Planes, Trains, and Co-Opetition: Evidence from China*

Marco Shaojun Qin
University of Minnesota

Job Market Paper

July 2, 2017

Work in Progress. Current version:
https://carlsonschool.umn.edu/faculty/marco-qin/

Abstract

The roll-out of the high speed train (HST) in China during the past decade offers travelers a realistic choice of rail versus air on overlapping routes. While the HST competes for traffic it also feeds traffic to the non-overlapping air routes that are connected to the HST network. To analyze the effects of the roll-out of the high speed train on the airline industry, I use a unique hand-collected dataset on airline networks and the timing of the HST roll-out. I exploit the variation in the timing of the HST roll-out across different regions to present descriptive evidence that suggests that the HST may have positive spillover effects on the airline industry. I then estimate a structural dynamic oligopoly model of air route entry and exit over time. The model allows me to quantify the negative and positive spillovers of the HST network roll-out on the airlines' route networks and to understand how airlines respond strategically to the roll-out of the HST by re-positioning themselves (in this case, through new network configurations). I find evidence of heterogeneous spillover effects of HST on air route networks which depend on the routes' characteristics and on their interaction with the HST network. Conditional on the airlines' decisions, the overall contribution of the positive spillovers from HST exceeds that of the negative spillovers, suggesting that airlines reposition their networks to avoid direct competition while taking advantage of the complementarity in routes.

Keywords: entry, dynamic games, continuous time, intermodal substitution and complementarity, network competition

*Ph.D. Candidate in Marketing, Carlson School of Management, University of Minnesota, E-mail: qinxx113@umn.edu. I am grateful to my advisors George John and Maria Ana Vitorino, and my committee Pradeep Chintagunta, Amil Petrin, and Linli Xu for their guidance and support. Thanks also to Thomas Holmes, Yi Zhu, and the participants of the Marketing Brown Bag and the Applied Micro Workshop at the University of Minnesota for their helpful comments and suggestions. I specifically thank Peter Arcidiacono for clarifications related with his work and helpful responses. This research was supported by computational resources provided by the Minnesota Supercomputing Institute (MSI) and the U-Spatial Department at University of Minnesota. All errors are my own.

1 Introduction

China has the fastest growing high speed train (HST) network in the world. During the past decade, China has built (or upgraded) over 20,000 km (12,500 miles) of high speed railways, more than the rest of the world combined . According to the Chinese government's plan, by 2030, the HST network is expected to expand to 45,000 km (30,000 miles) of railroad tracks, connecting most of the major cities in the country (see Figure 1).[1]

The roll-out of the HST, which is government-owned and operated, offers challenges to domestic air carriers' operations. HST service and air travel are often acknowledged to be reasonable substitutes and, by the end of 2015, almost 50% of air routes in China were also served by HST.

At the same time, the HST service might offer an opportunity for air carriers. Similar to the historical evidence on inter-modal effects between truck and rail freight traffic in the U.S., air carriers and HST might feed traffic to each other, thus increasing the attractiveness of both operations. Consistent with this, China's 13th five-year plan emphasizes the need to improve the connectivity between different modes to transportation. Also, online travel brokers often offer online bookings that combine flight and HST segments. Notwithstanding these examples of efforts aimed at exploiting inter-modal complementarities, there is no empirical evidence about the relative size of the substitution and/or complementarity effects between HST and flight services.

In this paper I study the effects of the roll-out of the HST on the airline industry in China. I empirically assess and quantify the impact of negative (substitution) and positive (complementary) spillovers of HST on air carrier profits and, in particular, the implications of these spillovers for the air carriers' route networks.

I use a unique hand-collected dataset on airline networks and on the timing of the HST roll-out from 2006 to 2016 to empirically test whether and how the presence of HST affects airlines' route configurations and profits. I use both descriptive and structural methods to analyze this issue and find evidence of both negative and positive spillovers from HST to airlines.

My research contributes to the ongoing policy debate as to whether HST acts solely as a substitute to existing modes of transportation (in this case air) or whether it can act as a complementary mode of transportation. By studying the roll-out of the HST in China, we can better understand what would be the consequences in terms of airlines' response to the HST in other locations which are considering the introduction of HST. More importantly, quantifying the negative and potential positive spillovers from HST to air travel (and their net effect) allows us to understand whether the resistance of the airline industry regarding the introduction of HST in some regions (such as California) is warranted. More generally, my approach contributes to the study of market entry and to the analysis of product assortment decisions in which there are substitution and complementarities in demand across firms and product categories.

Previous work has studied firm's entry and location choice in settings in which there are positive

[1]Source: http://www.economist.com/news/china/21714383-and-theres-lot-more-come-it-waste-money-china-has-built-worlds-largest

and negative spillovers between firms.[2] I extend this work by considering spillovers in settings in which it is important to consider inter-connected markets comprising a network. For example, the entry of Walmart has been shown to help small retailers while having large negative effects on regional and national chains (Arcidiacono et al. 2016). *Didi*, a Chinese "ride-hailing" service is generally considered an alternative to the traditional taxi service, but it is possible that this platform may feed traffic to traditional taxi carriers for certain routes within their overall network.[3] Also, traditional brick-and-mortar stores seem to be adapting to reap the positive spillovers associated with online commerce as more consumers today are "webrooming" (browsing online before making an in-store purchase) than "showrooming" (going to a store to make their selection and then searching online for the best price) across a variety of retail categories.[4]

If we consider that an air carrier's network is an "assortment of routes," this paper also informs the literature on product assortment. Most of the extant work on product assortment treats the individual products offered by a firm as substitutes, so firms' (re)positioning of products is driven mostly by incentives to soften competition and to avoid cannibalization within their assortments.[5] In the case of firms that offer complementary products (such as the case of airlines with naturally interconnected routes), firms should also consider the positive spillovers in demand that are generated among theirs and their competitors' products.

To analyze the net effect of the roll-out of the HST on the airline industry, I hand-collect daily flight information for the four major passenger airlines and their subsidiaries in China. I use this data together with detailed information on the timing of the HST roll-out.[6] The data covers over two thousand city-pairs (i.e., routes) for a period of ten years. I exploit the variation in the number of flights offered and in the airlines' route choices over time and across regions, together with the evolution of the HST's route network to identify how airlines' route choices respond to the presence of the HST.

I present descriptive evidence that suggests that the HST may have both negative and positive spillover effects on the airline industry. Then, I build and estimate a structural dynamic oligopoly model of airlines' route entry and exit decisions over time that allows me to estimate the magnitude of the spillover effects.

First, descriptive panel-data regressions that model the presence or absence of air service in a given route, and which control extensively for route and airline characteristics, are consistent with HST acting as a substitute to air travel in overlapping routes while being complementary to air travel

[2]For example, Vitorino (2012) studies the negative and positive spillovers among stores of different formats in the context of shopping centers, Yang (2012) studies the channel (fixed costs versus variable profits) through which spillovers affect firms' decisions in the fast-food industry, and Datta and Sudhir (2011) look at the trade-off between co-location and spatial differentiation in the context of big-box grocery stores.

[3]The main difference between *Didi* and *Uber* is that the *Didi* platform allows one not only to request the services of drivers that have signed up with *Didi*, but also to book traditional taxi-cabs.

[4]http://www.chainstoreage.com/article/how-smart-digital-strategy-can-make-physical-retail-store-prime-asset

[5]Examples include Draganska, Mazzeo, and Seim (2009), Eizenberg (2014), Jeziorski (2014a), Jeziorski (2014b), Sweeting (2010), and Sweeting (2013).

[6]The four major airlines in China (including their subsidiaries) account for almost 90% of the domestic air passenger traffic in China.

in connected routes. Further, the results from these regressions suggest that there are heterogeneous spillover effects of HST on air route networks which depend on the routes' characteristics, including the length of the route and whether a route is government regulated. A matching algorithm, that looks at a route as being "treated" or "not treated" depending on its level of connectivity to the HST, and which alleviates potential functional form concerns associated with the panel-data regressions, provides similar insights.

Second, in order to account for possible unobservable differences across routes, such as demand or cost shocks, which the panel-data regressions and the matching approach cannot account for, I use a difference-in-differences regression.[7] I exploit the presence of HST on a focal route and create "control" routes that are connected with the two airports at the ends of the focal route, but which do not overlap with HST. Here, the identification assumption is that both demand side and supply side shocks in the focal route are captured by one of the control routes. Thus, by comparing air service on the control routes against the focal route before and after the treatment (controlling for other observable differences), we can identify the impact of HST. The results from the difference-in-differences analysis are consistent with the presence of both positive and negative HST effects on air service, and provide additional insights. While the presence of HST seems to have a negative impact on shorter air routes that overlap with HST service, for longer routes, the effect of HST seems to be positive, thus suggesting the existence of market expansion effects.

Finally, and because the descriptive evidence does not explicitly model airlines' entry and exit decisions, I construct and estimate a structural dynamic oligopoly model of airlines network configurations. This approach allows me to account for the endogeneity of firms' entry end exit decisions while contributing to the understanding of the underlying mechanism through which spillovers occur. Also, by linking airlines observed decisions to unobserved latent profits I can quantify the negative and positive spillover effects from the HST to airlines and among the different airlines.

A structural model of airlines' network decisions poses several challenges. First, in this setting, there are multiple players and each firm's route decisions are dependent on their own and other firms' networks. Second, airlines' decisions are inherently dynamic (i.e., firms are forward-looking) because there is a cost associated with entry in each route. Further, the market conditions (demographics and train network) change over time leading airlines to have to reassess their network often, which results in frequent airline entry and exit from the numerous city-pair routes. Allowing for the forward-looking behavior of firms with a very large state space invalidates the use of traditional approaches to estimate dynamic games of entry and exit (e.g., Ericson and Pakes 1995, Bajari, Benkard, and Levin 2007, and Aguirregabiria and Mira 2007). To overcome these challenges I use a new framework (Arcidiacono et al. 2016) for estimating and solving dynamic discrete choice models in continuous time (as opposed to the most commonly used discrete-time framework) that can be applied to dynamic games.

I find strong evidence of both negative and positive spillover effects from the HST on air carrier

[7] Goolsbee and Syverson (2008) use a similar methodology in their study of how incumbents respond to the threat of entry by a competitor in their study of entry of Southwest Airlines.

routes. The effects depend on the routes' characteristics and on their interaction with the HST network. Specifically, I conclude that the HST is a strong competitor of air travel, especially in shorter routes. The substitution effects of HST relatively to air travel dissipate for longer routes, however. More connections to HST lines bring positive spillovers to air travel; these effects amount to roughly one fifth of the negative spillovers associated with the overlap with short routes. Interestingly, the interaction of HST presence and connectivity to HST train lines is negative, suggesting a negative moderating effect of HST presence on the spillovers associated with a higher HST connectivity.

Further, a profit decomposition analysis shows that airlines enjoy more positive spillovers from HST and less negative spillovers in newly entered routes than in well-established ones. At the same time, airlines tend to exit routes where the negative spillovers from HST are very large. On average, conditional on the observed airlines' route decisions, the HST substitution effects (i.e., negative spillovers) amount to approximately 6.7% of airlines' profits, while the complementarity traffic effects (i.e., positive spillovers) represent about 8.5% of airlines' profits.

Taken together, the evidence found is consistent with a positive net impact of the HST on the airline industry over the sample period and suggests that airlines strategically reposition their networks to avoid direct competition while taking advantage of the complementarity in routes.

The rest of the paper is organized as follows: Section 2 reviews the literature. Section 3 describes the industry background. Information on the data, and descriptive evidence are provided in Sections 4 and, 5 respectively. Section 6 presents the structural model, while Section 7 describes the structural model's empirical specification. The estimation strategy and results for the structural model are described in Sections 8 and 9. Section 10 concludes.

2 Literature

This paper is related to several streams of literature. First, this paper contributes to the transportation economics literature that studies intermodal competition. The earlier literature generally views different modes of transportation as substitutes, and has focused mostly on the competition between trains and trucks. For example, Oum (1979a) and Oum (1979b) derive a demand model for freight transportation and study the substitution patterns between trains and trucks in terms of price, quality and distance. Other papers look at the supply side and study how truck and train prices change in response to competition or other policies that affect competition (MacDonald 1987, Wilson, Wilson, and Koo 1988). Most of these papers, however, take the train or truck routes as given and abstract away from the possibility of positive spillovers between modes of transportation. More similar to this paper, Viton (1981) studies price and quality competition between trains and buses in the San Francisco Bay Area. Using demand and cost data, the author numerically solves for the market equilibrium and finds that both modes of transportation can survive as differentiated services.

The only two papers that study the benefits of intermodal transportation coordination are Roberts (1969) and Friedlaender and Harrington (1979). Roberts (1969) uses an analytical model to demonstrate that gains from coordination between train and truck transportation are best achieved through market-free prices and not through the use of regulation. Using both theory and empirics, Friedlaender and Harrington (1979) show that the lack of intermodal coordination in the U.S. during the 1970s was caused by restrictions on intermodal ownership. Neither of the previous two papers address the question of whether coordination affects the transportation networks.

Some recent literature on transportation policy has looked at the competition between high speed trains and airlines, but this literature tends to look at these two modes of transportation strictly as substitutes (e.g. Dobruszkes 2011, Behrens and Pels 2012, Clewlow, Sussman, and Balakrishnan 2014). Most papers that discuss the coordination between airline and railway systems do so based on anecdotal evidence (e.g. Givoni and Banister 2006 and Clewlow, Sussman, and Balakrishnan 2012) or theory models (e.g. Jiang and Zhang 2014 and Xia and Zhang 2016). Empirical studies on the positive spillovers between airlines and HST are sparse. Albalate et al (2014) is the only paper that empirically studies this question. Specifically, this paper focuses on the HST networks in European countries and studies the impact of the introduction of HST on service frequency and number of seats offered by airlines. They find that airports with an on-site HST station are less negatively affected by the introduction of HST, which is suggestive of positive spillovers from HST to airlines. However, this paper is mostly descriptive and does not study how airlines reposition their route services from a network perspective in response to the introduction of HST.

This paper also contributes to the empirical industrial organization literature on firm entry, by studying firms' entry decisions in a setting where markets are interconnected. Earlier literature on the airline industry has explored the influence of airline hubs on firms' entry decisions (e.g., Berry 1992, Ciliberto and Tamer 2009), but the possible positive spillovers from competitors are less studied. The coexistence of both negative and positive spillovers from competitors is similar to the idea of the agglomeration-competition phenomenon in the retail industry, which has been discussed in Vitorino (2012), Datta and Sudhir (2011) and Yang (2012), for example.[8] Vitorino (2012) studies stores' entry decisions in shopping centers and finds strong evidence of both positive and negative spillovers among stores of different formats. Datta and Sudhir (2011) develop a model of entry and location choice games among stores and use detailed store level data and spatial zoning data to disentangle the trade-off between co-location and spatial differentiation. Yang (2012) studies the fast food industry and explores the channel (i.e., variable profits versus fixed costs) through which the spillovers affect firms' entry decisions.[9] This paper contributes to this stream of literature by extending the entry game to a network setting, where negative and positive spillovers may exist across markets and industries.

As location matters in an industry with interconnected markets, this paper is also related to the empirical entry literature on spatial competition.One stream of literature on this topic focuses on the spatial differentiation among stores (e.g., Seim 2006, Zhu and Singh 2009, Datta and Sudhir 2011,

[8]For papers that have a similar question but that use non-structural methods, see Clapp, Ross, and Zhou (2016) and Shen and Xiao (2014).

[9]Some other papers explore the spillovers from a demand perspective, e.g., Sen, Shin, and Sudhir (2011) and Murry and Zhou (2016).

