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

Marco Shaojun Qin
University of Minnesota

Job Market Paper
Feb 02, 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. This paper analyzes the impact of HST roll-out on air carriers’ route network decisions. Using a unique hand-collected dataset on airline networks and the timing of the HST roll-out, I first present reduced-form 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. With the structural estimates, I carry out policy experiments in which the HST had not been introduced. Comparing the air carriers’ route networks in the counterfactual scenario with those in a scenario with HST roll-out, I find that the introduction of HST has reduced air carriers’ route presence by about 15%, and it is usually the cities located in the geographically central area of China that experience the largest decrease. In another policy experiment, I improve the positive spillover effects from HST and predict the corresponding route networks of the air carriers. I find that, to compensate for the overall negative impact of HST on the airline industry, the HST needs to feed at least 40% more traffic to the air 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

Firms with new technologies and/or business models usually bring disruption to the existing industry. Examples include the expansion of Walmart in the retail industry, the introduction of “e-commerce” represented by Amazon.com and its impact on brick-and-mortar stores, and the recent emergence of “sharing economy” such as Uber and Airbnb and the disruption they have brought to the transportation and hotel industries, respectively.

What is often ignored, however, is the fact that while these new companies bring competition to the existing industries, they may also create opportunities for the incumbent firms. For example, although Didi, a Chinese “ride-hailing” service is generally considered as an alternative to the traditional taxi service, it is possible that this platform may feed traffic to traditional taxi carriers for certain routes within their overall network.\footnote{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.}\footnote{In fact, an incumbent Taxi company in Shanghai has already started collaborating with Didi to provide better hailing services for consumers. Source: http://www.szdaily.com/content/2016-05/03/content_13191679.htm.} Also, Best Buy has been 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.\footnote{Source: http://www.chainstoreage.com/article/how-smart-digital-strategy-can-make-physical-retail-store-prime-asset} And even Walmart has been found to drive customers to nearby stores.\footnote{Source: http://www.cbsnews.com/news/the-myth-of-the-walmart-effect/}

Do such positive spillovers affect incumbent firms’ marketing mix decisions, and if so, how? Despite the ubiquitous phenomenon and the importance of the question, empirical studies on this topic is sparse. In this paper, I look at the airline industry in China, which faces increasing competition from High-speed trains (HST) over the past decade. I empirically quantify the negative and positive spillovers from HST introduction to the airline industry and study how airline carriers reposition their route network in response.

I choose this setting for two reasons: first, China has the fastest growing high speed train (HST) network in the world. As is shown in Figure 1, China has built (or upgraded) over 20,000 km (12,500 miles) of high speed railways over the past decade. The length is 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.\footnote{Source: http://www.economist.com/news/china/21714383-and-theres-lot-more-cope-it-waste-money-china-has-built-worlds-largest}
and opportunities to domestic air carriers’ operations. On one hand, HST service and air travel are often acknowledged to be reasonable substitutes. On the other hand, however, 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. In fact, online travel brokers often offer bookings that combine flight and HST segments.5

This paper has two objectives. The first objective is to assess the impact of the HST roll-out on airline carriers’ route network decisions. How strong are these positive and negative spillovers from HST, and how do air carriers make their decisions accordingly with respect to where to provide air services? Despite the heated discussion regarding the challenges and opportunities HST has brought to the airline carriers, there is no empirical evidence that quantifies the effects. More generally, it has largely been under-studied how firms reposition their products to avoid competition while taking advantage of positive spillovers. This paper thus seeks to contribute to this gap. Finally, by studying the case of 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.7

A related research question is whether and how much improvement in the positive spillovers is needed for the airline carriers to compensate for the negative impact of HST roll-out. This is an important question because China has specifically emphasized the need to integrate inter-modal services between airplanes and trains in its recent five-year plan.8 The answer to this question thus may help local government determine what level of services are needed to facilitate the growth of the airline industry. Also, depending on their characteristics, different cities may experience different levels of increase in airline presence as a result of such improvement. It is therefore interesting to pinpoint the cities that gain the most airline routes due to the change.

I hand-collect daily flight information for the four major passenger airlines and their subsidiaries in China and use this data together with detailed information on the timing of the HST roll-out.9 The data covers over two thousand city-pairs (i.e., routes) for a period of nine years. I exploit the variation 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 use both reduced-form analysis and structural modeling to answer the research questions. I first use a difference-in-differences regression to establish the existence of both negative and potential positive spillovers from HST to the airline industry.10 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

---

6For example, in United States, local airline carriers have been opposing the idea of HST introduction for years (For references, see https://www.citylab.com/transportation/2015/05/southwest-airlines-hasnt-decided-whether-or-not-to-oppose-texas-high-speed-rail/392462/). 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.
8The four major airlines in China (including their subsidiaries) account for almost 90% of the domestic air passenger traffic in China.
9Goolsbee 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.
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.

Next, to assess the impact of HST introduction on airline carriers’ route network decisions, I construct and estimate a structural dynamic oligopoly model of airlines network configurations. 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’s routes decisions. The effects depend on the routes’ characteristics and on their interaction with the HST network. Specifically, I conclude that the HST is a strong substitute for 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 sixth 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.

Using the estimates from the structural model, I explore a counterfactual scenario in which the HST had not been introduced. Comparing the predicted route networks for the counterfactual scenario with those for the baseline scenario, in which HST is introduced, I find that overall, the introduction of HST has reduced the air carriers’ route presence by about 15%. This suggest that the negative spillovers from HST outweigh the positive spillovers. I also find that it is usually cities located at the geographically central area of China that experience the most reduction route presence, because HST network is usually denser in these areas and that the average distance from the cities to those from the rest of the country is small, which makes HST a more favorable substitute
for traveling. Correspondingly, cities located in the geographically peripheral area of China are less affected, with some even experience an increase in the airline route presence.

In the second counterfactual analysis, I increase the air carriers’ profits associated with HST line connection by different levels, and study how much improvement is needed to compensate for the overall negative impact of HST on the airline industry. I find that, when the profits from connecting to HST lines are increased by 40%, the corresponding air carriers’ route presence is comparable to that in the scenario where HST had not been introduced. This implies that, assuming air carriers’ marginal costs of providing inter-modal services are small, the number of customers using inter-modal travel need to increase by at least 40% in order to shield the airline industry from the negative impact from HST. I also find that such improvement does not affect all route uniformly. In fact, apart from cities from the geographically central area of the country, cities from the North and Northeast part of China experience the largest increase in airline route presence when the positive spillovers from HST are improved.

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 11 concludes.

2 Literature

This paper is related to several streams of literature. First, to the extent that airlines can be regarded as multi-product firms that differentiate themselves through their route network configurations, this paper is 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.

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.\(^{11}\) 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.\(^{12}\) 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, 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.

Finally, this paper contributes to the transportation economics literature that studies intermodal

---

\(^{11}\)For papers that have a similar question but that use non-structural methods, see Clapp, Ross, and Zhou (2016) and Shen and Xiao (2014).

\(^{12}\)Some other papers explore the spillovers from a demand perspective, e.g., Sen, Shin, and Sudhir (2011) and Murry and Zhou (2016).
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.

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.

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.\textsuperscript{13}

\textsuperscript{13}Here and henceforth, all numbers are calculated for flights that connects the top 68 cities in China in terms of airport passenger volume in 2015.
The airline industry in China is growing at a steady rate. From 2006 to 2016, the total number of flights has increased by about 87%, and the total number of city-pairs served by the industry 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, air 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. 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.

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 pursues social-welfare related objectives and is not a profit maximizing firm. 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.

\[^{14}\text{Henceforth, I will use routes, markets and city-pairs interchangeably.}\]

\[^{15}\text{The decision to drop a flight does not require any permits.}\]

\[^{16}\text{To see a complete description of the company’s goals, please refer to http://www.china-railway.com.cn/en/aboutus.}\]
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.

The first fast train routes and bullet train routes were established in 2007 and 2008 respectively, connecting major cities such as Beijing, Tianjin, Wuhan, Shanghai and Guangzhou. 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. 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.

3.3 Impact of HST Roll-out on the Airline Industry

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. 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 30% of the airline routes are also served by fast trains, and more than 15% are also served by bullet trains.

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 presents 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, which increases the time needed for commute.

17 Upgraded existing rails can only accommodate fast trains and not bullet trains. There is no difference in terms of speed for fast trains operating on newly built rails and on upgraded existing rails.
18 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.
19 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.
Further, 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, I expect that HST has more strategic advantage in shorter routes.

---

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 but 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,\textsuperscript{22} which accounted for about 5.5% of the total passenger volume of the airport in that year.\textsuperscript{23,24}

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 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.\textsuperscript{25} This

\textsuperscript{22}Source: http://news.ifeng.com/a/20161227/50482906_0.shtml (in Chinese).
\textsuperscript{24}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.
accounts for only about 0.6% of total passenger traffic of the China Eastern Airlines in the HongQiao airport.\footnote{Source: https://www.cybersource.com/content/dam/cybersource/zh-APAC/documents/China_Eastern_Airlines_Case_Study_SC.pdf (in Chinese).} 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\footnote{Source: https://static.fraport.de/ONLINE/zdf/zadafa_e_2015/} 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 GDP, 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.\footnote{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.} I further refer to the China City Statistical Yearbook to collect GDP and GDP data for cities in China.\footnote{Unlike the first two data sources, the GDP data is at a yearly level. I therefore treat the GDP and GDP as constant within a year.} Finally, data for city coordinates come from Google Maps and is used to calculate the air distance between cities.

