Duso, Tomaso; Harrington, Joseph E.; Kreuzberg, Carl; Sapi, Geza

Working Paper

Public communication and collusion: New screening tools for competition authorities

DICE Discussion Paper, No. 430

Provided in Cooperation with:
Düsseldorf Institute for Competition Economics (DICE), Heinrich Heine University Düsseldorf

Suggested Citation: Duso, Tomaso; Harrington, Joseph E.; Kreuzberg, Carl; Sapi, Geza (2025) : Public communication and collusion: New screening tools for competition authorities, DICE Discussion Paper, No. 430, ISBN 978-3-86304-429-9, Heinrich Heine University Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Düsseldorf

This Version is available at:
https://hdl.handle.net/10419/329636

DICE

DISCUSSION PAPER

NO 430

Public Communication and Collusion: New Screening Tools for Competition Authorities

Tomaso Duso
Joseph E. Harrington Jr.
Carl Kreuzberg
Geza Sapi

October 2025

hhu
Heinrich Heine
Universität Düsseldorf

IMPRINT

DICE DISCUSSION PAPER

Published by:
Heinrich-Heine-University Düsseldorf,
Düsseldorf Institute for Competition Economics (DICE),
Universitätsstraße 1, 40225 Düsseldorf, Germany
www.dice.hhu.de

Editor:
Prof. Dr. Hans-Theo Normann
Düsseldorf Institute for Competition Economics (DICE)
Tel +49 (0) 211-81-15125, E-Mail normann@dice.hhu.de

ISSN 2190-9938 (online) / ISBN 978-3-86304-429-9

Public Communication and Collusion: New Screening Tools for Competition Authorities*

Tomaso Duso,† Joseph E. Harrington, Jr.,‡ Carl Kreuzberg§ and Geza Sapi¶

October 2025

Abstract

Competition authorities increasingly rely on economic screening tools to identify markets where firms deviate from competitive norms. Traditional screening methods assume that collusion occurs through secret agreements. However, recent research highlights that firms can use public announcements to coordinate decisions, reducing competition while avoiding detection. We propose a novel approach to screening for collusion in public corporate statements. Using natural language processing, we analyze more than 300,000 earnings call transcripts issued worldwide between 2004 and 2022. By identifying expressions commonly associated with collusion, our method provides competition authorities with a tool to detect potentially anticompetitive behavior in public communications. Our approach can extend beyond earnings calls to other sources, such as news articles, trade press, and industry reports. Our method informed the European Commission's 2024 unannounced inspections in the car tire sector, prompted by concerns over price coordination through public communication.

Keywords: Communication, Collusion, NLP, Screening, Text Analysis.
JEL Codes: C23, D22, L1, L4, L64

*We thank Maria Jaspers, Sari Suurnäkki and Martin Amegah for helpful comments and suggestions. Tomaso Duso is the chairman of the German Monopolies Commission. Geza Sapi is senior economist in the Cartels Directorate of DG Competition of the European Commission, where he has been involved in the AT.40863 Replacement Tyres cartel investigation, mentioned in the article. The views expressed in this article are those of the authors and may not in any circumstances be regarded as stating an official position of the German Monopolies Commission, DG Competition, or of the European Commission. The inclusion of any company in this article does not imply any wrongdoing or collusion. All examples are presented solely for illustrative purposes, reflecting the author's academic freedom and responsibility. In preparing this manuscript, we used the language model ChatGPT and DeepL to refine and improve the clarity and flow of the writing. We thank Felix Aubele for excellent research assistance.

†Deutsches Institut für Wirtschaftsforschung (DIW Berlin), Technische Universität Berlin, CEPR, and CESifo. E-mail: tduso@diw.de.

‡Department of Business Economics & Public Policy, The Wharton School, University of Pennsylvania. E-mail: harrij@wharton.upenn.edu.

§Deutsches Institut für Wirtschaftsforschung (DIW Berlin), Technische Universität Berlin. E-mail: ckreuzberg@diw.de.

¶European Commission and Düsseldorf Institute for Competition Economics, Heinrich-Heine Universität, Düsseldorf. E-mail: geza.sapi@ec.europa.eu.

I Introduction

Collusion between firms poses a significant challenge to the effective functioning of competitive markets. Despite being a top priority for competition authorities worldwide, detecting collusion remains a difficult task. To address this challenge, agencies have increasingly employed economic screens to identify markets where firms' behavior deviates from competitive norms. Such empirical screens are typically statistical tests rooted in economic theory that can be used to focus the attention of competition authorities on markets or firms that warrant further scrutiny. Screening methods are important elements in the toolbox of competition authorities to generate enforcement leads *ex-officio*.[1]

Screening for cartels is crucial for effective cartel enforcement for two key reasons. Firstly, the ability to initiate *ex-officio* investigations avoids situations where competition authorities only rely on the success of their leniency programmes to detect possible cartels.[2] Experience has shown that the number of leniency applications can in general fluctuate significantly (between time and jurisdictions) and applications may be more common in certain industries or for certain types of cartel conduct. An over-dependency on leniency programmes as the de facto sole detection tool can therefore result in a less diversified and deterrent enforcement or less detection overall.

Second, leniency programs rely on effective *ex-officio* enforcement to create an incentive for leniency applicants to come forward, as the risk of detection is a key factor in their decision to cooperate. By increasing *ex-officio* detection probabilities, screening tools indirectly reinforce leniency programs and improve the overall effectiveness of cartel enforcement.

Traditionally, cartels have formed and collusion consummated through private communications and secret agreements because of their illegality. Consequently, to our knowledge, all proposed cartel screens assume a hidden agreement and aim to identify the existence of such a secret agreement, including its potential start and end date, indirectly, through data analysis. However, recent research highlights the possibility that collusion occurs through public communication. Harrington (2022) suggests that firms can use public announcements, such as earnings calls, trade press interviews, annual reports, media appearances, and other public forums, to signal their willingness to coordinate strategies, such as raising prices or limiting production, without explicit agreements, potentially evading regulatory scrutiny.

Public text documents with corporate communication are a largely unexploited source of screening for indicia of collusion. This is surprising, as public communication is a relevant dimension to collusion for at least two reasons. First, public communication may act to *complement* an underlying secret agreement. For example, executives publicly stating in the press their intent to reduce capacities

[1] An *ex-officio* investigation refers to a procedure initiated and carried out by the competition authority on its own accord (including based on information provided by informants), without reliance on complaints or leniency applications.

[2] Leniency programs offer companies or individuals participating in a cartel reduced penalties or immunity from prosecution in exchange for cooperating with the investigation. These programs have been highly effective in triggering cartel investigations in all jurisdictions where such policy tools are available. Between 2015 and 2021, the number of leniency applications experienced a decline worldwide, consistently across all regions (Organisation for Economic Co-operation and Development (2023), Figure 2.8).

can act as commitment device that lends credence to a pre-existing secret agreement with competitors to curb capacity.

An example for public communication potentially complementing secret coordination is the US Generic Pharma case (Cuddy et al., 2024), where more than 20 drug manufacturers are being privately sued for allegedly colluding on the prices of generic pharmaceuticals. The complaint claims there were private communications among the defendants which resulted in coordinated price increases and adoption of a market allocation scheme. At the same time, Harrington (2022) documents a series of explicit public announcements made by one of the defendants, Lannett Company, during the alleged collusion period. Its CEO publicly called for less price competition.[3] Such statements can complement an existing, secret collusive scheme, as is suspected to have been the case here.

Second, public communication can act as a *substitute* for secret private communication to collude.[4] By communicating publicly, executives can convey messages similar to those that would be exchanged secretly in a smoke-filled room, but with the added benefit of easier deniability, as they appear to be addressing a broader audience than rivals.[5] Cartel cases where public communications constitute the primary source of evidence of an infringement are rare.[6]

This paper builds a method to screen public announcements for communication among firms that can be inductive of collusion. We propose a novel approach using natural language processing (NLP) techniques to analyze the textual content of earnings calls. Earnings calls, as a primary channel of communication with investors and analysts, also serve as a platform for indirect communication with competitors. By systematically examining the language used in these calls, our aim is to identify patterns and signals that may indicate attempts to coordinate behavior for the purpose of reducing competition.[7]

Our analysis draws on a comprehensive dataset of approximately $367,000$ earnings call transcripts from around $15,000$ publicly traded companies worldwide, spanning the period from 2004 to 2022.

[3] The earnings calls of Lannett Company, held on September 10, 2013, September 8, 2004, and February 4, 2015, included the following statements: "*We're not a price follower. We tend to be a price leader [...] We have more price increases planned for this year within our budget. And hopefully, our competitors will follow suit.*" "*We took a chance that our competitor, Impax, would raise theirs [prices]. [...] no one created a price war.*" "*I think you're going to find [...] less competition, in a sense. You won't have price wars.[...] I just don't see the prices eroding like they did in the past.*"

[4] The limited empirical evidence suggests that public communication substitutes for private communication. Bourveau et al. (2020) find that increased cartel enforcement in the US led firms to shift from secret collusion to public communication channels. Similarly, Kepler (2021) shows that competitors with direct private communication channels rely less on public disclosures for coordination, indicating that public and private communication channels can act as substitutes rather than complements.

[5] Rosenblatt and Nilsson (2012) and Rabinovici (2017) provide an overview of the evolving antitrust law surrounding the public exchange of information in European competition law, while Harrington (2022) reviews the antitrust treatment of this issue in the United States.

[6] An example is the Dutch telecom case, in which the mobile network operators KPN, T-Mobile, and Vodafone were accused of having used public announcements in the trade press and conferences to communicate their intention to increase rates. The case ended in a commitment, with the telecom operators agreeing to *"not make any oral or written announcements in the public domain about future prices and other commercial conditions for mobile-communication services in the Dutch market that would leave consumers worse off"* (Autoriteit Consument en Markt, 2014). There are also some private cases pursued in the United States; see Harrington (2022).

[7] For the record, we do not accuse any company mentioned in this article of being part of a collusive scheme.

These transcripts, sourced from the S&P Capital IQ database, capture structured discussions of financial results, operational performance, and strategic priorities, offering a rich basis for applying NLP to detect communication patterns that are consistent with collusion.[8] We employ a dictionary-based approach, focusing on unigrams (single words) and bigrams (two words) potentially indicative of collusive practices, selected through a review of existing academic literature, antitrust guidelines, the framework proposed in Harrington (2022) and consultations with experts. We aggregate the selected terms into two main indicators: one measuring communication about strategies (STRAT) and the other about making references to competition, competitors, or the industry (COMP). Our findings show that firms frequently discuss strategy, but less often address competition. To measure communication that could facilitate collusion, we analyze both the focal firm's communication and its rivals' responses. We create a comprehensive measure of communication based on these criteria, allowing us to capture the dynamics of language that can be associated with potential collusion among firms in various markets.

The vast majority of cartel screens are based on data from previously detected cartels, with only a handful of them having directly informed actual enforcement actions.[9] Our approach aims to provide competition authorities with a new screening tool for potentially collusive behavior. Its practical relevance is demonstrated by the fact that the results of our early analysis informed the European Commission before it decided to conduct inspections at replacement tyre manufacturers in 2024.[10] The Commission suspected that price coordination may have occurred among the inspected companies, potentially facilitated by public communications (European Commission, n.d.). Our early analysis highlighted suspicious patterns in public statements, contributing to the Commission's interest in reviewing tire manufacturer communication in earnings calls.[11]

In Section III, we describe how public communication can and has been used to engage in unlawful collusion which then provides the theoretical underpinning to our screening method which is described in Section IV. In Section V we bring our method to the data. Section VI summarizes the data requirements for our analysis and examines the applicability of our method when competition authorities have access to alternative data sources. Section VII concludes.

[8] Transcripts of earnings calls are widely available through several provides, such as Seeking Alpha, MarketBeat, investing.com, The Motley Fool and possibly others.

[9] Section II provides examples of such successful screens.

[10] In July 2025, the European General Court held in a judgment that the European Commission's screening method provided sufficient indications of collusion to justify conducting inspections in response to those suspicions (*Compagnie générale des établissements Michelin v European Commission*, n.d.). Hirst and Comte (2025) reported on the court hearing in the case that *"at the heart of the investigation is the novel claim that the tiremakers may have used 'public communications,' in particular earnings calls, as part of a strategy to collude over prices,"* adding also that *"the Commission had run 'bi-gram' analyses to scan hundreds of thousands of earnings calls, eventually reducing their focus to tier-1 manufacturers in the tire sector."*

[11] This inspection was a landmark, as it marked the first time the European Commission launched an investigation solely based on its screening of public data. We are not discussing the tyre sector further in this article as it is part of the ongoing European Commission investigation.

II Background on Cartel Screening

The effective enforcement of competition laws against collusion is a multi-stage process. First, a suspected cartel must be found. Second, it must be prosecuted and convicted. And, third, it must be punished. Cartel screening is the use of market data to assist with that first stage. Its primary aim is to narrow down markets or firms that can be reasonably regarded as suspicious from a collusion perspective, and where further investigation may be justified. It can provide the facts to rationalize investing agency resources and aid in meeting the legal standard to obtain documents and data from the suspected companies whether through a request or an unannounced inspection. In practical terms, competition authorities are typically legally required to reach at least a certain level of *reasonable suspicion* in order to deploy formal investigative measures.[12] The evidence developed through cartel screening can assist with this crucial investigatory step.

There are two approaches to cartel screening: structural and behavioral (Harrington, 2008). Structural screens identify markets with characteristics conducive to collusion. These characteristics include high market concentration, entry barriers, homogeneous products, price transparency, and stable costs and demand, all of which are widely considered to facilitate cartel profitability or stability (Ivaldi et al., 2003; Motta, 2004). However, it is generally recognized that structural screening is not very effective because it is difficult to predict cartel formation. We know of factors that contribute to making collusion more effective, such as those mentioned above, but it takes something more for companies' managers to decide to engage in illegal activity. The omitted or poorly measured variables affecting the cartel formation decision continues to elude us which prevents structural screening from being a prudent use of resources by a competition agency.