Orhun 2013). For example, Seim (2006)'s pioneer work uses an incomplete information framework to study firms' location choices in the video rental industry. Her framework is extended by Zhu and Singh (2009) and Orhun (2013) to allow for firm heterogeneity and location specific unobservables respectively. Another stream of the literature focuses on the "chain effect" of firms with multiple stores (e.g., Jia 2008, Holmes 2011, Aguirregabiria and Ho 2012, Nishida 2014). For example, Jia (2008) studies a location choice game between Walmart and Kmart and allows for positive spillovers among nearby stores of the same company. Nishida (2014) extends Jia (2008)'s framework to allow for multiple stores in the same market and applies it to the convenience store industry in Japan. These two papers use a static oligopoly model with two players. Holmes (2011) studies the dynamic network decisions of Walmart, but abstracts away from the competition between the firm and other chain stores. Aguirregabiria and Ho (2012) study the airline industry in the U.S. and account for the dynamic decisions of firms and interconnectedness of markets, which is closer to this paper. There are two key differences though. First, while their focus is to explore the sources of benefits from the hub and spoke business model, this paper emphasizes the spillovers from different modes of transportation. Second, the structural model in this paper allows for direct strategic interactions among airlines in the same route, while their model does not.

To the extent that airlines can be regarded as multi-product firms that differentiate themselves through their route network configurations, this paper is also related to the literature on product assortment decisions (e.g., Draganska, Mazzeo, and Seim 2009, Sweeting 2010, Sweeting 2013, Jeziorski 2014a, Jeziorski 2014b and Eizenberg 2014). For example, Draganska, Mazzeo, and Seim (2009) study the competition between firms in both product choices and prices and find that incorporating product assortment decision as a strategic variable is important for policy simulations. Eizenberg (2014) estimates a model of supply and demand in the PC industry in which both price and PC types are endogenously determined, and then uses the model to assess the welfare implications of the introduction of new upstream components. Other papers focus on the "repositioning" aspect of product assortment decisions and employ structural models to assess firms' product strategies in response to some change in the market structure. For example, Sweeting (2013) studies the impact of fees for musical-performance rights on radio station formats and finds that the impact of such a policy change is larger in the long run than in the short run. Jeziorski (2014b) develops a dynamic model to estimate the cost efficiency of mergers in the U.S. radio industry while accounting for the repositioning of the products (radio station) and merger choices. In these previous papers, the major motivation behind the product (re)positioning is either to avoid competition or to reduce cannibalization. My setting is different in that there could be complementarity among products of different firms. This difference may play a key role in determining firms' product assortment decisions.

3 Industry Background

3.1 The Airline Industry in China

Figure 2 shows the major airline companies in China as well as their parent firms. Although there are more than 30 airline companies in the industry, the majority of them are subsidiaries of the top four airline companies in China, namely Air China (CA), China Southern Airlines (CZ), China Eastern Airlines (MU) and Hainan Airlines (HU). All four companies are publicly traded firms. Including the subsidiaries, from 2006 to 2016, the top four airlines together have a market share of about 90% in terms of total number of flights, and cover more than 94% of the existing air routes in China.

========================

Insert Figure 2 about here

========================

The airline industry is growing at a steady rate. From 2006 to 2016, the total number of flights increased by about 87%, and the total number of city-pairs served by the industry[10] has increased by 57%. This corresponds to an annual growth rate of 8.1% in the total number of flights and 5.7% in the total number of routes provided.

Until 2006, the airline industry in China was heavily regulated by the Civil Aviation Administration of China (CAAC). An airline's flight decisions (i.e. adding or dropping flights) need to be examined and approved by the administration before taking effect. The process takes about 60 days and has many restrictions on airlines' performance such as seat occupancy rate, on-time rate and customer satisfaction.

In 2006, the central government issued a new policy, with the intent of reducing the regulatory oversight in the airline industry. Since then, airline routes can be classified into two types: regulated and nonregulated. For the nonregulated routes, airlines are not required to request a permit to operate on (only registration is needed). Regulated routes involve the airports of Shanghai, Beijing and Guangzhou as one of the end points. These airports are very crowded airports, and the government keeps these routes regulated mainly for traffic-control purposes. As a result, airlines still need to apply for permits to add flights on these routes.[11] Procedure is greatly simplified though. For example, the government's response time for the application has reduced from 60 days to 30 days, and there is no longer a constraint on seat occupancy rate for increasing flights. There are some important exceptions regarding the regulations with these main airports. If an airline is headquartered in one of the three cities mentioned above, then for the airline, the routes that connects its headquarter city and other cities (except for Beijing, Shanghai and Guangzhou) are exempt from the regulation.

[10] Henceforth, I will use routes, markets and city-pairs interchangeably.

[11] The decision to drop a flight does not require any permits.

3.2 The Railway Industry in China

The railway industry in China is operated by a monopoly government-owned enterprise called China Railway. Unlike the firms in the airline industry, this company peruses social-welfare related objectives and is not a profit maximizing firm.[12] In 2004, the government issued a long-term network plan for railways, aiming at building a widespread network of HST in China. The government has detailed plans regarding the railway network that should be completed by 2020. Specifically for the HST, there are two major plans: one is to build about 12,000 km (about 7500 miles) of HST rails. Another plan is to upgrade about 16,000 km (about 10,000 miles) of existing rails so that they can accommodate trains of higher speed.

There are two types of HST. The "fast train" achieves a maximum speed of 250 km/h (about 155 mph). The "bullet train" can achieve a speed of 350 km/h (about 215 mph). Although the networks for these two types of HST overlap, the trains operate on different rails. In terms of pricing, for the same route, a bullet train ticket is about 60% more expensive than a fast train ticket.[13]

The first bullet train route was established in 2008, connecting Beijing and Tianjin. Ever since, the network has expanded rapidly. By 2016, about 20,000 km (about 12,400 miles) of high speed train rails has been built or upgraded, more than the rest of the world combined.

Unlike the airline industry, where companies can modify their routes relatively frequently, the expansion of route network of HST is for the most part predetermined.[14] Because of the social welfare objectives of the railway industry, the operation of different routes is not exclusively based on profit maximization motives. For example, until 2016, the high speed rail between Shanghai and Beijing was the only high speed train line that was able to break even.[15]

3.3 Competition between Airlines and High Speed Trains

The introduction of high speed trains has brought some disruption to the airline industry. Many news articles report that airlines reduce the number of flights or even exit the market to avoid competition with HST.[16] Also, there has been an increasing overlap in the route network between HST and airlines. As Figure 3 shows, from 2007 to 2015, despite the growth in the number of routes served by the airline industry, the proportion of routes which face direct competition from HST has increased. By the end of 2015, more than 50% of the airline routes are also served by fast trains, and more than 25% are also served by bullet trains.

[12] To see a complete description of the company's goals, please refer to http://www.china-railway.com.cn/en/aboutus.

[13] This number is calculated based on the price data (for a seat in the economy class) collected on November 15th 2016 for routes where both fast train and bullet train operate.

[14] However, there is some uncertainty associated with when each specific HST route will be finished and regarding whether a specific route will be served by HST.

[15] Source: http://view.163.com/special/resound/chinahsr20160721.html (in Chinese).

[16] For example, see http://news.xinhuanet.com/fortune/2011-04/12/c_121293247.htm (in Chinese).

==========================

Insert Figure 3 about here

==========================

According to Civil Aviation Big Data, a Chinese based consulting company, consumers' preferences for a mode of transportation is a function of travel time. Figure 4 shows the relation between door-to-door travel time and travel distance for each mode of transportation. As we can see from the figure, for short routes (<600 km), the door-to-door time of HST is in fact shorter than that of airplanes. This is because, although airplanes are faster than HST, airports are usually located far away from the downtown area of the city. What's more, high-speed trains require less security check and are more punctual. All above add to the attractiveness of HST. Together with the fact that the ticket prices of HST are in general lower than those of flights, we can infer that HST has more strategic advantage in shorter routes.

==========================

Insert Figure 4 about here

==========================

3.4 Anecdotal Evidence of Demand for Intermodal Transportation

While much attention has been devoted to discussing the competition between HST and airlines, there is some anecdotal evidence of positive spillovers between these two modes of transportation. Such spillovers are explored not only by airlines by also by other firms. For example, since the introduction of the HST, both airports in Shijiazhuang and Tianjin began to provide various services to facilitate the transportation between airports and railway stations, such as free shuttle bus and lounge rooms for commuting passengers. In 2016, the Shijiazhuang airport reported 40,000 passengers utilized intermodal transportation,[17] which accounted for about 5.5% of the total passenger volume of the airport in that year.[18,19]

Travel agencies also benefit from the introduction of HST because they can now use the network of both modes of transportation to offer travel packages with more variety. For example, in 2013, five cities in the north eastern part of China, (Dalian, Shenyang, Changchun, Haerbin and Changchun) jointly introduced a travel package which combines the tourism resources of the five cities. Their slogan reads "fly to Dalian, and take a HST to tour around the North Eastern China". The HST makes the travel between the five cities more convenient and increases the attractiveness of the combined package, which in turn brings more passenger traffic to the flights. In 2015, *Ctrip*, one of the largest online travel platforms in China (the equivalent of *Priceline* in the U.S.), launched a

[17]Source: http://news.ifeng.com/a/20161227/50482206_0.shtml (in Chinese).

[18]Source: http://www.caac.gov.cn/XXGK/XXGK/TJSJ/201702/t20170224_42760.html (in Chinese).

[19]The surge in the volume of passengers with intermodal transportation at Shijiazhuang Airport is largely due to the collaboration between the airport and Spring Airlines, a local privately owned airline company.

service called "air-rail combo", which allows consumers to mix and match flights and HST to make more convenient travel arrangements.

There are also examples that illustrate how the airline and train industry also attempt to explicitly benefit from intermodality. For example, in April 2012, China Eastern Airlines signed a contract with the Shanghai Railway Bureau and started offering tickets that cover routes for both flights and trains often at a lower bundled price. This practice was soon followed by Hainan Airlines (July 2012) and Air China (Dec 2012) that started offering a similar service in Shanghai and Haikou, respectively.

Despite their attempts, the impact of such collaboration seems quite limited. For example, in 2015, the China Eastern Airlines sold about 12,000 tickets for the intermodal services.[20] This accounts for only about 0.6% of total passenger traffic of the China Estern Airlines in the HongQiao airport.[21] The lack of success of this program may be due to multiple factors such as inadequate advertising, difficulty in purchasing the intermodal tickets and poor customer services. The fact that in the Frankford Airport in Germany, 17.5% of passengers utilize both airplane and long distance trains[22] is also suggestive of the importance of intermodal transportation.

4 Data

4.1 Data Sources

I create an original dataset of flight schedule and the relevant market characteristics (presence of HST, city population, route length) from multiple sources. The dataset spans a 11-year period from January 1st 2006 to Dec 31st 2016. The flight schedule information was collected from a website that provides historical flight data. A record in the dataset includes the date, airline, flight number, departure and destination cities. The information about the HST network comes from government websites and news reports, from which I collect the exact dates when cities were connected with HST.[23] I further refer to the *China City Statistical Yearbook* to collect population and GDP data for cities in China.[24] Finally, data for city coordinates come from Google Maps and is used to calculate the air distance between cities.

4.2 Data Cleaning and Sample Selection

Selection of Airlines: I look at the airlines on a parent-company level so that an airline and its subsidiary airlines are regarded as one company. I focus on the competition between the top four airlines only, as together they account for about 90% of market share in number of flights.

[20]Source: http://finance.sina.com.cn/hy/20160219/150024312707.shtml (in Chinese).

[21]Source: https://www.cybersource.com/content/dam/cybersource/zh-APAC/documents/China_Eastern_Airlines_Case_Study_SC.pdf (in Chinese).

[22]Source: https://static.fraport.de/ONLINE/zdf/zadafa_e_2015/.

[23]I supplement the HST dataset with data scraped from *www.12306.com*, the official website for purchasing train tickets in China. I use the data to double check whether a route is served by HST.

[24]Unlike the first two data sources, the GDP and population data is at a yearly level. I therefore treat the GDP and population as constant within a year.

Selection of Markets: I define a route to be a non-directional city pair. I focus on the routes that connect the top 68 cities in China in terms of airport passenger volume.[25] In case that two airports belong to the same city, I aggregate the airports together. The major reason of doing so is to avoid the complication arisen from the competition between two airports from the same city.

Selection of Flights: I define a fight on a round-trip level such that if an airline company provides a flight from city A to city B, the same flight is provided from city B to city A.[26] I focus on flights that are provided regularly, so that seasonal flights only provided for occasions such as Chinese Festival, Christmas are not included. I also exclude infrequent flights, as these flights are mostly between non-popular cities and firms might provide such flights simply for the local government subsidies.[27] Specifically, I use two criteria to select flights: 1. The flight must be provided for more than one year and 2. The frequency of the flight must be at least once every two days.

A potential issue of the first criteria is that some flights might last for more than one year, but still be excluded because of the limitation in the time span of the data. For example, if a flight (operated on daily basis) starts from April 2005 and ends on May 2006, then it has lasted for 13 months and should be included in my data. However, since I only have data starting from 2006, I don't have enough information to decide whether to include this flight. Similarly for flights that were introduced less than one year before the end date of my dataset. This creates some data selection problems because some important flights might be dropped. To deal with this issue, I exclude the first and last year of the observations in my dataset, so that each flight in the remaining dataset satisfies the two criteria for sure.

4.3 Summary Statistics

For clearer illustration, I aggregate data from the daily level to the yearly level. After the aggregation, I have 82008 observations of flight decisions across four airlines, 2278 routes and nine years (4*2278*9=82008). Table 1 provides the summary statistics of the data. On average, there are 0.5 airlines that provide one flight in each route. However, the medians of these two variables are both zero, indicating that the airlines and flights are not evenly spread across routes, but tend to concentrate on a smaller set of city-pairs.

For each airline, I distinguish between two types of decisions: one is the adjustment without entering or leaving the route, and the other is the decisions which involve entering a new route or exiting an existing route. There are adjustments on both levels. On average, about 0.05 (0.03) airlines enter (exit) a route within a year, and 0.06 (0.08) flights are added (dropped) to a route in a year. For HST, on average, fast trains are present in 11% of the routes while the corresponding number for bullet trains is 3% in a year. These numbers are quite comparable to the average presence of airlines and demonstrate that there is a fair share of airline routes that have HST

[25] This rank is based on the airport passenger volume in year 2015. Altogether, the cities have 70 airports, which account for more than 95% of passenger volume in the airline industry in 2015.

[26] This is usually reflected in the closeness in the flight number for the two trips. For example, the flight number from city A to city B is 1231, then the return flight number is 1232.

[27] Source: http://news.sohu.com/20151201/n429001116.shtml (in Chinese).

services (especially given the fact that there were no HST until 2007).

==========================

Insert Table 1 about here

==========================

For the city-pairs and years studied, population grew at an average rate of 1% per year and GDP has grown at a more accelerated rate of 16% per year. The average length of routes is about 1500 km.

Figure 2 summarizes how the HST presence changes over time. Special attention needs to be paid to year 2013 through 2015, when there was a large increase in the stock of the HST network. Specifically, in 2013, the average presence of bullet train increased from 2% to 7% of the markets. There is also a spike in the presence of fast train in 2014, in which the average presence went from 11% to 19% of the markets. Starting from the same year, despite the increase in the average number of airlines/flights in each route, the average number of exits as well as decrease in number of flights has gone up since 2013. This is suggestive that airlines adjust their capacity in response to the introduction of HST.

==========================

Insert Table 2 about here

==========================

5 Descriptive Evidence

In this section I provide some descriptive evidence regarding the entry patterns of airlines and HST, and explore how the HST network structure influences airlines' entry decisions.

