The determinants of hidden-city ticketing: Competition, hub-and-spoke networks, and advance-purchase requirements

Luttmann, Alexander and Gaggero, Alberto A

The MITRE Corporation, University of Pavia

26 July 2022

The determinants of hidden-city ticketing: Competition, hub-and-spoke networks, and advance-purchase requirements*

Alberto A. Gaggero†
University of Pavia

Alexander Luttmann‡
The MITRE Corporation

March 3, 2023

Abstract

We offer a comprehensive empirical study on hidden-city ticketing (HCT), a pricing phenomenon in the airline industry that occurs when the fare for a nonstop trip from A to B is more expensive than a connecting trip from A to B and B to C. Exploiting a unique panel of over 473 thousand fares for flights departing between October 1st, 2019 and December 31st, 2019, we find that HCT depends on route competition (both on A-B and A-C routes), largely occurs in the last week to departure, is less likely when airport C is a hub, and primarily occurs on carriers that operate large hub-and-spoke networks (e.g., American, Delta, and United).

JEL classification: L11, L13, L93, D40.

Keywords: advance-purchase, airline pricing, competition, hidden-city ticketing, hub airports, price discrimination.

Conflict of Interest: the authors declare that they have no conflict of interest.

*We are grateful to three anonymous referees, Graziano Abrate, Jan Brueckner, Branko Bubalo, Antonio Majocchi, Marcella Nicolini, and to the participants of the 2021 International Transportation Economics Association conference and the 2021 TEM workshop for helpful comments. This manuscript is a revised version of the 2022 MPRA WP, N.113960. All errors are our own. The author's affiliation with The MITRE Corporation is for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions, or viewpoints expressed by the author.

†Department of Economics and Management, University of Pavia, Via S. Felice 5, 27100 Pavia, Italy; email: alberto.gaggero@unipv.it. *Corresponding author*.

‡Transportation Performance and Economic Analysis Department, The MITRE Corporation, 7525 Colshire Drive, McLean, VA 22102; email: aluttmann@mitre.org.

1 Introduction

Hidden-city ticketing (HCT) is a pricing phenomenon in the airline industry that occurs when the price for a nonstop trip from A to B is more expensive than the price for a connecting trip from A to C that connects at B (i.e., the "hidden city").[1] When this phenomenon occurs, passengers traveling from A to B can save money by purchasing the connecting A-B-C trip. These HCT passengers would take the first flight from A to B, then deliberately forego the trip's second flight from B to C.

From a price discrimination perspective, HCT can be viewed as a unique, and perhaps counterintuitive, case of mixed bundling.[2] In the typical case, the bundle of products is sold at a cheaper price than the sum of the prices of the individual components (e.g., a vacation package that includes flight and hotel). However, in the HCT case, the bundle of products (i.e., the flights from A to B and B to C) is sold at a cheaper price than one component of the bundle (i.e., the flight from A to B).

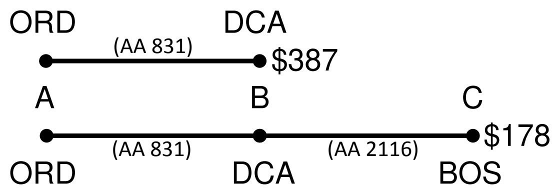
Figure 1 presents an example of HCT on American Airlines. In this example, the A-B route is a nonstop trip from Chicago O'Hare to Reagan National (DCA) in Washington, D.C. The A-C route is a connecting trip from Chicago O'Hare (city A) to Boston (city C) that connects at DCA (i.e., "hidden-city" B). In this instance, the price of the connecting A-B-C trip ($178) is $209 cheaper than the price of the nonstop A-B trip ($387). Hence, passengers whose final destination is Washington, D.C. will save money if they purchase the connecting Chicago to Boston trip and then, after deplaning at DCA, ending their journey by not boarding the second flight to Boston.

Although passengers can save money by purchasing hidden-city tickets, only one-way passengers are eligible to take advantage of these opportunities. For instance, failure to show up for the second flight on the outbound portion of a roundtrip will result in the cancellation

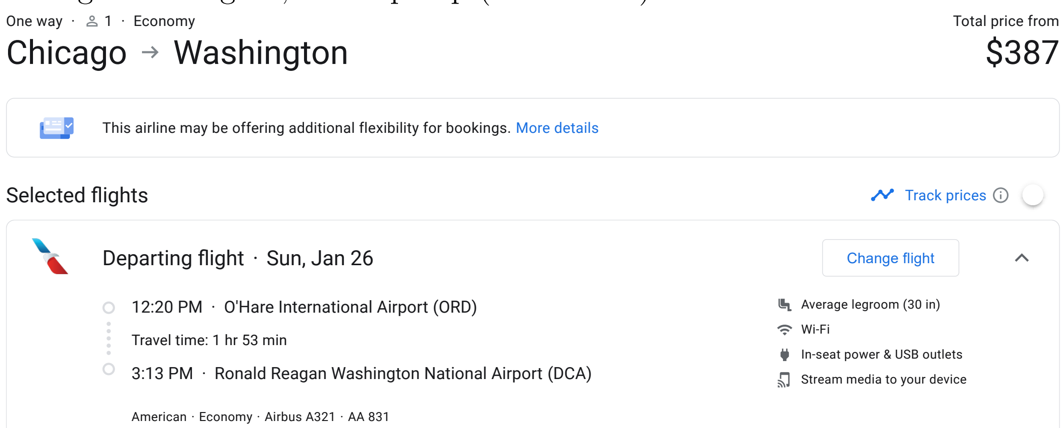
[1]HCT is also referred to as "skiplagging". For a comprehensive review on different aspects of HCT, see Meire and Derudder (2022). For some theory behind the cause and impact of HCT, see Wang and Ye (2016) and Oh and Huh (2022).

[2]For a review on bundle pricing, see Venkatesh and Mahajan (2009).

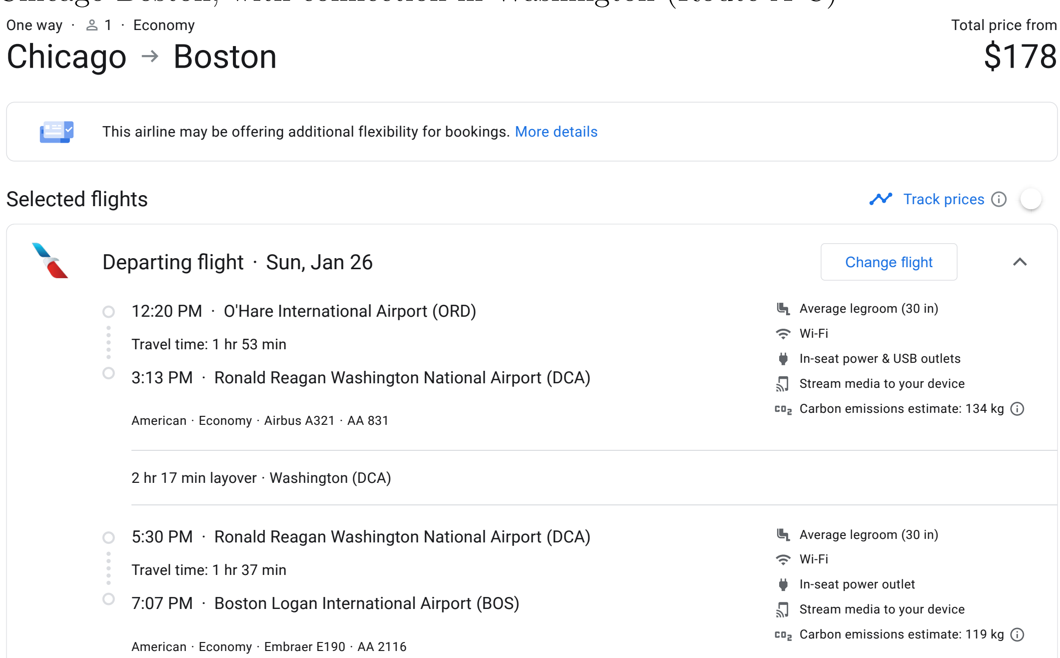
Figure 1: Example of Hidden-City ticketing



Chicago-Washington, nonstop trip (Route A-B)



Chicago-Boston, with connection in Washington (Route A-C)



of the rest of the roundtrip ticket. In addition, only passengers with carry-on luggage may engage in HCT because checked luggage will not be transferred to baggage claim at the connecting city on a hidden-city ticket.

There is another key risk that prospective HCT passengers should be aware of. Specifically, most airlines prohibit HCT in their contract of carriage (e.g., American, Delta, and United explicitly state that a passenger must complete all segments of a purchased ticket). As a result, passengers engaging in HCT may suffer retaliatory consequences including receiving a lifetime ban from the airline or having their frequent flyer membership revoked. In rare instances, airlines have even sued HCT passengers.[3]

Even though there are risks associated with HCT, the focus of this article is on the potential factors (e.g., network, route, and ticket characteristics) that contribute to the existence of HCT opportunities. One obvious factor is the extensive hub-and-spoke network structure of the large full-service carriers (e.g., American, Delta, and United). By funneling passengers through a hub, carriers are able to exploit economies of traffic density, resulting in a lower cost per passenger (Caves et al., 1984; Brueckner et al., 1992; Brueckner and Spiller, 1994). However, by controlling a large fraction of flights and gates at their hubs, carriers are also able to exercise market power and charge a "hub premium" to passengers who originate or terminate their trips at the hub (Borenstein, 1989; Lederman, 2008; Ciliberto and Williams, 2010; Escobari, 2011; Bilotkach and Pai, 2016). In other words, fares for A-B trips may be high due to the hub premium while fares for A-B-C trips may be low due to the density savings that are passed on to passengers who connect or "flow through" the hub.

A second factor that likely contributes to the existence of HCT opportunities is an airline's yield management strategy.[4] For example, airlines employ a variety of mechanisms (e.g., advance-purchase requirements and other ticket restrictions such as Saturday night stay, minimum stay, and non-refundability) to segment passengers with different price elasticities of demand.[5] All else equal, HCT opportunities will likely arise if passengers on the A-C route

[3]For example, Lufthansa sued a passenger in 2019 for missing the last leg of his ticketed journey. See `https://www.cnn.com/travel/article/lufthansa-sues-passenger-scli-intl/index.html`.

[4]For background on airline yield management practices, see Talluri et al. (2004) and Belobaba (2009).

[5]For specific examples of price discrimination in the airline industry, see Dana (1998), Stavins (2001), Bischoff et al. (2011), Puller and Taylor (2012), Aslani et al. (2014), Escobari and Jindapon (2014), Wang

are more price-elastic (i.e., have a higher price elasticity of demand) than passengers on the A-B route.

In the sections that follow, we examine how various route and ticket characteristics affect the prevalence of HCT opportunities. Related to the first factor mentioned above, we hypothesize that the level of competition within an airline's hub-and-spoke network is a key driver of HCT. In particular, the level of competition on A-B and A-C routes should have countervailing effects on the frequency of HCT opportunities. Since HCT occurs when the nonstop A-B fare is more expensive than the connecting A-C fare (i.e., $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$), additional competition on A-C routes should decrease Fare_{AC}, increasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. In contrast, additional competition on A-B routes should decrease Fare_{AB}, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds.

In addition to competition, we hypothesize that advance-purchase requirements are another key driver of HCT. Assuming that passengers who purchase tickets closer to departure are more price-inelastic and have higher search costs than passengers who book further in advance, then HCT opportunities are expected to be more frequent closer to departure. In other words, passengers who purchase tickets further in advance are more likely to seek out HCT opportunities given their low search costs and high price elasticity. In contrast, passengers who purchase tickets closer to departure are less likely to seek out these opportunities given their high search costs and low price elasticity. Armed with this knowledge of the customer base, airlines may respond by ensuring that HCT opportunities are scarce during the early booking period.

Although HCT opportunities may be common in hub-and-spoke networks,[6] the lack of sufficient data has likely been the reason why we are aware of only two prior empirical studies conducted on this topic (Liu, 2020; Sun et al., 2022).[7] Sun et al. (2022) use aggregated data

and Ye (2016), Escobari et al. (2019), and Luttmann (2019b), among others.

[6]For example, a study conducted by Hopper in 2015 found that HCT opportunities exist in 26% of U.S. domestic routes. See `https://media.hopper.com/research/hidden-city-ticket-opportunities-common-think`.

[7]The Airline Origin and Destination Survey (DB1B) released by the United States Department of Trans-

at the annual level to identify airports that are more prone to HCT. The authors find that intercontinental routes involving large hubs (especially those in the Middle East and China) are major sources of HCT. Liu (2020) is the closest study to ours in terms of data structure since she examines HCT in the United States (U.S.) using fares collected two months prior to departure. She finds that HCT only occurs on airlines that employ hub-and-spoke networks. However, contrary to our sample, she considers a single departure date with corresponding fares collected only sixty days prior to departure. We take a step forward by examining the dynamics of HCT during the booking period which was not possible with the data used in Liu (2020) or Sun et al. (2022). Compared to Sun et al. (2022), our study also offers a more refined analysis at the day-flight level rather than the year-airport level.

