Public Communication and Collusion in the Airline Industry[*]

Gaurab Aryal[†]   Federico Ciliberto[‡]   Benjamin T. Leyden[§]

February 12, 2018

Abstract

We investigate whether the top management of all legacy U.S. airlines used their quarterly earnings calls as a mode of communication with other airlines to coordinate output reduction (fewer passenger seats) on competitive routes. We build an original and novel dataset on the public communication content from the earnings calls, and use Natural Language Processing techniques from computational linguistics to parse and code the text from earnings calls by airline executives to measure communication. Then we determine if mentioning terms associated with "capacity discipline" is a way to sustain collusion on capacity. The estimates show that when all legacy carriers in a market communicate "capacity discipline," it leads to a substantial reduction in the number of seats offered in the market. We find that the effect is driven entirely by legacy carriers, and also that the reduction is larger in smaller markets. Finally, we leverage our high-dimensional text data to develop novel approaches to implement falsification tests and check conditional exogeneity, and confirm that our finding —legacy airlines use public communication regarding capacity discipline to collude —is not spurious.

JEL: D22, L12, L41, L68.

Keywords: Airlines, Communication, Collusion, Capacity Discipline, Text Data.

1

1 Introduction

In all OECD countries, there are two legal paradigms that are meant to promote market efficiency but that are potentially at odds with each other. On one hand, antitrust laws forbid firms from communicating their strategic choices with each other to deter collusion. On the other hand, financial regulations promote open and transparent communication between publicly traded firms and their investors. While these latter regulations are intended to level the playing field among investors, policy makers have raised concerns in recent years that they may also facilitate anticompetitive behavior. For example, in 2010, the OECD Competition Committee noted that while there are pro-competitive benefits from increased transparency, increased transparency can also facilitate collusion because "information exchanges can ... offer firms points of coordination or focal points," while also "allow[ing] firms to monitor adherence to the collusive arrangement" [OECD, 2011].[1] Thus, firms can be transparent about their future strategies in their public communications to investors — for example, by announcing their intention to rein in capacity — which, in turn, can spur and sustain collusion on capacity.[2]

In this paper, we contribute to this overarching research and policy debate by investigating whether the top managers of all legacy U.S. airlines used their quarterly earnings calls to communicate with other legacy airlines in reducing the number of seats sold in the U.S.[3] More specifically,

---

[1] Similar situations, where one set of laws is at odds with another, generating unanticipated consequences, often in the form of antitrust violations, are observed in many industries. For example, in the U.S. pharmaceutical industry, the tension between the FDA laws and patent law, led to the Drug Price Competition and Patent Term Restoration Act (colloquially known as the Hatch-Waxman Act). This Act was intended to reduce entry barriers for generic drugs, but it incentivized incumbent firms to Pay-for-Delay of generic drugs and stifle competition. For more, see Feldman and Frondorf [2017]. In another example, Byrne and de Roos [2017] document that gasoline retailers in Australia used the price transparency program called *Fuelwatch* to initiate and sustain collusion.

[2] There is a subtle difference between economists and lawyers when it comes to the use of the term "tacit" to refer collusion; see Green, Marshall, and Marx [2014]. For lawyers, collusion is explicit only if there was an "agreement," which has a special meaning, otherwise it is tacit. For economists, a collusion is explicit if it involves communication, which includes cheap-talk, and/or transfers, otherwise the collusion is tacit. Thus, we are in the world of explicit collusion, henceforth collusion. We thank Leslie Marx for clarifying this distinction.

[3] Earnings calls are teleconferences in which a publicly traded company discusses its performance and future expectations with financial analysts and news reporters. Legacy carriers are Alaska Airlines (AS), American Airlines

we test the hypothesis that these airlines used keywords associated with the notion of "capacity discipline" in their earnings calls to communicate to their counterparts their willingness to reduce offered seats in markets where they compete head-to-head.[4]

Our empirical analysis is theoretically founded on the recent work by Awaya and Krishna [2016, 2017], who show that firms can use cheap talk (unverifiable and non-binding communication) to sustain collusion even when demand is stochastic and monitoring is hard as long as the sales across firms are affiliated, i.e., their joint density is log-supermodular, under collusion and less correlated when firms are not colluding.[5] In Awaya and Krishna [2016, 2017] framework, the airline industry is characterized by stochastic demand as well as *private* and *noisy* monitoring, without the ability to communicate collusion would otherwise be infeasible. In our institutional framework, the basic idea in Awaya and Krishna [2016, 2017] is developed as airlines having access to a communication technology (i.e., earnings calls) that allows them to signal to others whether their demand was high or low. When all airlines simultaneously communicate that their (residual) demand is low, it signals to everyone that their individual revenue was low due to low demand and not because some firm cheated.

Such a communication strategy can potentially allow airlines to circumnavigate difficulty they face when trying to coordinate that is particularly strong in airline industry where demand is affected by exogenous local events, such as weather or unforeseen events at the airport, and cross-

_____

(AA), Continental Airlines (CO), Delta Airlines (DL), Northwest Airlines (NW), United Airlines (UA) and US Airlines (US), and the low cost carriers (LCC) are AirTran Airways (FL), JetBlue (B6), Southwest (WN) and Spirit Airlines (NK).

[4] The idea of using "capacity discipline" as a message sent by airlines to signal their intention to restrict supply is also applied in the recent class action lawsuits filed against a few airlines (c.f. Section 2.1). Sharkey [2012] and Glusac [2017] provide the coverage of this concept in popular press. Also see Rosenfield, Carlton, and Gertner [1997] for antitrust issues related to communication among competing firms, and for law and economics of collusion and the use of communication in collusion see Kaplow [2013].

[5] There is a vast research on firms' market conduct and behavior of cartels; see, for example, Harrington [2006] and Marshall and Marx [2014]. This literature finds that cartel stability depends inversely on the extent of demand uncertainty [Green and Porter, 1984] and the frequency with which firms observe each other's actions [Sannikov and Skrzypacz, 2007], but proportionally on the monitoring technology [see Mailath and Samuelson, 2006, Chapter 12].

market events like political events and oil price shocks. Moreover, because airlines use connecting passengers to manage their load factors, monitoring is especially difficult, as it is impossible to break down a competitor's ticket fare by the segments of the trips.

To test our hypothesis we build an original and novel dataset on the public communication content in the earnings calls. The Securities and Exchange Commission (SEC) requires all publicly traded companies in the U.S. to file a quarterly report, which is usually accompanied by an earnings call where the top executives discuss the content of the report with analysts and financial journalists. First, we collect transcripts of these calls for 11 airlines from 2002:Q4 to 2016:Q4. Then we classify each earnings call as pertinent or not pertinent depending on whether the executives on the call declared their intention of engaging in capacity discipline.[6]

We estimate the effect of communication on the carriers' market-level capacity decisions using data from the T-100 domestic segment for U.S. carriers at the monthly and non-stop route level. To that end, we run a fixed-effect regression of the number of seats, offered by an airline in a market in a month, on an indicator of whether *all* legacy carriers that are operating in a market discuss capacity discipline. Given that airlines' capacity decisions depend on a wide variety of market-specific and overall economic conditions, our analysis includes a rich set of covariates to control such variation across markets, carriers, and over time.

We find that when all legacy carriers operating in an airport-pair market communicated about capacity discipline in a given quarter, the average number of seats offered in those markets decreased by 1.45% in the next quarter. Moreover, if we decompose the average effect by the type

---

[6] For example, consider the following statement by Alaska in 2003:Q3:

```
''I think what we've concluded is that there's enough noise in the markets with adjustments to
capacity in many of the markets that we serve that we are seeing strength in demand, which is
more a function of the changes in capacity than it is changes to the price.''
```

Clearly, there is a fine line between managing capacity to provide adequate service to satisfy demand while engaging in capacity discipline, whereby the airlines restrict the number of seats made available in a market even when there would be demand for more seats. We will return to this issue in Section 2.2.

of the airlines (legacy or LCC) we find no evidence of capacity restriction by the LCCs, and that all effects are due to legacy carriers. Thus, we find evidence to support the hypothesis that legacy airlines used public communication to reduce their offered capacity. To put this 1.45% decrease in perspective, consider the following fact. The average change in capacity among legacy carriers in our sample is 3.78%. So, the 1.45% decline in capacity associated with the use of the phrase capacity discipline accounts for more than a third of this average change. In this light, it is clear that the effect is economically significant.

We further explore whether this effect varies by market size. We would expect the estimate to differ by market size if the profitability of collusion varies by market size, which would be true if, for example, the size of a market affects the feasibility of collusion.[7]

On one hand, if smaller markets are more conducive to collusion than larger markets, either because smaller markets are less volatile and therefore easy for monitoring or because smaller markets are less contestable than bigger markets and therefore the cartel does not have to worry about potential entry, then we should estimate a larger (negative) effect in smaller markets than larger markets. On the other hand, if larger markets have a disproportionately high fraction of for-business travelers, who tend to be less price sensitive [Berry, Carnall, and Spiller, 2006; Berry and Jia, 2010; Ciliberto and Williams, 2014; Aryal, Ciliberto, Murry, and Williams, 2017], the larger markets might be more conducive to collusion and, therefore, we would estimate an effect that is increasing in market size.[8] Which of these two effects dominate is ex-ante ambiguous, making it an empirical question. The answer is important because it provides a newer and nuanced

---

[7] Following the literature [Berry and Jia, 2010], we use the geometric mean of the population at the two ends of the market to define the size of a market, and say that a market is small if the size is less than 25th percentile of the overall size distribution in the sample. Likewise, a market is large or medium based on whether the population is greater than or less than the 75th percentile.

[8] We use the same thresholds to classify a market as low, medium, or high business market as we did for the population designations; see Footnote 7.

understanding of the role of market size on the efficacy of communication on collusion.

To that end, we estimate how the effect of the communication differs by market sizes. The estimates suggest that whenever legacy carriers communicated with each other in small markets, they also reduced the number of seats by 4.21%. The effect for medium and larger markets are smaller, -1.95% and -1.25%, respectively. If we allow the effect of communication to vary by the proportion of business travelers, we find that the effect of communication is -2.74% in low-business markets and is not statistically different from the effect in medium-business markets. In fact, we find that communication has a positive effect on the number of seats offered in markets with a high fraction of for-business travelers. These results suggest that for collusion among legacy carriers, the ease of monitoring and the threat of entry is more important than the slope of demand.[9]

Next, we propose a novel approach to develop multiple placebo falsification tests to show that the only channel through which airlines coordinate is through the use of keywords associated with the concept of capacity discipline. This is a complex undertaking because the placebo falsification exercise should consist of keywords from all earnings calls that are unrelated to "capacity discipline" and that are discussed approximately as frequently. To find the keywords that satisfy these requirements we employ the `word2vec` model, a neural network model that is commonly used in computational linguistics [Mikolov, Chen, Corrado, and Dean, 2013].

The `word2vec` model takes our collection of transcripts as input, and then maps all of the one- and two-word phrases (henceforth, tokens) to a set of numerical vectors, where tokens that are similar in use/meaning are located "close" to each other, and tokens that are more dissimilar are located "farther" away from each other.[10] Essentially, the method captures the meaning and

---

[9] The results are largely the same when we define markets as city-pairs.

[10] As in much of the research on natural language processing, we use *cosine similarity* as our definition of the similarity between two tokens. Cosine similarity measure the cosine angle between any two vector associated with tokens so that it is 1 for identical vectors and 0 for orthogonal vectors. See Section 5 for further discussion.

the context by using word associations — the regular proximity of certain keywords to capacity discipline. Using the `word2vec` model, we select 40 tokens as placebos. For each of these placebo tokens, we estimate the same fixed-effect model for capacity discipline and find that none of these tokens lead to fewer passenger seats. This exercise supports our finding that is the use of keywords associated with capacity discipline that drive our estimated effect.

Related Literature. We contribute to a very rich literature in economics on collusion that goes back to at least Stigler [1964]. For a comprehensive overview, see Viscusi, Harrington, and Vernon [2005] and Marshall and Marx [2014]. One important class of models, including Green and Porter [1984] and Abreu, Pearce, and Stacchetti [1986], considers collusion when the output of individual firms is not observed by other firms, and instead a noisy signal, in the form of market clearing price, is publicly observed. In an important empirical paper, Porter [1983] tests the prediction from Green and Porter [1984] using data from the Joint Executive Committee railroad cartel. In this regard, our paper is similar in spirit to Porter [1983] because we rely on the prediction from Awaya and Krishna [2016, 2017] to test whether there is evidence of collusion maintained by the use of public communication in the U.S. airline industry.[11] And, as far as we know, this is the first empirical paper that links the theory of communication with collusion in capacity using field data.

We also complement the literature on law and economics of collusion, such as Miller [2010], that studies the airline industry in the context of the DOJ's litigation of collusion against 8 airlines and a clearing house that publishes airfares and restrictions among all airlines. As described Borenstein [2004], the DOJ alleges that the airlines used the electronic fare system from the aforementioned clearing house to communicate and sustain collusion.[12]

---

[11] In Porter [1983] and Green and Porter [1984], all firms observe the same (noisy signal) price, and access to communication technology does not change anything because the profits from public perfect equilibrium (a solution concept used in those papers) is the same with and without communication.

[12] The use of a clearing house to communicate about fares means Miller [2010] is closer to Porter [1983] than us.

There is a rich literature in game theory that studies the role of communication in noncooperative games; see [Myerson, 1997, Chapter 6]. In particular this literature shows that if players can use communication then they achieve higher payoff than they would without communication.[13] Ability to communicate can be even more beneficial under imperfect monitoring, where without communication collusion would be infeasible. In this paper we provide an empirical evidence supporting that claim for the airline industry.

