The Network of US Airports and its Effects on Employment

CoPS Working Paper No. G-313, December 2020

Nicholas Sheard
Centre of Policy Studies,
Victoria University

ISSN 1 921654 02 3 ISBN 978-1-921654-21-3

The Centre of Policy Studies (CoPS), incorporating the IMPACT project, is a research centre at Victoria University devoted to quantitative analysis of issues relevant to economic policy.
Address: Centre of Policy Studies, Victoria University, PO Box 14428, Melbourne, Victoria, 8001
home page: www.vu.edu.au/CoPS/ email: copsinfo@vu.edu.au Telephone +61 3 9919 1877

About us

Researchers at the Centre of Policy Studies have a 45-year history of continuous achievement in the development, application and dissemination of large-scale economic models. Our models and software are used around the world to analyse a diverse range of economic issues. CoPS' funders include: Australian federal and state government departments; private firms and universities in many parts of the world; central government agencies such as finance and trade ministries in many countries; and international development organisations. The Centre's GEMPACK software, used for solving large economic models, is used at more than 700 sites in over 95 countries.

Citation

Sheard, Nicholas (2020), "The Network of US Airports and its Effects on Employment", Centre of Policy Studies Working Paper No. G-313, Victoria University, December 2020.

The Network of US Airports and its Effects on Employment*

Nicholas Sheard†

December 2020

Abstract

This paper estimates the effects of airport infrastructure on employment and the distribution of the labor force in US metropolitan areas. The analysis is based on models for the air network and for its effects on employment, which are estimated using US data. Air traffic is found to have a positive effect on the population of the local area, with an elasticity of 0.010, so airport improvements induce a reallocation of workers between regions. Air traffic is also found to have a positive effect on employment in the local area with an elasticity of 0.036 and a weakly positive effect on the employment rate in other places within 400 miles. Simulations suggest that for each job created in the local area by an airport expansion, two and a half jobs are created elsewhere in the US due to the changes in the air network and the distribution of employment. Expanding the average airport adds one job in the US for roughly each $78,000 invested. The results further suggest that the US air network is less centralized than would be optimal.

Keywords: Airport, Network, Transportation infrastructure, Urban growth

JEL classification: H54, L93, R11, R42

*The author thanks Jan Brueckner, Wulong Gu, Bjarne Strøm, Maame Esi Woode, and seminar participants at the 2016 CEA Annual Conference and the 2019 ITEA Annual Conference for helpful comments and suggestions.

†Centre of Policy Studies, Victoria University, 300 Flinders Street, Melbourne, VIC 3000, Australia. E-mail: nicholas.sheard@vu.edu.au. Website: https://sites.google.com/site/nicholassheard/.

Contents

1. Introduction	3
2. Data	6
3. Relationship between air traffic and the regional distribution of the labor force and employment	9
4. Relationship between airport infrastructure and traffic on the air network	19
5. Simulation	28
6. Conclusion	34
References	36
Appendix A. Approximating the values of runways and terminals	39
Appendix B. Airline and aircraft categories used for instruments	41
Appendix C. First-stage results for the CBSA-level estimation	44
Appendix D. Elasticities of trip cost to the costs of using individual airports	44

1 Introduction

Aviation plays a major role in the transportation system of the United States, with almost 400 commercial service airports hosting flights that carry around 2.5 million passengers each day. The construction and maintenance of these airports is largely funded by federal, state, and local governments, which raises the question of whether the public spending on them is justified. Recent studies have estimated the economic effects of airports on the local areas they serve (Brueckner, 2003; Sheard, 2014, 2019; Blonigen and Cristea, 2015), which corresponds to the goals of local governments. The current paper estimates the effects of airport infrastructure on employment in the local area as well as other parts of the country. This exercise is more relevant to the goals of the federal government, which provides almost half of the public funding for airports in the United States.[1]

Expanding an airport may affect employment and the labor force in places outside of the local area through two main channels. The first is by affecting traffic elsewhere in the national air network. Airports may complement one another, as every airport is a potential destination for the other airports, or they may be in competition with each other for travelers. As a result, a change in the size of an airport may have positive or negative effects on traffic elsewhere in the network. The second channel is the effect of airport size on the regional allocation of employment. Sheard (2019) found that airport size has a positive effect on local employment, though partly by drawing workers from other areas. The current paper extends that work by studying the effects on employment in other parts of the country.

The analysis in this paper is in two parts, with a separate model and estimation for each. The first part is on the causal effects of air traffic at a given set of airports on the labor force and employment. The second part is on the relationship between airport infrastructure and traffic throughout the air network. The full parameterized model is used for simulations that link changes in local airport infrastructure to changes in the allocation of the labor force and employment throughout the country.

The first part of the analysis concerns how exogenous changes in air traffic affect the allo-

[1] The Federal Aviation Administration currently has an annual budget of around $16 billion (United States Department of Transportation, 2016). Airport spending by state and local governments is currently around $21 billion per year (United States Census Bureau, 2014).

cation of the labor force and employment between regions. The effects of changes in air traffic are identified using instruments analogous to those proposed by Bartik (1991) and adapted to air travel by Sheard (2019). The results for the effects of air traffic on local population and employment are similar to those in Sheard (2019): positive effects on the local population and employment with elasticities of around 0.010 and 0.036, respectively. Further estimates show that changes in air traffic in a given metropolitan area have positive effects on employment rates in other metropolitan areas up to 400 miles away and negative effects on employment in some metropolitan areas further away.

The second part of the analysis concerns how an exogenous change in the infrastructure at a given airport affects the traffic in all parts of the network. This is done by estimating a model for the air network adapted from that used by Allen and Arkolakis (2019) to study the welfare effects of the United States Interstate Highway System. The model accounts for the returns to scale in airline operations, the differences in airport capacities, and the observed heterogeneity in the network.

Once the model has been estimated, simulations are run to estimate the effects of changes to airport infrastructure on traffic and employment throughout the United States. The simulation results show positive effects of the infrastructure at most airports on traffic in other parts of the national air network. Although there is a negative effect of local air traffic on populations elsewhere, due to the migration it induces, it has positive and negative effects on employment rates in some other locations. The overall effect on employment elsewhere is on average two and a half times as large as the effect on employment in the local area.

The simulations are based on rough aggregations and thus the quantitative results should be treated as approximate. That said, the main result is that for the average airport, one job is created for approximately each $78,000 added to its infrastructure. There is wide variation between the airports in how effective their expansion is for creating jobs. The airport with the highest return to investment has one job created for roughly each $9,400 of infrastructure improvements, while for a handful of airports an expansion in infrastructure leads to a decrease in overall employment. There is a positive overall relationship between the size of an airport and the payoff to a given dollar investment in infrastructure, which suggests the air network in the United States is less centralized than would be optimal.

This paper contributes to the recent literature on the economic effects of airports that uses modern econometric techniques to achieve reliable identification. This work was pioneered by Brueckner (2003) and Green (2007), who used instruments to identify the effects of air traffic. Most of the subsequent work has used instruments for identification and data from metropolitan areas in the United States. Green (2007), Blonigen and Cristea (2015), McGraw (2017), and Sheard (2019) found positive effects of airport size on local population and employment for cities and metropolitan areas in the United States. Brueckner (2003) and Sheard (2014, 2019) studied how airports affect local activity in specific sectors in the United States and found the largest positive effects on services but no measurable effects on manufacturing. Lakew and Bilotkach (2018) found that airline delays have negative effects on local services, manufacturing, and leisure and hospitality. Gibbons and Wu (2020) studied the effects of air access in China and found a positive effect on productivity. Campante and Yanagizawa-Drott (2018) studied global air networks and exploited a discontinuity in the operating ranges of aircraft to identify a positive effect of air connections on local economic activity. Although this literature studies a range of outcomes, it focuses mostly on the effects an airport has on its local area. The current paper extends this literature by studying the effects of air traffic on economic activity in other regions.

A related body of literature has studied the effects of other types of transportation infrastructure. Baum-Snow (2007), Michaels (2008), Duranton and Turner (2012), Duranton, Morrow and Turner (2014), and Allen and Arkolakis (2019) estimated the effects of roads and highways on urban development and economic activity in the United States. Duranton (2015) estimated the effects of roads within and between cities on trade in Colombia. The welfare and productivity effects of historical railways were studied in the context of colonial India by Bogart and Chaudhary (2013) and Donaldson (2018) and for the United States in the late 1800s by Donaldson and Hornbeck (2016). Gonzalez-Navarro and Turner (2018) studied the effects of subways on urban population growth. Baum-Snow, Brandt, Henderson, Turner and Zhang (2017) and Protsiv and Sheard (2020) estimated the effects of several modes of transport infrastructure, in China and Norway respectively.

The current paper makes two main contributions to this literature. The first is that it produces estimates of the effects of airport infrastructure on economic activity in other regions, whereas the existing literature focuses on the local effects. Local governments are naturally concerned with the effect of airport size on local employment. However, the interests of the federal government and possibly also state governments are primarily at the level of overall rather than local employment. If an airport improvement increases local employment but does so simply by shifting jobs from other parts of the country, then this may actually be a net loss from the national perspective.

The second main contribution of this paper is that it goes further than existing work in quantifying the effects of airport investments in a way that can be applied to decisions about whether and where to invest. The estimates are given in terms of the effects of given dollar investments in infrastructure as well as the standard elasticities to the level of air traffic. The simulation results also show the level of investment required at each airport to create each additional job.

The rest of this paper is arranged as follows. Section 2 outlines the data used in the analysis. Section 3 presents the estimation of the relationships between air traffic and the spatial distribution of population and employment. Section 4 presents the estimation of the relationship between airport infrastructure and air traffic. Section 5 details the simulations that relate changes in airport infrastructure to the labor force and employment. Section 6 concludes the paper.

2 Data

The estimation is conducted using data on airports, air traffic, population, and employment in the United States (henceforth the "US") for 1991 to 2018. The data are aggregated by airport or Core Based Statistical Area (CBSA), which is the official definition of a metropolitan area maintained by the Office of Management and Budget.[2] Separate datasets are constructed to conduct the two parts of the estimation. The estimation of the effects of air traffic on the labor force and employment uses an annual panel of air traffic, population, and employment data aggregated by CBSA for the period 1991 to 2018 that is henceforth known as the 'CBSA-level panel'. The estimation of the effects of airport infrastructure on traffic throughout the network uses a cross

[2] The CBSAs are defined as sets of counties that each represent a distinct metropolitan area with a dense urban core. This paper uses the December 2009 definitions.

section of airport statistics and the route network aggregated by CBSA in 2015 that is henceforth known as the 'route dataset'. The two samples are aligned so that both include the same sets of airports and CBSAs.

The CBSA-level panel includes data on the population, labor force, employment, GDP, and air traffic, aggregated by CBSA. The population data are from the US Census Bureau, the labor-force data are from the US Bureau of Labor Statistics, the employment data are from the County Business Patterns, and the GDP data are from the FRED database of the Federal Reserve Bank of St Louis. The data for the total traffic at each airport are from the T-100 segment data published by the Bureau of Transportation Statistics (BTS).[3] The panel is defined for the years 1991 to 2018, which is the longest period for which all of the variables are available.

The route dataset includes data on the air traffic between the sample airports, the total traffic levels at the sample airports, the replacement values of runways and terminals at each of the sample airports, and local population levels. The data are aggregated by CBSA, which Brueckner, Lee and Singer (2014) and Sheard (2020) found to be the appropriate geographical extent for most airport markets in the US. The variables are all measured in 2015. The data on air traffic by route are from the DB1B coupon data from the BTS, which are a 10% sample of all domestic tickets by quarter and detail the origin, destination, and connection airports for each trip. The data on air traffic are from the T-100 segment data. The data on runways are from the Airport/Facility Directory for 25 June 2015 to 20 August 2015 published by the Federal Aviation Administration (2015) and the sizes of airport terminals are from OpenStreetMap. The replacement values of runways and terminals are estimated from their physical characteristics using methods detailed in Appendix A. The population data are from the US Census Bureau.

The samples are restricted to CBSAs in the contiguous US (the District of Columbia and all states except Alaska and Hawaii). The criteria for an airport to be included are that it must be within a CBSA in the contiguous US, have had at least 2,500 departing passengers in each year from 1991 to 2018, and be the endpoint of at least one segment in the route sample. The criteria for the inclusion of route segments are that they must link two sample airports and have had

[3]The employment figures in the County Business Patterns are measured in the week including the 12th of March and the population in the Census is measured at the beginning of April. To match the timing of the air-travel data to that of the employment and population data, the air-travel data are aggregated by 12-month periods from April 1st to March 31st. Thus the air-travel figures for '1991' are the traffic from April 1st, 1990 to March 31st, 1991.

at least 1,000 passengers in 2015 in the T-100 data. The routes are the combinations of flight segments used by at least one passenger in the DB1B coupon data for either direct flights or flights with one connection.[4]

These restrictions result in a sample of 196 airports in 179 CBSAs. The route dataset includes 65,242 routes that are combinations of 2,305 segments. The sample CBSAs represent 70% of the population, 75% of employment, 92% of the departing flights, and 99% of the departing passengers in the contiguous US from 1991 to 2018.[5]

Table 1 presents summary statistics for the CBSA-level panel. The CBSAs are shaded on the map in Figure 1.

	Mean	Std. dev.	Minimum	Maximum
Population	1,132,408	2,055,054	15,239	19,341,770
Labor force	573,770	1,029,894	8,559	9,768,602
Number of employees	472,408	868,136	5,520	8,607,832
Employment rate	0.40	0.07	0.14	0.73
Mean wage ($'000)	33.72	10.32	13.83	130.28
Gross domestic product (GDP) ($'b)	54.16	125.22	0.20	1,811.20
Number of firms	28,762	55,244	794	566,050
Number of airports	1.10	0.45	1	5
Number of departing flights	43,679	86,764	119	629,481
Number of seats on departing flights	5,068,009	11,005,751	7,000	82,871,004
Number of departing passengers	3,606,728	8,079,667	3,022	67,976,306

Note: 5,012 observations of each variable, in a balanced panel of 179 CBSAs

Table 1: Summary statistics for the main variables in the CBSA-level panel for 1991 to 2018.

Summary statistics for the route dataset are shown in Tables 2 and 3. Table 2 summarizes information on air travel aggregated by route and Table 3 summarizes information on air travel, population, and employment aggregated by CBSA. Again these data are all for 2015. The sample routes, airports, and CBSAs are illustrated on the map in Figure 1.

	Mean	Std. dev.	Minimum	Maximum
Distance (miles)	1,564.6	800.2	29.5	5,288.0
Travel time (hours)	3.13	1.60	0.06	10.58
Number of connections	0.973	0.162	0	1
Number of passengers	650.4	7,523.1	1	818,487

Note: 65,242 observations of each variable

Table 2: Summary statistics for the air routes in the route dataset.