5.1 Factors Influencing Flight Decisions

I first explore the factors that could potentially affect airlines' flight decisions. I am especially interested in the following variables: the average number of flights provided by an airline, entry probability, exit probability, probability of increasing flights, and probability of reducing flights. The measure of the first variable is straightforward. To calculate the entry probability for a group of routes, I first count the total number of entries, defined by the case where an airline provides flight services in a route in the current but not previous year. I then divide this number by the total number of observations where entry is possible. Exit probability, probability of increasing flights, and probability of decreasing flights are similarly defined. One thing to note is that entry and flight increase are defined as mutually exclusive decisions in the sense that if the number of flights of a given airline in a route changes from 0 to a positive number, it is regarded as an entry, but not a flight increase. On the other hand, flight increase is only conditional on an airline having flights in the previous period. There is a similar distinction between exit and decrease in the number of flights.

Market Characteristics

Table 3 reports the break down of airlines' flight decisions by average gross domestic production (GDP) of city-pairs across years. Specifically, all observations are divided into five quantiles by the city-pair's average GDP. Perhaps not surprisingly, the higher the city pair average GDP, the more flights are provided and the higher the likelihood of entry and increase in number of flights. However, airlines' exit and flight decrease decisions also follow a similar pattern, with higher probability being associated with routes with larger average GDP in a year. This is probably due to some other factors such as competition, as the more attractive is the route, the more airlines will choose to enter.

Similar patterns can be found for the number of flights as we break down routes by population and income per capita, as shown in tables 4 and 5.

========================

Insert Table 3, 4 and 5 about here

========================

Table 6 explores the relation between route length and the number of flights. Routes of small and median length are associated with a higher number of flights and a higher probability of entry and increase in the number of flights. There is also a higher probability of exit and number of flights reduction for shorter routes. Note however that, while the largest number of flights and probability of entry/flight increase is found at the second quantile, the highest probabilities of exit and decrease in number of flights are for routes with the shortest length. This could be because airlines face more competition from HST in shorter routes.

========================

Insert Table 6 about here

========================

Airlines' Network Structure

A typical feature of the airline industry is that markets/routes are interdependent. The decision of an airline to enter or exit a given route is affected by its operations in other routes, especially those ones that are connected to that route. I explore how the presence of an airline in both ends of a city-pair affects the probability that the airline enters that route. I run a probit regression in which I regress an airline's decision to enter a given route in year t on the categorical variable "airport presence", which captures the number of end points of the route which the airline operates in year $t-1$, while controlling for year and airline fixed effects.

Table 7 presents the marginal effects of airport presence on airlines' entry probabilities. Consistent with the results from Goolsbee and Syverson (2008), having operated in the end points of a route in the previous year is associated with a higher probability of entering that route, and the marginal effect of operating in both airports on the entry decision is almost ten times larger than that of operating in only one of the airports.

==========================

Insert Table 7 about here

==========================

5.2 Entry Patterns of HST

I next explore the patterns associated with the introduction of the HST. Table 8 shows the average presence of fast trains and bullet trains at groups of routes with different average GDP. A clear pattern is that the presence of both types of HST increases with the level of a route's GDP. Tables 9 and 10 report the presence of HST for different route groups broken down by population and income per capita, respectively. The pattern observed in these tables is similar to that on Table 8 for GDP. This implies that the introduction of HST is not random, as larger markets are more likely to be served by HST. This also shows the importance of taking these market variables into consideration when studying firms' entry decisions in this setting, because both airlines and HST tend to enter markets with higher GDP and population.

==========================

Insert Table 8, 9 and 10 about here

==========================

Table 11 reports the routes by distance. The HST tends to serve more short-distance routes than long-distance ones. This contrasts with the airlines, which tend to provide more flight services on median-length routes and is consistent with the differences in the competitive advantage between the two modes of transportation.

==========================

Insert Table 11 about here

==========================

5.3 How Do Airlines Respond to the Introduction of HST?

Here I analyze how airlines respond to the introduction of HST. I first focus on airline routes that overlap with the HST routes and then look at airline routes that do not overlap with but connect the HST routes.

Airline Routes that Overlap with HST

Motivated by the previous discussion regarding the differences in the competitive advantage of airlines and HST, I divide all routes into three categories depending on their length. Specifically, a route is defined as short if its length is below 600 km, median if its length is between 600 km and 1200 km, and long if the length is above 1200 km.[28] For each length category, I further divide routes into two groups: one is the treatment group-the routes that face or will face direct competition from HST,[29] and the other is the control group, which consists the routes that have not yet faced competition from HST during the period of the data. By comparing the number of flights across categories, I examine how the presence of HST affect airlines flight choices and how such effects differ by the length of the route.

The results are shown in Figure 5 through 7. For each figure, the dashed line represents the average number of flights provided by each airline in the treatment group, while the dotted line represents the corresponding metric for the control group routes. For each group, I look at the changes in the number of flights relative to their initial values to facilitate comparison between groups. Also, because the treatment (i.e., the presence of HST) occurs at different periods, I use a bar graph to represent the proportion of routes that face competition from HST in the treatment group.[30] These figures are suggestive of HST presence having an heterogeneous effect on flights depending on the length of the route: airlines seem to reduce their flights when faced with competition from HST in short routes, but less so for median and long routes. These graphs demonstrate the importance of controlling for the moderating effect of route length when studying the impact of HST.

========================

Insert Figures 5, 6, 7 about here

========================

Airline Routes that Connect but Don't Overlap with HST

The introduction of HST may also potentially bring opportunities to the airline industry, as airlines can utilize the non-overlapping network from the HST to feed traffic to its own routes. If such positive spillover effects exist, then we should expect airlines to offer more flights on non-overlapping routes with a higher level of HST connectivity (defined as the number of HST lines that are connected to the two end-cities of the route) than in routes with a lower level of HST connectivity.

There is a selection problem, though, as cities with higher HST connectivity also tend to have larger population and larger number of flights. Directly comparing the average number of flights for routes with different levels of HST connectivity would then not be correct. To alleviate such concerns, I attempt to control for other factors that could affect airlines' flight decisions. As discussed above, airlines' route decisions are likely to be affected by factors such as population, the number of own connecting routes and route length. I therefore compare the number of flights provided by an

[28]The choices of the cutoff values are motivated by an industry report such that HST has more advantage for routes shorter than 600 km and Airlines has more advantage for routes longer than 1200 km. Source: http://www.pinchain.com/article/95181 (in Chinese).

[29]Here I don't distinguish between fast trains and bullet trains.

[30]The range is therefore from 0 to 1.

airline between two routes that are similar in all aspects (population, the number of own connecting routes and route length) and only differ regarding the level of HST connectivity to see whether the higher the level of HST connectivity in a given route the larger the number of flights offered.

To do so, I focus the city of Kunming. I choose this city because this is the only city which is a major hub of an airline but not connected to HST during the period of the study. Being a hub city ensures that there are enough flight connections to other cities, which secures enough variation in levels of HST connectivity (all variation comes from other cities) and not connected to HST helps avoid the cases where the route faces direct competition from HST, which would further complicate the analysis. The analysis here is at the airline-route-year level, as the same route may have different numbers of connecting HST lines over the two years. I first divide all routes that connect the city of Kunming into two groups. Specifically, I assign a route into the treatment group if it is connected to at least to three HST lines. The control group therefore consists of routes with less than three connections. I develop a matching algorithm which compares routes in the two groups along several dimensions. Two routes from different groups are matched as long as i) They are provided by the same airline, ii) They exist in the same year, iii) They belong to the same length group (i.e., short, median or long), iv) The routes are either both regulated or nonregulated and v) The population and number of own connecting airline routes differ by less than 10%. Table 12 shows the results. I have tried the matching algorithm both with and without replacement, and the results are very similar. On average, the routes in the treatment group have about 0.2 more flights than the control group. A T-test shows that the difference is significant at 5% level.

========================

Insert Table 12 about here

========================

To further explore the spillovers from HST on airlines, I conduct a regression analysis and control for other factors that could possibly affect airlines' flight decisions. Here I focus on airlines' decisions regarding whether to operate flights in a route and how many flights to provide in a route at a yearly level. The unit of analysis is therefore a combination of airline, route and year. There are two dependent variables I explore: one is the number of flights offered and the other is whether the airline operates at all in a given route. A linear regression and a logistic regression are used for the two dependent variables respectively.

Columns 1 through 4 of Table 13 present the results for the linear regressions, in which the dependent variable is the number of direct flights an airline provides in a route in a given year. For all specifications, I distinguish between fast trains and bullet trains. In Specification 1, I regress the dependent variable on the characteristics of HST, including the presence of HST, number of HST lines connected as well as the interactions between these variables. I find that the coefficient for the presence of bullet train is negative and significant, while that for the fast train presence is significantly positive. This suggests that there might be some endogeneity issues, as both airlines and HST tend to enter markets with high demand. Specification 2 therefore controls for the market characteristics such as the average city-pair population, length of the route as well as whether the route is regulated by the government. Motivated by the descriptive evidence from before, I interact the presence of HST with the length of route to allow the spillovers to depend on the route length. The coefficients for both the presence of fast train and bullet train are negative and significant, which is consistent with what we observed before that airlines tend to decrease the number of flights offered for short routes that overlap with HST. As in Specification 1, the coefficient for the number of connections to HST are both positive and significant for both bullet trains and fast trains. This suggests that there are positive spillovers in airline routes that connect to HST routes.

========================

Insert Table 13 about here

========================

In Specification 3, I include the characteristics of airlines as additional controls. The estimation results show that an airline tends to have more flights in routes that are connected to its base city (where it headquartered). Also, having more connecting routes is associated with a larger number of flights in the focal routes, while having an indirect flight connection for the same route reduces the number of direct flights. However, the number of flights provided by competitors in the same route increases the number of flights provided by the focal airline, which does not sound reasonable. This indicates that there might be some unobserved market characteristics that Specification 3 does not control for.

To alleviate this concern, in specification 4, I further include city-pair dummies in the regression. The results now look more reasonable, namely the coefficient on the number of competitors' flights is now negative. Also, the coefficient on the interaction between the presence of bullet train and number of connections to bullet train rail now becomes negative, which indicates that the positive spillovers from connecting HST lines go away when there is an overlap between the airline and HST in the focal route. The coefficient for the corresponding interaction for fast trains is not significantly different from zero.

The estimation results are qualitatively similar for the logistic regression reported in Columns 5 and 6, where the dependent variable is a dummy which equals to one if flights are provided by an airline in a route in a given year and zero otherwise. This shows that there is enough variation in an airline's entry/exit decisions to identify the major spillovers from HST to the airline industry.

One potential concern regarding the previous regression results is that, although the city-pair dummies control for all unobserved characteristics of the routes that do not change over time, there may exist some additional shocks over time which drive the airlines' route decisions. If the introduction of HST correlates with such shocks, this may lead to spurious correlation between the number of airline flights and the presence HST.

To address this concern, I follow Goolsbee and Syverson (2008)'s insight and exploit the presence/absence of treatment (i.e. overlap with HST) between the focal route and other routes involve the same airports for one of the cities in the city-pair but which do not overlap with the HST routes. I illustrate the principles behind the selection of the control routes in Figure 8.

==========================

Insert Figure 8 about here

==========================

Suppose the focal route is city pair AB, which overlaps with HST. I use routes AC and BD as control routes. To be included in the difference-in-differences regression, the routes must satisfy the following criteria:

1. The firm must have provided flight services in all three routes at some point in the data.
2. Each control route must share one of the end cities with the focal route.
3. The distance between city B and C must be less than 600 km (375 miles). Same for the distance between A and D.
4. Both routes AC and BD do not overlap with HST.

Criterion 1 makes sure that providing flight services is feasible for the airline in all three routes. Criterion 2 establishes that airport-specific operating cost shocks are embodied in either of the control routes. Criterion 3 makes sure that the control routes are reasonably close to the focal route so that they can also absorb potential demand shocks. Criterion 4 helps to determine whether there is no treatment in the control routes.

The identification assumption underlying this analysis is that both demand side and supply side shocks in the focal route should be captured by either one of the control routes. Thus, by comparing the number of flights offered in the control routes with those offered in the focal route over time (while controlling for other observable differences), we can identify the impact of the presence of HST.

In practice, I proceed by assigning a group ID that is the same for the focal route and its two corresponding control routes. I regress the dependent variable (number of flights) on the treatment (the presence of HST), market characteristics, airline characteristics, and city-pair dummies, airline dummies and group dummies interacted with year dummies. Therefore, while the city-pair dummies control for the time invariant differences across routes, the group-year dummies allow each group to have its own flexible trend in terms of the number of flights thus capturing any unobserved shocks to the focal route. I combine fast train and bullet train and define treatment as a dummy variable that is equal to 1 if either fast train or bullet train is present in the route. Thus, what is measured here is the average treatment effect of the presence of fast trains and bullet trains on the number of airline flights offered.

Table 14 reports the regression results. The dependent variable is the number flights an airline provides in a route for each year. I present the results from four different specifications in the table.

The first specification uses all observations, while the rest of the three breaks them down into three groups depending on the length of the route. Specification 2 is for short routes, while Specifications 3 and 4 are for median length routes and long routes, respectively.

=========================

Insert Table 14 about here

=========================

The results show that, the presence of HST (either fast train or bullet train) has a negative impact on the number of flights, but only for short routes.[31] For median length and long routes, there is a positive effect for the HST overlap, which suggests there may be market expansion effects associated with the introduction of HST.

6 Structural Model

So far I have demonstrated the existence of spillovers from HST to the airline industry using reduced form evidence. Such evidence is however not sufficient to explore the profit implications for the industry and to carry out policy experiments. In this section, I present a structural model that studies dynamically the entry decisions of airlines. I model airlines' decisions to enter a route or not. Since the purpose of this paper is to evaluate the implications of spillovers from HST on airlines' configuration of networks, the intensity of entry (i.e. the number of flights provided in a route) is less relevant. It also comes at a cost of considerably increasing the dimension of the state space, adding burden to the computation. Moreover, conditional on entry, about 60% of the routes only have one flight per airline, therefore the quantity change should be reflected in the airlines' entry/exit decisions. Also, as is shown in the previous regression analysis, there is enough variation of entries and exits in the data to help identify parameters of key interest. Further, I model the flow payoff of an airline in a reduced form in the same manner as Arcidiacono et al. (2016) and Igami and Yang (2016).

6.1 Setup

There are N airlines in the industry that provide flight services between different cities. Denote the total number of cities in the industry by C. A route $r \in \{1, ..., R\}$ is defined as a non-directional city pair, with R denoting the total number of possible routes between the cities. Let $x_{irt} \in \{0, 1\}$ be an indicator variable where $x_{irt} = 1$ if airline i provides direct flights in route r at time t and $x_{irt} = 0$ otherwise.[32] The network of an airline i at t is therefore defined by the collection of x_{irt}

[31] The coefficient for the population for short routes is significantly negative. This could be because, conditional on other characteristics, there is more variation in the number of flights for routes with larger population. For example, airlines may provide more flights in routes with more population, and when HST appears, the reduction in the flights is also larger for more populated routes. This effect is to some extent captured by the coefficient on population.

[32] As routes are non-directional, I assume that if an airline provides flights from one city to the other, it also provides flights in the opposite direction.

for all routes, i.e. $x_{it} \equiv \{x_{irt} : r = 1, ..., R\}$. Further denote $x_t = \{x_{it} : i = 1, ..., N\}$ as the network for the whole industry at time t.

Time is continuous, indexed by $t \in [0, \infty)$. I assume that the flow payoff airline i gets from route r at time t is a function of both the network of the industry and some exogenous characteristics of the route, z_{rt}. Denote that payoff by $f_{irt}(x_t, z_{rt})$.

