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

August 2022

Marco Shaojun Qin† Maria Ana Vitorino‡ George John§

Abstract

This paper analyzes the impact of the high-speed train (HST) roll-out on air carriers' route network decisions in China. Using unique hand-collected data on the timing of the HST roll-out and airline networks, we estimate a dynamic oligopoly model of air-route entry and exit that accounts for the competition between airlines and the (positive and negative) spillovers from the HST. We use the model to assess the impact of the HST on airlines' decisions by simulating a scenario without HST. We find that, despite significant positive spillovers from the HST on airlines, derived from intermodal connectivity, the HST reduced airlines' route presence by about 14%, and airline profits by 23%. We find considerable heterogeneity across route-types and regions. Airlines readjusted their networks substituting towards longer routes and more peripheral regions in China thus improving connectivity among regions. We also use the model to explore the benefits of improving intermodal transportation.

Keywords: Entry, Dynamic Games, Continuous Time, Intermodal Substitution and Complementarity, Network Competition, High-speed Train, Airline Industry

---

*We are grateful to Frederico Belo, Pradeep Chintagunta, Katja Seim, Thomas Holmes, Amil Petrin, Joel Waldfogel, Linli Xu, and Yi Zhu for their helpful comments and suggestions. We also thank the participants of the Applied Micro Workshop and of the Marketing Brown Bag at the University of Minnesota. We thank the participants of the Marketing Dynamics Conference 2019 (University of Maryland), and seminar participants at Fudan University, Peking University, Tongji University, the University of Texas at Austin, Temple University, and the UCLA Anderson School of Management for their comments. We 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 our own.

†Temple University, Fox School of Business, Department of Marketing, Philadelphia, PA 19122, USA, shaojun.qin@temple.edu.

‡INSEAD, 77300 Fontainebleau, France, maria-ana.vitorino@insead.edu (corresponding author).

§University of Minnesota, Carlson School of Management, Department of Marketing, Minneapolis, MN 55455, USA, johnx001@umn.edu.

1 Introduction

Transportation has a key role in the spatial distribution of economic activity. The increasing availability of geographic data on transportation networks together with better retail data has led to a renewed interest in the efficiency of urban transportation markets (e.g., Fréchette, Lizzeri, and Salz 2019, Barwick, Kalouptsidi, and Zahur 2021). Changes in transportation infrastructure and its policies have effects on economic variables such as trade costs, congestion, and unemployment (e.g., Allen and Arkolakis 2014, Barwick et al. 2022, Fajgelbaum and Gaubert 2020) and can differ across transportation technologies (e.g., roads versus rail). Further, as argued in Redding (2021), investments in a transportation technology may have effects that vary across locations and could lead to changes in the entire (endogenous) transportation network.

This paper analyzes the impact of the high-speed train (HST) roll-out on air carriers' route network decisions in China. The airline industry is an industry characterized by intense competition. However, cooperation among competitors (i.e., co-opetition) is also prevalent in this industry: airlines often engage in cooperation either by entering explicit alliances or by feeding traffic to each other through a complex network of route connections.[1] The introduction of the HST, a viable substitute for air travel, increases competition but also generates further potential complementarities for airline companies by creating more nodes and connections in the transportation network and thus facilitating travel and expanding the market. This reorganization in the airline transportation network can have significant distributional effects given the changes in transportation coverage across the country.

As there is an increasing demand for more efficient, faster and cleaner transportation, the HST industry has been rising globally. HST networks across Europe and Asia continue to grow, with new lines underway or planned in countries such as France, Germany, Spain, India, Japan and, on a much bigger scale, in China, where the high-speed network is expected to reach 50,000 kilometers by 2025.[2]

In this paper we empirically quantify the positive and negative spillovers generated by the introduction of the HST on the airline industry and study how airlines readjusted their networks in response to the HST. Despite the intense discussion regarding the challenges and opportunities that the HST has brought to airline carriers, there is no empirical evidence that quantifies these effects. More generally, the topic of how firms reposition their products

[1] Co-opetition can be defined as simultaneous competition and collaboration between two or more organizations (see Laamanen (2016)). The concept of "co-opetition" was first coined by Ray Noorda, founder of Novell, and popularized in the strategic management field by Adam M. Brandenburger and Barry J. Nalebuff in their book Co-opetition (Brandenburg and Nalebuff 1996).

[2] Source: https://lite.cnn.com/en/article/h_1bf22f56108ae8a03cc59f022d5a8e21

to avoid competition while taking advantage of positive spillovers is understudied. Further, our analysis can help guide the decision to introduce the HST in other regions or countries.[3]

We hand-collect daily flight information for the four major passenger airlines and their subsidiaries in China (which account for almost 90% of the domestic air passenger traffic) and use this data together with detailed information on the timing of the HST entry from 2006 to 2016. The data covers over two thousand city-pairs (i.e., routes) over a period of eleven years. We exploit 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.

We use both reduced-form analysis and a structural model to answer the research questions. First, we use a difference-in-differences approach to provide quasi-experimental evidence of the causal effect of the introduction of the HST on airlines' network adjustments. We measure the change in airline presence in routes which experience the entry of the HST against untreated routes (which do not experience the entry of the HST). The results from the difference-in-differences analysis are consistent with both positive and negative effects from the HST on air service, and provide further additional insights. While the HST reduces airline presence on shorter air routes that overlap with HST service, in longer routes we observe the opposite effect, thus suggesting the existence of market expansion effects. In our difference-in-differences analysis, we also accommodate potential concerns with endogeneity that may originate from the fact that the roll-out of the HST does not proceed randomly.[4]

Next, to quantitatively assess the impact of the HST introduction on airline carriers' route network decisions and conduct counterfactual experiments, we setup and estimate a structural dynamic oligopoly model of airlines' network configurations. Because demand side data such as prices and seat occupancy is not available, we formalize the observed airline route presence as the equilibrium outcome of an entry game (e.g., Berry 1992, Ciliberto and Tamer 2009).

A structural model of airlines' network decisions poses several challenges. First, airlines operate in a network structure; in our setting, the typical airline serves 323 routes and could serve potentially $2,278$. Solving for the optimal network by complete enumeration is unattainable. Second, there are multiple large airline players and each firm's route decisions

[3]For example, in the United States, local airline carriers have opposed 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.

[4]With this purpose, we rely on an insight from Goolsbee and Syverson (2008) in their study of how incumbents respond to the threat of entry by a competitor in their study of entry of Southwest Airlines.

are dependent on their own and the other firms' networks. Lastly, airlines' decisions are inherently dynamic (i.e., firms are forward-looking) because there is a sunk cost associated with entering each route and there is uncertainty regarding the market conditions (demographics and train network) which change over time. Airlines reassess their networks often, which results in frequent entry and exit from the numerous city-pair routes.

We account for both the dynamic nature of airlines' competition and for the network structure that firms exhibit in this context in a similar manner to Aguirregabiria and Ho (2012), where airlines make decisions regarding their networks in a decentralized manner by keeping track of the state of the network through a sufficient set of statistics instead of considering the entire network configuration. Different from Aguirregabiria and Ho (2012), however, we adopt a continuous time framework (Arcidiacono et al. 2016, henceforth ABBE) for the airlines' dynamic entry game which allows us to take advantage of the lack of regularity in airlines' decision making and to make use of our high frequency data. (Modeling 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 and Aguirregabiria and Mira 2007.) In addition, a continuous time approach makes it possible to conduct counterfactuals despite the high-dimensionality of the dynamic game.

Our model extends ABBE's empirical setup in two novel ways. First, while in ABBE markets are independent, we allow the model to capture airline and train networks interdependencies. Second, because airlines do not exit routes in a permanent fashion (airlines often exit routes which they re-enter later on) we cannot apply ABBE's finite dependence representation which is frequently used to simplify the solution of dynamic models. We proceed by following a suggestion made in Rust (1996) and apply a GMRES (generalized minimal residuals) method (Saad and Schultz 1986) to solve our model. Although this makes the estimation computationally more intensive, our approach is still doable given the continuous-time formulation of the model.

The estimation of the structural model provides strong evidence of both negative and positive spillover effects from the HST on airlines, consistent with the results from the difference-in-differences analysis. These spillover effects depend on the routes' characteristics and on the interaction between the characteristics of the routes and the HST network. Specifically, we conclude that the HST is a strong substitute for air travel, especially in shorter routes. The substitution effects of HST relative to air travel dissipate for longer routes, however. Air routes with a larger number of connections to HST lines benefit more from positive spillovers.

We perform counterfactual analyses using the structural model to quantify the effects of the entry of the HST on the airline industry. Using the model to simulate the endogenous equilibrium outcomes is necessary because the evolution of the airline network during the period studied naturally depends on the market environment. First, we simulate the network configurations of the airline carriers in a scenario in which the HST is not present, and compare the endogenous airlines' route decisions and profits to the baseline case in which the HST is introduced. Because the HST has a heterogeneous effect on the airline industry, we also explore the sources and impact of such heterogeneity.

We find that, overall, despite the existence of significant positive spillovers from the HST on the airline industry, the introduction of the HST reduced airlines' route presence by about 14% and airline profits by 23%. Even though the overall net impact of the introduction of HST is negative, the results reveal considerable heterogeneity across cities and route-types in how the HST impacted the airline industry. Airlines readjusted their networks by substituting towards longer routes and more peripheral regions in China. Airline entry in these routes would not have occurred in the absence of the HST, which made them profitable for airlines. Thus, we find that the HST did not have a pure crowding-out effect but rather led to an expansion of the size of the overall network through complementarities. This highlights a potential indirect benefit of the HST entry in shifting airlines to more remote and underserved areas, improving connectivity among regions and reducing inequality.

In another experiment, we simulate an increase in the positive spillover effects between the HST and the airline industry to explore the possible benefits from services that facilitate the complementarity between the two modes of transportation. This experiment is motivated by the government's goal to increase people's mobility through a better integration of air and rail travel, and allows us to assess how the efforts to improve the degree of connectivity between the two modes of transportation may compensate the negative effects of the HST on airline entry.

We find that, although a uniform increase in the strength of positive spillovers from HST can compensate the negative impact of HST on air carriers' overall network size, the increase in airline profits associated with this improvement does not affect all air routes uniformly. The higher elasticity of the airline service provision in some regions with respect to the improvement in the positive spillovers from the HST makes them especially important targets for providing intermodal connectivity services.

The rest of the paper is organized as follows: Section 2 reviews the literature. Section 3 describes the industry background. Information on the data is provided in Section 4. Section 5 provides reduced-form evidence of the effects of the HST on the airline industry. Section 6 presents the structural model and the model's empirical specification and estimation strategy. The results from the structural model are presented in Section 7. Finally, Section 8 presents the counterfactual analyses and Section 9 concludes.

2 Literature

This paper is related to several streams of literature. First, to the extent that this study uses route entry and exit information to estimate airline's profit functions, this paper contributes to the empirical industrial organization literature on firm entry and spatial competition. Empirical static discrete games have been used to capture firms' strategic interaction behavior (e.g., Bresnahan and Reiss 1990, Bresnahan and Reiss 1991,Seim 2006).[5] Several studies in this literature have focused specifically on the coexistence of both negative and positive spillovers from competitors. This makes them closer to our research given that we empirically quantify the positive and negative spillovers generated by the introduction of the HST on the airline industry and among airlines. For example, 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 spillovers affect firms' entry decisions.[6]

A subset of the literature on firm entry focuses on the "chain effects" of firms with multiple stores (e.g., Jia 2008, Holmes 2011, Aguirregabiria and Ho 2012, Nishida 2014 and Zheng 2016). 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 estimate static oligopoly models with two players. Holmes (2011) allows for dynamic network decisions of Walmart, but abstracts away from the competition between Walmart and other chain stores. Similar to these papers we allow for interdependencies across markets to capture the features of the transportation network that characterizes the airline industry. However, we use a dynamic model (with multiple players) which helps us to better understand the evolution of the airline industry over time as a result of the introduction of the HST.

Our paper also contributes to the growing literature on transportation. Recent research

[5]As opposed to other transportation markets, such as the taxi market (e.g., Fréchette, Lizzeri, and Salz 2019), where there are tens of thousands of players and thus entry is competitive and players only keep track of the aggregate state of the market, in our context, airline firms are strategic as in most studies in the entry literature.

[6]In a network transportation context, Cao et al. (2021) study the dock-less bike sharing industry and find reduced-form evidence that the entry of a competitor may improve the incumbent's market coverage and profitability. In their follow-up theoretical model, they show that this could happen due to the complementarity between the spatial networks of the two firms in the market.

has developed richer transportation and spatial equilibrium models which address a variety of topics.[7] Spatial models have been used, for instance, to study the role of network effects on trade costs (e.g., Brancaccio, Kalouptsidi, and Papageorgiou 2020), the consequences of transportation policies on congestion (e.g., Barwick et al. 2022), the spatial sorting of labor (e.g., Fajgelbaum and Gaubert 2020), and search frictions in decentralized spatial markets (e.g., Fréchette, Lizzeri, and Salz 2019, Liu, Wan, and Yang 2021, Brancaccio et al. 2022, Buchholz 2022). Our research is especially related to a subset of the literature that studies the impact of infrastructure development. In particular, the papers by Heblich, Redding, and Sturm (2020), Donaldson and Hornbeck (2016) and Donaldson (2018) study the impact of developments in the railway network. Using general equilibrium models, Heblich, Redding, and Sturm (2020) quantify the impact of the invention of the steam railway on the separation of workplace and residence in London. Donaldson and Hornbeck (2016), and Donaldson (2018) evaluate the welfare implications of railroad development in the US and India, respectively. Our paper differs from these these studies because, while these papers focus on the impact of changes in infrastructure on economic output variables such as trade costs, congestion, and unemployment, our ultimate interest is on the heterogeneous and dynamic response of airlines across geographic space to the development of a new means of transportation (the HST). Further, the HST may act not just as a substitute but also as a complement to air travel. Our model also allows the airline industry to adjust their network configuration (endogenously) in response to the HST entry in a context of oligopolistic competition.[8]

Methodologically, we build on the literature on dynamic estimation of equilibrium entry games. Most of this literature has used conditional choice probability estimators (CCP) (Hotz and Miller 1993 and Hotz et al. 1994) to study a wide range of dynamic discrete choice problems including simultaneous-move dynamic games (Aguirregabiria and Mira 2007; Bajari, Benkard, and Levin 2007; Pakes, Ostrovsky, and Berry 2007; Pesendorfer and Schmidt-Dengler 2008). From this literature, the paper closest to our research is Aguirregabiria and Ho (2012) who were the first to propose a dynamic game of network competition and to estimate it in the context of the US airline industry.[9]

[7] Redding and Rossi-Hansberg (2017) and Redding (2021) offer two reviews on the broader literature on economic geography and spatial economics.

[8] Fajgelbaum and Schaal (2020) and Allen and Arkolakis (2022) have recently studied optimal (and endogenous) transport network design within a general equilibrium spatial trade framework. Different from our focus, Allen and Arkolakis (2022) measure the impact of road infrastructure investments on congestion and transportation costs, while Fajgelbaum and Schaal (2020) study optimal network investments subject to congestion in a context in which a social planner is responsible for building the road infrastructure.

[9] There are other recent papers which have studied the the airline industry but Aguirregabiria and Ho (2012) is the most closely related to ours. For example, Wei (2018) focuses on airlines' network value from the perspective of traffic density (as opposed to the hubbing effect) and provides evidence that consumers

Different from Aguirregabiria and Ho (2012), who model airline's decisions in discrete time, we adopt a continuous time framework for the airlines' dynamic entry game. Further, 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. We adopt Arcidiacono et al. 2016's (henceforth ABBE) continuous time framework for the airlines' dynamic entry game to accommodate our high frequency data (airlines reassess their networks often, which results in frequent entry and exit) and high-dimensional game. ABBE developed a general framework for estimating and solving dynamic discrete choice models in continuous time which has been used to study dynamic entry games in several settings.[10] We further extend the ABBE methodology to accommodate our context by allowing for interdependencies across markets and to allow for airline re-entry into routes.

Also, 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, Eizenberg 2014, Fan and Yang 2020, Viswanathan, Narasimhan, and John 2021). 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. Fan and Yang (2020) examine the relationship between oligopolistic competition and product offering and find that a reduction in competition decreases both the number and variety of products.

prefer routes with higher flight frequencies. His paper structurally models preferences in consumer demand and solves for air carriers' optimal pricing strategies in interconnected air routes, but abstracts away from firms' decisions regarding their network configuration and from industry dynamics. Ciliberto, Murry, and Tamer (2021) and Li et al. (2022) jointly model airline entry and pricing but their focus is on airline mergers and each route is taken as an independent market.

[10]Recent papers have used the ABBE framework to study the dynamic entry games of firms in various scenarios. For example, Cosman (2017) studies the entry and exit decisions of firms in night life venues and finds evidence of strong customer preferences for variety. Smith (2018) applies a continuous-time entry model to estimate the net impact of globalization on the clothing industry, and finds that direct imports account for at least 14% in the decrease in the number of small clothing stores. Both papers approximate the value function based on a finite dependence property, which assumes that firms' exit decisions are permanent. Zhang (2020) models banks' branching decisions and assesses the long-run implications of banking service digitization and competition from fintech mortgage lenders. Other literature has also used ABBE's methodology to study single agent discrete choice problems. For example, Deng and Mela (2018) use set-top box viewing data to develop an instantaneous show and advertisement viewing model and find that device level advertising targeting is more effective than show-level targeting. Nevskaya and Albuquerque (2019) model consumers' gaming decisions with high-frequency data and proposes strategies for firms to manage excessive product use.

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. Our setting is different in that there can be complementarity among the different firms. This difference may play a key role in determining firms' product assortment decisions.

Finally, this paper contributes to the growing studies on the economic impact of the HST. The existing transportation and economic geography literatures have examined the effect of the HST on various outcomes, such as economic growth and regional equity (e.g., Qin 2016, Yao et al. 2019, Banerjee, Duflo, and Qian (2020), Zhang et al. (2020), business development and job creation (e.g., Heuermann and Schmieder 2018, Shi et al. 2020, Chen et al. 2022), and the re-distribution of healthcare resources (e.g., Chen, Hao, and Chen 2021, Yoo, Vitorino, and Yao 2022). Studies of the impact of the HST on the airline industry, however, tend to view the two modes of transportation as substitutes and therefore emphasize the negative spillover effects of the HST (e.g., Wang, Bonilla, and Banister 2016, Chen 2017, Li and Loo 2017. Also see Zhang, Wan, and Yang 2019 for a more detailed review of the literature). In contrast, our paper focuses on the co-existence of negative and positive spillovers from the HST to the airline industry. In this sense, our paper is closer to Liu et al. (2019) and Zhang, Wan, and Yang (2019), who find that the introduction of HST may have both positive and negative impact on airports depending on their types (i.e., international versus domestic) and levels (more versus less) of air connections. While the previous studies rely on reduced-form analysis to quantify the impact of HST on the airport level, our paper builds a structural model, which allows us to assess the impact of the HST on the network configuration of the airline industry.

3 Industry Background

We use data from the airline and HST industries in China. Before describing the data, we provide a brief overview of these industries to help provide context for the subsequent analyses. We focus mostly on the period from 2006 to 2016 which corresponds to the data period used in the empirical analysis.

3.1 The Airline Industry in China

Air-passenger volume in China went from 237.1 billion passenger-km in 2006 to 837.8 billion passenger-km in 2016.[11] This rapid growth was in large part the result of an increase in the number of routes served and in the number of flights offered by airlines, possible due to the reduction of regulatory oversight on the airline industry.[12] In 2006, a significant deregulation effort lifted constraints imposed on airlines regarding their ability to add or drop flights from a route and, since then, airlines only need to apply for permission to add flights to routes that have the cities of Beijing, Guangzhou and Shanghai, as one of the route endpoints. Routes that do not involve these cities as endpoints require only a simple registration procedure to add a flight. Shanghai, Beijing, and Guangzhou are subject to stricter regulatory constraints because, as the busiest airports in China, the government carefully moderates the air traffic volume that can be accepted into these airports. An exception to this rule applies to airlines that are headquartered in Beijing, Guangzhou or Shanghai which are exempt from government approval whenever they wish to add a flight to a route that connects the airline's headquarters-city to other cities (except for Beijing, Guangzhou and Shanghai). Furthermore, since 2006, dropping a flight from any route requires only a simple cancellation request.

Figure 1 lists the major airline companies in China grouped by parent company. 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). These four airlines are publicly traded companies. Between 2006 and 2016, their combined share of passenger volume (including their subsidiaries) was approximately 90%.

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

Insert Figure 1 about here

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

3.2 The Railway Industry in China

Over the past decade, China has built (or upgraded) over $20,000$ km ($12,500$ miles) of high speed railways, more than the rest of the world combined. In 2016, more than 1.47 billion

[11] Source: National Bureau of Statistics of China, 2017.

[12] Wang, Bonilla, and Banister (2016) and Yang, Zhang, and Wang (2018) provide a detailed account of the history of regulatory oversight of the airline industry in China.

passenger trips, corresponding to 404 billion passenger–kilometers, were made by HST, which is about half of those made by air.[13]

The central government of the People's Republic of China, directly or indirectly through state-owned companies, develops and operates the entire train infrastructure in China including both the traditional and the high-speed rail networks. This means that the decisions regarding the operations of the different train routes are not driven by profit maximization motives (until 2016, the only HST route that was able to operate profitably was the route between Shanghai and Beijing[14]). The government views the Chinese train infrastructure as a source of international prestige, with a significant impact on people's lives through their mobility.[15]

In 2004, the government adopted an ambitious long-term rail development plan outlining the expansion of the HST rail network up to the year 2020. The plan included both the construction of about 12,000 km (7,500 miles) of new HST rail routes and the upgrading of about 16,000 km (10,000 miles) of existing rail routes to accommodate trains of higher speed. The expansion of the HST route network was, for the most part, predetermined even though there was some uncertainty regarding when and whether specific routes would be served by the HST.[16]

There are two types of HST. The "fast train" is capable of achieving a maximum speed of 250 km/h (about 155 mph) while the "bullet train" can achieve a speed of up to 350 km/h (about 215 mph). While the routes for these two types of trains often overlap, they operate on different rails.[17] In terms of pricing, for the same route, a bullet-train ticket is about 60% more expensive than a fast-train ticket.[18]

The first fast-train and bullet-train lines started operations in 2007 and 2008, respectively, connecting major cities such as Beijing, Guangzhou, Shanghai, Tianjin, and Wuhan. From 2007 to 2016, the network expanded rapidly and China spent an estimated 2.4 trillion yuan (353 billion USD) building 22,000 km (13,670 miles) of high-speed rail lines, more

[13]Source: `https://www.qianzhan.com/analyst/detail/220/181031-61b71943.html` (in Chinese and last accessed on July 5, 2022).