Behavioral screens instead focus on firm conduct and market outcomes to look for evidence of collusion. Rather than trying to predict where a cartel *will* form - which is the approach of structural screening - behavioral screening is designed to identify where a cartel *has* formed. Behavioral screens look at whether patterns in prices, bids, market shares, and other variables are contrary to competition or are distinctly more consistent with collusion than competition. It asks whether there is an abrupt change in such variables in the absence of any comparably abrupt change in demand or cost, and whether the direction of that change is consistent with collusion. Examples of collusive markers include a simultaneous increase in the level of firms' prices and a decrease in the variability of firm's prices (as is easily measured by a fall in the coefficient of variation of price) (Abrantes-Metz et al., 2006) or patterns in bids at procurement auctions consistent with a cartel using cover bidding (Chassang et al., 2022). While behavioral screening can be performed using rich and extensive data along with sophisticated econometric models, a critical element of its appeal is that it can be effective with easily available

[12] See Article 21(1) of European Union (2003) for the EU. The US legal system also knows the concept of "*reasonable suspicion*," but to conduct a dawn raid the Department of Justice may obtain a search warrant from a magistrate judge after showing "*probable cause*" for a search. While "*probable cause*" may require a higher standard of proof than "*reasonable suspicion*," the dividing lines are not clear cut.

data - such as prices and bids - and using simple summary statistics or regression models. Recent advancements using machine learning has further enhanced the effectiveness of behavioral screening. Leveraging data generated both when firms are colluding and when they are competing, supervised learning algorithms have enhanced predictive power (Huber and Imhof, 2019). Harrington and Imhof (2022) offers a review of many behavioral screening methods.

Due to its ease and effectiveness, behavioral screening has become a critical tool for dozens of competition agencies. Some agencies have been engaging in it for 10-15 years including those of Brazil, Sweden, and The Netherlands. Brazil's CADE developed its screening program - Project Cérebro - around 2014 and subsequently discovered a retail gasoline cartel. A few examples among many include the South Korea Fair Trade Commission which found collusion in tenders for subway construction, the South African Competition Commission which uncovered a cement cartel, and Italy's AGCM which found a bidding ring in tenders for school cleaning services (Giannino, 2016). Recently, the Danish Competition and Consumer Authority publicly announced its use of behavioral screening with a white paper (Danish Competition and Consumer Authority, 2022) and the National Commission on Markets and Competition in Spain has developed a bid-rigging detection system which is being applied to data obtained from public procurement information systems. There is no shortage of examples of competition agencies actively engaging in screening and doing so with success.

Despite its potential relevance, the role of public communication in facilitating collusion and how it can be screened for collusion is largely unexplored. Our approach fits into the growing strand of research using *text as data* (Gentzkow et al., 2019).

The novel perspective that our approach adds to behavioral screening is to apply it to a new data source: public communications. Just as much as firms' conduct through the prices they set and the bids they submit can inform us of latent coordination, firms' conduct through their public announcements can inform us of their intent to coordinate and their efforts to maintain that coordination once achieved. Supracompetitive prices and bids are the product of a coordinated effort to restrict competition but that coordinated conduct can only be achieved through communication. We have developed a screening method to detect such communication in companies' public statements. This screening method looks for anomalous communication patterns in corporate statements, messages that can serve as covert invitations to competitors to coordinate and restrict competition, messages that implicitly convey acceptance of that invitation to form an agreement, and concerted practices that fall short of an agreement but still lead to supracompetitive outcomes.

III How Public Communication Facilitates Collusion and is a Violation of Competition Law

Companies routinely make public announcements in order to inform shareholders, capital markets, customers, and suppliers. It is a company's fiduciary responsibility to properly report its current and

future financial status. A company may want to publicly forecast sales growth to inform investors but also input suppliers as to likely future purchases. Customers can benefit from a company announcing future prices, product offerings, and possible supply disruptions. This information can be conveyed through a variety of media including press releases, earnings calls, financial statements, annual reports, industry meetings, and interviews.

The sharing of information with investors, suppliers, and customers is generally viewed as conducive to well-functioning product and capital markets. It is also true that any public information is likely to end up being shared with rival firms for it is their business to know their competitors. A firm then faces a tension with its public announcements because a better informed competitor is often a tougher competitor. While there are exceptions, the default policy for many companies is to keep information from competitors to the best of its ability because, more often than not, a competitor's gain is a firm's loss.

There is one notable exception which is when public announcements are for the purpose of coordinating conduct to restrain competition. In that case, competitors are not an ancillary audience but rather the intended audience and, if successful, this information exchange can benefit all firms. There are two challenges to such communications being effective. First, public announcements can create uncertainty as to who is the intended audience. For example, if a firm announces a future price increase, is it intended to inform customers or instead to signal to competitors to coordinate on such a price increase?[13] A firm may intend the latter but its competitors infer the former and, consequently, the attempt to coordinate fails. Second, given a firm would not be so daft as to announce the precise details of a collusive scheme, the messages will tend to convey only general properties of a simple collusive scheme. While this may limit the likelihood of success and the extent of collusion, it can still lead to substantively restraining competition. There are many simple collusive outcomes upon which firms can coordinate through public announcements including a common price increase, one firm acting as a price leader, adopting a common surcharge, eliminating discounts, and not competing for each other's customers ("no poaching") or market share.

Harrington (2022) identifies three classes of public announcement that carry a particularly high risk of surmounting these challenges and facilitating collusion. The first class has a firm announce how it will behave in response to the conduct of competitors. For example, the Canadian dollar store retail chain Dollarama Inc. made the following statement regarding its pricing strategy on the earnings call of December 4, 2024:[14]

> *"[A]s we've also said many times, we're a price follower, not a price leader. So we will watch what the market does. And if the market absorbs inflation or weaker exchange rates or tariffs then that is what Dollarama will do too."*

[13] Who was the intended audience – customers or competitors – was a key point of contention in the advance price announcements case involving U.S. airlines. See Borenstein (1994).

[14] This and all following excerpts were flagged by the text screening program which we developed.

Such a public statement facilitates the creation of a price leader-follower arrangement.[15] If this announcement created a common understanding among firms that Dollarama (which had the largest market share) would follow their prices as they go up and down, it would ultimately cause prices to be higher as it incentivizes competitors to increase prices and disincentivizes them to decrease prices.[16]

The second class comprises public announcements stating how rival firms, or the industry at large, *should* behave. It might entail a firm publicly communicating that all firms should price higher, produce less, or limit capacity investment. For example, an executive of Czech beverage maker Kofola CeskoSlovensko explained in an earnings call on the 2 September 2022:

> *"But we are working very intensively on the price increase from the beginning of next year. I mean, like, all the competitors should do the same basically because the energy prices are the same for all. So that's the current task we have, right, to increase properly the prices for the beginning of next year."*

This statement is the basis for a mutual understanding emerging among firms that prices are to rise. This class of announcements is the most egregious for there is no procompetitive rationale for a firm to prescribe such conduct to its competitors. Furthermore, these announcements have a high chance of success because, as explained in Harrington (2022), they only have value for competitors so it is clear they are the intended audience.[17]

The third and final class encompasses public announcements stating how rival firms, or the industry at large, *will* behave. For example, an executive of insulation maker Rockwool A/S explained the following in their earnings call on 10 February 2022:

> *"**Analyst**: Okay. Understood. And how about the competition within stone wool market? Is there any of them not behaving the way they would have done it with these cost increases?*
>
> ***Executive**: We do what we do. And that we have done every year the last 6, 7 years. This is a different level and there will be some people waiting. There were some people not doing it. There will be people doing it. Our experience is that everyone would see – feel these costs, there's nowhere to hide. So I would expect that price increases – and we have seen that in many markets will need to follow for everyone producing stone wool [...]"*

Stating "I would expect ... price increases" could prove to be a self-fulfilling prediction when it causes

[15] The statement was made in response to an analyst asking "*what are some of the key margin levers that you can pull to offset some of these potentially higher cost pressure for next year.*"

[16] Several compelling cases are provided in Harrington (2022) involving this type of public announcements. Of particular note is the previously mentioned Dutch mobile telecom case where the medium was an interview in a trade publication by one of the company's executives.

[17] Several compelling cases using this type of public announcements are provided in Harrington (2022). Of particular note are those in the airlines and steel markets, both of which involved coordinating on limiting capacities. Messages were conveyed through earnings calls and industry meetings.

all firms to coordinate on those "predicted" price increases.[18] In theory such statements may also confirm in public the monitoring of ongoing collusion.

Some public statements combine multiple elements. For example, the earnings call of Lufthansa of 4 March 2021 left little doubt about the intent of the company to avoid price competition.

> *"[W]e just want to avoid the price war out there. So I think some capacity discipline on behalf of the leading carriers like us will help us all to create a healthy industry. And I think we're all sending the right signals this way."*

The executive expressed his desire for the competitors to limit capacity and confirmed that the competitors were signaling in a manner consistent with such a plan.

Thus far, we have described how certain types of public announcements can facilitate collusion. Sections IV and V deliver the main contribution of this paper which is providing a systematic and efficient approach to screening sources of public communication for such announcements. But before doing so, we need to explain why these communications are associated with *unlawful* collusion. More to the point, if a competition authority pursues a case based upon detecting such communications then it has a reasonable prospect of obtaining a conviction or a commitment.

One scenario is when a cartel uses public announcements to supplement private communications. Upon having coordinated on a collusive scheme by directly communicating at private meetings or through other private channels, a cartel may use public announcements to maintain it by, for example, expressing continued adherence to the agreed-upon outcome. As every private meeting means more evidence should the cartel be discovered, public announcements are a less risky means to convey certain content. By detecting suspicious public communications and then engaging in an investigation, a competition authority may ultimately find evidence of private communications which is generally sufficient to obtain a conviction.[19]

The more challenging scenario to prosecute is when firms communicated exclusively using public announcements. Could a competition authority win a case with only public communications? There is certainly content for which the answer is indisputably "yes". For example, suppose firms were to have press releases that contain the precise content typically conveyed through private communications such as the prices that all firms are to charge, when they are to charge them, and each firm's allocated market share. Clearly, the law does not provide a loophole in the form of public statements. Hence, all firms issuing such press releases is certainly sufficient evidence to prove they have an unlawful agreement. Of course, as already noted, inevitably firms will obscure the content of their communications when

[18] At present, there are no documented cases of these announcements being used to facilitate collusion. This could be because no such episodes exist or episodes exist and are difficult to identify. Consistent with both explanations, these announcements are less likely to be effective because forecasts about future conduct would also be of interest to capital markets. That there is a credible alternative audience makes it difficult - for the observer but also competitors - to infer a firm's intent when, for example, it predicts the end to a price war or industry capacity will not rise.

[19] As described in Harrington (2022), public communications supplemented the private communications in the generic drugs cartel and the (suspected) broiler chicken and pork cartels.

they are public, which raises the question: Are public announcements that could plausibly facilitate collusion also sufficient for courts to conclude there is a violation of competition law? While the answer may vary across jurisdictions – for it can depend on the specifics of the law, jurisprudence (particularly pertaining to information sharing), and the general receptivity of courts to expanding the boundaries of acceptable evidence – let us explain why we believe the answer is yes.

An unlawful agreement exists when a firm invites its competitors to constrain competition and those competitors accept this invitation, either explicitly or implicitly. To determine whether the content of the announcements is consistent with there being an agreement, one can pose the following question. Based on jurisprudence, would the firms be in violation of the law if the same content had been conveyed privately? If the answer is "*yes*" then the content is supportive of a by object or per se offense.

With private communications, courts in some jurisdictions have ruled that acceptance can be inferred from the lack of explicit refusal to participate; in other words, silence is acquiescence. While it is generally a firm's choice to communicate privately with competitors, it is often not a choice to listen to public information. Consequently, evidence of an agreement between firms based on public announcements will require something more than a passive act. The most compelling evidence will be when all firms publicly express their support to restrain competition. For example, all firms announce: "*The industry should reduce capacities.*"

In place of some firms' announcements could be conduct that is consistent with the acceptance of another firm's recommendation. For example, one firm announces: "*The industry should reduce capacities*"; and all firms subsequently reduce capacities in a manner that is not in their individual interests but for a coordinated plan of capacity reduction.

In some cases, public communications may not rise to the level of an agreement because not all firms have communicated. However, depending on the jurisdiction, there may still be a violation of competition law. In the European Union, conduct such as some information exchanges between competitors and advance price announcements, even if they should fall short of an explicit agreement, may still violate Article 101 of the Treaty of the EU.[20] Public announcements of the type described here could fall into that category. For example, a firm announcing it will follow rival firms in terms of their pricing or stating that the current price war should end can lead to coordinated conduct even though the evidentiary requirements of agreement are not satisfied.

Another avenue applies to jurisdictions where an invitation to collude is illegal. One firm stating that the industry should raise prices or limit supplies or engage in capacity discipline is an invitation to collude though the lack of supporting public announcements from competitors may prevent proving there is an agreement (and, indeed, there might not even be an agreement). Recently, Greece has modified its competition law to make invitations to collude illegal (Lianos and Wagner-von Papp, 2022). This approach is compelling for reasons of both legal consistency and deterrence; if an act is unlawful then an attempt to commit that act should be as well.

[20] See the chapter on information exchanges in the Horizontal Guidelines of the European Commission, (European Commission, Directorate-General for Competition, n.d.).