[4]Though some trips involve more than one connection, this only represents 1.7% of the tickets in the DB1B data. The sheer number of possible combinations of three segments makes it infeasible to run the estimation on routes with more than one connection.

[5]As the sample does not cover the entire US, the estimates for 'national' employment exclude largely rural areas with almost a third of the country's population. It is possible that this may bias the results in either direction, but there is no practical way of accounting for these areas in the analysis.

	Mean	Std. dev.	Minimum	Maximum
Total runway length (feet)	22,817.3	17,545.1	6,300	117,392
Total runway area (square feet)	3,413,767	2,833,815	673,000	18,516,225
Total runway value ($'m)	1,969.73	1,782.21	384.96	10,866.60
Terminal area (square feet)	598,695	1,038,134	21,805	7,371,596
Terminal volume ('000 cubic feet)	53,086	123,451	240	808,197
Terminal value ($'m)	530.87	1,234.51	2.40	8,081.97
Total airport value ($'m)	2,500.60	2,942.41	388.15	18,719.29
Number of departing passengers	4,284,328	9,908,748	4,282	62,459,460
Population	1,266,728	2,241,963	33,700	19,542,840

Note: 179 observations of each variable

Table 3: Summary statistics at the CBSA level for the route dataset.

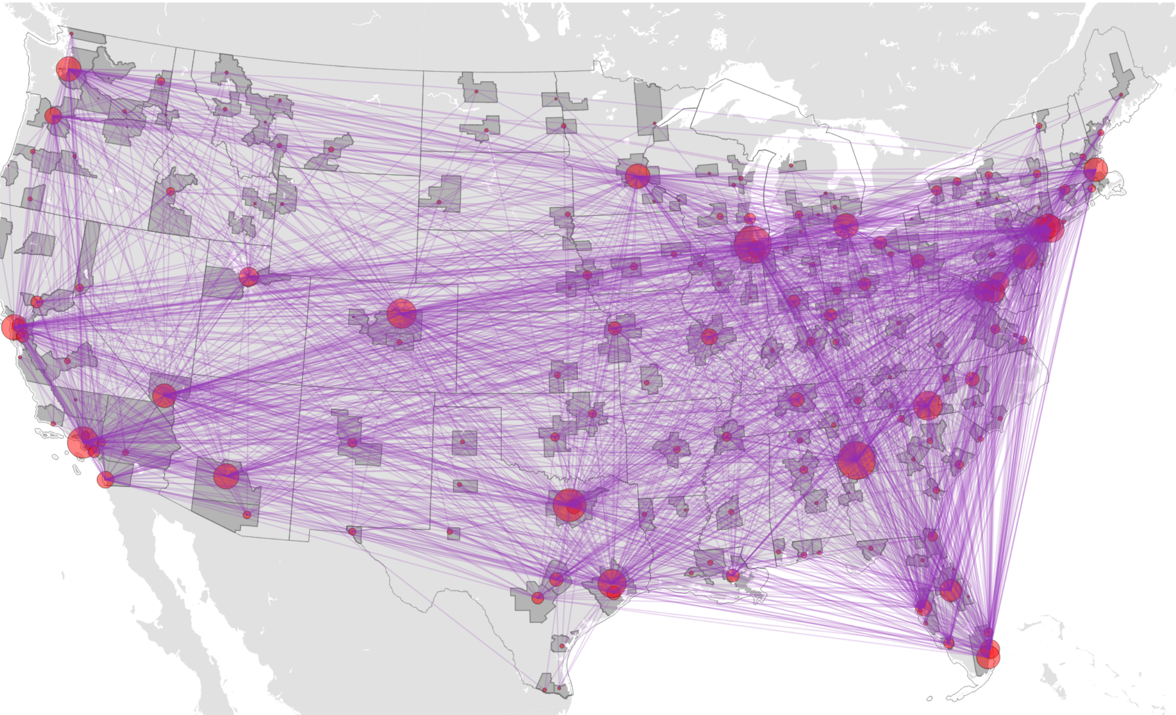

Figure 1: Map of the CBSAs, airports, and the network of air routes in the two datasets. The shaded areas of land are the CBSAs. The lines represent the route segments in the route dataset, which are exclusively domestic routes due to the limitations of the DB1B data. The circles represent the sample airports and the radius of each is proportional to the number of flights in 2015.

3 Relationship between air traffic and the regional distribution of the labor force and employment

This section presents the estimation of the relationships between airport traffic levels and the distribution of the labor force and employment. The amount of air traffic is intended to represent the opportunities for travel by air, the idea being that an airport with more flights provides more convenient schedules to a wider range of destinations. A higher level of traffic may also mean lower ticket prices due to tougher competition.

The main purpose of the estimation is to understand how an exogenous change in traffic at an airport affects the population and employment throughout the US. To that end, this section estimates how the changes in air traffic affect population and employment, both in the local area and in other metropolitan areas in the US. The following section estimates another essential component: the effect of local air traffic on traffic in other parts of the network.

The estimation is given structure by a simple theoretical model that relates air traffic to population and employment. As the labor force and population are likely to be simultaneously determined with air traffic, the estimation uses instruments for the growth in air traffic.

The availability of air connections from a metropolitan area may affect local employment through either the productivity of local firms or the amenity value of the opportunities for travel. Either channel would affect the quality of life of local residents and thereby induce migration. In addition, an airport with more flight operations may be beneficial to a degree for firms and individuals in places outside of the local metropolitan area. To estimate how a change in air traffic affects aggregate employment, it is therefore necessary to understand how the labor force and employment adjust between regions due to migration, as well as estimating the effects on local employment.

This part of the analysis is done by estimating the effects of a change in local air traffic on the local population and employment, then estimating how it affects population and employment in other places. The theory and estimation for each of these are presented in separate subsections below.

3.1 Model for effects of air traffic on the local labor force and employment

The estimation for the local effects is done by fitting the following type of relationship between the growth in local air traffic $A_{m,t}$ and the growth in the outcome variable $Q_{m,t}$ for CBSA m at time t:[6]

$$\frac{Q_{m,t+1}}{Q_{m,t}} = e^{\phi_{2,m}+\chi_{2,t}} A_{m,t}^{\alpha_2} Q_{m,t}^{\beta_2} \left(\frac{A_{m,t+1}}{A_{m,t}} \right)^{\eta_2} \tag{1}$$

[6]A full derivation of this type of equation from a model of growth and convergence to the natural level of population or employment is presented in Sheard (2019).

The outcome variable can be population, the labor force, employment, or GDP in the local area. The terms $\phi_{2,m}$, $\chi_{2,t}$, α_2, β_2, and η_2 are parameters that are fitted in the estimation. The basic idea underlying (1) is that the rate of change in say population or employment can be influenced by the current levels of that variable and air traffic – as the population or employment may be converging to a 'natural' level that those variables partly determine – and by changes in the level of air traffic. It may also be influenced by fixed factors for CBSA m or year t, which are represented by $\phi_{2,m}$ and $\chi_{2,t}$, respectively.

The size of the parameter η_2 indicates how much local population or employment will change when there is a given change in the traffic at the local airports. For example, if the estimated η_2 for population is positive, then there is migration to the local area when the local airports are expanded.

The main difficulty with estimating η_2 is that changes in air traffic in m may be simultaneously determined with changes in population or employment. That is, air traffic $A_{m,t}$ could feasibly be influenced by $Q_{m,t}$, or $A_{m,t}$ and $Q_{m,t}$ could be influenced by some common but unobserved factor. To address this issue, instruments are used for the changes in air traffic, the derivation of which is described below. The instrument for the change in air traffic in CBSA m between periods t and $t+1$ is denoted $\frac{\hat{A}^L_{m,t+1}}{A_{m,t}}$, where the L indicates that these are the instruments used in the estimation for the local area. The instruments are assumed to affect $\frac{A_{m,t+1}}{A_{m,t}}$ according to:

$$\frac{A_{m,t+1}}{A_{m,t}} = e^{\phi_{1,m}+\chi_{1,t}} A^{\alpha_1}_{m,t} Q^{\beta_1}_{m,t} \left(\frac{\hat{A}^L_{m,t+1}}{A_{m,t}} \right)^{\zeta_1} \tag{2}$$

The logic behind the inclusion of the level terms $A_{m,t}$ and $Q_{m,t}$ and the fixed CBSA and time factors $\phi_{1,m}$ and $\chi_{1,t}$ in (2) is similar to that for (1): air traffic may be converging to its natural level and there may be factors specific to the CBSA or year that apply across years or CBSAs.

The subsequent algebra is simplified by the notation $a \equiv \ln(A)$ and $q \equiv \ln(Q)$. Taking logs of both sides of (2) and adding the error term $\varepsilon_{1,m,t}$ yields the first-stage equation, which determines the instrumented values of $[a_{m,t+1} - a_{m,t}]$:

$$a_{m,t+1} - a_{m,t} = \alpha_1 a_{m,t} + \beta_1 q_{m,t} + \zeta_1 \left[\hat{a}^L_{m,t+1} - a_{m,t} \right] + \phi_{1,m} + \chi_{1,t} + \varepsilon_{1,m,t} \tag{3}$$

The second-stage equation is derived by adding the error term $\varepsilon_{2,m,t}$ to the log of (1):

$$q_{m,t+1} - q_{m,t} = \alpha_2 a_{m,t} + \beta_2 q_{m,t} + \eta_2 \left[a_{m,t+1} - a_{m,t}\right] + \phi_{2,m} + \chi_{2,t} + \varepsilon_{2,m,t} \tag{4}$$

The system of equations (3) and (4) is identified if the relevance and exogeneity conditions hold. The relevance condition requires at least one of the instruments in the first-stage regression to be significant, so the estimated $\zeta_1 \neq 0$. It is reasonable to expect the relevance condition to be satisfied, as the instruments capture overall changes in air travel that determine local traffic. It is confirmed statistically in the empirical section.

The exogeneity condition requires the instruments not to be correlated with the error term $\varepsilon_{2,m,t}$, given the controls, so $Cov\left(\hat{a}^L_{m,t+1} - a_{m,t}, \varepsilon_{2,m,t}\right) = 0$. It is plausible that the exogeneity condition is satisfied, as the instruments are driven by changes in overall air traffic and by design they exclude idiosyncratic changes in the local area. While it is not possible to test the exogeneity condition explicitly, a partial test is provided by the overidentification tests, which are shown to pass in the empirical section.

3.2 Model for effects of air traffic on the labor force and employment elsewhere in the US

The estimation of how local air traffic affects the labor force and employment in other parts of the US is conducted using a simple extension of the model for the local area. The ideas underlying the framework are similar to those presented above for the local area and thus a reasonably simple model could be written to ground the equations, but to avoid repetition the exposition here is limited to a sketch of the key ideas.

The goal of this step in the estimation is to understand how a change in air traffic in one CBSA affects the population and employment in all CBSAs. The estimates of how a change in air traffic affects the local population imply the effect on total net migration to or from the local area. The estimation described in this subsection first quantifies the changes in population by distance from the local area, which sheds light on the geographical patterns of the induced migration. It then quantifies the effects on employment rates in other places, to show how employment is affected given the changes in population.

Formally, the estimation is run by comparing the population and employment rate in a given CBSA m with the air traffic in CBSA m and in other places within given ranges of crow distances from CBSA m. The distances are measured between the central cores of the CBSAs. The data on the population and employment in other places is limited to the sample CBSAs so that the analysis is balanced and consistent, though most US air traffic is at airports in these CBSAs in any case. The subscript $b \in B$ is used to represent the set of all CBSAs within a given distance band from CBSA m. For the variables it indicates an aggregate of all CBSAs within the set, so for example $A_{m,b,t}$ represents the total air traffic in the distance band b from CBSA m.

The estimation equations differ from the theory for the local effects represented by (3) and (4) mainly by the inclusion of variables for air traffic in CBSAs by distance band and the corresponding instruments. The superscript D indicates that the instruments are designed for the estimation by distance band, with the L and O indicating that the instruments apply to the growth in traffic in the 'local' and 'other' areas. The calculation of the instruments is explained below. The estimation equations for the first stage are:

$$\begin{aligned} a_{m,t+1} - a_{m,t} = \;& \alpha_1 a_{m,t} + \sum_{\tilde{b} \in B} \beta_{1,\tilde{b}} a_{m,\tilde{b},t} + \gamma_1 q_{m,t} + \zeta_1 \left[\hat{a}^{D,L}_{m,t+1} - a_{m,t} \right] \\ & + \sum_{\tilde{b} \in B} \theta_{1,\tilde{b}} \left[\hat{a}^{D,O}_{m,\tilde{b},t+1} - a_{m,\tilde{b},t} \right] + \phi_{1,m} + \chi_{1,t} + \varepsilon_{1,m,t} \end{aligned} \quad (5)$$

$$\begin{aligned} a_{m,b,t+1} - a_{m,b,t} = \;& \alpha_2 a_{m,t} + \sum_{\tilde{b} \in B} \beta_{2,\tilde{b}} a_{m,\tilde{b},t} + \gamma_2 q_{m,t} + \zeta_2 \left[\hat{a}^{D,L}_{m,t+1} - a_{m,t} \right] \\ & + \sum_{\tilde{b} \in B} \theta_{2,\tilde{b}} \left[\hat{a}^{D,O}_{m,\tilde{b},t+1} - a_{m,\tilde{b},t} \right] + \phi_{2,m,b} + \chi_{2,t} + \varepsilon_{2,m,b,t} \end{aligned} \quad (6)$$

And the estimation equation for the second stage is:

$$\begin{aligned} q_{m,t+1} - q_{m,t} = \;& \alpha_3 a_{m,t} + \sum_{\tilde{b} \in B} \beta_{3,\tilde{b}} a_{m,\tilde{b},t} + \gamma_3 q_{m,t} + \eta_3 \left[a_{m,t+1} - a_{m,t} \right] \\ & + \sum_{\tilde{b} \in B} \iota_{3,\tilde{b}} \left[a_{m,\tilde{b},t+1} - a_{m,\tilde{b},t} \right] + \phi_{3,m} + \chi_{3,t} + \varepsilon_{3,m,t} \end{aligned} \quad (7)$$

As with the estimation for the local area, it is important that the relevance and exogeneity conditions hold for the system (5), (6), and (7). The relevance condition is tested statistically. The instruments are designed so it is plausible that the exogeneity condition holds. In addition, overidentification tests are run as a partial test of the exogeneity condition and are shown generally to be satisfied.

