What do Firms Say When Colluding? Evidence from Past Antitrust Cases

by

Chang (Margaret) Cheng

Department of Economics

University of Virginia

Supervised by

Professor Gaurab Aryal

Submitted to the Distinguished Majors Program

April 29, 2021

Contents

Abstract

This paper explores the use of Natural Language Processing methods to determine a dictionary of keywords and phrases –messages– that firms could use to initiate, or sustain a collusion. In particular, I propose to recent antitrust cases and the quarterly earnings calls of the defendant firms as my "training dataset" to build the dictionary of "collusive messages." To this end, I use four recent cartels that are under investigation for price-fixing, and I apply a combination of manual review and text analysis to parse the earnings call transcripts during the collusion period to build a dictionary that can help answer: What do firms say when they (want to) collude?

1 Introduction

Price-fixing behaviors take place a lot more frequently than people think, and early detection of such behaviors can stop and prevent huge welfare loss. Ever since the introduction of the Sherman Antitrust Act more than a century ago, the antitrust enforcement agencies in the United States have spent countless efforts trying to detect collusive behaviors. The topic has also attracted many economists to conduct theoretical and empirical research on related topics, and discuss the conditions under which communication amongst firms could yield extra profit. Harrington Jr. (2006) draws on past empirical evidence and suggests a list of "collusive markers" that could indicate the presence of collusive behaviors. Friedman (1971) models the role of communication in the case of perfect monitoring. Green and Porter (1984) introduces public monitoring and show that communication might not be feasible with the introduction of any market uncertainty. Awaya and Krishna (2016) and Spector (2020) argue that in case of market uncertainty and private monitoring, being able to communicate with each other can allow firms to differentiate between different causes of the market uncertainty and can help avoid unnecessary price wars against each other, thus improving the level of price and profitability.

While cartels have traditionally left people with the impression of smoke-filled rooms and under-the-table deals, there has been a major shift in paradigm in terms of the means by which firms communicate in recent decades. Modern cartels are motivated to adapt their way of communication and adopt more means to exchange information that are less likely to invite scrutiny from the antitrust enforcement agencies. Awaya and Krishna (2016), for example, find that even "cheap talks", unverified communication about past sales, can facilitate collusion given market uncertainty and private monitoring, and propose a set of equilibrium strategies that could facilitate collusion via public communications.

While communication is usually desirable given market uncertainty and imperfect monitoring, and the use of public communications has proved to be able to facilitate collusion, it is important to note that not all kinds of public communication are useful in sustaining collusion. Therefore, in order to make sure we learn about messages that indeed indicate collusive behavior, this paper looks at past cases where firms are accused of using public communication to fix prices. Using NLP techniques to study the quarterly earnings calls of these firms, I seek to answer the questions: What are the types of messages that firms say when they are colluding? Are there actually "cheap talks" carried out in practice as is proposed in Awaya and Krishna (2016)? Can we build a dictionary containing all the frequently used words that could indicate the presence of price-fixing behaviors?

It is extremely important that we answer these questions. With these questions answered, the antitrust enforcement agencies can be enabled to test and detect collusion at an earlier stage, increase the level of competition in the market, and reduce the welfare loss induced by price-fixing behaviors. To the best of my knowledge, however, there has, so far, not been any systematic study of the language that firm use when facilitating collusion. My main contribution in this paper, therefore, is to address the lack of knowledge in this specific field of research, and use NLP methods to study systematically the messages that firms send while facilitating collusion.

To this end, I conduct detailed case study on four past antitrust cases involving price-fixing behaviors (*U-Haul International* case, *Valassis Communications* case, *Delta/AirTran* case, and *Generic Drugs* case) where quarterly earnings calls are used as a means of facilitating price-fixing behavior. Specifically, I select the four cases by combining the information obtained in the Private International Cartels (PIC) data set, a comprehensive cartel database containing detailed information on contemporary price-fixing cartels from 1999 to 2019, and two papers, OECD (2012) and Harrington Jr. and Kashfipour (2020), which review recent cases investigated by the Federal Trade Commission (FTC) and the Antitrust Division of Department of Justice (DOJ) that involve public disclosure of information through quarterly earnings calls. For each of the cartels I study, I first collect and manually review the transcripts of all participating firms in the cartel during the collusion period to obtain a preliminary dictionary of related keyword, and then perform text analysis on the preliminary dictionary to obtain an expanded dictionary of keywords that has similar contextual meanings as the preliminary dictionary using quarterly earnings calls as the "training data".

Analyzing the four cartels indicates the following key insights. First, I find that all four cartels take place in concentrated markets dominated by only a few major players. This finding echoes the previous theory proposed in Markham (1951), where the author identifies as one key characteristics of a cartel the fact that "firms must be few in number and each firm must be sufficiently large." Furthermore, I find that the firm initiating the conversation tends to be the market leader (or claims to be so). An analysis of the messages indicates that firms in all four cases demonstrated anticipation of their competitor's reactions. Additionally, many key words and phrases, such as "cost", appear in more than one dictionary, and are frequently cited in multiple cartels as reasons to increase prices. Finally, since all four cartels focus on price-fixing rather than capacity restriction, I remove industry-specific words and phrases and combine the expanded dictionary obtained to form a larger generic dictionary containing words and phrases that could indicate price-fixing behaviors.

The rest of the paper is organized as follows. Section 2 offers an overview of the previous theories on the role of communications in facilitating collusive behaviors. Section 3 introduces the sources of data that are used and the selection process for cartels for further analysis. Section 4 describes the background and incentives of collusion, analyzes the details of the communication, and proposes a dictionary of related keywords for each of the four case studies. Section 5 introduces the text analysis method, `word2vec`, and analyzes the expanded dictionary obtained from the text analysis. Section 6 concludes the paper, and points out the potential limitations that could constrain the applicability of the key findings of this paper.

2 Where do we stand

This section provides an overview of previous theories on the role of communications in facilitating collusion. The contents are divided in two categories: the first subsection describes the level of communication that is necessary and effective for firms to coordinate, and the second subsection provides some insight on the level of communication that is allowed under today's antitrust legal system in the US.

2.1 Communication in Game Theory

This subsection seeks to answer two main questions. When does communication provide incentives for participating firms to coordinate? What type of communication can achieve such impact?

The first problem has been discussed in the context of oligopoly by a series of scholars, and the effectiveness of communication between firms is based on a variety of different assumptions of the level of monitoring. The topic is first discussed in Friedman (1971). His model is based on the assumption of perfect monitoring, where all past actions of the participating firms are commonly observed without any noise. In such setting, Friedman (1971) shows that for any fixed discount factor, there is no role for communication.

In practice, however, firms may face uncertain demand where firms are unable to monitor each other perfectly. Such an environment with public monitoring was introduced by Green and Porter (1984). Here firms are able to observe a common noisy signal (the market clearing price) but nothing more (firms are unable to observe each others' sales data). Even in this case, Green and Porter (1984) show that for a fixed discount factor, communication does not play any role if firms' strategies only depend on the history of past prices.

The main difficulty with public monitoring, however, is that firms are unable to distinguish the source of a fluctuation in market price, as the price can change either because someone cheated or because of a negative aggregate demand shock. Hence, once either colluding firm observes a fluctuation in price and profits below a certain threshold, it lowers its price to punish its competitor, and firms are likely to end up in a price war.

It is precisely in such situations that communication may become a useful tool to coordinate. If firms are able to send some (payoff irrelevant) signals, then firms could coordinate their actions more effectively. This idea is illustrated in a Awaya and Krishna (2016). The authors adopt and adapt the private monitoring model in Stigler (1964), where different firms observe different noisy signals (such as their own sales), and argue that being able to communicate with each other can allow firms to differentiate between different causes of the fluctuations of signals and can help avoid unnecessary price wars against each other.[1] Similarly, Spector (2020) models a market where demand is uncertain and sales data become available with a delay. The author shows that early communication on sales volumes allows firms to quickly identify the cause of an unexpected market share swing and compensate each other for the losses through short phases of market share reallocation, thus reducing the chance of price wars. Since such communication improves the level of monitoring in the market, it helps the firms maintain higher prices and improve profits, and hence, reduces welfare.

While it is clear that communication amongst firms are usually desirable when

[1] This argument is based on the important assumption that the correlation between firms' (log) sales is high when they charge similar prices and low when they charge dissimilar prices.

the level of monitoring is imperfect, the question that follows is what level of communication would firms be willing to engage in to achieve such effects. There are several papers that analyze various means of communication that have been employed amongst firms to gain more information.

Harrington and Skrzypacz (2011) model a communication scheme seen empirically in many industries, which is based on an allocation of sales and enforced by a guaranteed "buy-back" system. Specifically, firms report their sales on a regular basis to the cartel leader, who compares the sales data with the previously agreed quota. Those firms that report sales above their quota are obligated to make a payment to those that undersell.

Many other authors focus on the role of a third party organization in facilitating the communication process. Rahman (2014) models a repeated Cournot oligopoly with mediated communication, and shows that a nonbinding mediator (such as a trade association) can assist the firms to collude by allowing firms to secretly monitor each other and by enabling the firms to coordinate infrequently to aggregate information better. Awaya and Krishna (2020) model a situation where firms' aggregate sales are made public by a third party, and show that even this level of communication can facilitate profitable collusion amongst firms. Real-life examples as such have been analyzed by a variety of scholars. Genesove and Mullin (2001) adopt a narrative approach to analyze the detailed notes on weekly meetings of the members of a sugar-refining cartel, and discusses the important role played by the trade association in facilitating the communication amongst firms. Doyle and Snyder (1999) analyze automobile producers' declaration of production forecasts through a leading industry trade journal, and discusses the complementary impact of the declaration: a high declared production level is associated with the rival's upward adjustment of production level.

While these are all examples of communications amongst firms to send across verifiable private information in order to facilitate the collusion, Awaya and Krishna (2016) and Awaya and Krishna (2019) find that even "cheap talks", unverified communication about past sales, can facilitate collusion. Based on the assumption of private monitoring (where all participating firms can only observe their own sales data), Awaya and Krishna (2016) propose a set of equilibrium strategies consisting of a grim trigger strategy based on the communication of the firm and a threshold sales-reporting strategy in a duopoly setting. Both firms start by setting monopoly prices, and the communication process includes a regular and simultaneous reporting of one of the two types: H (which means that their sales is above a commonly known threshold) or L (otherwise). Since the communication process is simultaneous, neither firm will be certain about the other firm's disclosure. The collusion can be sustained only if the two types reported by the two firms match (both H or both L); otherwise, the two firms will end collusion immediately and permanently as a punishment. Awaya and Krishna (2019) extends the above strategy so that it applies to a market with more than two firms. In both papers, the authors prove that if this specific set of strategies is carried out successfully, all firms will be able to achieve higher profits than without communication, even if there is no way for either firm to verify the truthfulness of the reporting. Harrington and Ye (2019) proposes a similar argument but with a different approach. The authors argue that when sellers engage in collusive practices through cheap talks (such as coordination on list prices and surcharges), the sellers' behavior can effectively influence the buyers' beliefs about the sellers' costs, resulting in supra-competitive prices.[2]

Prior theories have shown that communication is indeed desirable for firms given market uncertainty and imperfect monitoring, and that the "cheap talk" modeled by Awaya and Krishna (2016) and Awaya and Krishna (2019) represent the minimum requirement (in theory) of the level of communication needed to increase profits and thus is desirable for firms to participate in. However, few existing theories goes into detail on what specific type of communication those "cheap talks" belong

[2]This paper assumes that the sellers are in intermediate goods markets, where buyer-seller negotiation is the norm, and thus the price of each transaction is not publicly available. Coordination on list prices and surcharges are considered as cheap talks because discounts and non-surcharge components could alter the actual transaction price, which means that coordinated action does not directly constrain competition.

to. What topics do firms usually communicate about? What keywords do firms use as signals? These questions remain unanswered. This creates difficulty for researchers and antitrust enforcement agencies to determine empirically whether firms are colluding via communication.

2.2 Communication in the Current Antitrust Legal Framework

Having examined the level of communication that is conducive to facilitating collusion amongst firms, this subsection examines the level of communication that is legally allowed today in the United States under the current antitrust legal framework. Specifically, this section starts by examining the Sherman Antitrust Act, analyzes some of the landmark cases stemming from the statute, and seeks to draw some insight on the current level of regulation on the extent of communication.

The 1890 Sherman Antitrust Act Section 1 marks the first antitrust legal framework in the United States history to confine horizontal agreements amongst firms. Specifically, the statute broadly prohibits anti-competitive agreements, and formally declares illegal "[e]very contract, combination in the form of trust or otherwise, or conspiracy[] in restraint of trade or commerce."[3] While the congressmen who passed the statute used three different terms ("contract", "combination", and "conspiracy") to describe the actions that ought to be prohibited by the law, the three terms are generally used to represent a single concept – an agreement.[4] However, the Sherman Act itself did not put further emphasis on what specific kinds of agreements should be prohibited or allowed, leaving the criteria much to courts' interpretation.