While a proper treatment of this issue requires its own separate study, we hope to have convinced the reader that public communications can be the basis for proving a violation of competition law or obtaining a commitment to prevent future public communications that facilitate collusion. Consequently, detecting such public communications is a worthwhile endeavor for a competition authority because it can screen markets for cases to pursue which ultimately might lead to a successful outcome.

IV Screening for Collusion-Facilitating Communication in Public Announcements

Although all forms of collusion are harmful, not all are unlawful. Only certain processes that result in a collusive outcome are unlawful. To provide evidence for unlawful collusion in practice, competition authorities therefore focus on observable communication that facilitates collusion rather than a collusive outcome in itself.

With the screening tool we provide in this paper, we aim to improve detecting unlawful collusion through *public* communication and thereby facilitate informed action against it. To this end, we present a user's manual for a novel behavioral screening tool designed to identify collusion-facilitating communication in any form of public announcements. We follow Harrington (2022), defining a public announcement as the *"conveyance of information by a firm or one of its employees using a medium that is widely accessible to individuals outside the firm."* Examples of public announcements include earnings calls, press releases, financial statements, annual reports, industry meetings, and interviews.

Our proposed screening method enables competition authorities to efficiently screen large volumes of public announcements. While the approach we suggest can flag suspicious cases, it remains the responsibility of competition authorities to manually assess flagged communications to determine whether they may constitute actual violations of competition law.

We first outline the development of a general tool, followed by an illustrative application to a dataset of around 370,000 earnings calls from 2004 to 2022. Then, we provide an example on how flagged cases of communication can be taken to a detailed, qualitative assessment by competition authorities.

A Screening Tool Development

The screening approach suggested here has the goal of flagging collusion-facilitating communication in public announcements. As described in detail in Section III, Harrington (2022) describes three main classes of public announcements that carry a high risk of facilitating collusion: a firm's statement that its future behavior will depend on how a rival behaves, a firm's suggestion or advice on how competitors or the industry should act going forward, and a firm's prediction about the future behavior of rivals or the industry as a whole. These three classes of collusion-facilitating public announcements typically feature strategic variables (e.g. "price", "cost", "output", "capacity", "margin") and reference competitors, either explicitly ("rival", "competitor") or implicitly ("industry", "peer"). Mentions of

strategic variables and competitors appear in combination with intention and action verbs, as firms announce plans (e.g. "intend", "plan", "will", "expect") or signal conditional responses (e.g. "match", "follow", "adjust"), potentially based on competitors' behavior. Further, words related to market conditions may be mentioned as firms use them to signal coordination indirectly through market assessments (e.g. "discipline", "rational", "inelastic", "oversupply").

Consequently, to identify such collusion-facilitating communication in public announcements, we suggest a dictionary-based screening approach. As collusion-facilitating communication is more likely to be present when multiple competing firms are engaging in it, we find it useful to consider both the public statements of a firm as well as those of its competitors. Our method offers a quantitative measure of a firm's own communication, which can be combined with similar metrics from its competitors. Additionally, we propose an index that aggregates these measures across groups of competitors into a single measure, which is particularly effective in highlighting groups of firms that use language indicative of potential collusion.

B The Dictionary

We first compile a list of unigrams that help identify collusion-facilitating communication for two categories: variables referring to strategy and variables referring to rivals or competition. We discussed these unigrams with a group of officials at DG Competition's Cartels Directorate, with extensive experience in cartel enforcement. The list is then enriched with unigrams mentioned in the empirical literature on collusion and public communication (Aryal et al., 2022; Harrington, 2022; Sheng and Vukina, 2024). In the second step, we combine unigrams related to strategy and competition with those expressing intent, calling for action, or describing the broader market context. This yields two sets of bigrams: one for strategy and another for competition. In the third step, we refine the lists by manually adding informative bigrams and removing those that are uninformative. The final result is a dictionary of bigrams, divided into two categories, strategy (J_{strat}) and competition (J_{comp}) that indicate communication that may facilitate collusion.

Step 1: Define unigrams for strategy and competition

In the first step towards constructing the dictionary, we build the following lists of unigrams that reference strategic variables, which may be the object of coordination, and references to rivals or competition.

1. **Strategy**: Related unigrams in our list include, among others, "price", "margin", "share", "profit", and "capacity". Broadly speaking, this group of terms ensures that we focus on statements that relate to key strategic choices, such as changes in price, capacity, and market share, or the need to react to cost or demand shocks.

2. **Competition**: Our list includes unigrams such as "competition", "competitor", "industry", and "peer", among others. This category captures instances where a company mentions its rivals,

identifies as part of a group, or collaborates with others. Such references are relevant because collusion requires some level of coordination with competitors, rather than merely a series of unilateral statements by a single firm.[21]

Step 2: Create bigram dictionaries by combining unigrams

Research on text analysis (Tan et al., 2002) highlights that measurement quality tends to improve by applying combinations of multiple words (so called *n-grams*) instead of searching for unigrams. We therefore combine each of the unigrams for strategy and competition variables with the following unigrams for intent, action, and market environment.

3. Expression of intent for the firm itself or regarding competitors. This group of keywords includes, among others, "should", "will", and "expect". It aims to capture instances of announcements related to how rivals *should* behave or are *expected* to behave. They also capture instances of a firm announcing how it *will* behave in response to rivals.

4. Calls for a certain type of action. Keywords in this group include, among others, "increase", "raise", "curb", "cut", "reduce". For example, announcements about price increases, price wars or capacity reductions, a firm's intent to follow a price leader, or increasing prices without losing market share would be captured by this group of words in our dictionary.

5. Unigrams related to the market situation. Keywords in this group include "excess", "oversupply", "rational", and "discipline". This group includes words that can describe a competitor, the industry, or a market situation in a way that allows the listening firm to understand that coordination among competing firms is ongoing or may be desired. For example, talking about "price discipline", "industry oversupply", "insane price cuts", or a "rational industry" are captured here.

This results in two lists of bigrams. The strategy dictionary J_{strat} combines strategy unigrams (1) with unigrams for intent (3), action (4), and market environment (5). This gives a list of bigrams that relate to potentially sensitive announcements about strategy. Examples include "price increase", "reduce capacity", "gain margin", and "rational pricing".[22] We would expect colluding firms to mention such bigrams in the process of coordinating on joint price increases or capacity reductions.

For a collusive scheme, we also expect competitors to explicitly call out to each other. For example, a firm would need to announce what it expects from competitors in terms of price actions, or that it was unsatisfied with the low prices or excess capacities in the industry. To capture this, the J_{comp} dictionary combines competition variables (2) with the same sets of unigrams (3–5). This results in a group of bigrams such as "competitor will", "price competition", or "industry should".

[21] This does not imply that a series of unilateral statements cannot be part of a collusive scheme. An example is firms' use of advance price announcements whereby a firm only refers to its own future prices. However, since the goal is to screen for and detect clearly collusive public communication, it is practical to explicitly focus on instances where companies refer to their rivals.

[22] We combine keys in both orderings. For example, we have "increase price" and "price increase" on our final list of bigrams.

Step 3: Manually adding more informative bigrams and removing uninformative bigrams

Finally, we review the resulting list of bigrams in Step 2 and add some bigrams that we find relevant. Examples include "win win" and "whole industry." Typically, these are two-word expressions that may be associated with a discussion that could be collusion-facilitating.

We also remove some bigrams from the lists resulting from Step 2. We do this where we believe a bigram is particularly likely to be used in a pro-competitive context. Examples include "capacity increase" or "reduce price". Clearly, there are examples where these bigrams could form part of a collusive discourse. For example, a statement along the lines "we have to avoid a further *capacity increase*" or "we should not *reduce the price*" could arise in a collusive context. The reason why we exclude them is mainly practical: these terms are most likely collusion-facilitating when they are negated. Taking into account negations in a meaningful way is a challenge. We therefore decided to omit these bigrams altogether.

Our final list of bigrams related to a potentially anticompetitive strategy contains 108 elements. The list of bigrams related to *competition* or *competitors* contains 503 elements. Table 7 provides further examples of keywords in each of the above categories.

C Identifying Competitors

Collusion through public communication is likely to include participation of multiple competing firms. For collusion-facilitating communication to take place, we expect that at least two rivals engage in such communication. Isolated statements by a single firm are less likely to form part of a collusive scheme than patterns of such communication between competitors. Therefore, it is important for the dictionary-based screening approach to identify both a firms' own collusion-facilitating communication as well as that of its' rivals. Competing firms are often identified using industry classifications such as NAICS or SIC codes. However, industries defined in this way do not necessarily correspond to markets in the antitrust sense, as firms within the same NAICS or SIC industry could be input suppliers, customers, or entirely unrelated to one another, as well as competitors. As a result, additional methods are typically employed to examine product market overlaps and geographic market presence in order to more accurately define competitive relationships (e.g. Affeldt et al., 2021).

A growing strand of empirical literature moves beyond these approaches and proposes using evidence from company disclosures, analyst reports, regulatory filings, patents, and online content to identify competitors (e.g. Werle and Laumer, 2022). These data-driven approaches provide a detailed view of competition by capturing direct and emerging rivals in a potentially more precise way and therefore offers a more granular and market-relevant view of competition than traditional classification methods. In Section V.B.2, we provide an illustrative example of how data-driven methods can be used to identify suppliers in the same market for a large dataset of earnings calls.

D Proposed Screening Tool: Index of collusion-facilitating public communication

We can now bring together the dictionary approach and competitor identification to flag collusion-facilitating communication in public announcements. In this context, communication is conceptualized as the interaction between different actors. Thus, we identify instances of potentially problematic communication by examining situations in which the following conditions are met.

1. The focal firm frequently discusses potentially anti-competitive strategies (STRAT is high),
2. makes frequent references to competitors (COMP is high),
3. has competitors who also exhibit similar communication patterns in the same quarter (i.e., STRAT or COMP are high for its competitors), and
4. displays this behavior consistently over multiple (consecutive) time periods.

We do not formalize a specific indicator in this context, as multiple approaches are possible. Defining what constitutes "frequent" or "high" behavior requires setting appropriate thresholds, which serve as tuning parameters. These thresholds can be calibrated by authorities based on the desired level of stringency and their tolerance for false positives and false negatives. A discussion of these errors is provided in Section IV.E.

Second, the way in which the amount of information along both the STRAT and COMP dimensions and for the focal firm and its competitors is aggregated into an indicator is also a modeling choice. For example, one could choose to assume that communication takes place when all indicators are high, both for the focal firm and its competitors. As another option, one could think that it is sufficient that one of the two indicators is high for either the focal firm or its competitors. Alternatively, one could consider the minimum and maximum of the two scores, respectively.

In the implementation of the method presented in Section V.B.3, we will make certain assumptions to illustrate these considerations.

E Potential Biases, Type I and Type II Errors

The suggested screening method for flagging collusion-facilitating public communication leaves room for two types of errors, as displayed in Table 1. First, *false positives* arise where we flag a firms' communication as collusion-facilitating despite it actually not being so. This would draw attention to public communication that is not relevant for competition authorities. Since the very aim of a quantitative screening exercise is to improve the efficiency of competition authorities' detection efforts, such Type I efforts are particularly harmful in our context. Second, *false negatives* occur when we fail to flag a firms' communication, although it may actually facilitate collusion.

There is an inherent tradeoff between Type I and Type II errors. Reducing one typically increases the other. Lowering Type I errors by making the flagging criteria stricter reduces false accusations

but increases the risk of missing relevant statements (higher Type II error). Conversely, relaxing the criteria to capture more truly suspicious statements decreases Type II error but increases the likelihood of flagging innocent firms. The optimal balance depends on the screening objective and the relative economic costs of each error type.

A competition authority can therefore calibrate the flagging threshold individually—such as by specifying a numerical cutoff for the screening index or by setting separate thresholds for STRAT and COMP in earnings calls deemed relevant—to achieve a desired balance between the two types of error. This allows the sensitivity of the flagging mechanism to be adjusted in line with the authority's screening priorities.[23]

Table 1: Type I and Type II Errors in Screening Approach

	Not Collusive	**Collusive**
Flagged	**Type I Error** (False Positive) Falsely flagged as collusive	True Positive Correctly flagged collusive firm
Not Flagged	True Negative Non-collusive firm	**Type II Error** (False Negative) Colluding firm not flagged

It is worth noting that in a screening exercise such as this one, the tolerance for Type I (false positives) and Type II (false negatives) errors may be—depending on the use case—highly asymmetric, with little tolerance for false positives and more tolerance for false negatives. For example, a competition authority may prioritize identifying only the strongest possible indications of collusion. In that case, Type II errors may be close to irrelevant, but Type I errors would count disproportionately: It did not matter whether the authority left some (even most) potential collusion indicia undetected, as long as it found (even a single) sufficiently strong indication that could result in successful enforcement action. The examples of public statements presented above testify to the usefulness of our measurement approach to uncover potentially relevant public communication that may be associated with collusion.

Measurement error is inherent in text analysis. Applying a similar text analysis methods to earnings calls for the purpose of quantifying corporate risk, Hassan et al. (2024) finds "*about half of the variation at the firm-quarter level is attributable to measurement error.*" For example, a mention of "*price increase*" may refer to raw materials, or past industry trends, rather than a signal about desired future conduct in the industry. Similarly, there is no difference between "*we will reduce capacity*" and "*we will not reduce capacity.*" This shortcoming applies to any dictionary-based text analysis screening, such as ours and (Aryal et al., 2022), and it is a difficult practical challenge to correct.

Our screening method is nevertheless relevant. Investigators at competition authorities have limited capacity to manually read earnings calls. They should focus attention to those documents that contain such references to language that can be associated with collusion. Manual review will reduce false

[23] For example, a competition authority may want to restrict attention to earnings calls that are in the top 25, 10 or even 1% of earnings calls in terms of STRAT or COMP, or in an index like the one we propose

positives (Type I error), but not false negatives (Type II errors). [24] At the same time, as mentioned above, in a screening exercise resource constrained competition authorities are likely to have larger tolerance for false negatives than false positives.

