DEPARTMENT
OF ECONOMICS

Modern Markets: Competition in the 21st Century

Philip Hanspach

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Florence, 25 September 2023

European University Institute
Department of Economics

Modern Markets: Competition in the 21st Century

Philip Hanspach

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Researcher declaration to accompany the submission of written work Department Economics - Doctoral Programme

I Philip Hanspach certify that I am the author of the work Modern Markets: competition in the 21st century I have presented for examination for the Ph.D. at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the Code of Ethics in Academic Research issued by the European University Institute (IUE 332/2/10 (CA 297).

I declare that this work consists of 54954 words.

Statement of inclusion of previous work:

I confirm that chapter 3 was jointly co-authored with Mr. Giacomo Calzolari and I contributed 50% of the work.

Philip Hanspach, 02.05.2023

Abstract

This thesis is composed of three independent chapters, the third of which consists of two separate but thematically related papers.

In Chapter 1, I introduce a theoretical model of vertical integration with a novel demand structure to investigate the effect of vertical integration into Internet infrastructure on competition in digital markets. I find that pure infrastructure providers have an incentive to accommodate vertically integrated firms by becoming "commoditized" suppliers of infrastructure. My model explains new trends in digital markets and has implications for competition policy, industrial policy and political economy.

In Chapter 2, I estimate the effects of crisis intensity and deregulation on home bias in procurement. Using a novel data set on the award of procurement contracts for medical supplies during the first wave of the Covid-19 pandemic in Europe, I study the propensity to award contracts internationally. I document a unique shift towards international procurement, driven by local spikes in infection rates and deregulation.

In Chapter 3, I study the role of pricing algorithms in online marketplaces. Its first part is a joint article with Giacomo Calzolari that describes the algorithmic repricing industry. Based on a novel sample of 130 repricing companies, we study the prices and claimed attributes of pricing algorithms. We find that turn-key algorithmic pricing services are widely available, and discuss product features, fees, and associated services. The second part of Chapter 3 is a literature review on algorithmic pricing. I summarize findings from the economics literature covering computational, experimental, and empirical methods as well as adjacent fields. I argue that a lack of understanding of buyer responses to algorithmic pricing cycles and endogenous adoption of algorithmic pricing are the main gaps in the literature.

Acknowledgments

This work would not have been possible without the support and inspiration of many others. First of all are my supervisors Giacomo Calzolari and David K. Levine who have guided my PhD research with their immense knowledge and insight. Their most important contribution, however, was in the trust and confidence that they placed in me. During the past 11 years, many other great teachers of economics in Bonn, Singapore, Toulouse and Florence helped to set me on my path and stoke my interest in different subjects. Some of these teachers who deserve special mention were Hendrik Hakenes, Jürgen von Hagen, Omkar Shrestha, Farid Gasmi, Thomas-Olivier Léautier, Paul Seabright, Jean Tirole, Bart Taub, and Russell Cooper. I thank the economists at the Chief Economist Team of DG Competition in Brussels, Belgium, who have facilitated a research stay that helped me a lot, especially in developing the first chapter, but also in improving as an economist. Thanks go also to colleagues and classmates at the EUI Department of Economics and other groups, including the Tech Cluster and the Competition Law Working Group, that helped me sharpen my thinking and made the PhD more interesting. Finally, thanks go to all the other professional colleagues who provided comments and suggestions at seminars, conferences or bilateral meetings, especially to those that I forgot to acknowledge individually at the end of each chapter.

Meinen Eltern gewidmet, ohne die ich Vieles nicht erreicht hätte.

Contents

1 Internet infrastructure and competition in digital markets

Abstract

Large digital platform companies increasingly integrate vertically by building Internet infrastructure. These proprietary infrastructures confer quality advantages in markets for digital services. I model investment incentives for an infrastructure firm and a vertically-integrated firm facing a rival without proprietary infrastructure. Small changes in the marginal cost of investment lead to a discontinuous jump in investment incentives both for the infrastructure firm and the vertically-integrated firm if the latter has more infrastructure than the former. The infrastructure firm benefits from "commoditization" when its infrastructure is smaller. I derive conditions under which the resulting increase in investment is socially efficient. As the market share of the rival firm decreases, a trade-off arises between efficiency and "contestability", a key objective in European competition policy for digital markets.
Keywords: vertical integration, competition policy, net neutrality, Internet, Internet infrastructure, commoditization
JEL Codes: L13, L42, L51, L63, L86

1.1 Introduction

Reigning in and regulating "big tech", a catch-all for diverse companies known by acronyms such as GAFAM or BAT, is on the political agenda in China, the EU, and the US. Economists have explained the size and power of "big tech" with features that generate "winner-takes-all" markets: (indirect) network effects, platform economics, or the use of big data analytics. One less well-researched aspect of "big tech" market power is the role of proprietary Internet infrastructures. This paper seeks to fill this gap by proposing a theory of vertical integration in digital markets to study the effect of proprietary Internet infrastructure on competition.