[14]Source: `http://view.163.com/special/resound/chinahsr20160721.html` (in Chinese and last accessed on January 24, 2019).

[15]Source: `https://www.straitstimes.com/asia/chinas-rail-ambitions-run-at-full-speed` (last accessed January 24, 2019).

[16]The plan adopted in 2004 was updated in 2008 and in 2016, mostly with the purpose to add more routes to the existing planned network.

[17]Upgraded existing rails can only accommodate fast trains, but not bullet trains. There is no difference in terms of speed between fast trains operating on newly built rails or on existing rails that have been upgraded.

[18]This number is calculated based on the price data (for a seat in economy class) collected on November 15th, 2016 for routes where both fast trains and bullet trains were in operation.

than the total high-speed railway length of the world combined.[19] Figure 2 shows the geographical expansion of the HST network in China between 2007 and 2016. The number of train passengers has also grown rapidly, at a nearly 90% annual growth rate (compounded), surpassing the number of domestic passengers in the airline industry in 2011.[20]

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

Insert Figure 2 about here

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

3.3 The Effect of the HST on the Airline Industry

The introduction of the HST had a significant impact on the airline industry. As reported in the press, airlines exited several routes and reduced the number of flights offered to avoid direct competition with the HST, consistent with the HST being a substitute for air transportation.[21]

Traveling by train can be more appealing than traveling by plane for several reasons. For example, trains are better for the environment, are more comfortable, tend to be more punctual and train tickets are in general cheaper. However, the extent to which air and rail are perceived as substitutes depends significantly on trip length. Figure 3 illustrates the relationship between door-to-door travel time and travel distance for both air and HST transportation.[22] For short routes (under 600 km in length), the door-to-door travel time by HST is shorter compared to air travel, even though airplanes are faster than HST. Airports are more remotely located, require longer check-in times and more security checks, all of which increase the door-to-door travel time. For these reasons, the HST has a larger strategic advantage in relatively shorter routes.

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

Insert Figure 3 about here

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

At the same time that the HST can be a substitute for air travel, it can also have a complementary role to air transportation. For example, the HST can connect a node within an airline network to follow-on destinations that may not be served by air.

[19]Source: `https://www.yicaiglobal.com/news/china-high-speed-rail-network-is-larger-than-the-rest-of-the-world` (last accessed January 24, 2019).

[20]Source: China/HK Airlines Sector Report, DBS Group Research, 2018

[21]Sources: `https://www.bloomberg.com/news/articles/2018-01-09/high-speed-rail-now-rivals-flying-on-key-global-routes` (last accessed November 6, 2019), and `http://news.xinhuanet.com/fortune/2011-04/12/c_121293247.htm` (in Chinese and last accessed on November 20, 2019).

[22]Source: `https://www.pinchain.com/article/95181` (in Chinese and last accessed on Oct 29, 2016)

Intermodal travel can be facilitated by explicit cooperation efforts such as the development of transfer connections between airports and railway stations, through ticketing involving flight and HST tickets, and end-to-end baggage handling. Despite these potential benefits, explicit cooperation between the two different modes of transportation has not been significant in China.

The complementarity between the two modes of transportation can also lead to a more indirect effect on the expansion of the travel market. With the entry of the HST, travel agencies have started to offer a larger variety of travel packages. For example, in 2013, five cities in the Northeastern part of China (Dalian, Shenyang, Changchun, Haerbin and Changchun) jointly introduced a travel package that combines the tourism resources of the five cities. Their promotional slogan read "Fly to Dalian, take a HST and tour Northeastern China". The HST makes travel between the five cities more convenient and increases the attractiveness of the combined package, which in turn leads to more flight sales.

Taken together, this evidence highlights that, in order to quantify the effect of the entry of the HST on the airline industry, one needs to consider both its negative and positive effects. Further, the heterogeneity of the different routes (in terms of, for example, their length and connectivity level) means that the effects are also likely to vary depending on the routes' characteristics.

4 Data

4.1 Data Sources and Sample Selection

It is difficult (if not impossible) to obtain comprehensive price and quantity data that encompass both the airline and rail industries in China. So, to study the effect of the entry of the HST on the airline industry, we rely mostly on market structure data. Specifically, we assemble a unique dataset with flight-schedule information and the detailed timeline of the HST introduction that spans the 11-year period from January 1, 2006 to December 31, 2016.

We obtained flight schedule data from a website that archives historical flight data. We extracted information for the top 70 airports in China, which translates into over 20 million observations.[23] Each record includes the date, airline, flight number, and origin/destination cities. This information allows us to uncover the routes covered by each airline and the

[23] The airport ranking we refer to is based on 2015 passenger volume provided in the Statistical Year Book Civil Aviation Administration of China 2015. The top 70 airports account for more than 95% of passenger volume in China and correspond to a coverage of 68 cities (Beijing and Shanghai each have two airports).

number of flights operated by each airline daily. We define a route to be a non-directional city-pair. Henceforth, we use route, market and city-pair interchangeably.

We focus on the major four airlines (including their subsidiaries) Air China (CA), China Southern Airlines (CZ), China Eastern Airlines (MU) and Hainan Airlines (HU). They cover about 96% of the number of the unique number of routes served between 2006 and 2016 (not tabulated) and account for about 88% of the number of flights operated during the period studied (not tabulated).

To capture the effective coverage of each airline, we focus on flights that operate regularly. Therefore, we exclude seasonal flights (such as those only provided in occasions such as the Chinese Festival or Christmas) and infrequent flights (which may reflect local government subsidies), hence we only consider flights that are offered on a schedule for longer than one year. Accordingly, a route is only served by an airline at a given point in time if a flight that operates regularly is in operation. To guarantee we satisfy this data-selection criterion, we drop the first and the last years in the dataset and thus focus on the time period from the beginning of 2007 to the end of 2015.[24]

To track the timing of the HST introduction on a route-by-route basis we supplement data from multiple government websites and from news reports with historical data scraped from `www.12306.com`, the official website used to purchase train tickets online in China. These data allow us to identify, at each point in time, which cities are part of the HST network and the type of HST that operates in each route (fast train or bullet train).

Lastly, we supplement the flight and train data with yearly demographic data obtained from the *China City Statistical Yearbook* published by the National Bureau of Statistics of China.

4.2 Summary Statistics and Data Patterns

In this section we provide summary statistics of the variables that characterize the routes (i.e., city-pairs) in our data and that can potentially help explain the airlines' network structure and the HST's introduction patterns. We also explore descriptively the relationship between the route characteristics and both the airlines' decisions and the HST expansion and investigate the overlap between the airline industry network and the HST rail coverage. The descriptive statistics and the relationships established in this section will be important for the next sections in which we model how the HST affects airlines' network decisions.

Tables 1 and 2 provide summary statistics for all the routes in our sample after aggregat-

[24]In several analyses in the paper we conduct robustness checks in which we relax this data-selection criterion. For example, we conduct several analyses that include all flights that operate for longer than three months.

ing the variables related to the airline data from daily to annual. Table 1 pools information across all years in the sample and Table 2 provides information regarding the variables' averages for each year. In these two tables we consider all potential routes, including those in which no flights or HST are in operation. In total there are $2,278$ routes in the data with an average length of about $1,500$ km. In the analyses that follow we categorize the routes into three groups in terms of route length: short, medium and long, using 600km and 1200km as cutoff points.[25] Most routes are long (60% of routes), followed by medium (28%) and short (13%).

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

Insert Tables 1 and 2 about here

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

Overall, our data reflects the country's and the transportation industry's growth discussed in Section 3.1. For the city-pairs and years studied, GDP grew at an average rate of 15% per year. Airline route coverage went from 25% in 2007 to 37% in 2015 at the same time that the average number of airlines operating in a route went from 0.44 in 2007 to 0.70 in 2015. Likewise, the average number of flights per route also went up from 1.06 in 2007 to 2.02 in 2015.[26] The statistics also show that airlines concentrate their operations in a subset from all of the possible city-pairs with 70% of routes (on average, across all years) not having any airline presence. The number of flights is also quite unevenly distributed across the routes with flights: 50% of the total number of flights are concentrated in about 30% of all routes that have flights (not tabulated).

Table 2 and Figure 2 show the rapid expansion of the HST network. HST route coverage went from 7% in 2007 to 21% in 2015 for fast trains and from no coverage in 2007 to 10% in 2015 for bullet trains. In 2013 and 2014 there was a particularly significant expansion of the HST network. The HST route coverage for bullet trains increased from 2% in 2012 to 7% in 2013, and the fast train coverage went from 11% in 2013 to 19% in 2014. There were significant changes in the level of overlap between the set of routes served by airlines and those served by the HST over time. As Figure 4 shows, between 2007 and 2015, the proportion of air-routes which faced direct competition from the HST increased. By the end

[25]These cutoffs are motivated by our discussion regarding the differences in competitive advantage of airlines and the HST; industry reports support the notion that the HST has more advantage for routes shorter than 600 km and airlines have more advantage for routes longer than 1200 km. Source: `http://www.pinchain.com/article/95181` (in Chinese and last accessed on June 9, 2022).

[26]Note that the average number of airlines and flights that operate in a route reported here can be less than one. This is because we are using the total number of *possible* routes (20,502) as the denominator in our average calculation. If we use as the denominator the total number of routes covered by airlines (564 in 2007 and 833 in 2015), the average number of airlines operating in a route is 1.79 for 2007 and 1.91 for 2015, and the average number of flights per route is 4.24 for 2007 and 5.42 for 2015 (not tabulated).

of 2015, more than 35% of the airline routes were served by fast trains, and more than 15% by bullet trains, compared to 15% and 0% in 2007 for fast and bullet trains, respectively.

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

Insert Figure 4 about here

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

Figure 5 shows that, even though there was an increase in overall airline route-presence, there was a significant number of air route exits as well (i.e., routes in which airlines had been operating for more than one year and then stopped operating). On average, for every four routes that experienced entry of one or more new airlines, there were about three routes that experienced airline exit.

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

Insert Figure 5 about here

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

According to Table 3, where we report airlines' route decisions and the presence of HST according to the distribution of route length across all years in the sample, most airline entry and exit occurred in routes of short and medium length. Further, shorter routes saw significant exit and, in net terms, were the ones that experienced the least growth in terms of airline presence.

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

Insert Table 3 about here

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

To better characterize the link between HST presence and airline entry and exit, in Figure 6, we display the evolution of airline route-presence as a function of the degree of overlap and connectivity with the HST.[27] Panel A describes the evolution in the number of airlines serving routes that overlap with the HST for short and medium/long routes, separately. The figures show that, while there is a decline in the number of airlines present in short routes that overlap with HST, there was significant airline entry in longer routes. This pattern is consistent with the fact that airlines have a strategic advantage relatively to the HST in longer routes. Turning to the airline routes which do not overlap with the HST, Panel B compares the evolution of airline presence in routes connected with the HST with routes not connected. We find that routes connected to the HST exhibit more airline presence (0.4

[27]Patterns are similar if, instead of looking at the evolution of airline route presence, we look at the evolution in the number of flights.

airlines are present on average in routes connected to the HST, compared to 0.1 for routes not connected). Further, even though in both types of routes (connected vs. no connected) airline exit was dominated by airline entry, in routes connected to the HST the growth in (net) airline entry was 3 times larger on average.

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

Insert Figure 6 about here

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

This descriptive evidence suggests that airline presence depends on both HST presence and connectivity to routes served by the HST, and points to the possible existence of both positive and negative spillovers from the HST to airlines.

Independently of the entry of the HST, because airline route decisions are interdependent, we expect an airline's decision to enter or exit a given route to be affected by the airline's operations in other routes, especially in (directly) connected routes that share the same endpoint airport (e.g., Berry 1992, Goolsbee and Syverson 2008, and Ciliberto and Tamer 2009). We descriptively assess the impact of endpoint-airport presence on an airline's decision to enter a new route by estimating a probit model in which the dependent variable is a binary variable equal to one if an airline enters a given route in year t and zero otherwise. As independent variables of interest we include dummy variables that capture whether an airline operates in none, one or both endpoints of a given route in year $t-1$. Year and airline fixed effects are also included in the regression. Only route-year combinations in which an airline is not present in year $t-1$ are considered. Table 4 reports the estimated marginal effects of airport presence on airlines' entry probabilities.

The results show that operating in one or both endpoints of a given route (as opposed to none) is associated with a significantly higher probability of entering that route. Further, the marginal effect on route entry of operating in airports located in both endpoints of a route is almost four times larger than that of operating in only one of the endpoint airports – when compared to the baseline, operating in both airport-endpoints of a route increases the probability of entry by about 11 percentual points per year. These results are directionally consistent with those that Goolsbee and Syverson (2008) obtain in a similar analysis and support the notion that route connectivity influences airline network decisions.

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

Insert Table 4 about here

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

5 Reduced-Form Evidence

We are ultimately interested in understanding and quantifying how the entry of the HST affects airline decisions. Before studying more formally the strategic interaction between the airlines and their response to the HST in the context of a structural model, in this section, we establish that the entry of the HST affects airlines' route decisions and thus that there is a causal link between the entry of the HST and the structure of the airline industry in China. In addition, we show evidence suggestive that the entry of the HST generates both positive and negative spillovers to the airline industry given its heterogeneous effects.

5.1 Difference-in-Differences Estimates

We use a difference-in-differences (DID) specification to assess the relationship between airline and HST presence. This allows us to exploit the variation in the introduction of the HST over time, and across routes with different characteristics. The DID compares route airline-presence before and after the entry of the HST while controlling for unobservables that are shared by the different routes and that change over time (by using control routes). Because we have multiple routes and time periods (and routes are not "treated" all at the same point in time), we use the following DID specification:

$$\text{Airline Presence}_{rt} = \lambda_r + \alpha_t + \beta \text{HST}_{rt} + \mathbf{X}_{rt}\boldsymbol{\gamma} + u_{rt}, \tag{1}$$

where r indexes route, and t indexes time (year).[28] We include year fixed effects, α_t, route fixed effects, λ_r, and route-specific covariates, $\mathbf{X}_{rt}$.[29] The term u_{rt} is a route-time specific error. The dummy variable "HST" indicates whether the high-speed train was present in route r by time t. The coefficient on "HST" is interpreted as the average increase/decrease in the number of airlines present in a route that is attributable to the entry of the HST. We allow for heterogeneous effects of the treatment across routes by interacting the variable HST with route-specific covariates such as route length and the number of HST connections. The dependent variable, Airline Presence, is defined as the number of airlines present in a given route in a given year; results are robust to using the total number of flights in a route instead.

The difference-in-differences estimation technique allows us to control for omitted vari-

[28] For example, see Bertrand, Duflo, and Mullainathan (2004) for a similar model setup to ours.

[29] Some of the route-specific covariates change over time (e.g., GDP) while others remain constant throughout the period studied (e.g., route length). Whenever route-fixed effects are included in the model, the route-specific covariates that do not change over time are naturally not identified (only their interaction with other covariates that change over time is identified).

ables. The year fixed effects control for nationwide shocks and trends that may affect the airline industry over time, such as business cycles and changes in regulation at the national level. The route fixed effects control for time-invariant, unobserved route characteristics that may explain why some routes are favored by airlines relatively to others. We estimate equation (1) allowing for route-level clustering of the errors, i.e. allowing for correlation in the error terms over time within routes.

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

Insert Table 5 about here

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

Table 5, column 3, reports the results for the DID model as specified in (1). The specification in column 1 does not account for time-varying nor route-varying factors. The second column accounts for time-varying but not for route-varying factors. Because the specifications in the first two columns do not include year and route fixed effects they do not correspond to implementing a difference-in-differences approach but they are useful for comparison purposes. Further, because these specifications do not include route-fixed effects, they allow us to get a better understanding of how routes of different lengths are affected by the entry of the HST.

Turning to the results for the specification in (1), reported in Table 5 column 3, the negative and significant coefficient on "HST" indicates that airline presence declines with HST presence. This points to negative spillovers from the HST to the airline industry. However, the negative spillovers disappear for medium and long distance routes, consistent with the HST having a larger strategic advantage relative to airlines in shorter routes.

The coefficient on the number of HST connections is positive and significant which suggests that there are positive spillovers from the HST to the airline industry in routes that are connected to the HST. These positive spillovers are negatively moderated by HST presence as reflected in the negative coefficient for the interactions of "HST" with the number of connections to the HST. These results seem sensible since we expect the benefits that airlines derive from serving routes connected to the HST to be reduced when those routes are also served by the HST; in the latter case, consumers may find it more convenient to not switch modes of transportation when traveling.

The coefficient on the number of airline connections is, as expected, positive and significant, consistent with an airline's operations in a given route being affected by its presence in connecting routes. While this variable allows us to control in a reduced-form way for the impact of airlines' networks on their presence in a focal route, in the structural model specified in the next section, airlines' strategic decisions will be explicitly modeled taking into account not only the evolution of each airline's network but also those of the competitors.

The identification of the effect of the HST in airline presence in the DID approach relies on the assumption that airline presence on treated and control routes follows the same trend in the absence of the HST. To test for parallel trends in the case of multiple treatment and control groups, we follow the literature (e.g., Angrist and Pischke 2009) and run falsification tests by adding "leads" to the set of independent variables in specification (1). Specifically, we add dummies that are equal to one for the first two pre-treatment years for each treated route. The estimated coefficients for the two dummy variables are not significantly different from zero (with both p-values greater than 0.20; not tabulated) suggesting that the parallel trends assumption is satisfied.

The DID analysis described above uses all routes without HST presence as control routes at each point in time. While this identification strategy captures overall country trends that could cause changes in airline presence even in the absence of HST entry, there could be a concern that different routes are subject to different local shocks. Consider, for example, the 2008 Olympics held in the city of Beijing. It would be reasonable to expect that airlines increased their presence in the routes connecting to this city due to this exogenous demand shock. If, around the same time, the HST was introduced in some of these routes, then it may seem that the HST has a positive effect on airline presence when that is not the case. More generally, if there are local shocks that are somehow correlated with both HST and airline presence this could bias our estimate of the causal effect of HST presence.

To address this concern, an alternative DID specification could use, for each individual treated route, a more refined control group of routes to try to capture systematic local differences due to reasons other than the entry of the HST. For example, Card and Krueger (1994), in their study of minimum wage effects, and Tuchman (2019), in her study of the effects of e-cigarette advertising, control for local trends by comparing boundary counties or states, respectively, that are exposed to different treatments but that are expected to experience the same unobservable shocks due to their geographical proximity. Unfortunately, because our analysis is at the level of a route (defined as a city-pair), the lack of well-defined "market" boundaries does not allow us to account for unobservable local trends by using a geographic-border strategy as the one pursued in the studies above.

Instead, we rely on an insight from Goolsbee and Syverson (2008) in their study of how incumbent airlines respond to the threat of entry. They use characteristics of routes that share the same endpoints as a focal route to control for potential unobserved shocks in the focal route. We adapt this idea to our DID setting by refining the definition of the control groups used in the analysis and selecting as control routes only those that share one of the same endpoints (i.e., airports) as the treated routes but which do not overlap with HST routes (which are "treated").[30]

We illustrate the idea behind the selection of the control routes in Figure 7. If we consider as the focal route the city-pair AB, which overlaps with the HST, we can use routes AC and BD as control routes. The identifying assumption is that the control routes AC and BD potentially experience the same unobserved demand and supply shocks as the focal treated route AB. Returning to the 2008 Olympics example mentioned above, the intuition behind this approach is that, by comparing routes that are connected to Beijing but which differ in terms of HST presence, we are able to measure the effect of HST presence on airline decisions after netting out the 2008 Olympics effect.

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

Insert Figure 7 about here

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

More specifically, for each route where the HST was introduced sometime between 2007 and 2015 we find two control routes that satisfy the following criteria. First, the two control routes (routes AC and BD in the diagram in Figure 7) cannot overlap with the HST (i.e., these routes do not experience HST entry at any point in time). In addition, to make sure that the control routes are not substitutes of the routes which overlap with the HST, which would make them "treated" to some degree, we restrict the distance between the non-shared endpoints of the treated and each of the control routes (cities A and D, or cities B and C, in the diagram in Figure 7) to be more than 300 km (186 miles).[31] Lastly, and more importantly, each control route must share one of the end cities with the treated route.

To implement this approach we adjust the set of fixed effects included in specification in (1) by adding route fixed effects to control for the time invariant differences across routes and allow each route to have its own average level of airlines present, and group-year fixed effects to allow each group of treated and two control routes to have its own flexible trend in airline presence and thus account for any unobserved local shocks. This analysis also has the added benefit of addressing the concern pointed out by de Chaisemartin and D'Haultfœuille (2022) which occurs in staggered DID designs with multiple time periods and treatment groups like in the two-way fixed effects design in equation (1), with results reported in Table 5, column 3. That is because we use as control routes the routes in which the HST did not

[30]More precisely, this can be considered a "Matching-DID approach" in which a matching method is used to select the control routes. (Further details on this method can be found in Blundell and Costa Dias 2000). We also used a matching approach without combining it with DID (more specifically, we used both exact matching and propensity-score matching) and the results obtained are consistent with the ones reported using the DID approach.