4.2 Sample Selection

Below I discuss the criteria I use for the 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 cover more than 95\% of the market and account for about 90\% of market share in number of flights.

**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.\footnote{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.} 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.\footnote{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.} 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. 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 excluded. To deal with this issue, I drop 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 ($4 \times 2278 \times 9 = 82008$). 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 services (especially given the fact that there were no HST until 2007).

For the city-pairs and years studied, GDP grows at the rate of 16% per year. The average length of routes is about 1500 km.

Table 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/flight 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.

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. 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 routes for all air carriers where entry is possible. Exit probability, probability of increasing flights, and probability of decreasing flights are similarly defined. 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.
Table 4 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, as well as 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.

Airlines' Route 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 5 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 eight times larger than that of operating in only one of the airports.

5.2 Entry Patterns of HST

I next explore the patterns associated with the introduction of the HST. Table 6 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. 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.

Table 7 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.

5.3 How Do Airlines Respond to the Introduction of HST?

Here I analyze how airlines respond to the introduction of HST. I first provide some descriptive evidence of the existence of both negative and positive spillovers from HST to the airline industry, and then move on to the regression analysis that further quantifies the effects of HST on airlines’ route decisions.

To begin with, I present the evolution of the number of airlines serving a route that is affected by HST introduction. Specifically, I divide airline routes into three categories: routes that overlap with HST, routes that does not overlap with but connect to HST lines, and routes that neither overlap with nor connect to HST. Motivated by the previous discussion regarding the differences in the competitive advantage of airlines and HST, I further divide routes that overlap with HST into two groups using 600 km as a cut-off value. The routes less than 600km are referred to as short routes, and the rest are referred to as long routes.

Figure 5 presents the evolution of the average number of airlines for routes from different groups. Specifically, sub-figures (a) and (b) correspond to short and long routes that overlap with HST, respectively. Sub-figure (c) corresponds to routes that does not overlap with HST but connect to HST lines, and finally, sub-figure (d) corresponds to routes that neither overlap with HST nor connect to HST lines. Comparing sub-figures (a) and (b), we can find that, although routes from both groups overlap with HST, the evolution of the average number of airlines follow very different patterns: specifically, while the number of airlines goes down gradually for short routes (sub-figure(a)), it goes

---

\[\text{A route is served by an airline if the airline provides direct flight on that route.}\]

\[\text{A route connects to HST line if either one of the end city has a HST station.}\]

\[\text{Because HST is introduced gradually to the market, I divide the routes into different groups based on the last year of the data, which is 2015. As a result, some routes may not yet overlap with HST in earlier years, but are still classified into the “overlapping” group.}\]
up for long routes over time. This is consistent with the industry report and reflects that there may be more negative spillovers from HST to the airline industry for short routes.

Turning to routes that does not overlap with HST, the comparison between sub-figures (c) and (d) reveals that the number of airlines are not only larger on average for routes that connect to HST lines than those which does not, but also experience larger growth in absolute terms. This is suggestive that HST may feed traffic to airlines, thus attracting more airline carriers to enter.

Clearly, these are just model-free evidences and has not accounted for other factors that could possibly affect airline carriers’ route decisions. For example, it is possible that the short routes that overlap with HST have been experiencing some slow-down in GDP growth and it could be the slow-down that contributes to the decrease in the number of airlines serving the routes. It is also possible that routes with more HST line connections are those that connect larger cities with existing airline hubs, and therefore if could be the hub size that drives the difference in number of airlines.

To alleviate these types of concerns and to further quantify the spillovers from HST to the airline industry, I conduct a regression analysis that controls for possible factors that affect airline carriers route decisions. Here, the unit of analysis is on the route-year level. The dependent variable is the number of airlines serving a route in a given year. I regress the dependent variable on factors related to HST, such as whether HST is present, the average number of HST lines connecting the end cities and the interaction between these two variables. I also control for the characteristics of the route, such as the average GDP (measured in logged millions of dollars) and length of the route. I categorize the routes into three groups in terms of route length: short, medium and long, using 600km and 1200km as cutoff points. Again, motivated by the industry report about the relative competitive advantage for HST and airlines for short and long routes, I interact the presence of HST with these distance dummies. I further include the total number of connecting airline routes to control for the impact of airlines’ networks on their presence in a focal route.36 Finally, depending on the specification, I include year and route fixed dummies to account for the unobserved effects that are constant within a route or a year.

Specifications 1 through 3 of Table 8 presents the results of the regression. Here I distinguish between fast trains and bullet trains, as they may have different impact on the airline industry. Compared with specification 1, specification 2 controls for year fixed effects, and specification 3 controls for both year and route fixed effects. Throughout the specifications, the basic pattern holds: overlapping with either fast trains or bullet trains is negatively correlated with the number of airlines in a focal route. This indicates negative spillovers from HST. However, the negative spillovers go away for medium and long routes, which is consistent with the industry report. Connecting to either fast train lines or bullet train lines is associated with a larger number of airlines in the focal route, which indicates positive spillovers from HST. However, the positive spillovers are negatively moderated if the focal route also overlaps with either type of the HST, as is shown in the negative

36To avoid concerns of reverse causality, I use total number of connecting airline routes in the previous year. This implies that the year 2007 is not included in the regression.
coefficients for the interactions of HST presence and connection to HST line for both fast trains and bullet trains. This is reasonable as it reflects that HST may no longer feed traffic to air travel if HST also provides service in the same route as airlines, because consumers may find it more convenient to stay with only one mode of transportation.

One potential concern regarding the 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 (specific to the city pair) 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 air carriers and the presence HST. For example, it is possible that the introduction of HST station in a given city is accompanies by some policy change (such as favorable policy for investment, tourism, etc) in that city, and it could be that policy change that affects the demand for travel to/from that city, which in turn influences airlines route decisions.

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 6.

---

Insert Figure 6 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 500 km (313 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 airlines) on the treatment (the presence of HST), market characteristics, city-pair 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 airlines 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.

Column 4 of Table 8 reports the regression results. The dependent variable is the number of airlines serving a route for each year. 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. Interestingly, for medium and long routes, there even seems to be a positive effect for the HST overlap. This suggests there may be market expansion effects associated with the introduction of HST. Also, connecting to HST lines on average has a positive impact on the number of airlines serving a route, which again provides from evidence of positive spillovers from HST.

To test whether the treated routes and control routes followed same trends before HST is introduced, I add two indicator variables, $HST_{t-1}$ and $HST_{t-2}$, which equal to 1 for one and two years prior to the introduction of HST in a given route, respectively. As is shown in column 4, the coefficients for both variables are not statistically significant, which implies that the common trend assumption cannot be rejected.

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 answer the research question regarding the impact of HST on airline carriers’ route network decisions. For example, how many more routes would be served if HST had not been introduced? And which are of the country has experienced the greatest impact of HST? The previous regression analysis takes the existing route network of airline carriers as given and it only measures the marginal impact of HST on airlines in the focal route. However, because routes are interconnected, an airline’s decision in one route may influence its decisions in other routes. In this section, I present a structural model that accounts for such interconnectedness of the markets.

Since the purpose of this paper is to evaluate the implications of spillovers from HST on airlines’ configuration of networks, I only model airlines’ decisions to enter a route or not. The intensity of

---

37 Here the t statistics are calculated using Huber-White robust standard errors that allow for arbitrary correlation of residuals across routes and years.
entry (i.e. the number of flights provided in a route) is less relevant and 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. As is shown in the appendix, there is enough variation of entries and exits in the data to help identify parameters of key interest.

6.1 Setup

There are \( N \) airlines in the industry that provide flights between different cities. Denote the total number of cities in the industry by \( C \). A route \( r \in \{1, \ldots, R \} \) 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.\(^{38} \) The network of an airline \( i \) at \( t \) is therefore defined by the collection of \( x_{irt} \) for all routes, i.e. \( x_{it} \equiv \{x_{irt} : r = 1, \ldots, R\} \). Further denote \( x_{t} = \{x_{it} : i = 1, \ldots, N\} \) as the network for the entire industry at time \( t \).

Time is continuous, indexed by \( t \in [0, \infty) \). 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 \( u_{irt}(x_{t}, z_{rt}) \).

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 \( j \in \mathcal{A} = \{0, \ldots, J - 1\} \) 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 \( z_{t} \equiv \{z_{rt} : r = 1, \ldots, R\} \), which together composes the industry state, and chooses actions for all routes to maximize its overall discounted expected payoff given by

\[
E \left[ \sum_{r=1}^{R} \left( \int_{0}^{\infty} e^{-\rho t} f_{irt}(x_{t}, z_{rt}) + \sum_{n_r=1}^{\infty} e^{-\rho T_{nr}} \phi_{irt}(j) \right) \right],
\]

where \( T_{nr} \) is the random time of the \( n \)th 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

\(^{38} \)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.
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.\(^{39}\)

**Assumption 2: Sufficient Statistics** Let $x_{rt} \equiv \{x_{irt} : i = 1, ..., N\}$ denote the states for all airlines in route $r$. Let $w_{rc}t \equiv \{w_{irt}c : i = 1, ..., N\}$ denote the vector that summarizes the state variables for each airline in all routes other than $r$. Each airline's entry decision for each route relies only on $(x_{rt}, w_{rc}t, z_{rt})$.