3.3 Instruments for changes in air traffic

The instruments for the changes in air traffic by CBSA are calculated using the technique proposed by Bartik (1991) to predict changes in local employment and adapted by Sheard (2019) to predict changes in airport size. The instruments are calculated by dividing up air traffic using a set of categories, then applying the changes in aggregate traffic for each category to the initial levels of traffic for those categories at each airport. The sets of categories used are the airlines and the types of aircraft, lists of which are given in Appendix B.[7]

The result is variables that describe the hypothetical increase in traffic at each airport that is driven entirely by overall changes in the operation of the national air network. For example, if more units of a certain model of aircraft are delivered then they may use airports anywhere in the country but will tend to operate at airports with the capacity to handle them. Many such changes are occurring at any point in time – to both the airlines and the types of aircraft – and the instruments combine their effects on local growth without the sources of those effects needing to be identified separately by the researcher.

To avoid the instrument values being driven by changes in the labor force and employment, air traffic in the relevant CBSAs is excluded from the calculation of the overall growth rate for each category. Thus the instrument for the local effects of air traffic in CBSA m applied in (3) excludes traffic that departs or arrives at the airports in CBSA m. Using $A_{c,m,t}$ to denote the traffic in category c at airports in CBSA m at time t, the instrument for the growth in air traffic in m, excluding traffic at the airports in m, is calculated using:

$$\hat{A}^{L}_{m,t+1} = \sum_{c} A_{c,m,t} \left(\frac{\sum_{n \neq m} A_{c,n,t+1}}{\sum_{n \neq m} A_{c,n,t}} \right) \tag{8}$$

The instruments for air traffic in the estimation of the effects of air traffic in other places that

[7] Sheard (2019) used these same two sets of categories along with a measure of the traffic by distance range. The distance-range instruments are not used in this paper as the variation in traffic by distance is part of what is estimated in the model and thus the exogeneity of the instruments may be put in question.

feature in (5) and (6) use a broader range of exclusions. The two types of instruments $\hat{A}^{D,L}_{m,t+1}$ and $\hat{A}^{D,O}_{m,b,t+1}$, which instrument for traffic in the 'local' area and 'other' places, both exclude traffic arriving or departing at airports in CBSA m and at airports in CBSAs in the relevant distance band b when calculating the overall growth rate for each category. The instruments are calculated as follows:

$$\hat{A}^{D,L}_{m,t+1} = \sum_c A_{c,m,t} \left(\frac{\sum\limits_{o \neq \{m,b\}} A_{c,o,t+1}}{\sum\limits_{o \neq \{m,b\}} A_{c,o,t}} \right) \tag{9}$$

$$\hat{A}^{D,O}_{m,b,t+1} = \sum_c A_{c,m,b,t} \left(\frac{\sum\limits_{o \neq \{m,b\}} A_{c,o,t+1}}{\sum\limits_{o \neq \{m,b\}} A_{c,o,t}} \right) \tag{10}$$

3.4 Estimated effects of air traffic on the local labor force and employment

Table 4 presents estimation results for the effects of a change in CBSA-level air traffic on the local population, labor force, employment, and GDP. These results are obtained by estimating the system (3) and (4). The first-stage results for the TSLS estimation are shown in Appendix C.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Population		Labor force		Employment		GDP	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
$\ln(A_{m,t+1}) - \ln(A_{m,t})$	0.005^a	0.010^a	0.007^a	0.012^b	0.014^a	0.036^a	0.017^a	0.026^b
	(0.002)	(0.003)	(0.002)	(0.005)	(0.003)	(0.007)	(0.004)	(0.011)
$\ln(A_{m,t})$	0.002^b	0.003^a	0.003^b	0.004^a	-0.002^c	0.002	0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)
$\ln(pop_{m,t})$	-0.025^a	-0.026^a						
	(0.005)	(0.004)						
$\ln(lf_{m,t})$			-0.049^a	-0.050^a				
			(0.006)	(0.006)				
$\ln(emp_{m,t})$					-0.060^a	-0.065^a		
					(0.006)	(0.006)		
$\ln(gdp_{m,t})$							-0.080^a	-0.082^a
							(0.010)	(0.010)
R^2	0.61		0.36		0.49		0.26	
First-stage statistic		64.86		66.72		68.25		66.89
Overid. p-value		0.19		0.13		0.98		0.98
Hausman test p-value		0.17		0.49		0.00		0.43

Note: the dependent variable in each regression is the change in the log population, labor force, employment, or GDP in CBSA m between t and $t+1$; 4,833 observations for each regression, representing 179 CBSAs; robust standard errors clustered by CBSA in parentheses; a, b, c denote significance at 1%, 5%, 10%; number of departing flights used as the measure of airport size; all regressions include CBSA and year fixed effects

Table 4: Estimation of the relationships between air traffic and the population, labor force, employment, and GDP in the same CBSA.

The results in Table 4 show positive relationships between the traffic at the airports in a CBSA and the population, labor force, employment, and GDP in the local area. The OLS results indicate positive correlations between air traffic and the levels of the local outcome variables, while the TSLS results indicate positive causal effects of air traffic. As the effects on the population and labor force are positive, an increase in the size of an airport causes net migration to the local area. To understand how air traffic affects overall employment, it is therefore necessary to understand how population and employment adjust across space.

The coefficients in Table 4 are elasticities. The magnitudes of the effects on employment and GDP are clearly greater than those for population and the labor force. However, as the levels of the outcome variables differ, it is not obvious from the elasticities how the absolute sizes of the effects compare. Table 5 shows the absolute effects of a 10% increase in the traffic at an airport from the coefficients in Table 4 and the mean levels of the variables from Table 1.

		Quantity		Proportion of
	Coefficient	Mean	Change	emp. change
Population	0.010	1,132,408	1,155	68.8%
Labor force	0.012	573,770	701	41.7%
Employment	0.036	472,408	1,680	

Note: the mean quantities are over all CBSAs in the sample and all years 1991 to 2018; the changes in quantity are the shifts implied by a 10% increase in air traffic

Table 5: Effects of a 10% change in air traffic on local population, labor force, and employment in absolute terms.

The numbers in Table 5 indicate that for every job created in the CBSA by an increase in local air traffic, the population increases by 0.69 residents and the labor force increases by 0.42. From the change in population it may appear that a majority of the new jobs are taken by in-migrants. However, the effect on the size of the local labor force is around two-thirds as large as the effect on the population, so roughly a third of the net migration is by individuals who are not part of the labor force. Those migrants could be family members who accompany workers, individuals who move because of the amenity value of the airport, or people attracted by better future job opportunities.

Table 4 also shows a larger elasticity for the effect on employment than GDP. This could be due to the sluggishness of capital adjustment relative to labor, for example with office space and some types of machinery, which leads firms to increase their share of labor when travel opportunities improve.

3.5 Estimated effects of air traffic on the labor force and employment elsewhere in the US

The next step is to estimate how the labor force and employment elsewhere in the US are affected when there is a change in air traffic in a given CBSA. This is done by estimating (5), (6), and (7) using growth in population and the employment rate as the outcome variables. For comparison, the estimation is also conducted using employment growth as the outcome. The estimation uses 200-mile distance bands from 0–200 miles to 1,000–1,200 miles. The results of the estimation are presented in Table 6.

The estimation with population as the outcome is used to show the pattern of internal US migration induced by changes in air traffic. For example, a negative effect of air traffic at a given distance on the population of CBSA m would suggest that migration is induced from places at those distances when an airport is expanded. The effects of air traffic elsewhere on employment in CBSA m are inferred from coefficients for the employment rate, given the changes to the population.

As international migration is influenced by many unobserved factors, of which air traffic is likely to play a minor role, it would be difficult to estimate the effect of US air traffic on the total US population from the available information. Furthermore, any changes in the total population would be captured by constants in the estimation. Therefore, for simplicity it is assumed that the total population of the sample CBSAs is held fixed, so the population is simply reallocated between them when air traffic changes.

The results for population growth in Table 6 show no clear pattern by distance from CBSA m. All OLS coefficients for traffic at the 'other' airports by distance band are close to zero and not significant. The TSLS coefficients are somewhat larger in magnitude, but are also not significant for any distance band. The OLS and TSLS coefficients on the air traffic in CBSA m in the regressions for population are similar to those shown in Table 4 for the regressions with no 'other' airports in the sample. This is reassuring as a large difference could indicate a bias from the air traffic levels in CBSA m and the distance bands being correlated.

	(1)	(2)	(3)	(4)	(5)	(6)
	Population		Employment		Emp. rate	
	OLS	TSLS	OLS	TSLS	OLS	TSLS
$\ln(A_{m,t+1}) - \ln(A_{m,t})$	0.005^b	0.012^a	0.012^a	0.020^a	0.006^b	0.006
	(0.002)	(0.003)	(0.003)	(0.008)	(0.002)	(0.007)
$\ln(A_{m,\{0-200\},t+1}) - \ln(A_{m,\{0-200\},t})$	0.002	−0.004	0.006	0.024	0.005	0.021
	(0.003)	(0.008)	(0.007)	(0.022)	(0.007)	(0.020)
$\ln(A_{m,\{200-400\},t+1}) - \ln(A_{m,\{200-400\},t})$	−0.000	−0.005	0.016^c	0.048	0.018^b	0.059^c
	(0.004)	(0.014)	(0.009)	(0.037)	(0.008)	(0.032)
$\ln(A_{m,\{400-600\},t+1}) - \ln(A_{m,\{400-600\},t})$	0.004	0.012	0.030^a	0.076	0.024^a	0.046
	(0.004)	(0.018)	(0.010)	(0.046)	(0.009)	(0.045)
$\ln(A_{m,\{600-800\},t+1}) - \ln(A_{m,\{600-800\},t})$	−0.001	0.009	0.002	-0.084^c	−0.003	-0.111^a
	(0.005)	(0.014)	(0.012)	(0.043)	(0.010)	(0.041)
$\ln(A_{m,\{800-1,000\},t+1}) - \ln(A_{m,\{800-1,000\},t})$	−0.005	−0.028	−0.018	0.019	−0.009	0.089
	(0.004)	(0.023)	(0.012)	(0.057)	(0.011)	(0.055)
$\ln(A_{m,\{1,000-1,200\},t+1}) - \ln(A_{m,\{1,000-1,200\},t})$	−0.001	0.027	0.002	−0.005	0.005	−0.003
	(0.003)	(0.020)	(0.011)	(0.057)	(0.010)	(0.051)
R^2	0.61		0.50		0.48	
First-stage statistic		8.25		8.24		8.07
Overid. p-value		0.10		0.80		0.82
Hausman test p-value		0.20		0.02		0.02

Note: the dependent variable in each regression is the change in the specified variable in CBSA m between t and $t+1$; 4,833 observations for each regression, representing 179 CBSAs; robust standard errors clustered by CBSA m in parentheses; a, b, c denote significance at 1%, 5%, 10%; number of departing flights used as the measure of airport size; all regressions include CBSA m and year fixed effects and controls for initial air traffic and the specified variable in CBSA m and air traffic in all distance bands that are not displayed in the interests of space

Table 6: Estimation of the relationships between air traffic in CBSA m and in various distance bands from m and the population, employment, and employment rate in CBSA m.

The estimates for employment in Table 6 show some significant relationships between employment in CBSA m and traffic at airports in 'other' places. The OLS coefficients for 'other' places are positive and significant for the distance bands between 200 and 600 miles for both employment and the employment rate, indicating a positive correlation between air traffic and employment across those distances.

The TSLS coefficients for employment and the employment rate indicate a positive effect of air traffic between 200 and 400 miles away and a negative effect of air traffic between 600 and 800 miles away. However, there are two caveats: (1) the TSLS coefficient for air traffic in CBSA m is much lower than when no 'other' airports were included and (2) there is substantial variation between the coefficients for the different distance bands. The first caveat suggests that there may be a degree of multicollinearity and the positive coefficients for 'other' places may in fact be partly attributable to the local area. Regarding the second caveat, the negative coefficient for the distance band for 600–800 miles is largely counteracted in magnitude by the coefficient for 800–1,000 miles, so it is plausible that it is spurious.

The first-stage statistics displayed at the bottom of Table 6 show that the instruments are somewhat weak. The overidentification test barely passes for the population levels and comfortably passes for the employment and employment rate.

4 Relationship between airport infrastructure and traffic on the air network

This section presents the second part of the theory and estimation, which relates physical airport infrastructure to air traffic on the network. The theoretical model for the air network shows how traffic on the network is determined by the airport infrastructure. The estimation is conducted using 2015 data on air traffic and the physical infrastructure at the airports.

4.1 Model of the air network

An improvement to the physical infrastructure at an airport may affect traffic throughout the network, as it facilitates flights to and from other places and, due to competition, it may reduce traffic on alternative routes. These phenomena are represented here with a model for the traffic on each route as a function of the travel time and the costs of using each of the airports on the route. The model is based on that proposed by Allen and Arkolakis (2019) to study road networks, which is adapted to air travel by defining the cost of travel as a function of the traffic and infrastructure at the nodes of the network rather than the links.

Individuals in the model make decisions about whether and where to travel by air based on the utility they gain from travel and the cost of travel. This usually involves a choice of route, as most destinations can be reached by flights that involve connections at different airports. For simplicity it is assumed that the costs of travel are borne entirely by the traveler.[8]

The estimation proceeds in three steps. The first is to estimate the parameters for the choice of route, given the choice of origin and destination CBSAs, which yields estimates of the cost of airborne travel time and of using each airport for a connection. The second is to estimate the parameters that determine how many trips are made and to which destinations, which yields

[8] Air travel is subject to taxes and subsidies, so the price paid for a ticket differs from the actual cost of the trip. However, lacking detailed data on the taxes and subsidies it is not feasible to adjust for them.

estimates of the cost of using each airport for the origin or destination of a trip. The third step is to estimate the contributions of the traffic and infrastructure levels to the costs of using each airport.

The first step is as follows. The cost of travel on a given route is assumed to be a function of the airborne travel time, the costs of using the origin, connection, and destination airports, and an idiosyncratic factor that is specific to each individual traveler and route. The CBSAs are indexed by $i \in I$. A route may be either a direct flight or an indirect flight via any number of connection airports. Formally, route r is defined to be a set of flight segments between CBSAs i and j via connections at $C(r) \geq 0$ airports and has airborne travel time $t_{i,j}(r)$. The direct airborne travel time between i and j is denoted $t_{i,j}$, so the airborne travel time for an indirect route is $t_{i,j}(r) = t_{i,k_1(r)} + \ldots + t_{k_{C(r)}(r),j}$.