Having outlined a general procedure for flagging collusion-facilitating communication in public announcements, we now turn to an empirical illustration of the screening method. Specifically, we use earnings calls as a form of public announcements.

V Empirical Application of the Screening Tool to Public Announcements

In this section, we apply the developed screening tool using Natural Language Processing (NLP) techniques to a comprehensive dataset of earnings call transcripts from publicly listed companies worldwide to identify language patterns indicative of collusion-facilitating communication. This serves as an example of how competition authorities can use the screening tool on public announcements to detect such communication.

A Earnings Calls Data

Typically held quarterly, an earnings call is a structured conference call where the management of a publicly traded company addresses the financial community, including investors, analysts, and journalists. The purpose of these calls is to disclose financial results, explain operational performance, and outline strategic priorities. Additionally, earnings calls often include forecasts of costs, demand, and other critical factors that impact a firm's future performance.

The structure of an earnings call comprises three main parts. First, there is an introductory segment, which frequently includes disclaimers such as the safe harbor statement for forward-looking information. This is followed by a detailed presentation by senior executives on financial results and strategy. The call concludes with a Q&A session where investors, analysts, and journalists engage directly with management. While the primary purpose of earnings calls is to provide information to the financial community, it is important to note that these calls are open to the public, including competitors.

Our analysis is based on transcripts from approximately $370,000$ earnings calls conducted by nearly $15,000$ companies worldwide between 2004 and 2022, using data from the Capital IQ database. As shown in Table 2, the largest number of earnings call documents comes from firms headquartered in North America, representing $266,380$ transcripts from $8,596$ firms, followed by Europe and Asia. Companies headquartered in other regions, such as Oceania, South America, and Africa, make up a smaller, but still significant share of the data.

[24] False positives and negatives can arise in our analysis through means other than misidentifying text snippets as potentially suspicious. In particular, as discussed in Section V.B.2, we utilize data to identify companies that are likely competitors in relevant product markets, which may be subject to Type I and Type II errors. These will be addressed separately in that section.

Table 2: Number of Documents and Firms by Region

Region	Number of Documents	Number of Firms
Africa	1,901	171
Asia	27,192	2,004
Europe	56,044	2,904
North America	266,380	8,596
Oceania	7,868	759
South America	6,224	333
Unknown	1,328	93
Total	**366,937**	**14,860**

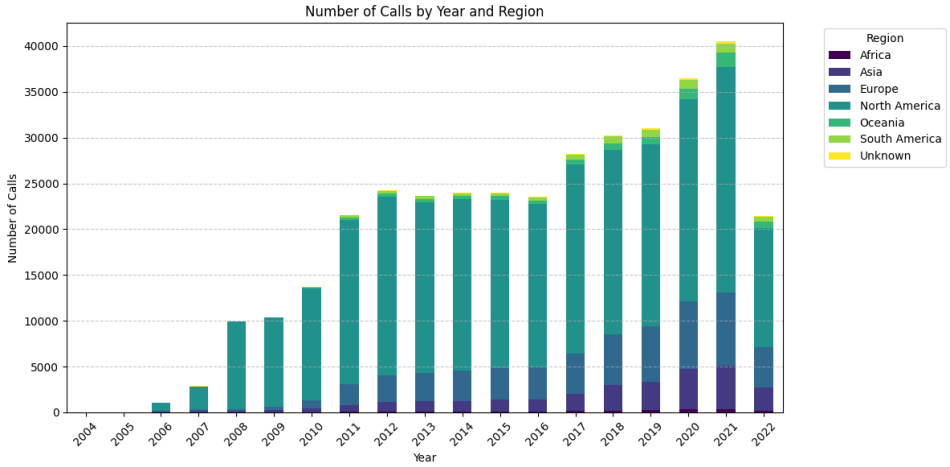

Figure 1: Number of earnings call transcripts by year and region

The number of available recording of earnings calls is shown in Figure 1. There is a clear, continuous increasing trend over time, with the number of earnings calls per year increasing fourfold between 2008 and 2021. The period between 2011 and 2022 includes about 85% of all calls. Until 2011, almost all earnings calls in the data set come from US companies. Typically, a firm has one earnings call per quarter, but there are exceptions where firms hold multiple conference calls close in time, even on the same day. For example, certain special events such as mergers may trigger special calls which depart from the normal quarterly reporting schedule.

We use the Natural Language Toolkit (NLTK) in Python on the earnings calls transcripts to build a data set that enables dictionary-based screening. This involves the following processing steps on the transcripts: converting all words to lowercase, removing stop words and punctuation, tokenizing the

text,[25] and discarding tokens or phrases with a length less than or equal to a single character.[26] The result of these processing steps is a document-level dataset of earnings calls with a *processed transcript*, which we use in the next step to screen for our dictionary of collusion-facilitating bigrams (see Section IV.B).

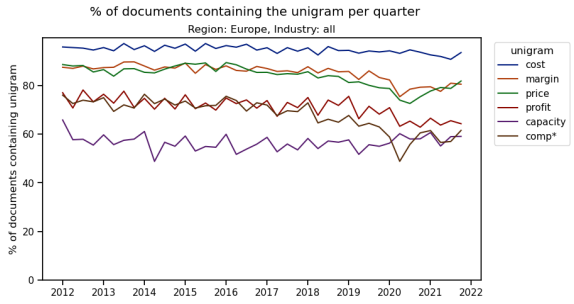

Figure 2: Share of documents mentioning unigram (Europe)

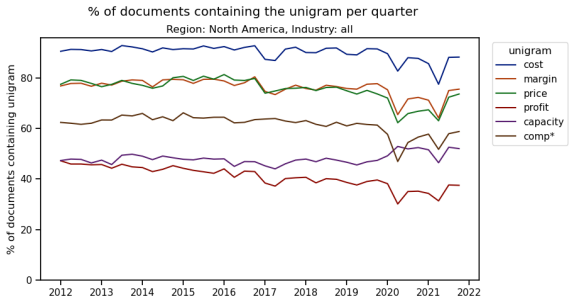

Figure 3: Share of documents mentioning unigram (North America)

We start our analysis by exploring how common is it for companies to talk publicly in a manner that may be relevant to a collusive discourse. We think of terms such as prices, capacities and margins. To understand this better, let us look at unigrams. As described in *step 1* of Section IV.B, Figures 2 and 3 show that talking about prices, profits, margins and capacity is a very common phenomenon in earnings calls. For example, over all industries in Europe and North America, the word "price" appears in some 90% of earnings calls, consistently over time. "Costs" are mentioned in nearly all earnings calls. Various (single-word) references to competition and competitors are made in more than half of earnings calls.

B Applying the Dictionary to Measure Collusion-Facilitating Communication

1 Measuring what firms talk about

Using the processed data set, we now turn to measuring collusion-facilitating language on earnings call document level. Based on the dictionary for strategy J_{strat} and competition J_{comp}, we create two indicators for each document, counting the number of occurrences of the bigrams. In particular, we calculate for each document $i \in I$ (the full list of documents), and bigram $b = 1...B_i$ of all bigrams contained in document i, the total number of occurrences of bigrams in each list $j \in (J_{strat}$ and $J_{comp})$, including duplicate occurrences:[27]

[25] Tokens are the basic unit of NLP. In our case, they correspond roughly to words.

[26] Hassan et al. (2024) provide an overview of pre-processing methods which we found useful in working with earnings call text data. Our implementation largely conformed to their suggestions with the main deviation that we remove stopwords.

[27] We take a simple additive approach to create this indicator, but other parametrizations are also possible. For example, one may decide to count duplicate occurrences of the same bigram as one, as is the case in Aryal et al. (2022), who use a dummy indicator equal to one if a transcript mentions capacity discipline, irrespective of the number of mentions in the document.

$$STRAT_i = \sum_{b}^{B_i} \mathbb{1}[b \in J_{strat}] \tag{1}$$

$$COMP_i = \sum_{b}^{B_i} \mathbb{1}[b \in J_{comp}], \tag{2}$$

where $\mathbb{1}[.]$ denotes the indicator function.

Hence, each earnings call document in our database is described by two variables, corresponding to the number of occurrences of keywords related to strategy (*STRAT*) and competition (*COMP*). Figures 4 and 5 provide an overview of the 30 most frequently encountered bigrams in both the strategy and competition groups, respectively.

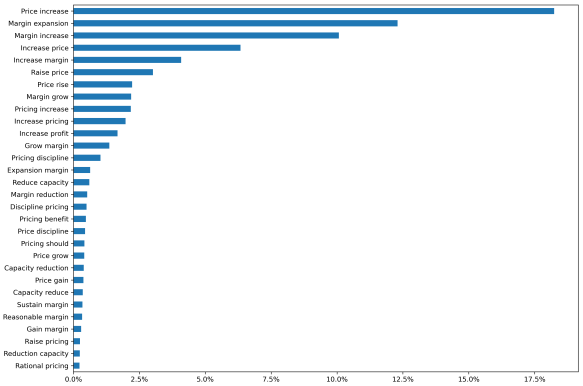

Figure 4: Share of documents mentioning top 30 strategy bigrams

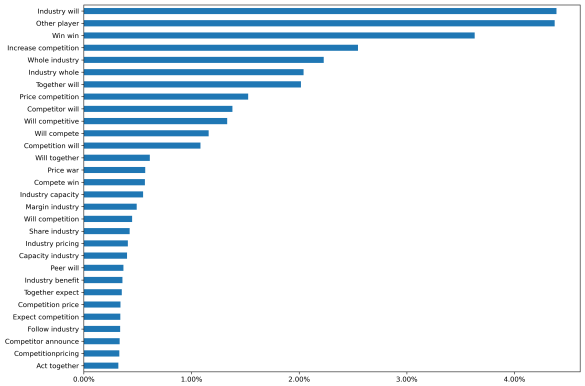

Figure 5: Share of documents mentioning top 30 competition bigrams

We aggregate STRAT and COMP on the firm-quarter level by summing these variables respectively in each given quarter. The intuition behind this is that our variables measure how often in a quarter a firm mentions bigrams related to strategy and competition.[28] We expect competitors to listen and pick up these signals, regardless of whether two mentions occur in a single earnings call or one mention takes place in each of two earnings calls.

Table 3 provides descriptive statistics for STRAT and COMP. It is worth noting that the median and mode of both variables are zero. In other words, more than half of the earnings calls in our sample contain no reference to any of the expressions included in our list of bigrams.

Figure 6 provides a detailed breakdown of these indicators, as well as the product of the two indicators, over time.[29] These figures confirm that only a minority of earnings calls had at least one

Our rationale is that more mentions of our bigrams, even if they are duplicative, provide a stronger indication of the text containing relevant language for flagging potential collusion.

[28] For example, if a company holds two earnings calls in a quarter, each containing one of our trigger bigrams related to *strategy*, it would have STRAT = 2 for that quarter. Similarly, if the same company had a single earnings call with two mentions of our trigger bigrams, it would also have STRAT = 2 for that quarter.

[29] We present the distribution of earnings calls over time, starting from 2012, as earnings calls were significantly less frequent

Table 3: Summary statistics for COMP and STRAT

Statistic	COMP	STRAT
Count	366,937	366,937
Mean	0.58	1.45
Std Dev	1.08	2.93
Min	0	0
25%	0	0
Mode	0	0
75%	1	2
Max	24	61

mention of the bigrams included on our lists. The only exception is the last quarter of our data, 2022Q1, where the number of documents mentioning STRAT reached 50%.

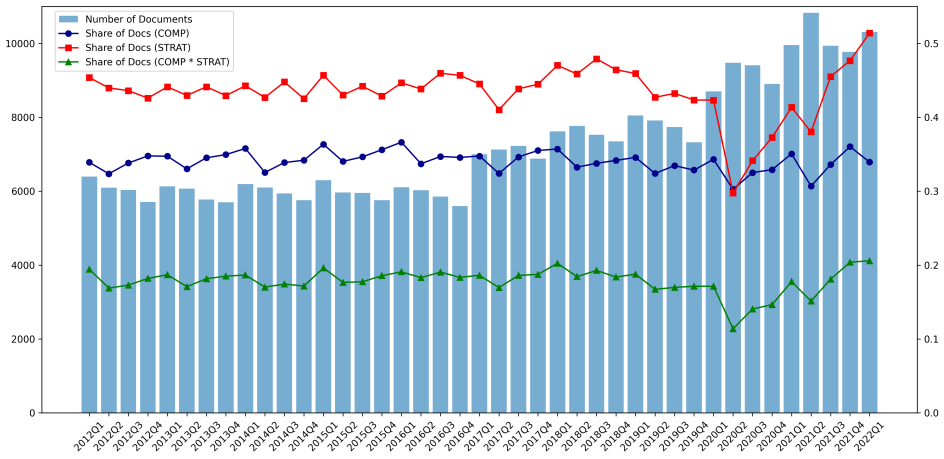

Figure 6: Number of Documents and Share Mentioning Competition and Strategy per Quarter by Region

Figure 6 also reveals a very visible drop in strategy-related talk on earnings calls in the second quarter of 2020. With the arrival of COVID, we would expect discussion in earnings calls to have shifted to topics related to how firms can cope with lockdowns and supply disruptions. Following the initial drop at the onslaught of COVID, STRAT gradually returned to the pre-COVID level, and the share of earnings calls mentioning strategy-related bigrams even surpassed 50%. We did not observe similar COVID-related development for COMP.