Many of the Supreme Court decisions early on reflect the tendency to take a broad interpretation of the level of "agreement" that should be deemed unlawful under Section 1 of Sherman Act, and decisions in some of the landmark cases are still widely cited today. In *Interstate Circuit*, the decision of the Court states that

[3] 15 U.S.C. §1

[4] See Kaplow (2013)

"[i]t is elementary that an unlawful conspiracy may be and often is formed without simultaneous action or agreement on the part of the conspirators...Acceptance by competitors, without previous agreement, of an invitation to participate in a plan, the necessary consequence of which, if carried out, is restraint of interstate commerce, is sufficient to establish an unlawful conspiracy under the Sherman Act."[5] In *American Tobacco*, the Court deems that the means used to communicate is largely irrelevant when it comes to determining whether an action constitutes unlawful conspiracy. In *Container*, the Court deems the implicit understanding to constitute a concerted action that could result in a violation of Sherman Act.[6]

The more recent trend, however, has been a gradual relaxation during the enforcement process of Section 1. Court decisions in recent decades reflect an attempt to differentiate firms' interdependent behaviors from mere parallel business actions. In *Theatre Enterprises*, the Court decision states explicitly that "[t]his court has never held that proof of parallel business behavior conclusively establishes agreement or, phrased differently, that such behavior itself constitutes a Sherman Act offense."[7] Moreover, courts have started to place more requirements on the evidence that the plaintiffs must provide in order to instigate the investigation process. In *Matsushita*, the Court states that "[t]o survive a motion for summary judgment or for a directed verdict, a plaintiff seeking damages for a violation of Section 1 must present evidence 'that tends to exclude the possibility' that the alleged conspirators acted independently."[8] In *Twombly*, the Court focused on erecting a nontrivial hurdle on motions to dismiss that plaintiffs must overcome, motivated by the same idea of strengthening pleading requirements.[9]

Although many recent decisions set seemingly higher standards to convict firms engaging in collusion, there were much fewer debates on whether the nature of the "agreement" (for example, whether there were explicit or tacit information ex-

[5]See Interstate Circuit v. United States and 306 U.S. 208, 226-27 (1939)

[6]See Kaplow (2013)

[7]346 U.S. at 540-41

[8]See Matsushita Elec. Indus. Co., Ltd. v. Zenith Radio Corp. and Monsanto Co. v. Spray-Rite Service Corp.

[9]See Kaplow (2013)

changes) constitutes a Section 1 violation. This lack of clarity in the Supreme Court decisions gave rise to a great deal of uncertainty in the lower courts when dealing with cases where the existence of an "agreement" is in dispute.[10] The uncertainty in the antitrust legal framework, in combination with the ability to sustain collusion via public communications, gives firms that could be able to obtain benefits from communication more incentives to rely more on indirect and subtle means of communication (such as unilateral conducts via public communications) to achieve coordination.

The potential for unilateral conducts to exert anti-competitive harm is broadly recognized by antitrust enforcement agencies in the United States. An FTC study published in 1985 found that even unilateral price signaling by companies without reaching any agreements can increase prices in the affected market. Under the current U.S. antitrust legal framework, unilateral conduct (such as a unilateral disclosure of information) alone does not violate Section 1 of the Sherman Act due to the lack of an "agreement". However, a unilateral disclosure of information may violate Section 5 of the Federal Trade Commission Act ("FTC Act"), which prohibits "unfair methods of competition", or Section 2 of the Sherman Act, which prohibits efforts to "monopolize, or attempts to monopolize," including acts to "combine or conspire" with another person to monopolize.[11][12] When enforcing these laws, the U.S. antitrust agencies adopt five general criteria to assess the legality of unilateral information disclosures, including the nature and quantity of information disclosed, the specificity and context of the information disclosed, the setting of the disclosure (private or public), the nature of the industry and the market involved, and whether there are pro-competitive business justifications for the disclosure of information.[13]

Having provided the economic and legal background for the role of communication in facilitating collusion, the next few sections describe the procedures and

[10] *Id.*

[11] Although violations of the Sherman Act are also deemed to be violations of Section 5 of the FTC Act, 15 U.S.C. § 45, the Supreme Court has held that Section 5 of the FTC Act also applies to some conduct that does not violate the Sherman Act.

[12] 15 U.S.C. § 2.

[13] See OECD (2012) for a more detailed review

results of four case studies, where I review quarterly earnings calls of firms that were under investigation for participating in price-fixing, and attempt to build a dictionary containing key words and phrases that could indicate such behaviors.

3 Data

There are three primary sources of data for this paper: a comprehensive cartel database that provides detailed information on each cartel and cartel members, as well as two papers that review a number of recent cases that involves public disclosure of information investigated by the FTC and the DOJ. This section first introduces the three sources of data in order, and then goes on to describe the selection process for a list of cartels and firms for further analysis.

The data set I use as a starting point of this paper is the Private International Cartels (PIC) Data Set. This data set is especially valuable in that it comprises the largest collection of legal-economic information on contemporary price-fixing cartels from 1990 to 2019, involving 1,528 international cartels that are either convicted or investigated and 12,852 firms that participated in collusive behavior.[14] There are two major components to the data set: the first set of data contains detailed information on each cartel, and the second set of data describes each of the member firm in each cartel. Combining the two portions of the data set, there are a total of more than 400 variables.[15]

After cleaning the data set and dropping all the rows with incomplete information about the cartel, a total of 1,303 cartels and 11,513 firms remain. Table 1 displays the ten regions with the largest number of cartels in the data set. Specifically, there are 194 (14.8%) international cartels involving 2,155 firms (18.72%), and 140 (10.74%) US cartels involving 1,076 firms (9.35%).

This paper focus primarily on cartels under investigation in the US mainly due to two reasons. First, there is an abundance of data for the US market, and it is rel-

[14]See Connor (2014)

[15]See a description of important variables is displayed in Appendix A.

Table 1: Top 10 Regions by Number of Cartels

Region	Number of Cartels
Global	194
United States	140
European Union	82
Spain	67
Italy	66
South Korea	61
Germany	59
South Africa	48
Brazil	47
France	45

atively easy to obtain and interpret the data for firms that were part of a previously convicted or investigated cartel. As long as a firm is publicly traded, all its necessary information, especially the earnings conference call transcripts, which I treat as the primary form of public communication within a cartel, are readily available in the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system provided by the US Security and Exchange Commission (SEC), and are easily accessible via Lexis Nexis Academic. Other countries and regions, on the other hand, might not offer access to all relevant data in a format as standardized as in the US. Second, international cartels are subject to drastically different jurisdictions and market landscapes in different regional markets, which might lead to different collusive strategies and behaviors of different subsidiaries during the same time frame, whereas members of the same cartel in the US market face a more homogeneous market landscape, making it easier to interpret and disentangle the potential collusive behaviors.

Looking at the US market only, the median duration of a cartel is 60 months, but there exist some extreme outliers, some cartels with abnormally long (which could mean error in data set) or short (which could indicate the cartel's lack of actual impact on the domestic market) durations. To standardize the data set, I get rid of all the cartels with a duration of longer than 2,000 months or with a duration of shorter than 12 months. The remaining data set contains 113 cartels involving 1,010 firms, which means an average number of 8.94 firms in each cartel. The average duration of the remaining cartel is 76.96 months, and the median duration remains

Table 2: Top 10 Industries by Number of Cartels

Industry	Number of Cartels	Number of Firms Involved
Pharmaceuticals, medicines, medical devices	27	128
Finance, insurance, banking	15	213
Stone, clay, graphite, glass products	7	57
Inorganic chemicals and fertilizers	5	45
Construction	5	29
Fabricated metal products	4	61
Food and beverage manufacturing	4	36
Machinery, including electrical and parts	4	30
Transport services	4	20
Organic chemicals, other than pharmaceuticals	3	53

60 months.

The data set also categorizes all cartels into 30 industries. Table 2 displays the number of cartels convicted or investigated in the US in each industry. The table shows that the industry with the largest number of cartels are pharmaceuticals with 27 cartels involving 128 firms. Since pharmaceutical companies usually requires large overhead to invest in facilities, laboratories, and various other R&D efforts, most players competing profitably in the market are multinational private-held firms, and the market is made up of a few large companies competing head to head against each other in multiple sub-sectors. In summary, the antitrust cases in this industry are mostly brought against multi-year cartels involving pharmaceutical giants.[16]

The finance industry presents a similar picture, with 15 cartels involving 213 firms documented in the data set. Although the industry has been subject to relatively tight regulations and has been under constant scrutiny of the public and regulators, the major players in the industry remained largely unchanged in the past century, allowing the major banks and payment service firms to build up the alliance amongst themselves against new entrants. The settlement of *In re Payment Card Interchange Fee and Merchant Discount Antitrust Litigation* in 2019, with a cash value of roughly $5.5 billion, marked the end of a decade-long litigation against Visa Inc. and Mastercard, along with most of the major Wall Street Banks, for

[16] An example would be the *Generic Drugs* case discussed in Section 4.

overcharging credit card interchange fee that was first brought up in 2005.[17]

The second source of data this paper uses is OECD (2012), a report submitted by the United States to the Organization for Economic Co-operation and Development (OECD) Directorate for Financial and Enterprise Affairs Competition Committee. The report reviews a total of nine past antitrust cases investigated by the FTC and the DOJ that involve unilateral disclosure of information that appeared to be invitations to collude. Additionally, Harrington Jr. and Kashfipour (2020) also serves as a source of data that complements the OECD report. The authors identify three classes of public announcements which facilitate coordination among competitors to restrict competition, investigate nine cases involving collusion, and assess the level of enforcement of the conduct of competition authorities and courts in these cases.

Table 3 combines the information about the cartel and the participating firms in OECD (2012), Harrington Jr. and Kashfipour (2020), as well as the PIC Data Set. In terms of industry, the 16 cases in the two documents combined are heavily concentrated across three major industries: five (31.25%) of the cases in Manufacturing, five (31.25%) in Transport Services, and two (12.5%) of the cases in Wholesale and Retail Distribution. In terms of the time frame of the communication process amongst firms, the cases are fairly evenly distributed across the past three decades, with two (22.2%) cases in the 2000s, three cases in the 1990s, and another three cases in the 1980s.[18]

Since the primary objective of this paper is to analyze the communication process amongst firms that involves public disclosure of information, I focus primarily on cases where earnings call is utilized by firms as a means of information disclosure for further analysis. In the next few sections, I conduct case studies on four different cartels in completely different industries using a combination of manual review and text analysis, and analyze the results.

[17] See In re Payment Card Interchange Fee & Merch. Disc. Antitrust Litig., 986 F. Supp. 2d 207 (E.D.N.Y. 2013)

[18] The remaining case is the AE Clevite case, whose detailed information is currently unavailable.

Table 3: Summary of Recent Cases involving Collusion via Public Disclosure of Information

Firms Involved	Industry	Start Year [b]	End Year [b]	Use Earnings Call?
U-Haul International, Avis Budget Group	Transport Services	2007	2008	Yes
Valassis Communications, News America	Paper, Printing, Publishing	2004	2004	Yes
Stone Container Corp.	Manufacturing	1993	1993	Yes
Precision Moulding Co., Inc.	Manufacturing	1995	1995	No
AE Clevite, Inc.	Manufacturing	N/A[c]	N/A[c]	No
YKK(U.S.A) Inc.	Manufacturing	1988	1988	No
Quality Trailer Products Corp	Manufacturing	1990	1990	No
American Airlines, Inc.	Transport Services	1982	1982	No
Airline Tariff Publishing Co.	Transport Services	1988	1990	No
KPN, T-Mobile, Vodafone	Communication Services	2009	2009	No
Tyson, Pilgrim's	Wholesale, Retail Distribution	2008	2019	Yes
Tyson Foods, JBS USA, Smithfield Foods, Hormel Foods, etc.	Wholesale, Retail Distribution	2009	N/A [c]	Yes
Lannett Company, etc.	Pharmaceuticals, Medicines, Medical Devices	2009	2016	Yes
ArcelorMittal, Nucor, U.S. Steel, etc.	Primary Metals and Alloys	2005	2007	No
AirTran, Delta	Transport Services	2007	2008	Yes
AirTran, Delta, etc.	Transport Services	2002	2016	Yes

[a] Only 16 cases are listed in this table because the *U-Haul International* case and the *Valassis Communications* case appear in both OECD (2012) and Harrington Jr. and Kashfipour (2020).
[b] The Start Year and End Year columns in the table describes the start and end years of the communication amongst firms.
[c] Detailed Information of this case is currently unavailable.