The cost of using an airport in CBSA i for the origin or destination of a trip is defined to be μ_i while the cost of using it for a connection is ν_i. The combined costs of airborne travel time and of using the origin, connection, and destination airports for a trip between i and j on route r is then $\tilde{\tau}_{i,j}(r)$, which has the following functional form:[9]

$$\tilde{\tau}_{i,j}(r) = t_{i,j}(r)\,\mu_i \left[\prod_{c=1}^{C(r)} \nu_{k_c(r)}\right] \mu_j \tag{11}$$

Note that the costs of airborne time sum together whereas the costs of using the origin and destination airports enter (11) multiplicatively, where the latter assumption is made to keep the model tractable.[10] The idiosyncratic component of the cost of traveling between CBSAs i and j on route r for individual ς is denoted $\xi_{i,j}(r,\varsigma)$. Its values are a set of independent observations from a Fréchet distribution with shape parameter $\psi > 0$, so that $\Pr\left(\xi_{i,j}(r,\varsigma) \leq z\right) = F(z) = e^{-z^{-\psi}}$ and $f(z) = \psi z^{-1-\psi} e^{-z^{-\psi}}$. The cost of travel between i and j on route r for individual ς is defined to be $\tilde{\tau}_{i,j}(r)\,\xi_{i,j}(r,\varsigma)$.

Each individual who travels between i and j chooses the route that has the lowest cost. The cost of the chosen route is thus denoted $\tau_{i,j}(\varsigma) = \min_r \left(\tilde{\tau}_{i,j}(r)\,\xi_{i,j}(r,\varsigma)\right)$. Using the derivations of

[9]For routes with no connections, the term in square brackets in (11) is defined to be equal to one.

[10]Allen and Arkolakis (2019) fit the costs of using each link in the highway network and to facilitate this they have them combine as a multiple. As the model here fits the costs of using the airports (i.e. the nodes) in the network and an overall airborne travel cost, the more intuitively satisfying assumption of the travel cost reflecting the total airborne travel time is able to be made.

Eaton and Kortum (2002) and Allen and Arkolakis (2019), the mean cost of travel between i and j given travelers' preferences for routes is:[11]

$$\tau_{i,j} = \Gamma\left(\frac{\psi - 1}{\psi}\right)\left(\sum_{r=1}^{R} \tilde{\tau}_{i,j}(r)^{-\psi}\right)^{-\frac{1}{\psi}} \tag{12}$$

The same derivations generate the probability $\pi_{i,j}(r)$ of route r having the lowest cost of the R routes from i to j for a given individual:

$$\pi_{i,j}(r) = \frac{\tilde{\tau}_{i,j}(r)^{-\psi}}{\sum_{s=1}^{R} \tilde{\tau}_{i,j}(s)^{-\psi}} \tag{13}$$

The relationship in (13) is used in the first step of the estimation. This yields estimates of the parameters ψ and $\{\nu_i\}_{i\in I}$, which explain the choice of route given the passenger is traveling between i and j. It cannot be used to estimate the parameters $\{\mu_i\}_{i\in I}$ for the origin and destination airports, as these cancel on the right-hand side of (13). The parameters ψ and $\{\nu_i\}_{i\in I}$ are estimated from the following gravity-style expression, which is found by substituting (11) into (13) and adding the error term $\varepsilon_{i,j}(r)$:

$$\begin{aligned} \ln\left(\pi_{i,j}(r)\right) = -\psi \ln\left(t_{i,j}(r)\right) - \psi \sum_{c=1}^{C} \ln\left(\nu_{k_c(r)}\right) \\ - \ln\left(\sum_{s=1}^{R}\left(t_{i,j}(s)\prod_{c=1}^{C(s)} \nu_{k_c(s)}\right)^{-\psi}\right) + \varepsilon_{i,j}(r) \end{aligned} \tag{14}$$

The value of ψ is estimated from (14) as the coefficient on $\ln\left(t_{i,j}(r)\right)$. Given the estimated value of ψ, the values of $\{\nu_i\}_{i\in I}$ are inferred from the fixed effects for the connection airports. The greater the share of passengers who connect at a given airport i, the lower the estimated cost ν_i of using that airport for a connection.

The components of $\tau_{i,j}$ estimated thus far are summarized in the term $\Omega_{i,j}$, which is defined to satisfy $\tau_{i,j} \equiv \mu_i \Omega_{i,j} \mu_j$ and can be derived from (11) and (12):

$$\Omega_{i,j} = \Gamma\left(\frac{\psi - 1}{\psi}\right)\left(\sum_{r=1}^{R} t_{i,j}(r)^{-\psi}\left[\prod_{c=1}^{C} \nu_{k_c(r)}\right]^{-\psi}\right)^{-\frac{1}{\psi}} \tag{15}$$

[11] Appendix D details the elasticities of the trip cost $\tau_{i,j}$ to the costs μ_i and ν_i of using the individual airports.

The second step of the estimation concerns the choices of destination and the frequency of trips. Individuals in all CBSAs are assumed to have identical preferences for travel and budgets to spend on travel. Furthermore, their choices of where to live are treated as being exogenous.[12] The utility that an individual in CBSA i gains from making $a_{i,j}$ trips to each other CBSA j is defined to be:[13]

$$U_i = \left[\sum_{j \neq i} a_{i,j}^{\frac{\rho-1}{\rho}} \omega_j^{\frac{1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \tag{16}$$

The parameter ρ is the elasticity of substitution between destinations and ω_j is a parameter for the overall attractiveness of CBSA j as a destination. The total number of passengers traveling between CBSAs i and j is denoted $A_{i,j}$ and the total number of passengers departing the airports in CBSA i is $A_i = \sum_{j \in I} A_{i,j}$. Each individual is assumed to have a budget of one unit for travel. Given there are L_i individuals in i and the mean cost of travel between i and j is $\tau_{i,j}$, the total number of trips between i and j is:

$$A_{i,j} = \frac{\tau_{i,j}^{-\rho} \omega_j}{\sum_{k \neq i} \tau_{i,k}^{1-\rho} \omega_k} L_i \tag{17}$$

The values of ρ, $\{\mu_i\}_{i \in I}$, and $\{\omega_i\}_{i \in I}$ are estimated from the observed numbers of trips between each origin and destination, given the populations of the CBSAs and the values of $\{\Omega_{i,j}\}_{i,j \in I}$. These estimates rely on the assumption that individuals have identical preferences for travel, as otherwise it is not possible to separate the cost of using an airport from the propensity of local residents to travel. The estimation equation is derived by substituting $\tau_{i,j} \equiv \mu_i \Omega_{i,j} \mu_j$ into (17) and adding the error term $\varepsilon_{i,j}$:

$$\ln\left(\frac{A_{i,j}}{L_i}\right) = -\rho \ln\left(\Omega_{i,j}\right) + \ln\left(\frac{1}{\mu_i \sum_{k \neq i} \mu_k^{1-\rho} \Omega_{i,k}^{1-\rho} \omega_k}\right) + \ln\left(\mu_j^{-\rho} \omega_j\right) + \varepsilon_{i,j} \tag{18}$$

The value of ρ is estimated from (18) as the additive inverse of the coefficient on $\ln\left(\Omega_{i,j}\right)$. The values of $\{\mu_i\}_{i \in I}$ and $\{\omega_i\}_{i \in I}$ are inferred from the fixed effects for the origins and destina-

[12] These assumptions are made as there is not sufficient detail in the data to separate the propensity to travel from the cost of travel.

[13] Work and private trips are not treated separately, so these preferences are assumed to capture all types of trips.

tions of the trips.[14]

The third step of the estimation relates the costs of using the airports in each CBSA for a trip endpoint or connection to the level of traffic A_i and the physical capacity x_i of the airports in CBSA i. The functional forms of the relationships are assumed to be the following:[15]

$$\mu_i = \frac{A_i^{\alpha}}{x_i^{\beta}} e^{\varepsilon_i^{\mu}} \tag{19}$$

$$\nu_i = \frac{A_i^{\gamma}}{x_i^{\delta}} e^{\varepsilon_i^{\nu}} \tag{20}$$

The terms α, β, γ, and δ are parameters while ε_i^{μ} and ε_i^{ν} are idiosyncratic factors that are independently drawn from normal distributions with zero mean. Two main ideas underlie the functional forms in (19) and (20). Firstly, the cost of using an airport of a given capacity may depend either positively or negatively on the level of traffic – so the parameters α and γ may have either sign – as congestion is costly but there may be returns to scale. Secondly, the cost of using an airport with a given level of traffic should be decreasing in its physical capacity due to congestion – so the parameters β and δ should be positive. The parameters for beginning or terminating a trip at an airport are allowed to differ from those for making a flight connection because they involve different activities, for example making a flight connection generally does not involve ground transportation or checking in.

The parameters α, β, γ, and δ are estimated by regressing the estimated values of $\{\mu_i\}_{i\in I}$ and $\{\nu_i\}_{i\in I}$ on the air traffic and airport capacity. The estimation equations are derived directly from the definitions (19) and (20):

$$\ln(\mu_i) = \alpha \ln(A_i) - \beta \ln(x_i) + \varepsilon_i^{\mu} \tag{21}$$

$$\ln(\nu_i) = \gamma \ln(A_i) - \delta \ln(x_i) + \varepsilon_i^{\nu} \tag{22}$$

[14]Both fixed effects have elements of $\{\mu_i\}_{i\in I}$ and $\{\omega_i\}_{i\in I}$. The values of $\{\mu_i\}_{i\in I}$ and $\{\omega_i\}_{i\in I}$ are solved for by iterating on their values until $\ln\left(\frac{1}{\mu_i \sum_{k\neq i} \mu_k^{1-\rho} \Omega_{i,k}^{1-\rho} \omega_k}\right)$ and $\ln\left(\mu_j^{-\rho}\omega_j\right)$ converge to the values of the fixed effects.

[15]The physical capacity x_i is the replacement value of the airport runways and terminals, which is explained in more detail in Appendix A.

As the level of traffic is likely to be influenced by the cost of using the airport, the parameters in (21) and (22) are estimated using an instrumental-variables technique. The instrument used for $\ln(A_i)$ is the log population of the CBSA in 2015. This instrument should be relevant as the level of traffic is driven by demand from the local area. It is credible that the exclusion restriction holds because, once the physical capacity of the airport is controlled for, the CBSA population is only likely to influence the cost of using the airport through its effect on the amount of travel.

4.2 Estimated effects of airport infrastructure on traffic throughout the network

Table 7 presents the results for the estimation of ψ and $\{\nu_i\}_{i\in I}$, which is conducted using (14). The three columns represent regressions with different sets of fixed effects for the connection and endpoint airports, though it is the specification in the third column – both connecting and endpoint airport fixed effects – that is used to derive the parameter estimates.

	(1) Poisson	(2) Poisson	(3) Poisson
$\ln(t_{i,j}(r))$	-1.662^a (0.052)	-0.249^a (0.015)	-0.159^a (0.016)
Connection fixed effects		Y	Y
Endpoint fixed effects			Y
Pseudo R^2	0.21	0.88	0.89

Note: 65,242 observations for each regression; robust standard errors in parentheses; *a*, *b*, *c* denote significance at 1%, 5%, 10%

Table 7: Estimates of effect of travel time on route choice. Columns 2 and 3 include fixed effects for the connection CBSAs. Column 3 also includes fixed effects for the endpoint CBSAs.

The value of the coefficient on airborne travel time is chosen from Column 3 of Table 7 to be $\psi = 0.159$, as the coefficient on $\ln(t_{i,j}(r))$ corresponds to $-\psi$. The values of $\{\nu_i\}_{i\in I}$ are then inferred from the 'connection' fixed effects, which take the form $\psi \ln(\nu_i)$ in (14). The values of $\{\nu_i\}_{i\in I}$ are illustrated in the map in Figure 2.

The estimated values of ψ and $\{\nu_i\}_{i\in I}$ provide enough information to calculate $\Omega_{i,j}$ for all combinations of i and j. The second step of the estimation is to identify the parameter ρ and the values of $\{\mu_i\}_{i\in I}$ and $\{\omega_i\}_{i\in I}$ from (18). The equation is estimated using a Poisson regression, with the number of passengers between i and j as the dependent variable and the population of

CBSA i as the 'exposure'. The results are presented in Table 8, with the columns representing specifications with different sets of fixed effects.

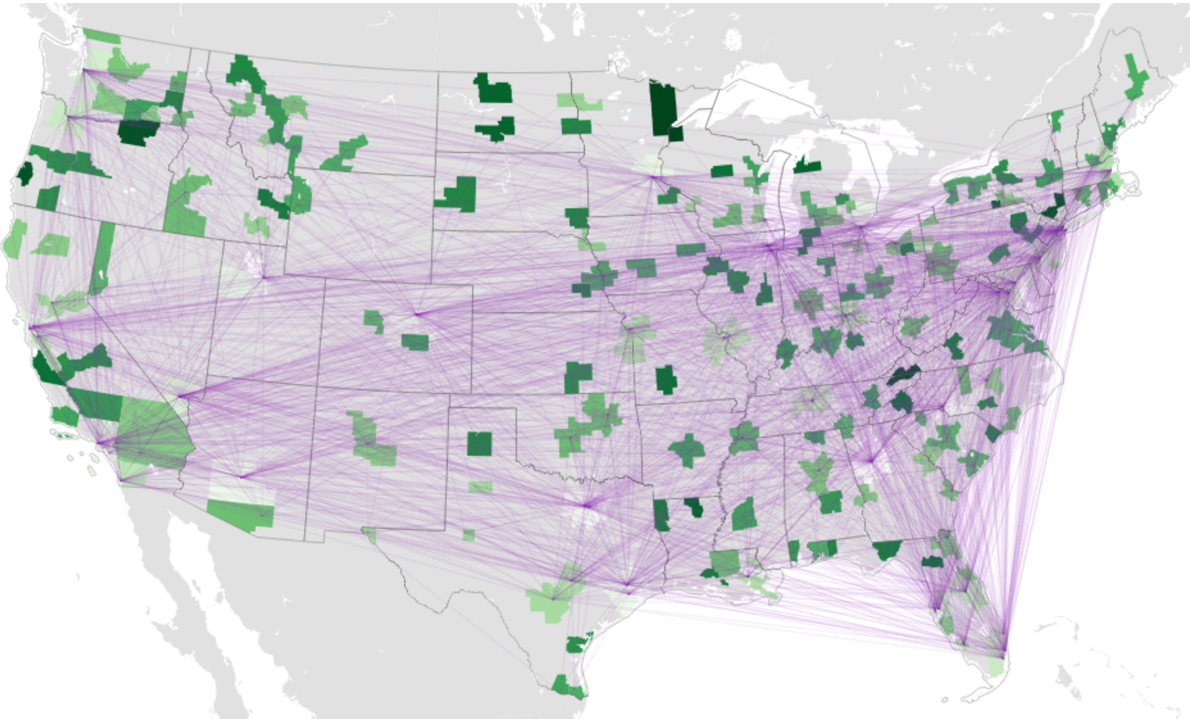

Figure 2: Map of the values of v_i, which represent the costs of connecting at an airport in each CBSA. Darker shading indicates a higher cost of using an airport. The lines represent the route segments in the route dataset.

