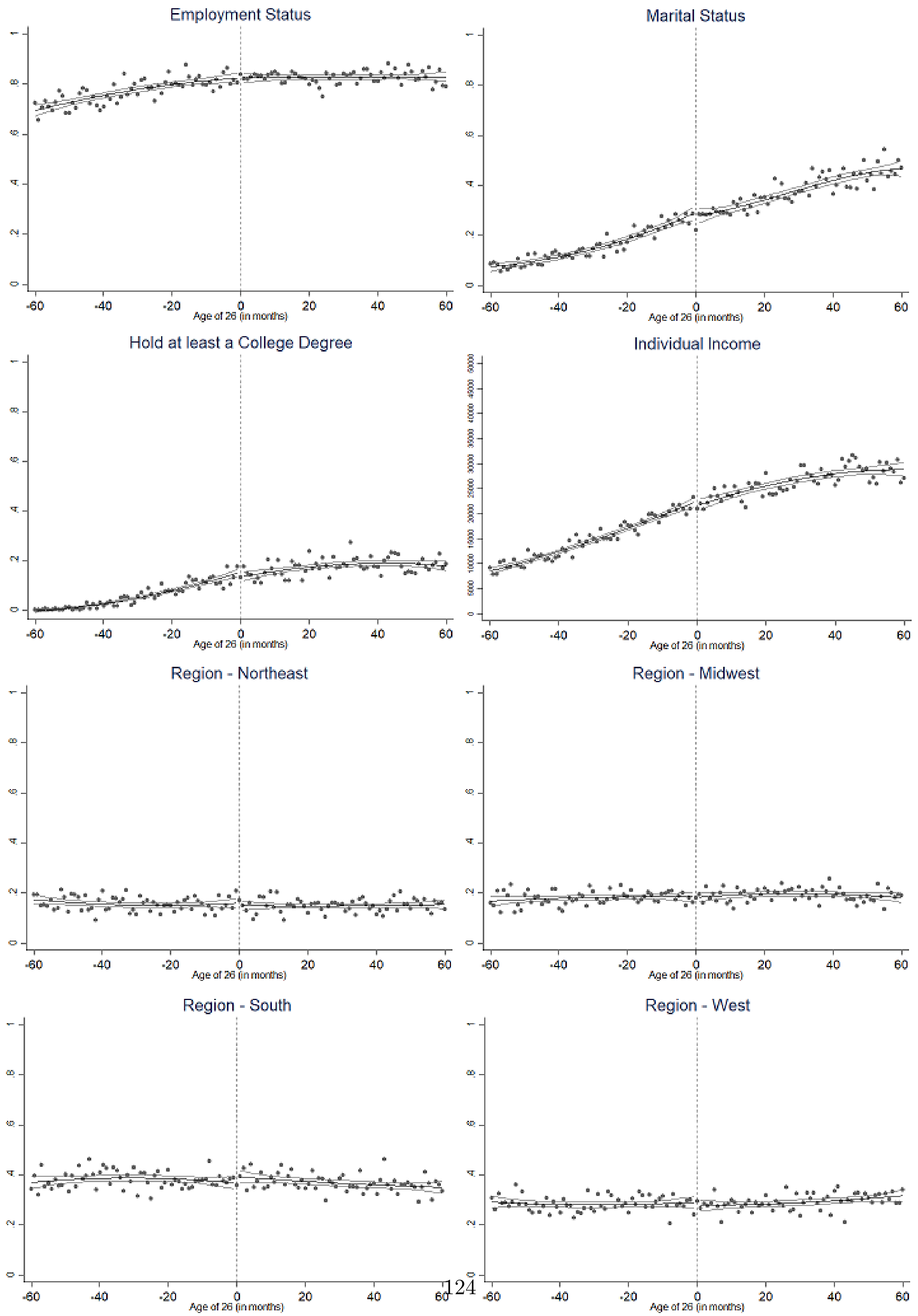


Figure 3.2: Balanced Covariates

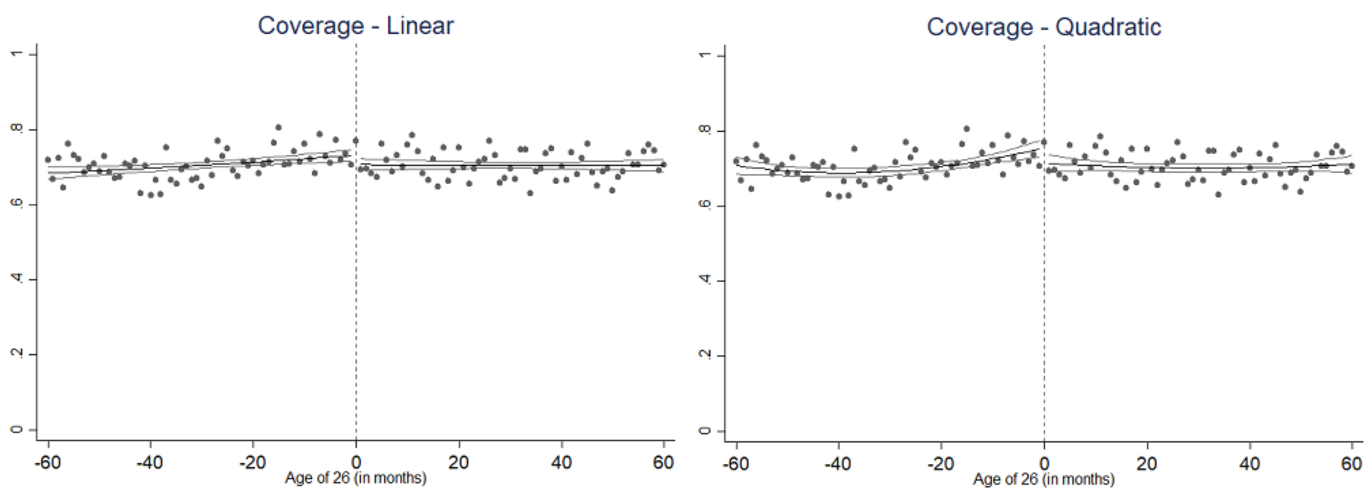


Figure 3.3: RD Treatment Effects on Insurance Coverage

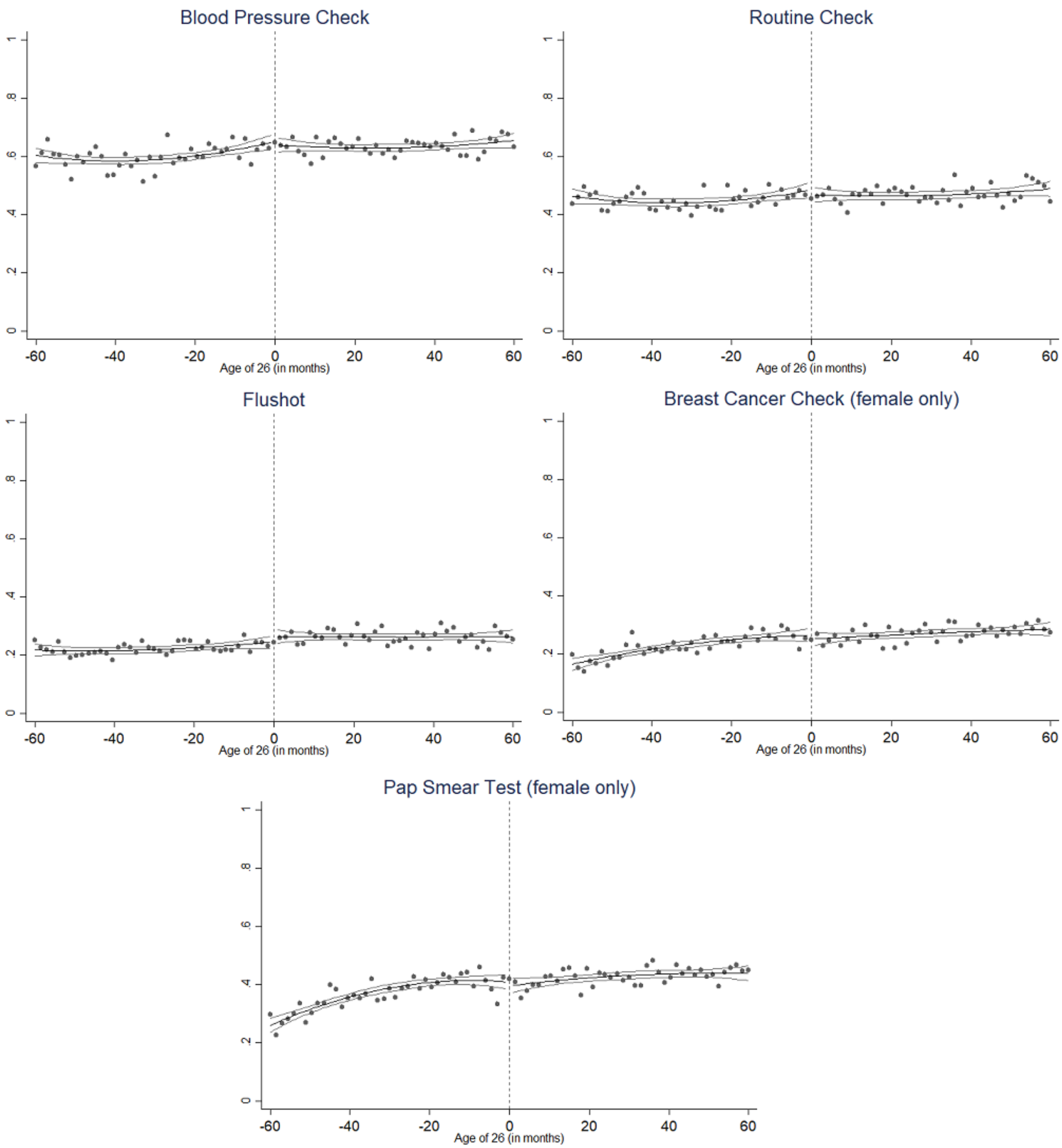


Figure 3.4: Treatment Effects

3.11 Appendix

Table 3.8: Insurance Coverage on Preventive Care Usages with Linear RD

Outcomes	OLS	RF	IV	IV
	(1)	(2)	(3)	(4)
Blood Pressure Check	0.311***	-0.00779	0.282	0.278
	(0.008)	(0.013)	(0.382)	(0.422)
Bandwidth	45	45	45	45
N	18,893	18,862	18,893	18,862
Routine Check	0.266***	-0.00486	0.187	0.252
	(0.007)	(0.013)	(0.755)	(0.663)
Bandwidth	57	57	57	57
N	24,338	24,300	24,338	24,300
Flu shot	0.159***	0.0196	-0.62	-0.687
	(0.007)	(0.015)	(0.606)	(0.649)
Bandwidth	44	44	44	44
N	18,457	18,426	18,457	18,426
Breast Cancer Check	0.209***	0.00276	1.514	3.063
	(0.014)	(0.022)	(6.166)	(60.910)
Bandwidth	66	66	66	66
N	8,403	8,392	8,403	8,392
Pap Test	0.182***	-0.0232	3.295	3.233
	(0.016)	(0.020)	(10.870)	(9.226)
Bandwidth	57	57	57	57
N	7,242	7,232	7,242	7,232