4 Case Studies

This section provides a detailed review of four past antitrust cases in four different industries. The first two of the cases (*U-Haul International* and *Valassis Communications*) concern the unilateral actions taken by one firm within each investigated cartel that appear to be an "invitation to collude"; the third case (*Delta/AirTran*) and the fourth case (*Generic Drugs*) concern the public disclosure portion of the investigated cartels that appear to reveal a "dialogue" between the colluding firms. Within each case, I describe the background and incentives of collusion, analyze the details of the communication, and propose a dictionary of related keywords.

4.1 U-Haul International Inc. and AMERCO

The *U-Haul International* case takes place in the truck rental industry. At the time, the industry was dominated by only three major US companies: U-Haul International Inc. ("U-Haul"), Avis Budget Group, Inc. ("Budget"), and Penske Truck Leasing Co. ("Penske"), ranked in the order of size of the firms.

The key players involved in the case are U-Haul and its closest competitor, Budget. While U-Haul was at the time the largest player in the industry in the United States in terms of truck numbers, rental locations, revenues, as well as market share, its profitability and pricing power was limited by the presence of Budget. While the two firms are each other's biggest competitors, U-Haul and Budget combined account for 70 percent of one-way truck rental transactions in the United States, and, if acting in coordination, the two firms could profitably impose higher prices upon consumers.[19] This creates incentives for U-Haul to initiate the conversation about price increase and invite its competitor to engage in collusion.

The specific communications process among U-Haul and Budget investigated by the FTC includes private communications from year 2006 to year 2007 and public disclosures of information in late 2007 and throughout the year of 2008, both

[19] See Liu v. Amerco, 677 F.3d 489, 493 (1st Cir. 2012). and Complaint at 3, U-Haul Int'l, Inc., Docket No. C-4294, 2010 F.T.C. LEXIS 61 (2010).

initiated by the executives of U-Haul. (This paper focuses primarily on the public communication portion.) In order to better understand the primary incentives and goal of U-Haul's disclosures of information in its earning calls, I manually reviewed all the eight transcripts of the earnings calls for each of the two firms that are held in years 2007 and 2008.

The preliminary review of the earnings calls of AMERCO (parent company of U-Haul) indicates that U-Haul has complained in its earnings call about the low pricing in the industry long before the start date of the communication investigated by the FTC. For instance, when asked about the possibility of pricing stabilization in the future during its third quarter fiscal year 2007 earnings conference call on February 8, 2007, U-Haul indicated that,

> "It sends the wrong message to a customer when you give them something below the true cost of the product, and that is what is going on right now. So some people tell me, well, the customer is benefiting. I don't believe the customer is benefiting, I believe the customer is getting confused and confused customers don't like the industry as a whole, is my experience. So we don't want this to continue one day."[20]

In this earnings call, U-Haul directly attached the pricing scheme at the time by describing the current pricing as "below the true cost of the product" and arguing that the consumers are not benefiting from the low price. The firm also expressed dissatisfaction of the pricing scheme by stating directly its desire to put the situation to an end.

Apart from that, U-Haul, for multiple times, indicated publicly its intention to demonstrate price leadership in the market and its intention to raise prices both before and during the investigated time frame. Some examples include:

> "The way we would like to function in this marketplace is that we act as the price leader, and if we're at 3% or something above [B]udget, then

[20] Q3 2007 AMERCO Earnings Conference Call - Final (February 8, 2007)

so be it."[21]

> "...We are very, very much trying to function as a price leader and not give away share and those are kind of contradictory strategies. So what that means is in...a market where I don't see a lot of competition I'm trying to exhibit some price leadership. And even in several corridor markets that are highly competitive I'm trying to exhibit some price leadership..."[22]

Sometimes, however, U-Haul went well beyond its intention to exert price leadership, to directly inviting its competitor to collude, often citing the rising cost as a reason.

> "We're attempting to, in certain areas, raise prices to a rate that supports the cost of the truck and the return. We would hope that they would follow. We have tried this in specific areas, and if they don't follow, we will come back down until they do."[23]

> "...I worked really hard starting about in January for – to tell my fuel people float the price up and give [Budget] at least a month to match the price, just in case they don't get it. In other words – And so I feel [forced thought that] they gave away a little bit of transactions by doing that."[24]

It is clear that U-Haul's intent to increase prices demonstrate price leadership emerged even as early as late 2006 and went all the way through 2008. More notably, the complaint filed by the FTC indicates that the Chief Executive Officer and Chairman of U-Haul was aware that Budget representatives would monitor its third quarterly earnings call in financial year 2008.[25] Hence, his repeated mentioning of

[21] Q4 2007 AMERCO Earnings Conference Call - Final (June 7, 2007)

[22] Q3 2008 AMERCO Earnings Conference Call - Final (February 7, 2008)

[23] Q2 2007 AMERCO Earnings Conference Call - Final (November 9, 2006)

[24] Q4 2008 AMERCO Earnings Conference Call - Final (June 5, 2008)

[25] Complaint at 3, U-Haul Int'l, Inc.,Docket No. C-4294, 2010 F.T.C. LEXIS 61 (2010). Available at: https://www.ftc.gov/sites/default/files/documents/cases/2010/07/100720uhaulcmpt.pdf

price leadership and intention to increase the prices in that meeting could be an example of U-Haul's continued attempt to signal its competitors.

While a preliminary analysis of the U-Haul earnings calls indicates clearly an intent to coordinate price increase, the Budget earnings call, on the other hand, is much less informative. Since Budget is only a subsidiary of the greater Avis Budget Group, Inc., the structure of the report is much more fixed, and the language seldom changes from quarter to quarter. Specifically, when it comes to reporting the performance of the truck rental industry, Budget tend to attribute the declining revenue and prices to macroeconomic trends, such as the soft demand across all rental segments as well as the increasingly high fuel prices. Budget also frequently link the perceived declining demand to the decline in housing sales, a trend prevalent at the time. An example of a most typical statement from Budget would be as follows:

> "Revenue declined due to a 9% decline in rental days and a 4% decline in time and mileage revenue per day. The rental day drop was driven by reduced demand across all rental segments, as well as our fleet being 9% smaller than in second quarter 2006. The decline in T&M revenue per day reflected a decrease in one-way rental rates, which we believe is consistent with market trends. We believe the volume decline reflects softness in consumer demand, in-line with the decline in housing sales and not helped by historically high fuel prices."[26]

Having analyzed and reviewed the earnings call transcripts from both U-Haul and Budget, it is clear that it is U-Haul's primary intention to exert price leadership and increase price. Some of the most commonly used keywords include {`price, price leadership`}. Meanwhile, on Budget's end, the intention is not always clear, but the executives of the company usually complains about the low pricing of the truck rental industry and attribute it to soft demand. Hence, keywords such as {`truck rental, demand`} could be of interest.

[26] Q2 2007 Avis Budget Group, Inc. Earnings Conference Call - Final (August 8, 2007). The language in other earnings calls in the period follow roughly the same structure.

4.2 Valassis Communications Inc.

Valassis Communications Inc. ("Valassis") and News America Marketing ("News America") are the only two US producers and distributors of cooperative free-standing inserts ("FSI's").[27] Following a sustained price war between the two firms since 1998, when the two firms each own approximately half of the market, Valassis attempted to initiate a price increase of 5% in 2001, expecting News America to follow suit. News America, however, did not follow the price increase and was able to obtain a lead in market share by capturing additional customers from Valassis. In February 2002, Valassis abandoned the new pricing scheme and the two firms returned to competitive pricing, causing the price to fall by nearly 20% by 2004. Regaining market share and easing the fierce competition within the industry have since become top strategic objectives of Valassis executives. In mid-2004, Valassis determined that its aggressive pursuit of greater market share was no longer serving the company's interests. Company executives developed a new strategy that aims at halting contest between the two firms and raising FSI prices, and involves coordination among the two firms to cease challenging for each others' customers.[28]

The communication process among Valassis and News America investigated by the FTC includes public disclosures of information via quarterly earnings call throughout the year of 2004, initiated by Valassis. In order to better understand the specific incentives of Valassis's action, I manually reviewed the transcripts of the four earnings calls held in the year of 2004.

A preliminary review of the Valassis's earnings call reveals the firm's dissatisfaction of the pricing scheme prior to proposing the new strategy in mid-2004. For example, during its first quarterly earnings call in fiscal year 2004, Valassis indicated that the low FSI pricing has had a negative impact on its own revenue. In addition, Valassis also accused News America's recent action of having a "backdoor price increase" to offset lost revenue and profit as a result of lower FSI pricing. In essence,

[27] An FSI is a multi-page booklet containing discount coupons for the products of various firms that is inserted into newspapers for distribution to consumers.

[28] Complaint at 3, Valassis Communications, Inc.,Docket No.C-4160, 2006. Available at: https://www.ftc.gov/sites/default/files/documents/cases/2006/04/0510008c4160valassiscomplaint.pdf

the two statements combined indicate Valassis's belief that the pricing scheme at the time has had a negative impact on both firms' revenues in the industry, and this could create Valassis's incentives to invite News America to join its new pricing scheme.

Furthermore, Valassis also revealed its belief that there is room for a price increase during the same earnings call. When asked whether News America's action to introduce a "backdoor price increase" gives Valassis a chance to also raise its own price, the executive from the company replied with a clear "Yes," and stated that:

> "I think clearly there are - you know, clients view there to be significant differences between the Valassis market list and the News America market list right now. We always felt as if there's differences between us and Valassis that allow us to get a premium. Clearly, this market list is viewed by many clients as a backdoor price increase, which basically allows us to attempt to get a larger price differential than what we would have been able to get in the past. And obviously it's our goal to get as much in price as we can possibly get depending on market conditions. Obviously, as you're well aware, the current pricing environment has been perpetuated by News America."[29]

Valassis officially introduced its new pricing scheme during its second quarterly earnings call in fiscal year 2004, when it clearly stated its strategic objectives: maintain current market share, demonstrate price leadership, and raise floor price to the place in 2001 (its previous attempt to raise prices). Specific examples include:

> "In essence we have been able to achieve our page volume objectives with less than a 50% market share due to industry strength...We believe we can achieve our 2005 target for pages produced with no further shifts in co-op FSI market share."[30]

> "...we believe that now is the time to create a low risk opportunity to

[29] Q1 2004 Valassis Communications Inc. Earnings Conference Call - Final (April 22, 2004)
[30] Q2 2004 Valassis Communications Inc. Earnings Conference Call - Final (July 22, 2004)

> change the long term pricing trends in the co-op FSI industry. Therefore, effective Monday, July 26th, we will quote all newswriter refusal customers at the floor price which was in effect in May of 2001...The reason I said this is a low risk opportunity is that we will defend our customers and market share and use whatever pricing is necessary to protect our share."[31]

Most importantly, Valassis's public disclosure in the earnings call reveals that it has been closely monitoring News America's actions, knowing that News America would do the same. For instance, the top executives of Valassis reveal their own monitoring of News America's behavior by stating that "...we're going to monitor that situation on a daily basis." during its second quarterly earnings call in fiscal year 2004. The executives also openly discussed there expectations about News America's reaction to their new strategy:

> "In the recent past News America has been quick to make their intentions known. We don't expect the need to read the tea leaves. We expect that concrete evidence of News America's attention or intentions will be available in the marketplace in short order. If News continues to pursue our customers and market share then we will go back to our previous strategy."[32]

> "...we think we're at a point where we believe both FSI companies can achieve significant volume with their current market share positions. Generally, this type of supply demand equation typically leads to increased pricing power. I think logic would suggest that this condition provides an opportunity to create a positive long-term pricing trend. And so clearly that pricing trend could lead to increased profitability... If it doesn't work we will continue to look for creative ideas to improve

[31] *Id.*
[32] *Id.*

> our– you know ideas and opportunities to improve our profitability. And we'll try to do so sooner as opposed to later."[33]

While Valassis made its intentions clear in the earnings calls reviewed above, a preliminary review of News America's earnings call reveals much less information. This could be because of the fact that News America is a subsidiary of News Corporation Limited and the parent company had much more to focus on (such as the acquisition of DirecTV Group and the upcoming acquisition of Fox Entertainment Group in 2005) than the changes in the FSI industry (which is not their main area of business).[34] Interestingly, however, the reaction of News America to Valassis's new pricing scheme was revealed during Valassis's subsequent quarterly earnings call, when Valassis claimed that News America was no longer challenging for Valassis first right of refusal customers.[35]

> "One example and this would be...a large symbolic Valassis client in the northeast who is about to extend their first right of refusal agreement with us. The client shared this fact and details of our proposal with the News America rep who informed the customer that...News America could offer lower pricing. If they were to move that business from Valassis to News America. And the News America rep requested that the client not sign the Valassis extension and give them time and an opportunity to put together a proposal. The News America rep called the client back two days later and said that the Valassis deal that was being offered was a very good one and that he should move forward and take advantage of the Valassis proposal."[36]

Another example that Valassis gave was about a News America first right of refusal customer, whose contract was ultimately won by News America, but at what is believed to be a significantly higher price than before:

---

[33] *Id.*

[34] News Corporation Limited is now News Corporation.