	(1) Poisson	(2) Poisson	(3) Poisson	(4) Poisson
$\ln(\Omega_{i,j})$	-0.188^a (0.003)	-0.213^a (0.002)	-0.114^a (0.002)	-0.110^a (0.002)
Origin fixed effects		Y		Y
Destination fixed effects			Y	Y
Pseudo R^2	0.50	0.61	0.82	0.89

Note: 23,560 observations for each regression; robust standard errors in parentheses; a, b, c denote significance at 1%, 5%, 10%

Table 8: Estimates of effect of the cost of travel between (but not including) the two endpoint airports on the level of traffic. Columns 2 and 4 include fixed effects for the origin CBSAs and Columns 3 and 4 include fixed effects for the destination CBSAs.

The value of the coefficient ρ is chosen from the fourth column in Table 8, which represents the full specification, to be $\rho = 0.110$. The values of $\{\mu_i\}_{i\in I}$ and $\{\omega_i\}_{i\in I}$ are then solved for by iterating on the values of $\{\mu_i\}_{i\in I}$ until each of the terms that contain $\{\mu_i\}_{i\in I}$ and $\{\omega_i\}_{i\in I}$ in (18) matches the respective estimated fixed effect. The resulting values of $\{\mu_i\}_{i\in I}$ are shown in the map in Figure 3.

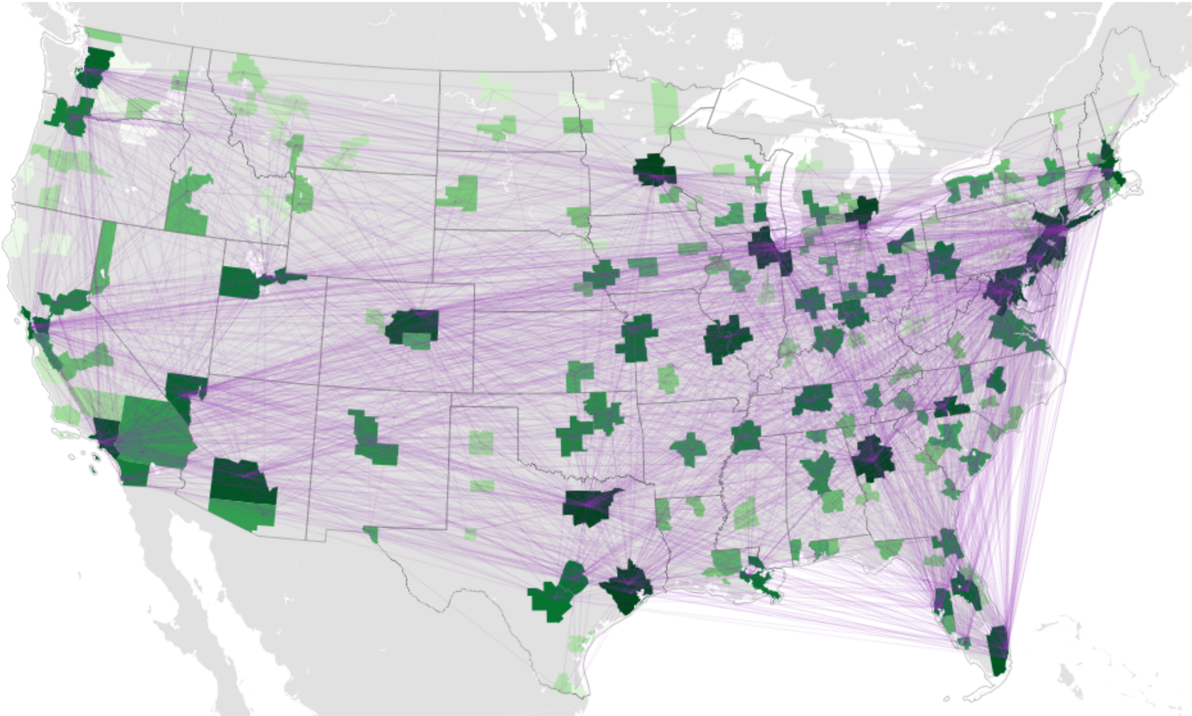

Figure 3: Map of the values of μ_i, which represent the costs of flying to or from the airports in each CBSA. Darker shading indicates a higher cost of using an airport. The lines represent the route segments in the route dataset.

The third step of the estimation is to identify the parameters α, β, γ, and δ, which determine how the cost using the airports in a given CBSA is affected by the airports' traffic levels and physical capacities. This is done by estimating (21) and (22). The results are presented in Table 9. Note that the coefficients β and δ are negative in sign in those equations.

Columns 1 to 5 of Table 9 present estimates run using ordinary least squares (OLS), with five measures of physical capacity: total runway length, total runway surface area, the value of the runways, the value of the terminals, and the sum of the runway and terminal values. Columns 6 to 10 present estimates run using two-stage least squares (TSLS) with instruments for log CBSA-level air traffic $\ln(A_i)$ and the same set of physical-capacity measures. The TSLS technique is employed because of the concern that the level of traffic could be influenced by the cost of using the airport or that both could be influenced by some external factor, in particular because the level of traffic is a factor in the estimates of $\{\mu_i\}_{i\in I}$ and $\{\nu_i\}_{i\in I}$. The instrument used for $\ln(A_i)$ is the log CBSA population in 2015.

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) TSLS	(7) TSLS	(8) TSLS	(9) TSLS	(10) TSLS
Panel A. Dependent variable: endpoint airport cost μ_i.										
$\ln(A_i)$	3.379^a	3.303^a	3.267^a	3.011^a	3.279^a	3.915^a	3.857^a	3.910^a	6.809^a	4.221^a
	(0.178)	(0.180)	(0.188)	(0.411)	(0.221)	(0.275)	(0.274)	(0.301)	(1.377)	(0.401)
$\ln(x_i)$ (log runway length)	1.189^c					−0.183				
	(0.613)					(0.822)				
$\ln(x_i)$ (log runway area)		1.406^b					0.042			
		(0.594)					(0.775)			
$\ln(x_i)$ (log runway value)			1.334^b					−0.132		
			(0.553)					(0.764)		
$\ln(x_i)$ (log terminal value)				0.681					-3.070^b	
				(0.435)					(1.366)	
$\ln(x_i)$ (log airport value)					1.032^c					−0.911
					(0.535)					(0.863)
R^2	0.83	0.83	0.83	0.83	0.83					
Panel B. Dependent variable: connection airport cost ν_i.										
$\ln(A_i)$	-2.825^a	-2.878^a	-2.901^a	-2.137^b	-2.197^a	-2.816^a	-2.901^a	-3.025^a	−3.860	-2.099^b
	(0.501)	(0.528)	(0.548)	(1.036)	(0.609)	(0.610)	(0.656)	(0.698)	(2.646)	(0.856)
$\ln(x_i)$ (log runway length)	-5.067^a					-5.089^a				
	(1.341)					(1.675)				
$\ln(x_i)$ (log runway area)		-4.419^a					-4.364^b			
		(1.339)					(1.704)			
$\ln(x_i)$ (log runway value)			-3.704^a					-3.423^b		
			(1.286)					(1.686)		
$\ln(x_i)$ (log terminal value)				-1.951^b					−0.249	
				(0.985)					(2.634)	
$\ln(x_i)$ (log airport value)					-4.972^a					-5.173^a
					(1.255)					(1.827)
R^2	0.56	0.56	0.55	0.54	0.57					
First-stage statistic						194.54	181.89	163.99	29.13	98.74

Note: 179 observations for each regression; robust standard errors in parentheses; *a*, *b*, *c* denote significance at 1%, 5%, 10%

Table 9: Estimates of the coefficients α, β, γ, and δ that define how the costs of using airports as trip endpoints or for connections depends on air traffic A_i and capacity x_i. Columns 1 to 5 are estimated using OLS and Columns 6 to 10 are estimated using TSLS, with the log population in 2015 as the instrument for log air traffic in 2015.

The values of the coefficients are chosen from Column 10 of Table 9, which shows the TSLS estimates with the value of the runways and terminals used as the measure of airport capacity. Thus the chosen values are $\alpha = 4.221$, $\beta = 0.911$, $\gamma = -2.099$, and $\delta = 5.173$. Positive values of β and δ mean that higher-capacity airports are less costly to use for a given level of traffic. The positive value of α indicates that for a given capacity, more traffic means a higher cost of using an airport, which is likely due to congestion. A similar effect of congestion should apply to flight connections, but the value of γ is negative. This suggests that busier airports have a substantial advantage for connecting flights, for example because of the greater possibilities for scheduling tickets with connections at airports with more frequent flights to more destinations, that exceeds the costs of congestion. This does not mean that connections should increase without bound, as they are counterbalanced by the costs of airborne travel time, but it is consistent with there being multiple equilibria in hub locations.

5 Simulation

This section presents a simulation of the model using the coefficients estimated above. The simulation exercise generates quantitative results not evident from the model coefficients. The exercise proceeds by first simulating the effect of the change in airport infrastructure on traffic in all parts of the network, then simulating the effects on employment that result from those changes in air traffic.

The effects of changes in airport infrastructure on air traffic across the network are simulated using the coefficients estimated from (15), (18), (21), and (22) in Section 4. Given the changes to runway and terminal values $\{x_i\}_{i\in I}$, the simulated levels of traffic $\{A_{i,j}\}_{i,j\in I}$ on the routes (and the corresponding values of $\{A_i\}_{i\in I}$ at the airports) are those that satisfy the equations, holding the values of the parameters α, β, γ, δ, ψ, ρ, and $\{\omega_i\}_{i\in I}$ and the residuals $\{\varepsilon_{i,j}\}_{i,j\in I}$, $\{\varepsilon_i^{\mu}\}_{i\in I}$, and $\{\varepsilon_i^{\nu}\}_{i\in I}$ constant. It is implicitly assumed that the preferences for travel on each route and to each destination remain unchanged and the airport infrastructure in each CBSA remains at the same level of efficiency.

Given the simulated changes in air traffic throughout the network, the changes in employment in all sample CBSAs are predicted using the coefficients estimated in Section 3. The effects of a change in local air traffic on the local population and employment are estimated using the coefficients estimated in Table 4. Given the change in the local population, the changes in population levels in other CBSAs are adjusted according to the migration patterns estimated in Table 6 to keep the total population constant. Given the population changes, the changes in employment in other CBSAs are inferred from the effects on employment rates in other places from Table 6.

The TSLS coefficients in Table 6 for the effects of air traffic by distance band on the population in CBSA m are all not significant and so they show no clear pattern. It is therefore assumed that the populations of all other CBSAs adjust in equal proportions of their populations such that they offset the population change in CBSA m. The TSLS coefficients for the effects of changes in air traffic elsewhere on the employment rate in Table 6 are taken as given where they are significant and set to zero otherwise. Thus the coefficients for 200-400 miles is 0.059 and the coefficient for 600-800 miles is -0.111.

It is convenient to express the simulation results as a set of elasticities. The relationship between changes in the value x_i of the airport runways and terminals in CBSA i and the level A_j of air traffic in CBSA j is represented by the elasticity $e^{x \cdot A}_{i,j} = \frac{dA_j/A_j}{dx_i/x_i}$. The relationship between changes in air traffic A_j in CBSA j and employment L_k in CBSA k is represented by the elasticity $e^{A \cdot L}_{j,k} = \frac{dL_k/L_k}{dA_j/A_j}$. The effect of a change in the value of the runways and terminals in CBSA i on employment in CBSA k is thus approximated as the elasticity $e^{x \cdot L}_{i,k} = \sum_j e^{x \cdot A}_{i,j} e^{A \cdot L}_{j,k}$. Table 10 presents summary statistics for the simulated elasticities $\left\{e^{x \cdot A}_{i,j}\right\}_{i,j \in I}$ and $\left\{e^{x \cdot L}_{i,k}\right\}_{i,k \in I}$.

	Elasticity $e^{x \cdot A}_{i,j}$ of air traffic to runway value		Elasticity $e^{x \cdot L}_{i,k}$ of employment to runway value	
CBSA pairs	$j = i$	$j \neq i$	$k = i$	$k \neq i$
Mean	0.590	0.033	0.022	0.001
Standard deviation	0.405	0.149	0.018	0.007
Minimum	0.100	0.000	0.004	−0.052
Maximum	3.546	5.170	0.170	0.269
10th percentile	0.139	0.000	0.005	−0.001
25th percentile	0.251	0.000	0.009	−0.000
Median	0.591	0.001	0.021	−0.000
75th percentile	0.820	0.007	0.029	0.000
90th percentile	0.899	0.039	0.032	0.001
Number of observations	179	31,862	179	31,862

Table 10: Summary statistics for the elasticities of passenger numbers and employment to the value of airport infrastructure at the CBSA level. The elasticities are presented first for the airport value paired to passengers or employment in the same CBSA and then for the variables paired between different CBSAs.

The map in Figure 4 illustrates the values of the elasticity $e^{x \cdot A}_{i,j}$ that defines the relationship between the value of runways and terminals in CBSA i and air traffic in CBSA j. As it is not practical to illustrate all pairs of CBSAs, the map represents the elasticity parameters for the change in the value of the runways and terminals in a single CBSA. Oklahoma City, OK is used for the illustration as it is in a relatively central location and not near either extreme of the size spectrum. The map also highlights the flight segments in the route dataset from Will Rogers World Airport in Oklahoma City, the only sample airport in the Oklahoma City, OK CBSA.

The simulation results can be used to estimate how much a given investment in the infrastructure at an airport affects employment across the country. This is straightforward to do from the elasticities $\left\{e^{x \cdot L}_{i,k}\right\}_{i,k \in I}$, the values of runways and terminals $\{x_i\}_{i \in I}$, and employment $\{L_k\}_{k \in I}$. Figure 5 illustrates the effects of investing one million dollars in the runways and terminals in Oklahoma City on employment in each sample CBSA. It is apparent from the map that there is a larger effect on employment in the CBSAs with direct connections to Oklahoma City.