Assumption 2 implies that, what happens in other routes affect an airline's decision in the focal route through $w_{rc}t$.\(^{40,41}\) This helps to reduce 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, $\lambda$, 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.

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

$$E \left[ \left( \int_0^\infty e^{-\rho u_{irt}(x_{rt}, w_{rc}t, z_{rt})} \, dt + \sum_{n_r=1} e^{-\rho T_r n_r} \phi_{irt}(j) \right) \right].$$  \hspace{2cm} (2) 

\(^{39}\)Similar assumptions have 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).

\(^{40}\)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).

\(^{41}\)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.
6.4 Value Function and Equilibrium

For notational simplicity, I summarize $(x_{rt}, w_{rc}, z_{rt})$ 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 $\chi$ 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.

The instantaneous payoff $\phi_{irk}(j)$ consists of two components and is given by $\psi_{ijk} + \epsilon_{ij}$. $\psi_{ijk}$ is the mean payoff (or cost) associated with action $j$ in state $k$ for airline $i$, with $|\psi_{ijk}| < \infty$, and $\epsilon_{ij}$ 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.\(^{42}\)

Let $\sigma_{ijkr}$ 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 $w_{ivc}$, 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 $w_{ivc}$ can be interpreted as the result of the action of an “aggregate player”, which I denote as $i^c \in \{N + 1, ..., 2N\}$.\(^{43}\) As $w_{ivc}$ is a function of airline $i$’s decisions on the other routes, this cumulative process still follows a Poisson arrival process. Denote the rate of that process as $\lambda_c$. We can similarly define $\sigma_{ivjkr}$ 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 $\zeta_{ivr}$ 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 $(2N) \times J \times K$ probabilities $\zeta_{mjkr}$ for each $m \neq i$, state $k$ and choice $j$. Finally, let $V_{irk}(\zeta_{ivr})$ 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 $\zeta_{ivr}$. For small increments $h$, under the Poisson assumption, the probability of an event with rate $\lambda$ occurring is $\lambda h$. Given the discount rate $\rho$, 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 $\zeta_{ivr}$ for brevity) is given by

\(^{42}\)I suppress the dependency of $\epsilon$ on $r$ and $t$ for simplicity.

\(^{43}\)I use $N + 1$ through $2N$ to distinguish the identities of aggregate players and those of airlines in a given route.
\[ V_{irk} = \frac{1}{1 + \rho h} \left[ u_{irk} + \sum_{l\neq k} q_{kl} V_{irl} + \sum_{m=1, m\neq i}^{N} \lambda h \sum_{j=0}^{J-1} \epsilon_{mj,k} V_{ir,l(m,j,k)} + \sum_{i' = N+1}^{2N} \lambda \rho h \sum_{j=0}^{J-1} \epsilon_{i'j,k} V_{ir,l(i',j,k)} \right] \]

\[
\lambda h E_{max} \left[ \psi_{ijk} + \epsilon_{irk} + V_{ir,l(i,j,k)} \right] + \left( 1 - N(\lambda + \lambda_c)h - \sum_{l\neq k} q_{kl} \right) V_{irk}.
\]

There are six terms in the square bracket. The first one is the flow profits airline \( i \) gets from being in state \( k \), and the rest of the terms enumerate all possible state changes within the increment time \( h \) and incorporates the corresponding expected discounted payoff associated with the changed state. Specifically, the second term corresponds to the value function of state change due to the nature move, the third and forth terms correspond to the value function associated with a state change due to the action of competitors and the aggregate players, respectively. The fifth term corresponds to the instantaneous payoffs associated with airline \( i \)'s route decision and the value function associated with the state change due to the action. Finally, the sixth term corresponds to the value function such that the route remains in state \( k \) after the increment time \( h \).

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

\[
V_{irk} = \frac{u_{irk} + \sum_{l\neq k} q_{kl} V_{irl} + \sum_{m=1, m\neq i}^{N} \lambda h \sum_{j=0}^{J-1} \epsilon_{mj,k} V_{ir,l(m,j,k)} + \sum_{i' = N+1}^{2N} \lambda \rho h \sum_{j=0}^{J-1} \epsilon_{i'j,k} V_{ir,l(i',j,k)} + E_{max} \left[ \psi_{ijk} + \epsilon_{irk} + V_{ir,l(i,j,k)} \right]}{\rho + \sum_{l\neq k} q_{kl} + N(\lambda + \lambda_c)}. \tag{3}
\]

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 \( \mathcal{A} \) to each state \( (k, \epsilon_{irk}) \in \chi \times \mathbb{R}^J \). Given beliefs \( \{\zeta^{ir} : i = 1, \ldots, N; r = 1, \ldots, R\} \), and a collection of model primitives, a Markov strategy for firm \( i \) in route \( r \) is a best response if

\[
\delta_{ir}(k, \epsilon_{irk}; \zeta^{ir}) = j \iff \psi_{ijk} + \epsilon_{irk} + V_{ir,l(i,j,k)} \geq \psi_{ij'k} + \epsilon_{ij'r} + V_{ir,l(i,j',k)} \quad \forall j' \in \mathcal{A}. \tag{4}
\]

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

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

A Markov perfect equilibrium is thus defined as a collection of stationary policy rules \( \{\delta_{ir} : i = 1, \ldots, N; r = 1, \ldots, R\} \) such that (4) holds for all \( i, r, k \) and \( \epsilon_{irk} \) given beliefs \( \zeta^{ir} = (\sigma_{ir} : i = 1, \ldots, N; r = 1, \ldots, R) \) generated by (5).\(^{44}\)

\(^{44}\)The standard arguments apply for existence of equilibrium. (See Doraszelski and Satterthwaite 2010, Doraszelski and Judd 2012).
7 Empirical Specification

7.1 State Variables and Action Variables

I focus on the top four airlines and the top 68 cities in China, so \( N = 4 \) and \( C = 68 \). In total, there are \( R = 68 \times 67/2 = 2278 \) possible routes. In a given route \( r \), the state variables for airline \( i \) are \( x_{rk}, w_{rk}, \) 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.\(^{45}\)

Let \( c_{irk} \) denote the number of connecting routes for airline \( i \) in route \( r \) in state \( k \).\(^{46}\) Then \( w_{rk} = \{c_{irk} : i = 1,...,4\} \).

The state variable \( z_{rk} \) contains the exogenous characteristics of route \( r \), namely, 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), the length of the route, \( d_{rk} \) and the average GDP for endpoint cities \( gdp_{rk} \). I further allow routes to have difference GDP growth rates. Specifically, depending on the GDP growth rate of the city pair, I divide the routes into three categories and assign one “growth type” to routes from each category. Finally, I assume that each route is characterized by a time-invariant unobserved type \( s \), which is observed by airlines but not by the econometrician.

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.\(^{47}\) 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 city pair average GDP, 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

\[
\begin{align*}
    u_{irk} &= \beta_0 + \beta_1 \sum_{m \neq i} x_{mrk} + \beta_2 c_{irk} + \beta_3 \sum_{m \neq i} c_{mrk} + \beta_4 Fast_{rk} + \beta_5 Bullet_{rk} \\
    &\quad + \beta_6 c_{Fast_{rk}} + \beta_7 c_{Bullet_{rk}} + \beta_8 Fast_{rk} \times c_{Fast_{rk}} + \beta_9 Bullet_{rk} \times c_{Bullet_{rk}} + \beta_{10} gdp_{rk} + \beta_{11} d_{rk} \\
    &\quad + \beta_{12} Fast_{rk} \times d_{rk} + \beta_{13} Bullet_{rk} \times d_{rk} + \beta_{14} s_r.
\end{align*}
\]

(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 routes.

\(^{45}\)A connecting route is a route which shares one of the endpoint cities with the focal route.

\(^{46}\)For example, if route \( r \) connects city \( A \) and city \( B \), and in state \( k \), airline \( i \) offers direct flights in \( X \) routes that connect city \( A \) and \( Y \) routes that connect city \( B \), then \( c_{ir} \) is \( X + Y \).

\(^{47}\)This is consistent with literature on firm entry, e.g., Bresnahan and Reiss (1990), Seim (2006).
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.

Airlines pay a sunk cost to enter a route. The sunk cost depends on the unobserved market type, $s$, on whether the route is a government-regulated air route, and on airline specific characteristics such as the number of connecting routes an airline provides for the focal route and whether the airline is exempt from the regulation. The scrape value associated with exiting a route, is assumed to be zero.\textsuperscript{48} Therefore, the choice-specific instantaneous payoffs $\psi_{ijk}$ is given by

$$\psi_{ijk} = \begin{cases} 
\eta_0 + \eta_1 \times s + \eta_2 \times \text{reg}_i + \eta_3 \times C_{irk} + \eta_4 \times HQ_i & \text{if } j = 1 \\
0 & \text{otherwise,}
\end{cases}$$

(7)

where $\text{reg}_i$ 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 end-cities, and zero otherwise. The structural parameters to be estimated are $\theta = (\beta_0, ..., \beta_{14}, \eta_0, ..., \eta_4)$.

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 GDP into five quantile-based bins.

\textsuperscript{48}Aguirregabiria and Ho 2012 have a similar assumption.
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 GDP and presence of HST. With the estimation results from the first stage, I estimate the structural parameters in the second stage.

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. To account for this challenge, 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 of the competitors’ states. Symmetry implies that we can use a common policy function to model each player’s behavior.