Figure 6 further shows the share of earnings calls that contained at least one reference to competition, as well as one related to a potentially sensitive strategy (STRAT and COMP strictly positive). We consider these earnings calls the most relevant from the perspective of screening for potential collusion,

before that, as shown in Figure 1. The product of the two indicators is a proxy of the share of documents containing both STRAT and COMP bigrams, allowing more occurrences of each.

because they are the most likely to be explicit about prices or capacities (STRAT>0), but talk about them in a manner that goes beyond purely unilateral statements and referring to their competitive environment in some way (COMP>0). In the full sample, around 20% of the earnings calls consistently contain references to both the potentially sensitive strategy and the competitive environment.

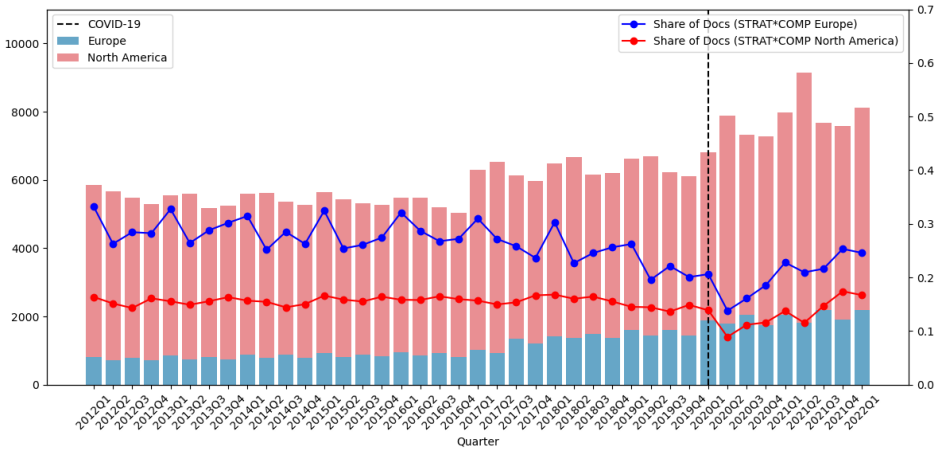

Figure 7: Number of Documents and Share Mentioning Strategy and Competition per Quarter by Region

Figure 7 shows the share of earnings calls that mention sensitive strategy and competitive terms over time for companies with headquarters in Europe and in North America. A surprisingly regular trend emerges, in which European companies appear significantly more likely to mention both of these issues than their North American peers.[30] European firms have a roughly 10% higher share of earnings calls containing bigrams related to strategy or competition than those headquartered in North America. This is somewhat surprising, given that financial transparency and reporting obligations appear to be more extensive for firms listed in the U.S., (Areno et al., 2020). In particular, the U.S. seems to have more detailed requirements for disclosing forward-looking information, including risks and uncertainties for financial performance. It is possible that U.S. firms are more careful in the language used in earnings calls than their European counterparts.

2 *Identifying Competitors*

Given that collusion requires the participation of at least two firms, it is natural to look for situations with at least two competing firms exchanging public messages.[31]

[30] t-tests comparing the shares of documents mentioning STRAT, COMP, and their product between Europe and North America all reveal statistically significant differences at the 5% level.

[31] Clearly, public communication may form *part* of a collusive scheme with only one firm resorting to public announcements, and even only on a single occasion. For example, with a secret price-fixing agreement in place, an executive publicly announcing that the company's pricing strategy remains unchanged may provide competitors with the necessary confirmation that the firm will stick to the agreement. This unilateral announcement may therefore *stabilize* a cartel.

We obtain detailed data from the S&P Capital IQ database on the competitive relationships between firms. In particular, under the earnings call functionality, the database has a proprietary feature that suggests a set of competitors to each focal firm. Capital IQ collects information from public and proprietary documents, such as annual and analyst reports and filings, about companies being mentioned as competitors. Although the precise method for determining competitors is not disclosed, upon inspecting several known companies, we find the suggested list of competitors to be intuitive.

Table 4: Competitor Descriptives

Statistic	Number of Competitors
Mean	20.23
Std. Dev.	21.92
Min	1.00
25th Pctl	10.00
Median	12.00
75th Pctl	22.00
Max	637.00

As displayed in Table 4, the median number of competitors per firm is 12, while the average is 20, which we consider reasonable and realistic.

	Competitors of Wizz Air
1	Air France-KLM SA
2	Asiana Airlines, Inc.
3	Deutsche Lufthansa AG
4	easyJet plc
5	Finnair Oyj
6	FirstGroup plc
7	International Consolidated Airlines Group
8	Jet2 plc
9	Ryanair Holdings plc
10	Spirit Aviation Holdings, Inc.
11	Stagecoach Group Limited
12	Thai Airways
13	The Go-Ahead Group Limited

Table 5: List of Competitors of *Wizz Air*

	Competitors of Piaggio
1	BMW
2	BRP Inc.
3	Eicher Motors Limited
4	Ferrari N.V.
5	Harley-Davidson, Inc.
6	Hero MotoCorp Limited
7	Loncin Motor Co., Ltd.
8	Ninebot Limited
9	Bajaj Auto Limited
10	PIERER Mobility AG
11	Polaris Inc.
12	Sanyang Motor Co., Ltd.
13	Yamaha Motor Co., Ltd.

Table 6: List of Competitors of *Piaggio & C. SpA*

Digital firms make up the top end in terms of competitor count, with Amazon (637 competitors), Google (444), Microsoft (438), IBM (393) and Apple (332) leading. A likely reason for this is their wide reach and multi-product nature. For example, Amazon operates as a marketplace, a seller competing

with many other sellers, and a cloud service provider.[32] Similarly, Google and Microsoft are active in the cloud service industry, beyond their core businesses in internet search and productivity apps.

We label two firms as competitors if they are mentioned as such in the competitor recommendation functionality of the Capital IQ database. We believe the main driving factor behind our data providers' approach to labeling two firms as competitors is whether they are mentioned as such in public documents, filings, publications, analyst reports, and other sources, as gathered by the data provider.

We believe our approach to classifying firms as competitors is reasonable and yields groups of competing firms that are intuitive and realistic. This classification offers some advantages over the widely used alternative based on industry classifications. The primary benefit lies in its apparent alignment with the concept of relevant markets in the antitrust context. When two companies are repeatedly referred to as competitors in various documents - as collected by our data provider - it is typically because they operate in the same real-world market, competing for the same customers.

Compared to industry classifications that often aggregate firms at a broad level, our method provides a more granular view of actual market interactions. We illustrate this with two firms close to the median in terms of competitor count. Both firms have 13 competitors according to Capital IQ. Table 5 displays competitors of *Wizz Air Holdings Plc.*, a European commercial airline, while table 6 displays competitors of *Piaggio & C. SpA*, a European manufacturer of motor vehicles.

By contrast, industry or sector-based classifications yield significantly higher counts: 1,984 at the sector level, 71 at the SIC 4-digit level, and 68 at the SIC 5-digit level. Furthermore, unlike industry-based methods, our data provider defines competitors at the firm level, allowing for firm-specific competitor sets. For example, different airlines may have distinct competitors, reflecting the nuances of their market positioning.

A closer look at the resulting competitor sets reveals that multi-product firms pose a distinct challenge when it comes to identifying their competitors. For instance, the car manufacturer Volkswagen, which has 73 competitors, is correctly identified as competing with firms such as Stellantis, Tesla, Volvo, BMW, Renault, and several other well-known car brands. However, its competitor set also includes less intuitive companies, such as Uber, Shell, and Deutsche Telekom. These firms are likely competitors (and possibly customer or suppliers) to Volkswagen in a niche market segment, and not in core car making.

A further example illustrating the complexity of classifying multi-product firms as competitors in the context of monitoring earnings calls is Microsoft, which has the third largest number of competitors

[32] A comprehensive review of these competitors is challenging due to their large number. However, the list of firms appears reasonable upon examination. Notable examples include Amazon's competition with Sennheiser in the audio device market, Spotify in the music streaming sector (Amazon Music), Zoom in video conferencing (Amazon Chime), Xiaomi in the tablet and smartphone market (Amazon Fire devices), Red Hat in cloud services (Amazon Web Services), and Electronic Arts in game development (Amazon Games). While some of these competitors may also have supplier or customer relationships with Amazon, they indeed appear to be genuine competitors. CapitalIQ also provides lists of suppliers and customers for Amazon. We cross-checked the list of competitors with these lists, based on the stock exchange ticker field that was available for 450 Amazon competitors. Of these firms we found 68 to be also on the list of suppliers and 60 on the list of customers. It appears therefore that the majority of these firms are indeed pure competitors.

(438) in our dataset. Notably, Logitech is listed as one of Microsoft's competitors, likely due to the fact that both firms offer hardware products that can be classified as substitutes, such as the *Microsoft Sidewinder Gaming Mouse* and Logitech's gaming mouses. However, it is unlikely that Microsoft would make sensitive references to Logitech in its earnings calls, given that the vast majority of Microsoft's revenue is generated from its business productivity and cloud software markets, rather than the computer accessories market where Logitech is a key player. This highlights a challenge of identifying competitors, as multi-product companies often have diverse businesses that may not be directly relevant to their competitors. As with any exercise involving the identification of competitors, this complication can be addressed so that our approach delivers valuable insights.

In summary, while our approach to classifying firms as competitors offers several advantages, it also carries an inherent risk of misclassification. We anticipate that this bias is more likely to result in Type I errors (false positives). Specifically, incorrectly labeling firms as competitors may suggest a relationship where none exists, particularly when these firms appear to communicate in a suspicious manner, as outlined in Criterion 3 of Section IV.D. It is less probable that we would fail to classify two firms as competitors if they are actively discussing their activities in earnings calls, which typically signals a genuine competitive relationship. As discussed in Section IV.E, avoiding Type I errors is particularly crucial for this type of screening. False positives can be costly for the authorities[33]. Therefore, robust methods to accurately classify firms as competitors are highly valuable and represent a promising avenue for further research.

3 Flagging Collusion-Facilitating Communication in Competitive Environment

Figure 8 illustrates the results of our screening approach. The variables STRAT and COMP for the full sample are depicted on the vertical and horizontal axes, respectively. Each observation corresponds to an earnings call by a company on a specific date (and time). It is easy to identify the earnings calls that can be regarded as outliers based on their intensity of use of bigrams in lists STRAT and COMP. The small dots highlight earnings calls in the top 1% range of STRAT and COMP. Circle markers highlight earnings calls that belong to a specific group of firms that were deemed competitors according to the method set out in Section V.B.2.

[33]Due to the subsequent manual review, it is unlikely that this type of error would impose costs on firms.

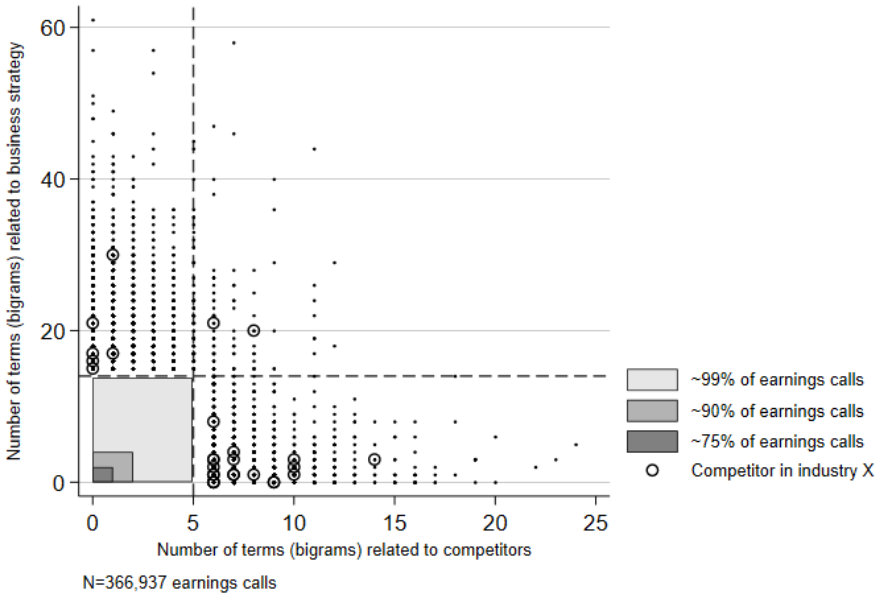

Figure 8: Earnings calls by number of sensitive terms related to business strategy and competitors - full sample and selected list of competitors

For this particular set of competitors, a disproportionately large percentage of earnings calls fall within the top 1% of earnings calls in terms of STRAT and COMP.[34] This indicates that these earnings calls are unusual in their frequency of trigger bigrams. These earnings calls would therefore be suitable candidates for further, detailed assessment by competition authorities. The competition authority may further refine the set of earnings calls for detailed assessment by focusing on those that are close in time or have a minimum number of competitors involved in the same quarter.

Competition authorities should carefully examine these earnings calls to understand the context in which these bigrams appear, the markets involved, and the potential impact on decision making. This requires meticulous manual review by one or more competition authority officials. By filtering out a few hundred documents from tens or even hundreds of thousands for closer inspection, quantitative screening provides assurance that enforcement efforts are concentrated on the most relevant documents. As a result, the manual review process becomes significantly more feasible, allowing authorities to focus their attention on a select few documents that are most likely to contain relevant information.

A final result of the screening exercise may be a graph such as Figure 8 or a table of companies that compete and have *"hot"* earnings calls. The resulting list may vary depending on the thresholds chosen for STRAT and COMP of the firm, its competitors, and the time horizon considered.