To examine *when* and *why* HCT opportunities occur, we rely on a unique panel of over 473 thousand published fares collected over a five-month period from a major online travel agency. Flights in our sample depart between October 1st, 2019 and December 31st, 2019 and encompass many of the most densely traveled routes across the continental United States. Notably, because we track the price of both nonstop (A-B trips) and connecting trips (A-B-C trips) in the sixty-day period before departure, we are able to examine how advance-purchase requirements affect HCT opportunities.

We have four primary findings. First, the level of competition on both A-B and A-C routes are key determinants of HCT. Consistent with expectations, we find that an additional carrier providing nonstop service on the A-C route increases the likelihood of HCT by 1.8%-4.7% while an additional nonstop carrier on the A-B route decreases the likelihood of HCT by 2.2%-3.8%.[8]

portation has been used in several previous empirical studies of the airline industry (e.g., see Brueckner et al. (1992), Gerardi and Shapiro (2009), Brueckner et al. (2013), or Dai et al. (2014), among others). However, the DB1B currently does not include information on the specific flight(s) purchased or the exact purchase and departure dates (only the quarter of travel is reported). As a result, the DB1B cannot be used to examine how factors such as advance-purchase requirements affect the frequency of HCT opportunities.

[8]As we mentioned earlier, HCT occurs when the nonstop A-B fare is more expensive than the connecting A-C fare (i.e., $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$). Therefore, additional competition on A-C routes should decrease Fare_{AC}, increasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. In contrast, additional competition on A-B routes should decrease Fare_{AB}, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds.

Second, we find that advance-purchase requirements are another key determinant of HCT. In particular, HCT opportunities are more frequent in the last week before departure because nonstop A-B fares increase at a higher rate than connecting A-C fares during this period. As we previously discussed, one possible explanation for this result is related to passenger heterogeneity during the booking period. Because most passengers purchasing tickets a few days before departure are price-inelastic customers with high search costs, airlines may be less concerned about passengers seeking out HCT opportunities during this period.

Third, we find that the major full-service carriers (i.e., American, Delta, and United) are responsible for majority of HCT, while HCT opportunities are relatively rare on low-cost carriers (e.g., Frontier, JetBlue, Spirit, and Sun Country). As alluded to earlier, the hub-and-spoke network structure provides passengers with more opportunities to exploit HCT. In contrast, the business models of low-cost carriers typically do not involve funneling passengers through large connecting hubs.

Fourth, we find that HCT opportunities are 4.8%-8.3% less likely when airport C is a hub. This finding is sensible considering that airlines are able to charge a premium for trips that originate or terminate at one of their hubs. As a result, A-C fares will be high when airport C is a hub, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds.

Although the focus of this article is on the airline industry, we believe our results are applicable to other transport modes that operate using hub-and-spoke networks. For example, Eurolines, Eurostar, and FlixBus all operate hub-and-spoke networks. These companies also offer discounted fares to early purchasers. Hence, HCT opportunities may also exist in the long-distance bus and passenger rail markets.

The rest of this article is structured as follows. Section 2 describes the data sources used in the analysis. Section 3 presents a descriptive analysis of HCT. Section 4 conducts the econometric investigation of HCT. Finally, Section 5 provides concluding remarks.

2 Data

The data we use are obtained from several sources. However, the data underlying our main empirical results are obtained from two sources: fare and itinerary information from a major online travel agency (OTA) and supplementary airline data from the U.S. Department of Transportation (DOT). Section 2.1 describes our primary source of fare and itinerary data, Section 2.2 the data sources used to construct instrumental variables, and Section 2.3 the source of our transacted fare data. Finally, Appendix Table A1 provides summary statistics and a brief description of the variables included in our empirical analysis.

2.1 Fare and Itinerary Data

Our primary source of fare and itinerary data information comes from a major OTA.[9] From the OTA, one-way economy-class fare quotes for both nonstop and connecting trips were obtained for flights departing between October 1st, 2019 and December 31st, 2019.[10] Our data encompasses over 100 of the most densely traveled routes in the continental U.S.[11] For each route, the lowest observed economy-class fare for each of the next sixty travel days were collected. This data collection procedure allows us to track the evolution of economy fares for an individual flight (or pair of flights for connecting trips) over the sixty-day period before departure.[12]

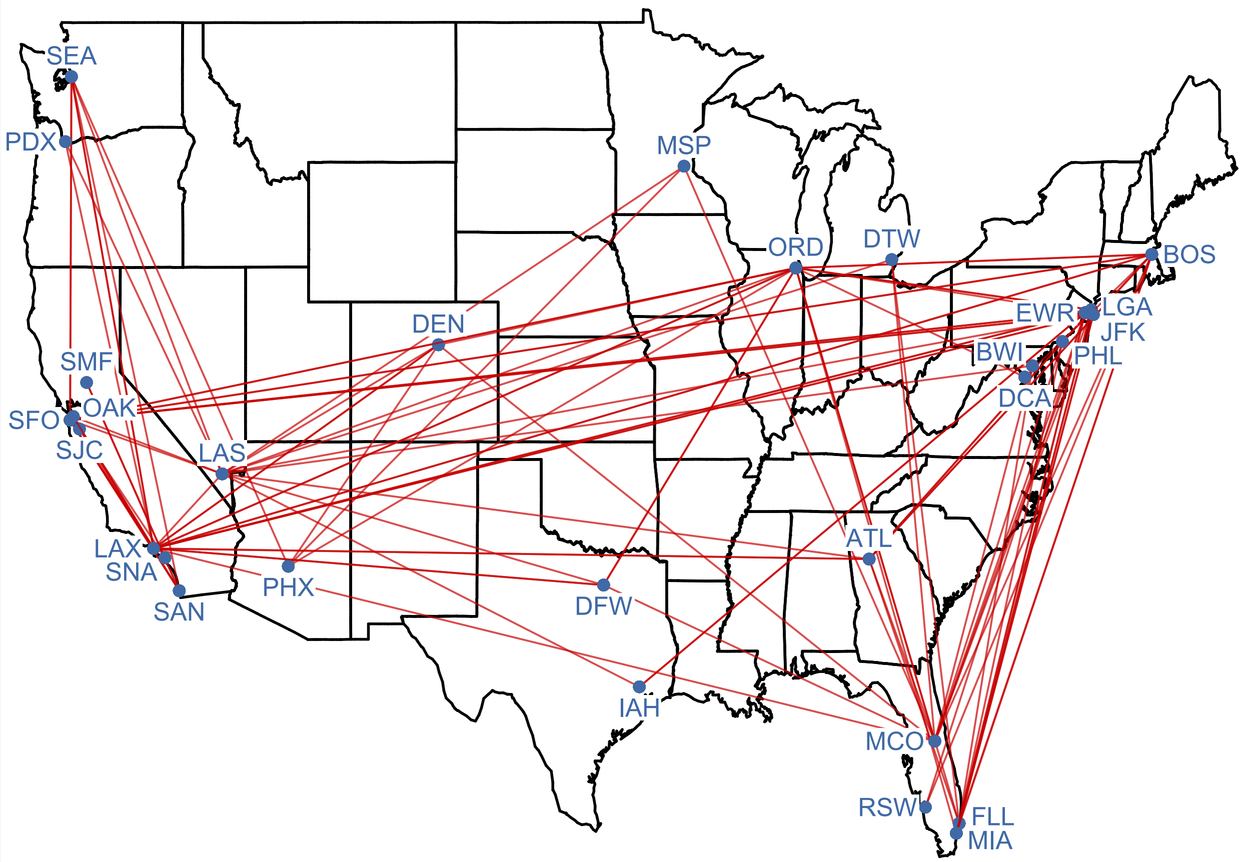
[9]Major OTAs include Expedia, Google Flights, Kayak, and Priceline. Several previous studies have relied on data from a major OTA. Among others, see Bergantino and Capozza (2015), Bilotkach et al. (2015), Escobari (2012), Gaggero and Piga (2010), Gaggero and Piga (2011), Koenigsberg et al. (2008), and Luttmann (2019a).

[10]Roundtrips are not included because only one-way passengers can take advantage of HCT opportunities. Because our analysis sample ends on December 31st, 2019, the COVID-19 pandemic does not impact our results. In the U.S., COVID-19 was declared a national emergency on March 13th, 2020. Moreover, California became the first state to issue a statewide stay-at-home order on March 19th, 2020.

[11]In lieu of collecting published fares for all possible routes in the U.S. market, we relied on the DOT's Airline Origin and Destination Survey from the third and fourth quarters of 2018 to identify the major airport-pairs within the continental U.S. ranked by total passenger traffic. A market in our analysis is defined as a directional pair of origin and destination airports. Therefore, Los Angeles (LAX)-New York City (JFK) and JFK-LAX are treated as separate markets.

[12]Roughly 60% of passengers purchase tickets in the sixty-day period before departure. For example, see Table 4 in Aryal et al. (2023) or Figure 1 in Williams (2022).

Figure 2: Hidden-City routes (i.e., A-B routes) in our analysis sample



To determine if HCT occurs within an airline-route combination on a given day, we matched the fare for each of our one-stop connecting trips (A-B-C trips) with the corresponding nonstop fare (A-B trips) for the first segment of the connecting trip. Our resulting dataset contains 473,642 fare observations. The airlines included in our sample include four full-service carriers (Alaska, American, Delta, and United) and four low-cost carriers (Frontier, JetBlue, Spirit, and Sun Country).[13] The total number of A-B routes in our sample is 101. Figure 2 presents a visual representation of the these routes (see Table 2 in Section 3 for the complete list).

[13]Although fare quotes for Southwest Airlines are not available from any of the major OTAs, the presence of Southwest is accounted for in our empirical analysis when we construct any variable controlling for the number of carriers serving a given route.

2.2 Instrumental Variables

In general, measures of market concentration such as the number of competitors or the Herfindahl–Hirschman Index are endogenous in analyses of airline pricing. For instance, markets with high fares may be attractive for new entrants. At the same time, these markets may be unattractive if high fares are a direct result of entry barriers such as limited slot or gate access at the endpoint airports. Accordingly, the potential simultaneity bias that results from an airline's decision to enter or exit a given route may bias coefficient estimates in regressions of airline pricing. To correct for this potential endogeneity, we employ an instrumental variables strategy (see Section 4 for specific details).

To construct our instruments, we rely on data from the U.S. DOT and the U.S. Census Bureau. From the U.S. DOT's T-100 Domestic Segment database, we retrieved the total number of nonstop passengers on each route and month between October 2018 and December 2018. From the U.S. Census Bureau, we obtained yearly population measures at the metropolitan statistical area for each endpoint airport in our analysis sample.

2.3 Transacted Fare Data

There exists substantial uncertainty regarding whether passengers actually exploit HCT in the U.S. domestic market. To demonstrate that a subset of passengers are likely engaging in HCT, we rely on transacted fare data from the U.S. DOT's Airline Origin and Destination Survey (DB1B). These data are released quarterly and represent a 10% random sample of tickets purchased for domestic air travel. To capture the same time period as our published fare data, we rely on DB1B data from the fourth quarter of 2019.

3 Descriptive Analysis

As discussed in Section 2.1, we are able to identify if a HCT opportunity occurs on a given day by matching an airline's connecting A-B-C fare with the airline's nonstop fare for the first segment of the connecting trip (A-B segment). HCT occurs if the fare for the connecting

A-B-C trip is cheaper than the nonstop fare for the A-B trip on the same airline.

Table 1 displays the probability of observing HCT across the four full-service and four low-cost carriers in our sample. Across all carriers, HCT occurs 13.8% of the time (21.7% on full-service carriers and 3.1% on low-cost carriers). This finding is consistent with a U.S. Government Accountability Office report from 2001. Analyzing fare data for six major U.S. airlines across 2,302 markets, GAO (2001) found that HCT opportunities occur approximately 17% of the time.

In Table 1, the three largest full-service carriers (American Airlines, Delta, and United) are responsible for the majority of HCT, as they jointly account for almost 91% of the instances of HCT observed in our sample.

Notably, we find that HCT rarely occurs on Frontier or JetBlue and almost never occurs on Sun Country (a small low-cost carrier). Nevertheless, these findings are expected. HCT opportunities are more likely to occur on carriers that operate large hub-and-spoke networks (e.g., American, Delta, and United) while they are less likely to occur on carriers that operate point-to-point networks (e.g., Frontier, JetBlue, Spirit, and Sun Country).

Table 1: Probability of HCT by airline

Airline	Type of airline	HCT	Total observations
Alaska	Full-service	1.9%	27,776
American Airlines	Full-service	20.9%	115,030
Delta	Full-service	27.5%	46,577
United	Full-service	26.1%	82,948
Frontier	Low-cost	0.8%	21,321
JetBlue	Low-cost	2.6%	17,343
Spirit	Low-cost	3.5%	160,000
Sun Country	Low-cost	0.1%	2,647
Overall Full-service		21.7%	272,331
Overall Low-cost		3.1%	201,311
Overall All carriers		13.8%	473,642

To illustrate how the probability of observing HCT evolves in the sixty-day period before departure, Figure 3 displays the probability of observing HCT (denoted by a gray bar) and, when HCT occurs, the average difference between the nonstop A-B fare and the connecting A-B-C fare (denoted by the connected solid blue line). The number above each gray bar indicates the probability of observing HCT while the number below the solid blue line indicates the average fare difference. For example, the gray bar at 60 days to departure in the top panel of Figure 3 indicates that the probability of HCT occurring 60 days before departure is 11.4% and the solid blue line indicates that the average fare difference is $26. Similarly, the gray bar at 29 days to departure in the bottom panel of Figure 3 indicates that the probability of HCT occurring 29 days before departure is 10% and the solid blue line indicates that the average fare difference is $26.