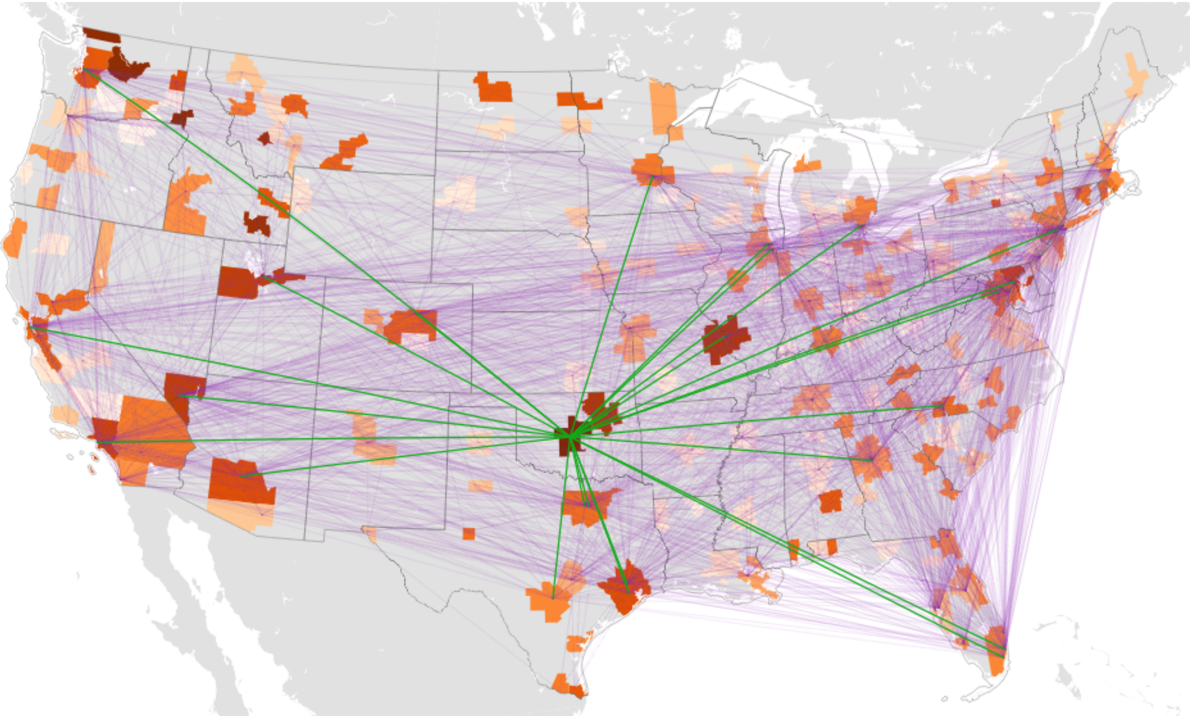

Figure 4: Map with the sample CBSAs shaded to represent the elasticity parameter $e_{i,j}^{x \cdot A}$ for the Oklahoma City, OK CBSA. The values of $e_{i,j}^{x \cdot A}$ represent the change in air traffic in each CBSA j that results from a change in the value of the airport infrastructure in CBSA i. Darker shading represents larger values of $e_{i,j}^{x \cdot A}$. The air routes from Will Rogers World Airport in Oklahoma City are also highlighted.

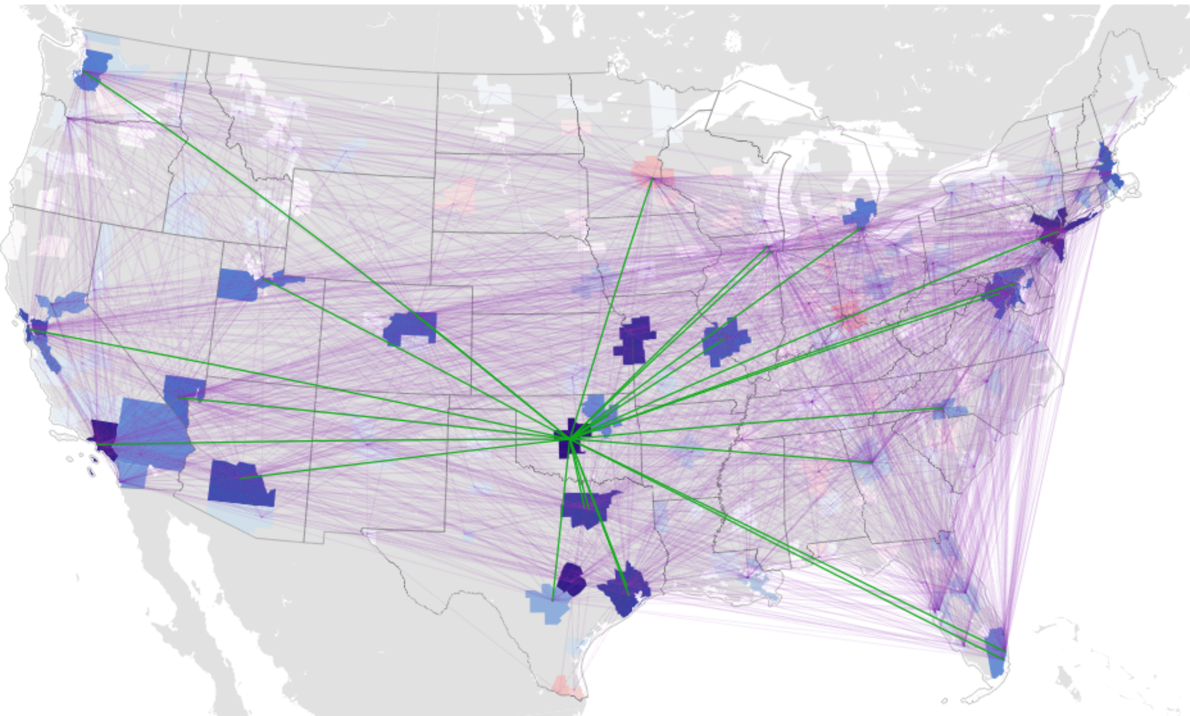

Figure 5: Map shaded for the changes in employment in each CBSA that result from adding $1 million of airport infrastructure to the Oklahoma City, OK CBSA. Positive changes are shaded in blue and negative changes are shaded in red, with darker shading representing values of larger magnitude.

A natural extension is to calculate the number of jobs created in the US by an investment at the airports in a single CBSA. This is done by summing the effects of a change to the airport infrastructure in a CBSA over all CBSAs in which employment is affected. Table 11 presents summary statistics – at the CBSA level – for the numbers of jobs created in the local CBSA and in the whole US by an investment of $1 million in the runways and terminals in each CBSA. The map in Figure 6 illustrates the numbers of jobs created in the US by investing $1 million in each of the CBSAs in the sample.

	Change in local employment for $1m investment	Change in US employment for $1m investment
Mean	3.57	12.75
Standard deviation	4.51	20.23
Minimum	0.07	−16.66
Maximum	45.56	105.91
10th percentile	0.40	0.05
25th percentile	1.14	1.07
Median	2.45	6.65
75th percentile	4.63	16.76
90th percentile	7.11	29.76

Note: 179 observations for each variable

Table 11: Summary statistics for the numbers of jobs created within the same CBSA and in the US as a whole by investing $1 million in the airport infrastructure in each CBSA.

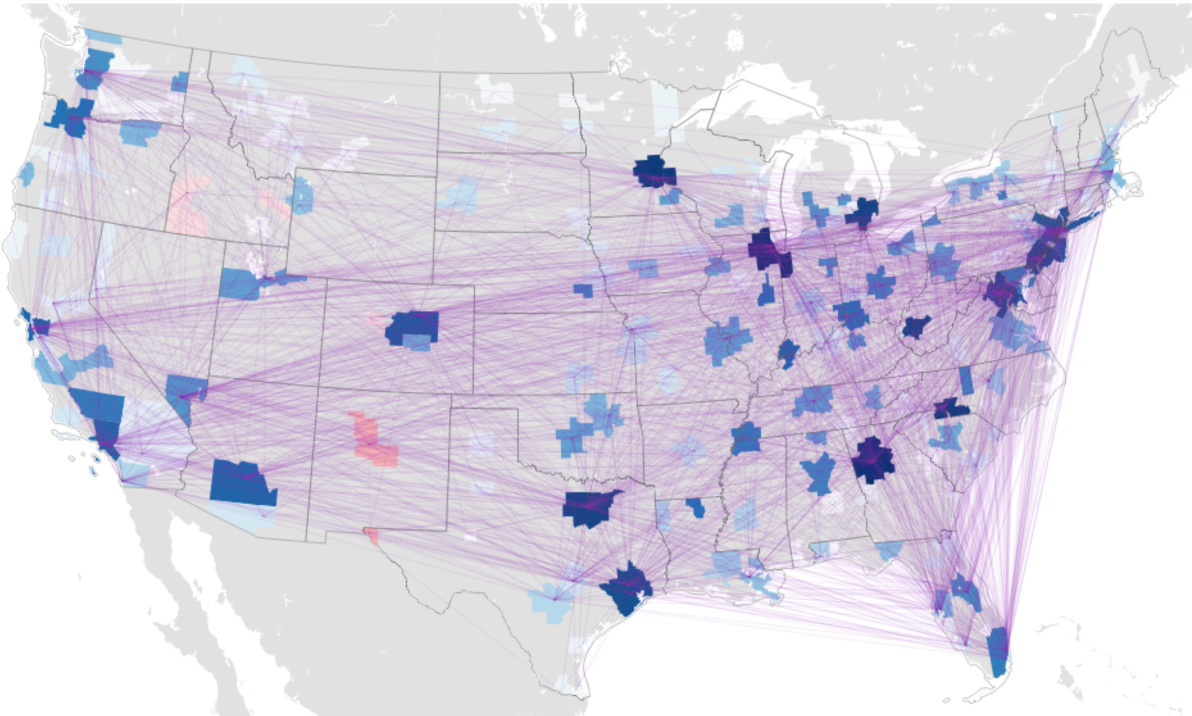

Figure 6: Map shaded for the changes in US employment resulting from adding $1 million of airport infrastructure to each CBSA. Positive values are shaded in blue and negative values are shaded in red, with darker shading for larger magnitudes.

As shown in Table 11, the mean number of jobs created by a $1 million investment in the airport infrastructure in a CBSA is 12.75, which corresponds to one job created for approximately each $78,000 invested. The results range from a small loss of jobs, due to the diversion of connecting traffic away from more productive CBSAs, to 105.91 jobs created per $1 million investment, which is approximately $9,400 of additional infrastructure for each job created.[16] On average, the number of jobs created in total is around three and a half times as large as in the local CBSA.

The results also indicate a positive correlation between the size of an airport and the number of jobs created by its expansion. Figure 7 plots the number of jobs created in the US by a $1 million investment against the actual value of the airport infrastructure in 2015. Table 12 lists the CBSAs for which investing in the airport infrastructure has the smallest and largest effects on national employment. The table shows the total number of jobs created by a $1 million investment and, where this is positive, the corresponding investment per job created.

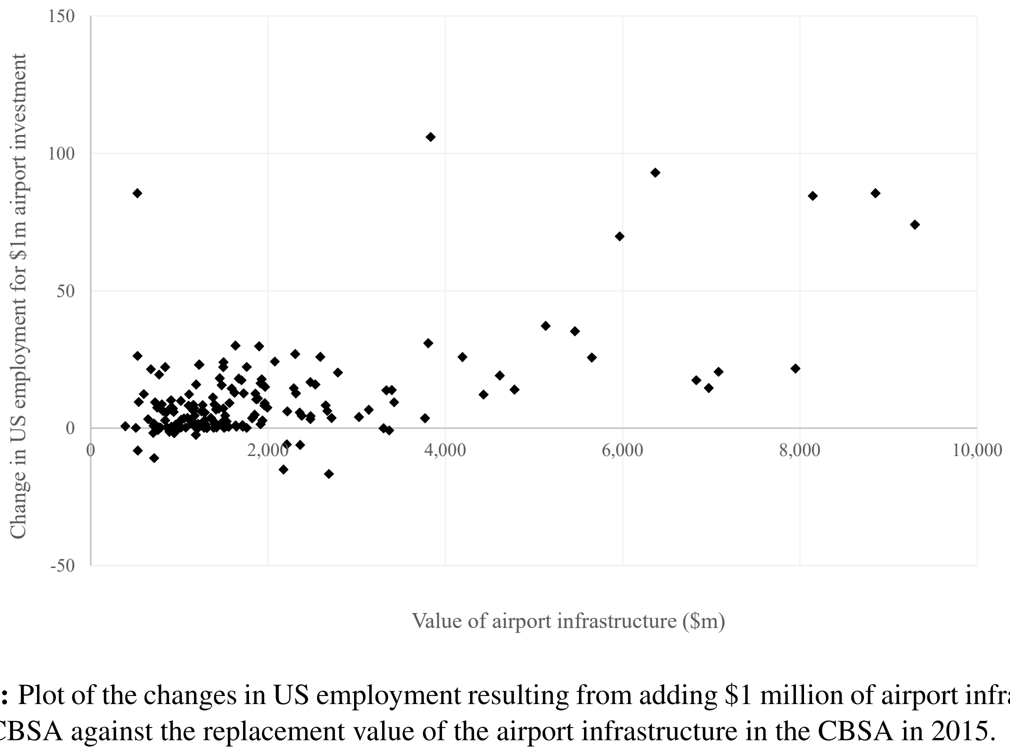

Figure 7: Plot of the changes in US employment resulting from adding $1 million of airport infrastructure to each CBSA against the replacement value of the airport infrastructure in the CBSA in 2015.

[16] It should be noted that the ongoing costs of future maintenance are not considered, so the costs of expanding an airport are likely to be understated in this analysis.

CBSA	Change in US employment for $1m investment	Investment ($) for each job created in US
Albuquerque, NM	−16.66	–
El Paso, TX	−15.04	–
Pendleton-Hermiston, OR	−10.89	–
Jackson, WY-ID	−8.18	–
Lincoln, NE	−6.15	–
Boise City-Nampa, ID	−5.93	–
Eugene-Springfield, OR	−2.45	–
Idaho Falls, ID	−1.89	–
Edwards, CO	−1.76	–
Bend, OR	−1.34	–
⋮	⋮	⋮
Houston-Sugar Land-Baytown, TX	56.47	17,707
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	69.73	14,340
Atlanta-Sandy Springs-Marietta, GA	74.03	13,508
Dallas-Fort Worth-Arlington, TX	75.03	13,328
Washington-Arlington-Alexandria, DC-VA-MD-WV	84.55	11,827
Detroit-Warren-Livonia, MI	85.44	11,704
Charleston, WV	85.44	11,704
Minneapolis-St. Paul-Bloomington, MN-WI	92.95	10,759
Chicago-Joliet-Naperville, IL-IN-WI	100.12	9,988
Charlotte-Gastonia-Rock Hill, NC-SC	105.91	9,442

Table 12: The numbers of jobs created in the US by investing $1 million in the airport infrastructure in a CBSA and the corresponding investment per job created. The top 10 and bottom 10 CBSAs are listed.