Lastly, our paper is also related to the growing economic and computational social science literature that uses text as data. As more and more communication and market interactions are recorded digitally, the use of large-scale, unstructured text data in empirical research in and outside of industrial organization is likely to become even more important. For instance, Leyden [2018] considers the problem of defining relevant markets for smartphone and tablet applications using text descriptions of the applications. Other examples of papers that use text as data include Gentzkow and Shapiro [2014], who use phrases from the Congressional Record to measure the slant of news media; Baker, Bloom, and Davis [2016], who use the frequency of keywords related to "policy uncertainty" in newspapers to construct a measure of policy uncertainty; and Hoberg and Philips [2016] who use the text descriptions of businesses included in financial filings to define markets. Also see Hassan, Hollander, van Lent, and Tahoun [2017]; Hoberg and Maksimovic [2015]; Koijen, Philipspn, and Uhlig [2016]; Tetlock, Saar-Tsechanksy, and Macskassy [2008] for other applications of text-based analysis, and Gentzkow, Kelly, and Taddy [2017] for a survey.

We proceed as follows. In Section 2.1 we introduce the legal case. In Section 2.2 we describe the earning calls data and explain how we measure communication. In Section 2.3 we describe the airline data, and in Section 3 we present the main results from Awaya and Krishna [2016]. In

---

[13] Some key theory papers in this topic include Forges [1986, 1990]; Barany [1992]; Compete [1998]; Kandori and Matsushima [1998]; Gerardi [2004],and Spector [2015], among others. Also see Goltsman and Pavlov [2014] for the role of third-party communication in an oligopoly market.

Section 4 we present the results, and in Section 5 we present results from falsification tests before concluding in Section 6.

2 Institutional Analysis and Data

In this section we introduce the legal cases that motivate our approach, explain how we use Natural Language Processing (NLP) techniques to quantify communication among airlines, and finally present our data on the airline industry.

2.1 Legal Case

On July 1, 2015, news agencies reported that the DOJ was investigating possible collusion to limit available seats and maintain higher fares in U.S. domestic airline markets by American, Delta, Southwest Airlines, and United/Continental [see Harwell, Halsey III, and Moore, 2015]. The news also reported that the major carriers had received Civil Investigative Demands (CIDs) from the DOJ requesting copies dating back to January 2010 of all communications the airlines had with each other, Wall Street analysts and major shareholders, concerning their plans for seat capacity and any statements to restrict it. The investigation was subsequently confirmed by the airlines in their quarterly reports.

Concurrently, several consumers filed lawsuits accusing American, Delta, Southwest, and United of fixing prices, which were later consolidated in a multi-district litigation. The case is currently being tried in the U.S. District Court for the District of Columbia. [14] Another case, filed on August 24, 2015, in the U.S. District Court of Minnesota against American, Delta, Southwest Airlines, and United/Continental, alleges that the companies conspired to fix, raise, and maintain the price of

---

[14] This is the "Domestic Airline Travel Antitrust Litigation," case number 1:15-mc-01404 in the US District Court for the District of Columbia.

domestic air travel services in violation of Section 1 of the Sherman Antitrust Act. [15]

The lawsuits allege that the airline carriers collusively impose "capacity discipline" in the form of limiting flights and seats *despite increased demand and lower costs*, and that the four airlines implement and police the agreement through *public signaling* of *future* capacity decisions.[16] In particular, one of the consumers' lawsuits reported several statements made by the top managers of American, Delta, Southwest, United, and other airlines (such as Alaska Airlines or Canada Air). The statements were made during quarterly earnings calls and various conferences. For example, during the US Airways 2012:Q1 earnings call, the CFO of US Airways Derrick Kerr and Delta's CEO Richard Anderson said, respectively,

> ".. mainline passenger revenue were $2.1 billion, up 11.4% as a result of the strong pricing environment and continued industry capacity discipline." – US Airways.
> "You've heard us consistently state that we must be disciplined with capacity." – Delta

These lawsuits provide a basic framework to build a vocabulary from the earnings call that can capture airlines' intention to restrict their offered capacity. To that end, we have to consider both the semantics (airlines' intention to rein in capacity) and the syntax (what keywords are used) of the earnings call reports. Next, we explain the steps we take to measure communication among legacy carriers.
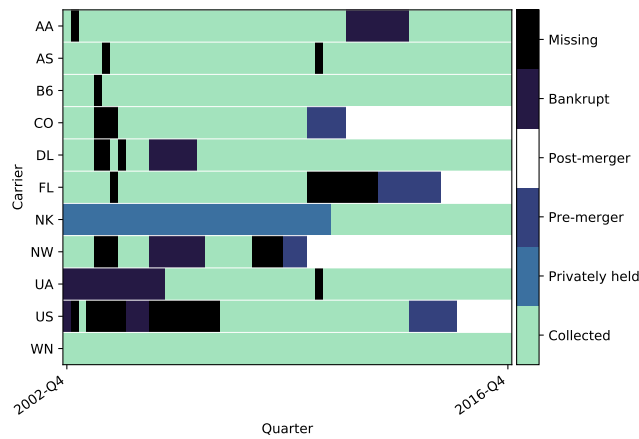
2.2 Earnings Call Text as Data

All publicly traded companies in the U.S. are required to file a quarterly report with the SEC. These reports are typically accompanied by an earnings call, which is a publicly available conference call between the firm's top management and the analysts and reporters covering the firm. Earnings

---

[15] This is Case 0:15-cv-03358-PJS-TNL, filed 8/24/2015 in the US District Court, District of Minnesota. In November 2015, this case was transferred to the District Court in DC.

[16] The consumers' lawsuits also stress the role of financial analysts who participate at the quarterly earnings call. See Azar, Schmalz, and Tecu [2017] for a recent important work on the role of institutional investors on market conduct, and Posner, Morton, and Weyl [forthcoming] for a proposal to limit the anti-competitive effect of having common institutional investors.

Figure 1: Transcript Availability

Notes. This figure shows the availability or non-availability of transcripts for 11 airlines. The x-axis denotes the time year and quarter, and the y-axis denote the name of the airline. Each color/shade denotes the status of the transcript.

calls begin with statements from some or all of the corporate participants followed by a question-and-answer session with the analysts on the call. Transcripts of calls are readily available, and we assume that carriers observe their competitors' calls.

We collected earnings call transcripts for 11 airlines, for all quarters from 2002:Q4 to 2016:Q4 from LexisNexis (an online database service) and Seeking Alpha (an investment news website). Figure 1 indicates the availability of transcripts in our sample for each of the 11 airlines. As the figure shows, transcripts are available for most of the periods except under (i) Bankruptcy — five carriers entered bankruptcy at least once during the sample period; (ii) Mergers and acquisitions — airlines did not hold earnings calls in the interim between the announcement of a merger and the full operation of the merger; (iii) Private airlines — Spirit Airlines, which was privately held until May 2011, neither submitted reports nor conducted earnings calls prior to its initial public offering; and (iv) Other reasons — there are a few instances when the transcripts were unavailable for an unknown reason. In all cases where a call is unavailable, we assume the carrier is unable to communicate to its competitors and engage in any potential cheap talk messaging.

11

The key step of our empirical analysis is to codify the informational content in these quarterly earnings calls into a dataset that can be used to see how capacity choices change over time in response to communication among legacy carriers. Before delving into the conceptual challenges that we face, there are two preliminary steps. Every statement made by the operator of the call and the analysts are removed from the transcripts, as are common English "stop words" such as "and" and "the."[17] Then, we tokenize (convert a body of text into a set of word or phrase "tokens") and lemmatize (reduce words to their dictionary form) the text from the earnings calls. For example, the sentence "The disciplined airline executive was discussing capacity discipline," would be reduced to the set $\{\texttt{discipline},\texttt{airline},\texttt{executive},\texttt{discuss},\texttt{capacity},\texttt{discipline}\}$. This process allows us to abstract from the inflectional and derivationally related forms of words in order to better focus on the substance/meanings of the transcripts.

The content of interest is of two types. First, using a combination of NLP techniques and manual review, we identify a list of words or phrases that are potentially indicative of managers communicating their intention to cooperate with others in restricting their capacity. Although in most cases managers specifically use the term "capacity discipline," there are instances where the managers use other word combinations when discussing the concept of capacity discipline. This identification is a time-consuming process, and it is the focus of the remainder of this section. Second, we use NLP to identify words that can be used for our placebo falsification test; we discuss this type of content in Section 5.

To codify the use of the phrase "capacity discipline" and other combinations of words that carry an analogous meaning, we begin by coding the phrase "capacity discipline" with a categorical variable $\texttt{Carrier-Capacity-Discipline}_{jt} \in \{0,1\}$ that takes the value 1 if that phrase appears in the earnings call transcript of carrier j in year-quarter t, and a value of 0 otherwise.

---

[17] The executives' responses to the analysts' questions are kept.

In many instances, however, airline executives do not use the exact phrase "capacity discipline," but the content of their statements are closely related to the notion of capacity discipline, as is illustrated in the following text:

> "We intend to at least maintain our competitive position. And so, what's needed here, given fuel prices, is a proportionate reduction in capacity across all carriers in any given market. And as we said in the prepared remarks, we're going to initiate some reductions and we're going to see what happens competitively. And if we find ourselves going backwards then we will be very capable of reversing those actions. So, this is a real fluid situation but clearly what has to happen across the industry is more reductions from where we are given where fuel is running." — Alaska Airlines , 2008:Q2.

Our view is that this instance, and other similar ones, should be interpreted as conceptually analogous to uses of the phrase capacity discipline.

Yet, in other cases it is arguable whether the content is conceptually analogous to the one of "capacity discipline," even though the wording would suggest so. For example, consider the following cases:

> "We are taking a disciplined approach to matching our plan capacity levels with anticipated levels of demand" — American Airlines, 2017:Q3

> "We will remain disciplined in allocating our capacity in the markets that will generate the highest profitability." — United Airlines, 2015:Q4

These statements, and others like these, cannot be easily categorized as a clear intention of the airline to reduce capacity below the GDP growth levels. On one hand, the "anticipated levels of demand" depend on the competitors' decisions, and thus one could interpret this statement as a signal to the competitors to maintain capacity discipline. On the other hand, an airline should not put more capacity than what is demanded because that implies higher costs and lower profits.

We take a *conservative* approach and code all these instances as ones where the categorical variable Carrier-Capacity-Discipline$_{jt}$ is equal to 1. This approach is conservative because it assumes that the airlines are coordinating their strategic choices more often than their words

would imply, and would work against finding a negative relation. In other words, we design our coding to err to find false negatives (failing to reject the null hypothesis that communication does not affect capacity), rather than erring on the side of finding false positives. The reason why we do this is that in our analysis we include variables that control for year, market, and year-quarter-market specific effects that control for any unobserved heterogeneity that might explain a reduction of capacity driven by a softening of the demand. Therefore, our coding approach *attenuates* the effect of "capacity discipline" and makes us *less* likely to find evidence of collusion when collusion is true.[18]

In practice, to identify all the instances where the notion of capacity discipline was present but the phrase "capacity discipline" was not used, we used NLP to process all transcripts and flag those transcripts where the word "capacity" was used *in conjunction with* either the word "demand" or "GDP." This filter identified 248 transcripts, which we read manually to classify them as either pertinent or not pertinent for capacity discipline. If the transcript was identified by all authors as pertinent, then we set the variable $\texttt{Carrier-Capacity-Discipline}_{jt} = 1$, and zero otherwise. Out of the 248 transcripts, we determined that 105 contained statements that we deemed pertinent.

Table 1 presents the summary statistics for the variable $\texttt{Carrier-Capacity-Discipline}_{jt}$.

We have 253 earnings calls transcripts for the legacy carriers, and 54.1% include content associated with the notion of capacity discipline. We have fewer transcripts for LCC, JetBlue and Southwest, and content associated with capacity discipline is much less frequent. Overall, we have 413 transcripts and $\texttt{Carrier-Capacity-Discipline}_{jt} = 1$ in 38.3% of them. The evidence in Table 1 suggests that the LCC, including Southwest (WN), are much less likely to publicly talk

---

[18] To further address the possibility that any association between "capacity discipline" and offered seats is driven by missing variables that are correlated with the former, in Section 5 we develop a placebo falsification test. In particular, we explore whether other words could exhibit a similar, negative, relationship as capacity discipline.

Table 1: Frequency of Communication

	Communication	N
Legacy	0.541	253
	(0.499)	
LCC	0.131	160
	(0.339)	
Jet Blue	0.111	54
	(0.317)	
Southwest	0.073	55
	(0.262)	
All	0.383	413
	(0.487)	

Notes. Fraction of earnings calls where `Carrier-Capacity-Discipline` is equal to one. The standard deviations are presented in the parentheses.

about capacity discipline. In view of this data feature, in our empirical exercise, we focus only on communication by legacy carriers.

2.3 Airline Data

We use two datasets for the airline industry: the T-100 Domestic Segment for U.S. carriers and a selected sample from the OAG Market Intelligence-Schedules dataset. We consider the periods between 2003:Q1 and 2016:Q3 (inclusive). The Bureau of Transportation Statistics' T-100 Domestic Segment for U.S. carriers contain domestic non-stop segment (i.e., route) data reported by U.S. carriers, including the *operating* carrier, origin, destination, available capacity, and load factor.

In many instances, there are also regional carriers, such as SkyWest or PSA, that operate on behalf of the *ticketing* carriers. The regional carriers might be subsidiaries that are fully owned by the national airlines, e.g., Piedmont, which is owned by American (and prior to that by U.S. Airways), or they might operate independently but contract with one or more national carrier(s), e.g., SkyWest. In order to allocate capacity to the *ticketing* carriers, we merge the information from the OAG Market Intelligence, which contains information about the operating and the ticketing

carrier for each segment at the quarterly level. Using this merged dataset, we allocate the available capacity in each route in the U.S. to the ticketing carriers, which will be the carriers of interest.

We consider only routes between airports that are located in the proximity of a Metropolitan Statistical Area in the U.S.[19] In our analysis, we use two methods of defining markets in the airline industry. The first follows Borenstein [1989]; Kim and Singal [1993]; Borenstein and Rose [1994]; Gerardi and Shapiro [2009]; Ciliberto and Tamer [2009]; Berry and Jia [2010]; Ciliberto and Williams [2010]; and Ciliberto and Williams [2014], and assumes that markets are defined by the origin and destination airport pairs. The second maintains that markets should be defined by the origin and destination *cities*, rather than airports. For example, consider two flights flying out of Reagan National Airport, located in Northern Virginia, with one flying to O'Hare International Airport and the other flying to Midway International Airport, both located in Chicago.