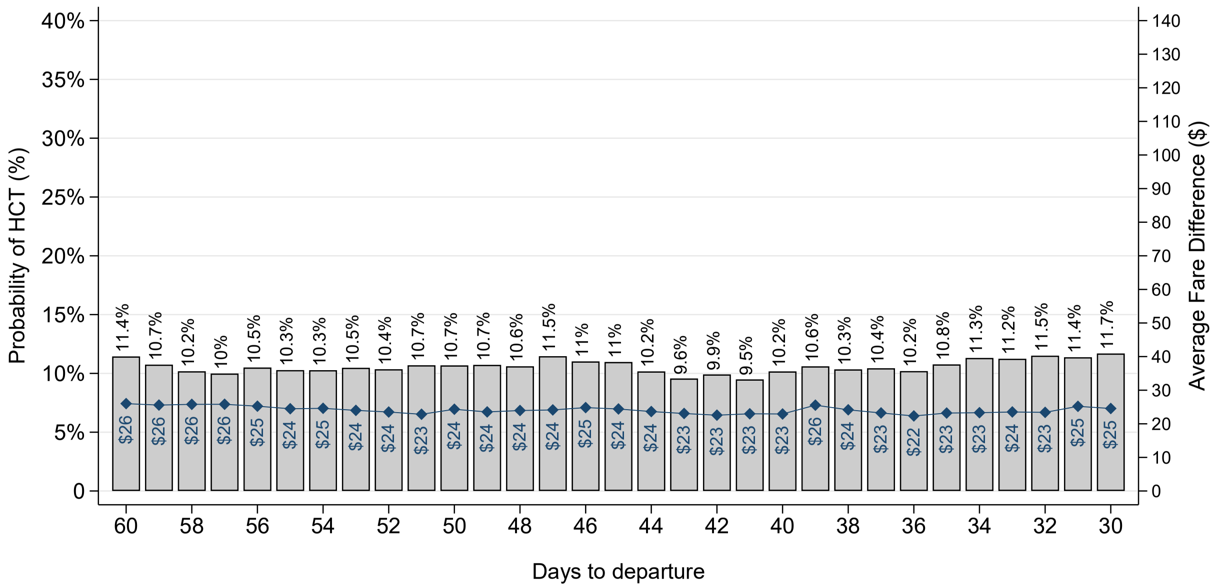
As depicted in the top panel of Figure 3, the probability of observing HCT is relatively unchanged during the early booking period, ranging from 9.5% to 11.7%. The likelihood of observing HCT remains relatively stable until two weeks before departure, when the probability of observing HCT begins to increase monotonically from 10.4% fourteen days before departure to 37.8% one day before departure.

Similarly, the average fare difference between the nonstop A-B fare and the connecting A-B-C fare, which is computed only under HCT and depicted by the connected solid blue line in the figure, is generally constant in the early part of the booking period, hovering around $24 until three weeks to departure. Then, the average fare difference increases to about $40 between two and three weeks to departure, and continues to increase until reaching a maximum of $134 three days before departure.

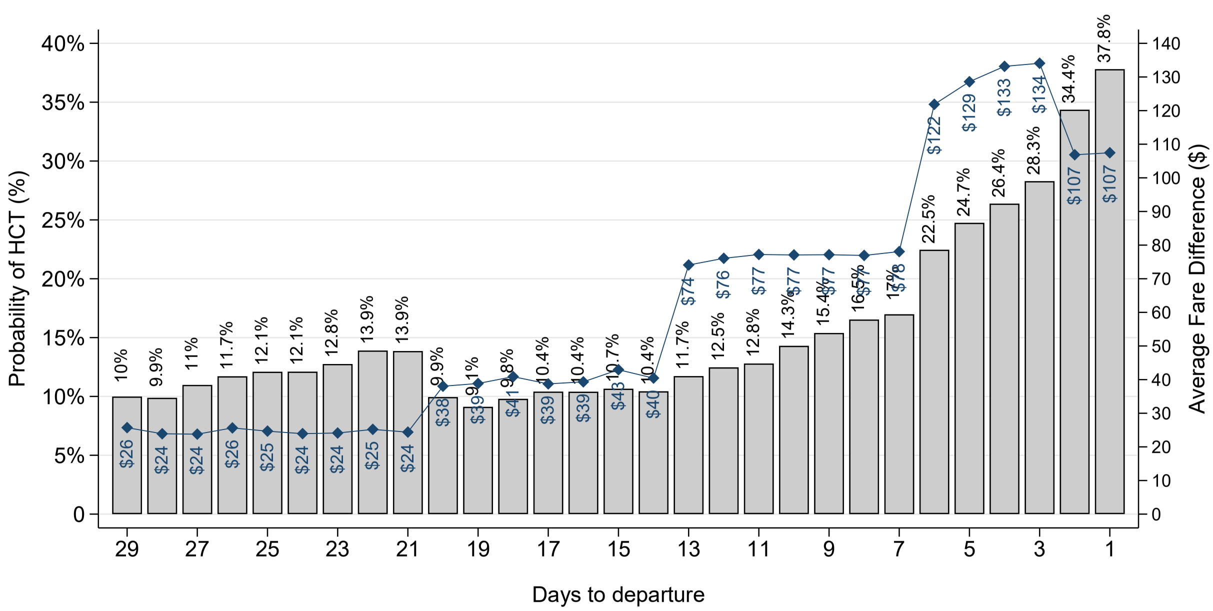
To provide a comprehensive summary of the hidden-city routes in our sample, Table 2 reports each A-B route (first column) with the probability of observing HCT on the route (second column). The last column of the table, displays the final destination(s) of the A-B-C tickets sorted in descending order by the percentage of HCT observed for each destination C on the given A-B route. For example, EWR-MIA, the last entry in the first panel of Table 2, may be the first leg of a connecting trip to Los Angeles (LAX), Orlando (MCO), or Chicago (ORD). Considering all fare observations from Newark (EWR) to one of these three destinations with a connection in Miami (MIA), the probability of observing HCT on EWR-

Figure 3: Probability of HCT and average fare difference during the booking period

(a) Early booking period



(b) Late booking period



Notes: The Average Fare Difference is the difference between the average nonstop fare from A to B and the average connecting fare from A to C with a connection at B. This difference is computed only under HCT instances (i.e., when Nonstop Fare_{AB} > Connecting Fare_{AC}).

MIA is 13%. However, if we only consider MCO (i.e., we only select the trips from EWR to MCO with connection in MIA), the probability of observing HCT is 50%. In particular, no instances of HCT are observed for connecting trips from EWR via MIA to the other two destinations of LAX and ORD.

The most common hidden-city route in our sample is Chicago O'Hare to Reagan National in Washington, D.C. (ORD-DCA). HCT occurs 87% of the time on ORD-DCA, and within this route, Boston is the most likely final destination on a hidden-city ticket. However, this finding is not entirely surprising considering that DCA is a hub for American Airlines.

Table 2: A-B routes and probability of HCT

A-B routes	HCT	Final destinations C, sorted by percentage instances of HCT within each final destination in parentheses
ATL-BOS	0%	LAS(0%), LAX(0%), LGA(0%)
ATL-FLL	1%	MCO(5%), BOS(1%), LGA(0%), LAS(0%), LAX(0%)
ATL-LAS	5%	FLL(100%), LAX(4%)
ATL-LAX	36%	LAS(36%)
ATL-LGA	60%	BOS(60%)
ATL-MCO	25%	FLL(55%), BOS(3%), LGA(1%), LAS(0%), LAX(0%)
BOS-ATL	22%	DCA(56%), ORD(50%), FLL(36%), RSW(26%), MCO(25%), MIA(21%), SFO(3%), LAX(2%)
BOS-DCA	15%	ORD(47%), MCO(22%), MIA(16%), RSW(5%), ATL(5%), SFO(3%), FLL(2%), LAX(0%)
BOS-FLL	9%	ORD(39%), ATL(11%), MCO(7%), SFO(3%), DCA(0%), LAX(0%)
BOS-LAX	21%	SFO(21%)
BOS-MCO	19%	ORD(48%), FLL(20%), ATL(1%), DCA(0%), LAX(0%)
BOS-MIA	15%	ATL(50%), MCO(11%), LAX(0%), SFO(0%)
BOS-ORD	2%	RSW(6%), FLL(5%), ATL(3%), MIA(2%), SFO(2%), LAX(1%), MCO(0%)
BOS-RSW	0%	ORD(0%)
BOS-SFO	41%	LAX(41%)
BWI-FLL	7%	MCO(7%), LAS(6%)
BWI-LAS	0%	FLL(0%)
BWI-MCO	1%	FLL(1%), LAS(0%)
DEN-LAS	2%	PHX(14%), LAX(2%), MCO(0%)
DEN-LAX	55%	PHX(67%), LAS(54%)
DEN-MCO	0%	LAS(0%)
DEN-PHX	38%	LAS(47%), LAX(26%)
DFW-LAS	9%	LAX(10%), ORD(6%), MCO(3%)
DFW-LAX	12%	ORD(89%), LAS(12%), MCO(0%)
DFW-MCO	38%	ORD(69%), LAS(0%), LGA(0%)
DFW-ORD	22%	LGA(32%), MCO(1%), LAS(0%), LAX(0%)
DTW-FLL	7%	MCO(7%), LAS(0%)
DTW-LAS	11%	FLL(12%), MCO(0%)
DTW-MCO	1%	FLL(1%), LAS(0%)
EWR-FLL	8%	LAX(13%), IAH(9%), MCO(7%), ORD(3%), SFO(1%)
EWR-IAH	12%	RSW(81%), SFO(44%), LAX(10%), MCO(7%), ORD(2%), FLL(0%), MIA(0%), PBI(0%)
EWR-LAX	6%	SFO(6%)
EWR-MCO	6%	ORD(18%), FLL(5%), IAH(2%), LAX(0%), SFO(0%)
EWR-MIA	13%	MCO(50%), LAX(0%), ORD(0%)

Continuing

Table 2 Cont.

A-B routes	HCT	Final destinations C, sorted by percentage instances of HCT within each final destination in parentheses
EWR-ORD	28%	FLL(85%), IAH(55%), RSW(52%), PBI(44%), MCO(38%), MIA(34%), SFO(12%), LAX(8%)
EWR-RSW	20%	ORD(20%)
EWR-SFO	15%	LAX(15%)
IAH-LAS	0%	EWR(0%)
JFK-FLL	0%	MCO(0%), LAS(0%), LAX(0%), SFO(0%)
JFK-LAS	13%	LAX(13%), SFO(8%)
JFK-LAX	21%	SFO(21%), LAS(21%)
JFK-MCO	1%	LAX(1%), FLL(1%)
JFK-MIA	16%	MCO(16%), SFO(0%)
JFK-SFO	14%	LAX(34%), LAS(10%)
LAX-ATL	36%	DFW(100%), LAS(100%), BOS(82%), MCO(62%), EWR(60%), ORD(40%), JAX(29%), JFK(0%)
LAX-BOS	13%	JFK(40%), EWR(1%), JAX(0%), ATL(0%), MCO(0%), ORD(0%)
LAX-DEN	12%	DFW(45%), SFO(39%), BOS(22%), ATL(14%), EWR(14%), MCO(11%), JAX(3%), ORD(2%), SEA(1%), LAS(0%), OAK(0%)
LAX-DFW	11%	ORD(39%), BOS(21%), ATL(17%), MCO(16%), EWR(8%), JFK(8%), JAX(7%), DEN(0%), OAK(0%), SEA(0%)
LAX-EWR	48%	ATL(91%), MCO(81%), BOS(49%), JAX(8%)
LAX-JFK	13%	MCO(68%), BOS(62%), ATL(37%), JAX(2%)
LAX-LAS	1%	DFW(3%), OAK(2%), ORD(1%), DEN(1%), JAX(0%), SEA(0%), EWR(0%), ATL(0%), BOS(0%), JFK(0%), MCO(0%), SFO(0%)
LAX-MCO	13%	JFK(29%), ATL(13%), EWR(2%), BOS(0%), DEN(0%), JAX(0%)
LAX-OAK	0%	LAS(0%), ORD(0%)
LAX-ORD	26%	LAS(100%), DEN(67%), BOS(40%), MCO(32%), ATL(30%), EWR(13%), JAX(5%), JFK(3%), DFW(2%), OAK(0%), SEA(0%)
LAX-SEA	4%	ATL(83%), ORD(39%), DEN(38%), MCO(35%), BOS(4%), OAK(0%), DFW(0%), EWR(0%), JFK(0%)
LAX-SFO	3%	LAS(48%), DFW(33%), BOS(1%), ATL(1%), MCO(1%), SEA(1%), EWR(0%), DEN(0%), JAX(0%), JFK(0%), ORD(0%)
LGA-ATL	24%	MIA(27%), MCO(9%), FLL(9%), ORD(0%)
LGA-FLL	29%	ORD(48%), ATL(32%), MCO(22%)
LGA-MCO	4%	ORD(28%), FLL(2%), ATL(2%), MIA(0%)
LGA-MIA	22%	ATL(44%), MCO(19%), ORD(0%)
LGA-ORD	16%	FLL(40%), MCO(26%), MIA(16%), ATL(4%)
MSP-LAS	0%	PHX(0%), MCO(0%)
MSP-MCO	17%	LAS(17%)
MSP-PHX	6%	LAS(6%)
OAK-LAS	1%	LAX(2%), SAN(1%), BUR(0%)
OAK-LAX	1%	LAS(1%), SAN(0%)
ORD-BOS	0%	LGA(3%), FLL(0%), DCA(0%), DEN(0%), DFW(0%), LAS(0%), LAX(0%), MCO(0%), MIA(0%), SFO(0%)
ORD-DCA	87%	BOS(100%), LGA(87%), MCO(84%), MIA(80%), SFO(67%), FLL(60%)
ORD-DEN	5%	PHX(18%), LAX(16%), LGA(3%), LAS(3%), DCA(1%), SFO(1%), DFW(1%), MIA(0%), FLL(0%), MCO(0%)
ORD-DFW	34%	SFO(69%), MIA(60%), LAX(46%), MCO(42%), DEN(42%), LAS(36%), FLL(31%), BOS(15%), PHX(12%), LGA(0%)
ORD-FLL	5%	BOS(11%), MCO(6%), LGA(6%), DEN(6%), DCA(0%), DFW(0%), LAS(0%), LAX(0%), PHX(0%)
ORD-LAS	16%	LAX(20%), DFW(6%), DEN(5%), FLL(5%), MCO(0%), PHX(0%), SFO(0%)
ORD-LAX	14%	PHX(15%), SFO(14%), LAS(13%), DFW(0%)

Continuing

Table 2 Cont.