[31]Robustness checks with 400 and 500 km cut-offs provide similar results and are available upon request from the authors.

enter. Our method is similar to the method proposed by Callaway and Sant'Anna (2021) with the difference that they used propensity score matching to specify the control routes and we do the matching using criteria more tailored to our context and more reflective of the analysis in Goolsbee and Syverson (2008).[32]

Column 4 of Table 5 shows the results. Note that the sample size decreases to 7,884 route-year observations when compared to the DID results reported in column 3. This is a direct result of having further refined the definition of what constitutes a control route which naturally restricts the set of routes included in the analysis.[33] Despite the smaller sample size, this analysis can be useful as a stricter test of whether and how HST presence affects airlines decisions. The results are directionally, and for the most part quantitatively, consistent with the results from the main DID specification reported in column 3. Namely, we find that the presence of HST may have a positive or negative impact on airline presence, depending on the length of the route, and that connecting to HST lines has a positive impact on the number of airlines serving a route. Again, we test for the parallel trends assumption in the same manner as in the main DID specification and we find evidence that the common trend assumption cannot be rejected.

We further conduct a variety of checks to test the robustness of the differences-in-differences results. First, we carry out the analysis at the airline-route-month level (as opposed to at the airline-route-year as in the main specifications) and restrict the set of treated routes such that we only include data for treated routes for one month before the entry of the HST and one month after the introduction of the HST. We do this to alleviate concerns related with route-specific trends which may generate a spurious relationship between the entry of the HST and airlines' decisions: for example, it could be the case that both the HST and airlines tend to enter routes where there is expected to be a growth in demand (alternatively, it could also be that the HST was designed to enter some routes to slow down or reverse the expected decrease in demand). If this is the case, and the entry/exit decisions of the two modes of transportation are truly independent of each other, then it is unlikely that their entry/exit decisions will happen at the same time. This can be tested by examining the airlines entry/exit decisions in a route right before and after the introduction of the HST. Therefore, by comparing the airlines' entry decisions shortly before and after the introduction of the HST, we can provide stronger evidence of the causal impact of the HST on air carriers.

[32]We have however tested the robustness of our analysis using the method proposed by Callaway and Sant'Anna (2021) and find that the main results are consistent with our specification.

[33]From the the 471 routes that experience HST entry during the sample period, we are able to find control routes for 292 of them. This means we have a total of 876 routes over a nine-year period, corresponding to 7,884 observations.

Second we use the number of flights operating in a route as a dependent variable in equation (1) instead of the number of airlines. Finally, we re-run our analysis using all flights that operate longer than three months (as opposed to including only those that operate longer than one year).

The results of the different robustness checks are reported in Appendix A. We find that our results are robust across the different specifications and sample definitions.

6 Structural Model

The previous section establishes that the entry of the HST had an impact on airlines' route decisions. Here, we structurally estimate a model in which airlines make route network decisions as a function of their competitors' actions and of the presence of the HST. This allows us to quantify the impact of the entry of the HST on different markets (i.e., routes) depending on their characteristics, and to determine how the heterogeneous impact of the HST affects airlines' networks configurations. The model also allows us to investigate counterfactual scenarios.

6.1 Overview

Airlines' route decisions are difficult to model. Airlines are forward-looking strategic players that maximize profits while anticipating the impact of current route decisions on current and future payoffs and their competitors' reactions. Further, because airlines operate in a complex structure of route connections, individual-route decisions have implications for the entire route network which need to be taken into consideration.

We setup an empirical entry model to account for the strategic nature of airlines' competition, following the seminal airline paper by Berry (1992). Here, firms play a game of strategic interaction and decide in which markets/routes to operate taking into account their competitors' decisions. Market structure data describing which routes the different airlines operate in is enough to extract relevant information about firms' profits; this is appealing because, in our setting, price and quantity data are not available.

We depart from Berry (1992), who uses a static model of entry with independent markets, in that we account for both the dynamic nature of airlines' competition and for the network structure of routes. We partly follow Aguirregabiria and Ho (2012), who were the first to propose a dynamic game of network competition and to estimate it with US airline industry data.

Different from Aguirregabiria and Ho (2012), who model airlines' decisions in discrete time, we adopt a continuous time framework. Doraszelski and Judd (2012) first showed the computational advantages of casting dynamic games in continuous time. ABBE further developed a continuous time framework for dynamic discrete choice models. In our context, a continuous time framework a la ABBE allows us to deal with the lack of regularity in airlines' decision making and to take advantage of our high frequency data: routes are not added/dropped at specific time intervals and can be frequently and relatively flexibly added and dropped by airlines. Also, this methodology makes it possible for us to conduct counterfactuals despite the high-dimensionality of our dynamic game.[34]

6.2 Setup

There are N airline companies, indexed by $i = 1, ..., N$, in the industry that provide flights between different cities. A route $r \in \{1, ..., R\}$ is defined as a non-directional city-pair, with R denoting the total number of possible routes between all of the different cities with airports in the data. Time is continuous and indexed by t. A stochastic independent Poisson arrival process governs when and which airline can move. With each move arrival, an airline may make decisions regarding the configuration of its network. In the context of airline entry and competition, a random move arrival process might reflect the stochastic timing of negotiations with airports for gate occupancy, delays in the processes for obtaining permits, and firms' staff recruiting and flight-capacity arrangements, for example.

There is also a Poisson process that controls changes in route characteristics exogenous to airline's decisions (e.g., HST presence and GDP), which we discuss in more detail below. We refer to such changes as moves by nature.

Let $x_{irt} \in \{0, 1\}$ be an indicator variable in which $x_{irt} = 1$ if airline i provides direct flights in route r at time t and $x_{irt} = 0$ otherwise.[35] The network of an airline i at time t is therefore defined by the collection of x_{irt} for all routes, i.e. $\mathcal{N}_{it} \equiv \{x_{irt} : r = 1, ..., R\}$. The vector $\mathcal{N}_t \equiv \{\mathcal{N}_{it} : i = 1, ..., N\}$ denotes the network for the entire industry at time t.

The flow payoff that airline i gets from route r at time t, denoted as $\Pi_{irt}(\mathcal{N}_t, \mathcal{Z}_t)$, is a function of both the network of the industry $\mathcal{N}_t$ and of the exogenous characteristics of

[34] A key advantage of a continuous time framework is that it reduces the computational burden associated with solving a dynamic game. In a continuous time framework, the probability of two events happening together is close to zero. This implies that players make decisions sequentially, as opposed to simultaneously. As a result, for any given state, the number of possible state changes in the next period increases only linearly with the number of players, rather than exponentially as is the case in discrete time frameworks. A smaller number of possible future states implies, in turn, that the transition matrix among states is sparse, which reduces the computational burden of solving for the equilibrium. For additional benefits of using a continuous-time framework in a discrete-choice setting please see ABBE (section 6.4).

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

the routes that affect the airline's demand and costs (i.e., nature), given by the vector $\mathcal{Z}_t$. The vector $\mathcal{Z}_t \equiv \{z_{rt} : r = 1, ..., R\}$ is a vector that includes the information on the variables exogenous to airlines' decisions for all routes at a given time t. Note that $\mathcal{Z}_t$ includes information on the HST network. HST presence is an exogenous variable because, consistent with what we discussed in Section 3.2, unlike in the airline industry, where companies can modify their routes relatively frequently, the expansion of the HST route network is, for the most part, predetermined.

When it is airline i's turn to move, the airline decides whether to make changes to its existing network, $\mathcal{N}_{it}$. Denote $j \in \mathcal{A} = \{-1, 0, 1\}$ as the action airline i can take in a given route r. Specifically, for each route, airline i can choose to do nothing ($j = 0$), to provide direct flights if the route is not being served (that is, to enter; $j = 1$), or to stop providing direct flights if the route is being served (that is, to exit; $j = -1$). Note that not all of the actions in $\mathcal{A}$ are available at each time; for example, if the airline is not operating in a given route, the only actions available are $j = 0$ and $j = 1$. Each action j is associated with an instantaneous payoff given by $\psi_{ijrt} + \epsilon_{ijrt}$, where ψ_{ijrt} is the mean cost associated with action j at time t for airline i in route r, with $|\psi_{ijrt}| < \infty$, and ϵ_{ijrt} is a private-information payoff shock, which is assumed to follow a type I extreme value distribution and is $i.i.d.$ distributed across airlines, actions, routes, and time.

Note that we do not model airlines' decisions on how many flights to operate in a given route as this would be computationally intractable. This means that we will not be able to make claims regarding how the presence of the HST affects flight intensity in different routes. However, studying the effect of the HST on flight intensity tends to be of second order when compared to the effect of the HST on airlines' network configurations. In fact, conditional on entry, in about 60% of the routes, airlines only operate one flight (not tabulated). Therefore, airlines' flight frequency decisions should be well captured by their decision to operate or not in a given route.

The forward-looking strategic behavior of the firms, together with the network considerations that are intrinsic to airlines, makes it infeasible to estimate the game described above without further assumptions given the size of the state space $(\mathcal{N}_t, \mathcal{Z}_t)$. Considering the possible network configurations captured in $\mathcal{N}_t$ alone, and because at each point in time an airline has the option to operate or not in each route, and given that there are 4 airlines and 2,278 routes, would result in a state space of size $2^{4\times2{,}278}$, which is computationally intractable.

To reduce the dimensionality of the problem, we adopt two assumptions previously used by Aguirregabiria and Ho (2012). First, we assume that each airline makes decisions regarding its network in a decentralized manner. That is, an airline makes decisions route by route, using local managers, who take the situation in other routes as given. This means that, instead of making decisions for all routes simultaneously, an airline only considers one route at a time.[36] In the continuous time setting, this implies that, with each Poisson move arrival, an airline may decide whether to serve a route or not.

Second, to further reduce the dimensionality of the state space, we adopt a sufficient statistics assumption in the same spirit of Aguirregabiria and Ho's that specifies that the information contained in the state variables $(\mathcal{N}_t, \mathcal{Z}_t)$ can be aggregated into a vector w_{irt} with a much smaller set of possible values. This implies that the strategy function of a local manager for a given route depends on the state variables $(\mathcal{N}_t, \mathcal{Z}_t)$ through a vector w_{irt}.[37] For example, instead of considering the entire network configuration $\mathcal{N}_t$, the local manager may consider solely the state of each individual competitor in the focal route together with the number of connections that each firm has between the focal route and other routes.[38] In a similar manner, the set of exogenous variables $\mathcal{Z}_t$ also gets summarized into a smaller set of variables. Section 6.4, where we discuss the empirical specification of the model, presents the detailed list of variables contained in w_{irt}.

Imposing the two assumptions above does not imply that airline decisions in different routes/markets are independent from each other, as is commonly assumed in strategic entry models (e.g., Berry 1992, ABBE). That is because there are variables in the vector w_{irt} that affect the decision of each local manager and that, as discussed in Aguirregabiria and Ho (2012), depend on decisions previously made by other local managers.

6.3 Value Functions and Equilibrium

Airlines are forward-looking and discount future payoffs at rate $\rho \in (0, \infty)$. For simplicity, and without loss of generality, we index the state of each route r at each instant t by an element k of some finite state space $\kappa = \{1, ..., K\}$. This means that the state of any route at each instant can be characterized by the vector w_k, presented in the previous section, which contains firm and route information. While in state k, airline i receives flow payoff Π_{ik}. Let σ_{ijk} denote the probability that airline i optimally chooses action j in state k. This action may result in a deterministic state change. Let $l(i, j, k)$ denote the continuation state that

[36]Similar assumptions have been made in other studies to deal with high-dimensional problems (e.g., Schiraldi, Smith, and Takahashi 2012, Sweeting 2013, He, Whited, and Guo 2021).

[37]Several other papers in addition to Aguirregabiria and Ho (2012) use similar assumptions to address the state-space dimensionality problem (e.g., Gowrisankaran and Rysman 2012, Weintraub, Benkard, and Van Roy 2006).

[38]Here we depart slightly from Aguirregabiria and Ho (2012) in that we allow for the state of each individual competitor to be considered by each player (as opposed to summary statistics on competition indicators, such as the total number of competitors present in the route, for example, as in Aguirregabiria and Ho 2012).

arises after airline i makes decision j in state k. We define $j = 0$ to be a costless continuation choice with instantaneous cost $\psi_{i0k} = 0$ and $l(i, 0, k) = k$, for all i and k.

We assume a common rate parameter, λ, for the Poisson arrival processes that govern when and which airline can move. The state of a given route can change not just due to the action of each of the local managers (if a local manager decides to pursue an action $j \neq 0$) but also due to other variables in w_k driven by nature's movements or the actions of local managers in other routes. Accordingly, in each state, there are, in addition to N independent, competing Poisson processes with rate λ, which generate move arrivals for each of the N players, a Poisson process with rate λ that governs movements by nature. We assume that moves by nature can be captured by a finite-state Markov jump process on κ by some $K \times K$ intensity matrix Q. Each element of Q, denoted by q_{kl}, is the hazard rate for transition from state k to state l and is nonnegative and bounded.[39] Further, and because decisions made by local managers in other routes can affect a given route's state, we allow for a competing Poisson process with some arrival rate λ' to generate state changes caused by decisions made in other routes. This arrival rate λ' is a function of the rates λ that govern the move arrivals for each of the players in other routes and of the mapping between those moves and the summary statistics considered by the local manager.

Let ζ_i denote the beliefs of airline i's local manager regarding the actions of rival airlines in a particular route, given by a collection of $(N-1) \times J \times K$ probabilities ζ_{imjk} for each player $m \neq i$, state k and choice j. Further, let ζ'_{ikl} denote the beliefs regarding the evolution of the summary statistics in the vector w_k which pertain to the decisions made by other local managers of airline i and by the local manager's competitors' in other routes.[40] Finally, let V_{ik} denote the expected present value for airline i being in state k and behaving optimally at all points in the future given beliefs ζ_i and ζ'_i. For small increments h, under the Poisson assumptions, the probability of an event with rate λ occurring is λh. Given the discount rate ρ, the discount factor for such increments is $1/(1+\rho h)$. Thus, for small time increments

[39] Note that the elements of Q which correspond to state changes that are not driven by nature's movements take the value zero.

[40] We use the notation ζ' to highlight that, in a given route, the beliefs regarding the summary statistics based on the actions of the players in connecting routes depend on the ζs in those routes. The term ζ' does not have subscript mj (as opposed to ζ) because it captures the expectation regarding the evolution of summary statistics which do not depend on the specific actions (or expectations regarding such actions) of individual players in other routes. Note that, even though, for each player i, there is a matrix K by K with elements ζ'_{kl}, $l = 1...K$, the elements in this matrix which correspond to state changes that are not driven by movements in the decisions made by players in other routes take the value zero.

h, the present discounted value of airline i being in state k in a particular route is given by

$$V_{ik} = \frac{1}{1+\rho h}\Bigg[\Pi_{ik}h + \sum_{l\neq k} q_{kl}hV_{il} + \sum_{m\neq i}\lambda h \sum_{j\in\{-1,1\}} \zeta_{imjk}V_{i,l(m,j,k)} + \lambda' h \sum_{l\neq k}\zeta'_{ikl}V_{il}$$
$$+ \lambda h\, \mathrm{E}\max_j \left[\psi_{ijk} + \epsilon_{ijk} + V_{i,l(i,j,k)}\right] + \Big(1 - \sum_{l\neq k} q_{kl}h - N\lambda h - \lambda' h\Big)V_{ik}\Bigg].$$

In the expression above, there are six terms within the square brackets. The first term refers to the flow profits that the local manager of airline i gets from being in state k, and the remaining terms correspond to the expected discounted payoffs associated with all possible states that may occur with a time increment h. More specifically, the second term corresponds to rate-weighted values due to state changes caused by nature's moves, and the third and fourth terms correspond to rate-weighted values associated with states that occur due the focal actions of the local manager's competitors and of the players in other routes, respectively. The fifth term corresponds to the instantaneous payoffs obtained when airline i's local manager makes a move in state k and the expected future value associated with the state change caused by that move. The expectation in this term is with respect to the joint distribution of ϵ_{ik}. Finally, the sixth term corresponds to the value associated with a route remaining in state k after a increment of time h.

Rearranging and letting $h \to 0$, V_{ik} can be written as

$$V_{ik} = \frac{\Pi_{ik} + \sum_{l\neq k} q_{kl}V_{il} + \sum_{m\neq i}^{N}\lambda \sum_{j\in\{-1,1\}}\zeta_{imjk}V_{i,l(m,j,k)} + \lambda' \sum_{l\neq k}\zeta'_{ikl}V_{i,l} + \mathrm{E}\max_j\left[\psi_{ijk} + \epsilon_{ijk} + V_{i,l(i,j,k)}\right]}{\rho + \sum_{l\neq k} q_{kl} + \sum_{m\neq i}^{N}\lambda\sum_{j\in\{-1,1\}}\zeta_{imjk} + \lambda + \lambda'\sum_{l\neq k}\zeta'_{ikl}}. \quad (2)$$

We focus on Markov perfect equilibria in pure strategies, as is standard in the empirical literature on dynamic entry games. A Markov strategy δ_i for player i is a mapping which assigns an action from $\mathcal{A}$ to each state (k, ϵ_i). Given beliefs for each player $\{\zeta_i : i = 1, ...N\}$, and a collection of model primitives, a Markov strategy for firm i is a best response if

$$\delta_i(k, \epsilon_i; \zeta_i) = j \iff \psi_{ijk} + \epsilon_{ij} + V_{i,l(i,j,k)}(\zeta_i) \geq \psi_{ij'k} + \epsilon_{ij'} + V_{i,l(i,j',k)}(\zeta_i) \quad \forall j' \in \mathcal{A}. \quad (3)$$

Given the distribution of choice-specific shocks, each Markov strategy δ_i implies the following response probabilities for each choice in each state

$$\sigma_{ijk} = Pr[\delta_i(k, \epsilon_i; \zeta_i) = j|k]. \quad (4)$$

A Markov perfect equilibrium is thus defined as a collection of stationary policy rules

$\{\delta_i : i = 1, ..., N\}$ and beliefs $\{\zeta_i : i = 1, ...N\}$ in which (3) holds for all i, k, and ϵ_i and the beliefs $\zeta_{mijk}, m \neq i$, are consistent with the best response probabilities generated by (4).

Note that here the beliefs of each local manager and its competitors are intrinsically linked to the beliefs of the local managers in other routes through the expectations that each local manager has to form regarding the summary statistics that govern the state of the route for which the manager is responsible. This means that the above defined equilibrium is an equilibrium at the network level which contrasts with other literature that regards markets as being independent from each other.

6.4 Empirical Specification and Implementation

6.4.1 State Variables and Action Variables

We focus on the top four airlines ($N = 4$) and the top 68 cities in China. So, in total, there are $R = 68 \times 67/2 = 2,278$ possible routes.

The state of a route at each instant can be summarized by a vector w_k containing information on two sets of variables. The first set relates to variables on the airlines' network configurations which include the airlines that operate in that route, $x_k = \{x_{ik} \in \{0, 1\} : i = 1, ...4\}$ and the number of connecting routes for each airline $x_k^c = \{x_{ik}^c \in \mathbb{N} : i = 1, ...4\}$.[41] The second set of variables relates to the exogenous characteristics of a given route, namely indicator variables for the presence of fast trains ($Fast_k$) and bullet trains ($Bullet_k$), the number of fast train lines and bullet train lines connecting to either of the route's endpoint cities ($Fast_k^c$ and $Bullet_k^c$, respectively), the length of the route, $length_k$, the average GDP for the route's endpoint cities gdp_k, the average growth-rate of the GDP in the endpoint cities $gdpgrowth_k$ whether a route is regulated Reg_k, and whether each airline is exempt from regulation on

[41] A connecting route is a route which shares one of the endpoint cities with the focal route. The total number of connecting routes of a given route is given by the sum of the routes that connect to each of the focal route's endpoints. For example, if a route connects city A and city B, and in state k, airline i offers direct flights in X routes that are connected to city A and in Y routes that are connected to city B, then x_i^c is equal to $X + Y$.

that route Exe_{ik}.[42] The vector w_k is then given by

$$\begin{aligned} w_k = &(x_{1k}, ..., x_{4k}, x^c_{1k}, ..., x^c_{4k}, \\ &Fast_k, Bullet_k, Fast^c_k, Bullet^c_k, length_k, gdp_k, gdpgrowth_k, Reg_k, Exe_{1k}, ..., Exe_{4k}). \end{aligned}$$

So, each value of k represents an encoded state vector, and the function $l(i, j, k)$ gives the continuation state that arises after airline i takes action j in state k. In addition, each route is characterized by a time-invariant unobserved route type s, which is observed by airlines but not by the econometrician.[43] It follows then that the full state vector at any instant can be written as (w_k, s).

6.4.2 Flow Profits and Choice-Specific Payoffs

We follow standard convention in the empirical entry literature and specify the flow payoff in terms of underlying latent profits because demand side data such as prices and seat occupancy is not available. The flow payoff of a local manager of airline i is thus specified as a linear function of the summary statistics that characterize an airline's and its competitors' route networks as well as of exogenous variables such as the city-pair average GDP, the length of the route and the summary statistics that characterize the HST network. The flow payoff also depends on the unobserved route-type s, which captures the unobserved tastes of consumers for different modes of transportation in a given route.[44]

[42] To reduce the dimension of the state space we discretize the number of airline connections into five bins ($[0, 5]$, $[6, 15]$, $[16, 25]$, $[26, 35]$, $[36, \infty)$). Similarly, the number of HST connections are discretized into three bins ($[0]$, $[1, 3]$ and $[4, \infty)$). The discretization was done such that there was a reasonable number observations in each category. To characterize route's length, we use three indicator variables to denote short routes ($\leq$ 600 km), medium routes (between 600 km and $1,200$ km) and long routes ($>1,200$ km). This allows airline payoffs to change non-linearly and non-monotonically with route length. Finally, we discretize the average city-pair GDP into five quantile-based bins and the average city-pair GDP growth-rate into three quantile-based bins (corresponding to low, medium and high growth).