Under the airport-pair market definition, these flights operate in separate markets — the first is in the Reagan-O'Hare market, and the second is in the Reagan-Midway market. But, under the city-pair method of defining markets, we treat these flight as operating in the same market, because they both serve the Washington D.C. to Chicago market.[20] This definition has been followed, among others, by Berry [1990, 1992]; Brueckner and Spiller [1994]; Evans and Kessides [1994]; and Bamberger, Carlton, and Neumann [20004].

How to define airline markets is of key interest for antitrust matters. While the airport-pair approach is often used in academic research on the airline industry, the city-pair approach is particularly important for antitrust practitioners. This is because using the city-pair approach leads to larger markets, which, for antitrust purposes, provides a stronger basis for government intervention if evidence of anticompetitive effects is found.

---

[19] We use the U.S. Department of Commerce's 2012 data to identify Metropolitan Statistical Areas in the U.S.

[20] In our empirical analysis, we follow Brueckner, Lee, and Singer [2014] to determine which airports should be grouped in the same city for the city-pair definition approach.

In light of these points, we will first present our results using the airport-pair market definition. Then, given the legal cases that are being brought against the airlines, we will also show the results when markets are defined using the city-pair method. We will then discuss how the two sets of results provide a lower and upper bound on the effect of communication on capacity.
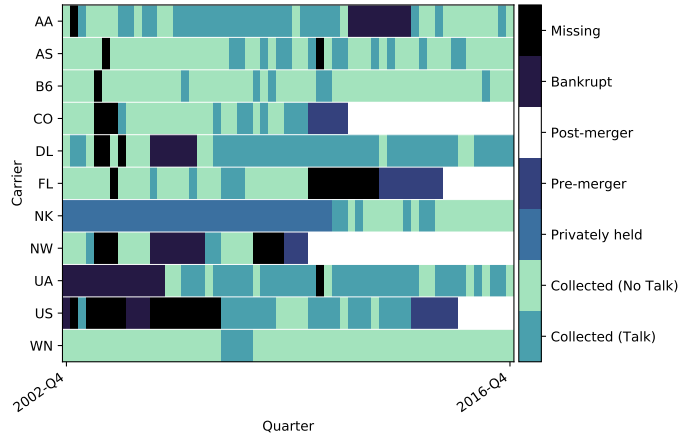
2.4 Variable Definitions

Legacy airlines are communicating with each other when *all* of those legacy airlines that are serving a non-monopoly market discuss capacity discipline. Defining $J_{mt}^{\texttt{Legacy}}$ as the set of legacy carriers in market m at time t, we define a new variable, only for the legacy carriers,

$$\texttt{Capacity-Discipline}_{mt} =$$
$$\begin{cases} \mathbb{1}\left\{\texttt{Carrier-Capacity-Discipline}_{jt} = 1 \; \forall j \in J_{mt}^{Legacy}\right\} & |J_{mt}^{Legacy}| \geq 2 \\ 0 & |J_{mt}^{Legacy}| < 2 \end{cases}$$

Thus, $\texttt{Capacity-Discipline}_{mt}$ indicates whether all of the legacy carriers in m discussed capacity discipline that quarter, conditional on two or more legacy carriers serving that market. In cases where one or fewer legacy carries serve a market, $\texttt{Capacity-Discipline}_{mt}$ is set equal to 0. While the variable $\texttt{Carrier-Capacity-Discipline}_{jt}$ varies by month and carrier, the variable $\texttt{Capacity-Discipline}_{mt}$ varies by market and year-month. This is an important distinction for the empirical analysis, where the observations will be at the market-carrier-year-quarter level.

Figure 2 shows the occurrence of the variable $\texttt{Carrier-Capacity-Discipline}_{j,t}$ in our data. Each row corresponds to one airline and shows the periods for which each carrier discussed capacity discipline. There is significant variation in communication across both airlines and time, which is necessary for identification. Even though the reports do not vary within a quarter, the composition of airlines operating markets vary both within a quarter and across quarters, providing enough variation in the dummy variable $\texttt{Capacity-Discipline}_{m,t}$.

Figure 2: Prevalence of "Capacity Discipline" in Earnings Call Transcripts

Notes. This figure shows the availability of transcripts and the prevalence of "Capacity Discipline" for 11 airlines. The x-axis denotes the time year and quarter, and the y-axis denote the name of the airline. Each color/shade denotes the status of the transcript. Collected (Talk) means the transcript is available and the airline uses "capacity discipline," and Collected (No Talk) means the transcript is available but the airline does not use "capacity discipline."

Table 2 provides a summary of this airline data. The top panel reports the summary statistics when we use the airport-pair market definition, while the bottom panel reports the statistics when we use the city-pair definition. The statistics are very similar, except for the ones for Capacity-Discipline$_{mt}$, and so we will mostly focus on the airport-pair ones.

Row 1 of Table 2 shows that legacy carriers offer, on average, 30,150.38 seats in a month, and Row 2 shows that LCCs serve, on average, 14,826.567 seats in a month. Thus, LCCs offer half may seats as the legacy carriers. Next, we introduce variables that are important to identify the effect of the variable Capacity-Discipline$_{mt}$.

Capacity-Discipline$_{mt}$ is equal to 1 for 8.7 percent of the observations in our sample. Consistent with our focus on the communication of legacy carriers, as opposed to LCCs, we find that legacy carriers are far more likely to be in a market where Capacity-Discipline is equal to 1. While there are quantitative differences in the frequency of observations where Capacity-Discipline$_{mt}$ = 1 between airport and city markets, the qualitative result holds: legacy carriers are more likely to

18

Table 2: Summary Statistics

(a) Airport-Pair Markets

	Seats			Cap. Discipline		Talk Eligible		Monopoly Market		Missing Report		
	Mean	SD	Median	Mean	SD	Mean	SD	Mean	SD	Mean	SD	N
Carrier Type												
Legacy	11,783.793	12,297.048	7,374.000	0.087	0.281	0.308	0.462	0.549	0.498	0.267	0.442	561,008
LCC	11,407.016	10,626.587	8,220.000	0.031	0.175	0.105	0.306	0.473	0.499	0.097	0.296	279,141
Total	11,658.608	11,769.699	7,809.000	0.068	0.252	0.241	0.428	0.524	0.499	0.210	0.408	840,149

(b) City-Pair Markets

	Seats			Cap. Discipline		Talk Eligible		Monopoly Market		Missing Report		
	Mean	SD	Median	Mean	SD	Mean	SD	Mean	SD	Mean	SD	N
Carrier Type												
Legacy	13,095.678	16,094.787	7,420.000	0.108	0.311	0.401	0.490	0.441	0.497	0.285	0.452	504,644
LCC	12,178.875	15,019.836	8,220.000	0.077	0.267	0.279	0.449	0.285	0.451	0.168	0.374	261,102
Total	12,783.069	15,742.496	7,809.000	0.098	0.297	0.359	0.480	0.388	0.487	0.245	0.430	765,746

Notes. Table of summary statistic for all key variables. Observations are at the carrier-market-month level. Panel (a) refers to airport-pair markets while panel (b) refers to city-pair markets.

be in markets where pertinent communications take place.

We define the categorical variable $\mathtt{Talk\text{-}Eligible}_{m,t} \in \{0,1\}$ to be equal to one if there are at least two legacy carriers in market m in period t and zero otherwise. This variable controls for the possibility that markets where legacy carriers *could* engage in coordinating communication are fundamentally different from markets where such communications are not possible. Not including this control variable would confound the effect of talking on seats. Table 2a shows that, on average, 24% of the observations in our sample have the potential for coordinating communications. In a similar vein, markets served by a single carrier could differ from non-monopoly markets. We account for this possibility by introducing the categorical variable $\mathtt{MonopolyMarket}_{mt}$, which is equal to 1 if market m in period t is served by only one firm and equal to zero otherwise. Table 2a shows that, on average, 52.4 percent of the observations are monopoly markets, and that legacy carriers are more likely to serve a monopoly market than LCCs.

Finally, the categorical variable $\mathtt{MissingReport}_{mt}$ is equal to one if at least one of the carriers

Table 3: Summary Statistics by Market Type

(a) Airport-Pair Markets

	Seats			Cap. Discipline		Talk Eligible		Monopoly Market		Missing Report		
	Mean	SD	Median	Mean	SD	Mean	SD	Mean	SD	Mean	SD	N
Market Participants												
Mixed Market	13,478.067	12,842.555	9,079.000	0.057	0.232	0.194	0.396	0.322	0.467	0.146	0.353	409,518
Legacy Market	9,928.354	10,357.287	6,260.000	0.079	0.270	0.285	0.451	0.715	0.451	0.272	0.445	430,631
Market Size												
Small	5,161.338	5,198.658	3,811.000	0.005	0.070	0.027	0.163	0.846	0.361	0.202	0.402	110,859
Medium	9,777.528	9,011.037	7,137.000	0.040	0.197	0.144	0.351	0.603	0.489	0.193	0.395	411,209
Large	16,354.890	14,496.748	11,794.000	0.126	0.332	0.441	0.496	0.308	0.462	0.236	0.425	318,081
Business Travel												
Low Business	11,291.462	11,442.104	7,562.000	0.065	0.246	0.214	0.410	0.447	0.497	0.202	0.401	175,179
Medium Business	12,092.010	12,241.082	8,000.000	0.088	0.283	0.294	0.456	0.463	0.499	0.231	0.422	294,836
High Business	11,643.653	11,514.403	7,900.000	0.057	0.231	0.216	0.411	0.601	0.490	0.230	0.421	149,833
Total	11,658.608	11,769.699	7,809.000	0.068	0.252	0.241	0.428	0.524	0.499	0.210	0.408	840,149

(b) City-Pair Markets

	Seats			Cap. Discipline		Talk Eligible		Monopoly Market		Missing Report		
	Mean	SD	Median	Mean	SD	Mean	SD	Mean	SD	Mean	SD	N
Market Participants												
Mixed Market	15,537.611	18,078.250	9,472.000	0.115	0.320	0.424	0.494	0.160	0.366	0.224	0.417	465,353
Legacy Market	8,515.877	9,771.681	5,382.000	0.070	0.255	0.259	0.438	0.741	0.438	0.277	0.448	300,393
Market Size												
Small	4,685.647	4,695.543	3,536.000	0.006	0.078	0.032	0.176	0.848	0.359	0.204	0.403	78,831
Medium	8,535.886	7,961.011	5,932.000	0.046	0.210	0.161	0.367	0.562	0.496	0.202	0.401	319,331
Large	18,209.271	19,932.548	11,820.000	0.162	0.368	0.602	0.489	0.138	0.345	0.291	0.454	367,584
Total	12,783.069	15,742.496	7,809.000	0.098	0.297	0.359	0.480	0.388	0.487	0.245	0.430	765,746

Notes. Table of summary statistic for all key variables separated by market types. Observations are at the carrier-market-month level. Panel (a) refers to airport-pair markets while panel (b) refers to city-pair markets.

serving market m in period t is not holding an earnings call at time $t - 1$. As discussed above, we take special note of markets where we were unable to collect an earnings call transcript.[21] Table 2a shows that legacy carriers are more likely to be missing a report — a result of the bankruptcy periods of many of the legacies.

2.5 Market Types

In order to gain a more nuanced understanding of the effect of cheap talk on capacity, we consider three methods for differentiating markets: the composition of carriers serving the market, the size

---

[21] See Section 2.2 for a discussion of when and why we were unable to collect a transcript. Transcripts are missing for legacy carriers more often than for LCCs, largely due to the increased prevalence of bankruptcies in the legacy carriers.

of the market, and the proportion of business travelers served by the market.

To begin, much of our focus will be on markets served by legacy carriers. This is because the theoretical framework on which our analysis is grounded does not provide empirical predictions for markets in which there are firms that do not publicly communicate information on their strategic decisions. Thus, it is unclear whether firms that are not publicly communicating their decisions will free ride on the ones that communicate and maintain their capacity unchanged, or whether they will increase their own capacities to fill the void left by the firms that are colluding. Our empirical analysis will provide some evidence on the behavior of the LCCs, which we hope will inform future theoretical research.

We distinguish in our analysis between mixed and legacy markets, where the former are all markets made up of both legacy and LCC carriers or just LCC carriers. Legacy markets are those that are composed entirely of legacy carriers. Table 3 shows descriptive statistics for the variables defined in Section 2.4. In particular, we see that the average number of seats in mixed markets is greater than in legacy markets, but again, consistent with our focus on the behavior of legacy carriers, $\texttt{Capacity-Discipline}_{mt}$ is more likely to be equal to 1 in legacy airport-pair markets. Notably, when we define markets using city-pairs, mixed markets are more likely to have $\texttt{Capacity-Discipline} = 1$, a result of combining multiple airports into single city-based designations, which makes the threshold of all legacy carriers in a market discussing capacity discipline more difficult to meet.

We also distinguish markets based on the population of the market. We follow Berry, Carnall, and Spiller [2006] and define market size as the geometric mean of the Core-based statistical area population of the end-point cities. The annual population data is from the U.S. Census Bureau. We define markets as those with a population that is larger than the 75th percentile of the market

21

population distribution as large, markets with a population in the range of $(25^{th}, 75^{th}]$ percentiles of the population as medium, and those below the 25th percentile as small markets.[22]

Table 3 shows that the average number of seats a carrier offers, the likelihood of `Capacity-Discipline` $=$ 1, and the likelihood of `Talk Eligible` $= 1$ are all increasing with the size of a market. Perhaps unsurprisingly, the likelihood that a market is a monopoly market is decreasing with the size of the market. These qualitative results are maintained under the city-pairs market definition.