A-B routes	HCT	Final destinations C, sorted by percentage instances of HCT within each final destination in parentheses
ORD-LGA	31%	BOS(71%), DFW(46%), MCO(36%), DCA(30%), MIA(27%), FLL(16%), DEN(0%)
ORD-MCO	18%	MIA(43%), BOS(35%), DEN(9%), DCA(7%), FLL(3%), DFW(1%), LAS(0%), LGA(0%), PHX(0%)
ORD-MIA	14%	MCO(28%), DCA(0%), DEN(0%), LAS(0%), PHX(0%), SFO(0%)
ORD-PHX	63%	LAX(76%), LAS(54%), SFO(47%), DFW(0%)
ORD-SFO	70%	LAS(84%), LAX(42%), PHX(41%), DEN(0%)
PDX-LAS	1%	LAX(4%), FLL(0%)
PDX-LAX	3%	LAS(12%), FLL(0%)
PHL-FLL	18%	MCO(18%)
PHL-MCO	7%	FLL(7%), SNA(0%)
SAN-SFO	61%	SMF(61%)
SAN-SMF	0%	OAK(0%)
SEA-LAS	7%	LAX(13%), SAN(1%), SFO(1%), PHX(0%)
SEA-LAX	56%	SAN(81%), LAS(65%), PHX(41%), SFO(35%)
SEA-PHX	36%	SFO(43%), SAN(43%), LAX(19%), LAS(14%)
SEA-SAN	0%	SFO(0%)
SEA-SFO	37%	SAN(38%), LAS(38%), PHX(36%), LAX(36%)
SFO-BOS	19%	EWR(22%), JFK(3%), ORD(0%)
SFO-EWR	21%	BOS(21%), ORD(0%)
SFO-JFK	26%	BOS(26%), ORD(17%)
SFO-LAS	1%	SEA(2%), ORD(1%), BDL(0%), BOS(0%), EWR(0%), JFK(0%), LAX(0%)
SFO-LAX	5%	LAS(40%), SAN(11%), SEA(4%), BDL(1%), JFK(1%), ORD(0%), BOS(0%), EWR(0%)
SFO-ORD	19%	EWR(49%), JFK(47%), BOS(33%), BDL(9%)
SFO-SAN	0%	BDL(0%), BOS(0%), EWR(0%), ORD(0%), SEA(0%)
SFO-SEA	1%	BOS(2%), ORD(1%), BDL(0%), EWR(0%), JFK(0%)
SJC-SAN	0%	SNA(0%)
SMF-SAN	0%	BUR(0%), SNA(0%)
SNA-SJC	0%	MCO(0%)

To provide additional context on the routes that are more affected by HCT, Table 3 displays means and standard deviations for several key route characteristics. These averages are computed when HCT occurs (column one) and when HCT is absent (column two). Column three displays the difference in means between columns one and two.

As hypothesized in Section 1, the number of competitors providing nonstop service on A-B and A-C routes are expected to have differential impacts on the likelihood of HCT. Specifically, additional competition on A-B routes should decrease Fare_{AB}, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. In contrast, additional competition on A-C routes should decrease Fare_{AC}, increasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. The means displayed in the first two rows of Table 3 are consistent with these expectations. In our sample, HCT is more prevalent when there are fewer carriers providing nonstop service on the A-B route (-0.40 fewer carriers on average) and when there are more carriers providing nonstop service on the A-C route (0.63 more carriers on average).

Given that American, Delta, and United account for the majority of HCT in our sample (see Table 1), the number and location of hubs on A-B-C trips are also expected to affect the prevalence of HCT opportunities. As Table 3 demonstrates, there are an average of 1.66 hubs involved on an A-B-C trip when HCT occurs compared to 0.97 hubs when HCT is absent.[14] This difference is also maintained when examining whether airports A, B, or C are hubs. When HCT occurs, airport A is a hub 62.6% of the time (compared to 34.4% when HCT is absent), airport B a hub 86.9% of the time (compared to 51.3% when HCT is absent), and airport C a hub 16.2% of the time (compared to 11.8% when HCT is absent). As expected, the largest mean value across the hub indicators is observed for Hub B, indicating that HCT opportunities are more likely when airport B (i.e., the hidden-city) is a hub. This finding is sensible considering that the "hub premium" will tend to increase fares on A-B routes if airport B is a hub, increasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. Applying a

[14]The hub airports for each airline are: Alaska (ANC, LAX, PDX, SEA, SFO), American (CLT, DCA, DFW, JFK, LAX, LGA, MIA, ORD, PHL, PHX), Delta (ATL, BOS, DTW, JFK, LAX, LGA, MSP, SEA, SLC), Frontier (DEN), Sun Country (MSP), and United (EWR, DEN, IAD, IAH, LAX, ORD, SFO).

similar argument, it is also sensible that the smallest mean value across the hub indicators is observed for Hub C. If airport C is a hub, the hub premium will tend to increase fares on A-C routes, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds.

The last few rows of Table 3 display the average nonstop distance between airports A and B (Distance A-B) and airports A and C (Distance A-C). In our sample, the average length of the A-B route is 1,298 miles when HCT occurs compared to 977 miles when HCT is absent. However, the mean A-C distance is similar with and without HCT, averaging 1,687 miles under HCT and 1,641 without HCT. Finally, $\mathbb{1}$(Dist. A-C > Dist. A-B) is a dummy variable that indicates whether the A-C route is longer than the A-B route. When HCT occurs, the A-C distance is longer than the A-B distance 84.2% of the time compared to 80.9% of the time in the absence of HCT.

Table 3: HCT and route characteristics

	HCT=1 (1)	HCT=0 (2)	Difference in means (3)
Competition A-B	3.551	3.953	-0.402
	(1.020)	(1.411)	
Competition A-C	2.987	2.355	0.632
	(1.523)	(1.646)	
Number of Hubs	1.656	0.974	0.682
	(0.754)	(0.942)	
Hub A	0.626	0.344	0.282
	(0.484)	(0.475)	
Hub B	0.869	0.513	0.356
	(0.338)	(0.500)	
Hub C	0.162	0.118	0.044
	(0.368	(0.322)	
Distance A-B, in miles	1,298	977	321
	(621)	(598)	
Distance A-C, in miles	1,687	1,641	46
	(701)	(808)	
$\mathbb{1}$(Dist. A-C > Dist. A-B)	0.842	0.809	0.033
	(0.365)	(0.393)	

Notes: Average values, standard deviation in parentheses. *Competition A-B* (*Competition A-C*) refers to the number of nonstop carriers serving the A-B (A-C) route on the itinerary's departure date.

4 Econometric Analysis

We aim to accomplish two primary objectives with our econometric analysis. Foremost, we wish to understand the main drivers of HCT (Section 4.2). Second, we would like to determine the possible savings that a passenger gains from HCT, but also the potential loss that an airline incurs if a passenger engages in HCT (Section 4.3). Before doing so, we first show that passengers are likely exploiting HCT opportunities in the U.S. market (Section 4.1).[15]

4.1 Exploitation of HCT

To determine if a subset of U.S. passengers are likely taking advantage of HCT opportunities during our sample period, we rely upon transacted fare data provided in the DOT's DB1B database. As discussed in Section 2.3, these data are released quarterly and represent a 10% random sample of all airline tickets purchased for travel in the domestic U.S. market. For the best correspondence of the DB1B with the time period of our published fare and itinerary data (October 2019–December 2019), we use DB1B data from the fourth quarter of 2019.[16]

Although the DB1B does not provide information on the specific date each ticket was purchased and, more importantly, the actual flight(s) each passenger boarded, it is still possible to test for potential exploitation of HCT by passengers. Specifically, we assume that a passenger cannot exploit HCT on a roundtrip ticket, since failure to show up for the second leg of the outbound portion of the trip typically results in cancellation of the rest of the roundtrip ticket. Therefore, to exploit HCT, a roundtrip passenger would need to purchase two separate one-way tickets.

[15]We are grateful to Jan Brueckner for this insightful suggestion.

[16]To prevent outliers from affecting results, we exclude tickets with prices below the 5^{th} and above the 95^{th} percentiles.

Based on this idea, we estimate the following regression,

$$HCT\%_{rca} = \beta_0 + \beta_1 \cdot AveragePriceDifference_{rca} + \delta_a + \epsilon_{rca} \tag{1}$$

where the main independent variable of interest is the average price difference (*AveragePriceDifference*$_{rca}$), computed as the difference between $\text{AverageFare}_{\text{AB}}$, the average one-way nonstop fare on airline a and route r (i.e., A-B routes), and $\text{AverageFare}_{\text{AC}}$, airline a's average one-way connecting fare that uses route r to a given final destination c (i.e., A-B-C routes).[17] The dependent variable ($HCT\%_{rca}$) is the percentage of tickets on route r, airline a, and final destination c that were purchased on a one-way basis, so as to exploit HCT. Formally,

$$HCT\%_{rca} = \frac{One\text{-}Way\ Tickets_{rca}}{One\text{-}Way\ Tickets_{rca} + One\text{-}Way\ Tickets_{ra}} \tag{2}$$

where $One\text{-}Way\ Tickets_{rca}$ is the number of one-way connecting tickets (i.e., A-B-C tickets) and $One\text{-}Way\ Tickets_{ra}$ is the number of one-way nonstop tickets (i.e., A-B tickets).

A positive coefficient on β_1 in equation (1) would indicate that an increase in the difference between the average one-way nonstop fare and the average one-way connecting fare is associated with an increase in the number of tickets purchased on a one-way basis (i.e., an increase in the number of passengers potentially exploiting HCT). Airline fixed effects (δ_a) are included as controls.

Note that β_1 may suffer from simultaneity bias because changes in the quantity of tickets sold affects fares, which is a determinant of the *AveragePriceDifference%* variable. In other words, *HCT%* may also cause the *AveragePriceDifference*. However, any bias that results will decrease the magnitude of the *AveragePriceDifference* coefficient because the sign of

[17] For example, ORD-DCA-BOS and ORD-DCA-MIA trips on American constitute two separate observations. In this example, ORD-DCA is the "A-B" route (i.e., route r) and ORD-DCA-BOS and ORD-DCA-MIA are two separate "A-C" routes that use route r (i.e., rc routes). Because *AveragePriceDifference* is intended to measure the savings from HCT, *AveragePriceDifference* is set to zero in the case of negative values (i.e., when HCT does not occur).

the bias is negative.[18] Since we are only interested in documenting the positive correlation between *AveragePriceDifference* and $HCT\%$ to demonstrate that a subset of passengers may be engaging in HCT, a biased β_1 that is estimated to be positive would be, a fortiori, still positive if the simultaneity bias were corrected.

The results of estimating equation (1) are reported in Table 4. The first column displays ordinary least squares (OLS) estimates while the second column displays fractional logit estimates. Because the dependent variable is a percentage that is bounded between zero and one, our preferred estimates are the fractional logit estimates in column (2).

The positive and statistically significant coefficient on *AveragePriceDifference* in both Table 4 columns indicate that passengers are likely exploiting HCT in the U.S. domestic market. Furthermore, consistent with our Table 1 findings, the positive coefficients on American, Delta, and United indicate that HCT typically arises on full-service carriers that operate large hub-and-spoke networks (although the coefficient on United is statistically insignificant).[19]

Having established that a subset of passengers are likely exploiting HCT opportunities, we now turn our attention to examining the main drivers of HCT in Section 4.2 and the potential savings that passengers may obtain from engaging in HCT in Section 4.3.