[43] We use $S = 5$ points of support for the route's unobserved type, $s \in \{-1.3998, -0.5319, 0.0, 0.5319, 1.3998\}$, based on a discrete approximation to a standard normal random variable.

[44] Note that there are no airline specific intercepts in the payoff function. This is driven by computational reasons (especially when implementing the counterfactual simulations) and by the fact that the purpose of the paper is not to determine the specific effect of the HST entry on each of the airlines but rather on the entire industry. This is equivalent to making a symmetry assumption which is commonly used in the literature (e.g., Doraszelski and Judd 2012 and Ryan 2012). Nonetheless, we relax the symmetry assumption in a robustness check in which we allow for airline-specific intercepts. See Appendix B. The estimated parameters when this assumption is relaxed are consistent with the estimates from the base model. To illustrate what the symmetry assumption implies, suppose there are 5 firms, and each firm has three states, represented by numbers 1, 2 and 3, respectively. We use a 5×1 vector to denote the industry state, of which the i's cell corresponds to the state of firm i. The symmetry assumption implies that the strategy of firm 1 in state $(3, 1, 1, 2, 1)$, will be the same as that of firm 3 in the state $(1, 1, 3, 2, 1)$, because firms share the same policy function. This assumption also implies that the strategy followed by firm 1 in state $(1, 1, 3, 2, 1)$ will be the same as the one in state $(1, 3, 2, 1, 1)$.

The flow payoff to airline i in state (k, s) can be written as

$$
\begin{aligned}
u_{ik} = {} & \beta_0 + \beta_1 \sum_{m \neq i} x_{mk} + \beta_2 x^c_{ik} + \beta_3 \sum_{m \neq i} x^c_{mk} + \beta_4 Fast_k + \beta_5 Bullet_k \\
& + \beta_6 Fast^c_k + \beta_7 Bullet^c_k + \beta_8 Fast_k \times Fast^c_k + \beta_9 Bullet_k \times Bullet^c_k + \beta_{10} gdp_k + \beta_{11} length_k \\
& + \beta_{12} Fast_k \times length_k + \beta_{13} Bullet_k \times length_k + \beta_{14} s.
\end{aligned}
\tag{5}
$$

The choices regarding the variables included in the flow payoff and its functional form follow mostly from the empirical analysis conducted in Section 5. Namely, we allow airlines' payoffs in a route to depend on the number of (own and competitors') routes that connect to that route. Further, to better quantify the impact of the HST on airlines' decisions, we allow for heterogeneous effects of different types of HST (i.e., Fast and Bullet trains), and we interact the presence of the HST with the number of connecting HST lines to capture the fact that we expect the benefits that airlines derive from serving routes connected to the HST to be negatively impacted if those routes are also served by the HST. Finally, we also interact the presence of the HST with route length to allow for the effects of the HST to be different depending on the route length. The flow payoff from not operating flights in a given route is normalized to 0.

Airlines pay a sunk cost to enter a route. This cost depends on the unobserved route type, s, on whether the route is a government-regulated air route, and on airline-specific characteristics, such as the number of routes connected to the focal route, and on whether the airline is exempt from route regulation. Airlines' entry costs may be higher in regulated routes because the application to enter a regulated route has to fulfill several additional requirements (when compared to a non-regulated route) and is thus costly. Also, when entering a route, airlines which operate in connecting routes, tend to obtain more advantageous entry terms (such as lower gate fees, for example). As in Aguirregabiria and Ho (2012), the value associated with exiting a route is assumed to be zero. Therefore, the choice-specific instantaneous payoffs ψ_{ijk} can be written as

$$
\psi_{ijk} = \begin{cases} \eta_0 + \eta_1 \times s + \eta_2 \times Reg_k + \eta_3 \times Exe_{ik} + \eta_4 \times x^c_{ik} & \text{if } j = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{6}
$$

where Reg is an indicator variable which equals one if the route is regulated and zero otherwise, and Exe_i is an indicator variable which equals one if the route is regulated but airline i is exempt from the regulation, and zero otherwise.[45]

[45] Similar to what we assume in the flow payoff function, here there are no airline-specific intercepts (per

The structural parameters of interest to be estimated are the coefficients of the flow payoff function together with the parameters of the instantaneous payoffs

$$\theta = (\beta_0, ..., \beta_{14}, \eta_0, ..., \eta_4).$$

6.5 Model Estimation and Identification

We follow ABBE and estimate the structural parameters using a conditional choice probability (CCP) based approach with continuous-time data. This approach proceeds in two steps. In the first step, we estimate the reduced-form hazards that capture the dynamics in entry and exit decisions for airlines in each route, as well as those related with the moves of the summary statistics, which capture the players' decisions in other routes, and with the moves of the exogenous nature variables such as GDP and presence of HST. At the same time we also obtain estimates for the unobserved route types. In the second step, we estimate the structural parameters taking the reduced-form hazards as given.

6.5.1 Step 1: Estimating the Reduced-form Hazards

We estimate the probabilities of entry (if the airline is not present), exit (if the airline is present) and doing nothing for an airline in a route using a multinomial logit sieve. Let $\tilde{\sigma}_{ij}(k, s, \alpha)$ denote the reduced form probability of airline i making choice j in state (k, s), where α is a vector with the parameters to be estimated. The probabilities $\tilde{\sigma}_{ij}(k, s, \alpha)$ take the following form

$$\tilde{\sigma}_{ij}(k, s, \alpha) = \frac{exp(\phi_j(k, s, \alpha))}{\sum_{j' \in \mathcal{A}} exp(\phi_{j'}(k, s, \alpha))}, \tag{7}$$

where $\phi_j(k, s, \alpha)$ is a flexible linear function of the state variables. The variables included in this function are a constant, the total number of competitors and its square, the number of own connecting air routes and its square, the total number of competitors' connecting routes and its square, the presence of fast trains, the presence of bullet trains and the interaction of these indicator variables with route length indicator variables, the number of connecting fast train lines interacted with the presence of fast trains, and the number of connecting bullet train lines interacted with the presence of bullet trains. We also include indicator variables for route length, the average city-pair GDP, the GDP growth-rate level (low, medium and high), and the unobserved route type. In addition, we control for the

the symmetry assumption discussed above). Nonetheless, we relax the symmetry assumption in a robustness check in which we allow for entry costs to be airline-specific. The estimated parameters are consistent with the estimates from the base model (results are available upon request from the authors).

following variables interacted with an indicator for route entry: constant, whether a route is subject to regulation, whether an airline is exempt from regulation in the case of a regulated route, the number of own connecting air routes, and the unobserved route type.[46]

The probabilities of movement of each of the exogenous variables caused by "nature" are modeled as follows. The probability of each of the HST (fast and bullet trains) changing from not present to present in a given route is estimated using two binary logit models (one for fast and one for bullet trains) with independent variables: constant, number of fast and bullet trains connections, average city-pair GDP, indicator variables for route length, and the GDP growth-rate level, and the unobserved route type. The probability of the number of HST (fast and bullet trains) connections increasing or decreasing in a given route is estimated using two binary logit models (one for fast and one for bullet trains) with independent variables: constant, indicator variables for whether fast or bullet trains are present, average city-pair GDP and GDP growth-rate level. These specifications for the probabilities of the HST being present and the HST's number of connections capture the uncertainty around the planning of the HST that we discussed in section 3.2. The transition probabilities of the average city-pair GDP (going up or down) are estimated using the frequency of city-pair GDP transitions between 2006 and 2016 for the markets in the sample. We use $\tilde{q}_l(k, s, \alpha_0)$ to denote the empirical estimate of the hazard rate for nature's transition from state k to state l with $l \neq k$ in a route with unobserved type s.[47]

We focus on each airline's total number of routes connected to a given route to capture the actions of the players in other routes in a summarized way, and estimate the transition probabilities of the number of connecting routes (going up or down) based on their empirical distribution. We use $\sigma'_l(k)$ to denote the transition probability of the total number of connecting routes when it transitions from state k to l. The empirical estimate of such transition probability is defined as $\tilde{\sigma}'_l(k)$.

The complete set of hazards estimated in Step 1 is given by

$$h(\alpha_0, \alpha) = (\tilde{q}_l(1, 1, \alpha_0), ..., \tilde{q}_l(K, S, \alpha_0), \lambda\tilde{\sigma}_{ij}(1, 1, \alpha), ..., \lambda\tilde{\sigma}_{Nj}(K, S, \alpha), \lambda'\tilde{\sigma}'_l(1), ..., \lambda'\tilde{\sigma}'_l(K)),$$
$$\forall l \neq k, \forall j \in \mathcal{A}.$$

In each route, the movement of airlines, nature and the summary statistics of the con-

[46]In Step 1, we jointly estimate the policy function and the probability of a route being of a specific unobserved type.

[47]Note that we proceed with a slight change of notation to accommodate for the time-invariant unobserved route types in our empirical specification. More specifically, q_{kl} in Equation (2) is replaced with $q_l(k, s)$ to denote the hazard rate for transition from state k to state l in a route with unobserved type s. Likewise, σ_{ijk} is replaced by $\sigma_{ij}(k, s)$ to denote the probability that airline i optimally chooses action j in state k in a route with unobserved type s.

necting routes can be captured by a joint Poisson process.[48] Therefore, in state k, conditional on the market being of unobserved type s, the probability of the next state change within τ units of time follows the CDF of an exponential distribution with rate parameter equal to the sum of the state transition rates for nature, the hazards of the non-continuation actions for each player, and the sum of the state transition rates for the summary statistics:

$$1 - exp\left[-\tau\left(\sum_{l\neq k} q_l(k,s) + \sum_i \lambda \sum_{j\in\{-1,1\}} \sigma_{ij}(k,s) + \lambda' \sum_{l\neq k} \sigma_l'(k)\right)\right]. \tag{8}$$

Differentiating with respect to τ gives the density of the time of the next state change:

$$\left(\sum_{l\neq k} q_l(k,s) + \sum_i \lambda \sum_{j\in\{-1,1\}} \sigma_{ij}(k,s) + \lambda' \sum_{l\neq k} \sigma_l'(k)\right) exp\left[-\tau\left(\sum_{l\neq k} q_l(k,s) + \sum_i \lambda \sum_{j\in\{-1,1\}} \sigma_{ij}(k,s) + \lambda' \sum_{l\neq k} \sigma_l'(k)\right)\right]. \tag{9}$$

Conditional on the state change, the probability that player i takes action j is given by

$$\frac{\lambda\sigma_{ij}(k,s)}{\sum_{l\neq k} q_l(k,s) + \sum_i \lambda \sum_{j\in\{-1,1\}} \sigma_{ij}(k,s) + \lambda' \sum_{l\neq k} \sigma_l'(k)}. \tag{10}$$

Taken together, the likelihood of the next stage change occurring after an interval of length τ while being the result of player i taking action j is the product of the previous two equations:

$$\lambda\sigma_{ij}(k,s) exp\left[-\tau\left(\sum_{l\neq k} q_l(k,s) + \sum_i \lambda \sum_{j\in\{-1,1\}} \sigma_{ij}(k,s) + \lambda' \sum_{l\neq k} \sigma_l'(k)\right)\right]. \tag{11}$$

We can similarly construct the likelihood of nature and the summary statistics going from state k to state l as

$$q_l(k,s) exp\left[-\tau\left(\sum_{l\neq k} q_l(k,s) + \sum_i \lambda \sum_{j\in\{-1,1\}} \sigma_{ij}(k,s) + \lambda' \sum_{l\neq k} \sigma_l'(k)\right)\right], \tag{12}$$

and

$$\lambda'\sigma_l'(k) exp\left[-\tau\left(\sum_{l\neq k} q_l(k,s) + \sum_i \lambda \sum_{j\in\{-1,1\}} \sigma_{ij}(k,s) + \lambda' \sum_{l\neq k} \sigma_l'(k)\right)\right], \tag{13}$$

respectively.

[48]The summary statistics reflect the decisions of airlines in connecting routes, therefore they also follow Poisson processes.

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

Denote $I_{rn}(i, j)$ as the indicator variable which equals one if player i takes action j in the nth event at route r. Further denote $I^q_{rn}(k, l)$ and $I'_{rn}(k, l)$ as the indicator variables which equal one if the state changed from k to l $(k \neq l)$ during the nth event at route r due to nature or due to changes in the summary statistics which capture the actions in connected routes, respectively. Now, conditional on a route being of unobserved type s, the likelihood for the single event n in route r is given by:

$$\tilde{L}_{rn}(h(\alpha_0, \alpha); s) =$$
$$\left(\sum_{l \neq k_{rn}} I^q_{rn}(k_{rn}, l)\tilde{q}_l(k_{rn}, s, \alpha_0) + \sum_i \lambda \sum_{j \in \{-1,1\}} I_{rn}(i, j)\tilde{\sigma}_{ij}(k_{rn}, s, \alpha) + \lambda' \sum_{l \neq k_{rn}} I'_{rn}(k_{rn}, l)\tilde{\sigma}'_l(k_{rn}) \right)$$
$$\times \, exp \left[- \left(\sum_{l \neq k_{rn}} \tilde{q}_l(k_{rn}, s, \alpha_) + \sum_i \lambda \sum_{j \neq 0} \tilde{\sigma}_{ijr}(k_{rn}, s, \alpha) + \lambda' \sum_{l \neq k_{rn}} \tilde{\sigma}'_l(k_{rn}) \right) \tau_{rn} \right]. \tag{14}$$

Following ABBE, we control for the unobserved route type using mixture distributions. We 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. We specify this probability as an ordered probit which we estimate using the observed value of the following variables at the beginning of the period studied: total number of flights, total number of connecting routes, average GDP growth rate of the city-pair, length of the route, and an indicator variable that captures whether the route is regulated.

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

$$(\alpha^*, \pi^*) = \arg\max_{\alpha, \pi} \sum_{r=1}^{R} ln \left(\sum_s \pi(s, k_{r0}) \prod_{n=1}^{T_r} \tilde{L}_{rn}(h(\alpha); s) \right). \tag{15}$$

6.5.2 Step 2: Estimating the Structural Parameters

In Step 2 we take the estimated hazards $h(\tilde{\alpha}_0, \tilde{\alpha})$ and the probabilities of the routes being in each of the unobserved states from Step 1 as given and use these to estimate the structural

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

parameters which underlie the flow payoffs u and the instantaneous payoffs ψ. To do so, we first approximate the value function in a closed-form inversion step. This approximate value function is then used in a likelihood criterion function to estimate the structural parameters.

We use Proposition 6 in ABBE to express the value function represented in equation (2) as a function of the structural parameters and the reduced form hazards $h(\tilde{\alpha}_0, \tilde{\alpha})$ from the first stage. This allows us to eliminate the value functions on the right-hand side of equation (2), such that no fixed point problem needs to be solved in the estimation. Applying this proposition, the empirical counterpart of the value function in (2) can then be re-written as

$$\tilde{V}_i(\theta; s) = \left[(\rho + N\lambda + \lambda')I - \sum_{m=1}^{N} \lambda \Sigma_m(\tilde{\sigma}_m(s)) - \lambda' \Sigma'(\tilde{\sigma}') - \tilde{Q}(s)\right]^{-1} [u_i(\theta) + \lambda E_i(\theta)], \quad (16)$$

where $\tilde{\sigma}_i(s)$ is defined as a set that collects $\tilde{\sigma}_{ij}(k, s, \alpha)$ for all $j \in \{-1, 0, 1\}$ and $k \in \kappa$, and $\Sigma_i(\tilde{\sigma}_i(s))$ denotes the $K \times K$ state transition matrix induced by the actions of player i given the estimated choice probabilities in Step 1. Similarly, $\Sigma'(\tilde{\sigma}')$ and $\tilde{Q}(s)$ represent the empirical transition matrices for the summary statistics and nature, respectively. I is a $K \times K$ identity matrix. $u_i(\theta)$ is a $K \times 1$ vector with the kth element being u_{ik}. Finally, $E_i(\theta)$ is a $K \times 1$ vector where each element k is the ex-ante expected value of the instantaneous payoff associated with choice j made in state k, $\sum_j \tilde{\sigma}_{ij}(k, s, \alpha)[\psi_{ijk} + e_{ij}]$, and e_{ij} is the expected value of ϵ_{ij} at state (k, s) given that choice j is optimal for player i.

Note that, to avoid inverting the $K \times K$ matrix in equation (16), ABBE use their Proposition 4 (instead of Proposition 6, as we are doing here). The finite dependence property of Proposition 4 allows ABBE to re-express the value function as a linear function of CCPs and structural parameters, which avoids both having to solve for a fixed-point and inverting (and storing) a $K \times K$ matrix, leading to significant computational gains. Proposition 4 is convenient but it can only be applied in models with a terminal choice (e.g., when firms exit the market in a permanent fashion). Because airlines often exit routes which they re-enter later on, applying ABBE's finite dependence representation would not be appropriate in our case. We thus do not take full advantage of the computational benefits associated with a continuous time representation which can be achieved when the finite-dependence property is applied. Nonetheless there are still several advantages to using a continuous time framework in our setting when approximating the value function. More specifically, the matrix inversion in equation (16) is significantly simplified due to its sparsity which comes from the continuous-time representation. This implies significant computational gains not only during the estimation but also when conducting policy simulations.

Because we cannot apply Proposition 4 as in ABBE, we are left with the challenge of inverting the matrix in equation (16). Due to the large state space in our model, this inversion cannot be done directly. Instead, we take advantage of the fact that the second part of the right-hand side of equation (16) can be re-written in matrix form and apply the GMRES (generalized minimal residuals) method (Saad and Schultz 1986).[50] More specifically, we rewrite the second part of the right-hand side of equation (16) as a linear function of the structural parameters such that $u_i(\theta) + \lambda E_i(\theta) \equiv Z_i\theta + e_i$, where Z_i is a $K \times O$ matrix, with O being the length of the vector θ, and e_i is a $K \times 1$ vector where the kth row corresponds to $\sum_j \tilde{\sigma}_{ij}(k, s, \alpha)e_{ij}$. We then obtain the products of the matrix in the first part of the right-hand side of equation (16) with each column of the matrix Z_i and the vector e_i. Due to the large state space, these products cannot be obtained using direct methods (e.g., Gaussian or Gauss-Jordan elimination); as an alternative we use the iterative method GMRES to calculate the resulting vectors. We set the number of iterations to 100 and allow for 20 re-starts of the algorithm. The tolerance level of the approximation is set to $1e^{-6}$ and all of our approximations converge at this level. After applying this procedure, we are able to approximate the value function as a linear function of the structural parameters, which are in turn used in the second step of the estimation.

Let $\tilde{V}_{ik}(\theta; s)$ be the approximated value function for player i at state k. With the assumption that the idiosyncratic error terms in equation (3) follow an i.i.d. type I extreme value distribution, the choice probabilities in equation (4) can be expressed as

$$\check{\sigma}_{ijk}(\theta; s) = \frac{exp\left(\tilde{V}_{i,l(j)}(\theta; s) + \psi_{ijk}(\theta; s)\right)}{\sum_{j' \in \{-1,0,1\}} exp\left(\tilde{V}_{i,l(j')}(\theta; s) + \psi_{ij'k}(\theta; s)\right)}.$$

Replacing $\tilde{\sigma}_{ijk}$ in $\tilde{L}_{rn}(h(\alpha); s)$ with $\check{\sigma}_{ijk}$ in expression (14), the new likelihood for the single event n in route r, which is now a function of the structural parameters, can be denoted as $\check{L}_{rn}(\theta; s)$. Also, let $\pi_r(s)$ be the likelihood of route r being of unobserved type s given the data. Using Bayes's rule, we have

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

[50] GMRES is essentially an iterative method that numerically approximates the solution to a system of linear equations. This method has been used in the chemistry and physics fields to solve extremely high-dimensional systems. Also, and as suggested in Rust (1996), this algorithm can be applied to solve dynamic models with large state spaces.

The second-step maximum-likelihood estimates therefore become

$$\theta^* = \arg\max_{\theta} \sum_{r=1}^{R} \sum_{s} \pi_r(s) \sum_{n=1}^{T_r} ln\check{L}_{rn}(\theta; s). \tag{17}$$

6.5.3 Identification

Here we briefly discuss the identification strategy for the structural parameters. Our strategy follows ABBE's closely and we refer readers to Section 5.3 of ABBE for more details. Specifically, for player i in each state k, given our assumption about ϵ, we can represent the differences in the choice-specific value functions in terms of the CCPs and nature's intensity matrix Q. That is, $V_{ik} - V_{i,l(ijk)-\psi_{ijk}} = ln(\sigma_{i0k}) - ln(\sigma_{ijk})$. Because we can further represent the value functions at each state using the specifications from equation (16), the LHS of the equation can also be expressed in terms of the CCPs and Q. We are therefore left with an equation with only payoffs that are unknown. Stacking the equations for player i across states and actions, we have $(J-1)K$ rows and JK unknown payoffs.

Following common practice in the existing literature on dynamic entry games, we impose two additional sets of restrictions to achieve identification: first, we explore the exchangeability in the flow payoff based on the symmetry assumption we made about the airlines (as discussed in section 6.4.2) so that we have $\Pi_{ik} = \Pi_{il}$ for some $l \neq k$ where l is the same as state k in every dimension (including the number of competitors) except for the identity of the competitors of player i. This allows us to add at least K linear restrictions. Second, we also impose exclusion restrictions in the instantaneous payoffs such that the entry cost of an airline does not depend on the states of other airlines. This again provides us with at least K additional restrictions.

7 Structural Model Results

This section presents the structural model estimates, including the parameters from the flow payoffs and from the instantaneous payoffs (i.e., the entry costs function). To provide an economic understanding of the relative importance of the HST presence given the estimated parameters, we conduct a decomposition analysis of the airlines' profits.

7.1 Parameter Estimates and Model Fit

Table 6 presents the structural parameter estimates. We first discuss the parameters which affect the airlines' flow payoffs. These include the strategic-effect parameters, the parameters related with the effect of the HST, and the parameters that are associated with the market characteristics.