A stochastic process governs when and which airline can move. When it is airline i's turn to move, it decides whether to make changes to its existing network, x_{it}. Specifically, for each route, airline i can choose to do nothing, to provide direct flights if the route is not being served (that is, to enter), or to stop providing direct flights (i.e. exit). Denote $j \in \mathscr{A} = \{0, ..., J-1\}$ as the action airline i takes in route r. Each action is associated with an additive separable instantaneous payoff, $\phi_{irt}(j)$.

Airlines are forward looking and discount future payoffs at rate $\rho \in (0, \infty)$. When an airline i gets a chance to move, the airline examines x_t and $z_t \equiv \{z_{rt} : r = 1, ..., R\}$, which together composes the industry state, and choses actions for all routes to maximize its overall discounted expected payoff given by

$$E\left[\sum_{r=1}^{R}\left(\int_0^\infty e^{-\rho t} f_{irt}(x_t, z_{rt}) + \sum_{n_r=1}^{\infty} e^{-\rho T_{n_r}} \phi_{irt}(j)\right)\right], \tag{1}$$

where T_{n_r} is the random time of the nth state change due to action of airline i at a given route r.

6.2 Assumptions

One complication of the model arises from the dimension of the state space and action space. Given the number of routes and airlines in the data, the total number of states for x_t alone is about $2^{NR} \approx 10^{2743}$, making it extremely challenging to tackle the model empirically. Estimating and computing dynamic games with very large state spaces has not been solved in general. To accommodate this issue, I follow Aguirregabiria and Ho (2012) and make the following two assumptions:

Assumption 1: Decentralized Decisions Each airline decides its network in a decentralized manner. That is, an airline makes decisions for one route at a time, taking the situation in other routes as given.

Assumption 1 says that, instead of making decisions for all routes simultaneously, an airline only considers one route at a time. This helps to reduce the action space from 2^R to 2, which greatly eases the computational burden.[33,34]

Assumption 2: Sufficient Statistics Let $x_{rt} \equiv \{x_{irt} : i = 1, ..., N\}$ denote the states for all airlines in route r. Let $w_{r^c t} \equiv \{w_{ir^c t} : i = 1, ..., N\}$ denote the vector that summarizes the state

[33] Note that, as will be discussed later, this assumption does not imply that airlines don't make decisions considering their entire network.

[34] A similar assumption has been made in studies of spatial competition (e.g., Schiraldi, Smith, and Takahashi 2012), product category management (e.g., Cachon and Kök 2007, Gajanan, Basuroy, and Beldona 2007, Basuroy, Mantrala, and Walters 2001), and international trade (e.g., Gopinath and Neiman 2014, Halpern, Koren, and Szeidl 2015, Blaum, Lelarge, and Peters 2016, Monarch 2016).

variables for each airline in all routes other than r. Each airline's entry decision for each route relies only on $(x_{rt}, w_{r^c t}, z_{rt})$.

Assumption 2 implies that, things that happen in other routes affect flight choices through $w_{r^c t}$.[35,36] This helps to reduces the state space as an airline no longer needs to track the states for all other routes when making decisions for a route.

6.3 Decisions and Payoffs

For each airline at each route, there is an independent Poisson arrival process that governs when each player can move. In the context of airline competition, a random move arrival process might reflect the stochastic timing of firms' staff recruiting, flight capacity arrangement, negotiation with airports for gate occupancy, delays in permitting processes, and so on. I assume a common rate parameter, λ, for these processes. In each route, there is also a Poisson process that affects the exogenous state z_{rt}, which I refer to as the move of nature.

Although I assume that airlines make route decisions in a decentralized manner, I acknowledge the fact that an airline may not just care about the flow payoffs from the focal route. Instead, the airline may also consider the impact of its decision for one route on the profits of the other routes. Consistent with this, I assume that such motivations are incorporated in the flow profits $u_{irt}(x_{rt}, w_{r^c t}, z_{rt})$, with $|u_{irt} < \infty|$. Note that, as in Aguirregabiria and Ho (2012), $u_{irt}(x_{rt}, w_{r^c t}, z_{rt}) \neq f_{irt}(x_t, z_t)$.

Conditional on a move, airline i chooses an action $j \in \mathcal{A}$ for route r that maximizes its expected net present value for that route, given by

$$E\left[\left(\int_0^\infty e^{-\rho t} u_{irt}(x_{rt}, w_{r^c t}, z_{rt}) + \sum_{n_r=1}^{\infty} e^{-\rho T_{n_r}} \phi_{irt}(j)\right)\right]. \tag{2}$$

6.4 Value Function and Equilibrium

For notational simplicity, I summarize $(x_{rt}, w_{r^c t}, z_{rt})$ for any route r at any time t into an element k of some finite state space $\chi = \{1, ..., K\}$. I assume that the nature's move at each route can be represented with a finite state Markov jump process on χ with a $K \times K$ intensity matrix Q_{0r}. The elements of Q_{0r}, denoted by q_{kl}, are the rates at which a particular state transitions occur and are nonnegative and bounded.

[35] Several other papers use a similar idea to tackle the space dimensionality problem (e.g., Gowrisankaran and Rysman 2012, Weintraub, Benkard, and Van Roy 2006).

[36] This is one of the key distinctions my model have with Aguirregabiria and Ho (2012). While in their model each local manager is faced with a single dynamic decision problem and the strategic interaction between them are only realized indirectly through time, my model specifically accounts for the strategic interaction between airlines in the same route. I believe such extension is natural and more reasonable because it is likely that when an airline makes decision on flight numbers in a given route, it pays more attention to the actions of other players in the same route, where the profits are more directly affected.

The instantaneous payoff $\phi_{irk}(j)$ consists of two components and is given by $\psi_{ijk} + \epsilon_{ij}$. ψ_{ijk} is the mean payoff (or cost) associated with action j in state k for airline i, with $|\psi_{ijk}| < \infty$, and ϵ_{ij} is a payoff shock, which is assumed to follow the type I extreme value distribution and is *i.i.d.* distributed across airlines, routes, actions and time.[37]

Let σ_{ijkr} denote the probability that airline i optimally chooses action j in state k for route r. This action may result in a deterministic state change. Let $l(i,j,k)$ denote the continuation state that arises after the focal airline makes decision j in state k. Following convention, I define $j = 0$ to be a costless continuation choice with $\psi_{i0k} = 0$ and $l(i,0,k) = k \; \forall k$.

Apart from nature's and the airline's decisions, the state k may also change as a result of the evolution of w_{ir^ck} which, as mentioned above, summarizes the states of airline i for all routes other than r. I assume that, at the moment of decision for route r, airline i does not have perfect foresight of what decisions it will make for other routes in the future. The evolution of this aggregate state variable thus can be regarded as stochastic. The change of state k due to w_{ir^ck} can be interpreted as the result of the action of an "aggregate player", which I denote as $i^c \in \{N+1, ..., 2N\}$.[38] As w_{ir^ck} is a function of airline i's decisions on the other routes, this cumulative process still follows a Poisson arrival process.[39] Denote the rate of that process as λ_c. We can similarly define σ_{i^cjkr} as the probability i^c "chooses" action j in state k in route r, and let $l(i^c,j,k)$ be the corresponding continuation state.

Define ζ^{ir} as airline i's beliefs regarding the actions of all other decision makers in route r (including nature and the "aggregate player"), given by a collection of $(2N-1) \times J \times K$ probabilities ζ^{ir}_{mjk} for each $m \neq i$, state k and choice j. Finally, let $V_{irk}(\zeta^{ir})$ denote the expected present value for airline i being in state k at route r and behaving optimally at all points in the future given beliefs ζ^{ir}. For small increments h, under the Poisson assumption, the probability of an event with rate λ occurring is λh. Given the discount rate ρ, the discount factor for such increments is $1/(1+\rho h)$. Thus, for small time increments h the present discounted value of being in state k for airline i at route r (suppressing dependence on ζ^{ir} for brevity) is given by

$$V_{irk} = \frac{1}{1+\rho h}\Bigg[u_{irk} + \sum_{l \neq k} q_{kl} V_{irl} + \sum_{m=1, m\neq i}^{N} \lambda h \sum_{j=0}^{J-1} \zeta^{ir}_{mjk} V_{ir,l(m,j,k)} + \sum_{i^c=N+1}^{2N} \lambda_c h \sum_{j=0}^{J-1} \zeta^{ir}_{i^cjk} V_{ir,l(i^c,j,k)}$$
$$\lambda h Emax_j \left[\psi_{ijk} + \epsilon_{ijrk} + V_{ir,l(i,j,k)}\right] + \Big(1 - N(\lambda+\lambda_c)h - \sum_{l\neq k} q_{kl}\Big) V_{irk}\Bigg].$$

Rearranging and letting $h \to 0$, V_{irk} can be defined recursively as

$$V_{irk} = \frac{u_{irk} + \sum_{l\neq k} q_{kl} V_{irl} + \sum_{m=1,m\neq i}^{N} \lambda h \sum_{j=0}^{J-1} \zeta^{ir}_{mjk} V_{ir,l(m,j,k)} + \sum_{i^c=N+1}^{2N} \lambda_c h \sum_{j=0}^{J-1} \zeta^{ir}_{i^cjk} V_{ir,l(i^c,j,k)} + Emax_j\left[\psi_{ijk} + \epsilon_{ijrk} + V_{ir,l(i,j,k)}\right]}{\rho + \sum_{l\neq k} q_{kl} + N(\lambda+\lambda_c)}. \quad (3)$$

[37] I suppress the dependency of ϵ on r and t for simplicity.

[38] I use $N+1$ through $2N$ to distinguish the identities of aggregate players and those of airlines in a given route.

[39] I have done some Monte Carlo experiments to show that this holds, but may need some formal proof.

I focus on Markov perfect equilibria in pure strategies, as is standard in empirical literature of dynamic games. A markov strategy for i in route r is a mapping which assigns an action from $\mathscr{A}$ to each state $(k, \epsilon_{irk}) \in \chi \times \mathbb{R}^J$. Given beliefs $\{\zeta^{ir} : i = 1, ...N; r = 1, ..., R\}$, and a collection of model primitives, a Markov strategy for firm i in route r is a best response if

$$\delta_{ir}(k, \epsilon_{irk}; \zeta^{ir}) = j \Longleftrightarrow \psi_{ijk} + \epsilon_{ijrk} + V_{ir,l(i,j,k)} \geq \psi_{ij'k} + \epsilon_{ij'rk} + V_{ir,l(i,j',k)} \quad \forall j' \in \mathscr{A}. \tag{4}$$

Given the distribution of choice-specific shocks, each Markov strategy δ_{ir} implies the following response probabilities for each choice in each state

$$\sigma_{ijrk} = Pr[\delta_{ir}(k, \epsilon_{irk}; \zeta^{ir}) = j|k]. \tag{5}$$

A Markov perfect equilibrium is thus defined as a collection of stationary policy rules $\{\delta_{ir} : i = 1, ..., N; r = 1, ..., R\}$ such that (4) holds for all i, r, k and ϵ_{irk} given beliefs $\zeta^{ir} = (\sigma_{ir} : i = 1, ..., N; r = 1, ..., R)$ generated by (5).[40]

7 Empirical Specification

7.1 State Variables and Action Variables

I focus on the top four airlines in China in terms of market share by number of flights, so $N = 4$. I focus on the top 68 cities in terms of flight frequency. In total, there are $R = 68 \times 67/2 = 2278$ possible routes. In a given route r, the state variables for airline i are x_{rk}, $w_{r^c k}$ and z_{rk}, where $x_{rk} = \{x_{irk} \in \{0, 1\} : i = 1, ...4\}$ denotes the identities of the airlines that operate flights in route r. I assume that i's decisions are affected by its network position only through its connecting routes.[41] Let c_{irk} denote the number of connecting routes for airline i in route r in state k.[42] Then $w_{r^c k} = \{c_{irk} : i = 1, ...4\}$.

The state variable z_{rk} contains the exogenous characteristics of route r, namely, the average population for endpoint cities pop_{rk}, indicator variables for the presence of fast train ($Fast_{rk}$) and bullet train ($Bullet_{rk}$), the number of fast train lines and bullet train lines connecting either of the endpoint cities ($c_{Fast,rk}$ and $c_{Bullet,rk}$, respectively), and the length of the route, d_{rk}. Also, I assume that each route is characterized by a time-invariant unobserved type s, which is observed by airlines but not by the econometrician.

[40] The standard arguments apply for existence of equilibrium. (See Doraszelski and Satterthwaite 2010, Doraszelski and Judd 2012).

[41] As defined before, a connecting route is a route which shares one of the endpoint cities with the focal route. As of now, I do not consider the impact of indirect flights on an airline's decision. However, descriptive evidence is suggestive that the absolute impact of indirect flights on a firm's entry decision of a route is relatively small.

[42] For example, if route r connects city A and city B, and in state k, airline i offers direct flights in a routes that connect city A and b routes that connect city B, then c_{ir} is $a + b$.

7.2 Flow Profits and Choice-Specific Payoffs

As I do not have demand side data such as price and seat occupancy, I specify the profit function in a reduced form.[43] The flow payoff to airline i for operating flights in route r is specified as a linear function of the airline's and the competitors' route networks as well as some exogenous variables such as population, length of the route and the network of high HST. The flow payoff also depends on an unobserved (to the econometrician) characteristic of the route, s, which captures the unobserved tastes of consumers in a given route. The full state of the model is therefore captured by (k, s).

Formally, the flow payoff to airline i in state (k, s) at route r is

$$\begin{aligned} u_{irk} = \beta_0 &+ \beta_1 \sum_{m \neq i} x_{mrk} + \beta_2 c_{irk} + \beta_3 \sum_{m \neq i} c_{mrk} + \beta_4 Fast_{rk} + \beta_5 Bullet_{rk} \\ &+ \beta_6 c_{Fast,rk} + \beta_7 c_{Bullet,rk} + \beta_8 Fast_{rk} \times c_{Fast,rk} + \beta_9 Bullet_{rk} \times c_{Bullet,rk} + \beta_{10} pop_{rk} + \beta_{11} d_{rk} \\ &+ \beta_{12} Fast_{rk} \times d_{rk} + \beta_{13} Bullet_{rk} \times d_{rk} + \beta_{14} s_r. \end{aligned} \tag{6}$$

I include c_{mrk} with $m \neq i$ to allow an airline's payoff to also depend on the network of its competitors. I interact the presence of HST, $Fast_{rk}$ and $Bullet_{rk}$ with their respective number of connecting HST lines to capture the moderating effect of competition on the positive spillovers. Finally, I also interact the presence of HST with the length of the route to control for the heterogeneous effect of the introduction of HST on routes of different lengths. The flow payoff for not operating flights in a given route is normalized to 0.

I assume that airlines pay a sunk cost from entering a route. The sunk cost depends on the unobserved market type, s, on whether the route is a government-regulated air route, and on whether the airline is exempt from the regulation. The scrape value associated with exiting a route, is assumed to be zero.[44] Therefore, the choice-specific instantaneous payoffs ψ_{ijk} is given by

$$\psi_{ijk} = \begin{cases} \eta_0 + \eta_1 \times s + \eta_2 \times reg_i + \eta_3 \times HQ_i & \text{if } j = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where reg_i is an indicator variable which equals to one if the route is regulated and zero otherwise, and HQ_i is an indicator variable which equals one if airline i is headquartered at one of the city-pair and zero otherwise. The structural parameters to be estimated are $\theta = (\beta_0, ..., \beta_{14}, \eta_0, ..., \eta_3)$.

[43] This is consistent with literature on firm entry, e.g., Bresnahan and Reiss (1990), Seim (2006).

[44] Aguirregabiria and Ho 2012 have a similar assumption.