[18] Consider a two-equation structural model where the first equation is equation (1) and the second equation is $AveragePriceDifference_{rca} = \alpha_0 + \alpha_1 HCT\%_{rca} + \delta_a + \varepsilon_{rca}$. It can be shown (e.g., see chapter 16.2 in Wooldridge, 2008) that Cov(*AveragePriceDifference*,ϵ) $= [\alpha_1/(1-\alpha_1\beta_1)]Var(\epsilon)$. Thus, the bias of the ordinary least squares (OLS) estimator of (1) has the same sign as $\alpha_1/(1-\alpha_1\beta_1)$. First, note that α_1 is negative. To prove this, start from the definition of $HCT\%$ in equation (2), which shows that the $HCT\%$ variable increases if $One\text{-}Way\ Tickets_{rca}$ goes up: if the demand for one-way connecting tickets increases, so do their prices $\text{AverageFare}_{\text{AC}}$, implying that the *AveragePriceDifference*$=\text{AverageFare}_{\text{AB}}-\text{AverageFare}_{\text{AC}}$ decreases. Second, the expectation from equation (1) is that β_1 is positive because a larger *AveragePriceDifference* will induce more passengers to exploit HCT opportunities. This implies that $(1-\alpha_1\beta_1)$ is positive and hence that $\alpha_1/(1-\alpha_1\beta_1)$ is negative, meaning that any simultaneity bias that results is negative. In other words, when we estimate equation (1) with OLS, we are likely underestimating the "true" effect of the *AveragePriceDifference* on $HCT\%$.

[19] The omitted airline fixed effect is Sun Country, a small low-cost carrier.

Table 4: Test for exploitation of HCT with DB1B data

	(1)	(2)
Estimator:	OLS	Fractional Logit
Dependent variable:	HCT%	HCT%
Average Price Difference	0.0003***	0.004***
	(0.000)	(0.000)
Alaska	-0.016***	-0.506***
	(0.006)	(0.172)
American Airlines	0.025***	0.533***
	(0.006)	(0.154)
Delta	0.022***	0.498***
	(0.006)	(0.150)
Hawaiian	-0.015***	-0.533***
	(0.006)	(0.191)
United	0.004	0.172
	(0.006)	(0.154)
Frontier	-0.019***	-0.705***
	(0.005)	(0.155)
JetBlue	-0.026***	-1.048***
	(0.006)	(0.210)
Spirit	-0.022***	-0.855***
	(0.005)	(0.153)
Southwest	-0.012**	-0.348**
	(0.005)	(0.149)
Allegiant	-0.018***	-0.682***
	(0.005)	(0.159)
R^2 or Pseudo-R^2	0.081	0.031
Observations	162,889	162,889

Notes: Data are from the DOT's DB1B database for the fourth quarter of 2019. Sun Country is the omitted airline fixed effect. Standard errors are provided in parentheses and clustered at the route A-B level. Constant is included but not reported. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

4.2 Determinants of HCT

4.2.1 Probability of Observing HCT

To determine how various route and ticket characteristics affect the prevalence of HCT opportunities, we model the probability of observing HCT as a function of route-level competition (both on A-B and A-C routes), hub airports, route distance, advance-purchase requirements, ticketing carrier, and other itinerary-specific characteristics such as the month-of-departure, day-of-the-week-of-departure, and the time-of-day-of-departure.

Specifically, we estimate equation (3) below,

$$\begin{aligned} \Pr(HCT_{ircdat} &= 1) = f(CompetitionA\text{-}B_{ird}, CompetitionA\text{-}C_{icd}, Hub_{irc}, \\ & Distance_{irc}, DaysToDeparture_{it}, Airline_{ia}, \delta_{id}) \end{aligned} \tag{3}$$

where the subscript i indexes the itinerary, r the A-B route, c the final destination for the itinerary that uses route r (i.e., the A-C route), d the departure date, a the airline, and t the time dimension, measured in the number of days to departure (i.e., how far in advance the itinerary is booked). Competition on A-B and A-C routes (*CompetitionA-B* and *CompetitionA-C*) are measured by the number of nonstop carriers serving the route on the itinerary's departure date. The effect of hub endpoints on HCT is measured using the *Hub A*, *Hub B*, and *Hub C* indicator variables that were previously defined in Table 3 while distance is accounted for using the $\mathbb{1}$(Dist. A-C > Dist. A-B) indicator that was also defined in Table 3.

To account for nonlinear fare changes that occur during the booking period, we follow Gaggero and Luttmann (2020, 2022) and split the days to departure variable into five categories: 1 to 2, 3 to 6, 7 to 13, 14 to 20, and 21 to 60; the indicator for 21 to 60 days to departure serves as the reference category. The ticketing carrier for each itinerary is represented by a separate indicator (*Airline*) with Sun Country serving as the reference category (Table 1 indicates that HCT opportunities are least prevalent on Sun Country). Finally, δ

is a matrix of fixed effects that control for each itinerary's month-of-departure, day-of-week-of-departure, and time-of-departure.

We recognize that there may exist some unobserved factor that is correlated with both the number of carriers serving A-B and/or A-C routes and the prevalence of HCT opportunities. To correct for the possible endogeneity of *Competition A-B* and *Competition A-C*, we employ a two-stage least squares (2SLS) approach with six instruments: (i) the number of nonstop passengers on route A-B during the same month of the previous year, (ii) the number of nonstop passengers on route A-C during the same month of the previous year, (iii) the natural logarithm of the arithmetic mean of the metropolitan statistical area (MSA) populations of the endpoint cities on route A-B, (iv) the natural logarithm of the arithmetic mean of the MSA populations of the endpoint cities on route A-C, (v) the natural logarithm of the geometric mean of the MSA populations of the endpoint cities on route A-B, and (vi) the natural logarithm of the geometric mean of the MSA populations of the endpoint cities on route A-C. These instruments are similar to those used in Gerardi and Shapiro (2009) and Dai et al. (2014). The rationale behind these instruments is straightforward. The past-year number of passengers and the population of the endpoint cities impact the suitability of a given route to a particular airline's fleet type and size, which directly affects an airline's route entry decision (and thus, the overall level of competition on the route). Furthermore, there is no reason to believe that city populations or previous passenger traffic levels are direct determinants of HCT.

In our baseline specification, we estimate equation (3) using 2SLS with standard errors that are clustered at the A-B route level. However, because our dependent variable is a binary indicator taking the values of zero or one, we also estimate equation (3) using instrumental variables (IV) probit.

The regression results are reported in Table 5. To ensure that the linear estimates of columns (1) and (2) are directly comparable with the output from the probit regressions, columns (3) and (4) report the marginal effects. The corresponding probit coefficients are reported in Appendix Table A2. The last two columns of Table A2 also report the first-stage estimates for *Competition A-B* and *Competition A-C*, respectively.[20] Notably, the statistically significant Kleibergen-Paap rk Wald F statistic in column (2) of Table 5 indicates that our instruments are both strong and relevant.[21]

As the Table 5 results indicate, competition is one of the primary drivers of HCT, especially on A-C routes. An additional nonstop carrier serving the A-C route increases the likelihood of HCT by 3.9%-4.7% under the linear estimates and by 1.8%-3.5% under the probit estimates. The effect of *Competition A-B* on HCT is slightly less pronounced, as the marginal effect is statistically insignificant in column (4). Considering only the statistically significant estimates, an additional nonstop carrier serving the A-B route decreases the likelihood of HCT by 2.2%-3.8% in the linear model and by 2.5% in the probit model.

The signs on both competition variables are consistent across all Table 5 specifications and in line with expectations of a negative effect of *Competition A-B* and a positive effect of *Competition A-C* on the likelihood of observing HCT. For instance, standard economic theory predicts that additional competition should result in lower market prices. Because HCT occurs when $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$, additional competition on A-C reduces Fare_{AC}, thereby increasing the likelihood that this inequality holds (expected positive sign on *Competition A-C*). In contrast, additional competition on A-B reduces Fare_{AB}, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds (expected negative sign on *Competition A-B*).

[20]The first-stage regressions for the 2SLS and IV probit models are identical (i.e., linear first-stage estimated by OLS with the same six instruments).

[21]Since we have more instruments (six) than endogenous variables (two), we are also able to test whether our overidentifying restrictions are valid using Hansen's J test. As evidenced by the statistically insignificant Hansen J statistic in column (2) of Table 5, we fail to reject the null hypothesis that our overidentifying restrictions are valid (i.e., we have valid instruments).

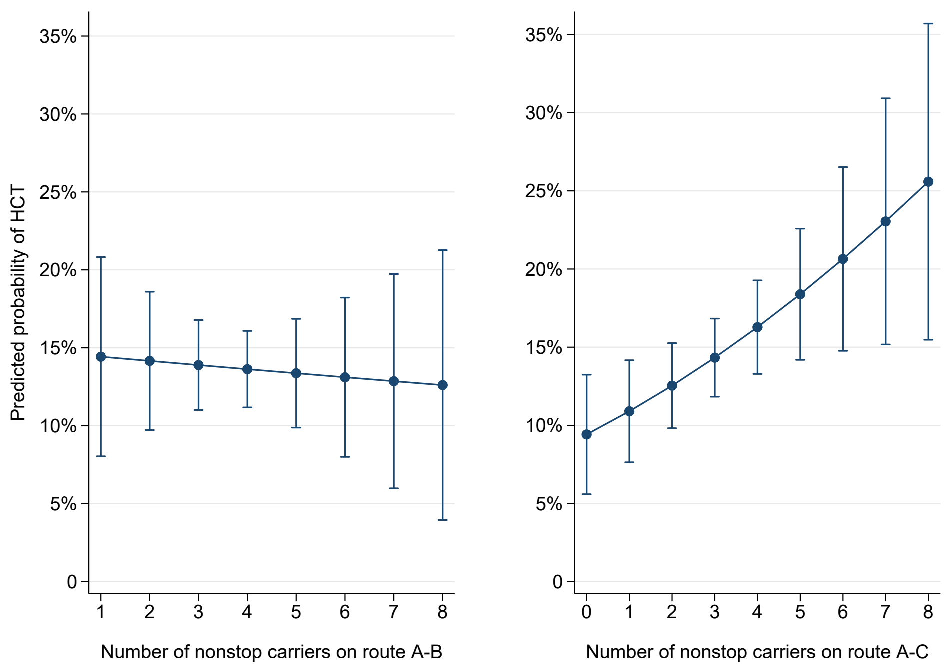
Table 5: Probability of observing HCT

	(1)	(2)	(3)	(4)
Estimator:	OLS	2SLS	Probit	IV-Probit
Dependent variable:	HCT	HCT	HCT	HCT
	Estimated coefficients	Estimated coefficients	Marginal effects	Marginal effects
Competition A-B	-0.022***	-0.038***	-0.025**	-0.003
	(0.008)	(0.013)	(0.010)	(0.011)
Competition A-C	0.039***	0.047***	0.035***	0.018***
	(0.006)	(0.007)	(0.005)	(0.006)
Hub A	0.048	0.051	0.032	0.029
	(0.050)	(0.048)	(0.032)	(0.031)
Hub B	0.036	0.017	0.017	0.003
	(0.039)	(0.036)	(0.039)	(0.036)
Hub C	-0.074*	-0.083*	-0.048*	-0.066***
	(0.041)	(0.043)	(0.025)	(0.023)
$\mathbb{1}$(Dist. A-C > Dist. A-B)	-0.029	-0.013	-0.035	-0.024
	(0.023)	(0.025)	(0.025)	(0.024)
DaysToDeparture 1-2	0.219***	0.217***	0.203***	0.206***
	(0.049)	(0.049)	(0.035)	(0.035)
DaysToDeparture 3-6	0.116***	0.116***	0.105***	0.105***
	(0.044)	(0.044)	(0.037)	(0.036)
DaysToDeparture 7-13	0.017	0.016	0.015	0.014
	(0.024)	(0.024)	(0.021)	(0.022)
DaysToDeparture 14-20	-0.017	-0.018	-0.016	-0.015
	(0.013)	(0.013)	(0.012)	(0.012)
Alaska	0.056	0.086	0.019**	0.024**
	(0.058)	(0.063)	(0.009)	(0.011)
American Airlines	0.249***	0.271***	0.202***	0.203***
	(0.062)	(0.065)	(0.031)	(0.029)
Delta	0.314***	0.343***	0.288***	0.309***
	(0.083)	(0.091)	(0.079)	(0.075)
United	0.264***	0.281***	0.209***	0.236***
	(0.070)	(0.069)	(0.039)	(0.043)
Frontier	0.049	0.081	0.009**	0.008**
	(0.066)	(0.071)	(0.003)	(0.003)
JetBlue	0.105*	0.126*	0.030	0.019
	(0.061)	(0.068)	(0.022)	(0.014)
Spirit	0.111**	0.117**	0.045**	0.038***
	(0.050)	(0.053)	(0.017)	(0.014)
Kleibergen-Paap rk LM stat.		14.198**		
Hansen J Statistic		2.885		
Kleibergen-Paap rk Wald F stat.		21.862***		
R^2 or Pseudo-R^2	0.165	0.161	0.224	
Observations	473,642	473,642	473,642	473,642

Notes: All specifications include month-of-year, day-of-week, and time-of-day-of-departure fixed effects. Sun Country is the omitted airline fixed effect. Columns (3) and (4) report the marginal effects for the Probit regressions. Probit coefficient estimates are reported in Appendix Table A2. The endogenous variables in columns (2) and (4) are *Competition A-B* and *Competition A-C* and the corresponding first-stage regressions are reported in Appendix Table A2. Standard errors are provided in parentheses and clustered at the route A-B level. Constant is included but not reported. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Using the IV-Probit estimates (our preferred specification), Figure 4 depicts the predicted probability of HCT as competition increases on A-B routes (left diagram) and A-C routes (right diagram). The bars stemming from the point estimates represent the 95% confidence interval. As the figure illustrates, the predicted probability of HCT monotonically increases as the number of nonstop carriers serving route A-C increases, in line with expectations. In contrast, as the number nonstop carriers serving the A-B route increases, the overall probability of HCT decreases. However, the slope of the line connecting the predicted probabilities is not very steep, pointing towards a relatively lower impact of *Competition A-B* on HCT, as already suggested by column (4) of Table 5.