The scale and the ownership structure of the physical aspect of the Internet - data centers, Internet exchange points, backbone - have undergone drastic changes in the past decade. Traditionally, large voice carriers connected local networks. These carriers deliver data packages based on principles of net neutrality and best-effort. The largest Internet Service Providers (ISP) interconnect with each other free of payment, creating a global network of networks, the Internet.

More recently, consumer-facing firms such as Google, Netflix or Meta, have increasingly complemented this so-called "public Internet" (operated by private companies nonetheless) with their own investments.[1] Some of the largest digital companies have made joint investments with traditional carriers and also created parallel, proprietary infrastructures.[2] These networks enable higher quality or guaranteed reliability contracts without violating net neutrality rules. This is essential for quality- or latency-sensitive applications, ranging from entertainment to corporate and security-related applications that require near 100% uptime.

Researchers and competition authorities are increasingly aware of proprietary Internet infrastructure but there is little knowledge about its implications. The EU's Digital Markets Act (DMA) recognizes that large platforms, which it calls "gatekeepers", can steer and block access

[1] An article by the Financial Times describes investors' diverging views on the competitive dynamics between data center operators that can be broadly summarized as belonging to the "public Internet" and those operated by "big tech groups". Financial Times, Will the cloud kill the data centre? Jim Chanos thinks so (2022), last accessed 31.10.2022.

[2] Examples include ocean-crossing submarine cables, such as *JUPITER*, connecting the United States, Japan, and the Philippines, owned by a consortium including Amazon Web Services, Meta, NTT, PCCW, PLDT, and Softbank Corp. A transatlantic example, *Havfrue/AEC-2* connects the United States, Ireland, Denmark and Norway and is owned by Aqua Comms, Bulk, Meta, and Google. Both cables became ready for service in 2020. See also Appendix A.2.

to certain infrastructures. The DMA calls for openness and free choice in its pursuit of "fairness" and "contestability" in digital markets.[3] The German competition authority in its report on "Competition 4.0" singles out content delivery networks (CDN) as a piece of Internet infrastructure which has been increasingly used by content firms.[4] However, neither text draws conclusions for the application of competition policy to Internet infrastructure based on economics principles.

This paper presents a theoretical model of vertical integration with a novel demand structure that represents Internet infrastructure as an (upstream) input to (downstream) digital services. A pure upstream firm, the infrastructure firm, invests in infrastructure and sells access to it to a pure downstream firm and to a vertically integrated firm. The latter also expands its available infrastructure through an investment in proprietary infrastructure. The resulting difference in infrastructure between the vertically integrated firm and the pure downstream firm allows the former to sell additional digital services for which it is a monopolist. When the cost structure of the industry results in this proprietary infrastructure being larger than that of the infrastructure firm, both firms' investment incentives jump upwards.

The model explains the incentives for "commoditization" of the infrastructure industry, a common fear in industrial policy, not only in telecommunications but also, for example, in automotive. It furthermore sheds light on the trade-off between contestability (the market share of the pure downstream firm) and efficiency (social welfare). The model suggests to account for the role of infrastructure in the review of vertical mergers in digital markets. It explains the side-payments that have occurred between vertically integrated firms and infrastructure firms such as Netflix and Comcast. Finally, it suggests that expanding net neutrality to certain kinds of infrastructure, such as CDN, would likely harm consumers. I show robustness to assumptions on congestion, the bargaining procedure, and product differentiation.

1.2 Literature

This paper relates to the emerging literature on the economics of Internet infrastructure (Greenstein, 2020). Wilson et al. (2021) analyze the investment decisions of ISP and find long-term effects of investment delays on infrastructure quality. Greenstein and Fang (2020) find that data centers are being built primarily where customers are located, rather than in locations with favorable (land- and energy-)cost structure. Chaturvedi et al. (2021) investigate ISP pricing, in the presence of complementarity between broadband and content. These papers typically focus on the monopoly standing of last-mile ISP with respect to residential connections. By contrast, the infrastructure I am describing appears at an earlier stage in the value chain where big tech firms are vertically integrating.

This paper is closest to Buehler et al. (2004) and Avenali et al. (2014). The former studies investment incentives by an upstream industry when network quality is not verifiable and the downstream (retail) industry is one-sided. They argue that vertical separation enhances incentives to invest in network quality most of the time. One main channel here is the quality sensitivity of retail demand. However, Buehler et al. (2004) study a chain of monopolies and do not consider vertical integration.

Avenali et al. (2014) analyze functional and ownership separation for broadband networks and find ambiguous effects of vertical integration. By contrast, my setup focuses on competing investment by a pure infrastructure firm (for example, Akamai, a CDN operator) and a vertically integrated firm (for example, Google). I do not analyze the last-mile connections in which

[3]The DMA discusses network access in recitals 14 and 51 of the preamble. Article 6(1)(e) proposes an unspecified obligation for "gatekeeper" firms not to restrict the choice of Internet access providers. However, it is not clear how the DMA will treat proprietary networks operated by gatekeepers.

[4]Bundeskartellamt (2016) Working Paper: Market power and platforms [in German]. The authors mention that on-demand server and network services allow small scale entry, while many large firms invest additionally in CDN to reduce response times. The report does not contain conclusions for the competitive assessment of these CDN.

broadband providers have monopoly access[5] but focus on the digital services for which large content firms integrate vertically.