7.3 Variable Discretization

To reduce the dimension of the state space, I discretize some of the variables. First, I discretize the number of airline connections into 5 bins as

$$N_{c_{irk}} = \begin{cases} 0 & \text{if } c_{irk} \in [0,5] \\ 1 & \text{if } c_{irk} \in [6,15] \\ 2 & \text{if } c_{irk} \in [16,25] \\ 3 & \text{if } c_{irk} \in [26,35] \\ 4 & \text{if } c_{irk} \geq 36 \end{cases}.$$

Similarly, the HST connections are discretized into 3 bins as

$$N_{c_{hrk}} = \begin{cases} 0 & \text{if } c_{hrk} \in [0,1] \\ 1 & \text{if } c_{hrk} \in [2,3] \\ 2 & \text{if } c_{hrk} \geq 4, \; h \in \{Fast, Bullet\} \end{cases}.$$

Because an airline's payoff does not necessarily change monotonically with the length of the route, I do not specify the utility as a linear function of route lengths. Instead, I use three indicator variables to denote short routes ($\leq$ 600 km), median routes (between 600 km and 1200 km) and long routes ($>$ 1200 km). Finally, I discretize the average city-pair population into five quantile-based bins.

8 Estimation

I now describe the steps I followed to estimate the dynamic oligopoly model. I follow Arcidiacono et al. (2016) and apply a two-step estimation approach similar to Bajari, Benkard, and Levin (2007) (henceforth BBL), Hotz et al. (1994) (henceforth HMSS) and Aguirregabiria and Mira (2007) but in a continuous-time framework. In the first step, I estimate the reduced form hazards that capture the dynamics in entry and exit decisions for airlines in each route, as well as the rate of change for the inclusive values for the network and exogenous states such as population and presence of HST. With the estimation results from the first stage, I estimate the structural parameters in the second stage (detailed below).

The finite dependence property used in Arcidiacono et al. (2016) greatly reduces the computation time when calculating value functions. However, this property requires the assumption of permanent exit, which does not apply in my case.[45] So I borrow from HMSS and BBL to approximate the value function using forward simulation in the second stage. To reduce the computational burden, I impose symmetry and anonymity in the model. Anonymity means that when an airline makes decisions, it does not care about the identity of a specific player, but only about the distribution

[45] In the sample, about one third of the airlines returned to the routes once they exited.

of the competitors' states. Symmetry implies that we can use a common policy function to model each player's behavior.[46]

8.1 Stage 1

Following Arcidiacono et al. 2016, I estimate the probability of entry, exit and doing nothing for an airline in a route using a multinomial logit sieve. Let $\tilde{\sigma}_{ijr}(k,s,\alpha)$ denote the reduced form probability of airline i making choice j in state (k,s) for route r, where α is the parameter to be estimated. I assume that $\tilde{\sigma}_{ijr}(k,s,\alpha)$ takes the following form

$$\tilde{\sigma}_{ijr}(k,s,\alpha) = \frac{exp(\phi_j(k,s,\alpha))}{\sum_{j' \in \mathcal{A}} exp(\phi_{j'}(k,s,\alpha))}, \tag{8}$$

where $\phi_j(k,s,\alpha)$ is a flexible function of the state variables. The variables included in the function are the number of competitors and its square, the number of own connecting routes and its square, the total number of competitors and its square, the total number of competitors' connecting routes and its square, the presence of fast train, the presence of bullet train and the interaction of these indicator variables with route length dummies, the number of connecting fast train lines and its interaction with the presence of fast train, and the number of connecting bullet train lines and its interaction with presence of bullet train. I also include the average city-pair population, dummies for length of route, and the average population growth rate of the city-pair[47] as controls. Further, I allow the first-stage policy function to depend on the unobserved type of the route.[48]

Entry costs are allowed to depend on the market unobserved type as well as on whether a route is under regulation and on whether a airline is exempt from regulation.

Define $\tilde{\sigma}_{i^c jr}(k,s,\alpha^c)$ as the reduced form probability that the "aggregate player" makes choice j in state (k,s) in r. I use the same functional form as in function 8 to model this probability. The function includes the following variables: an indicator variable for whether an airline is the incumbent in the focal route, the number of competitors in the focal route and its square, the number of own connecting routes and its square, the total number of competitors' connecting routes and its square, the presence of fast train, the presence of bullet train, the number of connecting fast train lines, and the number of connecting bullet train lines. Again, I include the average city-pair population and the average population growth rate of the city-pair as controls. I also allow the "aggregate player" probabilities to depend on whether the focal route is regulated and on whether a given airline is exempt from regulation.

The exogenous state variables include population, the presence of fast train, the presence of bullet train, the number of fast train connections and the number of bullet train connections. I treat the

[46] For example, if there are 5 firms, numbered 1 through 5. A vector summarizes the state of each firm. Suppose under the state (1,1,3,2,1), firm 1 choose to enter a route, then anonymity means that firm 1 will make the same decision under the state (1,3,2,1,1). Symmetry implies that under the state (3,1,1,2,1), firm 3 will make the same decision as firm 1 in the state (1,1,3,2,1).

[47] Specifically, I discretize these growth rates into three quantile-based bins.

[48] In the first stage, I jointly estimate the policy function and the probability of a route being of a specific unobserved type.

movement of all these variables as caused by "nature". As discussed above, I model the changes in the state of the HST's network as stochastic. However, I allow the "entry" probability of HST in a given route to depend on factors such as the average city-pair population, the average population growth rate of the city-pair, route length, as well as the number of connecting HST lines. Also, the movement of connecting HST lines in a focal route is allowed to depend on the average city-pair population and the average population growth rate of the city-pair. For the "entry" probabilities and the probabilities of the movement of connecting HST lines for both fast trains and bullet trains, I use multinomial logit model for the approximation. $\tilde{q}_j(k, s, \alpha_0)$ is used to denote the probability that nature takes action $j \in 0, ..., J-1$ in state (k, s).

8.1.1 Likelihood Function

Since we have assumed that airlines make decisions for each route individually and that they only take into account the states in the focal route as well as the summary statistics of other routes when making decisions, the likelihood function can be constructed at the route level. In what follows, I suppress the dependence on r for brevity. Because the Poisson processes governing the movements of all players (including nature) are assumed to be independent, the joint process can be modeled as an aggregate Poisson process with accumulated arrival rate. Specifically, let $h_{ijk} = \lambda\sigma_{ijk}$ denote the hazard of player i choosing action j in state k, and let

$$h = (q_{12}, q_{13}, ..., q_{K-1,K}, \lambda\sigma_{111}, ..., \lambda\sigma_{1J-1k}, ...\lambda\sigma_{N11}, ...\lambda\sigma_{NJ-1K}, ...\lambda_c\sigma_{N+111}, ..., \lambda_c\sigma_{N+1Jk}, ...\lambda_c\sigma_{2N11}, ...\lambda_c\sigma_{2NJK})$$

denote the vector of hazard rates of all players including nature for state-specific non-continuation choices. In state k the probability that player i takes action j after τ units of time can be decomposed into two parts: the first part is the probability that state changes after τ units of time. Because the aggregate process follows a Poisson process, the interval between state changes follows a exponential distribution. Therefore the probability that state changes after τ units of time is given by

$$\left(\sum_{l\neq k} q_{kl} + \sum_{i} \lambda \sum_{j\neq 0} \sigma_{ijk} + \sum_{i^c} \lambda_c \sum_{j^c\neq 0} \sigma_{i^c jk}\right) exp\left[-\tau\left(\sum_{l\neq k} q_{kl} + \sum_{i} \lambda \sum_{j\neq 0} \sigma_{ijk} + \sum_{i^c} \lambda_c \sum_{j^c\neq 0} \sigma_{i^c jk}\right)\right]. \tag{9}$$

The second part is the probability that player i takes action j, conditional on the state change. Again, by the property of independent Poisson arrival processes, the probability that player i takes action j is given by

$$\frac{\lambda\sigma_{ijk}}{\sum_{l\neq k} q_{kl} + \sum_i \lambda \sum_{j\neq 0} \sigma_{ijk} + \sum_{i^c} \lambda_c \sum_{j\neq 0} \sigma_{i^c jk}}. \tag{10}$$

Taken together, the probability becomes:

$$\lambda\sigma_{ijk}exp\left[-\tau\left(\sum_{l\neq k}q_{kl}+\sum_{i}\lambda\sum_{j\neq 0}\sigma_{ijk}+\sum_{i^c}\lambda_c\sum_{j\neq 0}\sigma_{i^cjk}\right)\right]. \tag{11}$$

For each route $r \in 1,...,R$, I observe W_r events over the continuous time interval $[0,\overline{T}]$.[49] Denote k_{rw} ($w \in \{1,...W_r\}$) as the state immediately prior to the wth event in route r and denote t_{rw} the corresponding time at which the even occurs. The holding time of wth event for route r, τ_{rw}, can therefore be defined as $\tau_{rw} = t_{rw} - t_{r,w-1}$.

Denote $I_{rw}(i,j)$ as the indicator variable which equals to one if player i takes action j in the wth event at route r. Now, conditional on a route being of unobserved type s, the likelihood for the single event w in route r is given by:

$$\begin{aligned}
&\tilde{L}_{rw}(h(\alpha);s) = \\
&\left(\sum_{j\neq 0}I_{rw}(0,j)\tilde{q}_j(k_{rw},s,\alpha^0)+\sum_{i}\lambda\sum_{j\neq 0}I_{rw}(i,j)\tilde{\sigma}_{ijr}(k_{rw},s,\alpha)+\sum_{i^c}\lambda_c\sum_{j\neq 0}I_{rw}(i^c,j)\tilde{\sigma}_{i^cjr}(k_{rw},s,\alpha^c)\right) \\
&\times exp\left[-\left(\sum_{j\neq 0}\tilde{q}_j(k_{rw},s,\alpha^0)+\sum_{i}\lambda\sum_{j\neq 0}\tilde{\sigma}_{ijr}(k_{rw},s,\alpha)+\sum_{i^c}\lambda_c\sum_{j\neq 0}\tilde{\sigma}_{i^cjr}(k_{rw},s,\alpha^c)\right)\tau_{rw}\right].
\end{aligned} \tag{12}$$

Following Arcidiacono et al. (2016), I control for the unobserved route type using mixture distributions. I discretize the standard normal distribution into five points and calculate the probability of each route being at each point as a function of the initial conditions of the routes. Specifically, I specify this probability as an ordered probit which depends on the total number of flights, the total number of connecting routes, the average population growth rate of the city-pair, length of the route as well as on an indicator variable that captures whether the route is regulated.

Denote k_{r0} as the initial state of route r. Let $\pi(s,k_{r0})$ be the probability of route r being type s given initial condition k_{r0}. The likelihood function therefore integrates over the unobserved types for each route. The maximum likelihood estimate then become

$$(\alpha^{0*},\alpha^*,\alpha^{c*},\pi^*) = \arg\max_{\alpha^0,\alpha,\alpha^c,\pi}\sum_{r=1}^{R}ln\left(\sum_{s}\pi(s,k_{r0})\prod_{w=1}^{W_r}\tilde{L}_{rw}(h(\alpha);s)\right). \tag{13}$$

8.2 Stage 2

In the second stage, I estimate the structural parameters, θ, based on the probabilities of being in each unobserved types and the hazards estimated in stage 1. Specifically, I take the following steps: first, based on the estimated hazards, I approximate the value function of airline i for route r at state (k,s) as a linear function of θ using forward simulation in the same spirit of BBL and HMSS.

Given the value functions $\check{V}_{irk}(\theta;s)$, together with the assumption that the idiosyncratic error

[49] This includes the "event" at time $\overline{T}$, where possibly nothing happens.

terms in equation (4) follow an i.i.d. type I extreme value distributions, the choice probabilities in equation (5) can be expressed as:

$$\check{\sigma}_{ijrk}(\theta; s) = \frac{exp\left(\check{V}_{irl(j)}(\theta; s) + \psi_{ijk}\right)}{\sum_{j' \in \mathcal{A}_k} exp\left(\check{V}_{irl(j')}(\theta; s) + \psi_{ij'k}\right)}.$$

Replacing $\tilde{\sigma}_{ijk}$ in $\tilde{L}_{rw}(h(\alpha); s)$ with $\check{\sigma}_{ijrk}$, the new likelihood function can be denoted as $\check{L}_{rw}(\theta; s)$. Also, denote $\pi_r(s)$ as the likelihood of route r being of unobserved type s given the data. Using Bayes's rule, we have:

$$\pi_r(s) = \frac{\pi(s, k_{r0}) \prod_{w=1}^{W_r} \tilde{L}_{rw}(h(\alpha); s)}{\sum_{s'} \pi(s', k_{r0}) \prod_{w=1}^{W_r} \tilde{L}_{rw}(h(\alpha); s')}.$$

The second stage estimates therefore become

$$\theta^* = \arg\max_{\theta} \sum_{r=1}^{R} \sum_{s} \pi(s, k_{r0}) \sum_{w=1}^{W_r} ln\check{L}_{rw}(\theta; s). \tag{14}$$

9 Results

9.1 Parameter Estimates from The Structual Model

Table 15 shows the results of the estimation of the structural parameters. For the characteristics regarding an airline's own industry, we see that the flow payoff of an airline decreases with the number of competitors. However, this negative impact can be compensated by the average number of connecting routes provided by the airline. In fact, the positive impact of connecting routes is almost twice as large as the negative impact associated with the presence of competitors. This implies that, even if all competitors are present in a given route (in our case the total number of competitors equals three), as long as the number of connecting routes takes a level of 2, which corresponds to 16 to 25 connecting routes, the negative effect from competitors on the flow payoff will be completely compensated. The number of connecting routes for competing airlines has little impact on the flow payoff of the focal airline.

=========================

Insert Table 15 about here

=========================

Turning to the impact of the HST, there is a negative impact of the presence of HST on the flow payoff of airlines for short routes. Interestingly, the presence of fast train has a larger effect on the airlines' flow payoff than that of the bullet train. This happens probably because, for shorter routes, the speed advantage of the bullet train is not as pronounced relative to the fast train, and the tickets for the fast train are cheaper.

Consistent with the results from the reduced form analysis in section 5, the negative effect of the presence of the fast train goes away when the length of the route increases. Specifically, the negative impact is reduced by 85% when the length changes from short to median, and even becomes positive when the length is long (i.e., greater than 1200 km). A similar effect, however, is not found for bullet trains. A possible explanation for the differential effect of fast and bullet trains on airlines' payoffs is that the introduction of HST has both positive and negative impact on the airlines in the same route. On one hand, there is a competition effect, and on the other hand there may also exist a "market expansion" effect. As reported by *the Economist*,[50] for many HST routes in China, about half of the traffic is "generated" such that passengers would not take the trip otherwise. This generated traffic expands the market, which would in turn possibly benefit the airlines in the same route. The strength of the two effects changes with the length of the route as well as with the type of HST. For the fast trains, the positive effect of their presence on the flow payoff of airlines could be a result of stronger positive effect. Without finer level data, however, there is no way to disentangle these two effects.

Connecting to more fast train lines and bullet train lines has a positive effect on the payoff of airlines, which shows that there are positive spillovers from the introduction of HST that result from the complementarity in the two modes of transportation in terms of connecting routes. On average, the negative impact of the presence of the fast train on the flow payoff for short routes can be compensated by having about 4 routes with a median level of connectivity to HST lines and no overlap with HST. The corresponding number of connections for bullet train is 3, suggesting a larger positive spillover from bullet trains than from fast trains. However, this positive effect seems to be moderated by the presence of HST in the same route, especially for fast trains. Specifically, when the fast train is present in a route, the impact of the connectivity to fast train routes changes from positive to negative. This seems reasonable because when the HST also connects the focal route, having connections to other HST lines is less likely to feed traffic to the airlines in the focal route because it is less convenient. This moderating effect is less strong for bullet trains than for fast trains.

Turning to the market characteristics, airlines have lower flow payoffs for longer routes, but enjoy higher payoffs in routes with higher values of the route unobserved type. The level of population does not seem to affect the payoffs.