[35] These are customers that basically give Valassis the first opportunity to place 100% of their business.

[36] Q3 2004 The News Corporation Limited Earnings Conference Call - Final (October 21, 2004)

> "One of those four [News America first right of refusal] clients continues to talk to us but the bottom line is that they are not happy with our new floor pricing which we will not go below. We also participated in an internet bid for a News America first right of refusal customer...News America won the business with pricing that was a few cents below our floor pricing. We believe that that price is significantly higher than what that customer paid in 2004."[37]

Based on the analysis of the earnings call of Valassis and News America, Valassis's objective to increase price and profitability, and some of the common words and phrases used by Valassis include: {`price, competition, profitability, difference`}.

4.3 Delta/AirTran Baggage Fee Antitrust Litigation

The *Delta/AirTran* case concerns the introduction of a fee on the first checked baggage at Hartsfield-Jackson Atlanta International Airport, where the two defendants in this case, Delta Air Lines, Inc. ("Delta") and AirTran Airways, Inc. ("AirTran) combined controlled 92% of the dominated route markets which had Atlanta as the origin or destination.[38] Of the two firms, Delta is the larger firm both in terms of routes and revenues, and is regarded by AirTran as its top competitor, but the strong presence of AirTran in the Atlanta market also posed competitive pressure on Delta, limiting its pricing power in the market. Therefore, even though most other airlines had already imposed fees for the first checked bag in addition to charging for additional checked bags by mid-2008, neither Delta nor AirTran planned to initiate the action for fear of the decreasing profitability in case the competitor does not follow suit. Both firms, however, tacitly understood that it would most likely follow suit if the competitor initiated the introduction of the first bag fee.[39]

[37] *Id.*

[38] See Harrington Jr. and Kashfipour (2020)

[39] See In re Delta/Airtran Baggage Fee Antitrust Litig., 245 F. Supp. 3d 1343 (N.D. Ga. 2017) for a detailed review of the background of the case.

The specific communications process among Delta and AirTran investigated by the DOJ includes Delta's internal documents assessing the risk of introducing a first bag fee in the market and the public disclosures of information by both firms in the year of 2008, including multiple earnings calls, as well as a number of press releases. (In this paper I am only interested in the latter portion.) In order to better understand the primary incentives and goal of both firms' actions, I reviewed the eight transcripts of the earnings calls for each of the two firms that are held in 2008, as well as the subsequent press releases held in the same year.

The preliminary review of the earnings calls of the two firms demonstrated a clear line of information exchanges on the introduction of a first bag fee.[40] The topic was initiated by Delta during its second quarterly earnings call in fiscal year 2008. When asked whether Delta will attempt to introduce a first bag fee following its merger with Northwest, Delta claimed that:

> "We are, we will study it. We will continue to study it but we have no plans to implement it at this point."[41]

Even though Delta did not clearly state the aim of introducing the fee immediately, this could be seen as a signal that Delta sent to AirTran regarding the introduction of the fee, as top executives of AirTran commented on the same topic in the subsequent quarter. When asked about the future plans of first bag fee, the top executive of AirTran clearly indicated his interest in the idea as well as its concerns about profitability in the case that Delta would not introduce the fee in the market.

> "We have the programming in place to initiate a first bag fee. And at this point, we have elected not to do it, primarily because our largest competitor in Atlanta where we have 60% of our flights hasn't done it. And I think, we don't think we want to be in a position to be out there alone with a competitor who we compete on, has two-thirds of our nonstop flights and probably 80 to 90% of our revenue is not doing the

[40] Since AirTran Airways is a subsidiary of AirTran Holdings, Inc., we analyze the earnings call transcripts of the parent company instead.

[41] Q2 2008 Delta Air Lines, Inc. Earnings Conference Call - Final (July 16, 2008)

> same thing. So I'm not saying we won't do it. But at this point, I think we prefer to be a follower in a situation rather than a leader right now."[42]

However, when asked whether it would consider introducing the fee if Delta initiated the action, AirTran replied that it would strongly consider doing so.[43]

Although Delta did not make any further comments regarding the matter in any of the subsequent quarterly earnings calls, it reacted to AirTran's above statement by immediately announcing the introduction of a first bag fee in its subsequent press release two weeks later, where Delta claimed that it is aligning the first bag policy (along with several other administrative fees) with Northwest post merger (while most other Northwest baggage policies and fees will be aligned to Delta's structure).

> "The increase in bags being carried on board Delta aircraft this year tells us that customers are not differentiating Delta as the only major airline not charging for a first checked bag...As we align customer policies and fees to simplify the travel experience for our customers throughout the merger, Delta is adopting proven practices from both Delta and Northwest that have been broadly accepted in the marketplace."[44]

> "Effective immediately, for travel on or after Dec. 5, customers flying within the United States will be charged $15 for the first checked bag and $25 for the second checked bag when traveling domestically, consistent with Northwest's existing policies."[45]

One week after Delta's press release, AirTran immediately followed suit and declared a first bag fee that goes in to effect on December.5th, 2008, the same day when Delta's first bag fee becomes effective.[46]

[42] Q3 2008 AirTran Holdings, Inc. Earnings Conference Call - Final (October 23, 2008)
[43] *Id.*
[44] GlobeNewswire. (2008, November 5). Delta Aligns Policies and Fees to Offer Consistency for Customers Traveling On Delta - and Northwest-Operated Flights; Fee structure reflects proven practices from both airlines that have been broadly accepted in marketplace [Press release]
[45] *Id.*
[46] Q4 2008 AirTran Holdings, Inc. Earnings Conference Call - Final (January 28, 2009)

Analyzing the earnings call and press releases by Delta and AirTran, the common words and phrases utilized to signal the competitor include: {`initiate, bag fee, competitor, follow`}.

4.4 Generic Drugs Pricing Antitrust Litigation

The *Generic Drugs* case is an ongoing investigation conducted by the DOJ since 2016.[47] Up to date, more than 26 corporate defendants and 10 individual defendants, seven of whom have already been charged, have been under investigation by a coalition of 51 states and territories for price-fixing, bid-rigging and market-allocation conspiracy regarding 80 topical generic drugs that account for billions of dollars of sales in the United States within the generic drugs industry.[48]

According to the FDA, a generic drug is "a medication created to be the same as an existing approved brand-name drug in dosage form, safety, strength, route of administration, quality, and performance characteristics."[49] Generic drugs usually have lower prices than their brand-name counterparts for two reasons. First, since generic drugs do not have to the repeat animal and clinical studies that were required of the brand-name medicines to demonstrate safety and effectiveness, their cost tend to be lower. Second, since the FDA usually approves multiple applications for generic drugs to market a single brand-name product, the existence of competition usually results in lower prices for generic drugs as compared to brand-name drugs. However, when drug companies conspire to raise the prices of generic drugs, the existence of generic drugs seems to have lose its meaning because consumers never reap the benefits that lower priced generics are supposed to provide.

The most recent complaint suggests that there had been extensive private communications amongst the firms via phone calls, text messages, emails, corporate conventions, and dinner parties from at least 2009 through 2016.[50] A review by

[47]The case was originally investigated by the state of Connecticut in 2014

[48]For more information, see https://portal.ct.gov/AG/Press-Releases/2021-Press-Releases/Court-Unseals-Latest-Generic-Drug-Complaint

[49]More information available at: https://www.fda.gov/drugs/generic-drugs/generic-drug-facts

[50]More information available at: https://portal.ct.gov/AG/Press-Releases/2021-Press-Releases/Court-Unseals-Latest-Generic-Drug-Complaint

Harrington Jr. and Kashfipour (2020), however, suggests that the public announcements made by one of the defendants, Lannett Company Inc. ("Lannett"), via quarterly earnings calls from 2013 to 2015 may have served to shore up an agreement made through private communications. In order to better understand the role of public communications in facilitating this collusion, I reviewed the transcripts of all the earnings calls for Lannett that were held between 2013 and 2015.

A preliminary review of Lannett's earnings calls indicates that the company reported an increase in prices in Levothyroxine, Digoxin, and Ursodiol during most of the quarters within the three years of interest, and that most of the company's increase in profit was a result of increased prices, rather than volume. The CEO claimed that the company is a price leader, and could aggressively lead a price increase whenever there is an opportunity to do so:

> "...we're very capable of raising prices and we tend to sometimes lead the market. We see opportunities to raise the price, we take it. We don't sit back and wait for someone else to do it. So you might say we're a little more aggressive in the pricing arena."[51]

> "We tend to be a price leader on price increasing and the credit goes to my sales vice president. He takes an aggressive stance towards raising prices. He understands one of his goals as objectives as a sales vice president is to increase profit margins for the Company and he's the first step in that process...With one or two exceptions, we've tended to lead in the way of price increases. We believe that these prices are important, we need to try raising them. Sometimes it doesn't stick and we have to go back and reduce our price and other times it does. I am finding a climate out there's changed dramatically and I see more price increases coming from our competitors than I've seen in the past. We're going to continue to lead. We have more price increases planned for this year within our budgets and hopefully our competitors follow suit. If they don't, that's

[51] Q2 2013 Lannett Company Inc. Earnings Conference Call - Final (February 7, 2013)

> their issue, but our plan is to raise prices on any product we think we can, or we haven't raised the price. We – you know our costs aren't going down. Someone has to pay for these things, unfortunately."[52]

In this above earnings call, the CEO of Lannett repeated the importance of raising the price, and the fact that he directly stated his expectation for the competitors to follow suit suggests that Lannett was conveying a plan for a coordinated price increase.[53]

Furthermore, Lannett often cites the increasing costs as a reason to raise prices, and commends those competitors who follow suit instead of grabs market share as a response to Lannett's increase in prices as "responsible" and "rational". Some examples include:

> "I'm always grateful to see responsible generic drug companies realize that our cost of doing business is going up as well. As everyone knows, the FDA has new requirements for stability work on generic drug products that are going to cost a lot of money, add the GDUFA fees on top of that. So, whenever people start acting responsible and raise prices as opposed to the typical spiral down of generic drug prices, I'm grateful because Lannett tends to be active in raising prices. We believe we have to sell our products for a price that we can make a profit, that profit has to cover all of the costs that we incurred to make the product as well as what we expect to incur for product development or enhancements to those products, so I'm grateful to see price increases."[54]

> "We're seeing more responsibility on the part of all of our competitors. I believe because all of us are facing the same costs... I would expect that all the companies are not going to behave like they have in the past. I suspect you're going to see more price increases in the generic

[52] Q4 2013 Lannett Company Inc. Earnings Conference Call - Final (September 10, 2013)
[53] This point is also illustrated in Harrington Jr. and Kashfipour (2020).
[54] Q4 2013 Lannett Company Inc. Earnings Conference Call - Final (September 10, 2013)

marketplace or certainly less price erosion in the marketplace because of that."[55]

"We were lucky that the authorized generic supplier was a rational competitor. And when they introduced their product, they introduced theirs at a higher profit – a higher-margin – excuse me, a higher WAC price, wholesale acquisition cost is the numbers I'm going to reflect here."[56]

"...since the companies we're looking at here are not irrational players, I don't see them just going out and trying to grab market share."[57]

Even though Lannett has reported continued increasing of prices in the three drugs, the increase in prices seemed to have very limited negative impact on the overall revenue and profit of the firm. The results of the price increase on the firm is clearly seen in the subsequent earnings calls following the price increase, and the CEO of Lannett even projected sustainability of the increased prices in the future.

"We do believe strongly that there's sustainability in some of the price increases that we are seeing, right now, that are in our third-quarter numbers."[58]

"If you're saying that the price increases that we've had in place, are they sustainable, and are they maintaining? My answer would be yes, they continue to hold up...We predict what our revenues will be for the year. We're not seeing any declines, generally speaking, on the price increase products. So they continue to, let's say, level off at their new pricing."[59]

" I think you're going to find more capital pricing – more – I'll say less competition, in a sense. You won't have price wars. You are still going

[55] *Id.*
[56] Q4 2014 Lannett Company Inc. Earnings Conference Call - Final (August 27, 2014)
[57] Q1 2015 Lannett Company Inc. Earnings Conference Call - Final (November 3, 2014)
[58] Q1 2014 Lannett Company Inc. Earnings Conference Call - Final (November 7, 2013)
[59] Q2 2015 Lannett Company Inc. Earnings Conference Call - Final (February 4, 2015)

> to have competition, because there's a lot of generic companies in the market. I just don't see the prices eroding like they did in the past. It's really unfortunate, but what they see some significant pricing – cost increases, I should say, that are driving this."[60]

Analysis of Lannett's earnings calls from 2013 to 2015 present the company's clear line of thought to raise the prices and increase profitability. The common words and phrases Lannett used that might amount to an intention to fix prices include: {`price increase`, `responsibility`, `profit margin`, `cost`, `competition`}.