Figure 9 illustrates one outcome of our screening process. The figure shows boxplots of a simple

[34] For instance, one may consider a group of firms to have a disproportionately large number of earnings calls in the top 1% in terms of STRAT or COMP if more than 1% of the total earnings calls from the company or the group of competing companies fall within this range.

additive screening index across the top 15 four-digit industries, ranked by median index value, for the year 2018. We selected 2018 as the basis for our illustration because that year was not yet affected by the COVID-19 crisis, which may have impacted public communication patterns. The index at the firm-quarter level is defined as the sum of four components: the average of STRAT in the focal firm's earnings calls, the average of COMP in those same calls, and the corresponding averages for the firm's rivals during the same quarter. We then sum this index over the four quarters of 2018. Boxplots display the interquartile range (IQR), with the median marked by the central line.[35] The figure therefore provides a straightforward measure of which industries are particularly prone to using potentially sensitive keywords in their public communications. For instance, the figure shows that the median value of the index for firms in the specialty chemicals industry is 40. This means that, in a given quarter, the sum of the average number of STRAT and COMP bigram mentions in the firm's own earnings calls and the corresponding average across its competitors' earnings calls is 40, with values ranging from 0 at the first quartile to nearly 100 at the third quartile of the distribution.

Once a specific set of firms has been identified, it becomes easy to examine whether there are periods during which these firms show heightened usage of sensitive terms. Such documents can then be subjected to manual review.

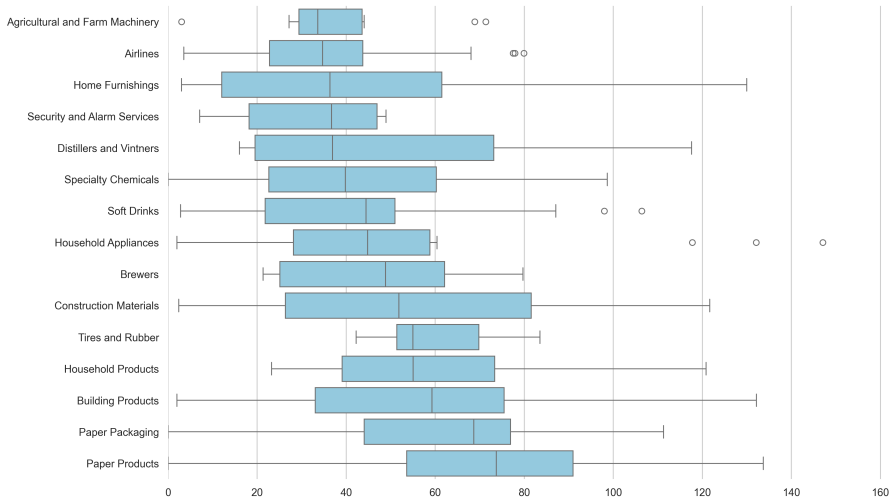

Figure 9: Top 15 Industries by Additive Communication Index (2018)

It is noteworthy that two of the top three sectors in Figure 9 are related to the paper industry, which was recently subject to cartel inspections by the European Commission (European Commission, 2018).[36] Furthermore, the airlines sector, where the use of collusive language in earnings calls has been well-documented (Aryal et al., 2022), appears on the list. This is evidence of the potential for finding

[35] Whiskers extend to the most extreme values within 1.5 interquartile ranges from the lower and upper quartiles, capturing the bulk of the distribution while excluding outliers, that are marked as dots.

[36] We note that, for illustrative purposes, this figure considers broad industry categories which likely encompass multiple relevant markets in the antitrust sense, each potentially involving distinct sets of competitors.

suspicious communications in other industries. The list in Figure 9 provides guidance on candidates for research.

To demonstrate the practical usefulness of our approach, we provide examples of excerpts from earnings calls that contain potentially collusive language.

4 Examples of Flagged Communication: The Security Solutions Market

Another way to demonstrate the effectiveness of our methods is through a more qualitative approach. Below, we present quotes from earnings calls in which focal firms disproportionately mention bigrams related to STRAT and COMP, as did their competitors during the same quarter. We will focus on the building products sector, that flag high up in the ranking, including two competitors, Assa Abloy and Allegion. Our attention has been drawn to these firms because they are indicated as competitors in our data and have several consecutive earnings calls in the - period that are among the top 5% in terms of STRAT and/or COMP, occasionally even in the top percentile.

In the following quotes, we merge statements from multiple executives and use [. . .] to indicate omitted parts of the original text for brevity.

The first case refers to two leading security solutions market players, Allegion Plc. and Assa Abbloy who showed up based on our analysis as deserving a closer look.[37] A cursory look at their transcripts reveal that these companies mentioned "price increase" in most earnings call between 2018 and 2022. Allegion regularly announces price increases with an explicit effective date, and talks about "disciplined market", "disciplined pricing environment", and "well-disciplined industry" in several conference calls. The examples below focus on the period of Spring-Autumn 2018.

Assa Abbloy, 26 April 2018 (STRAT: top 1%, COMP: top quartile)

> *"[In] EMEA we have strong price improvement. And we I would say overcompensate material costing in cases with pricing. So I don't see too many issues there. It was more challenging in the Americas and then in U.S. in particular, where we had tried to increase prices and where then the market was not always following. [...] But over time, material prices are the same for everybody. And as such, it's a good thing because, as such, everybody can increase prices. It's more the timing when the different players in the market do what. Now the good thing is there that we have seen now in quarter 1 also competition announcing price increases and some of them also announcing significant price increases. [...] Overall, we are confident that we will be able to further increase prices now in quarter 2. And that situation should ease up a bit on the pricing side."*

Competitor Allegion held an earnings call on the same day where the executive made the following statement in response to a question from an analyst who explicitly referred to having listened to the comments about pricing at the Assa Abbloy earnings call that morning.

[37] Security solutions include door hardware and locks.

Allegion, 26 April 2018 (STRAT: top 5%)

> *Executive: "I think we are in a better position [than Assa Abbloy]. As a price leader here, I think we've always worked extremely hard to get the value that our specification and products demand. [...] Demand is high, but I like our position to be able to get the price that we deserve in the marketplace [...]"*

The two earnings calls on April 26, 2018 contain elements consistent with collusive coordination. In the morning call Assa Abbloy provides reassurance that it intends to increase prices. To a competitor, this may reveal information about the expected sequence of pricing actions. In particular, Assa explains its past price increases have not always been followed. Such a message can serve to highlight the need for cooperation among competitors to successfully implement profitable price increases. Assa Abbloy also confirms it expects competitors to follow suit with price increases, because the cost increases "are the same for everybody."[38]

Assa Abloy asserting that cost increases are a "*good thing*" is consistent only with collusion, not with competition. Under competitive conditions, one would expect companies to absorb part of a cost increase due to competitive pressure, thereby reducing their margins. A cost increase can only be considered good news if all companies raise their prices in response. Doing that requires coordination.

Allegion in turn claims its price leader role, which may help resolve ambiguity Assa Abbloy expressed on the very morning of the same day regarding the sequence of price moves. A few days later the public messages about pricing continued.

Allegion, 8 May 2018 (STRAT and COMP in the top 5%)

> *"We announced a price increase effective July 1, which is the largest maybe in the last 15 years, and I know that's an industry that's disciplined. There's no reason that my competitors won't move and should move. It's important that we protect profitability [...]"*

Allegion's executive confirms that it expects the industry to behave in a "*disciplined*" manner. Competitors can understand this as confirmation that they are expected to continue moving up with prices. Using first person plural in the statement "*we protect profitability*" is consistent with joint action between competitors and Allegion. A few weeks later, Assa Abloy's earnings call flashes up on our indicators.

Assa Abbloy July 18, 2018 (Top percentile on STRAT, top quartile on COMP)

> *"We definitely don't have the ambition to buy market share. [...] We definitely are not a company that buys market share with reducing the price. That would be clearly the wrong strategy."*

Competitors may understand this statement as Assa Abbloy confirming its lack of interest in unleashing strong competition. "Buying market share" likely refers to price reductions. The statement leaves

[38] While a common input cost shock may indeed put pressure on competitors to increase prices, such a reaction is not automatic in a competitive market. Firms typically have some latitude to pass on costs in the form of higher prices.

little doubt that Assa Abbloy intends to avoid price cuts, which would be the expected behavior if the company was part of a collusive agreement. In little more than a week later, Allegion continues.

Allegion, 26 July 2018 (STRAT and COMP in the top 1%)

> *"We've seen our major global competitors fall in line. I think you've heard us talk that this industry is disciplined. I would describe Allegion as led the way, in terms of our price announcements [...]. And I think it reflected in our results versus the competitors. So I think there will be good price discipline and almost required with the amount of inflation this industry has seen over the last year."*

Competitors can understand the statement above as Allegion expecting them to remain disciplined with their prices. In light of the inflation mentioned, this likely corresponds to further price increases. The statement may also lend reassurance about the expected sequence of price moves, at least how it was in the recent past, when Allegion "*led the way*" with price increases. The executive praises the results of the price increase ("*reflected in our results versus the competitors*"), which both sends the message that price increases are profitable, and that competitors could do better, in order to have results like those of Allegion.

In early fall, the two companies follow up with several earnings calls that rank high on STRAT (less on COMP). An example is the following from Assa Abbloy, November 14, 2018:

> *"I already mentioned price is the easiest way from top line to bottom line. What is good in our industry is that it is an industry where you can increase prices. We, being the market leader, often, we are the first one coming with price increase, and we see that the market follows. If we increase price because of inflation, other places in the market have similar challenges. When we increase prices today, they follow."*

Assa Abloy refers explicitly to "*the industry*," which suggests that competitors are the main targets of the content. The company explains it is "*good*" that prices can increase in the industry, which clarifies to competitors its desire to go along with price increases. This statement furthermore is particularly explicit in clarifying the expected sequence of price moves. This time Assa Abbloy claims price leadership, which is particularly relevant information in light of the earlier statement of Allegion in a similar vein. While the Spring-Autumn 2018 statements of Assa Abbloy and Allegion make it unclear which of these companies is the price leader, that both are seeking to convey such a message only heightens collusive concerns.

C Summary: Flagging Collusion-Facilitating Communication in Earnings calls

The examples above are selected transcripts that came to our attention as a result of our screening exercise. We conclude that our method to use natural language processing (NLP) techniques is capable of picking up earnings calls with suspicious messages. Our STRAT and COMP measures are relevant indicators, especially if they are used in conjunction with knowledge of who are competitors.

Our approach to measuring communication is straightforward: we simply count the frequency of bigrams based on a predefined dictionary. We view this simplicity as a strength from a practical perspective, as it makes the method easy to explain to lawyers, judges, managers, and others involved in cartel investigations. Several modifications to this approach are possible. For instance, instead of relying on raw counts, one could assign weights to specific bigrams or documents.[39] While weighting schemes can offer advantages in some applications, they also introduce an element of arbitrariness and may reduce the intuitiveness of the analysis. The advancement of large language models presents a natural opportunity to enhance the screening of public texts for indications of collusion. Recent research has begun exploring these models as screening tools(Hassan et al., 2024).

VI Other Data Sources

We used two main data inputs to screen for indications of collusion. First, a corpus of public texts containing corporate communication. Second, information on the competitive relationship between firms. Competition authorities may not have access to the same data sources as we do, or their needs may require the use of other datasets.[40] We discuss here the information needs and provide some advice.

Our focus has been on public communication in earnings calls, which are particularly valuable for corporate surveillance. Hassan et al. (2024) describes them as a "*marketplace of information*," noting that their "*wide availability ensures that researchers and policymakers have easy access to this rich source of corporate information.*" While earnings calls are becoming increasingly accessible in the U.S. and Europe, they remain far less common in other jurisdictions (Figure 1). Our method can be extended to any other publicly available text sources that may aid detection efforts, including trade press articles, news websites, interviews, and more. Competition authorities would need access to digital archives of texts that feature public statements of competing firms.

A potential challenge when using text sources other than earnings calls lies in the difficulty of reliably identifying companies within such documents. Earnings calls are explicitly tied to a specific company, making attribution straightforward. In contrast, general news articles may not mention company names in direct proximity to relevant quotes, requiring the use of natural language processing techniques to extract such entities from the broader context. While this adds a layer of complexity, it is manageable as robust named entity recognition tools are readily available in standard programming environments such as Python.

A second challenge involves identifying groups of competing firms. For instance, news articles may contain questionable or ambiguous statements made by various executives. To assess their significance, it is very useful to cross-reference these statements with contemporaneous public communications

[39] For example, text analysis often employs measures such as *tf-idf* (term frequency-inverse document frequency), which is a weighting method that highlights important words by balancing their frequency in a document against their rarity across the entire collection.

[40] For example, analysing competition within national markets may require data of more local nature than our global database.

from competing companies. However, determining which firms qualify as competitors is not always straightforward. Relevant information can often be obtained from trade journals, industry publications, business directories, and company surveys, which provide useful resources for regulatory authorities or analysts conducting such investigations.

Screening public texts beyond earnings calls for signs of collusion can be both valuable and effective. Traditional print news serves as a clear and accessible alternative source of information. To illustrate the potential of such an approach, we present a preliminary example based on a simple keyword search in the Factiva news database, offering a glimpse into the kinds of insights such a screening exercise can uncover.

In 2011-2012, a series of articles in the Financial Times reported on the intense price war in the container shipping industry. In November 2011, the press wrote about *"panic"* in the container shipping industry due to market leader Maersk having gained market shares after earlier price reductions (Financial Times, 2011).

> ***Container slump hits Møller-Maersk*** *[...] "There's a little bit of panic in the industry. Yes, we're following the rates down when they drop, to make sure we keep our share." [...] But Mr Andersen insisted rates could not carry on at their current, "totally unsustainable" levels – which are mostly too low to cover lines' operating costs – for long.*

In this press release, the Maersk executive conveys two key points. First, he explains how Maersk responds to competitor price cuts, stating that the company "follows the rates down if they drop" with the explicit goal of maintaining market share. This statement can be interpreted as a warning to competitors that Maersk will retaliate against any deviation from a tacit agreement on higher rates. Second, his characterization of current rates as "totally unsustainable" suggests a desire for industry-wide price increases, reinforcing the idea of coordinated behavior.