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 \( \alpha \) is the parameter to be estimated. I assume that \( \tilde{\sigma}_{ijr}(k, s, \alpha) \) takes the following form

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

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 GDP, dummies for length of route, and the average GDP growth rate of the city-pair as controls. Further, I allow the

---

49In the sample, about one third of the airlines re-enter the routes once they exited.

50For 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 \((1,1,3,2,1)\).
first-stage policy function to depend on the unobserved type of the route.\textsuperscript{51} Entry costs are allowed to depend on the airlines’ number of connecting routes, the market unobserved type as well as on whether a route is under regulation and on whether a airline is exempt from regulation.

Define $\tilde{\sigma}_{ijr}(k,s,\alpha^c)$ as the reduced form probability that the “aggregate player” makes choice $j$ in state $(k,s)$ in $r$. Airlines know on average, depending on their current number of connecting routes in the focal route, the probability of the aggregate player’s choices. The probabilities are estimated non-parametrically.

The exogenous state variables include GDP, 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 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 GDP, the average GDP 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 GDP and the average GDP 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

$$h = (q_{12}, q_{13}, ..., q_{K-1,K}, \lambda \sigma_{111}, ..., \lambda \sigma_{1,t-1K}, ..., \lambda \sigma_{t-1K}, ..., \lambda \sigma_{N+111}, ..., \lambda \sigma_{N+1,t-1K}, ..., \lambda \sigma_{N+1K}, ..., \lambda \sigma_{2N+111}, ..., \lambda \sigma_{2N+1,t-1K}, ..., \lambda \sigma_{2N+1K}, ..., \lambda \sigma_{2NK})$$

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 $\tau$ units of time can be decomposed into two parts: the first part is the probability that state changes after $\tau$ units of time. Because the aggregate process follows a Poisson process, the interval between state changes follows an exponential

\textsuperscript{51}In the first stage, I jointly estimate the policy function and the probability of a route being of a specific unobserved type.
distribution. Therefore the probability that state changes after $\tau$ units of time is given by
\[
\left(\sum_{l \neq k} q_{kl} + \sum_i \lambda \sum_{j \neq 0} \sigma_{ijk} + \sum_{i'} \lambda_c \sum_{j' \neq 0} \sigma_{i'j'k}\right) \exp \left[-\tau \left(\sum_{l \neq k} q_{kl} + \sum_i \lambda \sum_{j \neq 0} \sigma_{ijk} + \sum_{i'} \lambda_c \sum_{j' \neq 0} \sigma_{i'j'k}\right)\right].
\]

(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
\[
\frac{\lambda \sigma_{ijk}}{\sum_{l \neq k} q_{kl} + \sum_i \lambda \sum_{j \neq 0} \sigma_{ijk} + \sum_{i'} \lambda_c \sum_{j' \neq 0} \sigma_{i'j'k}}.
\]

(10)

Taken together, the probability becomes:
\[
\lambda \sigma_{ijk} \exp \left[-\tau \left(\sum_{l \neq k} q_{kl} + \sum_i \lambda \sum_{j \neq 0} \sigma_{ijk} + \sum_{i'} \lambda_c \sum_{j' \neq 0} \sigma_{i'j'k}\right)\right].
\]

(11)

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

Denote $I_{rw}(i, j)$ as the indicator variable which equals to one if player $i$ takes action $j$ in the $w$th 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:
\[
\hat{L}_{rw}(h(\alpha); s) = \\
\left(\sum_{j \neq 0} I_{rw}(0, j) q_j(k_{rw}, s, \alpha) + \sum_i \lambda \sum_{j \neq 0} I_{rw}(i, j) \sigma_{ijk}(k_{rw}, s, \alpha) + \sum_{i'} \lambda_c \sum_{j \neq 0} I_{rw}(i', j) \sigma_{i'jk}(k_{rw}, s, \alpha)\right) \\
\times \exp \left[-\left(\sum_{j \neq 0} q_j(k_{rw}, s, \alpha) + \sum_i \lambda \sum_{j \neq 0} \sigma_{ijk}(k_{rw}, s, \alpha) + \sum_{i'} \lambda_c \sum_{j \neq 0} \sigma_{i'jk}(k_{rw}, s, \alpha)\right) \tau_{rw}\right].
\]

(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 GDP growth rate of the city-pair, length of the route as well as on an indicator variable that captures whether the route is regulated.

\[^{52}This includes the “event” at time $T$, where it is possible that nothing happens.\]
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

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

8.2 Stage 2

In the second stage, I estimate the structural parameters, $\theta$, 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 $\theta$ using forward simulation in the same spirit of BBL and HMSS.\(^\text{53}\)

Given the value functions $\tilde{V}_{ir(k)}(\theta; s)$, together with the assumption that the idiosyncratic error terms in equation (4) follow an i.i.d. type I extreme value distributions, the choice probabilities in equation (5) can be expressed as:

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

Replacing $\tilde{\sigma}_{ijrk}$ in $\tilde{L}_{rw}(h(\alpha); s)$ with $\tilde{\sigma}_{ijrk}$, the new likelihood function can be denoted as $\tilde{L}_{rw}(\theta; s)$. 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

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

8.3 Identification

Now I discuss the data pattern that helps the identification of the structural parameters. First of all, time-series and cross-sectional variation in the state variables and the corresponding airlines’ route decisions help identify the parameters associated with the profit function. For example, if everything else is constant, yet we observe that the number of airlines serving a given route goes down when HST is introduced, then we can infer that HST has a negative impact on airlines’ profits in a given route.

\(^{53}\)Details of the forward simulation procedure can be found in the appendix. (Work in progress)
Entry costs are identified by the time-series variation in airlines’ route decisions. Intuitively, suppose there are two routes that are identical in all market conditions except that the focal airline is present in one route but not in the other at time $t$. Without entry costs, the probability that the airline offers flight services in both routes after time interval $\delta t$ should be the same and be independent of whether the airline is present or not. If instead, the probability of serving the route is higher (lower) for the route where the airline is (not) present, then we can infer that there are costs of entering a route.

Finally, the unobserved types of routes are identified by comparing the entry probabilities of routes with similar characteristics. The intuition is that if two routes have the same characteristics, yet there are more air carriers in one route than the other, we can infer that the difference is due to the unobserved market characteristics. I use finite mixtures to model the unobservables and allow the probability of a route being a specific type to depend on the initial market characteristics of the route.

9 Results

9.1 Parameter Estimates from The Structural Model

Table 9 shows the results of the estimation of the structural parameters. Here the standard errors are calculated using a subsampling algorithm based on 50 subsamples using half of the observations. I begin with the strategic effects, which is related to the route network of the airline 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 not only the airline itself, but also its competitors. This is interesting because it shows that there might be some positive spillover among the route networks of competing airlines. In fact, the negative effect brought by one competitor can be almost compensated if the number of connecting routes for the airline takes a level of 1, which corresponds to 6 to 15 connecting routes, or if the number of connecting routes for its competitor takes a level of 2, which corresponds to 16 to 25 connecting routes.

Insert Table 9 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 compared to the fast train, and because 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 HST goes away when the length of the route increases. However, there exists some
heterogeneity between fast trains and bullet trains regarding how the negative spillovers change with the length of the route. While for fast trains, the negative impact almost goes away when the length of the route changes from short to medium, it only goes down by 67% for bullet trains for the same change. This is probably because bullet trains have some speed advantage than fast trains. However, for both fast trains and bullet trains, the spillovers turn positive for long routes. A possible explanation for this positive net effect 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,\(^\text{54}\) for many HST routes in China, about half of the traffic is “generated” such that passengers would not travel 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. 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 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 5 routes with a medium level of connectivity to fast train lines and no overlap with fast trains. The corresponding number of connections for bullet train is 6, suggesting a smaller positive spillover from bullet trains than from fast trains. However, for both types of HST, the positive effect seems to be moderated by the presence of HST in the same route. Specifically, when HST is present in a route, the impact of the connectivity 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.

Turning to the market characteristics, airlines have lower flow payoffs for longer routes, but enjoy higher payoffs in routes with large average city-pair GDP and with higher values of the route unobserved type.

Finally, there are significant costs associated with entering a route. On average, the entry costs are about 2.6 times the annual flow payoff of an airline.\(^\text{55}\) However, there is heterogeneity in the entry costs among routes. Specifically, the cost of entry goes down for routes with a higher value of the unobserved type, but it goes up for routes that are regulated by the government. Interestingly, having more connecting routes or having headquarter in one of the city-pair do not seem to reduce the entry cost for a route. This could possibly be because airlines usually provide many routes


\(^\text{55}\)This number is higher than what is estimated in Aguirregabiria and Ho (2012), which concludes that the average entry costs for air carriers are about 1.75 times their average annual profits. This is probably due to the differences in the operating efficiency between airlines in U.S. and in China. For example, in 2014 and 2015, the average operating margins for U.S. airlines are 8.4% and 15.2% respectively, while those for Chinese airlines are 4.4% and 6.4% respectively.
connecting their headquarters, having headquarter in a given route is highly correlated with the number of connecting routes an airline has for the same route. Such high correlation makes both coefficient non-significant.

10 Policy Experiments

In this section, I carry out counterfactual analysis to answer the research questions. Using the recovered structural parameters, I carry out two policy experiments: first, I remove the HST from the market and assess how the airline carriers’ route networks would have changed accordingly. In the second policy experiment, I increase the positive spillovers from HST and study whether and how much improvement of the spillover is needed to compensate the air carriers for the negative impact of HST.