Finally, we investigate whether the composition of the market demand in business and tourist travelers affects the degree to which carriers respond to cheap talk. We follow Borenstein [2010] and Ciliberto and Williams [2014] and use a business index that is constructed using the 1995 American Travel Survey (ATS). The ATS was conducted by the Bureau of Transportation Statistics (BTS) to obtain information about the long-distance travel of people living in the U.S., and it collected quarterly information related to the characteristics of persons, households, and trips of 100 miles or more for approximately 80,000 American households. We use the survey to compute an index that captures the percentage of travelers out of an origin that are traveling for business.

We define a market's business travel index to be the computed travel index for the market's origin airport. In classifying markets based on their level of business travel, we follow the same approach as in our market size classifications. Low business markets are those with an index value at or below the 25th percentile, medium business markets have an index value in the $(25^{th}, 75^{th}]$ percentiles, and high business markets are those with an index above the 75th percentile. The average number of seats offered in a market is fairly constant across our business travel classifications, but communication is more common in low and medium business markets than in high business markets.

---

[22] When classifying small, medium, and large markets, we use the average market population over our sample period, so that a market's size classification is fixed across time.

2.6 Flexible Capacity

A prerequisite for firms to coordinate capacity decisions is that the capacity is non-binding and airlines have sufficient flexibility across markets. To get a quantitative sense of the ability of carriers to change capacity and move planes across markets, we used the OAG dataset to count the number of unique markets that each aircraft serves in a month. We found that, on average, an aircraft (identified by its tail number) operates in *79* unique markets in a month. This suggests that firms do not face capacity constraints at the quarterly level. Airline carriers can change the capacity across markets in multiple ways. They can remove a plane from a domestic market and park it in a hangar, they can move that plane to serve an international route, or they can reallocate that plane to another domestic market. The airlines can also change the "gauge" of an aircraft, i.e., increas or decrease the number of seats or by changing the ratio of business to coach seats.

3 Model

There are few salient features of the data that we want to capture with a theoretical model. First, airlines have long-run (repeated) interactions with each other. Second, airlines can change their capacity, the number of seats to make available in a market with ease. Third, in the airline industry it is hard for firms to monitor each other because firms see each other's capacities and listed prices, but do not observe the sales made by the competitors. More specifically, the firms do not know whether a seat was sold for a nonstop or connecting flight. Finally, and this is the key novelty here, the model must allow for the firms to publicly communicate.

We summarize the main and relevant results from the model developed by Awaya and Krishna [2016] and Awaya and Krishna [2017], and show that it captures all the features described above. The basic environment of repeated games is standard in the literature.

Stage Game. Suppose there are N airlines, and an airline i can choose an action a_i from a finite set A_i. Furthermore, suppose an airline may observe some information about the market, which we denote as $s_i \in S_i$, where S_i is a finite set of all signals. The actions chosen by all airlines determine the information received by each of them. In our context, a_i is firm i's capacity and s_i is the sales.

Let $\Delta(B)$ denote the set of all probability distributions on any set B. Then the signal that airlines receive is distributed as $f_s(\cdot|\mathbf{a}) \in \Delta(S_1 \times \cdots \times S_N)$, where $\mathbf{a} := (a_1, \ldots, a_N)$ is the vector of actions chosen by all airlines. In particular, suppose $N = 2$, then the choices (a_1, a_2) determine the signals s_1 and s_2 from a conditional probability distribution $f_s(s_1, s_2|a_1, a_2)$.

The ex-ante expected profit, before observing the the signal, is $\Pi_i(\mathbf{a}) = \sum_{s_i \in S_i} \pi_i(a_i, s_i) f_i(s_i|\mathbf{a})$, where $f_i(s_i|\mathbf{a})$ is the marginal distribution of airline i's signal, where $\pi_i(a_i, s_i)$ is the ex-post profit after the airline observes the signal s_i. The conditional probability $f_s(\cdot|\mathbf{a})$ is the monitoring technology available to the airlines. The monitoring is poor if it is hard for i's competitors to detect deviation by i.[23]

Repeated Game. Now suppose the game is repeated infinitely many times, and let $\delta \in (0, 1)$ be the discount factor.[24] Time is discrete, and in each period the airlines play the stage game and the profits are discounted by δ. Every period airlines know their action a_i^t, and at the end of the period they observe their signal s_i^t, and they can also perfectly recall their past actions and signals. We denote the past history of a firm by a set H_i^{t-1}. Let $h_i^{t-1} = (a_i^1, \sigma_i^1, s_i^1 \ldots, a_i^{t-1}, s_i) \in H_i^{t-1}$ denote firm i's private history. The strategy of airline i is a vector $\sigma_i = (\sigma_i^1, \sigma_i^2, \ldots)$ of functions

---

[23] For instance, if the marginal density on the signals i's competitors get $f_{-i}(\cdot|\mathbf{a})$ when i chooses a_i is very close to the marginal density when i chooses a_i' then i's competitors cannot "detect" whether i is using a_i or a_i'. The measure of monitoring quality depends on how we define the distance between two probability distribution, such as the total variation norm, which is related to the Kullback-Leibler divergence, a widely used concept in economics.

[24] This infinitely repeated game can also be viewed as a metaphor multi-market relationships between airlines [Bernheim and Whinston, 1990; Evans and Kessides, 1994]. For instance Ciliberto and Williams [2014] show that multi-market contacts among airlines can preserve their incentives to collude.

$\sigma_i^t : H_i^{t-1} \longrightarrow \Delta(A_i).$

Awaya and Krishna [2017] provide a bound on equilibrium payoff when communication is impossible. The bound depends on (i) the trade-off between the incentive to deviate in any period and the efficiency with which the cartel is able to achieve high profits; (ii) quality of monitoring; and (iii) the discount factor δ. The bound is tight when the quality of monitoring is poor –in our context, airlines hardly observe their opponents' all past actions. Then they construct equilibrium with communication that generates payoff that exceeds the bound.

Communication. Now suppose the airlines can communicate with each other by sending a message from a set $\mathcal{M}$. For example, this set of messages could be {`high`, `low`}, or it could be {`capacity-discipline`, `expansion`}. Every period, firms simultaneously choose **a**, then observe their signal s_i, and then send a non-binding message. Now, with communication the history h_i^{t-1} will include firm $i's$ own choices of past actions, past signals, and past messages sent $\{\mathbf{m}_i^1, \ldots, \mathbf{m}_i^{t-1}\}$ and messages received, $\{\mathbf{m}_j^1, \ldots, \mathbf{m}_j^{t-1}\}$. Therefore, $i's$ strategy is now a pair (s_i, r_i) where s_i is the vector of pricing strategies $s_i^t : H_i^{t-1} \to \Delta(A_i)$ and r_i^t is the communication strategy $r_i^t : H_i^{t-1} \times A_i \times S_i \to \Delta(\mathcal{M})$, with the interpretation that $r_i^t(h_i^{t-1}, a_i, s_i)$ is the message sent in period t by airline i who observes the history h_i^{t-1} chooses an action a_i and receives a signal s_i.

Equilibrium. Characterizing a collusive equilibrium with communication for this general environment is difficult and requires new concepts and notations that would be too much of a digression from the main theme of the paper. Instead, we take a middle-ground approach and present the main result from Awaya and Krishna [2016] where they consider a model of secret price cutting by Stigler [1964], so that the actions are the prices and signals are own sales. They characterize an equilibrium strategy that sustains collusion with cheap talk in a duopoly.

Awaya and Krishna [2016] maintain that the signal density $f_s(\cdot, \cdot | a_1, a_2)$ is log-normal and sat-

isfies the property that the sales are more correlated under collusion than under price wars.[25] In particular, they show that there are two threshold values of sales, which we refer to as μ_1 and μ_2 such that the monopoly (collusive) price can be sustained using the following grim-trigger strategy with communication:

1. Both firms start period 1 by choosing the monopoly price (p_1^M, p_2^M).

2. If, in period 2 onwards, both firms together announce "high" or "low," then they both continue charging monopoly price.

   (a) In any period $t \geq 1$, if the price set was $p_i = p_i^M$ then report "high" if sales are greater than μ_1; otherwise, report "low."

   (b) In any period $t \geq 1$, if the price set was $p_i \neq p_i^M$ report "high" if the sales are at least as large as μ_2; otherwise, report low.[26]

3. If in the previous period the messages do not match, then the firms revert to the Nash equilibrium permanently.

In the context of the airlines we can replace the message "high" with "capacity discipline" and "low" with (say) "capacity expansion" without changing the conclusion. When we dispense with the log-normality assumption, and allow for more than two firms, determining the thresholds (μ_1, μ_2) becomes difficult. We refer the interested reader to Awaya and Krishna [2017].

For our empirical analysis the key takeaways are a) the message space used in communication does not have to be rich to sustain collusion; b) there is an equilibrium where collusion is possible only because of firms' ability to communicate with each other by sending some messages; and c)

---

[25] Firms with high δ are indifferent between colluding in quantities or prices [Lambertini and Schultz, 2003].

[26] The threshold μ_1 is such that it increases with the mean of sales and decreases with its variance, and the threshold μ_2 depends on μ_1, the correlation between sales of the two airlines and their individual mean values.

these messages, in the form of some keywords, do not have to be too complicated. Next, we test if there is negative association between communication and available seats.
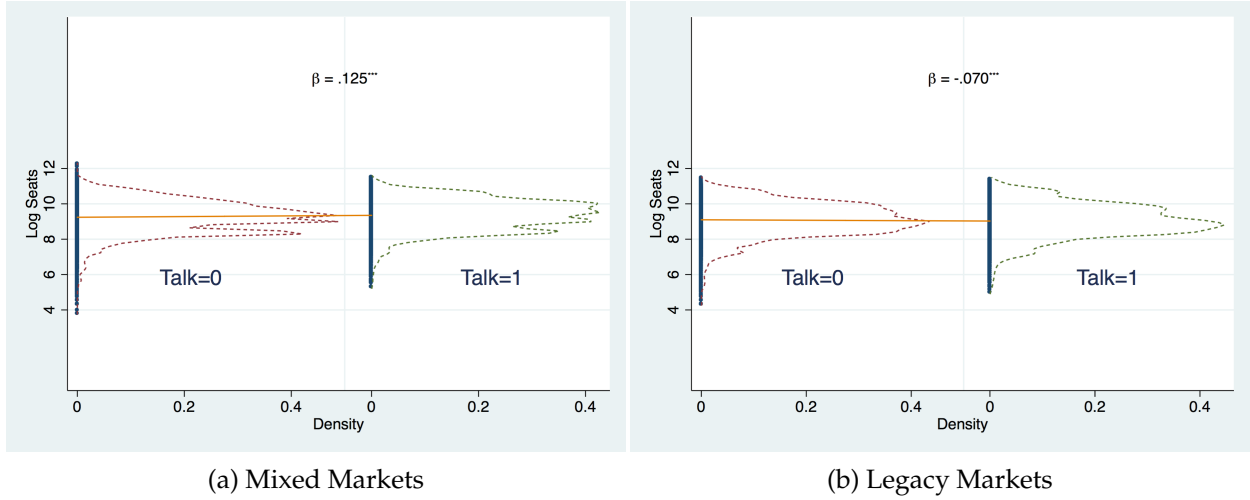
4 Empirical Analysis

In this section we examine the relationship between communication among airlines and the seats they offer between 2003:Q1 and 2016:Q3 (inclusive). We begin by considering the relationship observed in the raw data between log-seats and whether every legacy carrier operating in a given market communicated their intention to engage in capacity discipline. Across all non-monopoly markets, we find that capacity is 3.2% lower when legacy airlines "talk." In Figure 3 we present the raw data on the relationship between log-seats available and whether every legacy carrier operating in a given market communicated their intention to engage in capacity discipline in the previous quarter, broken down by whether a market consists of both legacy and LCC carriers ("mixed markets"), or just LCC carriers ("legacy markets"). As can be seen (Figure 3a), when all legacy airlines "talk" in mixed-markets it is correlated with a 13% increase in seats offered.[27] If, however, we consider markets served only by legacy carriers, i.e., legacy markets, then "talk" is correlated with approximately 7.0% fewer seats, on average. This exercise suggests that the collusion, if present, is not all-inclusive and occurs only among legacy carriers. Next, we estimate these effects after controlling for all relevant confounding factors.

To that end, we use the airline panel to estimate the following model for airline j in market m

---

[27] If the estimate of the coefficient of a dummy variable in a semilogarithmic regression is $\hat{\beta}$ then the percentage impact of the dummy on the outcome variable equal to $100 \times [\exp(\hat{\beta}) - 1]$ [Halvorsen and Palmquist, 1980].