In February 2012 the press reported again on the container shipping industry, where the company's financial director made the following statements in an interview (Financial Times, 2012):

> ***Maersk profits sink 33% on container woes*** *[...] Trond Westlie, the group's finance director, pointed to Maersk Line's decision on February 17 to cut Asia to Europe capacity by 9 per cent because rates had reached "unsustainably low levels". The decision was seen as the end of a damaging battle for market share between Maersk and its competitors including Switzerland's Mediterranean Shipping Company, operator of the second-biggest container ship fleet by capacity. Mr Westlie said the line was considering the need for further capacity cuts beyond those announced. "We are using all our tools in the tool box in looking at [the problem of profitability]" [...] "We're looking at our pricing models, to find a better stability in the market going forward," Mr Westlie said.*

This statement is notable because, as in the 2011 press interview, Maersk once again describes rates as "unsustainably low." This language can apply to the industry as a whole, implicitly signaling to

competitors how Maersk expects cooperation to unfold, that is, through price increases. Additionally, Maersk confirms that it has begun preparing for such increases, stating that it is "*looking at our pricing models*" to achieve greater market stability. The article explicitly mentioned that the last capacity move can be regarded as the "end of a damaging battle" between competitors, suggesting the beginning of a new, cooperative period, possibly triggered by the interviews. Public statements carry weight.

Around two weeks later, Financial Times Deutschland (2012) reported an interview with the Maersk leadership, who left little doubt about the company's intention to handle the situation in cooperation with rivals (text translated from the German original).

> ***Container shipping companies renounce predatory pricing*** *[...] Market leader Maersk announces abandonment of aggressive strategy. [...] In recent weeks, shipping companies such as Maersk Line and Hapag-Lloyd have successfully raised prices on the main route between Asia and Europe. [...] With the price increase, a ruinous battle for market share, initiated by market leader Maersk Line and the second-placed MSC, is coming to an end. Price competition probably caused losses for almost the entire industry last year [...] Now, Maersk has apparently offered the industry a truce.*
>
> *"We're where we want to be in terms of size, and therefore our focus is changing from 'wanting to grow faster than the market' to 'growing with the market,'" Sören Skou told the shipping newspaper "Lloyd's List."*
>
> *Maersk intends to use this to obtain higher rates. However, if the competition doesn't accept the offer, the company will fight back:*
>
> *"We hope our competitors are also satisfied with this. If they aren't, we will defend our position at all costs," said Skou [Maersk Line CEO]*

The article contains multiple indications that Maersk is signaling its intention to coordinate with rivals on pricing. In the article Maersk openly renounces its former "*aggressive strategy*," announces a departure from the "*vicious battle for market share*." A "*truce*" to the industry amounts to a desire to cooperate with rivals.

Several subsequent articles in various maritime trade journals feature interviews with Maersk executives discussing efforts to raise rates while maintaining market share, an approach that, all else being equal, suggests some degree of industry coordination and weak competition (Dow Jones International News - Factiva, 2012; Maritime Gateway - Factiva, 2012; Tradewinds - Factiva, 2012). This series of public communications occurred shortly before the European Commission initiated proceedings against several major container shipping service providers, including Maersk, in the AT.39850 — Container Shipping case (Rabinovici, 2017).[41]

[41] We are not aware that news statements directly influenced the European Commission's decision to undertake investigative measures in the container shipping industry at the time. When looking into the earnings calls of Maersk of the period mentioned here, we however do see several statements related to the "*price war*" that are in line with the tone in the trade

The fact that public statements by container shipping executives preceded a cartel investigation suggests that actively screening for such communications in news outlets may provide valuable indicators of potential collusion.

Regardless of the source, the core steps of a text screening exercise remain the same. First, the investigator must compile a corpus of public documents to serve as the main database. This may involve web scraping, with resources such as the Internet Archive (archive.org) offering multiple ways to access and download material. Additionally, various trade and general press archives are available worldwide, such as Arcanum (arcanum.com) and the GDELT Project (gdeltproject.org), the latter providing digitized news along with pre-extracted entity recognition from thousands of news websites worldwide.

After finalizing the corpus, the next step is to develop a dictionary and define a keyword list (Section IV.B). Our two-pronged approach, focusing on bigrams related to strategy (STRAT) and references to competitors or competition (COMP), offers a structured path rooted in economics. Next, we count the frequency of these bigrams in documents and assign these counts to firms over time (Section V.B). Identifying competitive relationships is particularly useful for selecting the most relevant documents for a detailed analysis (Section V.B.2). Finally, documents are ranked according to the number of trigger bigram mentions (Section V.B.3). The outcome is a set of firms operating in the same market whose documents contain the highest frequency of potentially collusive language. Then, this refined set of documents must undergo a careful manual review to fully understand the communication and market context.

VII Conclusions

Our study highlights the growing importance of screening public corporate communications for potential collusion. Traditional cartel detection methods have focused primarily on secret agreements, yet recent research suggests that public communication can serve as a substitute for secret agreements, allowing firms to align strategies while maintaining plausible deniability. Our screening approach aims to detect companies that use public announcements as a means of coordination. We screen for these patterns in publicly accessible corporate texts by identifying key linguistic patterns associated with collusion. Although collusion-related keywords are not uncommon, they are far from the norm: less than half of earnings calls contain such language. By applying natural language processing methods to a vast dataset of earnings call transcripts, we provide competition authorities with a novel tool to detect signals indicative of anti-competitive behavior. Our method has already proven its practical value, contributing to the 2024 European Commission inspections in the automotive tire sector.

Looking ahead, our approach can be extended to other sources of corporate communication, such as industry reports, media interviews, and the trade press, to further enhance cartel detection efforts.

press. For example, Maersk mentions in the earnings call of the 11th of May, 2011:"*I don't really believe that this calls for a price war, and since we don't have a lot of tonnage, we're definitely not going to lead a price war.*"

As machine learning and natural language processing techniques continue to evolve and large language models advance, the ability to detect subtle collusive signals in vast amounts of textual data will only improve. However, automated screening must be complemented by careful manual analysis to contextualize findings and avoid false positives. By integrating these tools into the enforcement arsenal, competition authorities can enhance their ability to proactively identify and deter collusion.

References

Abrantes-Metz, Rosa M, Luke M Froeb, John Geweke, and Christopher T Taylor, "A variance screen for collusion," *International Journal of Industrial Organization*, 2006, *24* (3), 467–486.

Affeldt, Pauline, Tomaso Duso, Klaus Gugler, and Joanna Piechucka, "Market Concentration in Europe: Evidence from Antitrust Markets," Technical Report, CEPR Discussion Paper DP15699 2021.

Areno, Filipe B., Brian V. Breheny, Adrian J. S. Deitz, Rajeev P. Duggal, Ryan J. Dzierniejko, Gregory A. Fernicola, P. Michelle Gasaway, David J. Goldschmidt, Stephan Hutter, Thomas J. Ivey, Laura A. Kaufmann Belkhayat, James A. McDonald, Gregg A. Noel, Michael J. Schwartz, Jonathan B. Stone, Kenji Taneda, Danny Tricot, Joseph (Yossi) Vebman, J. Mathias von Bernuth, Dwight S. Yoo, Andrew J. Brady, and Kenneth Chase, "Key Considerations for Non-US Companies Listing in the US," *Skadden Publication / Capital Markets Alert*, June 2020.

Aryal, Gaurab, Federico Ciliberto, and Benjamin T Leyden, "Coordinated capacity reductions and public communication in the airline industry," *The Review of Economic Studies*, 2022, *89* (6), 3055–3084.

Autoriteit Consument en Markt, "Decision of the Board of the Netherlands Authority of Consumers Markets within the meaning of Section 49a of the Dutch Competition Act, Case number 13.0612.53," `https://www.acm.nl/sites/default/files/old_publication/publicaties/14326_commitment-decision-regarding-mobile-operators.pdf` 2014. Accessed: 2025-03-20.

Borenstein, Severin, "Rapid communication and price fixing: The airline tariff publishing company case," in John Kwoka and Lawrence White, eds., *The Antitrust Revolution: Economics, Competition, and Policy*, 4th ed., Oxford University Press, 1994.

Bourveau, Thomas, Guoman She, and Alminas Žaldokas, "Corporate disclosure as a tacit coordination mechanism: Evidence from cartel enforcement regulations," *Journal of Accounting Research*, 2020, *58* (2), 295–332.

Chassang, Sylvain, Kei Kawai, Jun Nakabayashi, and Juan Ortner, "Robust screens for noncompetitive bidding in procurement auctions," *Econometrica*, 2022, *90* (1), 315–346.
Compagnie générale des établissements Michelin v European Commission

***Compagnie générale des établissements Michelin v European Commission*.**

Cuddy, Emily, Robert H. Porter, Amanda Starc, and Thomas G. Wollmann, *"Entry Barriers, Personal Relationships, and Cartel Formation," in Jr. Joseph E. Harrington and Maarten Pieter Schinkel, eds., Cartels Diagnosed: New Insights on Collusion, Cambridge University Press, 2024.*

Danish Competition and Consumer Authority, *"Collusion Detection in Public Procurement using Computational Methods," 2022. Paper No.56.*

Dow Jones International News - Factiva, *"Maersk Chairman Sees Shipping Loss Despite Successful Rate Hikes," 2012. Accessed: 2025-03-15.*

European Commission, *"Antitrust: Commission confirms unannounced inspection in kraft paper sector,"* https://ec.europa.eu/commission/presscorner/detail/en/ip_24_561 *2018. [Accessed 24-11-2024].*

_ , https://ec.europa.eu/commission/presscorner/api/files/document/print/en/statement_18_264/STATEMENT_18_264_EN.pdf, *note = [Accessed 05-05-2025], title = Commission carries out unannounced antitrust inspections in the tyres sector — ec.europa.eu, year = 2024,.*

European Commission, Directorate-General for Competition, *"Guidelines on the applicability of Article101of the Treaty on the Functioning of the European Union to horizontal co-operation agreements," Official Journal of the European Union, C259/01. CELEX52023XC0721(01). Available in all EU official languages.*

European Union, *"Council Regulation (EC) No 1/2003 of 16 December 2002 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty," 2003. Official Journal of the European Union, L 1, 4.1.2003, p. 1-25.*

Financial Times, *"Container slump hits Møller-Maersk,"* https://www.ft.com/content/b158a49c-0ab3-11e1-b9f6-00144feabdc0 *2011. Accessed: 2025-03-15.*

_ , *"Maersk profits sink 33% on container woes,"* https://www.ft.com/content/f36d4958-6113-11e1-a738-00144feabdc0 *2012. Accessed: 2025-03-15.*

Financial Times Deutschland, *"Containerreeder schwören Kampfpreisen ab,"* https://web.archive.org/web/20120317050829/http://www.ftd.de:80/unternehmen/handel-dienstleister/:logistik-containerreeder-schwoeren-kampfpreisen-ab/70008237.html *2012. Accessed: 2025-03-15.*

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, *"Text as data," Journal of Economic Literature, 2019, 57 (3), 535–574.*

Giannino, Michele, *"Collusion in Public Contracts Procurement: Suppliers of School Cleaning Services Fined for Bid Rigging (Italy)," Journal of European Competition Law Practice, 08 2016, 8 (4), 247–252.*

Harrington, Joseph E., *"Detecting Cartels," in Paolo Buccirossi, ed., Handbook of Antitrust Economics, The MIT Press, 2008, pp. 213–258.*

Harrington, Joseph E, *"Collusion in plain sight: Firms' use of public announcements to restrain competition," Antitrust Law Journal, 2022, 84 (2), 521–563.*

_ **and David Imhof**, *"Cartel screening and machine learning," Stan. Computational Antitrust, 2022, 2, 133.*

Hassan, Tarek Alexander, Stephan Hollander, Aakash Kalyani, Laurence van Lent, Markus Schwedeler, and Ahmed Tahoun, *"Economic Surveillance using Corporate Text," Technical Report, National Bureau of Economic Research 2024.*

Hirst, Nicholas and Jean Comte, *"Michelin tells court there's 'no evidence' underpinning EU cartel raids," MLex, 2025, (5 Mar 2025).*

Huber, Martin and David Imhof, *"Machine learning with screens for detecting bid-rigging cartels," International Journal of Industrial Organization, 2019, 65, 277–301.*

Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole, *"The economics of tacit collusion," 2003.*

Kepler, John D, *"Private communication among competitors and public disclosure," Journal of Accounting and Economics, 2021, 71 (2-3), 101387.*

Lianos, Ioannis and Florian Wagner von Papp, *"Tackling invitations to collude and unilateral disclosure: the moving frontiers of competition law?," Journal of European Competition Law & Practice, 2022, 13 (4), 249–253.*

Maritime Gateway - Factiva, *"Maersk to defend market share gains," 2012. Accessed: 2025-03-15.*

Motta, Massimo, *Competition policy: theory and practice, Cambridge University Press, 2004.*

Organisation for Economic Co-operation and Development, *OECD Competition Trends 2023, OECD Publishing, 2023.*

Rabinovici, Itai, *"Public exchange of information after container shipping," Journal of European Competition Law & Practice, 2017, 8 (3), 149–156.*

Rosenblatt, Howard and Tomas Nilsson, *"Analyst Calls and Price Signaling Under EU Law," The Antitrust Source, 2012, (6).*

Sheng, Qiwen and Tomislav Vukina, *"Public communication as a mechanism for collusion in the broiler industry," Review of Industrial Organization, 2024, 64 (1), 57–91.*

Tan, Chade-Meng, Yuan-Fang Wang, and Chan-Do Lee, *"The use of bigrams to enhance text categorization," Information processing & management, 2002, 38 (4), 529–546.*

Tradewinds - Factiva, *"Focussing on profitability will pay off,"* `https://www.tradewindsnews.com/weekly/focussing-on-profitability-will-pay-off/1-1-279675` *2012. Accessed: 2025-03-15.*

Werle, Marcel and Sven Laumer, *"Competitor identification: A review of use cases, data sources, and algorithms," International Journal of Information Management, 2022, 65, 102507.*

Appendix

Dictionary approach

Variable	Keywords
Key 1 **(Strategy)**	price, pricing, margin, [...]
Key 2 **(Competitors)**	competition, competitive, compete, [...]
Key 3 **(Intent)**	should, will, expect, [...]
Key 4 **(Action)**	grow, raise, rise, increase, [...]
Key 5 **(Triggers)**	excess, insane, aggressive, oversupply, [...]
Combination Rules	Keys 1 (Strategy) with 3, 4, 5; Keys 2 (Competition) with 3, 4, 5

Competition and strategy indicators **COMP** and **STRAT** are then constructed. **STRAT** is the combination of *keys1* and *keys3, keys4, keys5* while **COMP** is constructed as the comnbination of *keys2* with *keys3, keys4, keys5*.