I find significant costs of entry for airlines. This demonstrates the importance of modeling the game in a dynamic framework. Interestingly, a route with a higher value of the unobserved type not only has a positive effect on the flow payoffs, but also reduces the entry cost as well. This contrasts to the results of Arcidiacono et al. (2016), where markets with higher unobserved type have larger flow payoff but at the same time have higher costs of entry. Finally, the regulated routes are associated with a higher entry costs for airlines, but when an airline is exempt from the regulation, this negative effect goes away.

[50]Source: http://www.economist.com/news/china/21714383-and-theres-lot-more-come-it-waste-money-china-has-built-worlds-largest

Profit Decomposition

Using the estimated structural parameters, we can study the sources of airlines' flow profits. Specifically, I decompose an airline's profits into several sources/categories: 1. Market characteristics, 2. The characteristics of the airline's own network, 3. The characteristics of competing airlines' networks, 4. Negative spillovers from fast trains, 5. Negative spillovers from bullet trains, 6. Positive spillovers from fast trains, and 7. Positive spillovers from bullet trains. To carry out the profit decomposition for a specific group of routes, I first calculate the average characteristics of these routes (conditional on an airline's market presence) and then, depending on the sources of the profits I assign these characteristics into different categories. Finally, using the structural estimates, I calculate the contribution of profits in each category. Table 16 describes the characteristics included in each category.

==========================

Insert Table 16 about here

==========================

Table 17 presents the results. The first row of the table reports the decomposition of average airline profits across all routes and all time periods. We can see that an airline's own network has the greatest impact on the its profits. This is consistent with the finding in Aguirregabiria and Ho (2012) that having a large hub can significantly reduce the fixed costs and thus increase the profits in a route. The market characteristics of a route also have a significant impact on the profits, followed by competitors' networks. The negative impact of fast trains is bigger than that of bullet trains. This is probably because fast trains have a larger network coverage than bullet trains. Together, the negative spillovers from HST account for about 6.7% of the profits brought by each airline's own networks. Interestingly, however, these negative spillovers seem to be compensated by the combined positive spillovers from the connection to HST lines, which are about 8.6% of the profits from airline's own networks. This indicates that, conditional on an airline's (endogenous) network configuration, HST has a net positive impact on the airline's profits.

==========================

Insert Table 17 about here

==========================

To investigate this further, I break down the routes into several groups along various dimensions. I first look at the profit decomposition for routes of different lengths. Across all groups on that dimension, the contribution of the airline's own network and that of competitors' networks to the profits are quite comparable. However, there are large differences in terms of spillovers from the HST across groups. While the negative spillovers from HST are the strongest for short routes and weakest for long routes, it is the opposite for the positive spillovers, which are strongest for long routes. This implies that more competition happens in short routes while more positive spillovers happen in longer routes.

Next, I examine the decomposition of profits for routes of different unobserved types. There are several things worth notice: first, while the overall profits of airlines go up monotonically with the unobserved route types, this is largely driven by the market characteristics. Second, although the airlines in the high type routes are the ones that enjoy the strongest positive spillovers from HST, at the same time they also suffer from the largest negative spillovers, leaving the net effect negative. Interestingly, it is the routes of median and low unobserved types where the overall spillovers from HST are positive. A similar patten can also be found if we look at the groups of routes by their average population. Together, this indicates that airlines benefit more from the presence of HST in less popular routes.

It is also interesting to break down the routes by airlines' (endogenous) entry/exit decisions. Specifically, I focus on two groups: one group contains the routes which airlines enter during the sample period and the other consists of those that airlines exit during the sample period.[51,52] Figure 9 shows the results. We can see that, for routes from the "entry" group, airlines enjoy more positive and less negative spillovers from both fast trains and bullet trains relative to the average level. For routes from the "exit" group, the opposite is mostly true: airlines suffer from a much larger negative spillover and get smaller positive spillovers from fast trains. Although the negative spillovers from the bullet trains are smaller than the average level, the positive spillovers are also smaller. The net effect therefore makes the routes less attractive. Together, these results imply that airlines reposition their networks to avoid competition but at the same time take advantage of the positive spillovers from the HST network.

=========================

Insert Figure 9 about here

=========================

10 Conclusion

This paper empirically assesses and quantifies the negative and positive spillovers of the HST network on the airline industry and studies the implications of such spillovers for firms' entry decisions (i.e., network configuration choices). I find strong evidence of heterogeneous spillover effects from the HST on airlines which depend on the routes' characteristics and their interaction with the HST network. Conditional on entry decisions, the overall contribution of positive spillovers from HST exceeds that of negative spillovers, suggesting that airlines reposition their networks to avoid competition while taking advantage of the complementarity in route connectivity.

[51] For tractability purposes, I don't include those routes in which an airline has both entry and exit during the sample period.

[52] The profits for airlines from the "exit" group are calculated for the period before they leave the market.

References

Aguirregabiria, Victor and Chun-Yu Ho (2012), "A dynamic oligopoly game of the US airline industry: Estimation and policy experiments," *Journal of Econometrics*, 168 (1), 156–173.

Aguirregabiria, Victor and Pedro Mira (2007), "Sequential estimation of dynamic discrete games," *Econometrica*, 75 (1), 1–53.

Arcidiacono, Peter, Patrick Bayer, Jason R. Blevins, and Paul B. Ellickson (2016), "Estimation of dynamic discrete choice models in continuous time with an application to retail competition," *The Review of Economic Studies*, 83 (3), 889–931.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin (2007), "Estimating dynamic models of imperfect competition," *Econometrica*, 75 (5), 1331–1370.

Basuroy, Suman, Murali K. Mantrala, and Rockney G. Walters (2001), "The impact of category management on retailer prices and performance: Theory and evidence," *Journal of Marketing*, 65 (4), 16–32.

Behrens, Christiaan and Eric Pels (2012), "Intermodal competition in the London–Paris passenger market: High-Speed Rail and air transport," *Journal of Urban Economics*, 71 (3), 278–288.

Berry, Steven T. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60 (4), 889–917.

Blaum, Joaquin, Claire Lelarge, and Michael Peters (2016), "The gains from input trade with heterogeneous importers," *Working paper*.

Bresnahan, Timothy F and Peter C Reiss (1990), "Entry in monopoly market," *The Review of Economic Studies*, 57 (4), 531–553.

Cachon, Gérard P. and A. Gürhan Kök (2007), "Category management and coordination in retail assortment planning in the presence of basket shopping consumers," *Management Science*, 53 (6), 934–951.

Ciliberto, Federico and Elie Tamer (2009), "Market structure and multiple equilibria in airline markets," *Econometrica*, 77 (6), 1791–1828.

Clapp, John M., Stephen L. Ross, and Tingyu Zhou (2016), "Retail Agglomeration and Competition Externalities: Evidence from Openings and Closings of Multiline Department Stores in the US," *Journal of Business & Economic Statistics*, 0 (forthcoming).

Clewlow, Regina, Joseph Sussman, and Hamsa Balakrishnan (2012), "Interaction of high-speed rail and aviation: exploring air-rail connectivity," *Transportation Research Record: Journal of the Transportation Research Board*, 2266, 1–10.

Clewlow, Regina R., Joseph M Sussman, and Hamsa Balakrishnan (2014), "The impact of high-speed rail and low-cost carriers on European air passenger traffic," *Transport Policy*, 33, 136–143.

Datta, Sumon and K. Sudhir (2011), "The agglomeration-differentiation tradeoff in spatial location choice," *Working paper.*

Dobruszkes, Frédéric (2011), "High-speed rail and air transport competition in Western Europe: A supply-oriented perspective," *Transport policy*, 18 (6), 870–879.

Doraszelski, Ulrich and Kenneth L. Judd (2012), "Avoiding the curse of dimensionality in dynamic stochastic games," *Quantitative Economics*, 3 (1), 53–93.

Doraszelski, Ulrich and Mark Satterthwaite (2010), "Computable Markov-perfect industry dynamics," *The RAND Journal of Economics*, 41 (2), 215–243.

Draganska, Michaela, Michael Mazzeo, and Katja Seim (2009), "Beyond plain vanilla: Modeling joint product assortment and pricing decisions," *Quantitative Marketing and Economics*, 7 (2), 105–146.

Eizenberg, Alon (2014), "Upstream innovation and product variety in the us home pc market," *The Review of Economic Studies*, 81 (3), 1003–1045.

Ericson, Richard and Ariel Pakes (1995), "Markov-perfect industry dynamics: A framework for empirical work," *The Review of Economic Studies*, 62 (1), 53–82.

Friedlaender, Alan F. and Ian Harrington (1979), "Intermodalism and integrated transport companies in the United States and Canada," *Journal of Transport Economics and Policy*, 247–267.

Gajanan, Shailendra, Suman Basuroy, and Srinath Beldona (2007), "Category management, product assortment, and consumer welfare," *Marketing Letters*, 18 (3), 135–148.

Givoni, Moshe and David Banister (2006), "Airline and railway integration," *Transport policy*, 13 (5), 386–397.

Goolsbee, Austan and Chad Syverson (2008), "How do incumbents respond to the threat of entry? Evidence from the major airlines," *The Quarterly Journal of Economics*, 123 (4), 1611–1633.

Gopinath, Gita and Brent Neiman (2014), "Trade adjustment and productivity in large crises," *The American Economic Review*, 104 (3), 793–831.

Gowrisankaran, Gautam and Marc Rysman (2012), "Dynamics of Consumer Demand for New Durable Goods," *Journal of Political Economy*, 120 (6), 1173 – 1219.

Halpern, László, Miklós Koren, and Adam Szeidl (2015), "Imported inputs and productivity," *The American Economic Review*, 105 (12), 3660–3703.

Holmes, Thomas J. (2011), "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 79 (1), 253–302.

Hotz, V. Joseph, Robert A. Miller, Seth Sanders, and Jeffrey Smith (1994), "A simulation estimator for dynamic models of discrete choice," *The Review of Economic Studies*, 61 (2), 265–289.

Igami, Mitsuru and Nathan Yang (2016), "Unobserved heterogeneity in dynamic games: Cannibalization and preemptive entry of hamburger chains in Canada," *Quantitative Economics*, 7 (2), 483–521.

Jeziorski, Przemysław (2014a), “Effects of Mergers in Two-Sided Markets: The US Radio Industry,” *American Economic Journal: Microeconomics*, 6 (4), 35–73.

— (2014b), “Estimation of cost efficiencies from mergers: Application to US radio,” *The RAND Journal of Economics*, 45 (4), 816–846.

Jia, Panle (2008), “What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry,” *Econometrica*, 76 (6), 1263–1316.

Jiang, Changmin and Anming Zhang (2014), “Effects of high-speed rail and airline cooperation under hub airport capacity constraint,” *Transportation Research Part B: Methodological*, 60, 33–49.

MacDonald, James M. (1987), “Competition and rail rates for the shipment of corn, soybeans, and wheat,” *The Rand Journal of Economics*, 151–163.

Monarch, Ryan (2016), “'It's Not You, It's Me': Breakups in US-China Trade Relationships,” *Working paper.*

Murry, Charles and Yiyi Zhou (2016), “Consumer Search and Automobile Dealer Co-Location,” *Working paper.*

Nishida, Mitsukuni (2014), “Estimating a model of strategic network choice: The convenience-store industry in Okinawa,” *Marketing Science*, 34 (1), 20–38.

Orhun, A Yeşim (2013), “Spatial differentiation in the supermarket industry: The role of common information,” *Quantitative Marketing and Economics*, 1–35.

Oum, Tae Hoon (1979a), “A cross sectional study of freight transport demand and rail-truck competition in Canada,” *The Bell Journal of Economics*, 463–482.

— (1979b), “Derived demand for freight transport and inter-modal competition in Canada,” *Journal of Transport Economics and Policy*, 149–168.

Roberts, Merrill J. (1969), “Transport Coordination and Distribution Efficiency: Pricing Norms and Profit Potential,” *Journal of Transport Economics and Policy*, 165–177.

Schiraldi, Pasquale, Howard Smith, and Yuya Takahashi (2012), “Estimating a dynamic game of spatial competition: The case of the UK supermarket industry,” *Working paper.*

Seim, Katja (2006), “An empirical model of firm entry with endogenous product-type choices,” *The RAND Journal of Economics*, 37 (3), 619–640.

Sen, Boudhayan, Jiwoong Shin, and K. Sudhir (2011), “Demand externalities from co-location: Evidence from a natural experiment,” *Working paper.*

Shen, Qiaowei and Ping Xiao (2014), “McDonald's and KFC in China: Competitors or Companions?” *Marketing Science*, 33 (2), 287–307.

Sweeting, Andrew (2010), “The effects of mergers on product positioning: evidence from the music radio industry,” *The RAND Journal of Economics*, 41 (2), 372–397.

— (2013), “Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81 (5), 1763–1803.

Viton, Philip A. (1981), "On competition and product differentiation in urban transportation: The San Francisco bay area," *The Bell Journal of Economics*, 362–379.

Vitorino, Maria Ana (2012), "Empirical entry games with complementarities: An application to the shopping center industry," *Journal of Marketing Research*, 49 (2), 175–191.

Weintraub, Gabriel Y., Lanier Benkard, and Benjamin Van Roy (2006), "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," in "Advances in neural information processing systems," 1489–1496.

Wilson, William W., Wesley W. Wilson, and Won W. Koo (1988), "Modal competition in grain transport," *Journal of Transport Economics and Policy*, 319–337.

Xia, Wenyi and Anming Zhang (2016), "Effects of air and high-speed rail transport integration on profits and welfare: The case of air-rail connecting time," *Working paper*.

Yang, Nathan (2012), "Burger King and McDonald's: Where's the Spillover?" *International Journal of the Economics of Business*, 19 (2), 255–281.

Zhu, Ting and Vishal Singh (2009), "Spatial competition with endogenous location choices: An application to discount retailing," *Quantitative Marketing and Economics*, 7 (1), 1–35.

Figures and Tables

Figure 1: Evolution of the HST Network

This figure shows evolution of the high speed train (HST) network over years. I use different color to distinguish between fast train and bullet train rails.