Figure 3: Communication and Log-Seats

(a) Mixed Markets
(b) Legacy Markets

Notes. Relationship between communication (Talk) and log-seats available in the raw data. The unit of observation is market-year-month-carrier. Talk $= 1$ when all of the carriers in a market discuss "capacity discipline." Mixed markets are served by both legacy and LCCs while legacy markets are markets served only by legacy carriers .

in period t

$$\ln(\text{seats}_{j,m,t}) = \beta_0 \times \text{Capacity-Discipline}_{m,t} + \beta_1 \times \text{Talk-Eligible}_{m,t}$$

$$+ \beta_2 \times \text{Monopoly}_{m,t} + \beta_3 \times \text{MissingReport}_{m,t} \tag{1}$$

$$+ \mu_{j,m} + \mu_{j,yr,q} + \gamma_{origin,t} + \gamma_{destination,t} + \varepsilon_{j,m,t},$$

where, the dependent variable $\ln(\text{seats}_{j,m,t})$ is the log of total seats made available by airline j in (airport-pair) market m in time t. We estimate this model using a within-group estimator The main variable of interest is $\text{Capacity-Discipline}_{m,t} \in \{0,1\}$, which is the dummy variable introduced in Section 2.2 that is equal to one if there are at least two legacy carriers in market m and they all communicate about capacity discipline in the previous quarter's earnings call. Another variable is the dummy variable $\text{Talk-Eligible}_{m,t} \in \{0,1\}$ that is equal to one if there are at least two legacy carriers in market m in period t, and zero otherwise. In order to investigate the role of communication among airlines, we differentiate the monopoly markets from markets with more

Table 4: Identification

market	market structure	DL reports	communicating	Cap-Dis	Report	Monopoly	Talk-Eligible	parameters
1	$\{DL\}$	no	n/a	0	1	1	0	$\beta_3 + \beta_2$
2	$\{DL\}$	yes	n/a	0	0	1	0	β_2
3	$\{DL, UA\}$	yes	$\{DL, UA\}$	1	0	0	1	$\beta_0 + \beta_1$
4	$\{DL, UA, US\}$	no	$\{US\}$ or $\{UA\}$ or $\{US, UA\}$	0	1	0	1	$\beta_3 + \beta_1$
5	$\{DL, UA, US\}$	yes	$\{US, UA\}$	0	0	0	1	β_1
6	$\{DL, UA, US\}$	yes	$\{DL, UA, US\}$	1	0	0	1	$\beta_0 + \beta_1$
7	$\{DL, UA, US, F9\}$	yes	$\{DL, UA, US\}$	1	0	0	1	$\beta_0 + \beta_1$
8	$\{DL, F9\}$	yes	n/a	0	0	0	0	-

Notes. An example to show identification from the perspective of Delta, i.e., when $j = DL$, and here UA and US are legacy carriers while F9 is an LCC.

than one airline, both because the notion of communication is moot with only one airline and because monopoly markets can be inherently different from non-monopoly markets. To capture this, we use the dummy variable $\texttt{Monopoly}_{m,t} \in \{0, 1\}$ that is equal to one if only one airline serves market m in period t. In some cases, earnings call reports are missing (for reasons that are unknown to us), which we control for by including a dummy $\texttt{MissingReport}_{m,t} \in \{0, 1\}$ that is equal to one if any of the carriers in market m are missing a report in period t.

The idea behind capacity discipline is that airlines restricted seats even when there was adequate demand, which itself can vary across both markets and time. To control for these unseen factors, we include airline-market, airline-year-quarter fixed effects. These fixed effects allow airlines to provide different levels of capacities across different markets and time. Lastly, to control for time-dependent changes in demand we use origin- and destination-airport specific time trends, $\gamma_{origin,t}$ and $\gamma_{destination,t}$. These controls are important in isolating the direct effect of communication on available seats.

Next, we explain the identification strategy behind our estimation. To highlight the key sources of variation in the data, we fix an airline, say, Delta (i.e., $j = DL$), and consider different potential market structures and communication scenarios in Table 4. In markets $m = 1, 2$, only DL operates, so the concept of communication is moot and $\texttt{Capacity-Discipline}_{1,t} = \texttt{Capacity-Discipline}_{2,t} =$

0. Then we can use variation in whether a report is available (for $m = 2$) or not (for $m = 1$) to identify β_3 and β_2, as shown in the last column. Market $m = 3$ is served by both DL and UA and both use "capacity discipline" in the previous quarter, so `Capacity-Discipline`$_{3,t} = 1$, which identifies $\beta_0 + \beta_1$. The same identification argument applies to identifying $\beta_0 + \beta_1$ in markets $m = 6, 7$ where every airline in the market talks and a report for DL is available, even when a LCC is present ($m = 7$). In contrast, for market $m = 4$, even when both US and UA use cheap talk, we identify $\beta_1 + \beta_3$, because DL did not have a transcript. Lastly, we identify the fixed-effects using the deviation from the mean. One of the key sources of variation is the variation in `Capacity-Discipline` across markets and over time; see Figure 2. We also assume that conditional on all control variables, `Capacity-Discipline` is uncorrelated with the error. In other words we assume conditional exogeneity of the treatment, which is sufficient condition for identifying effects of `Capacity-Discipline` on log-seats [Rosenbaum, 1984]. We verify that this assumption holds in section 5.

Results. We present the estimation results from Eq. (1) in Column (1) of Table 5. Recall that we showed above that when legacy carriers engaged in discussion about capacity discipline, capacity was 3.2% lower. Using our model to control for a rich set of potentially confounding factors, we find that when all of the legacy carriers in a talk-eligible market communicate with each other about capacity discipline, there is a 1.45% decrease in the number of seats offered. This effect is an average effect across all markets, time, and types of carriers. The standard errors we report are the robust standard error, and, as can be seen, the decline is statistically significant at 1%. To get a sense of whether this effect is economically meaningful or not, it is helpful to compare it to the average percentage change in capacity for legacy airlines in our sample. The average percentage change is 3.78%, while the use of the phrase capacity discipline results in a 1.45% percentage drop

30

in capacity. This means whenever legacy airlines communicate, their capacity drops by 38% of the average change in capacity, which is a significant effect. To interpret this estimate from a regression with fixed effects but without a plausible instrumental variable as a causal effect we have to rule out the possibility of false positive and counfoundedness of `Capacity-Discipline`. In section 5 we address both these issues, respectively, by developing and implementing placebo tests and verifying conditional exogeneity.

Interestingly, we find that if a market is `Talk-Eligible`, that also leads to a 12.55% decrease in number of seats offered on average. This shows that it is important to control for market heterogeneity that treats markets with at least two *legacy* carriers differently from other markets, and the estimate shows that in some markets, the offered capacity can be low for reasons that are not associated with communication. In summary, we can reject our null hypothesis that communication regarding capacity discipline does not affect carriers' capacity decisions. That is, we find evidence in support of the claim that carriers are using this communication to coordinate capacity decisions.

The features of the raw data that a) the effect is negative only for the legacy markets (see Figure 3); and b) legacy carriers communicate about capacity discipline more frequently than LCCs (see Table 1) suggest that the average effect we find among all airlines is driven primarily by the legacy carriers, possibly because the cartel includes only the legacy carriers. To determine that, we extend the basic model and allow the effect of public communication to vary by carrier type and by whether the market is a legacy-only or mixed market, made up of both legacy and LLC carriers or

31

Table 5: The Effect of Communication on Available Seats

	(1) Log Seats	(2) Log Seats
Capacity Discipline	-0.01495*** (0.00241)	
Legacy Market × Capacity Discipline		-0.01462*** (0.00305)
Mixed Market × Capacity Discipline (Legacy)		-0.01838*** (0.00433)
Mixed Market × Capacity Discipline (LCC)		-0.00741 (0.00471)
Talk Eligible	-0.13230*** (0.00317)	-0.11811*** (0.00327)
Market Missing Report	0.01723*** (0.00234)	0.01923*** (0.00234)
Monopoly Market	0.05392*** (0.00233)	0.07723*** (0.00273)
Legacy Market		-0.05417*** (0.00335)
R-squared	0.866	0.866
N	840149	840149

Notes. Standard errors are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

just LCC carriers. With this in mind, we estimate the following model:

$$\ln(\texttt{seats}_{j,m,t}) = \beta_0^{\texttt{legacy}} \times \texttt{Capacity-Discipline}_{m,t} + \beta_0^{\texttt{LCC}} \times \texttt{Capacity-Discipline}_{m,t}$$

$$+\beta_1 \times \texttt{Talk-Eligible}_{m,t} + \beta_2 \times \texttt{Monopoly}_{j,m,t} + \beta_3 \times \texttt{MissingReport}_{j,m,t}$$

$$+\mu_{j,m} + \mu_{j,yr,q} + \gamma_{origin,t} + \gamma_{destination,t} + \varepsilon_{j,m,t}. \qquad (2)$$

For those markets where only legacy carriers operate, i.e., the legacy markets, we set $\beta_0^{\texttt{LCC}} = 0$. The identifying assumption for (1) applies verbatim to (2).

We present the results in Table 5, column (2). The three variables of importance are in the second, third, and fourth rows. As we can see, in markets that are served by only legacy carriers,
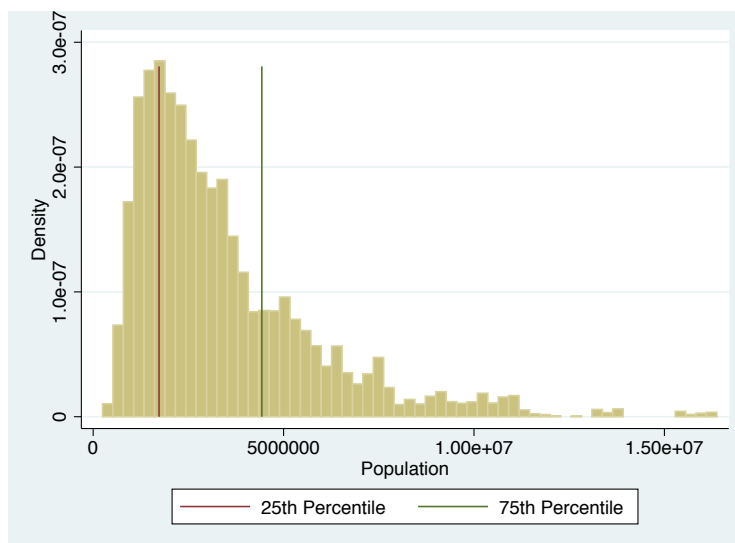
communication leads to a 1.45% decrease in the number of seats offered. This result is statistically significant at 1% and is also similar in magnitude to the estimates above. This also suggests that the average effect we found earlier must be entirely driven by the effect among legacy carriers. To assess that hypothesis, consider the third and the fourth rows. We find that, indeed, the effect of communication among legacy carriers increases to a 1.82% decrease in seats offered by legacy carriers, whereas we find no evidence of a significant effect on seats offered by LCCs.

In summary, we find evidence that supports the hypothesis that there is collusion only among legacy carriers. In fact, the LCCs in general not only do not communicate about capacity discipline, but they also do not respond to the communication by legacy carriers.

This leads us to our next question: If LCCs are neither communicating nor colluding with the legacy carriers, do legacy carriers respond (to communication) differently by market size to balance the threat from LCC? Below we show that the estimated effect above is primarily driven by legacy carriers reducing seats in small and medium-sized markets, where competition from LCC tends to be weaker.

Market Sizes. The ability of the legacy airlines to collude depends on how well they can monitor each other and the contestability of the markets. As noted by Stigler [1964], it is easier to collude in some markets than others. If larger markets have larger demand volatility than the smaller markets, it is easier to sustain collusion in the latter markets, ceteris paribus. As defined in Section 2.4, we categorize markets into three categories: small, medium, and large, depending on whether the (geometric mean of) the population at the two ends of the market is less than 25th percentile, between 25th and 50th percentile, or greater than 75th percentile of the population distribution, respectively. Figure 4 shows the histogram of the population with markers for 25th and 75th percentiles. When we consider the distribution of passengers transported by these three
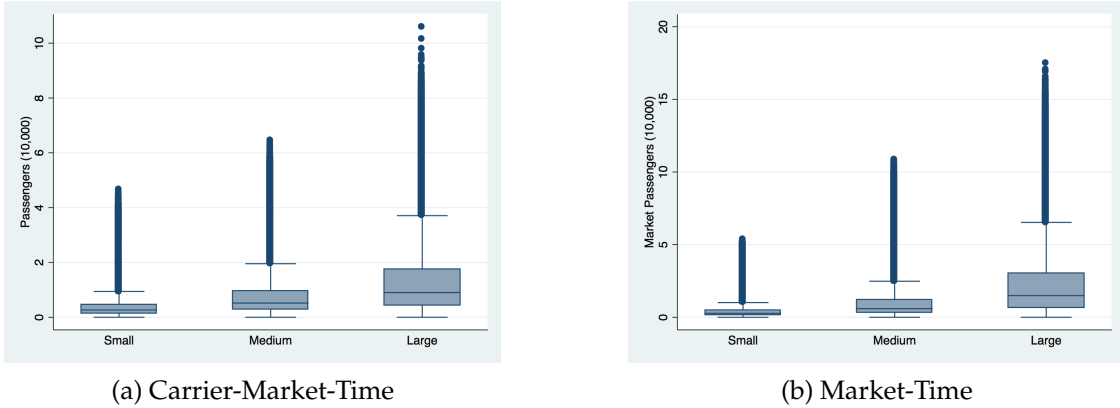
Figure 4: Histogram of Market Sizes

Airport-Pair

Notes. Market size is defined as the geometric mean of the MSA population of the end-point cities. Source for population data is the U.S. Census Bureau.

categories (see Figure 5), we find that markets with larger populations are more dispersed than in smaller markets. This is true both when the unit of observation is carrier-market-time, as in Figure 5a, and when we aggregate it to the market-time level, as in Figure 5b. Larger markets not only have a wider inter-quartile range, but they also have longer whiskers (outliers) than smaller and medium markets, which is consistent with the demand uncertainty increasing with market sizes.

Furthermore, larger markets can also accommodate more firms [Bresnahan and Reiss, 1991], and, given that our estimates so far suggest that this is not an all-inclusive cartel, legacy airlines might not reduce their supply because LCC would then meet the excess demand. To assess the role of market size on the intensity of collusion, we estimate the following model that allows the

Figure 5: Box plot of Passengers by Market Size

(a) Carrier-Market-Time

(b) Market-Time

Notes. These are the box-plots with whiskers of sales of tickets by market sizes. On the x-axis are the market sizes, small, medium, and large, and on the y-axis is the total number of passengers transported in that market. The unit of observations in subfigure (a) is carrier-market-time, whereas the unit of observation in subfigure (b) is market-time.

effect of communication to differ by market size, i.e.,

$$
\begin{aligned}
\ln(\texttt{seats}_{j,m,t}) \;=\; & \beta_0^{\texttt{small}} \times \texttt{Capacity-Discipline}_{m,t} + \beta_0^{\texttt{medium}} \times \texttt{Capacity-Discipline}_{m,t} \\
& +\beta_0^{\texttt{large}} \times \texttt{Capacity-Discipline}_{m,t} + \beta_1 \times \texttt{MissingReport}_{j,m,t} + \beta_2 \times \texttt{Monopoly}_{j,m,t} \\
& +\beta_3 \times \texttt{Talk-Eligible}_{m,t} + \mu_{j,m} + \mu_{j,yr,q} + \gamma_{origin,t} + \gamma_{destination,t} + \varepsilon_{j,m,t}.
\end{aligned} \tag{3}
$$

We present the estimation results from (3) in column labeled (4) in Table 6. The only difference between this model and our primary result is that now communication can have a different effect on supply depending on the size of the markets. In the second column of Table 6, we find that communication among legacy carriers leads to a large 4.21% reduction in seats supplied in smaller markets on average. This effect is statistically significant at 1%. The fact that we find that the effectiveness of communication is stronger in smaller markets is consistent with colluding being easier and more profitable in smaller markets. Moreover, we find that the negative effect of communication on available seats decreases to 1.95% and 1.25% in medium and large markets. Thus, we find evidence that the level of collusion is inversely proportional to the size of the markets.