5 Text Analysis

Having determined a preliminary list of words and phrases to focus on in each case above, I then apply text analysis in order to find all the tokens that have a similar contextual meaning with the given tokens and form a more comprehensive dictionary of relevant keywords in each case. This section proceeds in the following order: first, it provides an overview of the existing text analysis techniques; next, it introduces in depth `word2vec`, the text analysis technique applied in this paper; finally, this section discusses the application of `word2vec` on the previously mentioned cases, respectively.

5.1 Text Analysis Techniques Overview

With the development of new theories and technologies that allow for access and analysis of previously unavailable data, the mindset of using text as data is increasingly widely adopted in economic analyses. Gentzkow, B. Kelly, and Taddy (2019) introduce the methodology of the use of text as inputs to economic research, and argue that the most important feature that differentiate the use of text from regular data is that text is inherently high-dimensional, thus allowing for the use of

[60] *Id.*

statistical methods (such as machine learning) that are commonly used to analyze high-dimensional data in other domains.

Gentzkow, B. Kelly, and Taddy (2019) provide a nice summary of a three-step procedure that the application of text analysis techniques generally follow. Step one involves representing raw text D as a numerical array $\mathbf{C}$ using a bag-of-words approach, which is to encode the index for location for each token (such as words and phrases) into a large numerical array. The primary purpose of this step is to pre-process and encode the raw texts, and reduce them to a simpler representation that is more suitable for statistical analysis. The second step involves mapping $\mathbf{C}$ to predicted values $\hat{V}$ of unknown outcomes $\mathbf{V}$ and involves the application of data mining and machine learning techniques in order to generate predictions of a variable of interest. This step is of primary importance because researchers need to select the algorithm and technique that aligns best with the purpose of the research. The last step involves using $\hat{V}$ in subsequent descriptive or causal analysis.

Gentzkow, B. Kelly, and Taddy (2019) classify the common text analysis techniques into four main categories based on the purpose of the analysis. *Dictionary-based* methods is the most common method used by social scientists. In most cases, researchers simply specifies some function $f(\cdot)$ that maps the result of the previous step, $\mathbf{C}$, to generate the predictions, $\hat{V}$. For example, given a numerical array representing a set of pre-processed documents, c_i, a researcher can use a predefined dictionary that provides a mapping between words and sentiments (such as General Inquirer and the Loughran-McDonald Sentiment Word Lists), and obtain a prediction of the outcome of interest, $\hat{v}_i$, based on some form of aggregation of the count of the words expressing each sentiment.[61][62][63]

The second broad category of methods is the *text regression* methods. The method is applicable when the underlying causal relationship runs from language to

[61] More Information about General Inquirer available at: http://www.wjh.harvard.edu/ inquirer/

[62] The Loughran-McDonald Sentiment Word List is proposed in Loughran and McDonald (2011)

[63] See Tetlock (2007) and Scott R. Baker (2016) for examples of application

outcome (a model of $p(v_i|c_i)$).[64] Given some training data where both v_i and c_i is observable, researchers can regress the training value of v_i on the respective c_i. However, since the data is high-dimensional, ordinary regression methods (such as ordinary least squares) are infeasible. Under such circumstances, common techniques used by researchers include penalized linear models and nonlinear text regression models (such as SVMs and deep learning).[65]

If the underlying causal relationship runs from outcomes to language (a model of $p(c_i|v_i)$), then the third category, *generative model*, is suitable.[66] Applying *generative model* accounts for various dependencies among words (c_i) and among attributes (v_i) and helps researchers learn about how the attributes influence word choice.[67] This school of model can be further divided into *unsupervised* methods, *supervised* methods, and *semi-supervised* methods, based on whether v_i is observed. Latent Dirichlet allocation, a machine learning algorithm for probabilistic topic modeling that decomposes documents in terms of the fraction of time spent covering a variety of topics, is an example of an *unsupervised* model.[68]

All the previous schools of methods rely on a bag-of-words approach. The final category of methods, *word embeddings*, on the other hand, represents tokens in a vector space. Specifically, tokens in the vector space are relationally oriented, meaning that words that are "close" to each other in meaning in the specific context are mapped to locations close to each other in the vector space. This allows the model to encode more information about each token (such as the relative similarity between words in the particular context) than simply recording the location of each token in a particular sentence.

Word embedding techniques are widely used in NLP and has many benefits. First, it helps researchers interpret, analyze, and visualize the previously "hidden"

[64] For example, the relationship between the likelihood of passing an policy targeting climate change and number of news reports on the topic. We can predict the former from the latter.

[65] SVMs were originally proposed by Boser, Guyon, and Vapnik (1996).

[66] For example. the relationship between the number of news reports related to COVID and the actual COVID case numbers. The latter has an impact on the former, while the former cannot impact the latter.

[67] See Gentzkow, B. Kelly, and Taddy (2019) for a more detailed illustration.

[68] See Hansen and Prat (2018)

relationship between different words and tokens. Common NLP techniques could perform tasks such as sentiment analysis and analyzing the "closeness" in meaning between words in a particular context, and provide some form of visualization, which displays the relationship between word tokens in a more straightforward manner. Second, it helps researchers perform resource-intensive tasks with much higher efficiency. Research show that utilizing NLP techniques instead of entirely manual reviews could allow researchers to spend more time interpreting the results and developing action plans, rather than spending the majority of the time setting up and encoding the data. Finally, the use of algorithms could minimize human bias in the research process, such as to avoid confirmation bias that may cause the individuals sorting to miss or misunderstand important information when manually reading and categorizing response from a survey.[69]

Popular methods within this school of techniques include `word2vec` and Global Vector for Word Representation (GloVe).[70][71] Such techniques have already been adopted in a number of papers to perform various tasks. Aryal, Ciliberto, and Leyden (2020) adopt a mixed approach of NLP techniques and manual review to determine a list of words and phrases similar to the phrase "capacity discipline". Cao et al. (2020) document and compare the aggregate frequency of appearance for words in two existing lexicons to conduct sentiment analysis. In this paper, I apply `word2vec` in order to obtain a more comprehensive dictionary of relevant keywords based on the preliminary keyword sets in each case study that I obtain from manual review.

5.2 Introduction to `word2vec`

`word2vec` is a neural network architecture originally proposed in Mikolov, Chen, et al. (2013) and Mikolov, Sutskever, et al. (2013), and is considered the state-of-the-

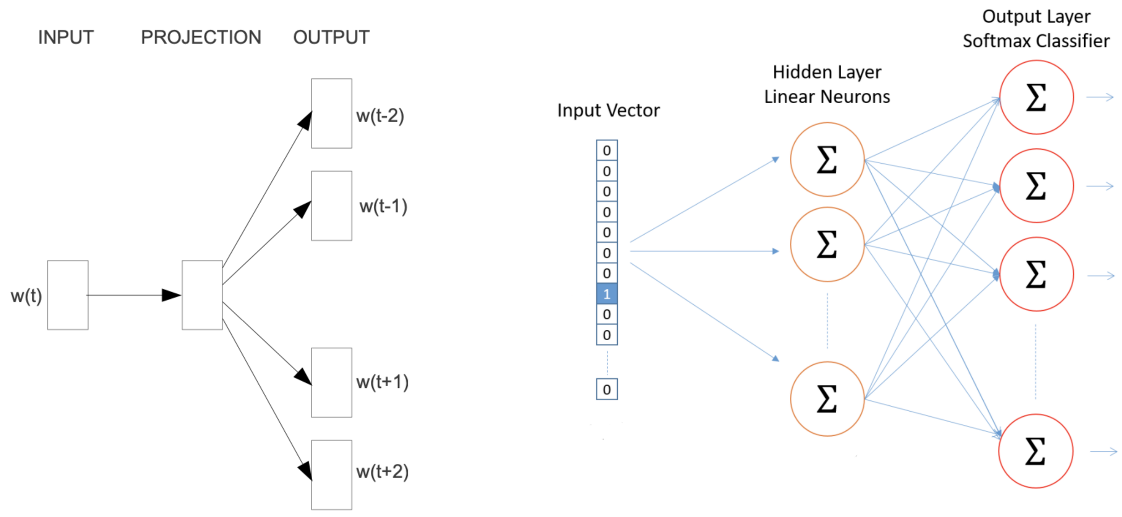
[69] See Chang (2020), Ignatow and Mihalcea (2018) and U. Kelly and Dr. Diane McDonald (2012)

[70] More information about `word2vec` can be found in Section 5.2.

[71] See Pennington, Socher, and Manning (2014) for more information about GloVe.

Figure 1: Skip-gram Model

(a) Purpose of a Skip-gram Model (b) Architecture of a Skip-gram Model



Source: Mikolov, Chen, et al. (2013), McCormick (2016)

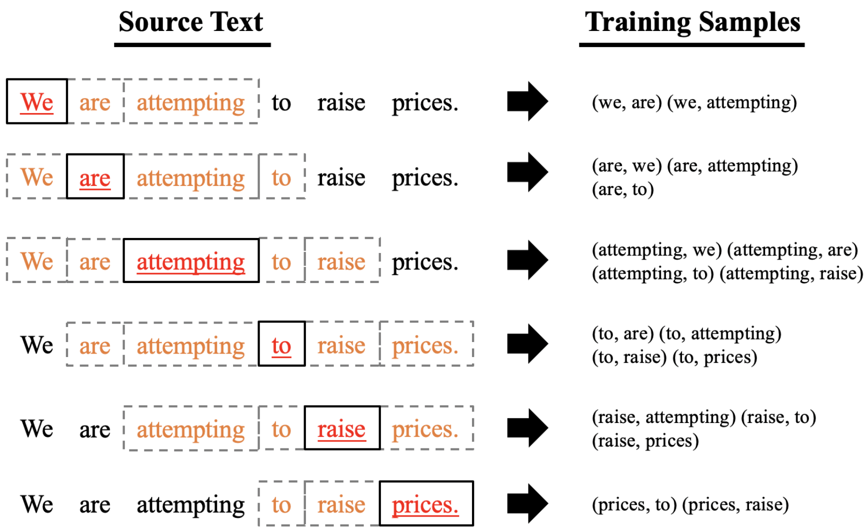
art embedding approach.[72] The basic intuition of `word2vec` is to assign the model to perform a fake task on a training data set, while the true motivation, which can be achieved while performing the fake task) is to obtain a vector that contains information on the relationship between the vocabularies in the training set. This subsection illustrates how the model is able to achieve this goal using an example based on the skip-gram model (presented in Figure 1).

As is displayed in Figure 1a, the original task of the model is to predict the surrounding words ($w(t-2)$,$w(t-1)$,$w(t+1)$, and $w(t+2)$) given a word $w(t)$, after training the model on a set of training documents. For example, given a training sentence "We are attempting to raise prices.", the model takes in training samples as illustrated in Figure 2, and outputs a probability distribution of surrounding words given a specific word input in the dictionary.

A detailed illustration of the structure of the skip-gram model is illustrated in Figure 1b. When training this network on word pairs, the input is a one-hot vector of size $1 \times V$, where V represents the number of unique words and phrases in the training data set. A "1" is placed in the position corresponding to the specific input word (such as "prices") in the input vector and the rest of the positions are filled

[72] See Malik (n.d.) for more on the development of `word2vec`.

Figure 2: Example of Skip-gram Model Input (window size = 2)



Note: This example illustrates the case of window size = 2, which means that for each word $w(t)$, only the previous two words and the following two words are considered neighboring words.

with zeros. The input is then fed into a hidden layer of linear neurons, where the input is multiplied by a weight matrix of size $V \times E$, producing an output of size $1 \times E$. Here, E is a hyper parameter representing the number of features we would like to learn the word pair with.[73] Notably, since in each round of training only one position in the vector has value "1", the operation performed in the hidden layer will effectively just select the matrix row corresponding to the "1". This means that the hidden layer of this model is really just operating as a lookup table. The output of the hidden layer is then fed into a Softmax regression classifier of size $E \times V$. This operation produces an output vector of size $1 \times V$ containing, for every word in our vocabulary, the probability that a randomly selected nearby word is that vocabulary word. For example, if our input word is "prices", then the output layer is a probability distribution of each word in the dictionary appearing near the word "prices".[74]

The output of the skip-gram model, however, is not of primary importance for our purpose of implementing `word2vec`, as this is our "fake task". Our true motivation is

[73] Google assigned $E = 300$ in its published paper using Google News as training data, but this hyper parameter is subject to further tuning to achieve best performance.

[74] See McCormick (2016) for a more detailed illustration of the model.

to obtain a “word vector” for representing the relationship between the input tokens, and this “word vector” happen to be the output of the hidden layer. Therefore, instead of taking the actual output of the skip-gram model, we need to obtain and utilize the output of the hidden layer as the output of our `word2vec` model.