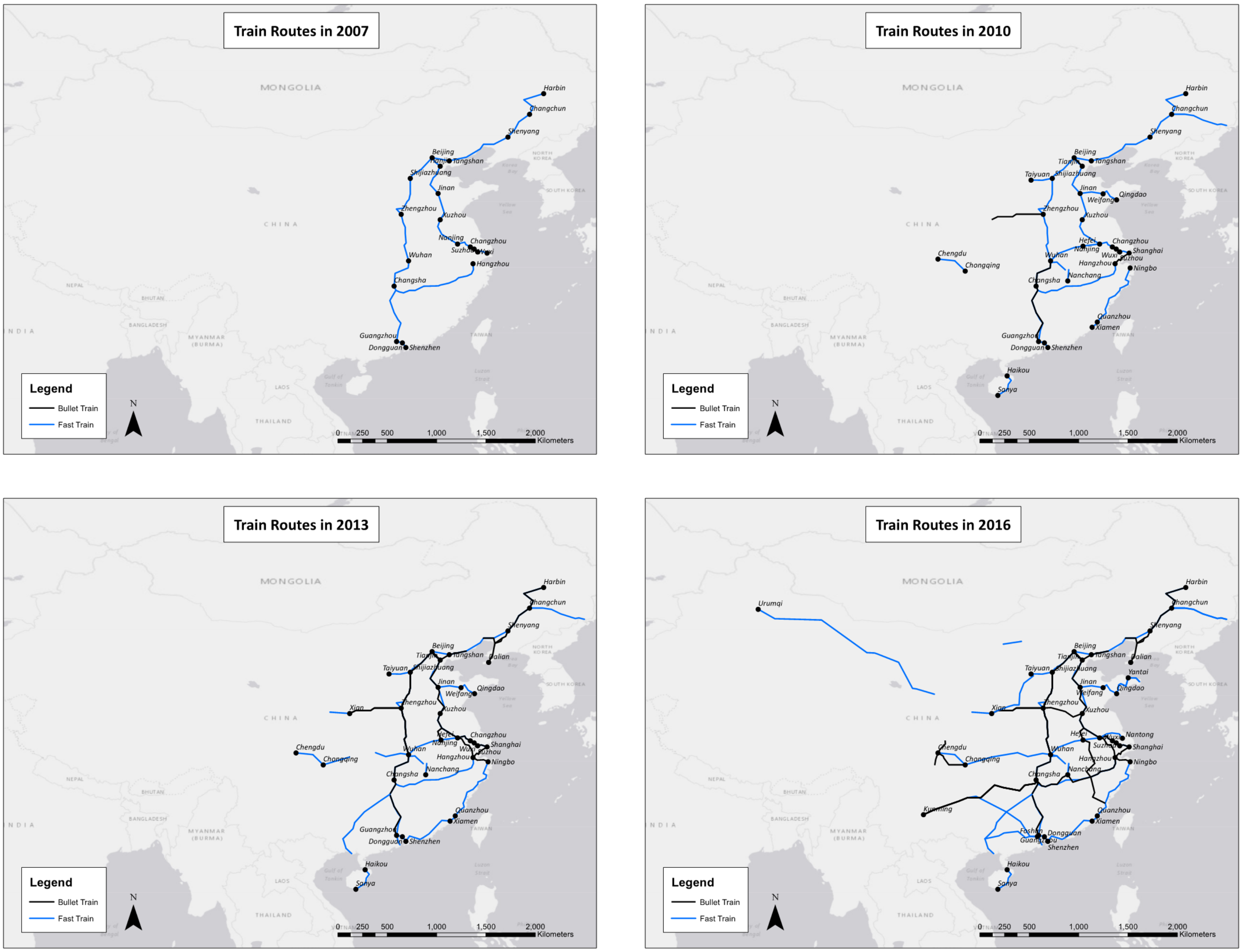

Figure 2: Airlines by Parent Company

This figure shows the airline companies in China. Most of the airlines are subsidiaries of the top four airlines. The market shares are calculated in terms of total number of flights among the top 68 cities in China from 2006 to 2016.

Figure 3: Number of Overlapping Routes by Year

This figure shows the number of routes provided by airlines and number of routes provided by both airlines and high speed trains over the years.

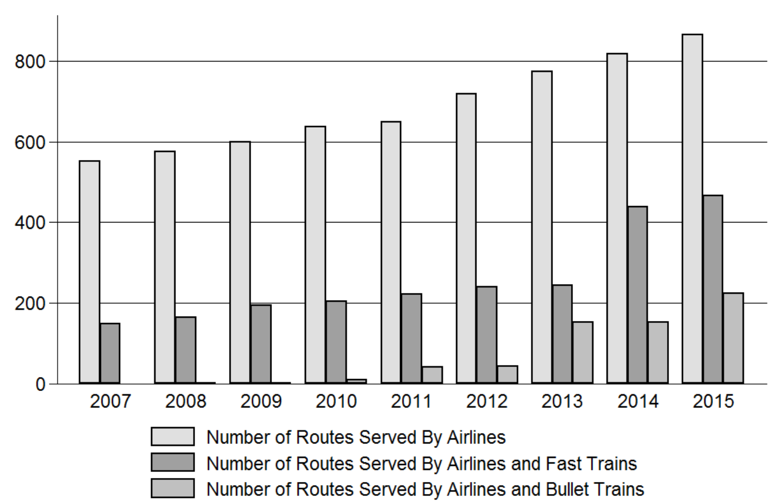

Figure 4: Travel Time as A Function of Distance by Transportation Modes

This figure shows the relation between travel time and travel distance for both bullet trains and airlines. (Source: "Civil Aviation Big Data")

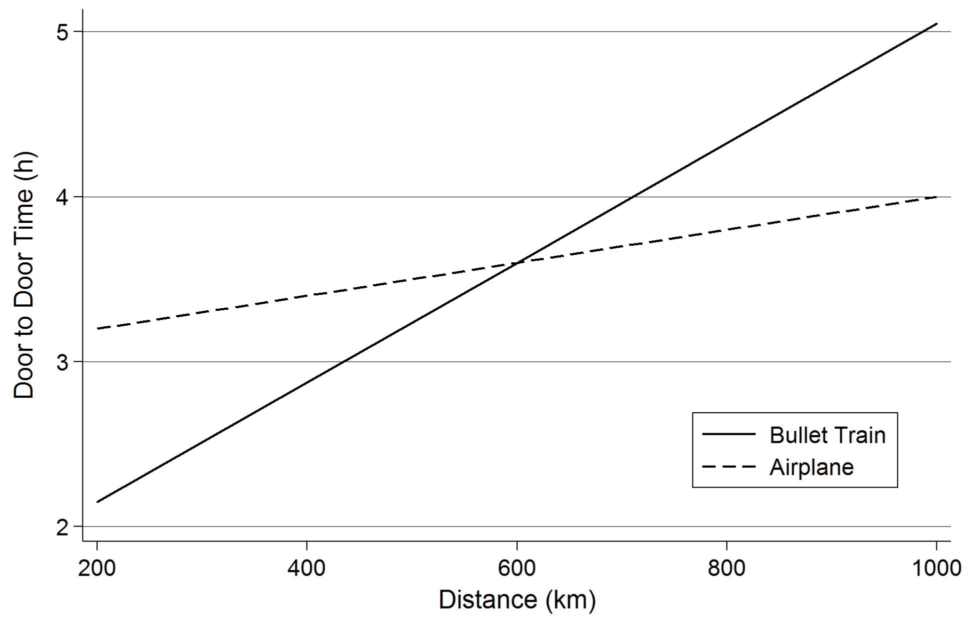

Figure 5: Number of Flights by Group: Short Routes

This figure shows the number of flights provided in routes with and without direct competition from high speed trains. The graph is for short routes only. The dotted line represents the number of flights provided in routes in the control group, where there is no HST competition, and the dashed line represents the corresponding number for routes in the treatment group which face competition from HST at different points of time. The bar graphs denotes the proportion of routes that face competition from HST in the treatment group.

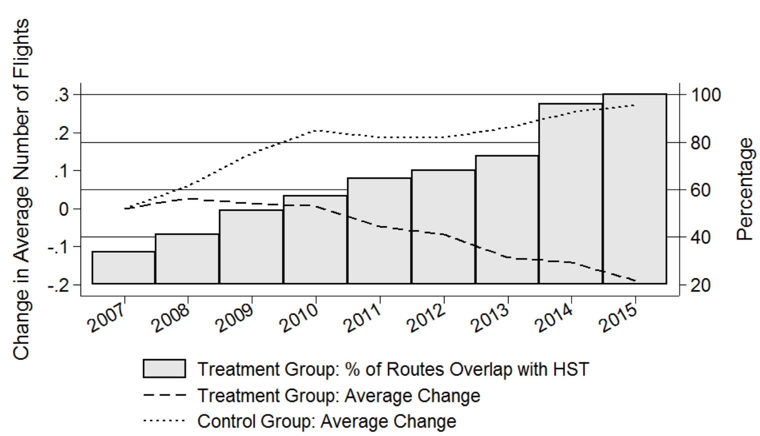

Figure 6: Number of Flights by Group: Median Length Routes

This figure shows the number of flights provided in routes with and without direct competition from high speed trains. The graph is for median length routes only. The dotted line represents the number of flights provided in routes in the control group, where there is no HST competition, and the dashed line represents the corresponding number for routes in the treatment group which face competition from HST at different points of time. The bar graphs denotes the proportion of routes that face competition from HST in the treatment group.

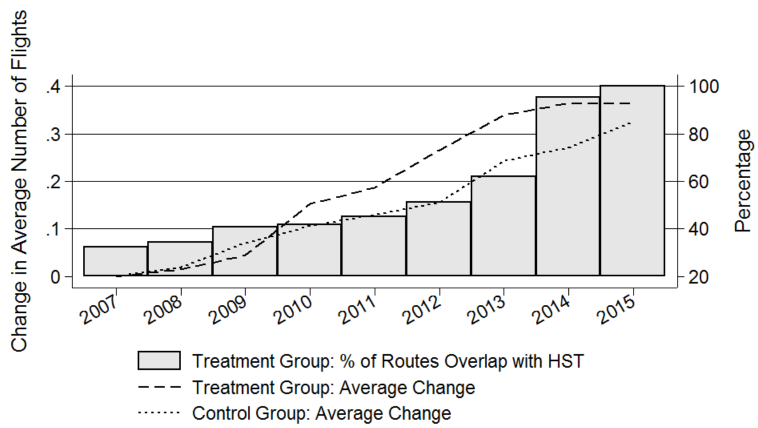

Figure 7: Number of Flights by Group: Long Routes

This figure shows the number of flights provided in routes with and without direct competition from high speed trains. The graph is for long routes only. The dotted line represents the number of flights provided in routes in the control group, where there is no HST competition, and the dashed line represents the corresponding number for routes in the treatment group which face competition from HST at different points of time. The bar graphs denotes the proportion of routes that face competition from HST in the treatment group.

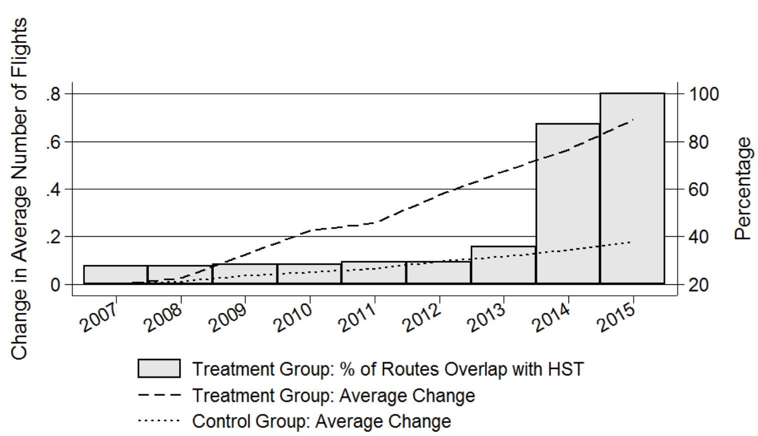

Figure 8: Selection of Control Routes in the Difference-in-difference Analysis

This figure shows the principles for selecting the control routes.

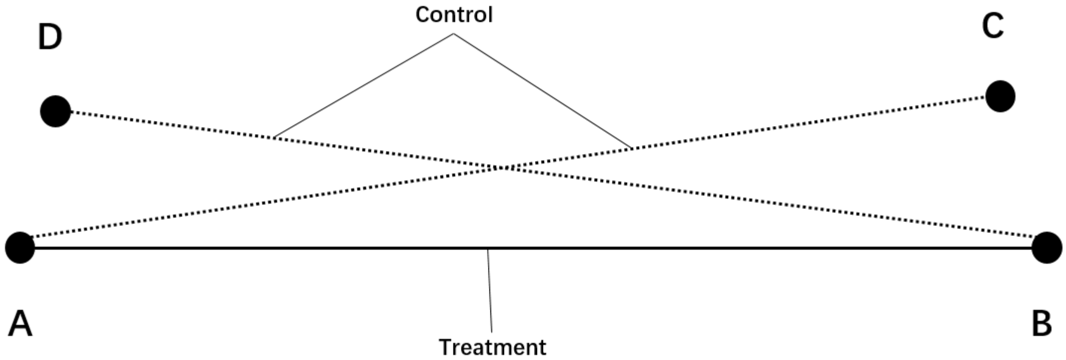

Figure 9: Profit Decomposition Conditional on Entry/Exit Decisions

This figure shows the decomposition of positive/negative spillovers from HST conditional on airlines' entry/exit decisions. I present the decomposition for three groups: all routes, routes where airlines enter, and routes where airlines exit. For the "eixt" group, the profits are calculated using periods before the airlines leave the market.

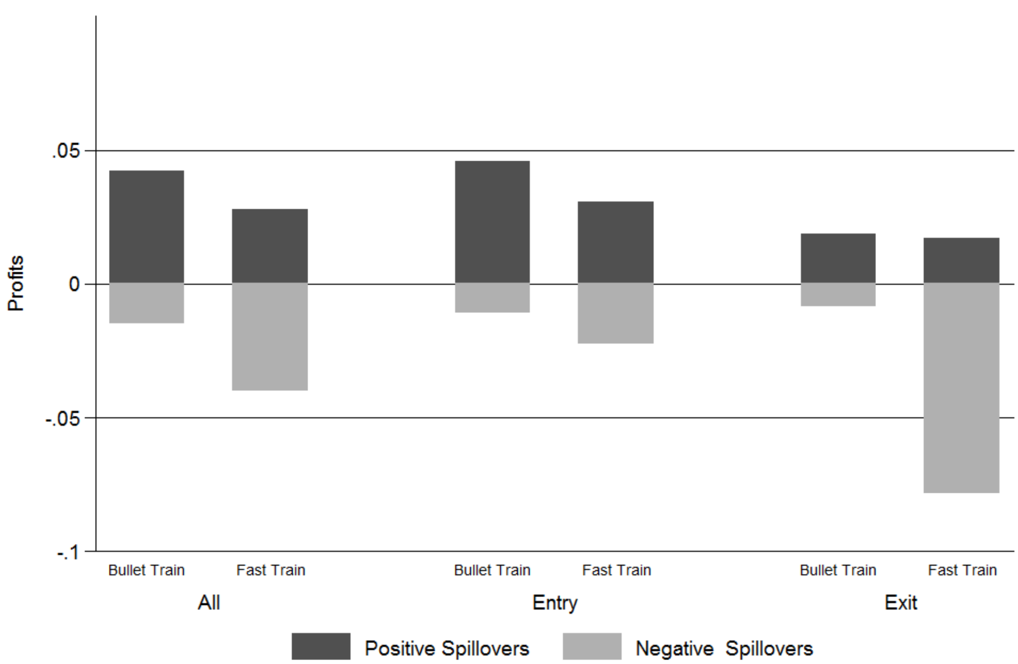

Table 1: Summary Statistics

This table reports the summary statistics of the dataset. The statistics are calculated at the route-year level.

	N	Mean	Median	S.D.	Min	Max
Number of airlines	20502	0.53	0.00	0.98	0.00	4.00
Number of flights	20502	1.05	0.00	2.96	0.00	52.00
Number of entries	20502	0.05	0.00	0.24	0.00	3.00
Number of exits	20502	0.03	0.00	0.18	0.00	3.00
Number of flight increases	20502	0.06	0.00	0.28	0.00	3.00
Number of flight decreases	20502	0.08	0.00	0.36	0.00	4.00
Fast train present	20502	0.11	0.00	0.32	0.00	1.00
Bullet train present	20502	0.03	0.00	0.17	0.00	1.00
City pair average population (Million)	20502	5.34	4.82	3.40	0.22	24.19
City pair average GDP (Billion)	20502	55.56	43.05	45.29	0.49	374.57
City pair average population growth rate	20502	0.01	0.01	0.01	-0.05	0.06
City pair average GDP growth rate	20502	0.16	0.15	0.02	0.11	0.27
City pair distance (00's km)	20502	15.26	13.89	8.71	0.54	44.06

Table 2: Summary Statistics by Year

This table reports the summary statistics by year. The statistics are calculated at the route level.