These infrastructure investments improve data management in the presence of net neutrality. Net neutrality is the imposition of zero-termination fees[6] and non-discrimination of data by carriers. Even though net neutrality is controversial and not uniformly enforced (for an early overview of the literature, see Schuett, 2010), it poses economic questions and trade-offs as described by Economides and Tåg (2012) or Greenstein, Peitz, et al. (2016). Current net neutrality regulation is uneven, focusing on ISP while leaving open bypass opportunities and loopholes for cloud services and content providers (Stocker, Smaragdakis, et al., 2020). This paper studies the investments into infrastructure that enables network management practices that arise to cope with the limitations of net neutrality.[7] It contributes to the debate on net neutrality by modeling investment incentives into the network management practices that already exist to adapt to net neutrality restrictions.

My setting where advertisement is served alongside digital services can be seen as a "reduced form" of Armstrong (2006)'s competitive bottleneck in two-sided markets. My paper introduces a true vertical structure with an intermediate input, in contrast with other papers on two-sided markets that call agreements with one market side "vertical" (Lee, 2013; D'Annunzio, 2017; Carroni et al., 2018). The infrastructure firm is not a platform member, but an input supplier. Its inability to commit against opportunistic renegotiation rules out exclusivity. Instead, I show how the infrastructure firm benefits when its network is smaller than the vertically integrated firm's (corresponding to a platform with market power in a two-sided market model). Even though the infrastructure firm loses its strategic role and essentially sells a "commodity", it ends up extracting a greater surplus from the vertically integrated firm.

The phenomena studied in this paper are well-documented in a growing descriptive literature that has been largely ignored by economists. Some large digital services firms have pursued vertical integration strategies through the construction of private backbone networks, edge computing facilities, and owned CDN that improve their ability to expand and change their digital infrastructure to improve the performance and quality of their services (Arnold et al., 2020; Arnold, 2020; Sermpezis et al., 2017; Motamedi et al., 2019). Depending on the business model, private infrastructure can result in cost decreases because of hardware that is fit for purpose or increasing connection quality from faster delivery of data packages at the router of the last-mile ISP.

This paper is a first approach to analyze the effects of this shift in Internet ownership structure. Stocker, Knieps, et al. (2021) document the geographic and virtual dimension of private networks and their implications for firm costs, service quality, and innovation. Lehr et al. (2019) and Balakrishnan et al. (2021) describe the functional disparities between services that rely on the public Internet versus services that are supported by proprietary networks and clouds. Using publicly available data, we add to the description of proprietary networks by showing that increasing numbers and shares of submarine cables have firms including Amazon, Meta, Microsoft, and Google among their owners (see Appendix A.2). By analyzing the previously overlooked competitive effect of a feature of digital markets that mainly concerns the largest digital firms, this paper contributes to the academic debate on regulation and antitrust towards large technology companies (see also Petit, 2020).

[5] Some attempts at competing in this area by content providers, such as Google Fibre, have failed in the past and do not play a large role in the marketplace.

[6] For example, an ISP such as AT&T cannot charge Netflix for traffic that terminates in an AT&T network such as a residential building where Netflix customers live. The price paid by the final consumer to the ISP is understood to compensate the terminating network, no matter where data packages originate.

[7] For example, infrastructure that ensures that one content provider's data packages arrive earlier and in a specific order at a router from where on they are treated on a first-come-first-served basis with the data packages of other content providers. Thus this infrastructure improves speed at the stage preceding the point where the net neutrality principle comes into play. See Easley et al. (2018).

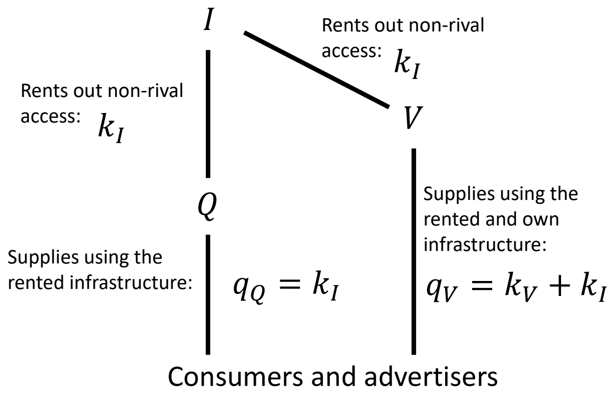


Figure 1: Infrastructure rental and downstream service supply

1.3 Model

This model has two key components:

1. Infrastructure, which is a capital investment either of an infrastructure firm or of a vertically integrated firm. The size of a firm's infrastructure is denoted k. An example are CDN, which are decentralized networks of servers that are used to distribute files, such as large media files, to consumers.

2. Digital services, which are provided to consumers by a digital services firm or by a vertically integrated firm using infrastructure as an input. Quantities of digital services are denoted q. An example of digital services are video streaming services which are offered both by vertically integrated firms and pure digital services firms.