Since `word2vec` maps every token into vector space, if two different tokens have very similar “contexts” (that is, what words are likely to appear around them), then our model needs to output very similar “word vectors” for these two tokens, which means that the two token are located very close when mapped into the vector space. In order to quantify the similarity between any two tokens, `word2vec` uses the cosine similarity to represent the “distance” between two tokens in a vector space. The metric is equal to the cosine of the angle between the vector representation of the two tokens, such that for any two normalized vectors associated with two tokens, k, and l, the measure of similarity is

$$d^{cos}(l, k) = \frac{k^T l}{||k|| \cdot ||l||},$$

where $|| \cdot ||$ is the L2 norm. Hence, when two vectors are the same, cosine similarity is 1, and when they are independent, cosine similarity is 0.[75]

5.3 Applying word2vec to Case Studies

Having explained the basic intuition behind `word2vec`, this section explains how `word2vec` is applied in my research to expand the preliminary keyword dictionary.

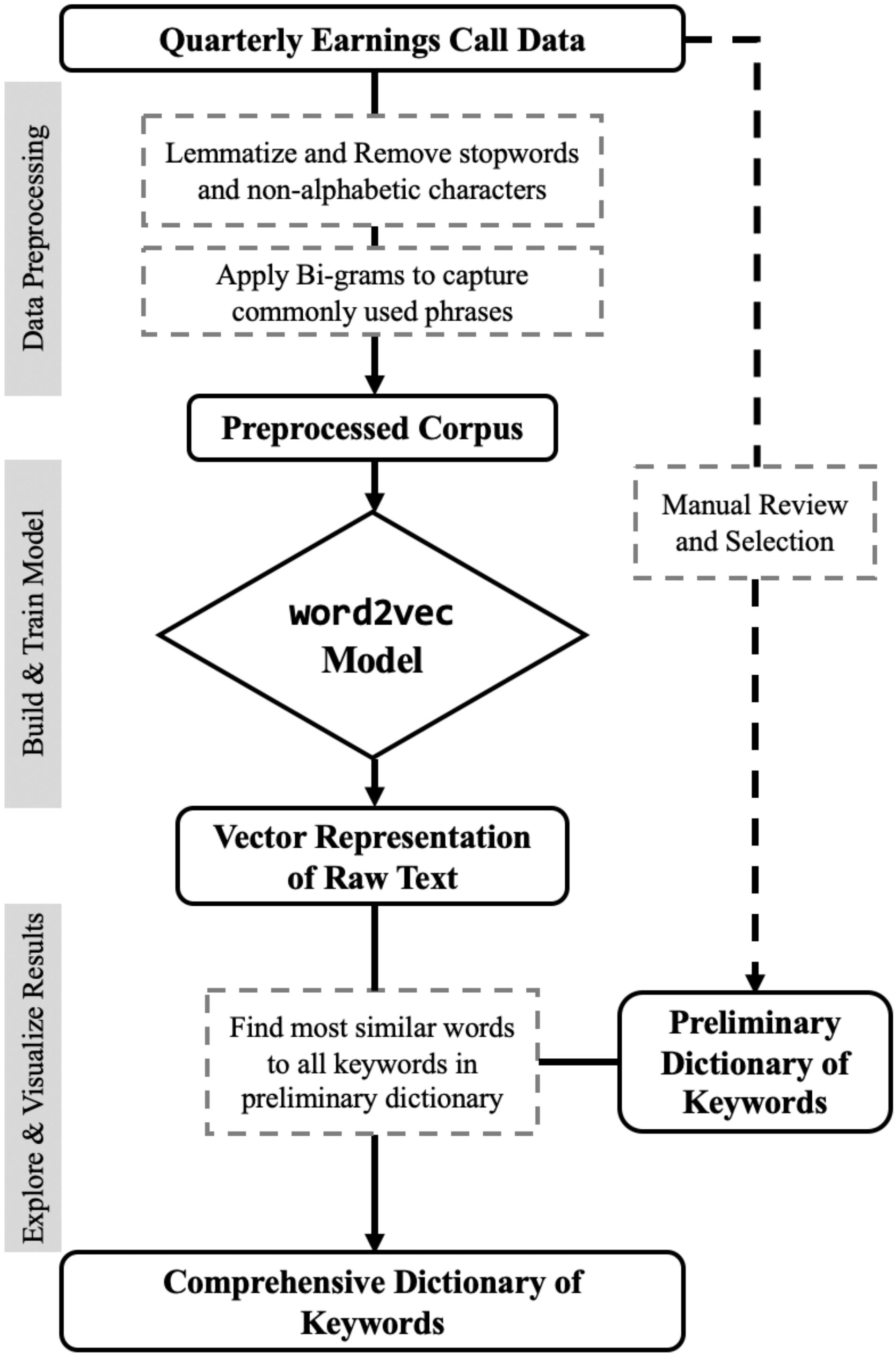
I use the `Gensim` library to implement the `word2vec` model in Python.[76] As is illustrated in Figure 3, the application of `word2vec` in my research could be separated into three phases: data pre-processing, building and training the `word2vec` model, and exploring and visualizing the results.

The input “training data” includes all the quarterly earnings call transcripts of all the colluding firms during the collusion period in each specific case under

[75] See Singhal 2001; see also Aryal, Ciliberto, and Leyden (2020) for an example.

[76] See Appendix B for example code; see Megret (2019) for similar implementation using a different data set.

Figure 3: Applying `word2vec` to expand keyword dictionary



Algorithm 1: `word2vec` **Model**

Input: Batch of pre-process corpus, *min_count*, *window*, *size*, *sample*, *alpha*, *min_alpha*, *negative*, *workers*, *sentences*, *progress_per*

Output: word vector

w2v_model = Word2Vec(*min_count*, *window*, *size*, *sample*, *alpha*, *min_alpha*, *negative*,*workers*, *epochs*);
w2v_model.build_vocab(*sentences*, *progress_per*);
w2v_model.train(*sentences*, total_examples=w2v_model.corpus_count, *epochs*, report_delay=1)

investigation. Since the purpose of the applying the `word2vec` model is to learn the relationship between specific keywords, I believe that stopwords and non-alphabetic characters (such as numbers) are unlikely to affect the results of the study. Therefore, the first step is to lemmatize the words in all the documents, and clean up the data set by removing all such words that do not contribute to the purpose of the model. Since `word2vec` uses context words to learn the vector representation of a target word, if a sentence too short, the benefit for the training is very small. Therefore I also remove all sentences that are less than or equal to two words long.

The next step of data pre-processing involves the application of Bi-grams. The concept of "n-gram" is discussed in Jurafsky and Martin (2019) as the idea to parse the document and extract information not only from one word, but also takes into account the broader context of the entire sentence where the word is located. The application of Bi-gram here allows us to detect and capture common phrases consisting of *two* words (hence "Bi"-gram) in the documents, such as "price increase" and "soft demand". Having cleaned up the raw text, we now have a list of pre-processed corpus to feed into the model.

Algorithm 1 illustrates the detailed building and training process of the `word2vec` model. The `Word2Vec()` function sets up the hyper parameters of the model one-by-one. The `build_vocab()` function builds the vocabulary from a sequence of sentences and thus initialized the model. The `train()` function trains the model on the given training data set, and hence, we obtain the vector representation of raw text that contains all information on the relationship between tokens in the dictionary.

Table 4: List of most similar tokens to specific keywords in *U-Haul International* Case

(a) Most Similar Tokens to "price"

Keyword or Phrase	Similarity
year	0.9487
think	0.9483
quarter	0.9435
go	0.9392
market	0.9388
truck	0.9383
increase	0.9372
fleet	0.9372
business	0.9361
look	0.9358

(b) Most Similar Tokens to "price leadership"

Keyword or Phrase	Similarity
continue	0.4323
area	0.4310
line	0.4213
progress	0.4170
step	0.4158
year	0.4137
Joe Shoen	0.4132
pricing	0.4119
applicable company	0.4106
deal	0.4087

(c) Most Similar Tokens to "truck rental"

Keyword or Phrase	Similarity
think	0.9019
increase	0.8997
quarter	0.8971
year	0.8970
time	0.8926
go	0.8920
continue	0.8889
market	0.8880
fleet	0.8873
business	0.8850

(d) Most Similar Tokens to "demand"

Keyword or Phrase	Similarity
think	0.8520
year	0.8446
market	0.8433
quarter	0.8391
increase	0.8374
look	0.8364
continue	0.8359
million	0.8355
go	0.8351
fleet	0.8348

Having obtained the vector representation, the following task is to expand the preliminary dictionary of keywords obtained from the manual review process in Section 4. Given a preliminary dictionary of keywords in each case, we apply the cosine similarity function (`most_similar()` in the `Gensim` Library) and obtain a list of 10 most similar tokens for each given keyword, each with a similarity score, ranging from 0 to 1. For each given keyword I only retain those tokens that: 1) has a cosine similarity of higher than 0.80, and 2) are among the top 10 in terms of the similarity to the given keyword. Then, I combine all the related tokens to keywords in the same case into a large dictionary after removing all duplicates. Following the above steps, I obtain an expanded list of keywords and phrases that firms frequently use to communicate the intention to collude. In the following subsections, I discuss the results after applying the `word2vec` model on each of the cases I study.

5.3.1 *U-Haul International*

The training data in this case is the quarterly earnings call transcripts by both AMERCO (the parent company of U-Haul) and Avis Budget Group (the parent company of Budget) that were held between year 2007 and 2008 (which is the time frame of communication as is indicated by FTC). The preliminary dictionary include {`price`, `price leadership`, `truck rental`, `demand`}.

Table 4 lists out the results after feeding the four preliminary keywords into the trained `word2vec` model. Applying the two criteria for combining the tokens, I obtain the expanded dictionary (in alphabetical order):

business	continue	demand	fleet	go
increase	look	market	million	price
price leadership	quarter	think	time	truck
truck rental	year			

Table 5: List of most similar tokens to specific keywords in *Valassis Communications* Case

(a) Most Similar Tokens to "price"

Keyword or Phrase	Similarity
business	0.9906
think	0.9898
customer	0.9892
client	0.9890
know	0.9888
look	0.9882
FSI	0.9880
year	0.9879
kind	0.9878
Alan Schultz	0.9875

(b) Most Similar Tokens to "difference"

Keyword or Phrase	Similarity
business	0.9187
client	0.9174
time	0.9163
year	0.9161
go	0.9157
think	0.9145
customer	0.9144
kind	0.9144
like	0.9132
expect	0.9128

(c) Most Similar Tokens to "profitability"

Keyword or Phrase	Similarity
know	0.9582
think	0.9582
product	0.9580
client	0.9578
year	0.9578
kind	0.9570
Robert Recchia	0.9566
business	0.9563
FSI	0.9561
customer	0.9559

(d) Most Similar Tokens to "competition"

Keyword or Phrase	Similarity
business	0.9115
customer	0.9097
price	0.9091
like	0.9075
look	0.9070
revenue	0.9067
know	0.9059
FSI	0.9054
perspective	0.9054
line	0.9050

5.3.2 *Valassis Communications*

The training data in the *Valassis Communications* case is the quarterly earnings call transcripts by Valassis Communications Inc. that were held in year 2004 (which is the time frame of communication as is indicated by FTC). The preliminary dictionary include {`price, difference, profitability, competition`}.

Table 5 lists out the results after feeding the four preliminary keywords into the trained `word2vec` model. Applying the two criteria for combining the tokens, I obtain the expanded dictionary (in alphabetical order):[77]

business	client	competition	customer	difference
expect	FSI	go	kind	know
like	line	look	perspective	price
product	profitability	revenue	think	time
year				

Table 6: List of most similar tokens to specific keywords in *Delta/AirTran* Case

(a) Most Similar Tokens to "initiate"

Keyword or Phrase	Similarity
unit cost	0.8892
say	0.8883
receive	0.8882
okay	0.8860
think	0.8860
cost	0.8858
expect	0.8858
continue	0.8856
growth	0.8852
liquidity	0.8852

(b) Most Similar Tokens to "bag fee"

Keyword or Phrase	Similarity
million	0.9674
Lines Inc	0.9661
year	0.9660
opportunity	0.9659
cost	0.9657
fourth quarter	0.9656
think	0.9655
AirTran Holdings	0.9655
capacity	0.9655
number	0.9655

(c) Most Similar Tokens to "competitor"

Keyword or Phrase	Similarity
think	0.9445
go	0.9444
Delta	0.9438
year	0.9436
quarter	0.9435
cost	0.9435
Lines Inc	0.9434
okay	0.9433
term	0.9423
million	0.9423

(d) Most Similar Tokens to "follow"

Keyword or Phrase	Similarity
Delta	0.9652
year	0.9647
quarter	0.9645
cost	0.9644
think	0.9644
month	0.9640
question	0.9634
look	0.9634
fuel	0.9631
go	0.9630

[77] Here, I get rid of all the names that appear in the list of most similar words, since specific names are unlikely to have an impact on the intention to collude.