Implementing the Policy Experiments

Carrying out these policy experiments requires solving for the market equilibrium under the corresponding counterfactual scenarios. This is no easy task because, although airlines are assumed to make route decisions in a decentralized manner, the equilibrium cannot be solved independently for each route as markets are interconnected. For example, when HST is removed, it affects an airline’s decisions not only for the focal route, but also for the rest of the network, and this may in turn affect the airline’s expectation with respect to the evolution of \( w_{\text{int}} \), or the inclusive value that summarizes the information for the rest of the network when an airline makes decision for the focal route.

In the similar spirit of Aguirregabiria and Ho (2012), I use a forward simulation method to incorporate the update in the expectation regarding the evolution of the inclusive value. The specific steps for solving for the equilibrium are described as follows:\textsuperscript{56}

1. Given belief for the inclusive values \( \sigma_{ic} \), solve for the equilibrium route by route
2. Use the recovered policy functions to simulate the evolution of the entire route network
3. With the simulation, update firms’ beliefs about the average evolution of the inclusive term, denoted by \( \sigma'_{ic} \)
4. Repeat steps 1-3 until \( |\sigma'_{ic} - \sigma_{ic}| < \epsilon \). Here I set \( \epsilon \) to be \( 1e^{-6} \)

Once the equilibrium in the counterfactual scenario is solved, I use airlines’ corresponding policy functions to simulate 1,000 times the evolution of airlines’ route networks over the period that is covered by the data (i.e. 2007-2015) and average across over them to compare the difference in the airline route networks between the counterfactual scenario and the baseline scenario, where HST is introduced as observed.\textsuperscript{57} To ensure a fair comparison, I also solve for the equilibrium in the

\textsuperscript{56}There is no guarantee that the equilibrium is unique. However, as is discussed in Arcidiacono et al. (2016), the continuous time frameworks helps to eliminate a likely source of multiplicity in the equilibrium. Besides, I tried solving for the equilibrium using different starting points for the value function as well as airlines expectation for the evolution of the inclusive value, and found that they always coverge to the same equilibrium.

\textsuperscript{57}For each simulation, I keep the initial network configurations the same as observed.
baseline scenario, and the corresponding route networks are simulated and averaged in the same manner.

**Goodness of Fit**

Before moving on to the results of the policy experiments, I first report the goodness of fit of the estimated model. Specifically, I compare the observed route network configurations with those predicted by the equilibrium in the baseline scenario at the end of year 2015. Column 1 of Table 10 shows the proportion of entries correctly predicted by the model by airline. We can see that, across airlines, the model correctly predicts about 62% of entries in the data, and the proportion varies from 56.5% (Hainan Airlines) to 68.5% (China Southern Airlines). The major reason that the fit rate is not as high is because in the structural model I imposes symmetry for all airlines and do not allow airline dummies for computational reasons. This means that I am only modeling average airlines; behavior, not a specific one. This is also reflected in the high hit rate when we look at the result on the industry as a whole (93.9%). Column 2 of the table reports the correlation in the airline presence across all routes by airline. We can see that using this metric, the model does a reasonably good job in terms of predicting airlines’ presence: across airlines, the correlation ranges from 0.63 (China Eastern Airlines) to 0.76 (China Southern Airlines). Again, if we look at the industry level the correlation is much higher. Together, this implies that the model does a reasonably good job in predicting airlines’ decisions on the industry level.

10.1 Policy Experiment 1: Removing the HST

To understand how the introduction of HST has changed the network configurations of airline carriers, in this policy experiment I study a scenario in which HST has never been introduced. The coexistence of both negative and positive spillovers from HST makes the question especially interesting as it is not clear overall which effect would dominate, and since the markets are interconnected, the question cannot be answered using reduced form techniques. Further, as the spillovers are heterogeneous across routes with different characteristics and routes that interact differently with the HST network (i.e. overlap with versus connect to HST lines), it is also interesting to investigate whether and how airlines’ networks from different areas of the country are affected differently.

**Net Impact of HST on The Airline Industry**

I first look at the impact of HST on the route decisions of the airline industry. Specifically I look at two measures: one is the total number of routes that are served by the airline industry, and the other is airline route presence, defined as the sum of the total number of routes served by each airline's presence corresponds to a number between 0 and 1.

58 Note that, because the route networks predicted by the model are averaged over 1000 simulations, in each route, each airline's presence corresponds to a number between 0 and 1.

59 For example, if in the data, there is only one route where an airline enters, and from simulation the airlines average presence in that route is 0.8, then the proportion is 80%. The proportion on the industry level does not differentiate among the identity of the airline. Specifically, for the observed sample, entry is defined as the route being served by at least one airline. For model prediction, entry is defined as the sum of presence of all airlines, capped by one.

60 For the industry level, I just calculate the correlation in total number of airlines across routes.

61 In Jia (2008)'s paper the corresponding correlation ranges from 0.61 to 0.75.
each airline carrier. Table 11 presents the results. Column 1 corresponds to the airline network configuration under the scenario where HST is removed while column 2 corresponds to the baseline scenario where HST is present. Comparing the numbers from column 1 and column 2, we can see that the introduction of HST has reduced the total number of routes served by the airline industry by about 12% and has reduced the total airline route presence by about 15%. This indicates that, despite the existence of positive spillovers from HST, there are stronger negative spillovers such that the net impact of the introduction of HST on the airline industry is negative. This reflects the challenge the airline industry in China is facing: as the network of HST continues to grow, it is very likely that the negative spillovers would get even stronger. Therefore, it is crucial for the airline carriers to adopt some new strategy to deal with the challenge, such as offering more international routes or exploring the opportunity of inter-modal services.

To further quantify the impact of the positive spillovers from inter-modal connectivity, I further explore a scenario where HST is introduced but the positive spillovers from connectivity is shut down. Specifically, I set the coefficients on the number of fast train/bullet train line connections to zero, adjust their interaction terms with the presence of HST accordingly, and then re-solve for the equilibrium to simulate the corresponding airline network evolution. The results are presented in Column 3 of Table 11. We can see that, under this scenario, the total number of routes served by the airline industry goes further down by about 38% and the total airline route presence goes down by 43%. Together, this implies that, although the net impact of HST introduction on the airline industry is negative, the positive spillovers from inter-modal connectivity still plays an important role in prevent the airline network size from shrinking much further.

**Zooming in: The Impact of HST by City**

Perhaps it is more interesting to look at how the introduction of HST has affected the airline industry differently for different parts of the country. To see this, Figure 7 reports the airline route presence for the top 20 cities in China. On the map, the black lines represent the HST rails, and each dot represents a city. The size of the dot corresponds to the size of airline route presence in that city for the baseline scenario and the color of the dot shows how airline route presence changes when HST is introduced. Specifically, I use warm colors to denote an increase in the airline route presence and cold colors to denote a decrease. Here we can see a clear pattern: it is usually the cities in the geographically central area (such as Wuhan, Zhengzhou and Nanjing) that experience the largest decrease in airline route presence. In contrast, cities in the geographically peripheral areas are less affected. In fact, some cities, such as Chengdu and Chongqing, even experience some increase in the airline route presence as a route of HST introduction.

The intuition behind this pattern is that: as is shown in the figure, cities in the central area

---

62 The key difference between these two measures is that, while the first measure does not distinguish between routes that are served by only one airline and those served by multiple airlines, the second measure does. For example, suppose there are only two routes in the entire market, and three airlines serve on of the routes and no airlines serve the other, then the total number of routes served by the airline industry in this case is 1, while the airline route presence is 3.

63 Because the pattern looks quite similar for the two measures, in the following analysis I only report the airline route presence.
of China usually connect densely to the HST network. This implies that the overlap between HST routes and airline routes is large. Besides, their central geographical location implies that the average distance from the cities to the rest of the places is relatively short, which gives HST more competitive advantage in terms of travel convenience. Together, these factors make airline carriers less likely to serve routes that connect the cities. Vice versa, for cities from the geographically peripheral areas, because the average distance to other cities is long and there is less overlap with HST routes, the negative spillovers from HST are relatively limited.

Table 12 reports the impact of HST on airline route presence for all cities aggregated into different parts of China. We can see that, consistent with Figure 7, cities from the East and Central part of China are the ones that are most affected by the introduction of HST. On average, cities from these two areas experience a decrease of almost 20 airline routes when HST is introduced. This contrasts with cities from the least affected area, where there is almost no decrease in airline route presence per city. Also consistent with the previous intuition, the difference is largely due to the fact that the average route length for cities in the East and Central areas are relatively shorter and there is a larger proportion of routes that overlap with HST.

**Zooming in Further: The Impact of HST by Route**

To further understand to what extent the introduction of HST has affected different types of routes, I look at the average number of airlines for each route type and compare the numbers over two scenarios. Specifically, I divide the routes into six groups along two dimensions: whether HST is present in the route and the length of the route. I calculate the average number of airlines in each cell for the scenario in which HST was not introduced and compare them with the numbers in the baseline scenario. The results are presented in Table 13. There are several observations: first, consistent with the parameters from the structural estimation, HST introduction reduces the number of airlines, but only for short and medium length routes. Long routes, in contrast, has experienced an increase in the average number of airlines when HST is present. This again implies that there is a market expansion effect of HST introduction on the airline industry, and the net impact is positive for long routes. Interestingly, for all length type, airline routes that does not overlap with HST experience a decrease in the number of airlines when HST is introduced. Although the level of decrease is not as much as short and medium routes that overlap with HST, this is still surprising because these routes does not compete directly with HST. The reason behind this is that routes are interconnected, so that when an airline chooses not to serve a route, it also makes serving other connecting routes less attractive.