Notes: Table reports local RD regressions with linear polynomials in months old using the optimal bandwidth $\hat{h}$ estimated by the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. Column 1 reports the OLS results using insurance coverage as the independent variable for an optimal bandwidth $\hat{h}$. Column 2 reports the reduced-form (RF) RD treatment effects of individuals age below 26 with a quadratic control function in the months old on each side of the discontinuity. Columns 3 and 4 report the two-stage least-squares (IV) RD treatment effects (by using treatment as an instrument for insurance coverage). Columns 1, 2 and 4 control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Column 3 does not include control variable. Pap test specification only includes sample of female respondents. Breast cancer exam and Pap test specifications only include sample of female respondents. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.9: Insurance Coverage on Preventive Care Usages using Static Bandwidth

Outcomes	OLS	RF	IV	IV
	(1)	(2)	(3)	(4)
Blood Pressure Check	0.290***	-0.0108	0.275	0.291
	(0.007)	(0.015)	(0.347)	(0.372)
Bandwidth	78	78	78	78
N	32,829	32,829	32,882	32,829
Routine Check	0.254***	-0.0131	0.315	0.352
	(0.007)	(0.015)	(0.380)	(0.403)
Bandwidth	78	78	78	78
N	32,829	32,829	32,882	32,829
Flu shot	0.134***	0.0172	-0.401	-0.464
	(0.006)	(0.017)	(0.480)	(0.521)
Bandwidth	78	78	78	78
N	32,829	32,829	32,882	32,829
Breast Cancer Check	0.192***	0.024	-1.752	-2.235
	(0.012)	(0.030)	(4.502)	(6.933)
Bandwidth	78	78	78	78
N	9,908	9,908	9,921	9,908
Pap Test	0.164***	0.00753	-0.392	-0.702
	(0.013)	(0.025)	(2.024)	(3.178)
Bandwidth	78	78	78	78
N	9,908	9,908	9,921	9,908

Notes: Table reports local RD regressions with quadratic polynomials in months old using the static bandwidth. Column 1 reports the OLS results using insurance coverage as the independent variable for an optimal bandwidth $\hat{h}$. Column 2 reports the reduced-form (RF) RD treatment effects of individuals age below 26 with a quadratic control function in the months old on each side of the discontinuity. Columns 3 and 4 report the two-stage least-squares (IV) RD treatment effects (by using treatment as an instrument for insurance coverage). Columns 1, 2 and 4 control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Column 3 does not include control variable.