5.3.3 *Delta/AirTran*

The training data in the *Delta/AirTran* case is the quarterly earnings call transcripts by Delta Air Lines, Inc. and AirTran Holdings, Inc. that were held in year 2008 (which is the time frame of communication as is indicated by DOJ), as well as the press releases of Delta following AirTran's indication of interest in introducing a first bag fee in October, 2008. The preliminary dictionary, on the other hand, include {`initiate, bag fee, competitor, follow`}.

Table 6 lists out the results after feeding the four preliminary keywords into the trained `word2vec` model. Applying the two criteria for combining the tokens, I obtain the expanded dictionary (in alphabetical order):[78]

AirTran Holdings	bag fee	capacity	competitor	continue
cost	Delta	expect	follow	fuel
go	growth	initiate	liquidity	look
million	month	number	opportunity	quarter
question	receive	say	think	term
unit cost	year			

5.3.4 *Generic Drugs*

The training data in the *Generic Drugs* case is the quarterly earnings call transcripts by Lannett Company, Inc. that were held in years 2013 to 2015, as is reviewed in Harrington Jr. and Kashfipour (2020). The preliminary dictionary include {`price increase, responsibility, profit margin, cost, competition`}.

Table 7 lists out the results after feeding the five preliminary keywords into the trained `word2vec` model. Applying the two criteria for combining the tokens, I obtain the expanded dictionary (in alphabetical order):[79]

[78] Here, I get rid of stopwords such as "okay" that were not filtered out in the pre-processing stage as well as "Lines Inc." that is a portion of the company name and duplicates "Delta".

[79] Here, I get rid of stopwords such as "okay" that failed to be filtered out in the pre-processing stage, as well as names, such as Arthur Bedrosian.

brief overview	company	competition	cost	expect
go	have	increase	know	Lannett Company
like	look	million	net sale	opportunity
price	price increase	product	profit margin	quarter
raise price	responsibility	risk	see	talk
think	year			

Table 7: List of most similar tokens to specific keywords in *Generic Drugs* Case

(a) Most Similar Tokens to "price increase"

Keyword or Phrase	Similarity
Arthur Bedrosian	0.9980
product	0.9979
million	0.9978
company	0.9976
Lannett Company	0.9973
year	0.9973
net sale	0.9973
look	0.9973
think	0.9973
increase	0.9972

(b) Most Similar Tokens to "responsibility"

Keyword or Phrase	Similarity
opportunity	0.9142
brief overview	0.9140
increase	0.9137
see	0.9137
like	0.9137
q	0.9133
analyst	0.9133
raise price	0.9132
risk	0.9132
Arthur Bedrosian	0.9131

(c) Most Similar Tokens to "profit margin"

Keyword or Phrase	Similarity
think	0.9800
Arthur Bedrosian	0.9798
product	0.9797
Lannett Company	0.9797
expect	0.9795
go	0.9795
net sale	0.9793
price increase	0.9792
million	0.9791
talk	0.9791

(d) Most Similar Tokens to "cost"

Keyword or Phrase	Similarity
million	0.9953
think	0.9952
company	0.9952
product	0.9952
Arthur Bedrosian	0.9952
price	0.9950
Lannett Company	0.9949
have	0.9947
know	0.9947
go	0.9946

(e) Most Similar Tokens to "competition"

Keyword or Phrase	Similarity
product	0.9968
Arthur Bedrosian	0.9967
company	0.9966
million	0.9965
net sale	0.9963
quarter	0.9961
price increase	0.9961
Lannett Company	0.9961
look	0.9960
go	0.9960

5.4 Results analysis

The previous subsections present a detailed analysis of four past antitrust cases in four different industries, including the truck rental industry, the FSI industry, the airline industry, and the generic pharmaceuticals industry. This section provides a brief analysis of the results.

1. All four cases happen in concentrated markets.

One similarity across the above cases is that all four cartels take place in very concentrated industries. In the *U-Haul International* case, U-Haul, Budget, and Penske are the only three firms in the truck rental industry; in the *Valassis Communications* case, Valassis and News America are close competitors, each occupying roughly half of the market; in the *Delta/AirTran* case, the two firms combined controlled 92% of the dominated route markets which had Atlanta as the origin or destination; and in the *General Drugs* case, even though there were more than two dozens firms under investigation, Lannett indicated that it was one of the few major players in all three of the markets where it proposed a price increase. A similar point was also brought in Markham (1951), where the author identifies as one key characteristics of a cartel the fact that “firms must be few in number and each firm must be sufficiently large.”

2. The firm that initiates the conversation tends to be the market leader (or claims to be so).

The fact that all four firms that initiate the talk on price increases tend to be the market leader is another common property of the four cases. In the *U-Haul International* case, U-Haul is a clear market leader in the truck rental industry, as it was at the time the largest player in terms of truck numbers, rental locations, revenues, as well as market share. Delta was also much larger than AirTran in terms of routes and revenues in the *Delta/AirTran* case. Interestingly, however, both Valassis and Lannett attempted to make the claim that they are the market/price leader even if they were commonly perceived as the follower. In the *Valassis Communications* case, even if Valassis had a smaller market share than News America when it initiated the price increase, Valassis branded itself is as the price leader, whose job is to "take on that responsibility...to look for ways to improve the long term pricing trend in the FSI industry."[80] In the *Generic Drugs* case, although an analyst at Lannett's earnings call indicated that he had usually thought of Lannett as a price follower, the CEO of Lannett countered the argument and claimed that "With one or two exceptions, we've tended to lead in the way of price increases."[81]

3. All four cases involve firms expressing anticipation towards competitors' behaviors.

Another property that all four cases above share is the fact that firms in all cases express anticipation of competitors' behaviors in the earnings calls immediately following a declaration of price increase, and executives at most firms express encouragement for competitors to follow suit. In the *U-Haul International* case, top executives of U-Haul, following a proposed price increase, explicitly indicated their hope that Budget will follow.[82] In the *Valassis Communications* case, leaders of Valassis also stated their expectations towards News America's response to its newly proposed pricing scheme.[83] In the *Generic Drugs* case, Lannett repeatedly emphasized its intention and action to raise prices, and encouraged its competitors to raise prices by commending competitors who followed suit as "responsible" and "rational".[84] Although neither company in the *Delta/AirTran* case expressed explicit encouragement for the competitors to introduce the bag fee, both indicated the interest to introduce the fee, and AirTran expressed its anticipation for Delta's strategy by declaring itself as a "follower" in the market.[85]

[80] Q2 2004 Valassis Communications Inc. Earnings Conference Call - Final (July 22, 2004)

[81] Q4 2013 Lannett Company Inc. Earnings Conference Call - Final (September 10, 2013)

[82] See Q2 2007 AMERCO Earnings Conference Call - Final (November 9, 2006)

[83] See Q2 2004 Valassis Communications Inc. Earnings Conference Call - Final (July 22, 2004), the same quarter when Valassis proposed its new pricing scheme.

[84] See a list of examples in Section 4

[85] See Q3 2008 AirTran Holdings, Inc. Earnings Conference Call - Final (October 23, 2008)

4. The focus on price-fixing behaviors allow us to develop a generic dictionary.

Since all of the cartels were focused on fixing price, rather than quantity, this allows us to combine the dictionaries obtained from the four cartels into a larger dictionary that contains keywords and phrases that could indicate price-fixing behaviors.[86] After removing all industry-specific keywords from the dictionaries, we obtain the set of generic words and phrases (listed below in alphabetical order):[87]

brief overview	business	capacity	client	company
competition	competitor	continue	cost	customer
demand	difference	expect	follow	go
growth	have	increase	initiate	kind
know	like	line	liquidity	look
market	million	month	net sale	number
opportunity	perspective	price	price increase	price leadership
product	profit margin	profitability	quarter	question
raise price	receive	responsibility	revenue	risk
say	see	talk	term	think
time	unit cost	year		

[86] While the *Delta/AirTran* case concerns the introduction of a new fee, all the other three cartels studied concern the increasing of prices of one or more existing products.

[87] Industry-specific keywords that are removed include "fleet", "truck", "truck rental" (specific to Truck Rental Industry); "FSI" (specific to FSI Industry); "AirTran Holdings", "bag fee", "Delta", "fuel" (specific to Airline Industry); "Lannett Company" (specific to Generic Drugs Industry). It is important to note that the keywords specific to industries, such as company names and industry names, constitute an important part of the communication that could indicate collusion. Here I remove it only for the purpose of building a generic dictionary that could be applied across different industries.

6 Conclusion

This paper studies quarterly earnings calls as a means of communication between colluding firms. By performing a combination of manual review and text analysis on four past antitrust cases involving price-fixing behaviors, I propose a dictionary that contains generic words and phrases that could indicate potential intention to fix prices.

It is important, however, to note the potential of selection biases in the process, as well as the need for other source of information to complement the findings from earnings calls. First, since earnings calls are only available for publicly traded firms, only the languages and information that is revealed by the public firm is available for interpretation and analysis. If most members of a cartel are not publicly traded, only analyzing the quarterly earnings calls might result in significant biases of information regarding the conducts and decisions of different firms.

Additionally, the size and structure of the firm could also be a source of bias. Take the *U-Haul International* case as an example. Although both U-Haul and Budget were subsidiaries of a parent company, U-Haul was the largest subsidiary of AMERCO, whereas Budget was only recently acquired by Cendant (owner of Avis), whose primary line of business was not in truck rental. As a result, the AMERCO earnings call was almost exclusively focused on U-Haul as well as the truck rental industry, whereas very few information could be extracted from the earnings calls of Avis Budget Group, making it hard to determine Budget's involvement in the collusion.

Another source of limitation comes from the fact that the appearance of the words in the dictionary in any earnings calls alone cannot guarantee the existence of collusion. While it is rather easy to make sense of how such keywords as "price increase" and "competition" could be used in facilitating collusion, these words are also fairly common words that are justified to appear in any quarterly earnings calls. This suggests the need for other supporting evidence, such as corresponding price increase by multiple competing firms in the same period of time, to determine the existence of collusive behaviors. Harrington Jr. (2006) lists out many characteristics in the market that could indicate the presence of a cartel, and combining the information in this paper as well as the methodology in this paper could be an interesting next step.

One interesting finding when comparing keywords in the four dictionary is the fact that many keywords appear in the dictionary of more than one cases. For example, the keyword "cost" appears in the dictionary of both the *U-Haul International* case and the *Generic Drugs* case. In the former case, U-Haul claims that the prices at the time was "below the true cost of the product", hinting at the need to raise the price to cover cost; in the latter case, top executives of Lannett Company emphasizes the increasing cost induced by new FDA requirements, and explains that the primary motivation of price increase is to cover cost.[88][89] This overlap in keywords across multiple dictionaries can indicate some common pattern of the use of keywords during engaging in price-fixing behaviors and could be a topic for further research.

[88] Q3 2007 AMERCO Earnings Conference Call - Final (February 8, 2007)

[89] Q4 2013 Lannett Company Inc. Earnings Conference Call - Final (September 10, 2013)

References

Aryal, Gaurab, Federico Ciliberto, and Benjamin T. Leyden. "Coordinated Capacity Reduction and Public Communication in the Airline Industry". Feb. 2020.

Awaya, Yu and Vijay Krishna. "Communication and Cooperation in Repeated Games". In: *Theoretical Economics* 14.2 (May 2019), pp. 513–553. DOI: 10.3982/TE3049.

— "Information Exchange in Cartels". In: *The RAND Journal of Economics* 51.2 (May 2020), pp. 421–446. DOI: 10.1111/1756-2171.12320.

— "On Communication and Collusion". In: *American Economic Review* 106.2 (Feb. 2016), pp. 285–315. DOI: 10.1257/aer.20141469.

Boser, Bernhard, Isabelle Guyon, and Vladimir Vapnik. "A Training Algorithm for Optimal Margin Classifier". In: *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory* 5 (Aug. 1996). DOI: 10.1145/130385.130401.

Cao, Sean et al. "How to Talk When a Machine is Listening: Corporate Disclosure in the Age of AI". NBER Working Paper Series. Oct. 2020.

Chang, Kevin. *The Powerful Benefits of Text Analysis (NLP) in Higher Education.* 2020. URL: https://www.kaianalytics.com/post/benefits-of-nlp (visited on 03/02/2020).

Connor, John M. "Price-fixing Overcharges". Feb. 2014.

Doyle, Maura P. and Christopher M. Snyder. "Information Sharing and Competition in the Motor Vehicle Industry". In: *Journal of Political Economy* 107.6 (Dec. 1999), pp. 1326–1364.

Friedman, James. "A Non-cooperative Equilibrium for Supergames". In: *The Review of Economic Studies* 38.1 (Jan. 1971), pp. 1–12. DOI: 10.2307/2296617.