Figure 7 shows a positive overall relationship between the size of an airport and the number of jobs created in the US by additional investment. Table 12 shows that many of the CBSAs where a given dollar investment in airport infrastructure would generate the most jobs are among the largest metropolitan areas in the country, while the CBSAs with the lowest payoffs are generally relatively small.[17] This suggests that it would be optimal for the national air network to be more concentrated than it currently is.

The result that the actual air network is less concentrated than would be optimal implies unrealized benefits from concentration, in particular for airports with high shares of connections. Table 9 shows that the effect of increased infrastructure on traffic is greater for connections than for the start or end of a trip. The returns to scale in the amount of traffic are also positive for connecting airports but negative for endpoint airports, as $\alpha > 0$ and $\gamma < 0$, which controlling for everything else means a larger increase in traffic at airports with more connecting traffic.

There are at least three potential explanations for why the actual air network would be less concentrated than the optimal network, despite the unrealized benefits this implies. One is the

[17]The second and third largest CBSAs in terms of air traffic in 2015 were Chicago-Joliet-Naperville, IL-IN-WI and Atlanta-Sandy Springs-Marietta, GA, which feature as the CBSAs with the second and eighth highest payoffs in Table 12. Most of the CBSAs in the bottom ten are smaller cities with regional airports.

role that local governments play in airport planning. As explained above, expanding an airport will generally lead to an increase in local employment, which motivates local governments to invest in their own infrastructure. However, this will cause traffic and therefore employment elsewhere to decline or increase, which is relevant to the interests of the federal government but the local government is not motivated to consider. As the effects on employment elsewhere tend to be positive and large for the larger airports, decisions being made at the local level leads to a network that is too dispersed to maximize national employment.

A second potential explanation for the air network being suboptimally concentrated is the imperfect information that voters have about infrastructure projects. Glaeser and Ponzetto (2018) argued that federal funding leads to overspending on local projects as voters in the rest of the country have less information about how the public money is being spent. Moreover, they argued that the degree of overspending is greater when the local area is smaller, which would increase the sizes of the smaller airports.

Thirdly, a simple explanation for the air network being suboptimally concentrated is the persistence in airport locations and sizes. In the past, the prevailing technology for air travel favored smaller aircraft and less concentrated operations than what exists today. In addition, air travel in the US was heavily regulated prior to the Airline Deregulation Act of 1978, one consequence of which a more dispersed network than market forces would have generated.[18] Even if the air network would naturally adjust to the degree of concentration that is optimal according to the analysis presented here, the process of adjustment would likely take many years and would thus still be incomplete.

6 Conclusion

This paper estimates the effect of airport infrastructure employment in the US. It extends the existing literature on the effects of airport infrastructure by estimating how a change to a given airport affects traffic across the network and the resulting shifts in population and jobs around the country. This exercise introduces a greater degree of uncertainty than the estimation of local

[18] The history of the 1978 Airline Deregulation Act is described in detail and its effects analyzed by Blonigen and Cristea (2015).

effects, but the topic has received far less attention in the literature than the local effects.

The analysis integrates a model for traffic on the air network with a model for the effects of that traffic on employment. To obtain reliable identification, the latter is estimated using an instrumental variables technique. The estimation is complemented by a simulation exercise that quantifies the effects of investments at each particular airport.

Though there is substantial uncertainty about the aggregate estimates, the results suggest that investing in airport infrastructure generates additional jobs with reasonable amounts of public spending. On average, around $78,000 of public spending is required for each job created. The benefits vary widely by metropolitan area. The airports with the highest returns to public investment require less than $10,000 for each job created, while expanding some airports actually decreases overall US employment by diverting air traffic from more important airports. These figures come with two caveats. The first is that there must be sufficient space to build more infrastructure and the cost of acquiring more land to expand an airport is not included in these figures. The second is that these figures are only for construction and do not consider the costs of maintenance, which would be higher for a larger facility.

The results further suggest that, from the national perspective, the optimal US air network would be more concentrated than the current network. That is, the number of jobs created by a given dollar investment in infrastructure is generally larger for larger airports and this is especially true for employment outside of the local area. This appears to be in part because of the benefits of scale in airport operations, mostly for connections. It would also be the logical result of local governments providing part of the public funding for airports, as they are primarily interested in local employment and discount the effects on employment elsewhere.

The analysis in this paper could be applied to other places or modes of transportation. The results for other air networks could be quite different, as most countries have significantly smaller networks with fewer airports than the US. The analysis could also be extended to include international flights, though this would require additional information that is at present not publicly available.

References

Allen, T. and Arkolakis, C.: 2019, The Welfare Effects of Transportation Infrastructure Improvements. Unpublished manuscript.

Bartik, T. J.: 1991, *Who Benefits from State and Local Economic Development Policies?*, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Baum-Snow, N.: 2007, Did Highways Cause Suburbanization?, *Quarterly Journal of Economics* **122**(2), 775–805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A. and Zhang, Q.: 2017, Roads, Railroads, and Decentralization of Chinese Cities, *Review of Economics and Statistics* **99**(3), 435–448.

Blonigen, B. A. and Cristea, A. D.: 2015, Air Service and Urban Growth: Evidence from a Quasi-Natural Policy Experiment, *Journal of Urban Economics* **86**, 128–146.

Bogart, D. and Chaudhary, L.: 2013, Engines of Growth: The Productivity Advance of Indian Railways, 1874–1912, *Journal of Economic History* **73**(2), 339–370.

Brueckner, J. K.: 2003, Airline Traffic and Urban Economic Development, *Urban Studies* **40**(8), 1455–1469.

Brueckner, J. K., Lee, D. and Singer, E.: 2014, City-Pairs vs. Airport-Pairs: A Market-Definition Methodology for the Airline Industry, *Review of Industrial Organization* **44**(1), 1–25.

Campante, F. and Yanagizawa-Drott, D.: 2018, Long-Range Growth: Economic Development in the Global Network of Air Links, *Quarterly Journal of Economics* **133**(3), 1395–1458.

Donaldson, D.: 2018, Railroads of the Raj: Estimating the Impact of Transportation Infrastructure, *American Economic Review* **108**(4-5), 899–934.

Donaldson, D. and Hornbeck, R.: 2016, Railroads and American Economic Growth: A "Market Access" Approach, *Quarterly Journal of Economics* **131**(2), 799–858.

Duranton, G.: 2015, Roads and Trade in Colombia, *Economics of Transportation* **4**(1–2), 16–36.

Duranton, G., Morrow, P. and Turner, M.: 2014, Roads and Trade: Evidence from the US, *Review of Economic Studies* **81**(2), 681–724.

Duranton, G. and Turner, M.: 2012, Urban Growth and Transportation, *Review of Economic Studies* **79**(4), 1407–1440.

Eaton, J. and Kortum, S.: 2002, Technology, Geography, and Trade, *Econometrica* **70**(5), 1741–1779.

Federal Aviation Administration: 2005, *Chicago's O'Hare Modernization Program*, U.S. Department of Transportation, Washington, DC.

Federal Aviation Administration: 2015, *Airport/Facility Directory – 25 June 2015 to 20 August 2015*, U.S. Department of Transportation, Washington, DC.

Gibbons, S. and Wu, W.: 2020, Airports, Access and Local Economic Performance: Evidence from China, *Journal of Economic Geography* **20**(4), 903–937.

Glaeser, E. L. and Ponzetto, G. A. M.: 2018, The Political Economy of Transportation Investment, *Economics of Transportation* **13**, 4–26.

Gonzalez-Navarro, M. and Turner, M.: 2018, Subways and Urban Growth: Evidence from Earth, *Journal of Urban Economics* **108**, 85–106.

Green, R.: 2007, Airports and Economic Development, *Real Estate Economics* **35**(1), 91–112.

Lakew, P. A. and Bilotkach, V.: 2018, Airport Delays and Metropolitan Employment, *Journal of Regional Science* **58**(2), 424–450.

McGraw, M. J.: 2017, The Heterogeneous Impact of Airports on Population and Employment Growth in Cities, *in* J. D. Bitzan and J. H. Peoples (eds), *The Economics of Airport Operations (Advances in Airline Economics, Volume 6)*, Bingley, UK: Emerald Publishing Limited, pp. 261–312.

Michaels, G.: 2008, The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System, *Review of Economics and Statistics* **90**(4), 683–701.

Protsiv, S. and Sheard, N.: 2020, Transport Infrastructure and Regional Welfare: The Benefits of Trade and Passenger Travel. Unpublished manuscript.

Sheard, N.: 2014, Airports and Urban Sectoral Employment, *Journal of Urban Economics* **80**, 133–152.

Sheard, N.: 2019, Airport Size and Urban Growth, *Economica* **86**(342), 300–335.

Sheard, N.: 2020, The Geography of US Airport Markets. Unpublished manuscript.

Transport Canada: 2016, *Airport Pavement Bearing Strength Reporting*, Government of Canada, Ottawa, Canada. Advisory Circular (AC) No. 302-011.

United States Census Bureau: 2014, *2014 State and Local Government Finances*, U.S. Government Printing Office, Washington, DC.

United States Department of Transportation: 2016, *Budget Highlights: Fiscal Year 2017*, U.S. Government Printing Office, Washington, DC.

A Approximating the values of runways and terminals

The replacement value of each airport in 2015 is derived from data on its runways and terminal buildings. The details of runways are from the Airport/Facility Directory for 25 June 2015 to 20 August 2015, a guide for pilots and navigators published by the Federal Aviation Administration (2015). The data on the sizes of airport terminal buildings are from OpenStreetMap.

The replacement value of each runway is calculated based on its dimensions and strength. The Airport/Facility Directory details the length and width of each runway along with its weight-bearing capacity for each type of wheel arrangement. The weight-bearing capacities (along with the runway lengths) were condensed into five broad categories of aircraft that the runway could handle: *small aircraft*, *medium 2-engine aircraft*, *large 2-engine aircraft*, *heavy 2-engine aircraft*, and *heavy 4-engine aircraft*. The criteria for each category are based on statistics for various models of aircraft published by Transport Canada (2016) and are detailed in Table 13.

		Weight bearing capacity by wheel arrangement ('000 lbs)			
Aircraft type	Runway length (ft)	S	D	DT	DDT
Small aircraft	3,000	10			
Medium 2-engine aircraft	6,000		65		
Large 2-engine aircraft	7,500			150	
Heavy 2-engine aircraft	9,000			400	
Heavy 4-engine aircraft	10,000				750

Table 13: Minimum runway requirements for each of the five categories of aircraft. The weight-bearing capacity is evaluated only for the relevant type of wheel arrangement. Each runway is assigned the largest category of aircraft that it has the capacity for.

Table 14 lists the factors used to calculate the value of each runway given the largest type of aircraft that it has the capacity to handle. The total value of each runway is the sum of the components calculated based on the length and area of the runway. The cost of constructing a runway naturally depends on its surface area, but due to the shoulders, markings, and taxiways there is also a component that depends on the length of the runway. Furthermore, higher-capacity runways have stronger surfaces and deeper subgrades and thus cost more to construct. The values in Table 14 were chosen to approximate the actual construction costs of recent Airport Improvement Program (AIP) runways detailed by the Federal Aviation Administration (2005). Due to the small number of projects these factors are estimated imprecisely, but the overall construction costs they generate match observed runway construction costs reasonably well.

Aircraft type	Per foot of length (USD)	Per square foot of area (USD)
Small aircraft	18,000	260
Medium 2-engine aircraft	23,400	338
Large 2-engine aircraft	27,000	390
Heavy 2-engine aircraft	30,600	442
Heavy 4-engine aircraft	33,300	481

Table 14: Factors used to calculate the value of a runway given its dimensions and the largest category of aircraft it has the capacity for. The total value assigned to each runway is the sum of the values calculated based on its length and its width.

As an example of how the figures in Tables 13 and 14 are applied, consider a hypothetical runway $8,000'$ long and $150'$ wide with a weight-bearing capacity of $400,000$ lb for DT (dual tandem) landing gear. According to Table 13, the runway is categorized for *large 2-engine aircraft* as it exceeds the $7,500'$ length threshold and the $150,000$ lb weight-bearing threshold for that category, but is too short for any larger category of aircraft (even though it has sufficient weight-bearing capacity for *heavy 2-engine aircraft*). Therefore, the replacement value of the runway is calculated using the factors in Table 14 to be $\$27,000 \times 8,000 + \$390 \times 8,000 \times 150 = \$684,000,000$.

The replacement values of the airport terminals are inferred from the sizes of the building footprints, which are calculated from the building shapefiles published by OpenStreetMap. As larger terminal buildings tend also to be taller structures that cost more per square foot to construct, the heights of the terminal buildings are approximated from the footprints and the value of each terminal is calculated as a multiple of its volume. The terminal heights are calculated as being proportional to the square root of the building footprint, with the values scaled so that the airport terminal in Lansing, MI has a height of 30 feet. The terminal value is then calculated as \$10 per cubic foot of volume, which approximates the construction costs of recent terminal projects at Boise International Airport,[19] O'Hare International Airport in Chicago (Federal Aviation Administration, 2005), Dallas/Fort Worth International Airport,[20] and Salt Lake City International Airport.[21] The figure of \$10 per cubic foot is approximate and the actual cost of

[19]The project to build the new terminal at Boise International Airport is described on the construction firm's website at https://www.cshqa.com/projects/boise-airport-terminal-replacement/ (accessed 2019-04-06).

[20]The Dallas News cited a cost of \$1 billion for the renovations to Terminal A at Dallas/Fort Worth International Airport in the article "DFW Airport shows off the first phase of its \$2.7 billion upgrade", published in January, 2017.

[21]The project to build a new terminal at Salt Lake City International Airport was discussed in the New York Times article "Building a Modern Airport in Salt Lake City From One Well Past Its Prime", published January 2nd, 2018.

building a terminal will depend on many architectural decisions that can yield vastly different costs for facilities that can handle similar numbers of passengers, but it serves as a reasonable guide based on these recent examples.

B Airline and aircraft categories used for instruments

B.1 Airlines

Table 15 lists the categories used to construct the 'airline' instrument. As it is necessary to track airlines between years, they are grouped by the *Unique Carrier Code* that is assigned by the BTS for exactly that purpose. The conditions for airlines to be included is that they must have had an average of at least 10 daily flights and 100 daily passengers in one year between 1991 and 2018. The exclusion of the airlines below this threshold actually makes little difference to the instruments, as they generally have small shares of the traffic at commercial airports. The table gives the aggregate amount of traffic for each airline between 1991 and 2018 with an origin or destination in the contiguous US and is presented in descending order of passenger numbers.