Dimensions of earnings calls data

Table 8: Overview of Data Set

Variables: Dimensional	
i	Focal firm
j	Competitor
t	Quarter
s	5th digit industry
r	Continent of firm i
d	Earning call of i
c	Earning call of j
Variables of Interest	
$Strat_{tsidc}$	Number of uni-/bigrams classified as strategic in d
$Comp_{tsidc}$	Number of uni-/bigrams classified as competitive in d
$StratRival_{tsidc}$	Number of uni-/bigrams classified as strategic in c
$CompRival_{tsidc}$	Number of uni-/bigrams classified as competitive in c

Descriptives

Desriptives for STRAT and COMP: While many firms also talk about competition, they do it to a much lesser extent than talking about strategy. Sharp decrease from 2017 onwards. Main message: Firms talk a lot about strategy and competition. Here we could add some descriptive analysis of heterogeneity: industries, countries, firm's size? First plots for Europe vs. US. Selected sectors could be added.

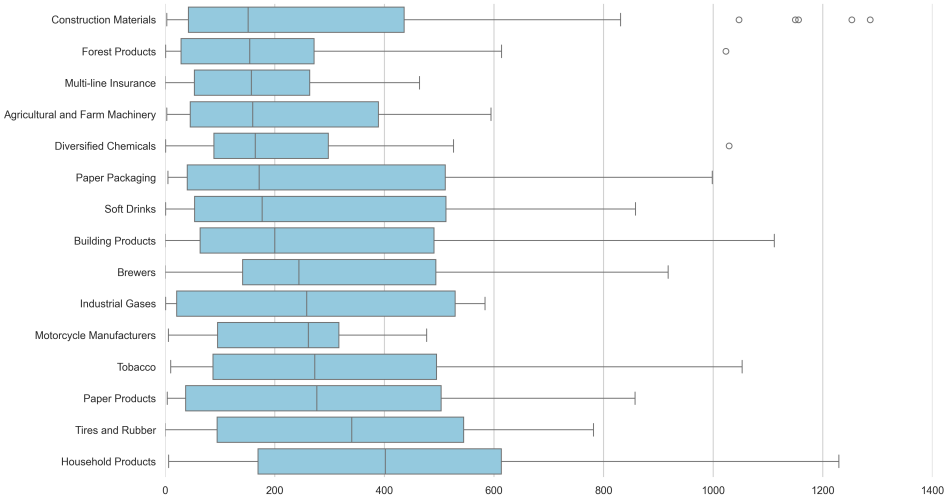

Figure 10: Top 15 Industries by Additive Communication Index (2004-2022)

PREVIOUS DISCUSSION PAPERS

430 Duso, Tomaso, Harrington, Joseph E. Jr., Kreuzberg, Carl and Sapi, Geza, Public Communication and Collusion: New Screening Tools for Competition Authorities, October 2025.

429 Benndorf, Volker, Kübler, Dorothea and Normann, Hans-Theo, Information Unraveling and Limited Depth of Reasoning, September 2025.
Forthcoming in: Games and Economic Behavior.

428 Martinez-Martinez, Ismael and Normann, Hans-Theo, Cooperation in Multiplayer Dilemmas, September 2025.
Forthcoming in: Journal of Economic Behavior and Organization.

427 Neyer, Ulrike, Stempel, Daniel and Stevens, Alexandra, Subsistence Consumption and Inflation Heterogeneity: Implications for Monetary Policy Transmission in a HANK Model, September 2025.

426 Dibiasi, Andreas and Erhardt, Katharina, Young Firms under Pressure: Heterogeneous Investment Responses to a Trade Shock, August 2015.

425 Dertwinkel-Kalt, Markus and Wey, Christian, Comparing the Effects of Subsidies on Target Goods, July 2025.

424 Dertwinkel-Kalt, Markus and Wey, Christian, The Economics of Fleet-Wide Emission Targets and Pooling in the EU Car Market, July, 2025.

423 Chowdhury, Shyamal, Puente-Beccar, Manuela, Schildberg-Hörisch, Hannah, Schneider, Sebastian O. and Sutter, Matthias, Spatial Patterns in the Formation of Economic Preferences, July 2025.

422 Breitkopf, Laura, Chowdhury, Shyamal, Kamhöfer, Daniel A., Schildberg-Hörisch, Hannah and Sutter, Matthias, The Right Timing Matters: Sensitive Periods in the Formation of Socio-Emotional Skills, June 2025.

421 Mori, Tomoya and Wrona, Jens, Culture, Tastes, and Market Integration: Testing the Localized Tastes Hypothesis, May 2025.

420 Marchal, Léa, Ourens, Guzmán and Sabbadini, Giulia, When Immigrants Meet Exporters: A Reassessment of the Migrant-Native Wage Gap, May 2025.

419 Kesler, Reinhold, Skiera, Bernd, Kraft, Lennart and Koschella, Tim, Existence, Antecedents and Consequences of Non-Compliance in Mobile App Markets, April 2025.

418 Oschmann, Sebastian, Vertical Market Structure Matters: The Case of a Horizontal Retail Merger in the German Gasoline Market, March 2025.

417 Breitkopf, Laura, Chowdhury, Shyamal, Priyam, Shambhavi, Schildberg-Hörisch, Hannah and Sutter, Matthias, Nurturing the Future: How Positive Parenting is Related to Children's Skills and Well-being, October 2024.

416 Rademacher, Philip, Forecasting Recessions in Germany with Machine Learning, September 2024.

415 Erhardt, Katharina and Gupta, Apoorva, Go Wide or Go Deep: Margins of New Trade Flows, August 2024.

414 Gupta, Apoorva and Stiebale, Joel, Gains from Patent Protection: Innovation, Market Power and Cost Savings in India, May 2024.

413 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A., and Schildberg-Hörisch, Hannah, Schooling and Self-Control, March 2024.

412 Huelden, Tobias, Jascisens, Vitalijs, Roemheld, Lars and Werner, Tobias, Human-Machine Interactions in Pricing: Evidence from Two Large-Scale Field Experiments, March 2024.

411 Hermes, Henning, Lergetporer, Philipp, Mierisch, Fabian, Schwerdt, Guido and Wiederhold, Simon, Does Information about Inequality and Discrimination in Early Child Care Affect Policy Preferences? January 2024.

410 Hunold, Matthias and Werner, Tobias, Algorithmic Price Recommendations and Collusion: Experimental Evidence, December 2023.

409 Herzog, Sabrina, Schildberg-Hörisch, Hannah, Trieu, Chi and Willrodt, Jana, Who is in Favor of Affirmative Action? Representative Evidence from an Experiment and a Survey, November 2023.

408 Stadelmann, David, Thomas, Tobias and Zakharov, Nikita, Too Hot to Play it Cool? Temperature and Media Bias, November 2023.

407 Dertwinkel-Kalt, Markus and Wey, Christian, Why "Energy Price Brakes" Encourage Moral Hazard, Raise Energy Prices, and Reinforce Energy Savings, September 2023.
Forthcoming in: RAND Journal of Economics.

406 Bertermann, Alexander, Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, More Education Does Make You Happier – Unless You Are Unemployed, September 2023.

405 Trieu, Chi, Who's Who: How Uncertainty About the Favored Group Affects Outcomes of Affirmative Action, August 2023.
Published in: Journal of the Economic Science Association, 9 (2023), pp. 252-292.

404 Fischer, Kai, Martin, Simon and Schmidt-Dengler, Philipp, The Heterogeneous Effects of Entry on Prices, July 2023.

403 García-Vega, María, Gupta, Apoorva and Kneller, Richard, Is Acquisition-FDI During an Economic Crisis Detrimental for Domestic Innovation?, July 2023.

402 Haucap, Justus and Stiebale, Joel, Non-price Effects of Mergers and Acquisitions, July 2023.

401 Simion, Ștefania and Sulka, Tomasz, Multidimensional Cognitive Ability, Intermediate Channels, and Financial Outcomes, May 2023.

400 Bartling, Björn, Cappelen, Alexander W., Hermes, Henning, Skivenes, Marit and Tungodden, Bertil, Free to Fail? Paternalistic Preferences in the United States, May 2023.

399 Kandelhardt, Johannes, Flexible Estimation of Random Coefficient Logit Models of Differentiated Product Demand, May 2023.

398 Hermes, Henning, Lergetporer, Philipp, Mierisch, Fabian, Peter, Frauke and Wiederhold, Simon, Discrimination on the Child Care Market: A Nationwide Field Experiment, April 2023.

397 Schmal, W. Benedikt, Haucap, Justus and Knoke, Leon, The Role of Gender and Coauthors in Academic Publication Behavior, March 2023.
Published in: Research Policy, 52 (2023), 104874.

396 Magin, Jana Anjali, Neyer, Ulrike and Stempel, Daniel, The Macroeconomic Effects of Different CBDC Regimes in an Economy with a Heterogeneous Household Sector, March 2023.

395 Dertwinkel-Kalt, Markus and Wey, Christian, Resale Price Maintenance in a Successive Monopoly Model, February 2023.
Forthcoming in: Journal of Industrial Economics.

394 Hermes, Henning, Krauß, Marina, Lergetporer, Philipp, Peter, Frauke and Wiederhold, Simon, Early Child Care and Labor Supply of Lower-SES Mothers: A Randomized Controlled Trial, December 2022.

393 Chowdbury, Shyamal, Schildberg-Hörisch, Hannah, Schneider, Sebastian O., and Sutter, Matthias, Information Provision Over the Phone Saves Lives: An RCT to Contain COVID-19 in Rural Bangladesh at the Pandemic's Onset, November 2022.

392 Normann, Hans-Theo and Sternberg, Martin, Human-Algorithm Interaction: Algorithmic Pricing in Hybrid Laboratory Markets, October 2022.
Published in: European Economic Review, 152 (2023), 104347.

391 Hunold, Matthias and Petrishcheva, Vasilisa, Foreclosure and Tunneling with Partial Vertical Ownership, September 2022.

390 Haucap, Justus and Heldman, Christina, The Sociology of Cartels, August 2022.
Published in: European Journal of Law and Economics, 56 (2023), pp. 289-323.

389 Döpper, Hendrik, Sapi, Geza and Wey, Christian, A Bargaining Perspective on Vertical Integration, May 2022.
Published in: Canadian Journal of Economics, 57 (2024), pp. 199-224.

388 Bachmann, Ronald, Gonschor, Myrielle, Lewandowski, Piotr and Madoń, Karol, The Impact of Robots on Labour Market Transitions in Europe, May 2022.

387 Fremerey, Melinda, Hörnig, Lukas and Schaffner, Sandra, Becoming Neighbors with Refugees and Voting for the Far-Right? The Impact of Refugee Inflows at the Small-Scale Level, April 2022.
Published in: Labour Economics, 86 (2024), 102467.

386 Fischer, Kai, Alcohol Prohibition and Pricing at the Pump, March 2022.

385 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, The Determinants of Population Self-Control, March 2022.
Forthcoming in: Economic Journal under the title "Surveillance and Self-Control".

384 Sulka, Tomasz, Planning and Saving for Retirement, March 2022.
Published in: European Economic Review, 160 (2023), 104609.

383 Cattan, Sarah, Kamhöfer, Daniel A., Karlsson, Martin and Nilsson, Therese, The Long-term Effects of Student Absence: Evidence from Sweden, March 2022.
Published in: Economic Journal, 133 (2023), pp. 888-903.

382 Martin, Simon and Rasch, Alexander, Collusion by Algorithm: The Role of Unobserved Actions, March 2022.
Published in: International Journal of Industrial Organization, 92 (2024), 103036
under the title "Demand Forecasting, Signal Precision, and Collusion with Hidden Actions".

381 Haucap, Justus, Nedic, Radivoje and Şimşek, Talha, An Empirical Analysis of German Casino Locations, March 2022.
Published in: European Journal of Law and Economics, 55 (2023) pp. 291-311.

380 Haucap, Justus, Heldman, Christina and Rau, Holger A., Gender and Collusion, March 2022.

379 Schain, Jan Philip, Foreign Institutional Investors and the Great Productivity Slowdown, November 2022 (First Version February 2022).

378 Neyer, Ulrike and Stempel, Daniel, How Should Central Banks React to Household Inflation Heterogeneity?, January 2022.

Older discussion papers can be found online at:
http://ideas.repec.org/s/zbw/dicedp.html

Heinrich-Heine-Universität Düsseldorf

Düsseldorfer Institut für Wettbewerbsökonomie (DICE)

Universitätsstraße 1, 40225 Düsseldorf

ISSN 2190-992X (online)
ISBN 978-3-86304-429-9