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

Insert Table 6 about here

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

The strategic effects capture the impact of the airlines' route networks on the airlines' flow payoffs. The coefficient associated with the number of own connections is positive and significant indicating that, the more connections an airline has to a route, the more profitable the route is. The parameter associated with the number of competitors is negative and significant implying that the presence of competitors in a same route leads to business stealing. However, this negative impact is dampened by the existence of competitors' connections consistent with the fact that air routes from competitor carriers may feed traffic to each other.

Turning to the impact of the HST, there is a negative impact of the presence of the HST on airlines' flow payoffs for short routes. Further, the presence of the fast train has a larger (negative) effect on airlines' flow payoffs than the bullet train. This is expected given that, in shorter routes, the speed advantage of the bullet train (relative to the fast train) becomes less relevant, especially considering that fast train tickets are cheaper.[51]

Consistent with the results from the reduced form analysis in section 5, the negative effect of the HST on airlines' payoffs is mitigated as the length of the route increases. However, the effects are heterogeneous depending on the train type. While the negative spillovers from fast trains become almost negligible in medium routes, in the case of bullet trains, the spillovers do not seem to get attenuated in longer-distance routes. This is most likely driven by the fact that the bullet train is a better substitute to air travel than its slower counterpart, the fast train, in relatively longer routes.

Interestingly, and especially in longer routes, overlapping with the HST (fast train) seems to have a positive net effect on airlines' profits. This could be due to market expansion effects that overcome the HST competition effect, consistent with surveys that describe that, for many HST routes in China, more than half of the traffic is "generated" – that is, traffic made up of people doing trips that they would not have made before.[52] It thus seems reasonable

[51]Note that, when compared to the reduced-form analysis in section 5, in the structural model, we are able to more easily analyze the heterogeneous effects of the different types of trains (bullet vs. fast). This is because, distinguishing between the two types of train in the reduced-form analysis would require analyzing two types of treatment effects (instead of a single one) which would reduce the number of degrees of freedom significantly and make it more difficult to identify the effects.

[52]Source: `http://www.economist.com/news/china/21714383-and-theres-lot-more-come-it-waste-money-china-has-built-worlds-largest` (last accessed June 16, 2022).

to assume that this expansion in the travel market can also have positive implications for airlines.

Having more connections to HST lines makes a route more attractive to airlines, highlighting the role of connectivity for intermodal transportation complementarities. However, for both types of HST, the positive effects of HST connectivity on airlines' payoffs are moderated by the presence of the HST in the same route. Specifically, when the HST is present in a route, the overall impact of having connections changes from positive to negative. This seems plausible because the HST is more likely to feed traffic to airlines when the air routes can take passengers to destinations that are not easily reachable by HST. However, for connecting routes that are also accessible via HST, customers are less likely to switch from rail to air travel because that would be less convenient.

The market (i.e., route) characteristics also affect airlines' profits. On average, airlines have lower flow payoffs in longer routes, and enjoy higher profits in routes with large values of the route unobserved state. The coefficient for the average GDP of a route not significantly different from zero. This could be because we focus on the main cities in China for which the GDP is already relatively high when compared to the rest of the country, and also because the GDP is growing steadily across all cities which may make the GDP less of a factor in driving airlines' decisions when compared to other factors.

Finally, there are significant costs associated with entering a route, which vary across routes, depending on their characteristics. An airline enjoys lower entry costs in routes in which the airline has more connections, suggesting that there are economies of density from operating multiple routes which originate in the same airport. Further, the cost of entry is higher in routes that are regulated by the government. Interestingly the coefficient on "Exempt" is not significant. This could be because the cost-reduction benefits associated with being exempt from regulation are already picked up by the coefficient on the number of connecting routes: airlines that are exempt from regulation are typically those operating out of their hubs which means they have several connections that originate from their airport hubs. Lastly, the cost of entry also varies depending on the route unobserved type, the cost being lower for routes with higher values of the unobserved state.

Taken together, the structural parameter results are consistent with the descriptive evidence and suggest that there are significant network and intermodal transportation effects which drive airlines' decisions. Airlines prefer entering routes with more air-route connections and suffer from business stealing from other airlines that share the same routes with them. Also, there are both positive and negative spillovers from the HST to the airline industry. Airlines benefit from being connected to the HST network, but face competition when serving routes that are also served by the HST. The type of HST and the length of a route moderate the magnitude of these spillover effects. Further, entering a route entails significant costs for airlines.

7.1.1 Model Fit

We compare the predicted with the observed airline network configurations using two different metrics. The first metric tries to capture the ability of the model to predict airline route presence (i.e., whether each route is served by at least one airline), while the second metric assesses the quality of the model with respect to airline route density (i.e., how many airlines operate in each route). More specifically, for the first metric we calculate the correlation between the vector (of length equal to the number of routes) of the expected number of airlines (capped at one if the expected number is greater than one) predicted by the model and the observed vector of indicator variables which take the value one if there is at least one airline present in a given route. We repeat this for each year-end and report the correlations between the predicted and the observed route-level values. In what concerns the second metric, we compare the vectors (with length equal to the number of routes) of the observed and the expected number of airlines serving each of the different routes. The expected numbers used in both metrics are calculated based on predictions about route presence decisions obtained from $1,000$ simulations of each airline's policy functions using the estimated parameters from the structural estimation. Table 7 present the results. The model does a reasonably good job at predicting both the route presence and the number of airlines present with average correlations (across all years in the sample) of 83 and 92 percent for the first and second metrics, respectively.

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

Insert Table 7 about here

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

7.2 Interpretation: Profit Decomposition

Because the structural estimates include coefficients associated with the entry of the HST with opposite signs, to get a better understanding of the structural model results, we decompose the airline flow profits into their different sources. This allows us to better assess the relative importance of the estimated effects on airlines' profits and, particularly, the weight of the HST. To this end, we calculate the average flow profits, across all routes and airlines, for the year 2015 (the last year in our sample) and decompose them into the effects from own and competitor's network presence, the effect from the different types of HST (fast and bullet) and the effect of market (i.e., route) characteristics. Table 8 presents the results of this decomposition.

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

Insert Table 8 about here

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

Airlines' own networks have a first-order effect on average airline profits with a magnitude that outweighs any other negative or positive effects. The magnitude of the market characteristics' effects is on average negative which is consistent with the fact that most routes on their own (89% of routes, not tabulated), i.e., without considering an airline's network of connections (and the HST network neither), are not profitable. Regarding the effect of the HST, we aggregate the spillovers from fast and bullet trains into positive and negative spillovers based on the sign of the coefficients of the structural estimates. While the positive and negative spillovers from the fast train tend to compensate each other, in the case of the bullet train, the positive spillovers are, on average, almost twice as large as the negative spillovers. This is due to the fact that airlines are present in routes in which they benefit from the positive spillovers associated with the existence of bullet train connections.

Even though this decomposition analysis may facilitate the interpretation of the structural model estimates, it has an important limitation because it is *conditional on the current airline network configuration*. Based on the results of this analysis one may be tempted to conclude that, overall, the HST has a positive effect on the airline industry. However, to properly assess the effect of the HST on the airline industry one needs to compare the current airline network configuration with the HST (and the profits associated with it) with the airline network configuration that would result from a scenario in which the HST were not present and the airlines were allowed to re-optimize their decisions accordingly. Put another way, the profits that airlines would make in the absence of the HST cannot be calculated based on the current configuration after simply removing the HST effects. This is because, in the absence of the HST, the entry and exit decisions of airlines would most likely be different leading to a different network configuration. In the next section, we conduct counterfactual policy experiments which allow us to properly quantify the impact of the entry of the HST on the airline industry.

8 Policy Experiments

We perform counterfactual analyses using the structural model to quantify the effects of the entry of the HST on the airline industry. Using the model to simulate the endogenous equilibrium outcomes is necessary because the evolution of the airline network between 2007 and 2015 depends on the market environment.

First, we simulate the network configurations of the airline carriers in a scenario in which the HST were not present, and compare the endogenous airlines' route decisions and profits to the baseline case in which the HST is introduced. Because the HST has a heterogeneous effect on the airline industry, we also explore the sources and impact of such heterogeneity.

In another experiment, we simulate an increase in the positive spillover effects between the HST and the airline industry to explore the possible benefits from services that facilitate the complementarity between the two modes of transportation. This experiment is motivated by the government's goal to increase people's mobility through a better integration of air and rail travel, and allows us to assess how government efforts to improve the degree of connectivity between the two modes of transportation may compensate the negative effects of the HST on airline entry.

8.1 Implementation

Carrying out policy experiments requires solving for the market equilibria under different scenarios. This is a particularly complex task in our setting because markets (i.e., routes) are interconnected. Different from most traditional entry models, airline decisions in each route depend not only on other players' decisions in that route but also on the (own and others') decisions made in other routes. Thus, any counterfactuals involving changes in exogenous market characteristics or model parameters will affect each airline's expectations with respect to the evolution of the industry's airline network and consequently its decisions and the equilibrium of the game. This means that the equilibrium cannot be solved independently for each route, as in more traditional entry model settings.

To compute the new equilibrium outcomes in the counterfactual scenarios, we proceed in a manner similar to Aguirregabiria and Ho (2012) and use a forward simulation method to update the beliefs regarding the transition of the summary statistics which capture information on the connecting routes. The specific steps for solving for the equilibrium are described as follows:[53]

1. Start with some initial belief of the transition probabilities for the summary statistics, denoted by σ_0'.

[53] There is no guarantee that the equilibrium is unique. However, as discussed in ABBE, the continuous time framework helps to eliminate simultaneity as a likely source of multiplicity in the equilibrium. In addition, we solve for the equilibrium using different starting points for the value function as well as for the airlines' expectations in what concerns the evolution of the summary statistics, and find that the results always converge to the same equilibrium.

2. Given σ_0', obtain the policy functions for each route (unobserved) type.

3. Use the recovered policy functions to simulate the evolution of the entire route network over the entire time period.

4. With the simulation results from the previous step, update airlines' beliefs of the transition probabilities of the summary statistics, denoted by σ_1'.

5. If $|\sigma_1' - \sigma_0' > \epsilon|$, replace σ_0' with σ_1' and obtain the policy functions for each route's (unobserved) type and repeat from step 3. Here, ϵ is set to $1e^{-6}$.

Once the equilibrium in the counterfactual scenario is found, we use the airlines' 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 then average across the equilibrium outcome variables of interest (such as entry probabilities and flow profits). We then compare these averages with those simulated for the baseline scenario, in which the HST is introduced as observed and the parameters are the ones estimated in the structural model. Note that, in each simulation, we keep the initial airline network configurations the same as observed. Further, to ensure a fair comparison, we also solve for the equilibrium in the baseline scenario using the steps listed above.

8.2 The Impact of the HST on the Airline Industry

To study the endogenous network configurations of the airline carriers in a scenario in which the HST were not present, we start with the airline network configuration at the beginning of our sample in 2007 (i.e., before the introduction of the HST) and simulate the economy forward to the end of our sample in 2015, assuming that the HST were not introduced. We then compare the resulting airline network outcomes in 2015 with and without the HST. In addition, we investigate the heterogeneous impact of the HST across different areas of the country and different types of routes (short vs. long, and overlap vs. connect to HST lines).

Overall Impact of the HST on the Airline Industry

To assess the overall impact of the HST on the airline industry's route decisions we look at two measures: the total number of routes that are served by at least one airline, and the total number of airline-routes, defined as the sum across airlines of the total number of routes served by each airline. Table 9 presents the results. Column "No HST" corresponds to the airline network configuration if the HST were not present, while column "Baseline" corresponds to the baseline scenario in which the HST is present. The table shows that the presence of the HST is associated with a 8% lower number of routes served by the airline industry and with a 14% lower number of airline-routes.

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

Insert Table 9 about here

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

In addition, we compare the airline industry's profits in both scenarios. Flow profits are computed based on the network configuration at the end of year 2015 and reported in annual terms. Table 9 shows that the lower number of routes served by airlines when the HST is present is accompanied by 23% lower profits. In addition, the lower overall profit in the baseline case is driven by a smaller airline network (number of routes served) when compared to the scenario without HST, as well as with a lower average profit per route.

Taken together, the previous results indicate that, despite the existence of positive spillovers from the HST on the airline industry, there are stronger negative spillovers such that the net impact of the introduction of HST on the airline industry is negative.

The positive spillover effects from the HST on the airline industry are substantial, however. To quantify the impact of the positive spillovers from intermodal connectivity, we simulate another scenario in which the HST is introduced but the positive spillovers from connectivity are shut down. Specifically, we 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. Column "No positive spill. from HST" in Table 9 reports the results, which we compare with column "Baseline". The total number of routes served by the airline industry is 52% lower, the total airline route presence is 63% lower, and the flow profits are 52% lower without the positive spillovers. These results imply that the positive spillovers from intermodal connectivity help to significantly mitigate the negative spillover effects from the HST and hence prevent the airline network size from being significantly smaller when the HST is present.

Heterogeneity Patterns

The previous analysis describes the overall effect of the HST on the airline industry. The effects can vary across different markets, however. In what follows we describe the impact of the HST on the airline industry across different geographical areas (cities and regions) and across different types of routes (short vs. long routes, and overlap vs. connect to HST lines). We then explain the sources of this heterogeneity.

Figure 8 shows a map with the top 20 cities by passenger volume at the end of the year 2015. The black lines represent the HST routes, and each dot represents a city. The size of each dot is proportional to the airline-route presence in a given city (defined as the number of unique airline-route combinations that either start or end at that city) for the baseline scenario and the color of each dot shows how the airline-route presence in the scenario without HST compares with the scenario in which the HST is introduced (baseline). We use warm colors to denote an increase in airline route presence when the HST is introduced relatively to the scenario without HST, and cold colors to denote a decrease.

The introduction of the HST has affected the airline industry differently across China. Examining Figure 8 reveals that the cities that experience the largest decrease in airline route presence are those geographically centrally located (such as Wuhan, Zhengzhou and Nanjing). In contrast, cities in geographically peripheral areas are less affected, and some of these, such as Chengdu and Chongqing, even experience an increase in airline route presence as a result of the introduction of the HST.

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

Insert Figure 8 about here

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

Table 10 reports in more detail the impact of the HST on airline route presence, for all regions in China. Air routes in Eastern and Central parts of China are the most negatively affected by the introduction of the HST. On average, cities from these two regions experience a decrease of about 20 airline-routes when the HST is introduced. This contrasts with cities from the western and northeastern parts of China, where there is an increase in the number of airline routes in each city associated with the introduction of the HST.

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

Insert Table 10 about here

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

The cities that observed an increase in airline presence have significantly lower average income level and population density (not tabulated), which suggests that the airline industry's response to the HST has the potential to help reduce inequality across China. Studies have established a link between transportation infrastructure and development and poverty in China (e.g., Démurger 2001, Zou et al. 2008).[54] Further, the World Bank's *Regional Economic Impact Analysis of High Speed Rail in China* report (World Bank 2014) discusses how the HST can have an impact on less developed regions in China.[55] We find that there is a

[54]Li and DaCosta 2013 provide an overview of this literature.

[55]See also an article published in 2017 in The Economist which also discusses the benefits of the HST for the economy in general, and for the poorer cities in particular (`http://www.economist.com/news/china/21714383-and-theres-lot-more-come-it-waste-money-china-has-built-worlds-largest`) (last accessed June 16, 2022).

potential link between the HST and China's development through the reorganization of the airline networks in response to the entry of the HST: airlines avoid head-to-head competition with the HST and choose to cater to more underserved and undeveloped areas. These areas become profitable for airlines only due to the complementarities with the HST, thus leading to airline entry.

Turning to the analysis of the effects across different types of routes, we first classify routes into six groups depending on whether the HST is present in the route (present and not present) and on the length of the route (short, medium and long). We then calculate the average number of airlines in each group for the scenario without HST and compare these with the corresponding averages in the baseline scenario.

Table 11 displays the results. Consistent with the parameters from the structural estimation, when we compare the scenarios with and without HST, there is a significantly lower airline presence in the short and medium routes which overlap with the HST. The effect is especially pronounced for shorter routes which have, on average, 72% lower number of airlines when the HST is present compared to when it is not. Long routes, in contrast, tend to have a higher average number of airlines when the HST is present. Interestingly, the entry of the HST also has a negative effect (albeit smaller) on the average number of airlines present in short and medium routes even when these do not overlap with HST. This is driven by the fact that routes are interconnected so that, when an airline decides not to serve a route, it makes serving connecting routes less attractive.

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

Insert Table 11 about here

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

Taken together, the results show there is considerable heterogeneity across cities and route-types in how the HST impacted the airline industry. Airlines readjusted their network by exiting shorter routes and substituting them for longer routes and routes in more peripheral regions in China. Airline entry in these routes would not have occurred in the absence of the HST, which made them profitable for airlines. This highlights a potential indirect benefit of the HST entry in shifting airlines to more remote and underserved areas thus improving connectivity among regions and reducing inequality. Different from other papers that have studied the crowding out effect from public investment (e.g., Berry and Waldfogel 1999, Sinai and Waldfogel 2005), we find, in addition to crowding out, positive effects from public investment: network complementarities lead to increased market coverage transportation-wise.[56] Wilson (2021) also finds public competition to provide a significant

[56]Note that, in addition to the expansion in airline routes to more remote and peripheral areas, the total

economic benefit to communities in the market for internet service despite the crowding out effect, but through a different mechanism from ours, namely via preemption through private investment in new technologies.

Heterogeneity Drivers

To help understand the patterns of reorganization of the airline networks and thus the sources of heterogeneity in the HST effects, we focus on routes which are the most affected (in terms of airline presence) by the entry of the HST. Specifically, we look at two groups with the top 10 routes which have the largest difference (positive or negative, depending on the group) in the predicted number of airlines between the two scenarios "Baseline" and "No HST". Results using a larger number of routes per group, for example 30, are consistent with those discussed here.

Table 12 compares itemized flow profits for each of the two groups of routes under the two scenarios. The table also reports the average number of airlines and flow profits for the two groups. Routes which are the most negatively affected by the HST in terms of airline presence have on average 2.75 fewer airlines in the "Baseline" than in the "No HST" scenario; this is consistent with these routes having lower profits on average when the HST is present compared to when it is not, as reported in the second row of Table 12. Conversely, the routes which are the most positively affected by the HST have on average 0.83 more airlines in the "Baseline" than in the "No HST" scenario and, as expected, have higher average profits when the HST is present.

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

Insert Table 12 about here

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

Panel A in Table 12 shows that the routes most negatively affected by the HST in terms of airline presence have a negative competitive effect with its origin in their overlap with the HST. Such competitive effect represents, on average, 88% of the profit losses of this group of routes and is five times larger than the positive HST spillover effects that result from connections to routes served by the HST. Considering that all of the routes with the largest negative difference in airline presence are short routes (not tabulated), these observations reflect the fact that airlines are at a disadvantage competing with the HST in relatively shorter routes. This also explains why cities in the central region of China were the most affected by airlines' exit after the entry of the HST. Cities in the East and Central areas

number of (unique) routes served either by train or air is larger when the HST is present, compared to when it is not: $1,030$ vs. 967 routes respectively (not tabulated).

of China are usually served by HST, which means that the overlap between HST routes and airline routes is large. In addition, their central geographical location implies that the average distance from the cities to other locations is relatively short, which gives the HST more competitive advantage in terms of travel convenience. Together, these factors make airline carriers less likely to serve routes that connect these cities when the HST is present.

As for the routes that are the most positively affected by the HST in terms of airline presence, panel B in Table 12 shows that the low route overlap with the HST together with a significant level of connectivity with the HST are the main drivers of airline presence. When the HST enters, routes that seemed less appealing otherwise, become more profitable due to the connection to other routes that are served by the train. This explains why airlines tend to relocate to cities in geographically peripheral areas, where the average distance between cities is large, and there is less overlap with the HST, therefore limiting the negative spillovers from the HST. Interestingly, there is also a preference for airlines to relocate to routes in which there is overlap with HST (fast train) as long as these are longer routes – about 70% of the positive spillovers from the fast train originate from long routes in which the fast train is present with the remaining 30% being associated with connections to routes served by this type of train (not tabulated). This is consistent with the existence of market expansion effects which make airlines want to explore new routes that are now served by the train, as discussed in section 7.

Finally, Table 12 also reveals that, even though the entry of the HST leads to considerable exit from airlines, the downsizing in airlines' own networks is not the main driver of airline network reconfiguration. The indirect effect of the HST presence on airline's profits, which is reflected in the downsizing of airlines' own networks, has the same magnitude in both the routes most negatively and the most positively affected by the HST, representing only 2% and 10% of the profits lost due to the HST's presence, respectively.

8.3 Improving the Positive Spillovers from the HST

The Chinese government consistently highlights the importance of increasing people's mobility in order to reduce inequality across the country.[57] Better integration of air and train travel can contribute to this objective by facilitating people's travel and has been encouraged by the government.[58] While such integration is likely to lead to more airline entry further enhancing mobility, and despite the government appeals, little has been done so far in an

[57] According to the 13th five-year-plan (2016 – 2020) Plan: "[We need to] establish sound mechanisms for the free movement of talent, improve horizontal and vertical social mobility, and encourage the orderly, free movement of talent between different kinds of organizations and between different regions".

[58] Source: `http://www.chinanews.com/cj/2017/02-28/8161929.shtml` (in Chinese and last accessed on July 5, 2022)

effort to integrate air and train transportation in China.

Better integration efforts could mirror what has been done in other countries. For example, in Germany, train stations are built closer to airports whenever possible and shuttle services are offered to connect railway stations and airports, when needed. Also, more cost-effective services have been implemented successfully by airlines such as Lufthansa including online single check-in, baggage drop-off directly in the terminal at the long-distance train station, and automatic rebooking onto the next flight or train in case of delays.