Table 3.10: Insurance Coverage Effects on Preventive Care Usages by Sex

Outcomes	Female (1)	Male (2)
Blood Pressure Check	-0.182	0.347
	(1.432)	(0.514)
Bandwidth	53	53
N	11,769	10,860
Routine Check	1.772	0.448
	(2.489)	(0.512)
Bandwidth	51	51
N	11,347	10,444
Flu shot	-0.255	-0.746
	(1.183)	(0.702)
Bandwidth	61	61
N	13,444	12,507

Notes: Table reports insurance coverage effects on preventive care usages by sex. Local RD regressions with quadratic polynomials in months old is usded using the optimal bandwidth $\hat{h}$ estimated by the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. All specifications control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.11: Insurance Coverage Effects on Preventive Care Usages by Race

	White	Black	Others
Outcomes	(1)	(2)	(3)
Blood Pressure Check	0.691	-7.726	6.666
	(0.937)	(7.349)	(15.770)
Bandwidth	53	53	
N	15,134	28,190	15,010
Routine Check	1.274	-6.684	2.872
	(1.348)	(6.409)	(7.022)
Bandwidth	51	51	
N	14,561	28,190	15,010
Flu shot	-0.766	-3.352	3.744
	(1.192)	(3.246)	(8.971)
Bandwidth	61	61	
N	17,319	28,190	15,010
Breast Cancer Check	0.0995	-2.681	0.674
	(3.419)	(2.566)	(1.259)
Bandwidth	79	79	79
N	6,577	9,627	359
Pap Test	2.021	(1.893)	0.642
	-6.325	-1.809	-0.952
Bandwidth	80	80	80
N	6,577	9,627	359

Notes: Table reports insurance coverage effects on preventive care usages by race. Local RD regressions with quadratic polynomials in months old is usded using the optimal bandwidth $\hat{h}$ estimated by the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. Breast cancer exam and Pap test specifications only include sample of female respondents. All specifications control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.12: INSURANCE COVERAGE EFFECTS ON PREVENTIVE CARE USAGES BY EDUCATION

	College Degree	Without College Degree
Outcomes	(1)	(2)
Blood Pressure Check	0.0244	79.81
	(0.477)	(3804)
Bandwidth	53	53
N	19,969	2,660
Routine Check	0.802	20.16
	(0.582)	(395.0)
Bandwidth	51	51
N	19,202	2,589
Flu shot	-0.312	18.49
	(0.504)	(340.3)
Bandwidth	61	61
N	22,995	2,956
Breast Cancer Check	-0.992	0.553
	(2.055)	(1.6)
Bandwidth	79	79
N	8,323	1,860
Pap Test	-0.0709	1.478
	(1.293)	(1.4)
Bandwidth	80	80
N	8,323	1,860

Notes: Table reports insurance coverage effects on preventive care usages by whether the respondent holds at least a college degree. Local RD regressions with quadratic polynomials in months old is usded using the optimal bandwidth $\hat{h}$ estimated by the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. Breast cancer exam and Pap test specifications only include sample of female respondents. All specifications control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.13: RD TREATMENT EFFECTS ON INSURANCE COVERAGE (WITH COVARIATES REPORTED)

	Quadratic RD $\hat{h}$ bandwidth (1)	Quadratic RD $\hat{h}$ bandwidth (2)	Quadratic RD $2\hat{h}$ bandwidth (3)	Linear RD $\hat{h}$ bandwidth (4)
Insurance Coverage	-0.0400**	-0.0454***	-0.0460***	-0.0300**
	(0.016)	(0.017)	(0.012)	(0.014)
Married	0.0585***		0.0507***	0.0448***
	(0.006)		(0.004)	(0.008)
College Degree	0.148***		0.143***	0.147***
	(0.007)		(0.005)	(0.009)
Student Status	0.119***		0.0922***	0.0999***
	(0.007)		(0.006)	(0.010)
Midwest	0.000327		-0.00464	0.000845
	(0.008)		(0.005)	(0.011)
South	-0.129***		-0.128***	-0.109***
	(0.007)		(0.005)	(0.010)
West	-0.0630***		-0.0587***	-0.0527***
	(0.008)		(0.005)	(0.011)
Employment Status	0.00147		-0.00424	-0.00465
	(0.007)		(0.005)	(0.012)
Income	0.00000313***		0.00000288***	0.00000361***
	(0.000)		(0.000)	(0.000)
Bandwidth	78	78	156	41
N	33,295	33,558	58,283	17,540
Year FE	YES	YES	YES	YES