**The Impact of HST on Airline Concentration**

Finally, I investigate how the introduction of HST has affected the concentration of airlines in a city. I define airline concentration in a given city as the Herfindahl Index for market share in terms of number of routes connecting the city.\(^{64}\) Table 14 reports the summary statistics of airline concentration for the 68 cities under the two scenarios. We can see that all cities experience an increase in the average number of airlines when HST is present. This again implies that there is a market expansion effect of HST introduction on the airline industry, and the net impact is positive for long routes. Interestingly, for all length type, airline routes that does not overlap with HST experience a decrease in the number of airlines when HST is introduced. Although the level of decrease is not as much as short and medium routes that overlap with HST, this is still surprising because these routes does not compete directly with HST. The reason behind this is that routes are interconnected, so that when an airline chooses not to serve a route, it also makes serving other connecting routes less attractive.

---

\(^{64}\)For example, if a city is served by two airlines, A, and B, with A having two routes connecting the city and B having only one, then the market share for A is 2/3 and that for B is 1/3. The corresponding airline concentration is therefore 
\[(2/3)^2 + (1/3)^2 = 5/9.\]
increase in the airline concentration when HST is introduced. It is reasonable that, when HST brings negative spillovers to airline carriers in a given city, usually the airlines with smaller route market share have a higher probability of exit, because the fewer connecting routes in a given city makes it relatively less attractive for them to keep proving the airline services. This in turn increases the airline concentration rate in a city. What is interesting is that, even for cities that experience net positive spillovers from HST, the airline concentration rate also increases. One possible reason is again the fact that routes are interconnected. As an airline already reduces routes that connects other cities, although there are positive spillovers from a given city, the negative impact of reducing connecting routes in other cities outweigh the positive spillovers, which in turn reduces the number of routes the airlines provides that connect the city. In other words, when HST is introduced, it is usually airlines with larger route market shares in a city that survives, which in turn increases their market power, and this market power can further carry over to other cities.

10.2 Policy Experiment 2: Improving the Positive Spillovers from HST

In the second policy experiment, I examine how much improvement in the positive spillovers from HST is needed to compensate for the overall negative impact of HST on the airline carriers. This is relevant because the local government has specifically emphasized the need to improve the services between plane and HST connections in its 13th five-year-plan.\textsuperscript{65} Answering the research question helps the government to better decide the level of inter-modal services to be provided.

To carry out the analysis, I improve the positive spillovers from HST to the airline industry by different levels and study their respective impact on airlines route decisions. Specifically, I increase the profits from connecting to fast trains and bullet trains by 5%, 10%,...50%, respectively and solve for the corresponding equilibrium.\textsuperscript{66} Similar to the first policy experiment, I run 1000 simulation for the evolution of the airline route network for each scenario and compare the resulting airline route networks at the end of 2015 with those from the baseline scenario.

**Overall Impact of the Improvement of Positive Spillovers**

Figure 8 reports the airline route presence in scenarios with different levels of improvement in the positive spillovers from HST to the airline industry. I also use the horizontal line to represent the airline route presence in the scenario where HST were not introduced. As we can see from the graph, when the profits from connection to HST lines increase by about 40%, the airline route presence would increase to a level which is similar to that in the scenario without HST roll-out. This implies that the HST need to feed at least 40% more customers to the airline industry to compensate for its negative overall impact.

**Zooming in: Impact of the Improvement of Positive Spillovers by City**

\textsuperscript{65}Examples of inter-modal services include increased frequency for shuttle buses between airports and HST stations, through ticketing and luggage services, as well as improved coordination in flight and HST schedules.

\textsuperscript{66}This policy experiment assumes the inter-modal service is provided for the entire airline industry. Apparently, many services such as coordination in schedule as well as through luggage handling require investment and time and may not be ready for the entire industry all at once. However, other services such as through ticketing can be achieved for the entire airline industry relatively quickly and easily.
The impact of the improvement in positive spillovers is not uniform across the country. Figure 9 presents the change in airline route presence for the top 20 cities compared to the baseline scenario at the end of year 2015. Each dot on the map represents a city and the size of the dot indicates the level of airline route presence for the city. Here, I use a warmer color to represent a larger increase in the airline route presence in a city. There are several interesting observations: first, cities located in the geographically central area of China experience a relatively large increase in airline route presence because of the inter-modal services. This contrasts sharply with the implications from Figure 7, which shows that these cities are the ones that lose most airline routes when HST is introduced. Together, the result suggests that the negative impact of HST on airline route presence for cities located in the central area of the country may be buffered to some extent with the provision of inter-modal services.

Interestingly, inter-modal services have also increased the airline routes presence in cities located in the North and Northeast part of China (such as Tianjin, Harbin and Dalian). This is largely because cities from these areas have moderate connection to HST lines. Also, because the average distance between a city in these areas and the rest of the cities are relatively large, the negative spillovers from overlapping with HST in a route is to some extent mitigated. As a result, inter-modal services might have a larger impact on airline carriers’ entry decisions in routes connecting the cities.

Perhaps surprisingly, although Beijing, Shanghai and Guangzhou are among the cities with the highest number of HST line connections, they do not seem to experience as much increase in airline route presence. One possible reason is that, for the two cities, there is a large proportion of airline routes that overlap with HST. In fact, by 2015, almost 50% of the airline routes connecting the three cities are served by HST. Together with the fact that the entry costs for these routes are higher because of the government regulation, this explains why the improvement in the positive spillovers has relatively little effect on the airline route presence of the two cities.

Table 15 presents the impact of the improvement in the positive spillovers from HST by area. The increase in the airline presence due to the improvement seems to be positively correlated with the number of HST lines that connect a city, but the relationship is not monotonic. For example, cities from the Northwest and Southwest area of China has the least number of HST line connections on average, and they also experience the least level of increase in airline route presence. On the other hand, however, cities from the geographically central area of China has the largest number of HST line connections on average, yet the increase in airline route presence is less than that for cities from the North and Northeast area of the country, where there are fewer HST connections but at the same time smaller proportion of airline routes that overlap with HST. Together, this implies that, when inter-modal services are provided, it is not always the case that cities with the
densest connection to HST lines that would experience the largest increase in the airline route presence, because larger number of HST line connections also imply a higher proportion of routes that overlap with HST. As a result, it is usually those cities with moderate connection to HST lines that experience the largest increase. Looking forward, we can also expect that, as the HST network continues to expand, the marginal benefit of providing inter-modal services may decline over time.

11 Conclusion

This paper empirically assesses and quantifies the negative and positive spillovers of the HST network on the airline industry in China, and studies the implications of such spillovers for firms’ entry decisions (i.e., network configuration choices). I hand-collected a unique dataset on flight schedules and timing of the introduction of the HST and use both reduced form and structural methods to answer the research questions. 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.

With the estimates from a dynamic oligopoly model of entry, I carry out a policy experiment in which HST had not been introduced and solve for the equilibrium in airline network configurations in this scenario. I find that the introduction of HST has heterogeneous impact on airlines’ route network configurations. Overall, the roll-out of HST has reduced the airline route presence by about 15%, and it is usually the cities that are located at the geographically central areas of the country that experience the largest decrease. This is because these cities are densely connected to HST lines and that the average distance of routes connecting the cities are short. Cities that are located at the peripheral areas of the country, on the other hand, experience less decrease or even slight increase in the airline route presence because the positive spillovers from HST are relatively stronger.

In the second policy experiment, I study the impact on air carriers’ route network decisions of a improvement in the positive spillovers from the HST. I find that, with 40% increase in the air carriers’ profits from HST line connections, the airline route presence is comparable to that in a scenario without HST roll-out. This implies that the HST needs to feed at least 40% more traffic to the air routes in order to compensate for its negative overall impact on the airline industry. The results also show that, it is usually the cities with moderate connections to HST lines that experience the largest increase in airline route presence, because for these cities, the number of air routes that overlap with HST is also smaller.
References


choice,” Working paper.


Vitorino, Maria Ana (2012), “Empirical entry games with complementarities: An application to the


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 served by the airline carriers and by both airline carriers and high speed trains (fast trains and bullet trains) in China from 2007 to 2015. A route is defined as a non-directional city pair. A route is served by airline carriers if there exists a direct flight that connects the corresponding city pair. A route is served by high speed trains if a passenger can take HST to travel between the city pair without changing trains.
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”)

![Travel Time as A Function of Distance by Transportation Modes](image)

Figure 5: Descriptive Evidence for Spillovers from HST

This figure presents the evolution of the average number of airlines serving different groups of routes. Sub-figures (a) and (b) correspond to short and long routes that overlap with HST (either fast train or bullet train) at the end of year 2015, respectively. Sub-figure (c) corresponds to routes that do not overlap with HST but connect to HST lines and sub-figure (d) corresponds to routes that neither overlap with or connect to HST lines.

![Descriptive Evidence for Spillovers from HST](image)
Figure 6: Selection of Control Routes in the Difference-in-difference Analysis

This figure shows the principles for selecting the control routes.