Given the Chinese government's emphasis on this topic, in this section we conduct a policy experiment to investigate the impact on the airline industry of an increase in the level of spillovers from the HST. This exercise allows us to quantify how much airline's profits can be affected by improving intermodal transportation, and to identify which cities would benefit the most from such coordination efforts.

We consider several scenarios in which we increase the value of the structural model coefficients associated with the positive spillovers from the HST on airlines (i.e., the parameters associated with the variable "number of HST line connections" for both the bullet and the fast trains) by different levels starting at 2% with 2% increments up to 30% and solve for the corresponding equilibria. As in the previous policy experiments, for each scenario, we run 1,000 simulations of the evolution of the airline route networks and compare the resulting airline route networks at the end of 2015 with those from the baseline scenario when the HST is present.

Figure 9 shows the impact of the different spillover levels on airline route presence (top two panels), measured as the total number of routes served and the total number of airline-routes, and on the airline industry total flow profits (bottom panel). The horizontal lines represent the levels of airline presence and flow profits in the scenario in which the HST is not introduced. The figure shows that, as expected, airline presence and profits increase substantially with the level of spillovers. Also, when the positive spillovers from connecting to HST lines increase by about 18%, the airline industry's flow profits match the profits in the scenario without HST. This increase in spillovers is associated with an increase in profits (per route and airline) of 10.75% relatively to the baseline scenario (not tabulated), which, assuming constant marginal costs, corresponds also to a 10.75% increase in the number of passengers. To put this number in perspective, this corresponds to about 84% of the average annual growth rate in air traffic volume in China (which was 13% between 2007 and 2015[59]).

[59]Source: Statistical Bulletin of Civil Aviation Industry Development, published by the Civil Aviation Administration of China, years 2017, 2012, and 2010.

==========================
Insert Figure 9 about here
==========================

We also investigate how the impact of improving the integration between train and air varies across the country. Figure 10 shows a map with the airline route presence for the top 20 cities by passenger volume. The black lines represent the HST routes, and each dot represents a city. The size of each dot is proportional to the airline-route presence in a given city (defined as the number of unique airline-route combinations that either start or end at that city) for the baseline scenario. The color of each dot shows how the airline-route presence in the scenario in which the positive spillovers from the HST are increased by 18% (the level at which the airline industry's flow profits match the profits in the baseline scenario without HST) compares with the scenario in which the HST is introduced (baseline). We use warm colors to denote an increase in airline route presence (and cold colors to denote a decrease) relatively to the scenario in which the level of spillovers is the one estimated in the data.

==========================
Insert Figure 10 about here
==========================

Table 13 reports in more detail the impact of the HST on airline route presence, for all regions in China, assuming the same level of positive spillovers as in Figure 10.

==========================
Insert Table 13 about here
==========================

The impact of an increase in the level of positive spillovers from the HST on the airline industry is not uniform across the country. As expected, the cities that benefit the most in terms of airline presence from improving intermodal transportation are those located in regions where there is a large number of HST connections. This is the case in the Central region, for which the difference in airline presence per city relative to the baseline scenario is the largest, followed by the Eastern region. As discussed in the policy experiment in Section 8.2, the cities in these regions are also the ones that are the most negatively affected in terms of airline presence due to the entry of the HST.

Despite the significant benefits from improved intermodal connectivity in the Central and Eastern regions, the benefits are not enough to compensate for the entry of HST in these two regions (the difference in airline presence per city relative to the case without HST presence is negative). This is because cities from these areas suffer more from the competition from the HST due to having a larger proportion of routes which overlap with the HST. In any case, given that airline service provision in cities from these regions is the most elastic to a change in spillovers from the HST, and without taking into account other economic considerations, this makes them important targets for providing intermodal connectivity services.

9 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). We setup and estimate a structural dynamic oligopoly model of airlines' decisions which accounts for the network structure in the data and that allows for multiple entry and exits of airlines into the same route.

We use a counterfactual exercise to compare the equilibrium airline network decisions with and without the presence of the HST. Despite the existence of significant positive spillovers from the HST on the airline industry, the introduction of the HST reduced airlines' route presence by about 14% and airline profits by 23%. Further, and even though the overall net impact of the introduction of HST is negative, there is considerable heterogeneity across cities and route-types in how the HST impacted the airline industry. Airlines readjusted their networks by substituting towards longer routes and more peripheral regions in China. This highlights a potential indirect benefit of the HST entry in shifting airlines to more remote and underserved areas thus improving connectivity among regions and reducing inequality.

The complexity of the effects of the HST on the airline industry, derived from the interplay between the positive and negative spillovers which differs across regions, must be taken into account when planning interventions from the government or airlines such as improvements in intermodal services.

References

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

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

Allen, Treb and Costas Arkolakis (2014), " Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129 (3), 1085–1140.

— (2022), "The Welfare Effects of Transportation Infrastructure Improvements," *The Review of Economic Studies*.

Angrist, Joshua D. and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

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

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

Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2020), "On the road: Access to transportation infrastructure and economic growth in China," *Journal of Development Economics*, 145, 102442.

Barwick, Panle Jia, Myrto Kalouptsidi, and Nahim Bin Zahur (2021), "Industrial Policy Implementation: Empirical Evidence from China's Shipbuilding Industry," *Working Paper*.

Barwick, Panle Jia, Shanjun Li, Andrew R Waxman, Jing Wu, and Tianli Xia (2022), "Efficiency and equity impacts of urban transportation policies with equilibrium sorting," *Working Paper*.

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

Berry, Steven T. and Joel Waldfogel (1999), "Public radio in the United States: does it correct market failure or cannibalize commercial stations?" *Journal of Public Economics*, 71 (2), 189–211.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics*, 119 (1), 249–275.

Blundell, Richard and Monica Costa Dias (2000), "Evaluation methods for non-experimental data," *Fiscal Studies*, 21 (4), 427–468.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou (2020), "Geography, Transportation, and Endogenous Trade Costs," *Econometrica*, 88 (2), 657–691.

Brancaccio, Giulia, Myrto Kalouptsidi, Theodore Papageorgiou, and Nicola Rosaia (2022), "Search Frictions and Efficiency in Decentralized Transport Markets," *Working Paper*.

Brandenburg, Adam M. and Berry J Nalebuff (1996), *Co-opetition*, New York: Doubleday.

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

Bresnahan, Timothy F. and Peter C. Reiss (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99 (5), 977–1009.

Buchholz, Nicholas (2022), "Spatial Equilibrium, Search Frictions, and Dynamic Efficiency in the Taxi Industry," *The Review of Economic Studies*, 89 (2), 556–591.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021), "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 225 (2), 200–230.

Cao, Guangyu, Ginger Zhe Jin, Xi Weng, and Li-An Zhou (2021), "Market-expanding or Market-stealing? Competition with network effects in bike-sharing," *The RAND Journal of Economics*, 52 (4), 778–814.

Card, David and Alan B. Krueger (1994), "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84 (4), 772–793.

Chen, Deqiu, Yujing Ma, Xiumin Martin, and Roni Michaely (2022), "On the fast track: Information acquisition costs and information production," *Journal of Financial Economics*, 143 (2), 794–823.

Chen, Fanglin, Xinyue Hao, and Zhongfei Chen (2021), "Can high-speed rail improve health and alleviate health inequality? Evidence from China," *Transport Policy*, 114, 266–279.

Chen, Zhenhua (2017), "Impacts of high-speed rail on domestic air transportation in China," *Journal of Transport Geography*, 62, 184–196.

Ciliberto, Federico, Charles Murry, and Elie Tamer (2021), "Market Structure and Competition in Airline Markets," *Journal of Political Economy*, 129 (11), 2995–3038.

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

Cosman, Jacob (2017), "Industry dynamics and the value of variety in nightlife: evidence from Chicago," *Working Paper.*

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

de Chaisemartin, Clément and Xavier D'Haultfœuille (2022), "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey," *The Econometrics Journal.*

Démurger, Sylvie (2001), "Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China?" *Journal of Comparative Economics*, 29 (1), 95–117.

Deng, Yiting and Carl F. Mela (2018), "TV Viewing and Advertising Targeting," *Journal of Marketing Research*, 55 (1), 99–118.

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

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

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

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

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

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

Fajgelbaum, Pablo D and Cecile Gaubert (2020), "Optimal Spatial Policies, Geography, and Sorting*," *The Quarterly Journal of Economics*, 135 (2), 959–1036.

Fajgelbaum, Pablo D. and Edouard Schaal (2020), "Optimal Transport Networks in Spatial Equilibrium," *Econometrica*, 88 (4), 1411–1452.

Fan, Ying and Chenyu Yang (2020), "Competition, Product Proliferation, and Welfare: A Study of the US Smartphone Market," *American Economic Journal: Microeconomics*, 12 (2), 99–134.

Fréchette, Guillaume R., Alessandro Lizzeri, and Tobias Salz (2019), "Frictions in a Competitive, Regulated Market: Evidence from Taxis," *American Economic Review*, 109 (8), 2954–92.

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

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

He, Li, Toni M Whited, and Ran Guo (2021), "Relative Performance Evaluation and Strategic Competition," *Working Paper*.

Heblich, Stephan, Stephen J Redding, and Daniel M Sturm (2020), "The Making of the Modern Metropolis: Evidence from London," *The Quarterly Journal of Economics*, 135 (4), 2059–2133.

Heuermann, Daniel F and Johannes F Schmieder (2018), "The effect of infrastructure on worker mobility: evidence from high-speed rail expansion in Germany," *Journal of Economic Geography*, 19 (2), 335–372.

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

Hotz, V. Joseph and Robert A. Miller (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60 (3), 497–529.

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

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

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

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

Laamanen, Tomi (2016), *Co-opetition*, London: Palgrave Macmillan UK.

Li, Linna and Becky P.Y. Loo (2017), "Railway Development and Air Patronage in China, 1993-2012: Implications for Low Carbon Transport," *Journal of Regional Science*, 57 (3), 507–522.

Li, Sophia, Joe Mazur, Yongjoon Park, James Roberts, Andrew Sweeting, and Jun Zhang (2022), "Repositioning and market power after airline mergers," *The RAND Journal of*

Economics, 53 (1), 166–199.

Li, Yan and Maria N. DaCosta (2013), "Transportation and income inequality in China: 1978–2007," *Transportation Research Part A: Policy and Practice*, 55, 56–71.

Liu, Shuli, Yulai Wan, Hun-Koo Ha, Yuichiro Yoshida, and Anming Zhang (2019), "Impact of high-speed rail network development on airport traffic and traffic distribution: Evidence from China and Japan," *Transportation Research Part A: Policy and Practice*, 127, 115–135.

Liu, Tracy, Zhixi Wan, and Chenyu Yang (2021), "The efficiency of a dynamic decentralized two-sided matching market," *Working Paper*.

Nevskaya, Yulia and Paulo Albuquerque (2019), "How Should Firms Manage Excessive Product Use? A Continuous-Time Demand Model to Test Reward Schedules, Notifications, and Time Limits," *Journal of Marketing Research*, 56 (3), 379–400.

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

Pakes, Ariel, Michael Ostrovsky, and Steven Berry (2007), "Simple estimators for the parameters of discrete dynamic games (with entry/exit examples)," *The RAND Journal of Economics*, 38 (2), 373–399.

Pesendorfer, Martin and Philipp Schmidt-Dengler (2008), "Asymptotic Least Squares Estimators for Dynamic Games," *The Review of Economic Studies*, 75 (3), 901–928.

Qin, Yu (2016), "No county left behind? The distributional impact of high-speed rail upgrades in China," *Journal of Economic Geography*, 17 (3), 489–520.

Redding, Stephen J (2021), "Trade and geography," *Working Paper*.

Redding, Stephen J. and Esteban Rossi-Hansberg (2017), "Quantitative Spatial Economics," *Annual Review of Economics*, 9 (1), 21–58.

Rust, John (1996), "Numerical dynamic programming in economics," in "Handbook of Computational Economics," Elsevier, volume 1, chapter 14, 619–729.

Ryan, Stephen P. (2012), "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 80 (3), 1019–1061.

Saad, Y. and M. Schultz (1986), "GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems," *SIAM Journal on Scientific and Statistical Computing*, 7 (3), 856–869.

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

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

Shi, Wenming, Kun-Chin Lin, Heather McLaughlin, Guanqiu Qi, and Mengjie Jin (2020), “Spatial distribution of job opportunities in China: Evidence from the opening of the high-speed rail,” *Transportation Research Part A: Policy and Practice*, 133, 138–147.

Sinai, Todd and Joel Waldfogel (2005), “Do low-income housing subsidies increase the occupied housing stock?” *Journal of Public Economics*, 89 (11), 2137–2164.

Smith, Dominic A (2018), “Concentration and foreign sourcing in the us retail sector,” *Working Paper*.

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

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

Tuchman, Anna E. (2019), “Advertising and Demand for Addictive Goods: The Effects of E-Cigarette Advertising,” *Marketing Science*, 38 (6), 913–1084.

Viswanathan, Madhu, Om Narasimhan, and George John (2021), “Economic Impact of Category Captaincy: An Examination of Assortments and Prices,” *Marketing Science*, 40 (2), 261–282.

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

Wang, Jiaoe, David Bonilla, and David Banister (2016), “Air deregulation in China and its impact on airline competition 1994–2012,” *Journal of Transport Geography*, 50, 12–23.

Wei, Yanhao (2018), “Airline networks, traffic densities, and value of links,” *Quantitative Marketing and Economics*, 16 (3), 341–370.

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

Wilson, Kyle (2021), “Does public competition crowd out private investment? Evidence from municipal provision of internet access,” *Working Paper*.

World Bank (2014), “Regional Economic Impact Analysis of High Speed Rail in China,” Technical Report ACS9734, The World Bank, Washington, DC.

Yang, Hangjun (Gavin), Qiong Zhang, and Qiang Wang (2018), “Airline Deregulation, Market Competition, and Impact of High-speed Rail on Airlines in China," in X. Fu and J. M. Peoples (Editors), "Advances in Airline Economics," Bingley: Emerald, 79–101.

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

Yao, Shujie, Fan Zhang, Feng Wang, and Jinghua Ou (2019), "Regional economic growth and the role of high-speed rail in China," *Applied Economics*, 51 (32), 3465–3479.

Yoo, Hyesung, Maria Ana Vitorino, and Song Yao (2022), "Hospital Competition and Quality: Evidence from the Entry of the High-Speed Train in South Korea," *Working paper*.

Zhang, Allen T. (2020), "Estimation of High-Dimensional Dynamic Games: Fintech Lenders and Bank Branch Closures," *Working paper*.

Zhang, Anming, Yulai Wan, and Hangjun Yang (2019), "Impacts of high-speed rail on airlines, airports and regional economies: A survey of recent research," *Transport Policy*, 81, A1–A19.

Zhang, Fangni, Zhiwei Yang, Jingjuan Jiao, Wei Liu, and Wenjie Wu (2020), "The effects of high-speed rail development on regional equity in China," *Transportation Research Part A: Policy and Practice*, 141, 180–202.

Zheng, Fanyin (2016), "Spatial competition and preemptive entry in the discount retail industry," *Working paper*.

Zou, Wei, Fen Zhang, Ziyin Zhuang, and Hairong Song (2008), "Transport Infrastructure, Growth, and Poverty Alleviation: Empirical Analysis of China." *Annals of Economics & Finance*, 9 (2), 345–371.

Figures and Tables

Figure 1: Airlines in China by Parent Company

This figure shows the logos of the major airline carriers (in terms of market shares) in China (Air China, China Southern Airlines, China Eastern Airlines, and Hainan Airlines) and their subsidiaries. It also lists other less significant airlines (labeled "other airlines"). Market shares were calculated using the authors' data and are based on the number of domestic flights operated by each airline for the top 70 airports in passenger volume between 2006 and 2016.

Figure 2: Evolution of the HST Network

This figure shows the evolution of the high speed train (HST) network in China from 2007 to 2016. Different colors are used to distinguish between fast train and bullet train routes.

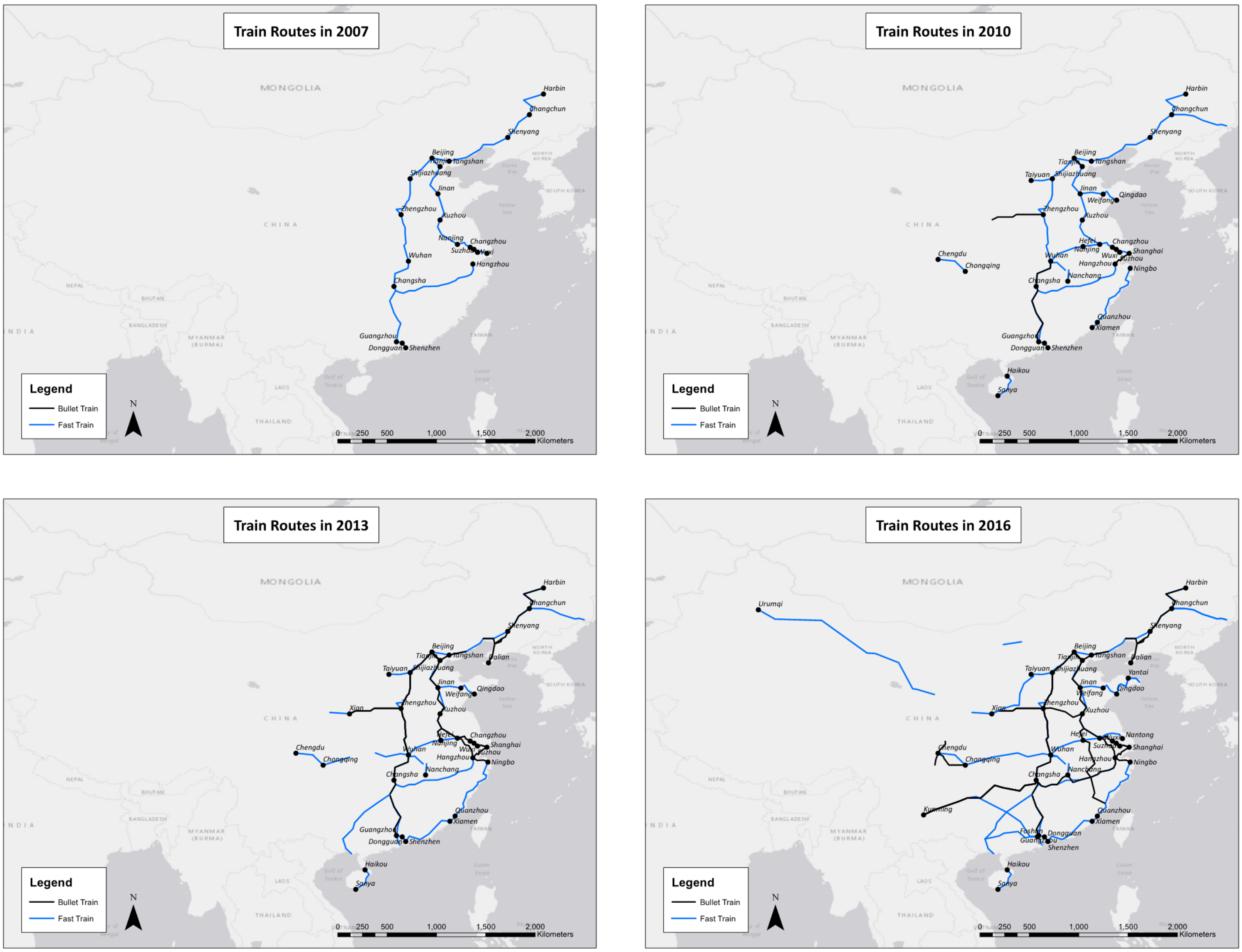

Figure 3: Travel Time as a Function of Distance by Mode of Transportation

This figure presents the relationship between travel time and travel distance for air travel and rail travel (bullet train). (Source: "Civil Aviation Big Data Report", 2016, in Chinese)

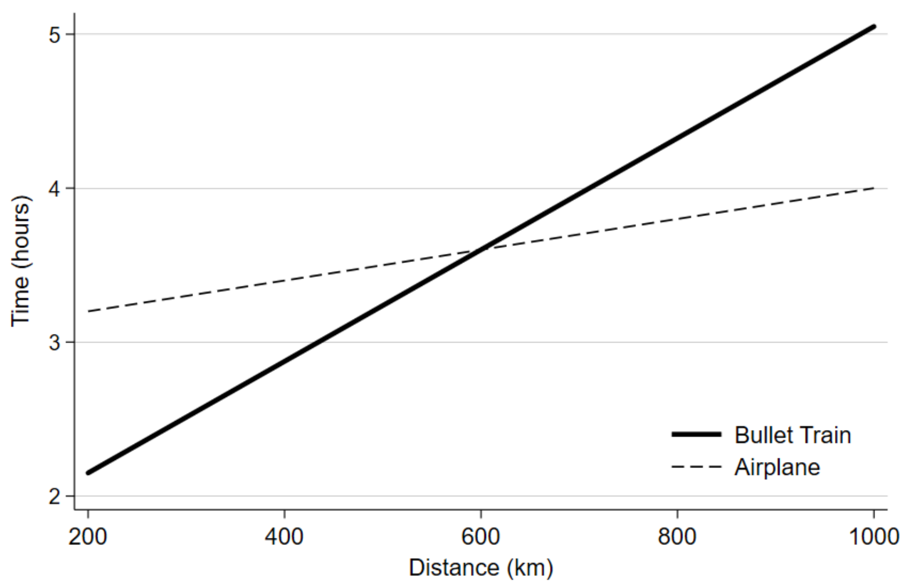

Figure 4: Number of Routes Served by Mode of Transportation

This figure shows the evolution of the number of routes served by airlines only and simultaneously by airlines and high-speed train (HST) (fast and bullet train) by year. 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 travel by HST between the cities in the city-pair without changing trains. The data used is based on our sample for the top 70 airports in passenger volume.