Firms sell digital services (in the "downstream" market) to consumers, generating revenue through these sales and advertisement. To sell digital services, a firm needs infrastructure as an input, either owned or rented (the "upstream" market). Firms can rent infrastructure from an infrastructure firm. In addition, we assume one firm to be vertically integrated and able to build some infrastructure for its own use.

There are three players:

1. A vertically integrated firm (V) that makes a costly investment in infrastructure, purchases access to additional infrastructure from an infrastructure firm, sells digital services to consumers and shows advertisement alongside these digital services. A well-known example would be Google's YouTube, a video-sharing platform.

2. A digital services firm (Q) that purchases infrastructure access from an infrastructure firm, sells digital services to consumers and shows advertisement alongside these digital services. Think of this as a smaller company offering web-based video hosting and sharing, such as Dailymotion.

3. An infrastructure firm (I) that makes a costly investment in proprietary infrastructure and sells access to its infrastructure to V and Q. Think of this as a CDN operator, such as Akamai or Limelight.

The vertical structure is represented in Figure 1.

Downstream there are potentially unlimited consumers, but firms are limited in their ability to sell services by their available infrastructure (either their own or rented infrastructure). For example, as a video-streaming service adds additional infrastructure, its average latency decreases so far that consumer watch one additional unit of video content. Adding one unit of infrastructure allows firms to serve exactly one more unit of services to consumers at zero marginal cost and to show one more unit of advertisement alongside these services.[8]

Firms compete over consumers for services that they both sell, while the firm with more infrastructure has a monopoly for those services that only it can sell. The idea is that the additional infrastructure allows the firm to offer a tiered service that cannot be offered by the firm with less infrastructure, and which is regarded by consumers as a separate product. In the example of video-streaming platforms, this could represent videos at a higher resolution which can only be delivered at acceptable speeds by the firm with the greater infrastructure. Therefore, the firms with more infrastructure can charge consumers for this service while offering a free, advertisement-funded service that is equivalent to the free, advertisement-funded service of its rival. In other words, it puts everything that is better about its product behind a paywall.

This results in a "competitive segment" and a "monopolistic segment" on which firms can set separate prices and face different demand. Firm Q can only choose whether to rent firm I's infrastructure k_I, so its downstream capacity is $q_Q = \{k_I, \emptyset\}$. Firm V can choose to rent access to k_I in addition to its own infrastructure k_V, so it offers $q_V = \{k_I + k_V, k_V\}$. Under this capacity structure, the size of the competitive segment is $min(q_V, q_Q)$ and the size of the monopolistic segment is $max(q_V, q_Q) - min(q_V, q_Q)$.[9] The relationship between available infrastructure and competition for consumers is visualized in Figure 2. In this example, $\bar{q}_Q = k_I$ and $\bar{q}_V = k_I + k_V$. Therefore, Q and V compete Bertrand-style over the first k_I units of demand, while over the remaining k_V, V acts as a monopolist.

In the base model, access to infrastructure is non-rival, so I can sell its infrastructure to both V and Q simultaneously. While infrastructure in reality is capacity-limited (for example, storage space in a data center, capacity on a fibre-optic cable), in the base model we focus on the type of infrastructure that enables new kinds of services (for example, enabling latency-sensitive applications by building additional servers closer to final users) rather than congestion.

In our example, I offers a CDN for storing large media files to video-streaming services. In this example, infrastructure investment would correspond to building additional servers in local networks closer to consumers. The model emphasizes the quality advantage from being able to reach consumers locally (and thus decreasing latency) which enables higher quality services (such as live-streaming, higher video resolutions) for which platforms can charge through premium services (such as YouTube Premium). We consider congestion in a robustness check in Section 1.6.1.

Firms compete in prices and their products (on the competitive segment) are perfect substitutes to consumers. Hence, consumers buy services from the cheapest firm up to their willingness-to-pay a. When two firms set identical prices for their services, they split consumer demand equally. The digital services firm and the vertically integrated firm sell advertisement space alongside services to advertisers at a constant price of r (irrespective of the segment on which the advertisement is displayed).[10]

I's objective is to maximize profits which are the sum of rental transfers t_V, t_Q minus the investment cost $c_I(k_I)$. V's objective is to maximize profits which are the revenue from selling digital services and advertisement downstream, minus the rental transfer t_V and investment cost $c_V(k_V)$. Q maximizes profits by selling digital services and advertisement downstream, minus

[8] Zero marginal cost is a common assumption to focus on pricing in two-sided markets (Hagiu and Lee, 2011; D'Annunzio, 2017).

[9] Firms never have an incentive to hold back capacity, so we do not allow for this possibility. The reason is that in contrast with a standard Cournot model, demand curves are not downward sloping.

[10] This is equivalent to the "competitive bottleneck" configuration (Armstrong and Wright, 2007) in two-sided markets. The pricing is similar to the "per-reader advertising charges" in Armstrong (2006). The downstream market of our model is a simplified version of this two-sided market. A secondary source of revenues enables Q to pay a positive transfer to the infrastructure firm, which is required for competition in equilibrium.