![Diagram of selection of control routes](image)

Figure 7: Impact of HST by City

This map shows how the introduction of HST has affected airline route presence in each city. The black lines on the map represent the HST rails for both fast trains and bullet trains. Each dot on the map represents a city. The size of the dot corresponds to the size of the airline route presence for a given city in the baseline scenario, where HST is introduced. And the color of the dot represents the impact of HST on the airline route presence. I use cold/warm colors to represent a decrease/increase in the number of airline route presence in a city. The numbers are calculated using the average of 1000 simulations for each scenario at the end of 2015.

![Map showing impact of HST by city](image)
Figure 8: Impact of HST by City

This graph presents the change in the airline route presence at a given level of improvement in the possible spillover from HST to the airline carrier. The horizontal line represents the level of airline route presence when HST were not introduced.
Figure 9: Impact of Inter-modal Services by City

This map shows the increase in airline route presence for each city for an improvement of complementarity between HST and airplanes. The black lines on the map represent the HST rails for both fast trains and bullet trains. Each dot on the map represents a city. The size of the dot corresponds to the size of the airline route presence for a given city in the baseline scenario, where HST is introduced. And the color of the dot represents the level of increase in the airline route presence in a city. I use a warmer color to represent a relatively larger increase in the airline route presence. The numbers are calculated using the average of 1000 simulations for each scenario at the end of 2015.
Table 1: Summary Statistics

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

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

Table 2: Summary Statistics by Year

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of airlines</td>
<td>0.57</td>
<td>0.58</td>
<td>0.63</td>
<td>0.66</td>
<td>0.68</td>
<td>0.72</td>
<td>0.77</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>Number of flights</td>
<td>0.80</td>
<td>0.84</td>
<td>0.95</td>
<td>1.04</td>
<td>1.09</td>
<td>1.18</td>
<td>1.27</td>
<td>1.34</td>
<td>1.40</td>
</tr>
<tr>
<td>Number of entries</td>
<td>N.A.</td>
<td>0.05</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.06</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>Number of exits</td>
<td>N.A.</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>Number of flight increases</td>
<td>N.A.</td>
<td>0.05</td>
<td>0.07</td>
<td>0.08</td>
<td>0.06</td>
<td>0.09</td>
<td>0.09</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td>Number of flight decreases</td>
<td>N.A.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Fast train present</td>
<td>0.07</td>
<td>0.07</td>
<td>0.08</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
<td>0.11</td>
<td>0.19</td>
<td>0.21</td>
</tr>
<tr>
<td>Bullet train present</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.02</td>
<td>0.07</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>City pair average population (Million)</td>
<td>5.13</td>
<td>5.19</td>
<td>5.22</td>
<td>5.30</td>
<td>5.36</td>
<td>5.40</td>
<td>5.43</td>
<td>5.49</td>
<td>5.55</td>
</tr>
<tr>
<td>City pair average GDP (Billion)</td>
<td>30.00</td>
<td>35.52</td>
<td>40.14</td>
<td>47.50</td>
<td>56.52</td>
<td>63.19</td>
<td>69.59</td>
<td>75.72</td>
<td>81.84</td>
</tr>
<tr>
<td>Observations</td>
<td>2278</td>
<td>2278</td>
<td>2278</td>
<td>2278</td>
<td>2278</td>
<td>2278</td>
<td>2278</td>
<td>2278</td>
<td>2278</td>
</tr>
</tbody>
</table>
Table 3: GDP Quantiles and Airline Decisions

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

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

Table 4: Route Length Quantiles and Airline Decisions

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

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

Table 5: 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. Standard errors are in parentheses.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Airline operates in one end point airport in the previous year</td>
<td>0.0069</td>
</tr>
<tr>
<td>(single presence)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Airline operates in both end point airports in the previous year</td>
<td>0.0531</td>
</tr>
<tr>
<td>(dual presence)</td>
<td>(0.0009)</td>
</tr>
<tr>
<td>N</td>
<td>65,185</td>
</tr>
</tbody>
</table>

Table 6: GDP Quantiles and High Speed Train Presence

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

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average GDP of the City Pair</td>
<td>1.15</td>
<td>2.76</td>
<td>4.34</td>
<td>6.62</td>
<td>12.91</td>
</tr>
<tr>
<td>Fast train present</td>
<td>0.00</td>
<td>0.02</td>
<td>0.06</td>
<td>0.15</td>
<td>0.34</td>
</tr>
<tr>
<td>Bullet train present</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.13</td>
</tr>
<tr>
<td>Observations</td>
<td>16404</td>
<td>16400</td>
<td>16404</td>
<td>16400</td>
<td>16400</td>
</tr>
</tbody>
</table>
Table 7: Route Length Quantiles and High Speed Train Presence

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

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>City pair distance (00's km)</td>
<td>4.90</td>
<td>9.86</td>
<td>13.89</td>
<td>18.61</td>
<td>29.09</td>
</tr>
<tr>
<td>Fast train present</td>
<td>0.27</td>
<td>0.18</td>
<td>0.08</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Bullet train present</td>
<td>0.09</td>
<td>0.05</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Observations</td>
<td>16416</td>
<td>16416</td>
<td>16380</td>
<td>16416</td>
<td>16380</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Fast Train present (Y/N)</td>
<td>-0.661***</td>
<td>-0.658***</td>
<td>-0.195***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-16.02)</td>
<td>(-16.00)</td>
<td>(-4.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bullet Train present (Y/N)</td>
<td>-0.482***</td>
<td>-0.413***</td>
<td>-0.233***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-6.88)</td>
<td>(-5.89)</td>
<td>(-5.45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of Fast Train line connections</td>
<td>-0.195***</td>
<td>-0.195***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-4.92)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of Bullet Train line connections</td>
<td>0.126***</td>
<td>0.052***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.68)</td>
<td>(4.23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast Train × No. of Fast Train line connections</td>
<td>-0.008</td>
<td>-0.008</td>
<td>-0.163***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.23)</td>
<td>(-0.22)</td>
<td>(-6.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bullet Train × No. of Bullet Train line connections</td>
<td>-0.226***</td>
<td>-0.258***</td>
<td>-0.088**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-4.11)</td>
<td>(-4.69)</td>
<td>(-2.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast Train × Medium Distance</td>
<td>1.056***</td>
<td>1.056***</td>
<td>0.424***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(21.64)</td>
<td>(21.72)</td>
<td>(9.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast Train × Long Distance</td>
<td>1.330***</td>
<td>1.335***</td>
<td>0.530***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(23.90)</td>
<td>(24.07)</td>
<td>(10.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bullet Train × Medium Distance</td>
<td>0.492***</td>
<td>0.502***</td>
<td>0.219***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.35)</td>
<td>(6.51)</td>
<td>(4.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bullet Train × Long Distance</td>
<td>0.618***</td>
<td>0.641***</td>
<td>0.292***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.55)</td>
<td>(5.77)</td>
<td>(4.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST present (Yes/No)</td>
<td></td>
<td></td>
<td>-0.202***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-3.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of HST line connections</td>
<td></td>
<td></td>
<td>0.084**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(3.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST × Medium Distance</td>
<td>0.478***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(6.76)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST × Medium Distance</td>
<td>0.403***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4.79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST present_{t-2}</td>
<td></td>
<td></td>
<td>0.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HST present_{t-1}</td>
<td></td>
<td></td>
<td>0.065</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of airline routes connected_{t-1}</td>
<td>0.062***</td>
<td>0.062***</td>
<td>0.029***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(94.68)</td>
<td>(93.10)</td>
<td>(15.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average GDP</td>
<td>-0.171***</td>
<td>-0.166***</td>
<td>0.373***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-11.45)</td>
<td>(-11.03)</td>
<td>(6.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium Distance</td>
<td>-0.040</td>
<td>-0.038</td>
<td>0.614**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.75)</td>
<td>(-1.67)</td>
<td>(2.88)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long Distance</td>
<td>-0.487***</td>
<td>-0.478***</td>
<td>1.608***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-23.43)</td>
<td>(-23.09)</td>
<td>(6.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.100***</td>
<td>-0.082**</td>
<td>-2.061***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.74)</td>
<td>(-2.77)</td>
<td>(-9.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year Dummies</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City Pair Dummies</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year Dummies × Group Dummies</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>18224</td>
<td>18224</td>
<td>18224</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.544</td>
<td>0.547</td>
<td>0.898</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(t\) statistics in parentheses

* \(p < 0.05\), ** \(p < 0.01\), *** \(p < 0.001\)
Table 9: Estimation Results of The Structural Parameters

This table reports the estimation results for the structural model.