Mergers and acquisitions are common in the airline industry and create an issue for the calculation of growth rates, as for example an airline that acquires a competitor would have a discrete increase in its operations that does not reflect organic growth in its operations. This issue is addressed by assigning the growth in aggregate traffic to all entities involves in the merger or acquisition over the period in which it takes place. That is, if airline A acquires airline B and continues operating under the code for airline A, then the overall growth rates applied to both A and B for the period overlapping the acquisition are calculated as the traffic of airline A at the end of the period divided by the sum of traffic of airline A and B at the start of the period. For some mergers, the traffic for the former entities was retained for subsequent traffic continued to be classified under the former codes and for those mergers no adjustment was made. The mergers and acquisitions where traffic was recoded were identified by hand and are listed in Table 16.

Unique Carrier Code	Airline name	Number of flights	Number of passengers ('000)	Unique Carrier Code	Airline name	Number of flights	Number of passengers ('000)
DL	Delta Air Lines	23,968,165	2,806,365	J7	Valujet Airlines	212,960	14,023
AA	American Airlines	23,110,525	2,605,576	SR	Swissair Transport	71,366	13,737
WN	Southwest Airlines	26,313,580	2,573,907	C5	CommutAir	643,462	13,698
UA	United Air Lines	18,231,745	2,123,829	TK	Turk Hava Yollari A.O.	60,161	13,677
US	US Airways	15,578,025	1,364,623	RV	Air Canada Rouge	83,904	13,036
NW	Northwest Airlines	10,198,850	982,796	CA	Air China	57,524	12,897
CO	Continental Air Lines	9,205,140	919,262	FI	Icelandair	87,605	12,765
AS	Alaska Airlines	4,508,752	445,868	PT	Capital Cargo International	426,673	12,461
B6	JetBlue Airways	3,537,754	396,121	U5	USA 3000 Airlines	95,257	11,900
OO	SkyWest Airlines	9,146,229	363,115	RG	Varig	71,009	11,605
HP	America West Airlines	3,577,033	334,334	NJ	Vanguard Airlines	147,636	11,046
MQ	American Eagle Airlines	10,094,955	325,515	KW	Carnival Air Lines	98,199	10,439
EV	Atlantic Southeast Airlines	8,646,970	318,461	9K	Cape Air	1,884,508	10,298
TW	Trans World Airways	3,192,912	274,421	3M	Silver Airways / Gulfstream Int'l	815,387	10,273
FL	AirTran Airways Corporation	3,079,364	273,342	F8	Freedom Airlines	277,230	10,261
HA	Hawaiian Airlines	1,798,121	193,706	UP	Bahamasair	159,663	9,672
XE	ExpressJet Airlines	6,053,023	189,996	QR	Qatar Airways (Q.C.S.C)	39,650	9,629
F9	Frontier Airlines	1,606,527	180,131	ML (1)	Midway Airlines (Chicago, IL)	170,052	9,611
NK	Spirit Air Lines	1,368,136	178,382	RD	Ryan International Airlines	71,318	8,408
YV	Mesa Airlines	4,176,737	171,080	OE	WestAir Airlines	651,950	7,994
QX	Horizon Air	4,568,632	158,884	2T	Canada 3000 Airlines Ltd.	49,719	7,643
BA	British Airways	744,993	158,727	HQ (1)	Business Express	620,932	7,627
9E	Pinnacle Airlines	3,793,651	157,543	MG	Champion Air	60,924	7,398
OH	Comair	3,801,660	147,670	N7	National Airlines	63,190	7,126
AC	Air Canada	1,687,140	132,000	DY	Norwegian Air Shuttle	26,073	6,871
YX	Republic Airlines	2,134,755	125,660	KN	Morris Air Corporation	66,201	6,828
ZW	Air Wisconsin	3,327,441	124,632	ZK	Great Lakes Airlines	918,629	6,795
LH	Lufthansa German Airlines	458,327	109,896	BF	MarkAir	127,673	6,736
G4	Allegiant Air	644,718	88,775	KP	Kiwi International	75,213	6,382
TZ	ATA Airlines	657,717	87,927	PD	Porter Airlines	146,190	6,217
KH	Aloha Air Cargo	1,140,736	86,173	4O	ABC Aerolineas SA de CV / Interjet	56,853	5,453
XJ	Mesaba Airlines	2,947,125	78,732	LGQ	Lineas Aereas Allegro	45,114	5,187
AF	Air France	339,569	77,963	RS	Sky Regional Airlines	89,616	5,022
VS	Virgin Atlantic Airways	268,733	75,942	T9	TransMeridian Airlines	38,443	4,772
JL	Japan Air Lines	301,264	73,121	NA	North American Airlines	38,566	4,505
RP	Chautauqua Airlines	2,121,262	71,788	W7	Western Pacific Airlines	58,319	4,503
AX	Trans States Airlines	2,778,364	66,518	5D	Aerolitoral	94,191	4,471
VX	Virgin America	538,899	60,371	AL	Skyway Airlines	297,374	4,405
S5	Shuttle America	1,254,914	59,415	ZX	Air Georgian	253,983	4,248
KE	Korean Air Lines	280,375	58,919	U2	UFS	136,516	4,207
MX	Mexicana	570,664	54,047	KS	Peninsula Airways	519,427	4,135
AM	Aeromexico	555,586	51,695	JR	Aero California	72,820	4,045
16	PSA Airlines	1,249,632	47,535	C8 (1)	Chicago Express Airlines	185,095	3,879
KL	KLM Royal Dutch Airlines	180,761	45,126	YR	Grand Canyon Airlines	299,777	3,683
OW	Executive Airlines	1,347,280	43,551	HRZ	Allegheny Airlines	183,669	3,422
CP	Compass Airlines	726,896	43,274	SLQ	Sky King	43,745	3,277
YX (1)	Midwest Airlines	729,720	40,580	WV (1)	Air South	50,628	3,087
WS	Westjet	334,431	39,944	ZV	Air Midwest	417,503	2,907
G7	GoJet Airlines / United Express	741,326	39,473	L3	Lynx Aviation / Frontier Airlines	62,578	2,873
SY	Sun Country Airlines / MN Airlines	320,593	38,300	PCQ	Pace Airlines	40,579	2,849
17	Piedmont Airlines	1,465,350	35,033	GD	Transp. Aereos Ejecutivos	30,077	2,593
CX	Cathay Pacific	121,853	30,653	KV	Sky Regional Airlines	41,331	2,370
QQ	Reno Air	351,058	30,156	0JQ	Vision Airlines	54,119	2,286
EI	Aer Lingus Plc	116,290	27,749	09Q	Swift Air, LLC	37,161	2,164
TA	TACA International Airlines	260,945	27,269	6A	Aviacsa Airlines	28,517	2,085
NH	All Nippon Airways Co.	135,362	26,769	KAH	Kenmore Air Harbor	321,845	1,405
AZ	Compagnia Aerea Italiana	130,522	26,499	JX	Southeast Airlines	12,677	1,313
BR	Eva Airways Corporation	108,732	26,085	5J	Private Jet Expeditions	12,139	1,277
QK	Air Canada Regional	785,383	25,782	PN	Pan American Airways (19982004)	20,365	1,270
SQ	Singapore Airlines Ltd.	109,370	25,729	P9	Pro Air	22,643	1,182
IB	Iberia	138,314	25,665	W9	Eastwind Airlines	25,763	1,142
JM	Air Jamaica	213,511	24,669	GQ	Big Sky Airlines	126,874	1,070
DH	Independence Air	769,234	22,788	EM	Empire Airlines	50,689	1,047
SK	Scandinavian Airlines Sys.	123,542	22,443	8N	Flagship Airlines	38,139	983
PA (1)	Pan American World Airways	198,026	22,288	K5	SeaPort Airlines	293,021	955
EK	Emirates	74,380	21,115	SX	Skybus Airlines	9,314	932
AV	Avianca	167,636	20,091	3C	Regions Air	90,874	676
CM	Compania Panamena	193,337	20,066	1DQ	Island Airlines	85,119	563
9L	Colgan Air	956,816	19,419	0MQ	Air Choice One	93,035	415
LX	Swiss International Airlines	91,272	17,371	1AQ	Charter Air Transport	26,262	411
EA	Eastern Air Lines	221,470	17,240	4B	Olson Air Service	78,805	391
CP (1)	Canadian Airlines	178,489	15,466	LF	Jettrain Airlines	30,216	330
JJ	Transportes Aeros Meridiona	78,411	15,306	5Y	Atlas Air	14,634	315
TB (1)	USAir Shuttle	219,841	15,263	APN	Aspen Airways	6,981	314
BW	Caribbean Airlines	136,551	15,025	1RQ	Sun Air Express	72,302	177
JI (1)	Midway Airlines (Morrisville, NC)	295,780	14,887	WST	West Isle Air	165,352	155
Y4	Volaris	116,800	14,832	2JQ	Delux Public Charter	6,064	109
LA	Lan-Chile Airlines	104,167	14,318				

Note: the traffic figures are for all flights originating or terminating in the contiguous US between 1 April 1990 and 31 March 2018

Table 15: List of airlines used as the categories for the 'airline' instrument.

Transition period	Airline retaining code and name		Airline made defunct	
1998–1999	FL	AirTran Airways Corporation	J7	Valujet Airlines
1999–2001	AA	American Airlines	QQ	Reno Air
2000–2002	AC	Air Canada	CP (1)	Canadian Airlines
2002–2003	AA	American Airlines	TW	Trans World Airways
2007–2009	US	US Airways	HP	America West Airlines
2009–2011	DL	Delta Air Lines	NW	Northwest Airlines
2011–2013	BW	Caribbean Airlines	JM	Air Jamaica
2011–2013	UA	United Air Lines	CO	Continental Air Lines
2014–2016	WN	Southwest Airlines	FL	AirTran Airways Corporation
2015–2017	AA	American Airlines	US	US Airways

Table 16: List of mergers and acquisitions where subsequent traffic was classified under a single code.

B.2 Aircraft classes

Table 17 lists the categories use to construct the 'aircraft class' instrument. The aircraft classes are based on the *Aircraft Type Group* variable defined by the BTS, which separates aircraft based on their types and numbers of engines. As large proportions of traffic are conducted using aircraft with 2, 3, or 4 jet engines, these groups are broken down further by the numbers of seats, which proxy for the sizes of the aircraft.

Index	Aircraft class	Number of flights	Number of pass. ('000)
0	Piston, 1-Engine / Combined Piston / Turbine	3,851,680	7,739
1	Piston, 2-Engine	3,150,023	13,425
2	Piston, 3-Engine / 4-Engine	2,895	0
3	Helicopter / STOL	32,278	1,012
4	Turbo-Prop, 1-Engine / 2-Engine	32,611,294	634,494
5	Turbo-Prop, 4-Engine	87,443	2,130
6.1	Jet, 2-Engine, 1-99 seats	57,873,883	2,504,266
6.2	Jet, 2-Engine, 100-149 seats	94,846,142	8,404,327
6.3	Jet, 2-Engine, 150-199 seats	40,434,412	5,204,068
6.4	Jet, 2-Engine, 200+ seats	9,403,550	1,801,822
7.1	Jet, 3-Engine, 1-99 seats	33,902	1,097
7.2	Jet, 3-Engine, 100-149 seats	9,998,032	901,465
7.3	Jet, 3-Engine, 150-199 seats	8,052	845
7.4	Jet, 3-Engine, 200+ seats	2,516,986	488,374
8.1	Jet, 4-Engine / 6-Engine, 1-99 seats	1,314,248	60,278
8.2	Jet, 4-Engine / 6-Engine, 100-199 seats	443,626	63,430
8.3	Jet, 4-Engine / 6-Engine, 200-299 seats	797,216	157,637
8.4	Jet, 4-Engine / 6-Engine, 300-399 seats	2,683,572	722,151
8.5	Jet, 4-Engine / 6-Engine, 400+ seats	446,012	139,549

Note: the traffic figures are for all flights originating or terminating in the contiguous US between 1 April 1990 and 31 March 2018

Table 17: List of aircraft classes used as the categories for the 'aircraft class' instrument.

C First-stage results for the CBSA-level estimation

Table 18 presents the first-stage coefficients from the estimation of (3). This corresponds to the estimation results displayed in Columns 2, 4, 6, and 8 of Table 4. The F-statistics demonstrate that the instruments clearly satisfy the relevance condition. Furthermore, as the coefficients on the two instruments are both positive and significant at the 1% level, they both contribute to explaining the variation in the growth in local air traffic.

	(1) OLS	(2) OLS	(3) OLS	(4) OLS
$\ln(A_{m,t+1}) - \ln(A_{m,t})$ ('airline' instrument)	0.439^a (0.057)	0.439^a (0.056)	0.439^a (0.056)	0.441^a (0.056)
$\ln(A_{m,t+1}) - \ln(A_{m,t})$ ('aircraft class' instrument)	0.317^a (0.069)	0.320^a (0.069)	0.318^a (0.068)	0.315^a (0.069)
$\ln(A_{m,t})$	-0.168^a (0.012)	-0.170^a (0.012)	-0.171^a (0.012)	-0.170^a (0.012)
$\ln(pop_{m,t})$	0.160^a (0.037)			
$\ln(lf_{m,t})$		0.191^a (0.038)		
$\ln(emp_{m,t})$			0.219^a (0.038)	
$\ln(gdp_{m,t})$				0.140^a (0.028)
R^2	0.33	0.33	0.33	0.33
F-statistic on the instruments	64.86	66.72	68.25	66.89

Note: the dependent variable in each regression is the change in air traffic in CBSA m between t and $t+1$; 4,833 observations for each regression, representing 179 CBSAs; robust standard errors clustered by CBSA in parentheses; *a*, *b*, *c* denote significance at 1%, 5%, 10%; number of departing flights used as the measure of airport size; all regressions include CBSA and year fixed effects

Table 18: First-stage estimation of the relationships between the instruments and the change in air traffic in metropolitan area m.

D Elasticities of trip cost to the costs of using individual airports

The framework of Allen and Arkolakis (2019) allows neat analytical solutions for the degrees to which the costs of using specific parts of the network affect the mean cost of travel. The difference is that in this version of the model the airport infrastructure affects the costs of travel at the nodes of the network, whereas their model is of a road network where the infrastructure determines the cost of using each of the links. The effects of μ_i, μ_j, and ν_k on the mean cost of travel between i and j can be derived from (11) and (12) and expressed as elasticities:

$$\frac{\partial \ln \tau_{i,j}}{\partial \ln \mu_i} = \frac{\partial \ln \tau_{i,j}}{\partial \ln \mu_j} = 1 \tag{23}$$

$$\frac{\partial \ln \tau_{i,j}}{\partial \ln \nu_k} = \left[\frac{\tau_{i,j}}{\tau_{i,j}^k}\right]^{\psi} \tag{24}$$

where $\tau_{i,j}^k$ is the hypothetical mean cost of travel between i and j if travel were restricted to routes with connections in k.