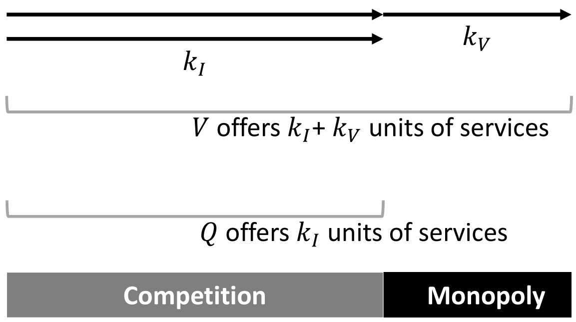


Figure 2: Market segmentation for differently-sized infrastructure

its rental transfer t_Q. Firms' outside options are normalized to zero, and they accept transfers that allow them to at least break even. We indicate the monopolistic and competitive segments by subscript $h = c$ (for competition), m (for monopoly).[11] Firm j's demand on h is $d_{j,h}$ and its price to consumers is $p_{j,h}$. Hence the payoffs:

$$\Pi_I = t_Q + t_V - c_I(k_I) \tag{1}$$

$$\Pi_V = \sum_{h=c,m} d_{V,h}(p_{V,h} + r) - t_V - c_V(k_V) \tag{2}$$

$$\Pi_Q = \sum_{h=c,m} d_{Q,h}(p_{Q,h} + r) - t_Q \tag{3}$$

The cost function for the creation of infrastructure is firm-specific and entails diminishing returns. We consider a convex cost function of the form $c_i(k) = k^\alpha + \beta_i(k - k_0), \alpha > 1, \beta_i > 0$ for some exogenous values $\beta_i, k_0 \in \mathbb{R}_+$, $k \geq 0$, $i = V, Q$.

The timing is as follows: First, V and I simultaneously choose non-negative levels of investment. Second, I simultaneously posts take-it-or-leave-it offers to buy access to k_I at transfers t_V, t_Q to V and Q, respectively. V and Q choose whether to accept or reject these offers. Player I cannot commit not to renegotiate prices opportunistically, so V and Q are accepting prices expecting competition by the other firm.[12] Third, firms set prices to consumers and advertisers. Fourth, consumers and advertisers decide simultaneously from which firm to buy and payoffs are realized.

[11]From the demand structure above it follows that only one firm, either V or Q, can offer services on the monopolistic segment and serve consumers thereon.

[12]More precisely, after V and Q have announced whether to accept the take-it-or-leave-it offer, if any of them rejected the offer, I can make one additional, revised take-it-or-leave-it offer. This prevents I from initially announcing a high transfer that can only be recouped under monopoly, but not under competition, and thereby attempting to create a downstream monopoly. Any firm facing such a transfer would anticipate an opportunistic offer to its rival and reject. In equilibrium V and Q correctly anticipate that the rival will enter. The idea is similar to franchising where the franchiser would like to commit to limit downstream competition but has an incentive to opportunistically award more franchises after a contract has been made (McAfee and Schwartz, 1994). With a similar logic a durable goods monopolist is disciplined by competition from its future self (Bulow, 1982). In our model, upstream competition could alternatively be a second infrastructure firm or a competitive fringe of infrastructure firms but given our demand structure, this is a convenient way of limiting upstream market power.

1.4 Analysis

We start by defining a subgame-perfect Nash equilibrium (SPNE) of this game.

Definition SPNE: A subgame-perfect Nash equilibrium of this game is a profile of prices $p_{j,h}$, transfers t_j $j = V, Q$, rental decisions, and investment levels k_i, $i = I, V$, such that these are optimal in every sub-game.

Now, we state the first main result regarding the types of equilibria of this game (Theorem 1). We state the main propositions regarding the existence of equilibria (Proposition 1) and an important feature of these equilibria (Proposition 2). Then, we state the second main result regarding social welfare in this model (Theorem 2). We then analyze the game, proving Propositions 1 and 2, which together imply Theorem 1, followed by an analysis of social welfare. The proof of Theorem 2 is not interesting in itself, however, so it appears in the Appendix.

Theorem 1: There are two types of equilibrium in which either $k_I > k_V$ or $k_V \geq k_I$. $k_V \geq k_I$ in equilibrium if $\beta_V - \beta_K \leq r/2$. The converse is true otherwise.

First, we show:

Proposition 1: Both I's and V's optimal level of investment are higher in the equilibrium when $k_V \geq k_I$ than in the equilibrium where $k_I > k_K$.

and

Proposition 2: When $\beta_V > \beta_I - \frac{r}{2}$ ($\beta_V < \beta_I - \frac{r}{2}$) there is a unique equilibrium where $k_I > k_V$ ($k_I < k_V$).

Theorem 2: Social welfare is decreasing in marginal costs β_I, β_V but has a discontinuity when $\beta_V - \beta_I = r/2$. Social welfare increases at this point if

$$(a+r)^{\alpha-1}\left[\frac{2a+r}{\alpha}\right] > \left[\frac{2a+r}{\alpha}\right]^{\alpha} + (\beta_I+\beta_V)^{\alpha-1}\left(\frac{a+\frac{r}{2}}{\alpha}\right) \tag{4}$$

and decreases otherwise.