	2008	2009	2010	2011	2012	2013	2014	2015
Number of airlines	0.43	0.47	0.50	0.51	0.55	0.58	0.61	0.65
Number of flights	0.81	0.91	0.99	1.06	1.14	1.20	1.27	1.32
Number of entries	0.04	0.06	0.06	0.06	0.06	0.07	0.07	0.08
Number of exits	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04
Number of flight increases	0.05	0.07	0.07	0.07	0.08	0.07	0.09	0.08
Number of flight decreases	0.03	0.02	0.03	0.03	0.03	0.05	0.05	0.07
Fast train present	0.07	0.08	0.09	0.10	0.10	0.11	0.19	0.21
Bullet train present	0.00	0.00	0.00	0.02	0.02	0.07	0.07	0.10
Observations	2278	2278	2278	2278	2278	2278	2278	2278

Table 3: GDP Quantiles and Airline Decisions

This table reports airlines' flight decisions by quantiles of average city-pair GDP.

	1	2	3	4	5
Average GDP of the City Pair	1.15	2.76	4.34	6.62	12.91
Number of flights provided by an airline	0.03	0.07	0.10	0.22	0.90
Probability of entry	0.00	0.01	0.01	0.03	0.04
Probability of exit	0.00	0.01	0.04	0.06	0.05
Probability of flight increase	0.01	0.02	0.04	0.08	0.16
Probability of flight decrease	0.03	0.06	0.16	0.27	0.21
Observations	16404	16400	16404	16400	16400

Table 4: Population Quantiles and Airline Decisions

This table reports airlines' flight decisions by quantiles of average city-pair population.

	1	2	3	4	5
Average Population of the City Pair	1.68	3.60	4.84	6.37	10.21
Number of flights provided by an airline	0.04	0.09	0.17	0.26	0.76
Probability of entry	0.00	0.01	0.02	0.03	0.04
Probability of exit	0.01	0.03	0.04	0.04	0.04
Probability of flight increase	0.02	0.03	0.05	0.07	0.12
Probability of flight decrease	0.03	0.07	0.09	0.16	0.20
Observations	16404	16400	16408	16396	16400

Table 5: Income perCapita Quantiles and Airline Decisions

This table reports airlines' flight decisions by quantiles of average city-pair income per capita.

	1	2	3	4	5
	mean	mean	mean	mean	mean
Average Income per Capita of the City Pair	0.00	0.01	0.01	0.01	0.02
Number of flights provided by an airline	0.06	0.11	0.20	0.31	0.63
Probability of entry	0.01	0.01	0.02	0.02	0.03
Probability of exit	0.01	0.03	0.04	0.06	0.05
Probability of flight increase	0.01	0.03	0.07	0.13	0.15
Probability of flight decrease	0.06	0.11	0.18	0.19	0.20
Observations	16404	16400	16404	16400	16400

Table 6: Route Length Quantiles and Airline Decisions

This table reports airlines' flight decisions by quantiles of route length.

	1	2	3	4	5
City pair distance (00's km)	4.90	9.86	13.89	18.61	29.09
Number of flights provided by an airline	0.38	0.43	0.29	0.17	0.04
Probability of entry	0.02	0.03	0.02	0.01	0.00
Probability of exit	0.05	0.04	0.03	0.02	0.01
Probability of flight increase	0.08	0.08	0.08	0.06	0.03
Probability of flight decrease	0.20	0.17	0.11	0.07	0.02
Observations	16416	16416	16380	16416	16380

Table 7: Entry Probability in A Route

This table reports the marginal effect of airport presence on an airline's probability of entering a route. Airline fixed effects as well as year fixed effects are included.

Airline operates in one end point airport in the previous year	0.0045
(single presence)	(0.0005)
Airline operates in both end point airports in the previous year	0.0429
(dual presence)	(0.0009)
N	65,185

Table 8: GDP Quantiles and High Speed Train Presence

This table reports the average presence of high speed trains by average GDP of the city-pair.

	1	2	3	4	5
Average GDP of the City Pair	1.15	2.76	4.34	6.62	12.91
Fast train present	0.00	0.02	0.06	0.15	0.34
Bullet train present	0.00	0.00	0.00	0.02	0.13
Observations	16404	16400	16404	16400	16400

Table 9: Population Quantiles and High Speed Train Presence

This table reports the average presence of high speed trains by average population of the city-pair.

	1	2	3	4	5
Average Population of the City Pair	1.68	3.60	4.84	6.37	10.21
Fast train present	0.01	0.04	0.07	0.16	0.29
Bullet train present	0.00	0.01	0.02	0.04	0.09
Observations	16404	16400	16408	16396	16400

Table 10: Income per Capita Quantiles and High Speed Train Presence

This table reports the average presence of high speed trains by average income per capita of the city-pair.

	1	2	3	4	5
Average Income per Capita of the City Pair	0.00	0.01	0.01	0.01	0.02
Fast train present	0.01	0.05	0.09	0.15	0.26
Bullet train present	0.00	0.00	0.01	0.03	0.12
Observations	16404	16400	16404	16400	16400

Table 11: Route Length Quantiles and High Speed Train Presence

This table reports the average presence of high speed trains by quantiles of route distance.

	1	2	3	4	5
City pair distance (00's km)	4.90	9.86	13.89	18.61	29.09
Fast train present	0.27	0.18	0.08	0.03	0.00
Bullet train present	0.09	0.05	0.02	0.00	0.00
Observations	16416	16416	16380	16416	16380

Table 12: Number of Flights by Group

This table reports the average number of flights provided in treatment and control groups. The routes in the treatment group are ones with at least three connections to HST lines. The observations in each group are paired using a matching algorithm such that two routes have similar route length, level of average population and size of connecting routes. Both results using the algorithm with and without replacement are reported.

	Number of Flights	All	Treatment	Control	P value
W/ Replacement	Average	0.42	0.51	0.32	0.01
	Std Dev.	0.75	0.88	0.58	0
	Observations	200	100	100	
W/O Replacement	Average	0.41	0.50	0.32	0.04
	Std Dev.	0.76	0.89	0.59	0
	Observations	164	82	82	

Table 13: Regression Results

This table reports regression results. The unit of analysis is at the airline-route-year level and the dependent variables are number of flights (Columns 1 through 4) and whether flights are provided (Columns 5 and 6). I use a linear regression and a logistic regression for the two dependent variables respectively.

	(1) Number of Flights	(2) Number of Flights	(3) Number of Flights	(4) Number of Flights	(5) Flights (Y/N)	(6) Flights (Y/N)
Fast Train present (Y/N)	0.291***	-0.387***	-0.265***	-0.390***	-1.743***	-1.728***
	(0.015)	(0.025)	(0.021)	(0.025)	(0.097)	(0.237)
Bullet Train present (Y/N)	-0.378***	-1.292***	-0.690***	-0.937***	-2.070***	-4.182***
	(0.034)	(0.049)	(0.041)	(0.030)	(0.220)	(0.368)
No. of Fast Train line connections	0.210***	0.048***	-0.056***	0.030***	0.293***	0.033
	(0.008)	(0.008)	(0.007)	(0.007)	(0.033)	(0.072)
No. of Bullet Train line connections	0.248***	0.085***	-0.005	0.230***	0.209***	0.211**
	(0.009)	(0.009)	(0.008)	(0.007)	(0.036)	(0.066)
Fast Train × No. of	0.494***	0.255***	0.118***	0.021	-0.482***	-0.489***
Fast Train line connections	(0.020)	(0.019)	(0.016)	(0.015)	(0.066)	(0.118)
Bullet Train × No. of	0.521***	0.448***	0.255***	-0.084***	-0.310**	-0.105
Bullet Train line connections	(0.033)	(0.032)	(0.027)	(0.020)	(0.116)	(0.172)
Average Population		0.000***	-0.000	0.001***	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Length of Route		-0.084***	-0.055***	0.133***	-1.027***	-4.209***
		(0.004)	(0.004)	(0.029)	(0.026)	(0.680)
Regulated Route		0.962***	-0.090***	-2.553***	-0.990***	0.694
		(0.013)	(0.014)	(0.114)	(0.057)	(0.784)
Fast Train × distance		0.680***	0.341***	0.555***	2.139***	2.059***
		(0.024)	(0.021)	(0.025)	(0.087)	(0.215)
Bullet Train × distance		1.512***	0.680***	1.283***	2.115***	3.899***
		(0.050)	(0.042)	(0.031)	(0.209)	(0.342)
No. of own routes connected			0.069***	0.052***	0.254***	0.303***
			(0.001)	(0.001)	(0.004)	(0.008)
Existence of own indirect flights			-0.090***	-0.055***	0.857***	0.637***
			(0.007)	(0.005)	(0.050)	(0.078)
No. of competitors' flights			0.161***	-0.549***	0.175***	-1.027***
			(0.002)	(0.002)	(0.008)	(0.025)
No. of competitors' routes connected			-0.011***	0.021***	0.004**	0.024***
			(0.000)	(0.001)	(0.002)	(0.006)
Existence of competitors' indirect flights			-0.096***	-0.034***	-0.019	0.411*
			(0.009)	(0.008)	(0.098)	(0.197)
Headquarter			1.446***	0.586***	0.206*	0.112
			(0.023)	(0.015)	(0.091)	(0.160)
Constant	0.016**	0.034*	-0.018	-0.875***	-5.445***	0.894
	(0.005)	(0.014)	(0.016)	(0.107)	(0.123)	(0.911)
Year Dummies	No	Yes	Yes	Yes	Yes	Yes
Airline Dummies	No	No	Yes	Yes	Yes	Yes
City Pair Dummies	No	No	No	Yes	No	Yes
Observations	82008	82008	82008	82008	82008	32832
R^2	0.100	0.209	0.443	0.784		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: Difference-in-Differences Regression

This table reports difference-in-difference regression results. The unit of analysis is at the airline-route-year level and the dependent variable is the number of flights an airline provides in a route in a given year. The key difference between this regression and the previous linear regression is that for each route that overlaps with HST, I include two control routes to capture the demand and cost shocks to the routes. I assign the same group ID for the treated route (i.e., the route that overlaps with HST) and its corresponding control routes. I include group-year dummies in the regression, which allows for each group to have their own flexible trend in number of flights.

	All	Short	Median	Long
HST present	0.176***	-0.430***	0.239***	0.591***
	(0.039)	(0.116)	(0.049)	(0.082)
No. of HST line connections	0.059*	-0.028	0.062	0.169**
	(0.030)	(0.092)	(0.035)	(0.065)
Average Population	0.001	-0.009***	0.000	0.005***
	(0.001)	(0.002)	(0.001)	(0.001)
Regulated Route	-1.534	-15.583***	18.310***	5.717***
	(1.033)	(3.657)	(1.125)	(1.548)
No. of own routes connected	0.089***	0.044*	0.096***	0.088***
	(0.005)	(0.019)	(0.007)	(0.010)
Existence of own indirect flights	-0.181**	-0.176	-0.169*	-0.228*
	(0.057)	(0.150)	(0.071)	(0.114)
No. of competitors' flights	-0.388***	-0.414***	-0.365***	-0.380***
	(0.010)	(0.024)	(0.013)	(0.020)
No. of competitors' routes connected	0.020***	0.019	0.010*	0.044***
	(0.004)	(0.013)	(0.005)	(0.008)
Existence of competitors' indirect flights	-0.140	-0.202	-0.123	0.008
	(0.101)	(0.218)	(0.122)	(0.300)
Headquarter	1.271***	2.021***	0.852***	1.735***
	(0.095)	(0.284)	(0.149)	(0.185)
Constant	0.217	18.178***	-18.654***	-11.682**
	(1.153)	(4.642)	(1.231)	(3.650)
Year Dummies	Yes	Yes	Yes	Yes
Group Dummies	Yes	Yes	Yes	Yes
Year Dummies×Group Dummies	Yes	Yes	Yes	Yes
City Pair Dummies	Yes	Yes	Yes	Yes
Airline Dummies	Yes	Yes	Yes	Yes
Observations	9039	1728	5157	2154
R^2	0.870	0.853	0.850	0.920

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Estimation Results of The Structural Parameters

This table reports the estimation results for the structural model.

		Coef.	std.err	t stat
Impact of HST	Bullet Train present (Y/N)	-0.18	0.10	-1.82
	Bullet Train × Median distance	0.12	0.10	1.19
	Bullet Train × Long distance	0.05	0.12	0.43
	No. of Bullet Train line connections	0.07	0.02	2.74
	Bullet Train × No. of Bullet train line connections	-0.03	0.05	-0.64
	Fast Train present (Y/N)	-0.27	0.07	-4.10
	Fast Train × Median distance	0.23	0.07	3.21
	Fast Train × Long distance	0.46	0.08	6.00
	No. of Fast Train line connections	0.06	0.02	2.84
	Fast Train × No. of Fast Train line connections	-0.09	0.03	-2.74
Own Network	No. of own routes connected	0.17	0.01	11.68
CompetitorsÔÇÖ Networks	No. of competitors	-0.10	0.05	-1.83
	No. of competitors' routes connected	-0.01	0.01	-0.70
	No. of competitors × Competitors' routes connected	-0.02	0.02	-1.01
Market Characteristics	Population	0.00	0.01	0.18
	Median distance	-0.07	0.03	-2.23
	Long distance	-0.21	0.03	-6.30
	Unobserved Type	0.15	0.02	9.06
	Constant	-0.05	0.05	-0.96
Entry Costs	Entry Cost	-3.99	0.06	-64.46
	Entry Cost × Unobserved Type	0.30	0.05	6.21
	Control	-1.39	0.14	-10.09
	Control-Exempt	1.32	0.14	9.49

Table 16: Variables Used in the Sources of Profits

This table reports the variables I include in the analysis of the sources of profits.

Source	Variables
Own Network	No. of own routes connected
Competitor's Network	No. of competitors No. of competitors' routes connected
Fast Train: Negative	Presence of Fast Train Presence of Fast Train × the length of the route Presence of Fast Train × No. of Fast Train line connections
Bullet Train: Negative	Presence of Bullet Train Presence of Bullet Train × the length of the route Presence of Bullet Train × No. of Bullet Train line connections
Fast Train: Positive	No. of Fast Train line connections
Bullet Train: Positive	No. of Bullet Train line connections
Market Specific	Constant Population Route length Unobserved market type

Table 17: Flow Profit Decomposition

This table reports the flow profit decomposition for airlines. The columns represent the sources of profits and rows represent the groups of routes that are divided along different dimensions.

		Average Profit	Own Network	Compeitors' Network	Fast Train Negative	Bullet Train Negative	Fast Train Positive	Bullet Train Positve	Market Specific
	Overall	0.225	0.810	-0.181	-0.040	-0.015	0.028	0.042	-0.420
Length	Low	0.190	0.741	-0.155	-0.107	-0.013	0.020	0.029	-0.324
	Median	0.227	0.793	-0.189	-0.055	-0.020	0.025	0.040	-0.367
	High	0.237	0.862	-0.181	0.007	-0.009	0.034	0.049	-0.526
Population Level	Low	0.040	0.584	-0.121	-0.005	0.000	0.013	0.001	-0.432
	MeidanLow	0.100	0.692	-0.148	-0.015	-0.002	0.015	0.008	-0.449
	Median	0.170	0.771	-0.163	-0.029	-0.006	0.024	0.028	-0.455
	MedianHigh	0.192	0.779	-0.166	-0.037	-0.009	0.027	0.037	-0.439
	High	0.303	0.886	-0.206	-0.054	-0.025	0.033	0.060	-0.390
Unobserved Market Type	Low	-0.531	0.704	-0.039	0.000	0.000	0.023	0.027	-1.247
	MeidanLow	-0.098	0.728	-0.056	0.000	-0.002	0.025	0.033	-0.825
	Median	0.128	0.762	-0.121	-0.019	-0.005	0.025	0.035	-0.549
	MedianHigh	0.322	0.855	-0.241	-0.057	-0.020	0.029	0.047	-0.290
	High	0.643	0.950	-0.341	-0.121	-0.063	0.040	0.070	0.107