35

Table 6: Fixed Effects Estimates of Communication on Available Seats Separated by Market Sizes

	(3) Log Seats	(4) Log Seats	(5) Log Seats
Capacity Discipline	-0.01465***		
	(0.00241)		
Small Population × Capacity Discipline		-0.04302**	
		(0.01323)	
Medium Population × Capacity Discipline		-0.01970***	
		(0.00404)	
Large Population × Capacity Discipline		-0.01256***	
		(0.00273)	
Low Business × Capacity Discipline			-0.02778***
			(0.00460)
Medium Business × Capacity Discipline			-0.02189***
			(0.00349)
High Business × Capacity Discipline			0.01510**
			(0.00526)
Log Population	1.32447***		
	(0.04570)		
Talk Eligible	-0.13410***	-0.13215***	-0.12767***
	(0.00316)	(0.00317)	(0.00361)
Market Missing Report	0.01436***	0.01724***	0.01203***
	(0.00234)	(0.00234)	(0.00273)
Monopoly Market	0.05356***	0.05384***	0.05296***
	(0.00233)	(0.00233)	(0.00262)
R-squared	0.866	0.866	0.863
N	840,149	840,149	619,848

Notes. Standard errors are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

An alternative way to control for market size is to treat market size as a continuous control variable and add it (after taking log) in the primary regression model, Eq. (1). The results from this fourth model are presented in the second column of Table 6. The variable of interest in this column is `Capacity Discipline`. As can be seen, we find that communication among legacy carriers reduces seats by 1.45%. Thus, we find that the legacy carriers reduced their capacity by a larger number in smaller markets than they did in medium or larger markets.

Business Markets. While we have focused only on reasons why smaller markets might be more

conducive to collusion, there is an counterargument against small markets, as follows. Larger markets tend to have a greater share of for-business travelers, who tend to have a higher willingness to pay for a ticket; ceteris paribus, i.e., they have (relatively) more inelastic demand for air travel than those who travel for leisure. This in turn implies that these markets should have higher mark-ups than smaller markets, and thus be more profitable for collusion.

To understand how business travelers might change the effect of communication on offered seats, we consider low, medium, and high Business markets, based on the proportion of for-business travelers originating from that market (c.f. Section 2.4). Then we estimate the following model

$$
\begin{aligned}
\ln(\texttt{seats}_{j,m,t}) \;=\; & \beta_0^{\texttt{low-business}} \times \texttt{Capacity-Discipline}_{m,t} + \beta_0^{\texttt{medium-business}} \times \texttt{Capacity-Discipline}_{m,t} \\
& + \beta_0^{\texttt{high-business}} \times \texttt{Capacity-Discipline}_{m,t} + \beta_1 \times \texttt{MissingReport}_{j,m,t} + \beta_2 \times \texttt{Monopoly}_{j,m,t} \\
& + \beta_3 \times \texttt{Talk-Eligible}_{m,t} + \mu_{j,m} + \mu_{j,yr,q} + \gamma_{origin,t} + \gamma_{destination,t} + \varepsilon_{j,m,t},
\end{aligned}
$$

that allows the effect of communication on offered seats to differ by the proportion of for-business travelers in that market. If we find that high-business markets have a larger effect of communication on offered seats, then we would have to reconsider our previous hypothesis that smaller markets are indeed more conducive to collusion.

We present the results from this regression in the third column, marked (5), in Table 6. The three variables of interest are in fifth, sixth, and seventh rows. The first corresponds to the effect on low-business markets, and, as we can see, we find that communication is associated with a 2.74% decrease in the number of seats offered. This decrease is also statistically significant at 1%. What is interesting is that the effects of communication are smaller for medium-business markets at -2.17%, and, in fact, they lead to an increase in the number of offered seats by 1.52% in

high-business markets. Although the effects on low and medium-business markets are statistically significant at 1%, the difference between the two are not statistically significant. Thus, we cannot reject the null that the effects in these two markets are similar.[28] On the other hand, the effect on high-business markets is statistically different from the other two, and the fact that we find a positive effect of communication means that when it comes to collusion, the differences in elasticity are less important than the threat of entry by LCCs and demand uncertainty.

City Pairs. So far we have used airport-pair as our definition of a market. An alternative definition of a market posits that we should consider city-pair as the market because consumers in a city might be served by multiple airports, and, given the hob-and-spoke network, access to airport(s) might affect an airline's market power [Ciliberto and Williams, 2010; Snider and Williams, 2015]. This means the difference between these two definitions of a market for our analysis is that while airport-pair always have only one airport there can be multiple airports under a city-pair method. This change can make (detection of) collusion more difficult because it can not only make monitoring more difficult and the demand faced by airlines operating from different airports less correlated it can also make airline's residual demand at the city level more uncertain because the market size is greater and served by more airlines than under the airport-pair definition of market.

As a consequence of these effects, we expect to estimate a smaller effect of communication on offered capacity. The effect on smaller markets, however, is ex-ante ambiguous. But to keep the analysis comparable, we have to "hold" the feasibility of collusion the same. To that end, we have to differentiate markets with three or more airports from markets with at most two airports because it is much more difficult to sustain collusion in markets with three airports.[29]

---

[28] When we allow the estimates to differ by both market size and carrier type, the qualitative results do not change.

[29] A rough analogy can be made with the spatial competition, where the equilibrium is stable if there are only two firms [Hotelling, 1929], but if there are three or more firms they have a strong tendency to disperse with no pure strategy Nash equilibrium [Cox, 1990; Eaton and Lipsey, 1975; Osborne, 1993], making demand uncertain.

For that we first use the same specification as (1) except with the city-pair definition of the markets. The results are in the first column (numbered 6) of Table 7. The interpretation of all variables is the same, and the coefficient of interest for us is the first row, which shows that communication does not decrease offered seats. In fact, the effect is slightly positive and statistically significant. This result is consistent with what we would expect with the city-pair markets. Next, we allow the effect to vary by market sizes and, as mentioned above, by whether the city is served by less than three airports. The results are in second column (numbered 7) of Table 7. The most important result is that in small markets that have less than three airports, we see that communication leads to 4.16% fewer offered seats, and this effect is statistically significant at 10%. What is important to note is that this effect is similar to the effect we found for the airport-pair markets. When we consider medium-sized markets with less than three airports, the effect is smaller at -1.36%, but it is still statistically significant at 10%. However, for larger markets or markets with more than three airports, we cannot reject the null that communication about capacity discipline has no effect on the number of offered seats in those markets.[30]

5 Falsification Tests

In this section we seek to verify that our result — legacy airlines use public communication regarding capacity discipline to collude — is not driven by spurious effects. We approach this in two ways: First, we conduct a series of placebo falsification tests, wherein we identify words that are unlikely to be a part of the collusive vocabulary, and test that the simultaneous use of those words by all legacy carriers in a market is not associated with a decline in capacity. Second, we consider our assumption that our regression model satisfies unconfoundedness (henceforth, conditional exogeneity), by conducting tests motivated by White and Chalak [2010].

---

[30] As our business index is calculated at the airport-level, we do not consider the level of business travel here.

Table 7: The Effect of Communication on Available Seats Separated by Market Sizes

	(6) Log Seats	(7) Log Seats
Capacity Discipline	0.00717***	
	(0.00213)	
Talk Eligible	-0.12151***	-0.12070***
	(0.00276)	(0.00276)
Market Missing Report	0.02162***	0.02140***
	(0.00218)	(0.00218)
Monopoly Market	0.04863***	0.04861***
	(0.00254)	(0.00254)
Small Population × Capacity Discipline (Cities w/ < 3 Airports)		-0.04251*
		(0.02107)
Medium Population × Capacity Discipline (Cities w/ < 3 Airports)		-0.01372*
		(0.00549)
Large Population × Capacity Discipline (Cities w/ < 3 Airports)		0.00028
		(0.00253)
Small Population × Capacity Discipline (Cities w/ ≥ 3 Airports)		0.29952***
		(0.05174)
Medium Population × Capacity Discipline (Cities w/ ≥ 3 Airports)		0.08708***
		(0.00698)
Large Population × Capacity Discipline (Cities w/ ≥ 3 Airports)		0.00520
		(0.00346)
R-squared	0.872	0.872
N	765,746	765,746

Notes. Standard errors are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
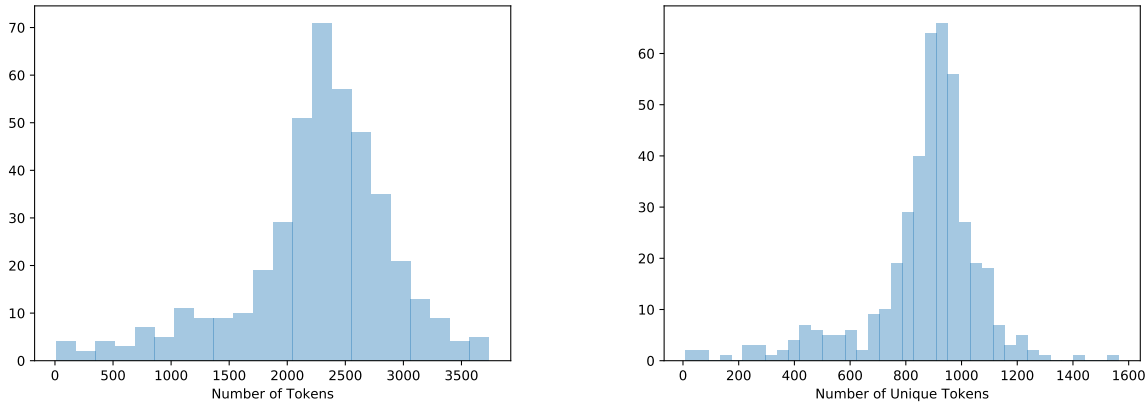
Placebos. Our maintained hypothesis is that legacy airline executives use earnings calls to coordinate on capacity discipline: discussion of capacity discipline translates into fewer seats in airline markets. Even though our results are consistent with the prediction of Awaya and Krishna [2016, 2017], our empirical strategy relies on the assumption that the only channel through which airlines coordinate is through the use of keywords associated with the concept of capacity discipline. In other words, our assumption implies that if we replace capacity discipline with other concepts, we would find no effect. This fact presents us with an opportunity to test the validity of our results.

To do so, we run multiple placebo falsification tests where we show that other keywords are not associated with declines in capacity. A placebo falsification test consists of choosing a token — a keyword or commonly used, two-word key-phrase (see Section 2.2 for a broader discussion of how tokens are constructed) — from the earnings calls that is i) unrelated to capacity discipline; and ii) discussed approximately as frequently as capacity discipline. Then, we regress our measure of offered capacity (log-seats) on that keyword and investigate systematically whether there are fewer seats offered when airline executives use that word.

The construction of such a placebo test poses two primary challenges. First, the words or phrases used in the placebo tests, henceforth "placebo tokens," cannot be close in meaning to the notion of capacity discipline. This restriction is important because it's possible that carriers are communicating via words that we have yet to identify as being part of the collusive vocabulary. Imagine, for example, that we had not identified "GDP" as a term that was often used when carriers discussed capacity discipline. If, in this hypothetical, we had chosen "GDP" as a placebo token, we would have found a non-zero effect on capacity — not because our specification is flawed or failed to account for a given confounding factor in carriers' capacity decisions, but instead because we had used a narrower set of messages sent by airlines.

Second, the placebo tokens should not be used in a way that substantially overlaps with the use of capacity discipline, even if they do not have the same meaning. The issue with such words results from the fact that our analysis uses a binary classification of "talk." That is, we define a carrier as "talking" in a given quarter if its executives discussed capacity discipline at least once at any point during the relevant earnings call. If, for example, carriers discussed the "holidays" in *every* quarter that they also discussed capacity discipline, then our analysis would also find an effect of the discussion of "holidays" on capacity.

Figure 6: Histogram of Transcripts Length and Unique Tokens

(a) Density of Transcript Length

(b) Density of Unique Tokens in Transcripts

Notes. Transcript length is measured in the number of tokens (one- or two-word phrases) in the transcript. Both the transcript length and the number of unique tokens are computed after removing common "stop words" and proper nouns from the transcripts.

Due to the large number of words in our dataset, addressing both of these challenges is not trivial. After removing commonly used "stop words" and proper nouns, we observe 26,680 unique tokens used across all of the transcripts in our dataset. On average, a transcript contains 2,215 total tokens, 852 of which are unique. Fig. 6 shows the full distribution of the length of the transcripts and the number of unique tokens in each transcript. Limiting our attention to just those transcripts we identified as discussing capacity discipline, we observe 18,427 unique tokens, accounting for 69% of the total vocabulary observed in earnings calls.

In light of these concerns, any restriction imposed to shorten the list of possible tokens is bound to be subjective. In an attempt to be as objective as possible, we employ the word2vec model from computational linguistics [Mikolov, Chen, Corrado, and Dean, 2013], to identify placebo tokens.[31] Broadly, the word2vec model maps each unique token we observe in the earnings call transcripts

---

[31] The model word2vec was developed at Google in 2013 [Mikolov, Chen, Corrado, and Dean, 2013] to analyze text data. For a more intuitive and accessible treatment of the model, see Goldberg and Levy [2014]. In practice, we use the gensim implementation of the word2vec model [Řehůřek and Sojka, 2010].

to an N-dimensional vector space (in our analysis, $N = 300$), in such a way as to preserve the contextual relationships between the tokens. The vector representation of each token is such that tokens that are similar in purpose/meaning are located "close" to each other, and tokens that are more dissimilar are located "farther" away from each other. We directly train the `word2vec` model using our transcript data, so the derived relationships between words are specific to the context of airlines' earnings calls, as opposed to a more general context. Thus, for example, if airline executives use the word "discipline" in a contextually different manner than it is used in in more general conversation or writing, our model will account for that.