We start by analyzing the game.

Proof of Proposition 1: At the third and final stage, firm $j = V, Q$ choose consumer prices $p_{j,h}$ for $h = c, m$. On the monopoly segment, firm j maximizes profits by setting $p^*_{j,m} = a$. On the competitive segment $p^*_{i,c} = 0$ is the only equilibrium: firms have zero marginal cost and products are perfect substitutes, so Bertrand pricing prevails.[13]

If both firms decide to rent infrastructure access from I (we'll verify next that they do so in equilibrium), the resulting available infrastructure is as in Figure 2 and resulting demand is:

$$d_{Q,c} = d_{V,c} = k_I/2 \tag{5}$$
$$d_{V,m} = k_V \tag{6}$$
$$d_{Q,m} = 0. \tag{7}$$

[13]In spite of the presence of capacities, this is not an example of Bertrand-Edgeworth pricing because in this model, prices are not declining in quantity. Every unit of infrastructure enables provision of an additional unit of services, valued at a constant value a. In this model, lower capacity corresponds to having a slower/less attractive web-service. It seems reasonable to assume that consumers do not have a higher willingness to pay for a worse service, so the assumption of downward sloping demand curves does not apply in this context.

Given these prices and demand, profits are

$$\Pi_Q = k_I(r/2) - t_Q \tag{8}$$
$$\Pi_V = k_I(r/2) + (r+a)k_V - t_V - c_V(k_V). \tag{9}$$

Recall that when setting transfers, I cannot commit not to renegotiate. For example, if Q rejects the initial offer, I can always offer $t'_Q = k_I(r/2)$ which allows Q to break even (anticipating $d_{Q,c} = k_I/2$ and revenues $rd_{Q,c}$) and is accepted. Anticipating that it will face competition for the first k_I units of services, the highest price that V is willing to pay to I is

$$t_V = \frac{r}{2}k_I + k_V(a+r) - k_V\frac{r}{2} - max(0, (k_V - k_I)(r/2 + a)) \tag{10}$$

which can be rewritten as

$$t_V = \begin{cases} (k_I + k_V)\frac{r}{2} + k_V a & \text{if } k_I > k_V \\ k_I(a+r) & \text{if } k_V \geq k_I. \end{cases} \tag{11}$$

Importantly, V's willingness-to-pay for access to k_I depends on whether its own infrastructure is greater than I's. Levels of infrastructure are selected in the first period and are fixed and observed when I makes its offers. When k_I is larger than k_V and V does not rent access to k_I, V can only serve k_V units of digital services and it faces competition from Q on this range. V's revenue in this case is $k_V\frac{r}{2}$, the third term in equation 10. Its willingness-to-pay is therefore the difference between the total revenue it makes with k_I minus this outside option.

If k_V exceeds k_I, V will be able to act as a monopolist for some part of demand $(k_V - k_I)$ even without access to I's infrastructure. At the margin, one additional unit of infrastructure allows V to serve one more unit of digital services at the monopolist's markup $a + r$.

At the first stage, V and I make their investment decisions. We now set up the profit functions and solve for the profit-maximizing level of investment. V's first period problem is

$$\max_{k_V} \Pi_V = d_{V,c}(p_{V,c} + r) + d_{V,m}(p_{V,m} + r) - t_V - c_V(k_V). \tag{12}$$

Now we can write

$$\frac{\partial \Pi_V}{\partial k_V} : a + r - c'_V(k_V) - \frac{\partial t_V}{\partial k_V} = 0 \tag{13}$$
$$\frac{\partial t_V}{\partial k_V} = \begin{cases} \left(\frac{r}{2} + a\right) & \text{if } k_I > k_V \\ 0 & \text{if } k_V \geq k_I \end{cases} \tag{14}$$
$$c'_V(k_V) = \alpha k_V^{\alpha-1} + \beta_V \tag{15}$$

which yields the first order conditions for k_V:

$$\begin{cases} \frac{r}{2} = c'_V(k_V) & \text{if } k_I > k_V \\ a + r = c'_V(k_V) & \text{if } k_V \geq k_I \end{cases} \leftrightarrow \begin{cases} k_V = (\frac{r/2 - \beta_V}{\alpha})^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_V = (\frac{a+r-\beta_V}{\alpha})^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{16}$$

The second order conditions are fulfilled from the convexity of $c_V(k_V)$.