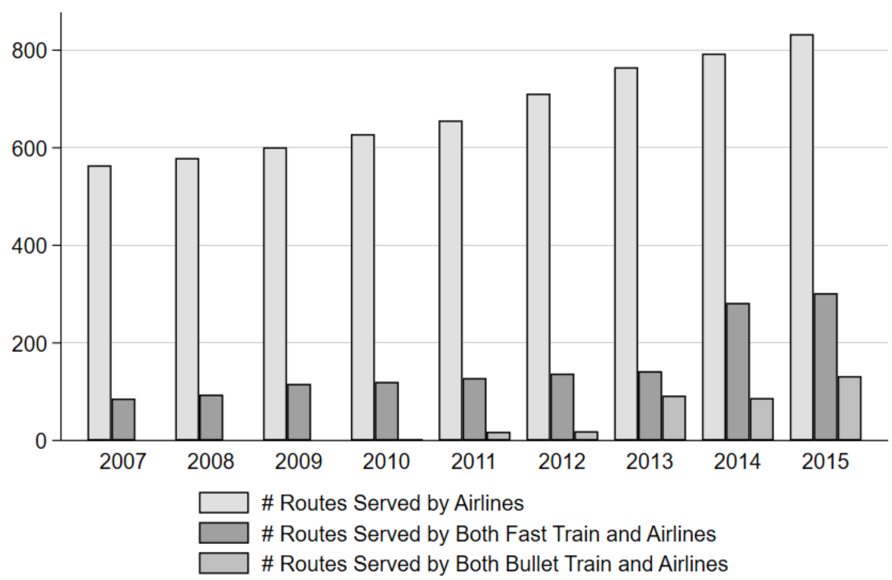

Figure 5: Airline Route Entry and Exit

This figure shows the evolution of the number of routes that experiencedx entry and exit by year. Entry and exit are defined independently of the number of entries and exits that occurred in a route. The data used is based on the authors' sample for the top 70 airports in passenger volume.

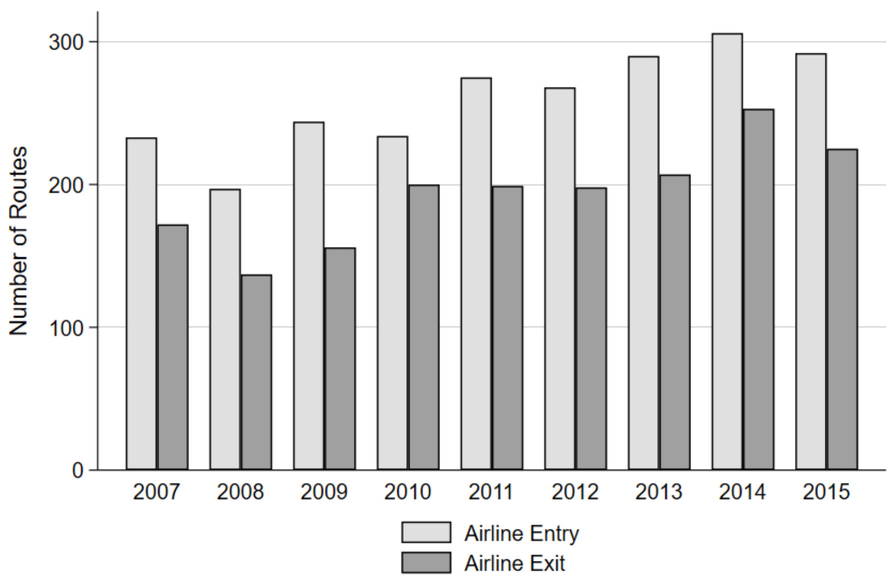

Figure 6: Average Number of Airlines per Route

This figure shows the evolution of the average number of airlines for different groups of routes. Panel A refers to short and medium/long routes that overlap with the HST and Panel B corresponds to routes connected or not to the HST that do not overlap with the HST. Routes are classified as overlapping with (or connecting to) the HST based on their status at the end of the year 2015.

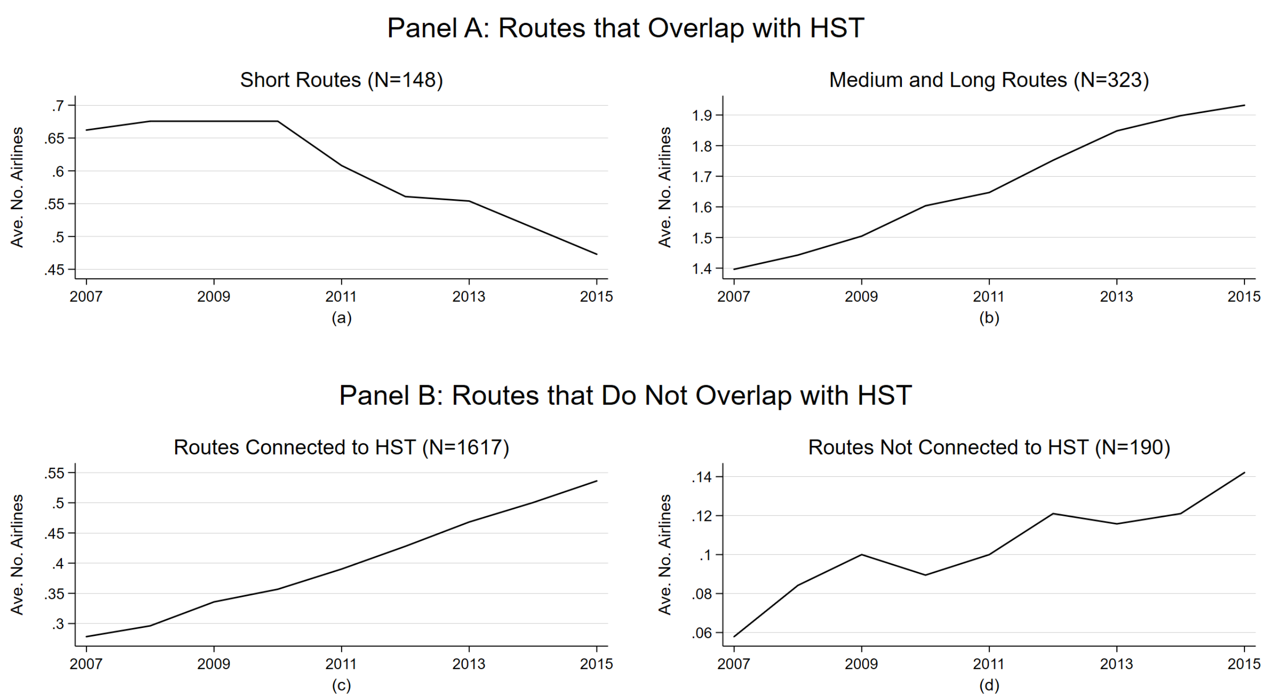

Figure 7: Graphical Representation of the Selection of the Control Routes

This figure graphically illustrates the criteria for selecting the set of two control routes per treated route in the difference-in-difference analysis. Here the treated route AB, with city endpoints A and B, is an air route that overlaps with HST, and the routes AC and BD are the control routes.

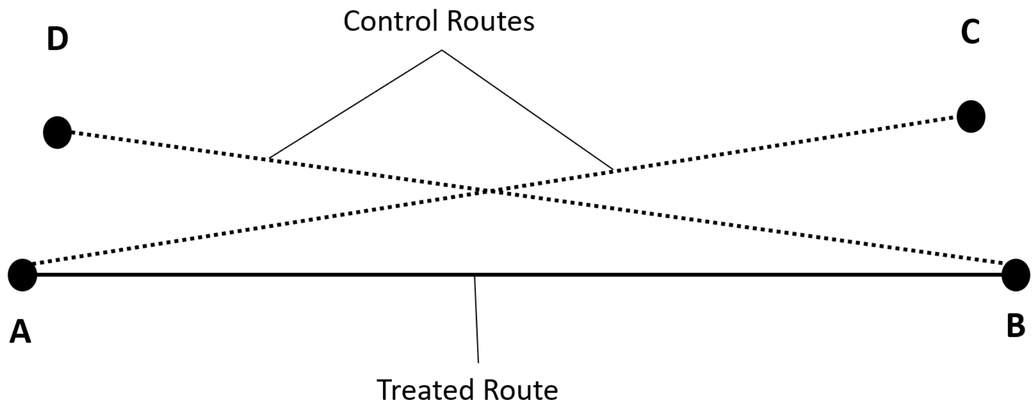

Figure 8: Impact of the HST on the Airline Industry by City

This figure shows a map with the top 20 cities by passenger volume at the end of the year 2015. The black lines represent the HST routes, and each dot represents a city. The size of each dot is proportional to the airline-route presence in a given city for the baseline scenario (with HST) and the color of each dot shows how the airline-route presence in a route in the scenario without HST compares with the baseline scenario. We use warm colors to denote an increase in the airline-route presence when the HST is introduced relatively to the scenario without HST, and cold colors to denote a decrease. Airline-route presence in a given city is defined as the number of unique airline-route combinations that either start or end at that city.

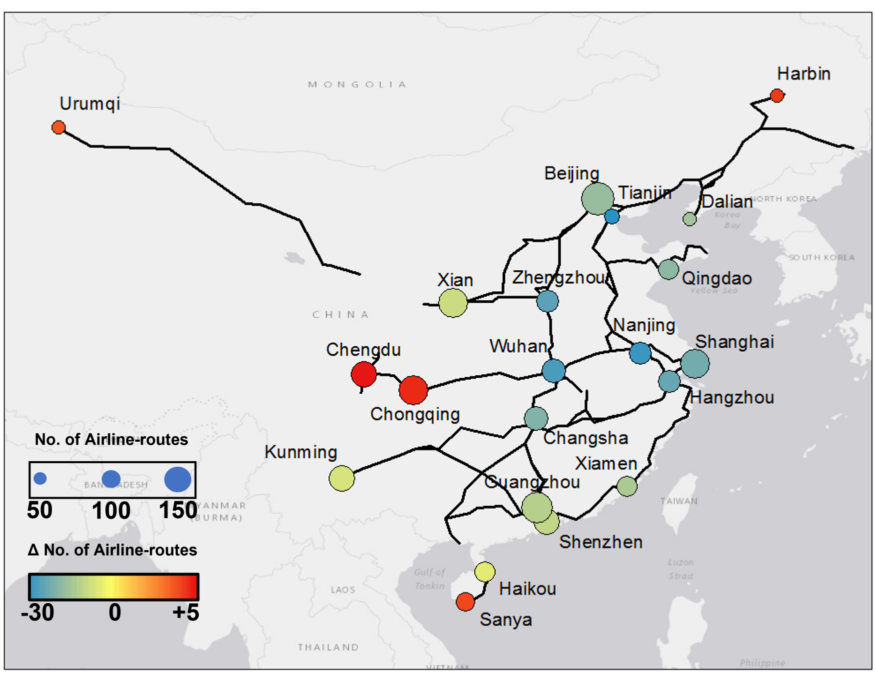

Figure 9: Impact of Improving the Positive Spillovers from the HST

This figure shows the impact of the effect of changing the value of the structural model coefficients associated with the positive spillovers from the HST on airlines (i.e., the parameters associated with the variable "number of HST line connections" for both the bullet and the fast trains) by different levels starting at 2%, with 2% increments, up to 30%. The top two panels show the impact on airline presence measured as the total number of routes served and the total number of airline-routes, and the bottom panel shows the impact on the airline industry total flow profits. The horizontal lines represent the levels of airline presence or of flow profits in the scenario in which the HST is not introduced.

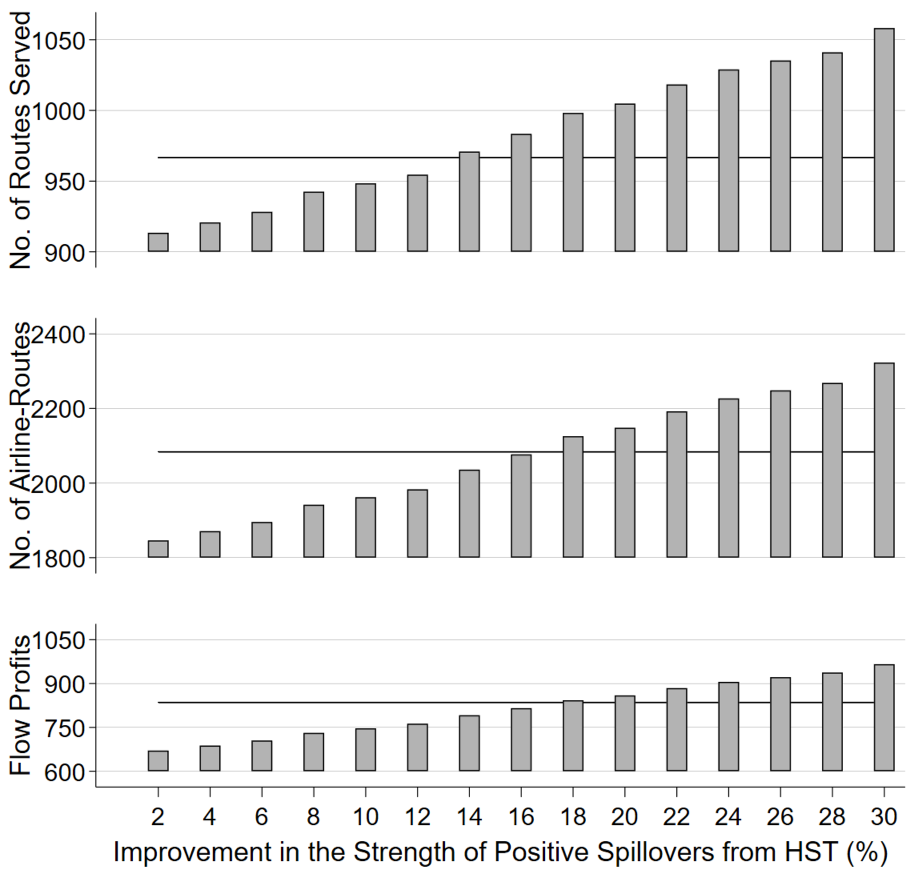

Figure 10: Impact of Improving the Positive Spillovers from the HST by City

This figure shows a map with the top 20 cities by passenger volume at the end of the year 2015. The black lines represent the HST routes and each dot represents a city. The size of each dot is proportional to the airline-route presence in a given city for the baseline scenario (with HST) and the shadow of each dot shows how the airline-route presence in a route in the scenario with an 18% improvement in the positive spillovers from the HST compares with the baseline scenario. Airline-route presence in a given city is defined as the number of unique airline-route combinations that either start or end at that city.

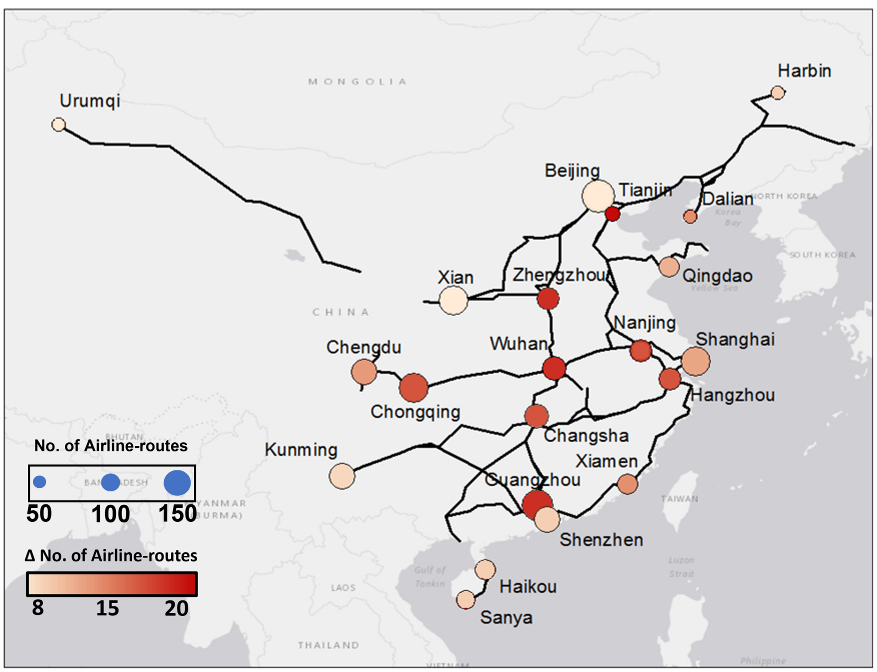

Table 1: Summary Statistics, Route Characteristics

This table reports the summary statistics for the route characteristics. A route is defined as a non-directional city-pair. The data is for the period from 2007 to 2015 (9 years) and includes all flights that operate for at least one year. The unit of analysis is a route-year combination. There are $2,278$ city-pairs in the data which makes a total of $20,502$ observations across all years. "Airline present" is an indicator variable that equals 1 if at least one airline is present in a given route and year (even if for part of the year), and 0 otherwise. "Number of airlines" is the number of airlines in a given route and year that have at least one flight that operates in that route during that year (even if for only part of the year). "Number of flights" is the number of unique flights (based on the flight numbers) that operate in a given route and year. "Number of airline connections" is is the sum across airlines of the number of distinct operating air routes that connect to either of the route endpoints. "HST present", "Fast train present", and "Bullet train present" are indicator variables that equal 1 if the HST train (either fast or bullet), fast train or the bullet train, respectively, are present in a given route and year, and 0 otherwise. "Number of HST connections" is the number of HST lines that pass through one (but not both) of the route endpoints. City-pair characteristics are calculated after averaging the characteristics of the two cities that constitute the route endpoints.

	N	Mean	Median	S.D.	Min	Max
Airline present	20,502	0.30	0.00	0.46	0.00	1.00
Number of airlines	20,502	0.57	0.00	1.02	0.00	4.00
Number of flights	20,502	1.52	0.00	4.12	0.00	74.00
Number of airline connections	20,502	87.97	78.00	56.16	2.00	318.00
HST present	20,502	0.12	0.00	0.32	0.00	1.00
Fast train present	20,502	0.11	0.00	0.32	0.00	1.00
Bullet train present	20,502	0.03	0.00	0.17	0.00	1.00
Number of HST connections	20,502	0.81	0.50	0.85	0.00	5.50
City-pair average population (Million)	20,502	5.34	4.82	3.40	0.22	24.19
City-pair average GDP (Billion USD)	20,502	55.56	43.05	45.29	0.49	374.57
City-pair average population growth rate	20,502	0.01	0.01	0.02	-0.16	0.23
City-pair average GDP growth rate	20,502	0.15	0.15	0.06	-0.21	0.61
City-pair distance (00 km)	20,502	15.26	13.89	8.71	0.54	44.06

Table 2: Average Route Characteristics by Year

This table reports the average route characteristics by year. The data includes all flights that operate for at least one year. There are $2,278$ city-pairs in the data. For the definition of the variables in the table please refer to Table 1's note.

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Airline present	0.25	0.25	0.26	0.28	0.29	0.31	0.34	0.35	0.37
Number of airlines	0.44	0.47	0.50	0.53	0.56	0.60	0.64	0.67	0.70
Number of flights	1.06	1.11	1.26	1.39	1.52	1.63	1.78	1.91	2.02
Number of airline connections	65.01	69.07	75.04	80.43	87.50	93.29	101.17	106.33	113.86
HST present	0.07	0.07	0.08	0.09	0.10	0.11	0.12	0.19	0.21
Fast train present	0.07	0.07	0.08	0.09	0.10	0.10	0.11	0.19	0.21
Bullet train present	0.00	0.00	0.00	0.00	0.02	0.02	0.07	0.07	0.10
Number of HST connections	0.26	0.31	0.40	0.65	0.86	0.89	1.11	1.23	1.54
City-pair average population (Million)	5.13	5.19	5.22	5.30	5.36	5.40	5.43	5.49	5.55
City-pair average GDP (Billion USD)	30.00	35.52	40.14	47.50	56.52	63.19	69.59	75.72	81.84
Observations	2,278	2,278	2,278	2,278	2,278	2,278	2,278	2,278	2,278

Table 3: Airline and High Speed Train Presence by Quantiles of Route Length

This table reports the mean values of select route characteristics for short, medium and long distance routes. Short routes have less than 600km in length, medium routes between 600km and 1200km and long routes more than 1200km. The unit of analysis is a year-route combination. The total number of observations is $20,502$, which is given by the number of routes $(2,278)$ times the number of years (9) in the data. "City-pair distance" is defined as the distance between the city-pair that constitute the endpoints of a route. "Number of flights" is the number of unique flights (based on the flight numbers) that operate in a given route and year. "Number of entries" in a given route and year is the number of flights (unique flight numbers) that started operating in that year (and which were not in operation before) and were in operation for at least one year after entry. "Number of exits" in a given route and year refers to the number of flights that ceased operations in that year and route (after having been in operation for longer than one year). "Fast train present" and "bullet train present" are indicator variables that equal 1 if fast train or bullet train, respectively, are present in a given route and year (even if for part of the year), and 0 otherwise.

	Short	Medium	Long
City-pair distance (00 km)	3.86	9.15	20.61
Number of flights	1.98	2.51	0.96
Number of entries	0.15	0.19	0.10
Number of exits	0.14	0.14	0.07
Fast train present	0.32	0.18	0.04
Bullet train present	0.10	0.05	0.01
Observations	2,682	5,643	12,177

Table 4: Probability of Entry in a Route

This table reports the marginal effect estimates from a probit regression of airline entry in a route in a particular year as a function of whether an airline operates in none, one or both endpoints of that route in the previous year. Year and airline fixed effects are included. The excluded category refers to observations in which an airline does not serve either endpoint airport at any point in time during the previous calendar year. Marginal effects are evaluated at the mean of all other variables. Robust standard errors are reported in parentheses. (***), (**) and (*) denote statistical significance at the 1%, 5% and 10% level, respectively.