Figure 4: Predicted probability of HCT as competition increases with 95% conf. interval



The coefficient on the airline fixed effects are consistent with the findings in Table 1, where HCT opportunities were found to be more prevalent on American, Delta, and United, the major full-service carriers in the U.S. domestic market. Relative to Sun Country, the omitted airline fixed effect in the regressions, HCT opportunities are approximately 20% more likely on American, 31% more likely on Delta, and 24% more likely on United. A smaller effect is found for Alaska. However, our sample excludes routes to Alaska (see Figure 2). In addition, Alaska's hubs are confined to cities on the west coast instead of being dispersed across the continental U.S. like the hub networks for American, Delta, and United.[22]

We believe the dispersed hub-and-spoke network structure of the three major full-service carriers provides passengers with more opportunities to exploit HCT. In contrast, HCT opportunities are less likely on low-cost carriers because their business models do not involve operating large connecting hubs. Consistent with this story, the coefficients for the low-cost carriers (Frontier, JetBlue, and Spirit) are substantially lower in magnitude than the coefficients for American, Delta, and United.

To further decompose the importance of hub-and-spoke networks, the coefficients on *Hub A*, *Hub B*, and *Hub C* in Table 5 test whether the specific location of hubs on an A-B-C itinerary affects the likelihood of HCT. Because carriers are able to charge a "hub premium" to passengers who originate or terminate their trips at a hub (Borenstein, 1989; Lederman, 2008; Ciliberto and Williams, 2010; Escobari, 2011; Bilotkach and Pai, 2016), hub location likely affects HCT opportunities. For example, if airport B is a hub, then fare levels on A-B routes will be high, increasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. In contrast, if airport C is a hub, then fare levels on A-C routes will be high, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. An interesting case occurs when airport A is a hub, since the hub premium applies to both A-B and A-C routes. In these instances, the hub status of airport A is not expected to significantly affect the likelihood of HCT.

Consistent with expectations, the coefficient on *Hub A* is statistically insignificant, the coefficient on *Hub B* is positive, and the coefficient on *Hub C* is negative in all Table 5 specifications. However, only the coefficient on *Hub C* is statistically significant. In our preferred specification in column (4), the coefficient on *Hub C* indicates that HCT is 6.6%

[22] Alaska currently has hubs at Anchorage (ANC), Los Angeles (LAX), Portland (PDX), San Francisco (SFO), and Seattle (SEA).

less likely when airport C is a hub. Although *Hub B* has the expected positive sign, we must mention that the legacy carrier fixed effects are likely absorbing some of the "Hub B effect" since almost all connecting itineraries on American, Delta, and United connect at one of the airline's hubs.

Since longer distances imply higher fares, HCT ticketing is expected to be less likely when the nonstop distance of the A-C route is greater than the nonstop distance of the A-B route. Consistent with expectations, the coefficient on $\mathbb{1}$(Dist. A-C > Dist. A-B) is negative. However, this coefficient is statistically insignificant in all Table 5 specifications.

Finally, the coefficients on the *DaysToDeparture* variables indicate that HCT opportunities are more prevalent in the last week before departure, consistent with the pattern previously displayed in Figure 3. The coefficients in Table 5 indicate that, relative to trips booked 21 to 60 days in advance, the likelihood of observing HCT increases by about 11% between three and six days before departure, and by about 21% in the last two days to departure. This finding may result from different pricing patterns of A-B and A-C fares as the departure date approaches, with a possible steeper trajectory for A-B fares. We investigate this presumption further in the next subsection.

4.2.2 Fare Regressions

To test the conjecture that the increased probability of observing HCT closer to the departure date is due to the steeper increase of nonstop A-B fares relative to connecting A-C fares, we regress the natural logarithm of fare on the same set of regressors deployed in the HCT regressions (i.e., we estimate equation (3) with the natural logarithm of fare as the dependent variable). Because the dependent variable is in logs, the estimated coefficients on the *DaysToDeparture* dummies represent the percentage change in fare relative to *DaysToDeparture 21-60*, the omitted days to departure category in the regressions.[23]

[23] Since the dependent variable is in logs and the *DaysToDeparture* variables are indicators, marginal effects are interpreted as the $100 \times (e^{\beta} - 1)\%$ change in fare.

Due to the potential endogeneity of the competition variables (see Section 2.2), we estimate our fare regressions using 2SLS with the same set of instruments used in equation (3). Table 6 reports results when the natural logarithm of the A-B fare (column 1) and A-C fare (column 2) are the dependent variables. Comparing the *DaysToDeparture 1-2*, *DaysToDeparture 3-6* and *DaysToDeparture 7-13* coefficients across columns, both coefficients are larger in magnitude when $\log(\text{Fare}_{\text{AB}})$ is the dependent variable. This finding implies that A-B fares increase at a higher rate than A-C fares, supporting the presumption that the increased likelihood of observing HCT in the last two weeks before departure is driven by a steeper growth rate of the nonstop A-B fare relative to the connecting A-C fare.

Another key finding emerges from Table 6. In addition to the expected result that additional competition on route A-B (A-C) decreases A-B (A-C) fares, we observe that *Competition A-B* is statistically insignificant in the $\log(\text{Fare}_{\text{AC}})$ regression, while *Competition A-C* is negative and marginally significant in the $\log(\text{Fare}_{\text{AB}})$ regression. We believe that A-C fares are less directly related to the extent of competition on A-B routes than the A-B fare is to competition on A-C routes. The rationale is that the effect of A-B competition on the A-C fare should be minimal, because the relevant competition measure for A-C routes is broader, not only involving route A-B, but all other routes that start in A, terminate at C, and connect at airports other than B.

Table 6: Fare regressions

	(1)	(2)
Estimator:	2SLS	2SLS
Dependent variable:	log(Fare_{AB})	log(Fare_{AC})
Competition A-B	-0.077***	0.018
	(0.023)	(0.020)
Competition A-C	-0.034*	-0.124***
	(0.019)	(0.018)
Hub A	0.195***	0.098
	(0.071)	(0.064)
Hub B	0.296***	0.163*
	(0.092)	(0.091)
Hub C	-0.047	0.087
	(0.053)	(0.070)
$\mathbb{1}$(Dist. A-C > Dist. A-B)	0.048	0.125***
	(0.062)	(0.036)
DaysToDeparture 1-2	0.868***	0.684***
	(0.053)	(0.040)
DaysToDeparture 3-6	0.524***	0.430***
	(0.058)	(0.031)
DaysToDeparture 7-13	0.201***	0.193***
	(0.045)	(0.025)
DaysToDeparture 14-20	0.035**	0.058***
	(0.018)	(0.016)
Alaska	0.123	0.511***
	(0.117)	(0.141)
American Airlines	0.325***	0.150
	(0.103)	(0.147)
Delta	0.506***	0.199
	(0.184)	(0.155)
United	0.334***	0.169
	(0.110)	(0.155)
Frontier	0.023	0.124
	(0.112)	(0.141)
JetBlue	0.747***	0.528***
	(0.138)	(0.142)
Spirit	0.182*	0.160
	(0.094)	(0.128)
Kleibergen-Paap rk LM stat.	14.198**	14.198**
Hansen J Statistic	7.394	6.168
Kleibergen-Paap rk Wald F stat.	21.862***	21.862***
R^2	0.494	0.435
Observations	473,642	473,642

Notes: All specifications include month-of-year, day-of-week, and time-of-day-of-departure fixed effects. Sun Country is the omitted airline fixed effect. The endogenous variables in columns (1) and (2) are *Competition A-B* and *Competition A-C*. The corresponding first-stage regressions are reported in columns (3) and (4) of Appendix Table A2. Standard errors are provided in parentheses and clustered at the route A-B level. Constant is included but not reported. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

4.3 Savings from HCT

The analysis thus far has shown *when* and *why* HCT is more likely to occur. Our next step is to examine the price differential due to HCT, which represents the possible savings that a passenger may accrue from engaging in HCT, or, alternatively, the airline's potential revenue loss from a HCT passenger. To do so, we construct a new variable, *PriceDifference*, which is set equal to the difference between Fare_{AB} and Fare_{AC}. If this difference is negative (i.e., HCT does not occur), *PriceDifference* is set equal to zero. Because *PriceDifference* is nonnegative and censored at zero, we estimate a Tobit model. We use the same set of regressors described in equation (3), as well as the same set of instruments to correct for the potential endogeneity of the competition variables. In other words, we estimate equation (3) using a Tobit model with *PriceDifference* as the dependent variable.

The Tobit results are presented in Table 7. The signs on the competition variables are consistent with expectations. In the same manner that additional competition on A-B routes decreases the likelihood of HCT, additional competition on A-B routes also decreases the price difference due to HCT. Furthermore, consistent with how additional competition on A-C routes increases the likelihood of HCT, additional A-C competition also increases the HCT price difference.

Considering the estimates from our preferred specification in column (2) of Table 7, an additional nonstop carrier on route A-C increases the average price difference by almost \$26, while an additional nonstop carrier on route A-B decreases the average price difference by almost \$28. In addition, the price difference due to HCT is higher on the major full-service carriers: Delta has the largest average price difference, followed by United, and then American.

The magnitude on the $DaysToDeparture$ indicators are also plausible because the HCT price difference increases as the departure date approaches. Consistent with Figure 3, the peak of the price difference occurs in the last two days to departure. Relative to trips booked

Table 7: Determinants of *PriceDifference*

	(1)	(2)
Estimator:	Tobit	IV-Tobit
Dependent variable:	PriceDifference	PriceDifference
Competition A-B	-13.700*	-27.515***
	(7.742)	(8.651)
Competition A-C	19.645***	25.623***
	(2.802)	(4.089)
Hub A	21.201	23.660
	(23.800)	(23.073)
Hub B	13.031	-4.586
	(27.379)	(23.954)
Hub C	-32.692*	-37.313*
	(18.572)	(19.883)
$\mathbb{1}$(Dist. A-C > Dist. A-B)	-13.081	0.838
	(15.010)	(15.764)
DaysToDeparture 1-2	125.694***	125.838***
	(33.058)	(33.101)
DaysToDeparture 3-6	97.980***	99.624***
	(30.985)	(31.197)
DaysToDeparture 7-13	29.191*	30.579*
	(17.373)	(18.064)
DaysToDeparture 14-20	-3.686	-3.861
	(9.441)	(9.528)
Alaska	153.073***	168.966***
	(37.328)	(39.693)
American Airlines	300.399***	306.814***
	(47.174)	(48.241)
Delta	357.117***	372.394***
	(81.300)	(91.130)
United	304.195***	310.789***
	(46.033)	(48.038)
Frontier	106.456***	125.557***
	(38.831)	(38.236)
JetBlue	183.357***	185.427***
	(51.700)	(51.536)
Spirit	201.235***	196.405***
	(45.680)	(44.563)
Pseudo-R^2	0.087	
Observations	473,642	473,642

Notes: The dependent variable (*PriceDifference*) is equal to max(0, $\text{Fare}_{\text{AB}} - \text{Fare}_{\text{AC}}$). All specifications include month-of-year, day-of-week, and time-of-day-of-departure fixed effects. Sun Country is the omitted airline fixed effect. The endogenous variables in column (2) are *Competition A-B* and *Competition A-C*. The corresponding first-stage regressions are reported in columns (3) and (4) of Appendix Table A2. Standard errors are provided in parentheses and clustered at the route A-B level. Constant is included but not reported. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

21 to 60 days in advance, the HCT price difference increases by almost $126 in the last two days to departure.

Finally, a robustness check based on a different dependent variable (the percentage price difference) and model (fractional logit) yields similar qualitative results to Table 7. These results are not reported here, but are available in Appendix Table A3.