Genesove, David and Wallace P. Mullin. "Rules, Communication, and Collusion: Narrative Evidence from the Sugar Institute Case". In: *American Economic Review* 91.3 (June 2001), pp. 379–398. DOI: 10.1257/aer.91.3.379.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. "Text as Data". In: *Journal of Economic Literature* 57.3 (Sept. 2019), pp. 535–74. DOI: 10.1257/jel.20181020.

Green, Edward J. and Robert Porter. "Noncooperative Collusion under Imperfect Price Information". In: *Econometrica* 52.1 (Jan. 1984), pp. 87–100. DOI: 10.2307/1911462.

Hansen, Stephen and Michael McMahon andAndrea Prat. "Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach". In: *The Quarterly Journal of Economics* 133.2 (2018), pp. 801–870. DOI: 10.1093/qje/qjx045.

Harrington, Joseph E. and Andrzej Skrzypacz. "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices". In: *American Economic Review* 101 (2011), pp. 1–25. DOI: 10.1257/aer.101.6.1.

Harrington, Joseph E. and Lixin Ye. "Collusion Through Coordination of Announcements". In: *The Journal of Industrial Economics* 67 (2 Dec. 2019), pp. 209–241. DOI: 10.1111/joie.12199.

Harrington Jr., Joseph E. "How do Cartels Operate?" In: *Foundations and Trends in Microeconomics* 2 (1 2006), pp. 1–105. DOI: 10.1561/0700000021.

Harrington Jr., Joseph E. and Ehson Kashfipour. "Collusion in Plain Sight: Firms' Use of Public Announcements to Restrain Competition". In: Competition & Regulation European Summer School and Conference (CRESSE). July 2020.

URL: https://www.cresse.info/wp-content/uploads/2020/09/Collusion-in-Plain-Sight_030720.pdf.
Ignatow, Gabe and Rada Mihalcea. *Text Mining and Text Analysis*. 2018. Chap. 1. ISBN: 9781506337005. DOI: 10.4135/9781506336985.n1.
Jurafsky, Daniel and James H. Martin. *N-gram Language Models*. 2nd ed. Oct. 2019. Chap. 3.
Kaplow, Louis. *Competition Policy and Price Fixing*. Princeton University Press, 2013. DOI: 10.2307/j.ctt2jc8k1.
Kelly, Ursula and Dr. Diane McDonald. *The Value and Benefit of Text Mining to UK Further and Higher Education*. JISC, 2012.
Loughran, Tim and Bill McDonald. "When is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks". In: *The Journal of Finance* 66.1 (Feb. 2011), pp. 35–65. DOI: 10.2139/ssrn.875399.
Malik, Usman. *Implementing Word2Vec with Gensim Library in Python*. URL: https://stackabuse.com/implementing-word2vec-with-gensim-library-in-python/.
Markham, Jesse W. "The Nature and Significance of Price Leadership". In: *American Economic Review* 41.5 (Dec. 1951), pp. 891–905.
McCormick, Chris. *Word2Vec Tutorial - The Skip-Gram Model*. Apr. 2016. URL: http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/.
Megret, Pierre. *Gensim Word2Vec Tutorial*. 2019. URL: https://www.kaggle.com/pierremegret/gensim-word2vec-tutorial#Acknowledgements.
Mikolov, Tomas, Kai Chen, et al. *Efficient Estimation of Word Representations in Vector Space*. 2013. URL: http://arxiv.org/abs/1301.3781.
Mikolov, Tomas, Ilya Sutskever, et al. "Distributed representations of words and phrases and their compositionality". In: *arXiv preprint arXiv:1310.4546* (2013).
OECD. *Unilateral Disclosure of Information with Anticompetitive Effects (e.g. Through Press Announcements)*. Submission to Working Party No.3 on Co-operation and Enforcement DAF/COMP/WP3/WD(2012)14. Version 2012-02-14. Organization for Economic Co-operation, Development Directorate for Financial, and Enterprise Affairs Competition Committee, Feb. 8, 2012.
Pennington, Jeffrey, Richard Socher, and Christopher Manning. "GloVe: Global Vectors for Word Representation". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1532–1543. DOI: 10.3115/v1/D14-1162. URL: https://www.aclweb.org/anthology/D14-1162.
Rahman, David. "The Power of Communication". In: *American Economic Review* 104 (11 Nov. 2014), pp. 3737–3751. DOI: 10.1257/aer.104.11.3737.
Scott R. Baker Nicholas Bloom, Steven J. Davis. "Measuring Economic Policy Uncertainty". In: *The Quarterly Journal of Economics* 131 (4 Nov. 2016), pp. 1593–1636. DOI: 10.1093/qje/qjw024.
Singhal, Amit. "Modern Information Retrieval: A Brief Overview". In: *IEEE Data Eng. Bull.* 24 (2001), pp. 35–43.
Spector, David. "Cheap Talk, Monitoring, and Collusion". 2020.
Stigler, George. "A Theory of Oligopoly". In: *Journal of Political Economy* 72 (1 Feb. 1964), pp. 44–61. DOI: 10.1086/258853.

Tetlock, Paul C. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market". In: *The Journal of Finance* 62.3 (May 2007), pp. 1139–68. DOI: 10.1111/j.1540-6261.2007.01232.x.

Appendices

A Description for Important Variables in the PIC Dataset

Table 8: Description for Important Variables in the PIC Dataset

Variables	Description
Cartel Number	Index for cartels in the dataset
Firm Number	Index for individual firms in each cartel
Location - Region	The region where the cartel operates
Location - Country	The country where the cartel operates
Industry Number	The broader industry where the cartel operates
Market	The specific markekt where the cartel operates
Firm - Subsidiary Name	The name of the subsidiary of the firm that engaged in collusive behavior, if any
Firm - Subsidiary Nation	The country where the subsidiary of the firm that engaged in collusive behavior is located
Firm - Intermediate Operating Company Name	The name of the intermediate operating company of the firm that engaged in collusive behavior, if any
Firm - Intermediate Operating Company Nation	The country where the intermediate operating company of the firm that engaged in collusive behavior is located
Firm - Parent Name	The name of the parent company of the firm that engaged in collusive behavior, if any
Firm - Parent Nation	The country where the parent company of the firm that engaged in collusive behavior is located
Cartel Start Year	The year in which the cartel started
Cartel Start Month	The month in which the cartel started
Cartel End Year	The year in which the cartel ended
Cartel End Month	The month in which the cartel ended
Cartel Diff. Months	The total number of months the cartel colluded
Firm Start Year	The year in which the firm started to participate in collusive behavior
Firm Start Month	The month in which the firm started to participate in collusive behavior
Firm End Year	The year in which the firm stopped to participate in collusive behavior
Firm End Month	The month in which the firm stopped to participate in collusive behavior
Firm Diff. Months	The total number of months the firm engaged in collusion

B Example Code

```
import re  # For preprocessing
import pandas as pd  # For data handling
from time import time  # To time our operations
from collections import defaultdict  # For word frequency
import spacy  # For preprocessing
import logging  # Setting up the loggings to monitor gensim
logging.basicConfig(format="%(levelname)s - %(asctime)s: %(message)s", datefmt= '%H:%M:%S', level=logging.INFO)
from google.colab import drive
drive.mount('/content/drive')
pip install PyPDF2
import PyPDF2
```

Data Preprocessing

```
# Loading Data
def read_earnings(path):
  pdfFileObj = open(path, 'rb')
  pdfReader = PyPDF2.PdfFileReader(pdfFileObj)
  content = ""
  for i in range(pdfReader.numPages):
    content += pdfReader.getPage(i).extractText()
  pdfFileObj.close()
  return content

df = pd.DataFrame(columns=['Contents'])
root = "/content/drive/MyDrive/Truck Rental Industry/"

for i in range(18):
    path = root + str(i) + ".pdf"
    content = read_earnings(path)
    df = df.append({'Contents' : content}, ignore_index = True)

# Data Cleaning
nlp = spacy.load('en', disable=['ner', 'parser'])
```

```

def cleaning(doc):
    txt = [token.lemma_ for token in doc if not token.is_stop]
    if len(txt) > 2:
        return ' '.join(txt)

brief_cleaning = (re.sub("[^A-Za-z']+", ' ', str(row)).lower() for row in df['Contents'])
t = time()
txt = [cleaning(doc) for doc in nlp.pipe(brief_cleaning, batch_size=5000, n_threads=-1)]
print('Time to clean up everything: {} mins'.format(round((time() - t) / 60, 2)))
df_clean = pd.DataFrame({'clean': txt})
df_clean = df_clean.dropna().drop_duplicates()
print(df_clean.shape)

# Bigrams
from gensim.models.phrases import Phrases, Phraser

sent = [row.split() for row in df_clean['clean']]
phrases = Phrases(sent, min_count=1, progress_per=10000)
bigram = Phraser(phrases)
sentences = bigram[sent]

# Display Most Frequent Words (Sanity Check)
word_freq = defaultdict(int)
for sent in sentences:
    for i in sent:
        word_freq[i] += 1
print(len(word_freq))
print(sorted(word_freq, key=word_freq.get, reverse=True)[:10])
```

Build and Train the Model

```
# Build Model
```

```python
import multiprocessing
from gensim.models import Word2Vec

cores = multiprocessing.cpu_count() # Count the number of cores in a computer

w2v_model = Word2Vec(min_count=2,
                     window=7,
                     size=300,
                     sample=0.8e-5,
                     alpha=0.036,
                     min_alpha=0.0001,
                     negative=5,
                     workers=cores-1)
# Note: all hyper parameters are subject to tuning to maximize performance

# Build Vocabulary Table
t = time()
w2v_model.build_vocab(sentences, progress_per=10000)
print('Time to build vocab: {} mins'.format(round((time() - t) / 60, 2)))

# Train the Model
t = time()
w2v_model.train(sentences, total_examples=w2v_model.corpus_count, epochs=10, report_delay=1)
print('Time to train the model: {} mins'.format(round((time() - t) / 60, 2)))

# Store the Model
w2v_model.init_sims(replace=True)
```

Use and Explore the Model

```python
# Most Similar To
```

```
print(w2v_model.wv.most_similar(positive=["price"]))

# Visualization
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
sns.set_style("darkgrid")
from sklearn.decomposition import PCA
from sklearn.manifold import TSNE

def tsnescatterplot(model, word, list_names):
    """ Plot in seaborn the results from the t-SNE dimensionality
   reduction algorithm of the vectors of a query word,
    its list of most similar words, and a list of words.
    """
    arrays = np.empty((0, 300), dtype='f')
    word_labels = [word]
    color_list  = ['red']

    # adds the vector of the query word
    arrays = np.append(arrays, model.wv.__getitem__([word]), axis
   =0)

    # gets list of most similar words
    close_words = model.wv.most_similar([word])

    # adds the vector for each of the closest words to the array
    for wrd_score in close_words:
        wrd_vector = model.wv.__getitem__([wrd_score[0]])
        word_labels.append(wrd_score[0])
        color_list.append('blue')
        arrays = np.append(arrays, wrd_vector, axis=0)

    # adds the vector for each of the words from list_names to the
```

```
array
    for wrd in list_names:
        wrd_vector = model.wv.__getitem__([wrd])
        word_labels.append(wrd)
        color_list.append('green')
        arrays = np.append(arrays, wrd_vector, axis=0)

    # Reduces the dimensionality from 300 to 50 dimensions with PCA
    reduc = PCA(n_components=19).fit_transform(arrays)

    # Finds t-SNE coordinates for 2 dimensions
    np.set_printoptions(suppress=True)

    Y = TSNE(n_components=2, random_state=0, perplexity=15).
fit_transform(reduc)

    # Sets everything up to plot
    df = pd.DataFrame({'x': [x for x in Y[:, 0]],
                       'y': [y for y in Y[:, 1]],
                       'words': word_labels,
                       'color': color_list})

    fig, _ = plt.subplots()
    fig.set_size_inches(9, 9)

    # Basic plot
    p1 = sns.regplot(data=df,
                     x="x",
                     y="y",
                     fit_reg=False,
                     marker="o",
                     scatter_kws={'s': 40,
                                  'facecolors': df['color']
                                 }
                    )
```

```

    # Adds annotations one by one with a loop
    for line in range(0, df.shape[0]):
         p1.text(df["x"][line],
                 df['y'][line],
                 '  ' + df["words"][line].title(),
                 horizontalalignment='left',
                 verticalalignment='bottom', size='medium',
                 color=df['color'][line],
                 weight='normal'
                ).set_size(15)


    plt.xlim(Y[:, 0].min()-50, Y[:, 0].max()+50)
    plt.ylim(Y[:, 1].min()-50, Y[:, 1].max()+50)

    plt.title('t-SNE visualization for {}'.format(word.title()))

tsnescatterplot(w2v_model, 'price', [i[0] for i in w2v_model.wv.
    most_similar(negative=["price"])]) # plot the words that are
    most similar to the keyword "price", as well as the words that
    are least similar
```