Airline operates in one endpoint airport in the previous year	0.0283***
	(0.0028)
Airline operates in both endpoint airports in the previous year	0.1115***
	(0.0028)
N	64,364

Table 5: Determinants of Airline Entry

This table reports the results for the difference-in-difference regressions. The dependent variable in columns (1) through (4) is the number of airlines operating in a given route. The unit of analysis is a year-route combination. “HST” is an indicator variable that equals 1 if the HST is present in a given route and year (even if for part of the year), and 0 otherwise. “No. of HST connections” is the number of HST lines that pass through one (but not both) of the route endpoints. “No. of airline connections” is the number of distinct operating air routes that connect to either of the route endpoints. “Average GDP” is the average of the GDP (in billions of US dollars) for a route’s endpoints. We categorize the routes into three groups in terms of route length: short, medium and long, using 600km and 1200km as cutoff points. The indicator variables “Medium Distance” and “Long Distance” correspond to the last two groups. Whenever route-fixed effects are included in the model, the route-specific covariates that do not change over time (i.e., route length) are naturally not identified (only their interaction with other covariates that change over time is identified). Robust and clustered (at the route level) standard errors are reported in parentheses. (***), (**) and (*) denote statistical significance at the 1%, 5% and 10% level, respectively.

	(1)	(2)	(3)	(4)
HST (Yes/No)	-0.697^{***}	-0.686^{***}	-0.172^{**}	-0.190^{**}
	(0.089)	(0.089)	(0.070)	(0.086)
No. of HST connections	-0.024^{**}	-0.003	0.030^{***}	0.071^{***}
	(0.011)	(0.012)	(0.008)	(0.017)
HST × No. of HST connections	-0.013	-0.013	-0.065^{***}	-0.049^{***}
	(0.017)	(0.017)	(0.013)	(0.019)
HST × Medium Distance	1.110^{***}	1.106^{***}	0.498^{***}	0.516^{***}
	(0.110)	(0.111)	(0.082)	(0.088)
HST × Long Distance	1.344^{***}	1.340^{***}	0.621^{***}	0.632^{***}
	(0.119)	(0.120)	(0.100)	(0.112)
No. of airline connections	0.014^{***}	0.014^{***}	0.010^{***}	0.010^{***}
	(0.001)	(0.001)	(0.001)	(0.002)
Average GDP	-0.002	0.001	0.010^{*}	-0.003
	(0.006)	(0.006)	(0.006)	(0.012)
Medium Distance	-0.021	-0.017		
	(0.059)	(0.060)		
Long Distance	-0.391^{***}	-0.380^{***}		
	(0.054)	(0.055)		
Year fixed effects	No	Yes	Yes	No
Route fixed effects	No	No	Yes	Yes
Group fixed effects	No	No	No	No
Year-group fixed effects	No	No	No	Yes
Observations	20,502	20,502	20,502	7,884
R^2	0.533	0.537	0.901	0.926

Table 6: Structural Model Estimation Results

This table reports the estimation results for the structural model. For details on the definition of the variables in the table please refer to Section 6.4. Standard errors are calculated via bootstrapping. (***), (**) and (*) denote statistical significance at the 1%, 5% and 10% level, respectively.

		Coef.	s.e.
Flow Payoffs' Parameters			
Strategic Effects	No. of own routes connected	0.20***	0.01
	No. of competitors	−0.12**	0.06
	No. of competitors' routes connected	0.02**	0.01
Impact of HST	Bullet Train (Y/N)	−0.32**	0.12
	Bullet Train × Medium distance	0.22	0.15
	Bullet Train × Long distance	0.15	0.13
	No. of Bullet Train connections	0.15***	0.05
	Bullet Train × No. of Bullet train line connections	−0.27***	0.09
	Fast Train (Y/N)	−0.53***	0.10
	Fast Train × Medium distance	0.42***	0.10
	Fast Train × Long distance	0.74***	0.13
	No. of Fast Train connections	0.10***	0.03
	Fast Train × No. of Fast Train line connections	−0.25***	0.06
Market Characteristics	Average GDP	−0.01	0.02
	Medium distance	−0.16***	0.05
	Long distance	−0.27***	0.05
	Unobserved Type	0.62***	0.12
	Constant	0.07	0.15
Entry Costs' Parameters			
	Constant	−4.63***	0.21
	Unobserved Type	1.36***	0.19
	Regulated	−1.42***	0.29
	Exempt	0.22	0.31
	No. of own routes connected	0.27***	0.06

Table 7: Goodness of Fit

This table reports the structural model's goodness of fit. Columns (1) and (2) report correlations between the predicted and observed values of two metrics, Route Presence and Route Density, respectively. Route Presence is measured as a vector of indicator variables which take the value one if there is at least one airline present in a given route (at the end of each year). Route Density is measured as a vector with the number of airlines present across the different routes. Predictions are obtained by averaging route presence and density across $1,000$ model simulations (using different draws of the error terms) of each airline's policy function using the estimated parameters from the structural estimation.

	Correlations between Model Predictions and Sample Observations	
Year	(1)	(2)
2007	0.923	0.953
2008	0.865	0.937
2009	0.825	0.919
2010	0.811	0.918
2011	0.821	0.916
2012	0.812	0.906
2013	0.806	0.907
2014	0.804	0.906
2015	0.815	0.905

Table 8: Profit Decomposition

This table shows the decomposition of airline flow profits into their different sources using the structural model estimates and the observed data. We report the average flow profits, across all routes and airlines, for the year of 2015. Average profits are calculated using only routes in which airlines are present ($1,388$ routes).

Source	Profit
Own Network	0.661
Competitors' Network	−0.001
Negative Spillover–Fast Train	−0.117
Negative Spillover–Bullet Train	−0.065
Positive Spillover–Fast Train	0.114
Positive Spillover–Bullet Train	0.109
Market Characteristics	−0.399
Constant	0.067
Total	0.369

Table 9: Overall Impact of the HST on the Airline Industry

This table presents the airline presence and flow profits under different simulated scenarios. Column "No HST" corresponds to the airline network configuration under the scenario where the HST were not present, column "Baseline" corresponds to the baseline scenario in which the HST is present, and column "No positive spill. from HST" corresponds to the scenario where HST is introduced but there are no positive spillovers from the HST. "Route presence" is defined as the total number of routes that are served by at least one airline. "Airline-Route presence" is defined as the sum across airlines of the total number of routes served by each airline. "Flow profits" are the total flow profits for the airline industry and are computed based on the network configuration at the end of year 2015 and reported in annual terms. The numbers reported are obtained by averaging across $1,000$ simulations for each scenario at the end of 2015. Specifically, "Route Presence" is calculated by counting, for each simulation, the number of unique routes served by at least one airline, and then taking the average across simulations. "Airline-Route presence" is calculated by calculating the total number of airlines across routes for each simulation, and then taking the average across simulations. "Flow profits" are calculated by taking the average across simulations of the total industry (taken across routes and airlines) flow profits.

	No HST	Baseline	No positive spill. from HST
Route presence	967	894	427
Airline-Route presence	2083	1785	654
Flow profits	835	639	306
Average flow profit per airline-route	0.40	0.36	0.47

Table 10: Impact of the HST by Region

This table presents the impact of HST on the airline industry for different regions in China defined according to the National Bureau of Statistics of China. Airline-Route Presence is defined as the sum across airlines of the total expected number of routes served by each airline in a given region. The numbers reported are obtained by averaging across $1,000$ simulations for each scenario at the end of 2015.

	Northeast	North	Northwest	East	Central	South	Southwest
No. of Cities	5	8	9	20	5	11	10
Average No. of HST Lines per City	1.8	2.3	0.7	2.3	3.0	1.9	0.7
Average Route Length (00 km)	18.6	14.4	20.6	12.9	11.3	15.2	16.2
Route Overlap with HST (%)	15.5%	19.4%	5.3%	34.0%	36.7%	18.3%	6.1%
Airline-Route Presence (Baseline)	220.6	450.0	339.6	1055.7	367.5	676.8	460.6
Airline-Route Presence (No HST)	230.4	543.2	333.0	1463.1	468.6	697.2	431.6
Difference	-9.7	-93.1	6.6	-407.4	-101.0	-20.4	29.0
Difference per City	-1.9	-11.6	0.7	-20.4	-20.2	-1.9	2.9

Table 11: Impact of the HST by Route Type

This table presents the impact of HST on airline presence across different types of routes. We classify routes into six groups depending on whether the HST is present in the route (present and not present) and on the length of the route (short, medium and long). The numbers reported are obtained by averaging across $1,000$ simulations for each scenario at the end of 2015.

Route type			Average No. Airlines			
Length	HST	N	Baseline	No HST	Difference	Difference (%)
Short	No	150	0.89	1.25	-0.35	-28.4%
Short	Yes	148	0.64	2.34	-1.70	-72.6%
Medium	No	424	0.96	0.98	-0.03	-2.6%
Medium	Yes	203	2.01	2.50	-0.49	-19.5%
Long	No	1,233	0.40	0.33	0.07	19.9%
Long	Yes	120	2.09	1.80	0.29	16.0%

Table 12: Routes Most Affected by the HST

This table reports simulated airline presence and profit decomposition for the two groups of top 10 routes which have the largest difference in the predicted number of airlines between the two scenarios "Baseline" and "No HST". Panel A reports results for the routes which are the most negatively affected by the HST in terms of airline presence, and Panel B for the routes which are the most positively affected by the HST. "Profit" refers to the average profits calculated across routes (10 routes) and airlines (4 airlines). The profit decomposition items are defined based on the structural model flow payoff variables listed in Table 6. "Own network" refers to the effect associated with the number of own routes connected. "Competitors' Network" includes the effects from the number of competitors and competitors' connections. The classification of the effects associated with the impact of the HST into "Negative" and "Positive" Spillover effects is done based on the sign of the estimated coefficients from the structural model. "Market characteristics" includes the effects associated with the constant term of profits, GDP, and the unobserved market type. The numbers reported are obtained by averaging across $1,000$ simulations for each scenario at the end of 2015.

Panel A: Routes most negatively affected by the HST

	Baseline	No HST	Difference
Average No. of Airlines	0.41	3.15	-2.75
Profit	-0.39	0.53	-0.92

Profit Decomposition

	Gains/Losses	Fraction of Total Gains (Losses)	Gains/Losses	Fraction of Total Gains (Losses)	Gains/Losses	Fraction of Total Gains (Losses)
Own Network	0.76	62%	0.79	100%	-0.04	(2%)
Competitors' Network	0.18	15%	-0.06	(24%)	0.25	46%
Negative Spillovers from HST	-1.41	(88%)	0.00	0%	-1.41	(98%)
Positive Spillovers from HST	0.29	23%	0.00	0%	0.29	54%
Market Characteristics	-0.20	(12%)	-0.20	(76%)	0	0%
Total Gains	1.23		0.79		0.53	
Total Losses	-1.61		-0.26		-1.45	

Panel B: Routes most positively affected by the HST

	Baseline	No HST	Difference
Average No. of Airlines	2.38	1.54	0.83
Profit	0.46	0.24	0.22

Profit Decomposition

	Gains/Losses	Fraction of Total Gains (Losses)	Gains/Losses	Fraction of Total Gains (Losses)	Gains/Losses	Fraction of Total Gains (Losses)
Own Network	0.73	54%	0.77	91%	-0.04	(10%)
Competitors' Network	-0.01	(1%)	0.08	9%	-0.09	(23%)
Negative Spillovers from HST	-0.26	(30%)	0	0%	-0.26	(67%)
Positive Spillovers from HST	0.61	46%	0	0%	0.61	100%
Market Characteristics	-0.61	(69%)	-0.61	(100%)	0.00	0%
Total Gains	1.34		0.85		0.61	
Total Losses	-0.88		-0.61		-0.39	

Table 13: Impact of Improving the Positive Spillovers from the HST by Region

This table presents the impact of an improvement of positive spillovers from HST (18%) on airline route presence for different regions in China. The last row in the table compares the figures of the scenario with improved spillovers with the scenario without HST which is reported in Table 10. The numbers reported are obtained by averaging across 1,000 simulations for each scenario at the end of 2015. For the definition of the variables in the table please refer to Table 10's note.

Area	Northeast	North	Northwest	East	Central	South	Southwest
No. of Cities	5	8	9	20	5	11	10
Average No. HST Lines per City	1.8	2.3	0.7	2.3	3.0	1.9	0.7
Route Overlap with HST (%)	15.5%	19.4%	5.3%	34.0%	36.7%	18.3%	6.1%
Airline-Route Presence (Improved Spill.)	276.4	540.5	394.2	1285.8	434.7	791.5	528.3
Airline-Route Presence (Baseline)	220.6	450.0	339.6	1055.7	367.5	676.8	460.6
Difference	55.8	90.4	54.7	230.1	67.2	114.7	67.8
Difference per City	11.2	11.3	6.1	11.5	13.4	10.4	6.8
Difference relative to No HST per City	9.2	-0.3	6.8	-8.9	-6.8	8.6	9.7

Appendix

A Robustness Checks for Difference-in-Differences Analysis

Here we check the robustness of the results from Table 5 by carrying out three additional analyses. Specifically, in Table A.1, we report the estimation results at the airline-route-month level (as opposed to at the airline-route year, as in the main specification) and restrict the set of treated routes such that we only include data for treated routes for one month before the entry of the HST and one month after the introduction of the HST.

Table A.1: Determinants of Airline Entry

	(1)	(2)	(3)	(4)
HST (Yes/No)	-0.457^{***}	-0.466^{***}	-0.081^{**}	-0.070^{*}
	(0.075)	(0.075)	(0.034)	(0.042)
No. of HST connections	-0.017^{*}	0.005	0.026^{***}	0.047^{***}
	(0.010)	(0.012)	(0.006)	(0.013)
HST × No. of HST line connections	0.050^{**}	0.059^{***}	-0.023^{*}	-0.027^{**}
	(0.022)	(0.022)	(0.013)	(0.013)
HST × Medium Distance	0.786^{***}	0.787^{***}	0.215^{***}	0.314^{***}
	(0.091)	(0.091)	(0.044)	(0.049)
HST × Long Distance	0.862^{***}	0.870^{***}	0.284^{***}	0.243^{***}
	(0.100)	(0.100)	(0.056)	(0.069)
No. of airline connections	0.014^{***}	0.014^{***}	0.010^{***}	0.012^{***}
	(0.001)	(0.001)	(0.001)	(0.002)
Average GDP	-0.003	-0.001	0.014^{***}	0.003
	(0.006)	(0.006)	(0.005)	(0.011)
Medium Distance	-0.010	-0.006		
	(0.056)	(0.056)		
Long Distance	-0.354^{***}	-0.343^{***}		
	(0.051)	(0.052)		
Year fixed effects	No	Yes	Yes	No
Route fixed effects	No	No	Yes	Yes
Group fixed effects	No	No	No	No
Year-group fixed effects	No	No	No	Yes
Observations	237,345	237,345	237,345	61,653
R^2	0.465	0.470	0.887	0.143

In Table A.2 we show estimation results when the dependent variable is the log of the number of flights (as opposed to number of airlines, as in the main analysis).

Table A.2: Determinants of Airline Entry

	(1)	(2)	(3)	(4)
HST (Yes/No)	-0.256^{***}	-0.251^{***}	-0.044	-0.114^{**}
	(0.029)	(0.029)	(0.028)	(0.048)
No. of HST connections	-0.018^{***}	-0.010^{***}	0.007^{*}	0.091^{***}
	(0.004)	(0.004)	(0.004)	(0.017)
HST × No. of HST connections	0.021^{***}	0.021^{***}	-0.021^{***}	-0.081^{***}
	(0.008)	(0.008)	(0.007)	(0.018)
HST × Medium Distance	0.311^{***}	0.309^{***}	0.120^{***}	0.316^{***}
	(0.038)	(0.038)	(0.040)	(0.047)
HST × Long Distance	0.418^{***}	0.416^{***}	0.179^{***}	0.449^{***}
	(0.052)	(0.052)	(0.054)	(0.057)
No. of airline connections	0.003^{***}	0.003^{***}	0.002^{***}	0.012^{***}
	(0.000)	(0.000)	(0.000)	(0.002)
Average GDP	0.009^{***}	0.010^{***}	0.014^{***}	0.264
	(0.002)	(0.002)	(0.003)	(0.167)
Medium Distance	-0.028	-0.026		
	(0.017)	(0.017)		
Long Distance	-0.087^{***}	-0.083^{***}		
	(0.016)	(0.016)		
Year fixed effects	No	Yes	Yes	No
Route fixed effects	No	No	Yes	Yes
Group fixed effects	No	No	No	No
Year-group fixed effects	No	No	No	Yes
Observations	20,502	20,502	20,502	7,884
R^2	0.308	0.313	0.588	0.963

In Table A.3, we check the robustness of our results with respect to our data sample definition and re-run our analysis using all flights that operate longer than three months (as opposed to including only those that operate longer than one year, as in the main specification).

Table A.3: Determinants of Airline Entry

	(1)	(2)	(3)	(4)
HST (Yes/No)	−0.780***	−0.768***	−0.204***	−0.181**
	(0.097)	(0.098)	(0.076)	(0.089)
No. of HST connections	−0.015	0.008	0.039***	0.079***
	(0.012)	(0.013)	(0.008)	(0.018)
HST × No. of HST line connections	−0.039**	−0.039**	−0.077***	−0.056***
	(0.018)	(0.018)	(0.013)	(0.020)
HST × Medium Distance	1.238***	1.234***	0.522***	0.553***
	(0.119)	(0.120)	(0.089)	(0.094)
HST × Long Distance	1.559***	1.555***	0.687***	0.603***
	(0.121)	(0.122)	(0.106)	(0.120)
No. of airline connections	0.014***	0.014***	0.009***	0.011***
	(0.000)	(0.000)	(0.001)	(0.002)
Average GDP	−0.003	−0.000	0.008	−0.006
	(0.006)	(0.006)	(0.006)	(0.013)
Medium Distance	−0.041	−0.036		
	(0.065)	(0.066)		
Long Distance	−0.470***	−0.459***		
	(0.060)	(0.061)		
Year fixed effects	No	Yes	Yes	No
Route fixed effects	No	No	Yes	Yes
Group fixed effects	No	No	No	No
Year-group fixed effects	No	No	No	Yes
Observations	20,502	20,502	20,502	7,884
R^2	0.539	0.543	0.893	0.929

For each analysis, we report the results for the same specifications as in Table 5 for ease of comparison. The definitions for the variables used in the tables can be found in Table 5's note. Robust and clustered (at the route level) standard errors are reported in parentheses. (***), (**) and (*) denote statistical significance at the 1%, 5% and 10% level, respectively.

B Robustness Checks for Structural Model Results

Here we check the robustness of the results from Table 6 by carrying out two additional analyses. Specifically, in Table B.1, we check the robustness of our results with respect to our data sample definition and re-run our analysis using all flights that operate longer than three months (as opposed to including only those that operate longer than one year, as in the main specification). In Table B.2 we report the estimation results for the case where we allow each airline to have both different baseline flow payoffs and different entry costs (note that, in this last case, compared to the main specification, we do not include the variable "exempt" because otherwise the state space becomes unmanageable). The definitions for the variables used in the tables can be found in Table 6's note. Standard errors are calculated via bootstrapping. (***), (**) and (*) denote statistical significance at the 1%, 5% and 10% level, respectively.

Table B.1: Structural Model Estimation Results

		Coef.	s.e.
Flow Payoffs' Parameters			
Strategic Effects	No. of own routes connected	0.58***	0.04
	No. of competitors	−0.82***	0.18
	No. of competitors' routes connected	0.09***	0.02
Impact of HST	Bullet Train present (Y/N)	−0.74***	0.20
	Bullet Train × Medium distance	0.51**	0.21
	Bullet Train × Long distance	0.73***	0.20
	No. of Bullet Train line connections	0.17***	0.05
	Bullet Train × No. of Bullet train line connections	−0.57***	0.11
	Fast Train present (Y/N)	−1.31***	0.24
	Fast Train × Medium distance	1.27***	0.22
	Fast Train × Long distance	2.06***	0.27
	No. of Fast Train line connections	0.26***	0.05
	Fast Train × No. of Fast Train line connections	−0.54***	0.08
Market Characteristics	GDP	0.08	0.06
	Medium distance	−0.56***	0.21
	Long distance	−1.18***	0.27
	Unobserved Type	1.84***	0.27
	Constant	−1.1 * **	0.28
Entry Costs' Parameters			
Entry Costs	Entry Cost	−2.48***	0.18
	Unobserved Type	0.75**	0.35
	Regulated	−1.44***	0.25
	No. of own routes connected	0.07*	0.04
	Exempt	0.20	0.28

Table B.2: Structural Model Estimation Results

		Coef.	s.e.
Flow Payoffs' Parameters			
Strategic Effects	No. of own routes connected	0.44***	0.11
	No. of competitors	−0.88***	0.33
	No. of competitors' routes connected	0.11 **	0.04
Impact of HST	Bullet Train present (Y/N)	−0.90***	0.31
	Bullet Train × Medium distance	0.56*	0.30
	Bullet Train × Long distance	0.95***	0.34
	No. of Bullet Train line connections	0.19**	0.08
	Bullet Train × No. of Bullet train line connections	−0.37**	0.15
	Fast Train present (Y/N)	−1.36***	0.29
	Fast Train × Medium distance	1.26***	0.26
	Fast Train × Long distance	2.01***	0.43
	No. of Fast Train line connections	0.23***	0.06
	Fast Train × No. of Fast Train line connections	−0.34***	0.11
Market Characteristics	GDP	0.16**	0.07
	Medium distance	−0.55**	0.26
	Long distance	−1.14***	0.34
	Unobserved Type	1.81***	0.31
	Constant	−1.27***	0.48
	CZ	−0.05	0.04
	MU	0.002	0.06
	HU	−0.08	0.06
Entry Costs' Parameters			
Entry Costs	Entry Cost	−3.10***	0.71
	Entry Cost × CZ	−0.33**	0.16
	Entry Cost × MU	0.14	0.11
	Entry Cost × HU	0.53***	0.19
	Unobserved Type	0.68***	0.22
	Regulated	−1.67***	0.45
	No. of own routes connected	0.41	0.41