5 Conclusion

This article has offered a comprehensive empirical analysis of hidden-city ticketing (HCT), which, to the best of our knowledge has never been conducted before. HCT is a pricing phenomenon that occurs when the fare for a nonstop trip from A to B (i.e., A-B routes) is more expensive than a connecting trip from A to C that connects at B (i.e., the "hidden city"). Exploiting a unique panel of over 473 thousand fares collected over a five-month period (flights in our sample depart between October 1st, 2019 and December 31st, 2019), we find that HCT opportunities arise approximately 14% of the time. In particular, the major U.S. carriers that operate large hub-and-spoke networks (i.e., American, Delta, and United) account for the majority of HCT.

Analyzing the determinants of HCT, we find that competition is one of the primary drivers, especially on A-C routes. An additional nonstop carrier on route A-C increases the likelihood of HCT by 1.8%-4.7% while an additional nonstop carrier on route A-B decreases the likelihood of HCT by 2.2%-3.8%. These findings are consistent with standard economic theory that predicts that additional competition results in lower market prices. Because HCT occurs when $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$, additional competition on A-C should reduce Fare_{AC}, thereby increasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds. Conversely, additional competition on A-B reduces Fare_{AB}, decreasing the likelihood that $\text{Fare}_{\text{AB}} > \text{Fare}_{\text{AC}}$ holds.

We also find that hub endpoints are another key determinant of HCT, with the likelihood of HCT decreasing by 4.8%-8.3% when airport C is a hub. This finding is sensible considering that carriers are able to charge a premium for trips that originate or terminate at their hubs (i.e., the "hub premium"). As a result, fares on A-C routes will tend to be high if airport C is a hub, decreasing the likelihood that $\mathrm{Fare_{AB}} > \mathrm{Fare_{AC}}$ holds.

We also find that advance-purchase requirements are another key driver of HCT, with HCT opportunities more likely closer to the date of departure. In particular, HCT is more prevalent in the last week to departure because nonstop A-B fares increase at a higher rate than connecting A-C fares during this period. One possible interpretation of this finding is related to the heterogeneity of passengers during the booking period. Because early purchasers are typically price-sensitive passengers with low search costs, they are more likely to seek out HCT opportunities. Accordingly, airlines may respond by ensuring that HCT opportunities are rare during the early booking period. In contrast, most passengers purchasing tickets a few days before departure are price-insensitive customers with high search costs (i.e., late purchasers who are less likely to seek out HCT opportunities). For this reason, airlines may decide to extract additional surplus by raising nonstop A-B fares at a higher rate than connecting A-C fares in the final week because they are less concerned about passengers taking advantage of HCT opportunities during this period.

In addition to examining the determinants of HCT, we also quantify the savings that a passenger receives from engaging in HCT. We find that an additional nonstop carrier serving the A-B (A-C) route leads to a \$28 reduction (\$26 increase) in average savings. Moreover, average savings from HCT increase by \$100 and \$126 for trips purchased three to six and one to two days before departure, respectively.

As internet search engines become more sophisticated, they are increasingly helping consumers quickly identify HCT opportunities. However, HCT is clearly detrimental to airline operations and profits. In addition to the revenue loss that results from lower fares paid by HCT passengers, HCT may also delay the departure of the B-C flight if the airline waits in vain for HCT passengers (Skorupski and Wierzbińska, 2015). There is also an opportunity cost associated with reserving a seat on the B-C flight for a HCT passenger when that seat could instead be sold to another customer.

It is also worth mentioning that if all connecting A-C passengers were HCT passengers at connecting city B, the B-C flight would fly empty. This is obviously an extreme and unlikely outcome, but it clearly demonstrates that HCT could have important environmental consequences that should be considered by regulators (Kang et al., 2022). In other words, HCT passengers are unnecessary polluters that should not only be discouraged by airlines, but also discouraged by regulatory authorities.

Given that HCT decreases profits, our findings should be of interest to airline yield managers. For example, yield managers may be able to improve their pricing strategy by using our results to identify when yield management algorithms should be altered to prevent HCT. In this way, our findings on advance-purchase requirements, competition, and hub endpoints can help yield managers identify specific routes and booking horizons where the risk of losses from HCT passengers are high (e.g., competitive A-B routes within one week of departure). When potential losses are high (i.e., above some specified threshold), a new pricing rule could be implemented that alters fares to either (i) remove HCT opportunities or (ii) reduce the severity of the potential revenue loss. Nevertheless, imposing a restriction that removes *all* HCT opportunities from the output of yield management algorithms is not likely to be optimal. For instance, ruling out HCT opportunities by decreasing nonstop A-B fares and/or increasing connecting A-C fares will often result in decreased revenues from A-B passengers who are charged lower fares, and decreased revenues from A-C passengers who decide not to purchase at the higher price.

An interesting extension to the analysis presented in this article would be to examine whether airlines attempt to circumvent HCT by applying differential pricing for one-ways and roundtrips that connect through attractive intermediate cities. To circumvent HCT, it is expected that the usual one-way premium would be higher for trips connecting at attractive destinations (e.g., Los Angeles, New York, Miami) than for trips connecting at relatively unattractive destinations (e.g., Atlanta, Houston, or Phoenix). In other words, airlines may raise one-way fares that connect in attractive cities, so that the gain from exploiting HCT on these routes is diminished.

More generally, future research could extend the present analysis to other countries or continents. The U.S. domestic market is quite consolidated, but elsewhere it is not. For example, the airline industry is at an earlier stage of consolidation in Europe, with almost every European country having its own flag carrier and few steps taken towards consolidation (e.g., the Air France/KLM merger in 2004 and the British Airways/Iberia merger in 2011). The European market is characterized by differences in airline network structures, with full-service carriers spatially operating around a small number of central hubs and low-cost carriers evenly spreading flights across their networks (Bubalo and Gaggero, 2021). Given these differences, it would be interesting to test whether the results we find on route competition also extend to the European airline market.

Furthermore, the present analysis could be extended to other modes of transport that, like airlines, operate using hub-and-spoke networks. Examples include passenger rail and long-distance bus, to see if these transport modes, which started applying rudimentary yield management techniques by offering discounted fares to early purchasers, have HCT opportunities.

References

Aryal, G., Murry, C., and Williams, J. W. (2023). Price discrimination in international airline markets. *Review of Economic Studies, forthcoming*.

Aslani, S., Modarres, M., and Sibdari, S. (2014). On the fairness of airlines' ticket pricing as a result of revenue management techniques. *Journal of Air Transport Management*, 40:56–64.

Belobaba, P. (2009). Fundamentals of pricing and revenue management. *The global airline industry*, pages 73–111.

Bergantino, A. S. and Capozza, C. (2015). Airline pricing behaviour under limited intermodal competition. *Economic Inquiry*, 53(1):700–713.

Bilotkach, V., Gaggero, A. A., and Piga, C. A. (2015). Airline pricing under different market conditions: Evidence from European low-cost carriers. *Tourism Management*, 47:152–163.

Bilotkach, V. and Pai, V. (2016). Hubs versus airport dominance. *Transportation Science*, 50(1):166–179.

Bischoff, G., Maertens, S., and Grimme, W. (2011). Airline pricing strategies versus consumer rights. *Transportation Journal*, 50:232–250.

Borenstein, S. (1989). Hubs and high fares: Dominance and market power in the U.S. airline industry. *The RAND Journal of Economics*, 20(3):344–365.

Brueckner, J. K., Dyer, N. J., and Spiller, P. T. (1992). Fare determination in airline hub-and-spoke networks. *The RAND Journal of Economics*, 23(3):309–333.

Brueckner, J. K., Lee, D., and Singer, E. S. (2013). Airline competition and domestic US airfares: A comprehensive reappraisal. *Economics of Transportation*, 2(1):1–17.

Brueckner, J. K. and Spiller, P. T. (1994). Economies of traffic density in the deregulated airline industry. *The Journal of Law & Economics*, 37(2):379–415.

Bubalo, B. and Gaggero, A. A. (2021). Flight delays in European airline networks. *Research in Transportation Business & Management*, 41.

Caves, D. W., Christensen, L. R., and Tretheway, M. W. (1984). Economies of density versus economies of scale: why trunk and local service airline costs differ. *The RAND Journal of Economics*, 15(4):471–489.

Ciliberto, F. and Williams, J. W. (2010). Limited access to airport facilities and market power in the airline industry. *The Journal of Law and Economics*, 53(3):467–495.

Dai, M., Liu, Q., and Serfes, K. (2014). Is the effect of competition on price dispersion nonmonotonic? Evidence from the U.S. airline industry. *The Review of Economics and Statistics*, 96(1):161–170.

Dana, Jr, J. D. (1998). Advance-purchase discounts and price discrimination in competitive markets. *Journal of Political Economy*, 106(2):395–422.

Escobari, D. (2011). Frequent flyer programs premium and the role of airport dominance. *Applied Economics Letters*, 18(16):1565–1569.

Escobari, D. (2012). Dynamic pricing, advance sales and aggregate demand learning in airlines. *The Journal of Industrial Economics*, 60(4):697–724.

Escobari, D. and Jindapon, P. (2014). Price discrimination through refund contracts in airlines. *International Journal of Industrial Organization*, 34:1–8.

Escobari, D., Rupp, N. G., and Meskey, J. (2019). An analysis of dynamic price discrimination in airlines. *Southern Economic Journal*, 85(3):639–662.

Gaggero, A. A. and Luttmann, A. (2020). Purchase discounts and travel premiums during holiday periods: Evidence from the airline industry. MPRA Paper 104863, University Library of Munich, Germany.

Gaggero, A. A. and Luttmann, A. (2022). How does COVID-19 affect intertemporal price dispersion? Evidence from the airline industry. MPRA Paper 111797, University Library of Munich, Germany.

Gaggero, A. A. and Piga, C. A. (2010). Airline competition in the British Isles. *Transportation Research Part E: Logistics and Transportation Review*, 46:270–279.

Gaggero, A. A. and Piga, C. A. (2011). Airline market power and intertemporal price dispersion. *The Journal of Industrial Economics*, 59(4):552–577.

GAO (2001). Restricting airline ticketing rules unlikely to help consumers. Technical report, U.S. Government Accountability Office, GAO-01-831.

Gerardi, K. S. and Shapiro, A. H. (2009). Does competition reduce price dispersion? New evidence from the airline industry. *Journal of Political Economy*, 117(1):1–37.

Kang, Y., Liao, S., Jiang, C., and D'Alfonso, T. (2022). Synthetic control methods for policy analysis: Evaluating the effect of the European Emission Trading System on aviation supply. *Transportation Research Part A: Policy and Practice*, 162:236–252.

Koenigsberg, O., Muller, E., and Vilcassim, N. J. (2008). easyJet® pricing strategy: Should low-fare airlines offer last-minute deals? *Quantitative Marketing and Economics*, 6(3):279–297.

Lederman, M. (2008). Are frequent-flyer programs a cause of the "hub premium"? *Journal of Economics & Management Strategy*, 17(1):35–66.

Liu, Q. (2020). Paying more for a shorter flight? - Hidden city ticketing. pages 1–41. In *Essays on Estimation of Microeconomic Models*. PhD dissertation, University of Pittsburgh.

Luttmann, A. (2019a). Are passengers compensated for incurring an airport layover? Estimating the value of layover time in the U.S. airline industry. *Economics of Transportation*, 17:1–13.

Luttmann, A. (2019b). Evidence of directional price discrimination in the U.S. airline industry. *International Journal of Industrial Organization*, 62:291–329.

Meire, S. and Derudder, B. (2022). Pirating the skies? A review of airline booking ploys. *Research in Transportation Business & Management*, 43.

Oh, J. and Huh, W. T. (2022). Hidden city travel and its impact on airfare: The case with competing airlines. *Transportation Research Part B: Methodological*, 156:101–109.

Puller, S. L. and Taylor, L. M. (2012). Price discrimination by day-of-week of purchase: Evidence from the US airline industry. *Journal of Economic Behavior & Organization*, 84(3):801–812.

Skorupski, J. and Wierzbińska, M. (2015). A method to evaluate the time of waiting for a late passenger. *Journal of Air Transport Management*, 47:79–89.

Stavins, J. (2001). Price discrimination in the airline market: The effect of market concentration. *Review of Economics and Statistics*, 83(1):200–202.

Sun, X., Wandelt, S., and Zhang, A. (2022). Price discrimination through hidden city options? A data-driven study on the extent and evolution of skiplaggability in the global aviation system. Available at SSRN: https://ssrn.com/abstract=4137806.

Talluri, K. T., Van Ryzin, G., and Van Ryzin, G. (2004). *The theory and practice of revenue management*, volume 1. Springer.

Venkatesh, R. and Mahajan, V. (2009). The design and pricing of bundles: A review of normative guidelines and practical approaches. *Handbook of Pricing Research in Marketing*, pages 232–257.

Wang, Z. and Ye, Y. (2016). Hidden-city ticketing: The cause and impact. *Transportation Science*, 50(1):288–305.

Williams, K. R. (2022). The welfare effects of dynamic pricing: Evidence from airline markets. *Econometrica*, 90(2):831–858.

Wooldridge, J. (2008). *Introductory Econometrics: A Modern Approach*. ISE - International Student Edition. Cengage Learning.

Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*. MIT Press Books. The MIT Press.