To measure the similarity of two tokens in the `word2vec` vector space, we use a commonly used metric called the cosine similarity metric, which is defined as the cosine of the angle between the vector representation of the two tokens; see, for example, Singhal [2001]. Given the normalized vectors for two tokens, k and ℓ, this measure of similarity is defined as
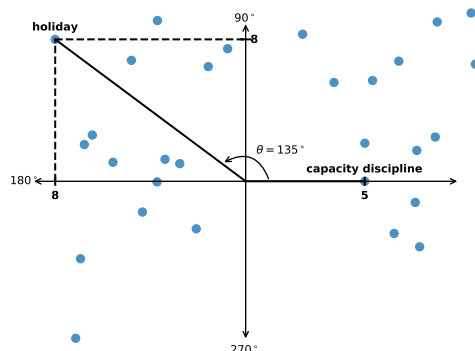
$$d^{\cos}(\ell, k) = \frac{k^T \ell}{||k|| \cdot ||\ell||},$$
(4)

where $|| \cdot ||$ is the L^2 norm of the vector. When two vectors are the same, cosine similarity is 1; when they are totally independent (perpendicular) to each other, then the similarity is 0; and when the angle is 180 degrees apart, the cosine similarity is -1.[32]

To understand our use of cosine similarity, consider Fig. 7, which displays a hypothetical example of training the `word2vec` model in a 2-dimensional space. The `word2vec` model maps all of the tokens in our vocabulary to this space. For example, the token "capacity discipline" is represented by the vector $(5, 0)$, and the token "holiday" is represented by the vector $(-8, 8)$. Our measure of similarity between these two tokens is the cosine of the angle between these two vectors, $\theta = 135°$.

[32] Note that the cosine metric is a measure of orientation and not magnitude. This metric is appropriate in our cases, as we are interested in comparing the contextual meaning of the words, not in comparing the frequency of the words.

Figure 7: Example of Placebo Token Selection Process

Notes. A schematic illustration of a hypothetical `word2vec` model. Tokens are mapped to a vector space, such that the cosine of the angle between two tokens represents the level of "similarity" between those tokens. In the case above, "holiday" is very dissimilar to "capacity discipline."

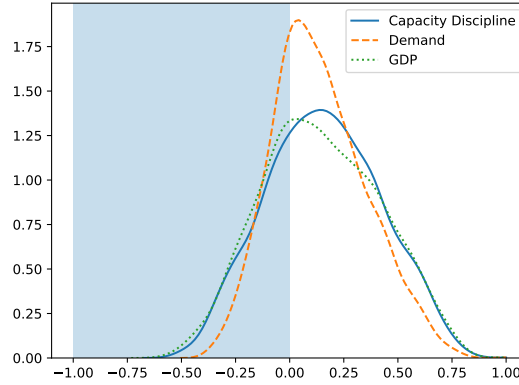In this example, $d^{cos}(\texttt{holiday}, \texttt{capacity discipline}) = -0.707$, so "holiday" is very dissimilar to "capacity discipline."

In order to construct a set of placebo tokens, we identify three tokens that are essential to the concept of capacity discipline: "capacity discipline," "demand," and "gdp." For each of these tokens $k \in \{\texttt{capacity discipline, demand, gap}\}$, we define the set:

$$L_k(\underline{d}, \overline{d}) = \left\{ \ell \in L : \underline{d} \leq d^{\texttt{cos}}(\ell, k) \leq \overline{d} \right\},$$

where L is the set of all tokens.

For our placebo tests, we select the tokens that are least similar to each of "capacity discipline," "demand," and "gdp" as our placebo tokens. For each of these tokens we set $(\underline{d}, \overline{d}) = (-1, 0)$, which captures tokens that fall in the shaded region of Fig. 7. In Fig. 8, we present the probability densities of the `cos` similarity, d^{cos}, of all tokens from "capacity discipline," "demand," and "gdp." The shaded region of 8 indicates the portion that fall in the region accounted for by the sets L_k.

44

Figure 8: Example of Placebo Token Selection Process

Notes. These are the Parzen-Rosenblatt Kernel densities of `cos` distance from "capacity discipline" (solid-line) "demand" (dotted line), and "gdp" (dashed-line).

Table 8: Placebo-Tokens

involve	distribution	proceed	slot	obligation
operator	security	negotiation	remind	handle
member	old	free	engine	negotiate
liability	worth	approximately million	equity	table
retirement	convert	accounting	marketing	flight attendant
rule	form	fix	final	budget
conference	directly	credit card	extend	requirement
president	apply	save	minute	government

Notes. The list of all placebo tokens that are dissimilar from either "capacity discipline," "gdp," or "demand."

Finally, to find the tokens least closely related to the notion of capacity discipline, we define our set of placebo tokens
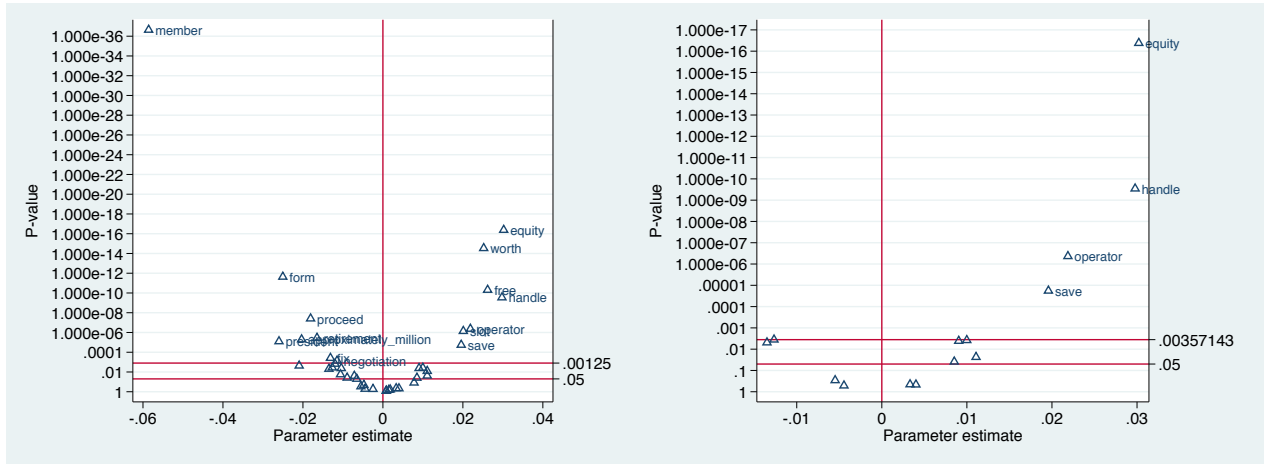
$$L^{placebo} := \bigcap_{k \in \{\texttt{cap-dis},\texttt{demand},\texttt{gdp}\}} L_k(\underline{d}, \overline{d}).$$

We find that 40 placebo tokens satisfy this criteria, all of which are listed in Table 8. For each placebo-token, we follow the same procedure as we did with "capacity discipline." Our primary econometric model of interest is (1), but we measure communication by a new dummy variable $\texttt{Placebo}_{m,t} \in \{0,1\}$ that is equal to one if there are at least two legacy carriers in market m in time t and they all use a placebo-keyword in the previous period.

Figure 9a displays a "Smile" plot of the results of our placebo regressions. This figure plots the estimated coefficients associated with the placebo tokens on the horizontal axis and the p-value for a 95% confidence level. The two horizontal lines represent the 5% significance level and the "corrected" 0.125% significance level, which has been (Bonferroni) corrected for multiple hypothesis testing; see Romano, Shaikh, and Wolf [2010]. We find that we can reject the null hypothesis of a placebo token having no effect on carriers' capacity in 15 of 40 placebo tests. However, this ignores that some of these placebo-tokens might overlap, or appear very frequently with our measure of communication. In other words, while the word2vec model ensures that we have selected tokens that are most dissimilar to capacity discipline, we still find significant overlap between $\texttt{Capacity-Discipline}_{m,t}$ and $\texttt{Placebo}_{m,t}$. Consider, for example, the placebo token "member," which is located in the top left of Figure 9a. In nearly 72.7% of observations where $\texttt{Placebo}_{m,t} = 1$, $\texttt{Capacity-Discipline}_{m,t}$ is also equal to 1. In Figure 10 we show the distribution of the percentage of cases where $\texttt{Capacity-Discipline}_{m,t} = 1$ conditional on $\texttt{Placebo}_{m,t}$ being equal to 1.

In view of this overlap it stands to reason that a valid placebo token should not only be dissimilar (in the sense of meaning and context) from our measure of communication, it should not

Figure 9: Fixed Effect Estimates of Communication of Placebo-Tokens on Seats

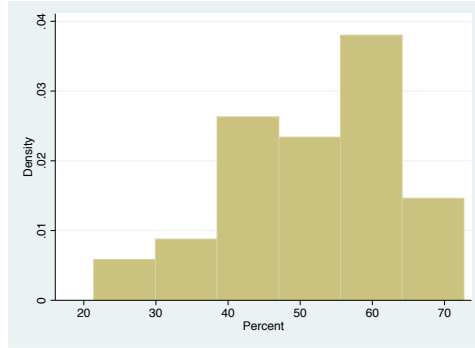(a) "Smile" Plot for all placebo tokens.

(b) "Smile" Plot for placebo tokens with relatively low overlap with "capacity discipline."

Notes. The table presents the fixed effect estimates of the coefficient on placebo-tokens in (1) with the (Bonferroni) corrected 95% confidence intervals. The vertical line at zero denotes the null of no effect.

also occur frequently with latter. Having said this, it is not clear what is the right cut-off, so we make a judgement call and say that a $\text{Placebo}_{m,t}$ in Table 8 is a valid placebo if it does not overlap with $\text{Capacity-Discipline}_{m,t}$ more than 50% of the time in our sample. Thus, in Fig. 9b we consider only those placebo-tokens with less than 50% overlap with $\text{Capacity-Discipline}$. Among tokens that have the lowest levels of overlap with capacity discipline, we find no evidence that the placebo tokens result in decreased capacity.

Conditional Exogeneity. Although we employ a rich set of fixed-effects and other covariates (henceforth, X) as control variables, it is desirable to verify that our finding is not driven by a missing variable that is positively related with the discussion of capacity discipline, and has a negative effect on offered seats. In other words, we want to verify that our data satisfies conditional exogeneity, i.e., given X, the "treatment" $\text{Capacity-Discipline}$ is uncorrelated with the error, because conditional exogeneity ensures unconfoundedness which is sufficient to identify the causal effect of communication on offered seats [Rosenbaum, 1984; Altonji and Matzkin, 2005].

47

Figure 10: Histogram of Placebo Token Overlap with Capacity Discipline

Notes. We define the level of overlap between Placebo$_{m,t}$ and Capacity-Discipline$_{m,t}$ as the percentage of Placebo$_{m,t} = 1$ observations where Capacity-Discipline$_{m,t} = 1$.

To test the conditional independence we conduct a test motivated by White and Chalak [2010], which relies on testing unconfoundedness, as rejecting unconfoundedness implies rejecting conditional exogeneity.[33] Suppose we can find another covariate $Z \in \{0,1\}$ that is positively related with Capacity-Discipline and is also a function of X. In other words suppose there is known function $\rho(\cdot)$ and an unobserved error ν such that $Z = \rho(\text{Capacity-Discipline}, X, \nu)$, where $\rho(\cdot)$ is a known structural equation with unobserved error ν. Then, White and Chalak [2010] show that *if* Capacity-Discipline $\perp \varepsilon | X$ *then* $\ln(seats) \perp Z | (\text{Capacity-Discipline}, X)$.

Therefore to implement this test we have to first determine a random variable Z. As we have seen in the placebo exercises, we have to be careful in choosing such a random variable that is positively related to Capacity-Discipline but has a negative effect on log-seats. To that end we first determine tokens or keywords that satisfy the following conditions: (i) the token has similar meaning as "capacity discipline", "gdp" and "demand"; and (ii) appears frequently with at least one of these three keywords in the earnings call reports. Then, similar to the placebo-tokens, for that token define a dummy variable Z_{mt} equal to one if all legacy carriers in market m use it in period t and include it as an additional regressor in (1). If the estimated coefficient for Z_{mt} is not

---

[33] See Rosenbaum [1987] and Heckman and Hotz [1989] for tests of uncounfoundedness.

Table 9: Estimates for Conditional Exogeneity

	Z's Coefficient	Capacity-Discipline
slow	-0.00514*	-0.01417***
	(0.00311)	(0.00246)
weakness	0.01520***	-0.01539***
	(0.00300)	(0.00241)
domestically	0.01914***	-0.01461***
	(0.00323)	(0.00241)
internationally	0.00525*	-0.01518***
	(0.00306)	(0.00242)
stable	0.00937*	-0.01551***
	(0.00524)	(0.00243)
pace	0.00264	-0.01525***
	(0.00426)	(0.00248)

Notes. Estimation results from including new tokens as additional regressors in (1). The table shows the coefficient estimates for each token and the corresponding estimate of Capacity-Discipline. Standard errors are in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

statistically different from zero, or positive then our estimates satisfy uncounfoundedness.

To find a token that satisfy the first criteria we choose tokens with cosine similarly measure, defined in (4), close to one for all three "capacity discipline", "gdp" and "demand." In order to satisfy the second criteria we restrict the token to be such that at least 50% of the time it appears in the same report as these three keywords. In Table 9, first column we present all the tokens that satisfy these two criteria. Similar to the placebo-tokens, for each token let $Z_{mt} = 1$ if all legacy carriers use that token in a given quarter and use it as an additional regressor in (4). The estimated coefficients are in the second column of the table, with their p-values in the third column. In the fourth column we present the estimates on Capacity-Discipline with their p-values in the last column. As can be seen, only one token, "slow," has slight negative effect on log seats but that is also statistically insignificant at 5%, while the rest are either insignificant or have positive effect. The cases that find a positive, non-zero relationship between Z_{mt} and capacity show that if anything, our results understate the true effect of the relationship between the discussion of capacity disci-

pline and capacity. What is also reassuring is the fact that the estimates for `Capacity-Discipline` are stable, negative and statistically significant, with effects very close to the primary regression, suggesting that our estimate is stable and, if anything, underestimates the role of communication.