I's first stage problem is

$$\max_{k_I} \Pi_I = t_Q - t_V - c_I(k_I) \tag{17}$$

with

$$\frac{\partial \Pi_I}{\partial k_I}: \quad \frac{r}{2} + \frac{\partial t_V}{\partial k_I} - c_I'(k_I) = 0 \tag{18}$$

$$\frac{\partial t_V}{\partial k_I} = \begin{cases} \frac{r}{2} & \text{if } k_I > k_V \\ a + r & \text{if } k_V \geq k_I. \end{cases} \tag{19}$$

which yields the first order conditions for k_I:

$$\begin{cases} r = c_I'(k_I) & \text{if } k_I > k_V \\ a + (3r/2) = c_I'(k_I) & \text{if } k_V \geq k_I \end{cases} \leftrightarrow \begin{cases} k_I = \left(\frac{r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_I = \left(\frac{a+(3r/2)-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{20}$$

Comparing the upper and lower branches of equations 16, 20, we see that there are two types of equilibrium and that these are characterized by either $k_I > k_V$ or $k_V \geq k_I$. In the former case, investment is lower than in the latter case, as stated in Proposition 1. □

One interpretation of the case $k_I > k_V$ is that I plays a strategic role by "pushing" V into the region where it has market power - without it, V faces tough competition and only reaps low profits. In the second case, k_I does not have such a strategic role: Q's ability to constrain V's pricing power is limited and additional infrastructure at the margin allows V to increase its revenue by $a + r$, independent of whether it makes this investment privately or whether I makes this investment (as long as $k_I < k_V$).

I call the intuition behind the case $k_V > k_I$ "commoditization": I's infrastructure becomes interchangeable with marginal investment for V because it only adds to the monopolistic segment. For this reason, V's willingness-to-pay is higher in this scenario.[14]

In equations 16, 20 we assume implicitly that investment is as in the case of $k_V > k_I$ if the condition holds with equality. This is purely for completeness, there is nothing in the model that tells us that this particular equilibrium will be chosen when the condition holds with equality. The following section shows that there is a set of parameters for which either outcome (the upper or lower branch of these equations) is possible.

1.4.1 Existence of equilibria

So far, we have computed the transfers and prices that maximize firm profits for given levels of investment as well as the optimal levels of investment as a function of model parameters and whether k_I or k_V is larger. Without further restrictions, there is no guarantee that the optimal levels of investment are indeed consistent with the size-ordering of the two infrastructures.

From equations 16 and 20 we see that I (V) will never choose a level of investment greater than 0 if $\beta_I > a + (3r/2)$ ($\beta_V > a + r$). In general, strictly positive levels of investment only arise in equilibrium if $\beta_V < r/2$ and $\beta_I < r$. To investigate only interesting cases, we are going to restrict the parameter space accordingly. Furthermore, for any set of parameters to result in an equilibrium, the parameters also need to result in the correct size ordering of the two infrastructures, i.e., $k_I > k_V$ or the converse must be true when computing equilibrium investment levels. We now show that such equilibria exist.

[14] The idea of commoditization, in the more general sense of converting a product into a standardized and interchangeable input, is common in the discussion of industrial policy surrounding vertical integration of some big tech firms, both for Internet infrastructure firms and also with regards to other industries such as the automotive sector. One concern is whether automotive firms will become mere suppliers of hardware as the value generation shifts to the digital sphere and the data generated by cars and drivers. A major German daily compares the fear of Internet infrastructure firms like Deutsche Telekom to become "dumb tubes", or commoditized suppliers to data-driven consumer-facing platform companies. The corresponding fear of car companies and other manufacturing companies is to become "dumb wheels" or "extended workbenches" as profit shifts to digital platform companies. FAZ, Autobranche im Spagat (2021) [in German], last accessed 31.10.2022.

Proof of Proposition 2: Take the first-order conditions from the upper cases of equations 16 and 20, impose $k_I > k_V$, and some algebra yields

$$\beta_V > \beta_I - \frac{r}{2}. \tag{21}$$

Therefore, given marginal costs that fulfill the above inequality, the infrastructure firm has indeed the larger network. Similarly, comparing the lower cases of equations 16 and 20 and solving for $k_V \geq k_I$ yields

$$\beta_V \leq \beta_I - \frac{r}{2}. \tag{22}$$

Now, setting the upper and lower cases of equations 16 and 20 equal yields:

$$\left(\frac{r/2 - \beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} = \left(\frac{r - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} \quad \leftrightarrow \quad \beta_V = \beta_I - \frac{r}{2} \tag{23}$$

$$\left(\frac{a + r - \beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} = \left(\frac{a + (3r/2) - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} \quad \leftrightarrow \quad \beta_V = \beta_I - \frac{r}{2}. \tag{24}$$

The profit functions are strictly concave ($c''_I(k_I) < 0$, $c''_V(k_V) < 0$), so the first-order conditions are sufficient to describe the unique values of k_I, k_V where Π_I, Π_V attain their global maxima. Also, the equilibrium value of k_I does not depend on k_V and vice versa except through the ordering of size of these two variables. Therefore, when either $\beta_V > \beta_I - \frac{r}{2}$ or $\beta_V < \beta_I - \frac{r}{2}$, the equilibrium is unique. □

It is not surprising that the firm with lower marginal costs will generally have the larger infrastructure. Also the "wedge" of $r/2$ that separates the marginal cost values which equalize investment is not interesting in itself - it is a consequence of the no-congestion-assumption which allows I to charge Q an additional $r/2$ for every unit of k_I added in addition to charging its contribution to V's profit.[15]

There is one combination of parameters where both orderings of the two infrastructure variables are an equilibrium. When $\beta_V = \beta_I - \frac{r}{2}$, both $k_V > k_I$ and $k_I > k_V$ can be optimal. For completeness, I have defined the conditional functions above such that that $k_V > k_I$ will be chosen. However, there is nothing inherent in the model which tells us that this equilibrium will be chosen. This is interesting because it gives us a configuration of parameters for which we can directly compare the two types of equilibrium. We will use this below to compare social welfare for the two types of equilibrium, holding parameters constant.