Notes: Columns 1 and 2 report local RD regressions with quadratic polynomials in months old using the optimal bandwidth $\hat{h}$. Column 1 includes control variables, column 2 does not. Column 4 reports the local RD regressions with linear polynomial in months old using the optimal bandwidth $\hat{h}$. Column 3 reports local RD regressions with quadratic polynomials in months old using twice of optimal bandwidth $2^*\hat{h}$. The optimal bandwidth is estimated using the Imbens and Kalyanaraman (2009) Mean Squared Error algorithm. All specifications control for if the person is married, hold at least a college degree, is current a student, is currently employed, is in one of the for regions: Northeast, Midwest, South, West, and the individual income. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.

Table 3.14: Insurance Effects on Preventive Care Usage (with covariates reported)

	Blood Pressure	Routine Check	Flu Shot	Breast Exam	Pap Test
Panel A: OLS Results					
Coverage	0.296***	0.263***	0.140***	0.189***	0.162***
	(0.008)	(0.008)	(0.006)	(0.012)	(0.013)
Married	0.0731***	0.0675***	0.0740***	0.0774***	0.118***
	(0.008)	(0.008)	(0.007)	(0.013)	(0.011)
College Degree	0.0621***	0.0412***	0.0657***	0.0591***	-0.00827
	(0.011)	(0.012)	(0.012)	(0.014)	(0.013)
Student Status	0.0350***	0.0465***	0.0186**	-0.0721***	-0.235***
	(0.010)	(0.012)	(0.009)	(0.015)	(0.019)
Midwest	0.0282**	-0.0619***	0.00818	-0.0174	0.0154
	(0.012)	(0.012)	(0.010)	(0.019)	(0.017)
South	0.00803	-0.0506***	-0.0057	0.0211	0.0292**
	(0.011)	(0.010)	(0.009)	(0.016)	(0.014)
West	-0.0334***	-0.105***	-0.0174**	-0.119***	-0.0871***
	(0.011)	(0.011)	(0.008)	(0.015)	(0.016)
Employment Status	-0.0127	-0.0461***	-0.0186**	0.02	0.0186
	(0.009)	(0.009)	(0.008)	(0.014)	(0.012)
Income	-0.000000291	-0.000000519***	0.000000305*	0.00000110***	0.000000809***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: IV Results					
Coverage	0.287	0.806	-0.484	-1.64	-0.572
	(0.505)	(0.603)	(0.643)	(4.658)	(2.511)
Married	0.0671**	0.0341	0.105***	0.0945	0.0986*
	(0.029)	(0.034)	(0.037)	(0.095)	(0.052)
College Degree	0.0564	-0.0409	0.153	0.254	0.0498
	(0.074)	(0.087)	(0.095)	(0.533)	(0.288)
Student Status	0.0631	0.007	0.106	0.137	-0.0548
	(0.054)	(0.066)	(0.075)	(0.374)	(0.200)
Midwest	0.0286**	-0.0587***	0.00498	-0.0428	0.00773
	(0.013)	(0.014)	(0.012)	(0.079)	(0.042)
South	0.00783	0.0161	-0.0858	-0.19	-0.0528
	(0.064)	(0.074)	(0.084)	(0.541)	(0.292)
West	-0.0338	-0.0697*	-0.0595	-0.222	-0.123
	(0.036)	(0.042)	(0.045)	(0.276)	(0.148)
Employment Status	-0.0119	-0.0466***	-0.015	0.0535	0.0299
	(0.009)	(0.012)	(0.011)	(0.084)	(0.046)
Income	-0.000000395	-0.00000251	0.0000023	0.00000565	0.00000214
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	22,629	21,791	25,951	10,183	10,183
Time FE	YES	YES	YES	YES	YES

Notes: Table presents the insurance effects on preventive care usage with covariates reported. Panel A reports the OLS results; panel B presents the IV results. Breast cancer exam and Pap test specifications only include sample of female respondents. Standard errors are clustered at the month-year of birth. ***, ** and * denote significance at the 1, 5 and 10 percent levels.