6 Conclusion

In this paper we investigate whether legacy airlines use public communication to sustain collusion offering fewer seats in a market. We say airlines were communicating whenever all legacy carriers serving a market communicated about capacity discipline in their earnings calls. Using methods from NLP, we convert the text data into numeric data to measure communication among legacy carriers. We estimate that communication leads to a negative, and statistically significant, effect of -1.48% on seats offered, on average, across airlines and markets. Furthermore, this effect is entirely driven by the legacy carriers, and the reduction is substantially greater in smaller markets at -4.21%, and the size of the effect decreases with market size. Even for smaller city-pair markets with at most two airports, communication still causes -4.16% decrease in seats.

Our finding is relevant for the current policy debate about the correct response to increasing information about firms in social media and increasing market concentration across industries. Thus, in the airline industry, the SEC's transparency regulations are at odds with antitrust laws — a fact that policy makers must be cognizant of. While the value of public quarterly earnings calls remains debatable, the public disclosure of information through these calls is generally viewed as beneficial for investors. At the same time, the competitive effects of this increased transparency are theoretically ambiguous and under-studied. In this paper we attempt to address this lacuna in the literature, and we hope that this paper will spur more research in this direction.

While it is known that, in some cases, communication helps in equilibrium selection, its broader

50

implications for prices and welfare are unknown. Answers to these questions will help designing laws that are related to public communication and antitrust. That, however, requires estimating structural model of dynamic oligopoly with incomplete information and communication, which is left for future research.

References

ABREU, D., D. PEARCE, AND E. STACCHETTI (1986): "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, 39, 251–269. 7

ALTONJI, J. G., AND R. MATZKIN (2005): "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, 73, 1053–1102. 47

ARYAL, G., F. CILIBERTO, C. MURRY, AND J. WILLIAMS (2017): "Price Discrimination in International Airline Markets," *Mimeo*. 5

AWAYA, Y., AND V. KRISHNA (2016): "On Communication and Collusion," *American Economic Review*, 106(2), 285–315. 3, 7, 8, 23, 25, 40

——— (2017): "Communication and Cooperation in Repeated Games," *Mimeo*. 3, 7, 23, 25, 26, 40

AZAR, J., M. C. SCHMALZ, AND I. TECU (2017): "Anti-Competitive Effects of Common Ownership," *Mimeo*. 10

BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2016): "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593–1636. 8

BAMBERGER, G. E., D. W. CARLTON, AND L. R. NEUMANN (20004): "An Empirical Investigation of the Competitive Effects of Domestic Airline Alliances," *The Journal of Law and Economics*, 47(1), 195–222. 16

BARANY, I. (1992): "Fair Distribution Protocols or How the Players Replace Fortune," *Mathematics of Operations Research*, 17, 327–340. 8

BERNHEIM, B. D., AND M. D. WHINSTON (1990): "Multimarket Contact and Collusive Behavior," *RAND Journal of Economics*, 21, 1–26. 24

BERRY, S., M. CARNALL, AND P. T. SPILLER (2006): "Airline Hubs: Costs, Markups and the Implications of Consumer Heterogeneity," in *Advances in Airline Economics: Competition Policy and Antitrust*, ed. by D. Lee, vol. 1, pp. 183–214. Amsterdam: Elsevier. 5, 21

BERRY, S., AND P. JIA (2010): "Tracing the Woes: An Empirical Analysis of the Airline Industry," *American Economic Journal: Microeconomics*, 2, 1–43. 5, 16

BERRY, S. T. (1990): "Airport Presence as Product Differentiation," *American Economic Review*, 80(2), 394–399. 16

——— (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), 889–917. 16

BORENSTEIN, S. (1989): "Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry," *RAND Journal of Economics*, 20(3), 344–365. 16

——— (2004): "Rapid Price Communication and Coordination: The Airline Tariff Publishing Case," in *The Antitrust Revolution: Economics, Competition and Policy*, ed. by J. K. Jr., and L. White, pp. 223–51. Oxford University Press, 4 edn. 7

BORENSTEIN, S. (2010): "An Index of Inter-City Business Travel for Use in Domestic Airline Competition Analysis," *NBER Working Paper*. 22

BORENSTEIN, S., AND N. ROSE (1994): "Competition and Price Dispersion in the U.S. Airline Industry," *Journal of Political Economy*, 102(4), 653–683. 16

BRESNAHAN, T. F., AND P. C. REISS (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99(5), 977–1009. 34

BRUECKNER, J. K., D. LEE, AND E. SINGER (2014): "City-Pairs Versus Airport-Pairs: A Market-Definition Methodology for the Airline Industry," *Review of Industrial Organization*, 44(1), 1–25. 16

BRUECKNER, J. K., AND P. T. SPILLER (1994): "Economies of Traffic Density in the Deregulated Airline Industry," *Journal of Law and Economics*, 37(2), 379–415. 16

BYRNE, D., AND N. DE ROOS (2017): "Learning to coordinate: A study in retail gasoline," *Mimeo*. 2

CILIBERTO, F., AND E. TAMER (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77(6), 1791–1828. 16

CILIBERTO, F., AND J. W. WILLIAMS (2010): "Limited access to airport facilities and market power in the airline industry," *Journal of Law and Economics*, 53(3), 467–495. 16, 38

——— (2014): "Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conduct Parameters in the Airline Industry," *RAND Journal of Economics*, 45(4), 764–791. 5, 16, 22, 24

COMPETE, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66, 597–626. 8

COX, G. W. (1990): "Multicandidate Spatial Competition," in *Advances in the Spatial Theory of Voting*, ed. by J. M. Enelow, and M. J. Hinich, pp. 189–198. Cambridge University Press. 38

EATON, B. C., AND R. G. LIPSEY (1975): "The Principle of Minimum Differentiation Reconsidered:

Some New Developments in the Theory of Spatial Competition," *Review of Economic Studies*, 42, 27–49. 38

EVANS, W., AND I. N. KESSIDES (1994): "Living By The 'Golden Rule': Multimarket Contact in the U.S. Airline Industry," *Quarterly Journal of Economics*, 109, 341–366. 16, 24

FELDMAN, R., AND E. FRONDORF (2017): *Drug Wars: How Big Pharma Raises Prices and Keeps Generics off the Market*. Cambridge University Press. 2

FORGES, F. (1986): "An Approach to Communication Equilibria," *Econometrica*, 54, 1375–1385. 8

——— (1990): "Universal Mechanisms," *Econometrica*, 58, 1341–1364. 8

GENTZKOW, M., B. T. KELLY, AND M. TADDY (2017): "Text as Data," *Journal of Economic Literature, forthcoming*. 8

GENTZKOW, M., AND J. SHAPIRO (2014): "What Drives Media Slant? Evidence from U.S. Daily Newspapers," *Econometrica*, 62(1), 35–71. 8

GERARDI, D. (2004): "Unmediated Communication in Games with Complete and Incomplete Information," *Journal of Economic Theory*, 114, 104–131. 8

GERARDI, K. S., AND A. H. SHAPIRO (2009): "Does competition reduce price dispersion? New evidence from the airline industry," *Journal of Political Economy*, 117(1), 1–37. 16

GLUSAC, E. (2017): "In the Air That Uneasy Feeling of Us vs. Them," The New York Times. 3

GOLDBERG, Y., AND O. LEVY (2014): "word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method," *ArXiv e-prints*. 42

GOLTSMAN, M., AND G. PAVLOV (2014): "Communication in Cournot Oligopoly," *Journal of Economic Theory*, 153, 152–176. 8

GREEN, E. J., R. C. MARSHALL, AND L. M. MARX (2014): "Tacit Collusion in Oligopoly," in *The Oxford Handbook of International Antitrust Economics*, ed. by R. D. Blaier, and D. D. Sokol, chap. 19. Oxford University Press. 2

GREEN, E. J., AND R. H. PORTER (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52(1), 87–100. 3, 7

HALVORSEN, R., AND R. PALMQUIST (1980): "The Interpretation of Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 70(3), 474–475. 27

HARRINGTON, J. E. (2006): *How do Cartles Operate?*, Foundations and Trends in Microeconomics. Now Publishers Inc. 3

HARWELL, D., A. HALSEY III, AND T. MOORE (2015): "Justice Dept. Investigating Potential Airline Price Collusion," The Washington Post. 9

HASSAN, T. A., S. HOLLANDER, L. VAN LENT, AND A. TAHOUN (2017): "Firm-Level Political Risk," *Working Paper*. 8

HECKMAN, J., AND V. HOTZ (1989): "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," *Journal of American Statistical Association*, 84, 862–874. 48

HOBERG, G., AND V. MAKSIMOVIC (2015): "Redefining Financial Constraints: A Text-Based Analysis," *Review of Financial Studies*, 28, 1312–1352. 8

HOBERG, G., AND G. PHILIPS (2016): "Text-Based Network Industries and Endogenous Product Differentiation," *Journal of Political Economy*, 124(5), 1423–1465. 8

HOTELLING, H. (1929): "Stability in Competition," *Economic Journal*, 39, 41–57. 38

KANDORI, M., AND H. MATSUSHIMA (1998): "Private Observation, Communication and Collusion," *Econometrica*, 102(1), 189–228. 8

KAPLOW, L. (2013): *Competition Policy and Price Fixing*. Princeton University Press. 3

KIM, E. H., AND V. SINGAL (1993): "Mergers and Market Power: Evidence from the Airline Industry," *American Economic Review*, 83(3), 546–569. 16

KOIJEN, R. S. J., T. J. PHILIPSPN, AND H. UHLIG (2016): "Financial Health Economics," *Econometrica*, 84(1), 195–242. 8

LAMBERTINI, L., AND C. SCHULTZ (2003): "Price or Quantity in Tacit Collusion?," *Economics Letters*, 78(1), 131–137. 26

LEYDEN, B. T. (2018): "There's an App (Update) for That: Understanding Product Updating Under Digitization," *Working Paper*. 8

MAILATH, G. J., AND L. SAMUELSON (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press. 3

MARSHALL, R. C., AND L. M. MARX (2014): *The Economics of Collusion: Cartels and Bidding Rings*. MIT Press. 3, 7

MIKOLOV, T., K. CHEN, G. CORRADO, AND J. DEAN (2013): "Efficient Estimation of Word Representations in Vector Space," *ArXiv e-prints*. 6, 42

MILLER, A. R. (2010): "Did the Airline Tariff Publishing Case Reduce Collusion?," *The Journal of Law and Economics*, 53(3), 569–586. 7

MYERSON, R. B. (1997): *Game Theory: Analysis of Conflict*. Harvard University Press. 8

OECD (2011): "Information Exchanges Between Competitors under Competition Law," https://www.oecd.org/competition/cartels/48379006.pdf. 2

OSBORNE, M. J. (1993): "Candidate Positioning and Entry in a Political Competition," *Games and Economic Behavior*, 5, 133–151. 38

PORTER, R. H. (1983): "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886," *The Bell Journal of Economics*, 14(2), 301–314. 7

POSNER, E. A., F. M. S. MORTON, AND E. G. WEYL (forthcoming): "A Proposal to Limit the Anti-Competitive Power of Institutional Investors," *Antitrust Law Journal*. 10

ŘEHŮŘEK, R., AND P. SOJKA (2010): "Software Framework for Topic Modelling with Large Corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pp. 45–50. ELRA, http://is.muni.cz/publication/884893/en. 42

ROMANO, J. P., A. M. SHAIKH, AND M. WOLF (2010): "Hypothesis Testing in Econometrics," *Annual Review of Economics*, 2(1), 75–104. 46

ROSENBAUM, P. (1984): "From Association to Causation in Observational Studies: The Tole of Tests of Strongly Ignorable Treatment Assignment," *Journal of American Statistical Association*, 79, 41–48. 30, 47

——— (1987): "The Role of a Second Control Group in an Observational Study," *Statistical Science*, 2, 292–306. 48

ROSENFIELD, A. M., D. W. CARLTON, AND R. H. GERTNER (1997): "Communication Among Competitirs: Game Theory and Antitrust Application of Game Theory to Antritrust," *George Mason Law Review*, 5(3), 423–440. 3

SANNIKOV, Y., AND A. SKRZYPACZ (2007): "Impossibility of Collusion Under Imperfect Monitoring with Flexible Production," *American Economic Review*, 97(5), 1794–1823. 3

SHARKEY, J. (2012): "Expect Fewer Seats, Even for Overseas Flights," The New York Times. 3

SINGHAL, A. (2001): "Modern Information Retrieval: A Brief Overview," *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 24(4), 35–42. 43

SNIDER, C., AND J. W. WILLIAMS (2015): "Barriers to Entry in The Airline Industry: A Multidimensional Regression-Discontinuity Analysis of Air-21," *The Review of Economics and Statistics*, 97(5), 1002–1022. 38

SPECTOR, D. (2015): "Cheap Talk, Monitoring and Collusion," *Mimeo*. 8

STIGLER, G. (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 74, 44–61. 7, 25, 33

TETLOCK, P., M. SAAR-TSECHANKSY, AND S. MACSKASSY (2008): "More than Words: Quantifying Language to Measure Firms' Fundamentals," *Journal of Finance*, 63, 1437–1467. 8

VISCUSI, W. K., J. E. HARRINGTON, AND J. M. VERNON (eds.) (2005): *Economics of Regulation and Antitrust*. MIT Press, 4 edn. 7

WHITE, H., AND K. CHALAK (2010): "Testing a Conditional Form of Exogeneity," *Economic Letters*, 109, 88–90. 39, 48