A Appendix

A.1 Summary statistics

Table A1: Summary statistics and a brief description of the variables included in the analysis

	Description	Mean	Std. Dev.	Min	Max
DEPENDENT VARIABLES					
HCT%†	$\frac{\textit{One-Way Tickets}_{rca}}{(\textit{One-Way Tickets}_{rca} + \textit{One-Way Tickets}_{ra})}$	0.055	0.094	0.000	0.907
HCT	Dummy=1 in case of Hidden-City Ticketing	0.138	0.345	0.000	1.000
$\log(\text{Fare}_{\text{AB}})$	Fare A-B, nonstop flight, in logs	4.697	0.628	2.708	7.955
$\log(\text{Fare}_{\text{AC}})$	Fare A-C with layover in B, in logs	5.149	0.535	3.555	7.901
PriceDifference	$\max(0, \text{Fare}_{\text{AB}} - \text{Fare}_{\text{AC}})$	8.418	38.212	0.000	2,277
PriceDifference%	$\max\left(0, \frac{\text{Fare}_{\text{AB}} - \text{Fare}_{\text{AC}}}{\text{Fare}_{\text{AB}}}\right)$	0.029	0.098	0.000	0.880
REGRESSORS					
AveragePriceDifference†	$\max(0, \text{AverageFare}_{\text{AB}} - \text{AverageFare}_{\text{AC}})$	30.905	60.011	0.000	600.0
Competition A-B	Number of nonstop carriers serving route A-B on the flight's day of departure	3.898	1.370	1.000	8.000
Competition A-C	Number of nonstop carriers serving route A-C on the flight's day of departure	2.442	1.644	0.000	8.000
Hub A	Dummy=1 if airport A is a hub	0.383	0.486	0.000	1.000
Hub B	Dummy=1 if airport B is a hub	0.562	0.496	0.000	1.000
Hub C	Dummy=1 if airport C is a hub	0.124	0.329	0.000	1.000
$\mathbb{1}$(Dist. A-C > Dist. A-B)	Dummy=1 if Distance A-C > Distance A-B	0.813	0.390	0.000	1.000
DaysToDeparture 1-2	Dummy=1 if DaysToDeparture $\in [1, 2]$	0.043	0.204	0.000	1.000
DaysToDeparture 3-6	Dummy=1 if DaysToDeparture $\in [3, 6]$	0.094	0.292	0.000	1.000
DaysToDeparture 7-13	Dummy=1 if DaysToDeparture $\in [7, 13]$	0.141	0.348	0.000	1.000
DaysToDeparture 14-20	Dummy=1 if DaysToDeparture $\in [14, 20]$	0.118	0.323	0.000	1.000
DaysToDeparture 21-60	Dummy=1 if DaysToDeparture $\in [21, 60]$, omitted category in the regressions	0.604	0.489	0.000	1.000
Alaska	Dummy=1 for Alaska	0.059	0.235	0.000	1.000
American Airlines	Dummy=1 for American Airlines	0.243	0.429	0.000	1.000
Delta	Dummy=1 for Delta	0.098	0.298	0.000	1.000
United	Dummy=1 for United	0.175	0.380	0.000	1.000
Frontier	Dummy=1 for Frontier	0.045	0.207	0.000	1.000
JetBlue	Dummy=1 for JetBlue	0.037	0.188	0.000	1.000
Spirit	Dummy=1 for Spirit	0.338	0.473	0.000	1.000
Sun Country	Dummy=1 for Sun Country, omitted category in the regressions	0.006	0.075	0.000	1.000
INSTRUMENTS					
Passengers A-B	Monthly number of nonstop passengers on route A-B, in thousands	88.194	32.035	19.194	168.542
Passengers A-C	Monthly number of nonstop passengers on route A-C, in thousands	51.834	35.120	0.000	168.542
$\log(\sqrt{PopA * PopB})$	Geometric mean of population of A and B, in logs	15.690	0.443	14.480	16.588
$\log(\sqrt{PopA * PopC})$	Geometric mean of population of A and C, in logs	15.613	0.471	14.480	16.588
$\log(\frac{PopA+PopB}{2})$	Arithmetic mean of population of A and B, in logs	15.798	0.450	14.687	16.606
$\log(\frac{PopA+PopC}{2})$	Arithmetic mean of population of A and C, in logs	15.777	0.452	14.687	16.606

Notes: Number of observations is 473,642, except 162,889 for the variables marked with † (DB1B data).

A.2 Estimated probit coefficients and first-stage regressions

The estimated probit coefficients corresponding to the last two columns of Table 5 are provided in columns (1) and (2) of Table A2 while the first-stage estimates for Table 5, Table 6, and Table 7 are reported in columns (3) and (4).

In the first-stage regressions, the days to departure indicators are generally statistically insignificant, indicating that the number of competitors on both A-B and A-C routes is unaffected by how far in advance airfare is purchased. Moreover, relative to Sun Country (a small low-cost carrier), the presence of other carriers reduces the total number of competitors operating on a given route.

With respect to our instruments, the number of passengers has the expected positive sign, indicating that denser routes are able to sustain more competitors. This passenger-traffic effect is also only observed for the route that the variable directly corresponds to. For example, *Passengers A-B* is only positive and statistically significant in column (3) when the dependent variable is Competition A-B. Similarly, *Passengers A-C* is only positive and statistically significant in column (4) when the dependent variable is Competition A-C. Finally, the effect of endpoint city populations is less clear. However, the arithmetic mean has the expected positive sign and is statistically significant in column (3) when Competition A-B is the dependent variable.

Table A2: First-stage regressions and estimated probit coefficients for Table 5

	(1)	(2)	(3)	(4)
Estimator:	Probit	IV-Probit	OLS	OLS
Dependent variable:	HCT	HCT	Comp. A-B	Comp. A-C
Competition A-B	-0.146**	-0.242***		
	(0.059)	(0.061)		
Competition A-C	0.207***	0.270***		
	(0.025)	(0.039)		
Hub A	0.182	0.216	0.042	-0.221
	(0.180)	(0.176)	(0.217)	(0.175)
Hub B	0.103	-0.017	0.001	0.206
	(0.231)	(0.210)	(0.272)	(0.198)
Hub C	-0.305*	-0.355*	0.264	-0.091
	(0.177)	(0.194)	(0.191)	(0.260)
$\mathbb{1}$(Dist. A-C > Dist. A-B)	-0.193	-0.082	0.287*	0.071
	(0.135)	(0.142)	(0.171)	(0.119)
DaysToDeparture 1-2	0.906***	0.884***	-0.013	0.133**
	(0.127)	(0.123)	(0.039)	(0.060)
DaysToDeparture 3-6	0.530***	0.526***	-0.020	-0.049
	(0.159)	(0.156)	(0.038)	(0.049)
DaysToDeparture 7-13	0.090	0.092	0.028	-0.020
	(0.125)	(0.127)	(0.029)	(0.035)
DaysToDeparture 14-20	-0.103	-0.105	0.002	0.044**
	(0.078)	(0.077)	(0.019)	(0.021)
Alaska	1.452***	1.504***	-1.100***	-1.527***
	(0.306)	(0.297)	(0.419)	(0.488)
American Airlines	2.836***	2.793***	-1.009**	-1.675***
	(0.273)	(0.253)	(0.417)	(0.498)
Delta	3.143***	3.168***	-0.815	-1.476***
	(0.332)	(0.350)	(0.512)	(0.508)
United	2.866***	2.809***	-0.959**	-0.959*
	(0.281)	(0.251)	(0.450)	(0.523)
Frontier	1.119***	1.168***	-0.396	-1.437**
	(0.324)	(0.274)	(0.506)	(0.594)
JetBlue	1.679***	1.646***	-0.169	-1.796***
	(0.351)	(0.339)	(0.537)	(0.517)
Spirit	1.875***	1.772***	-0.649	-1.280**
	(0.212)	(0.185)	(0.416)	(0.529)
Passengers A-B			0.026***	0.000
			(0.004)	(0.002)
Passengers A-C			0.003	0.040***
			(0.002)	(0.002)
$\log\left(\sqrt{PopA * PopB}\right)$			-5.771***	0.543
			(1.307)	(0.431)
$\log\left(\sqrt{PopA * PopC}\right)$			-1.154**	-0.261
			(0.472)	(0.482)
$\log\left(\frac{PopA+PopB}{2}\right)$			4.798***	-0.134
			(1.289)	(0.390)
$\log\left(\frac{PopA+PopC}{2}\right)$			1.246***	-0.187
			(0.464)	(0.456)
Wald χ^2 test		24.519***		
R^2 or Pseudo-R^2	0.224		0.667	0.666
Observations	473,642	473,642	473,642	473,642

Notes: Columns (1) and (2) report the estimated probit coefficients. Columns (3) and (4) report the first-stage regressions. All specifications include month-of-year, day-of-week, and time-of-day-of-departure fixed effects. Sun Country is the omitted airline fixed effect. Standard errors are provided in parentheses and clustered at the route A-B level. Constant is included but not reported. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

A.3 Robustness: percentage price difference

As a robustness check, we replicate the analysis reported in Table 7 using, in place of *PriceDifference*, the price difference in percentage, i.e., *PriceDifference%*$= \max\left(0, \frac{\mathrm{Fare_{AB}} - \mathrm{Fare_{AC}}}{\mathrm{Fare_{AB}}}\right)$. We then estimate a fractional logit model (see Table A3).

The potential endogeneity of the competition variables is accounted for using a control function approach described in Wooldridge (2001), where each endogenous variable (i.e., *Competition A-B* and *Competition A-C*) is first regressed on the instruments and the exogenous variables to obtain the residuals, $\widehat{v}_{AB}$ and $\widehat{v}_{AC}$, which are then included as additional regressors in fractional logit model to produce unbiased estimates.[24]

[24] Because the residuals are used as regressors in the second-stage, standard errors are bootstrapped.

Table A3: Determinants of *PriceDifference%*

	(1)		(2)	
Estimator	Fractional logit		Fractional logit	
Dependent variable	PriceDifference%		PriceDifference%	
	Estimated coefficients	Marginal effects	Estimated coefficients	Marginal effects
Competition A-B	-0.220	-0.006	-0.431***	-0.011***
	(0.168)	(0.005)	(0.005)	(0.000)
Competition A-C	0.306***	0.008***	0.413***	0.011***
	(0.075)	(0.002)	(0.004)	(0.000)
Hub A	0.131	0.003	0.194***	0.005***
	(0.371)	(0.010)	(0.012)	(0.000)
Hub B	0.204	0.005	-0.016	-0.000
	(0.477)	(0.013)	(0.032)	(0.001)
Hub C	-0.500*	-0.013*	-0.564***	-0.015***
	(0.302)	(0.008)	(0.014)	(0.000)
$\mathbb{1}$(Dist. A-C > Dist. A-B)	0.028	0.001	0.219***	0.006***
	(0.291)	(0.008)	(0.014)	(0.000)
DaysToDeparture 1-2	1.546***	0.057***	1.551***	0.057***
	(0.258)	(0.017)	(0.015)	(0.001)
DaysToDeparture 3-6	1.397***	0.048***	1.405***	0.048***
	(0.250)	(0.014)	(0.015)	(0.001)
DaysToDeparture 7-13	0.693***	0.017**	0.708***	0.017***
	(0.220)	(0.007)	(0.014)	(0.000)
DaysToDeparture 14-20	-0.038	-0.001	-0.050***	-0.001***
	(0.235)	(0.004)	(0.017)	(0.000)
Alaska	4.176***	0.003**	4.478	0.004***
	(0.838)	(0.002)	(3.036)	(0.000)
American Airlines	6.674***	0.039***	6.863**	0.039***
	(0.732)	(0.009)	(3.045)	(0.001)
Delta	7.388***	0.074**	7.755**	0.087***
	(0.798)	(0.032)	(3.044)	(0.001)
United	6.787***	0.043***	6.961**	0.043***
	(0.790)	(0.013)	(3.046)	(0.001)
Frontier	2.708***	0.001***	3.100	0.001***
	(0.791)	(0.000)	(3.022)	(0.000)
JetBlue	4.391***	0.004	4.593	0.004***
	(0.814)	(0.003)	(3.061)	(0.000)
Spirit	4.766***	0.006**	4.845	0.006***
	(0.656)	(0.003)	(3.045)	(0.000)
$\widehat{v}_{AB}$			0.459***	
			(0.008)	
$\widehat{v}_{AC}$			-0.236***	
			(0.006)	
Pseudo-R^2	0.176		0.185	
Observations	473,642		473,642	

Notes: The dependent variable (*PriceDifference%*) is equal to max $\left(0, \frac{\text{Fare}_{\text{AB}} - \text{Fare}_{\text{AC}}}{\text{Fare}_{\text{AB}}}\right)$. All specifications include month-of-year, day-of-week, and time-of-day-of-departure fixed effects. Sun Country is the omitted airline fixed effect. Model (1) originates from a standard fractional logit regression. Model (2) originates from a fractional logit regression with a control function approach, where each endogenous variable (*Competition A-B* and *Competition A-C*) is first regressed on the instruments and the exogenous variables to obtain the residuals, $\widehat{v}_{AB}$ and $\widehat{v}_{AC}$, which are then included as additional controls in the fractional logit model to produce unbiased estimates (Wooldridge, 2001). Standard errors are provided in parentheses and clustered at the route A-B level in Model (1) and bootstrapped in Model (2). Constant is included but not reported. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.