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>Std.Err</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strategic Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of own routes connected</td>
<td>0.42***</td>
<td>0.03</td>
</tr>
<tr>
<td>No. of competitors</td>
<td>-0.67***</td>
<td>0.13</td>
</tr>
<tr>
<td>No. of competitors' routes connected</td>
<td>0.11***</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Impact of HST</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bullet Train present (Y/N)</td>
<td>-0.62***</td>
<td>0.18</td>
</tr>
<tr>
<td>Bullet Train × Median distance</td>
<td>0.42**</td>
<td>0.20</td>
</tr>
<tr>
<td>Bullet Train × Long distance</td>
<td>0.77***</td>
<td>0.17</td>
</tr>
<tr>
<td>No. of Bullet Train line connections</td>
<td>0.09**</td>
<td>0.04</td>
</tr>
<tr>
<td>Bullet Train × No. of Bullet train line connections</td>
<td>-0.37***</td>
<td>0.07</td>
</tr>
<tr>
<td>Fast Train present (Y/N)</td>
<td>-1.03***</td>
<td>0.18</td>
</tr>
<tr>
<td>Fast Train × Median distance</td>
<td>0.97***</td>
<td>0.15</td>
</tr>
<tr>
<td>Fast Train × Long distance</td>
<td>1.56***</td>
<td>0.18</td>
</tr>
<tr>
<td>No. of Fast Train line connections</td>
<td>0.2***</td>
<td>0.04</td>
</tr>
<tr>
<td>Fast Train × No. of Fast Train line connections</td>
<td>-0.4***</td>
<td>0.07</td>
</tr>
</tbody>
</table>

| **Market Characteristics**    |       |         |
| GDP                           | 0.11*  | 0.06    |
| Median distance               | -0.43**| 0.19    |
| Long distance                 | -0.92***| 0.24    |
| Unobserved Type               | 1.46***| 0.19    |
| Constant                      | -1.12***| 0.27    |

| **Entry Costs**               |       |         |
| Entry Cost                    | 2.75***| 0.17    |
| Unobserved Type               | -0.8** | 0.33    |
| Regulated                     | 1.41***| 0.26    |
| No. of own routes connected   | -0.05  | 0.04    |
| Headquarter                   | -0.16  | 0.14    |

Standard errors calculated using 50 subsamples with half of observations.

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

Table 10: Goodness of Fit

This table reports the model goodness of fit. Column 1 presents the proportion of entries correctly predicted by airline, and column 2 presents the correlation in airline presence between model prediction and observed sample by airline.

<table>
<thead>
<tr>
<th></th>
<th>Proportion of Entries Correctly Predicted</th>
<th>Correlation between Model Prediction And Sample Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air China</td>
<td>63.1%</td>
<td>0.73</td>
</tr>
<tr>
<td>China Southern Airlines</td>
<td>68.5%</td>
<td>0.76</td>
</tr>
<tr>
<td>China Eastern Airlines</td>
<td>62.3%</td>
<td>0.63</td>
</tr>
<tr>
<td>Hainan Airlines</td>
<td>56.5%</td>
<td>0.67</td>
</tr>
<tr>
<td>Industry</td>
<td>93.9%</td>
<td>0.93</td>
</tr>
</tbody>
</table>
Table 11: Impact of HST on The Airline Industry

This table presents the total number of routes/airline route presence under different scenarios. Column 1 corresponds to the scenario where HST were not introduced, column 2 corresponds to the baseline scenario where HST is introduced as observed in the data and column 3 corresponds to the scenario where HST is introduced but there is no positive spillovers from HST. The numbers are calculated using the average of 1000 simulations for each scenario at the end of 2015.

<table>
<thead>
<tr>
<th></th>
<th>No HST</th>
<th>Baseline</th>
<th>No Complementarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Routes Served</td>
<td>1048</td>
<td>923</td>
<td>654</td>
</tr>
<tr>
<td>Total Airline Route Presence</td>
<td>2225</td>
<td>1893</td>
<td>1259</td>
</tr>
</tbody>
</table>

Table 12: Impact of HST by Area

This table presents the impact of HST on the airline industry for different parts of China. The numbers are calculated using the average of 1000 simulations for each scenario at the end of 2015.

<table>
<thead>
<tr>
<th>Area</th>
<th>Northeast</th>
<th>North</th>
<th>Northwest</th>
<th>East</th>
<th>Central</th>
<th>South</th>
<th>Southwest</th>
</tr>
</thead>
<tbody>
<tr>
<td># Cities</td>
<td>5</td>
<td>8</td>
<td>9</td>
<td>20</td>
<td>5</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td># HST Lines Per City</td>
<td>1.8</td>
<td>2.3</td>
<td>0.7</td>
<td>2.3</td>
<td>3.0</td>
<td>1.9</td>
<td>0.7</td>
</tr>
<tr>
<td>Average Route Length (00's km)</td>
<td>18.6</td>
<td>14.4</td>
<td>20.6</td>
<td>12.9</td>
<td>11.3</td>
<td>15.2</td>
<td>16.2</td>
</tr>
<tr>
<td>% Routes Overlap with HST</td>
<td>15.5%</td>
<td>19.4%</td>
<td>5.3%</td>
<td>34.0%</td>
<td>36.7%</td>
<td>18.3%</td>
<td>6.1%</td>
</tr>
<tr>
<td># Airline Route Presence (Baseline)</td>
<td>247</td>
<td>511</td>
<td>363</td>
<td>1065</td>
<td>371</td>
<td>744</td>
<td>485</td>
</tr>
<tr>
<td># Airline Route Presence (No HST)</td>
<td>278</td>
<td>616</td>
<td>386</td>
<td>1412</td>
<td>461</td>
<td>798</td>
<td>500</td>
</tr>
<tr>
<td>Difference</td>
<td>-31</td>
<td>-105</td>
<td>-23</td>
<td>-347</td>
<td>-90</td>
<td>-54</td>
<td>-15</td>
</tr>
<tr>
<td>Difference per City</td>
<td>-6</td>
<td>-13</td>
<td>-3</td>
<td>-17</td>
<td>-18</td>
<td>-5</td>
<td>-1</td>
</tr>
</tbody>
</table>

Table 13: Impact of HST by Routes

This table presents the impact of HST on different routes. Specifically, I divide the routes along two dimensions: the length of the routes and whether the route overlaps with HST. The numbers are calculated using the average of 1000 simulations for each scenario at the end of 2015.

| Route Length | HST | # Routes | Average # Airlines (Baseline) | Average # Airlines (No HST) | Difference | |
|--------------|-----|----------|-------------------------------|-------------------------------|------------|
| Short        | No  | 150      | 1.07                          | 1.35                          | -0.28      |
| Short        | Yes | 148      | 0.61                          | 1.82                          | -1.21      |
| Medium       | No  | 424      | 0.98                          | 1.09                          | -0.11      |
| Medium       | Yes | 203      | 2.06                          | 2.46                          | -0.40      |
| Long         | No  | 1233     | 0.44                          | 0.45                          | -0.01      |
| Long         | Yes | 120      | 2.15                          | 1.93                          | 0.22       |

Table 14: Impact of HST on Airline City Concentration

This table presents the impact of HST on the concentration of airlines' route services on the city level.

<table>
<thead>
<tr>
<th>HHI</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>HST (Baseline)</td>
<td>68</td>
<td>0.278</td>
<td>0.024</td>
<td>0.251</td>
<td>0.372</td>
</tr>
<tr>
<td>No HST</td>
<td>68</td>
<td>0.271</td>
<td>0.020</td>
<td>0.251</td>
<td>0.350</td>
</tr>
<tr>
<td>Difference</td>
<td>68</td>
<td>0.007</td>
<td>0.006</td>
<td>0.000</td>
<td>0.032</td>
</tr>
<tr>
<td>Difference (%)</td>
<td>68</td>
<td>2.51%</td>
<td>2.03%</td>
<td>0.09%</td>
<td>11.68%</td>
</tr>
</tbody>
</table>
Table 15: Impact of Inter-modal Services by Area

This table presents the impact of inter-modal services on the airline industry for different parts of China. The numbers are calculated using the average of 1000 simulations for each scenario at the end of 2020, where HST network is assumed to be expanded as scheduled.

<table>
<thead>
<tr>
<th>Area</th>
<th>Northeast</th>
<th>North</th>
<th>Northwest</th>
<th>East</th>
<th>Central</th>
<th>South</th>
<th>Southwest</th>
</tr>
</thead>
<tbody>
<tr>
<td># Cities</td>
<td>5</td>
<td>8</td>
<td>9</td>
<td>20</td>
<td>5</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td># HST Lines Per City</td>
<td>1.8</td>
<td>2.3</td>
<td>0.7</td>
<td>2.3</td>
<td>3.0</td>
<td>1.9</td>
<td>0.7</td>
</tr>
<tr>
<td>% Routes Overlap with HST</td>
<td>15.5%</td>
<td>19.8%</td>
<td>5.3%</td>
<td>34.0%</td>
<td>36.7%</td>
<td>18.3%</td>
<td>6.1%</td>
</tr>
<tr>
<td># Airline Route Presence (Baseline)</td>
<td>246.6</td>
<td>510.7</td>
<td>363.0</td>
<td>1064.8</td>
<td>370.9</td>
<td>744.3</td>
<td>484.3</td>
</tr>
<tr>
<td># Airline Route Presence (Intermodality)</td>
<td>310.4</td>
<td>608.9</td>
<td>419.6</td>
<td>1284.7</td>
<td>426.3</td>
<td>861.3</td>
<td>544.3</td>
</tr>
<tr>
<td>Difference</td>
<td>63.8</td>
<td>98.2</td>
<td>56.6</td>
<td>219.9</td>
<td>55.4</td>
<td>117.0</td>
<td>66.0</td>
</tr>
<tr>
<td>Difference per City</td>
<td>12.8</td>
<td>12.3</td>
<td>6.3</td>
<td>11.0</td>
<td>11.1</td>
<td>10.6</td>
<td>6.0</td>
</tr>
</tbody>
</table>
Online Appendix

- robustness check regression
  1. DV changes to number of flights
  2. (Not sure) Unit of analysis changed to airline route year combination to show indirect flights does not affect too much an airline’s decisions in the focal route

- Estimation procedure: forward simulation method to approximate the value function