Figure 3 illustrates equilibrium investment values as a function of β_I for some set of parameters. k_V does not depend on β_I, so it can only take two values for a given set of parameters, represented by horizontal black lines: one corresponding to the case $k_V > k_I$ and one where $k_I > k_V$. We denote V's optimal level of investment k_V^{**} in the former case and k_V^* in the latter case. k_I depends on I's cost parameter β_I as well as other parameters and on which firm has the larger infrastructure, so it is represented by two decreasing functions. Given all other parameter values, a value of β_I admits an equilibrium in which infrastructure levels are k_V^{**}, k_I^{**} (k_V^*, k_I^*) whenever it is true that $k_V^{**} > k_I^{**}$ ($k_I^* > k_V^*$). These regions are marked with solid black lines. When $\beta_I = \beta_V + r/2$, both equilibria are possible.

The model is silent on which equilibrium to expect in case of indifference. However, we are going to compare the two equilibria and show that $k_V > k_I$ dominates the other in terms of firm profits and sometimes on social welfare. The presence of the two different kinds of equilibria is consistent with the increasingly large role played by vertically integrated firms in infrastructure

[15] Accordingly, this wedge decreases in robustness checks (in Section 1.6) that afford less bargaining power to I or that consider congestion.

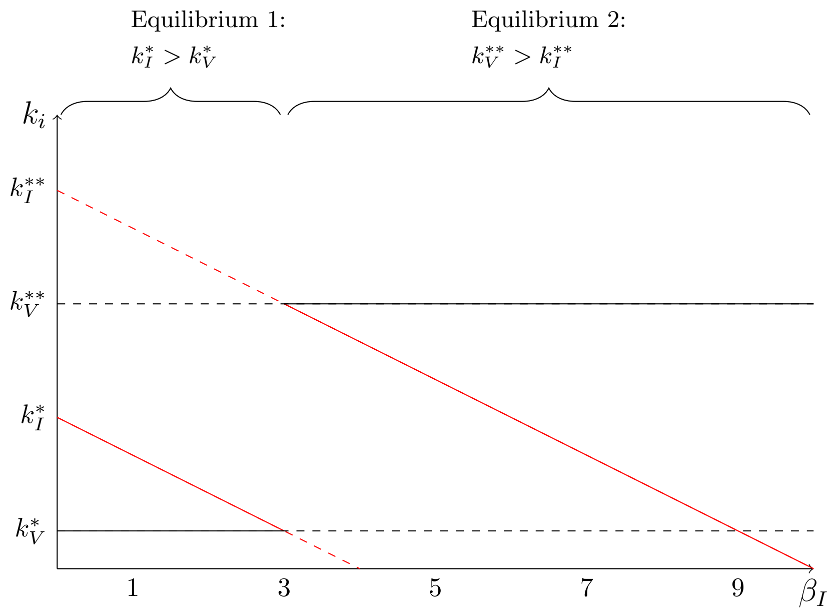


Figure 3: Equilibrium investment decisions as a function of β_I for $\beta_V = 1, a = 4, r = 4, \alpha = 2$.

which is documented in the literature (see Section 2.2) and Appendix 2.3.

1.4.2 Social welfare

We consider whether the equilibria we just found maximize social welfare. We define social welfare as the total sum of surplus minus costs. This implies that both parts of the network are used to supply services. In terms of the game we presented in the previous section, this is the case whenever $q_V = k_I + k_V$. We define social welfare S as follows:

$$S = (k_I + k_V)(a + r) - c_I(k_I) - c_V(k_V) \tag{25}$$

It is easy to show that social welfare is maximized when

$$c_I'(k_I) = c_V'(k_V) = a + r \tag{26}$$

Comparing this optimum to the outcome of the decentralized model, we see that social welfare is generally not maximized. Furthermore, social welfare is a function both of the parameters (higher costs clearly imply lower social welfare) as well as the two different types of equilibrium described in Theorem 1. As mentioned before, we want to compare social welfare when the parameters admit both types of equilibria. This allows us to ignore the levels of marginal cost parameters β_I, β_V to analyze the role of the two different cases where either V or I owns more infrastructure.

Proof of Theorem 2: In the Appendix.

It is unsurprising that social welfare decreases when marginal costs increase. The interesting question is what happens with social welfare when we jump between the two equilibria of the decentralized game that we have identified earlier, either because we have marginal cost parameters that admit both types of equilibria as shown above in Propositions 1 and 2, or if we consider a small change of β_V, β_I such that we move between these two equilibria (while having a negligible direct effect of the change). This proposition shows that the benefit from the increased investment incentives in the $k_V > k_I$-equilibrium dominates additional investment cost if β_I, β_V are low or if α, a and r fulfill certain conditions jointly.

Now compare the socially optimal investment with the Nash-equilibrium investment levels which we have computed in equations 16 and 20. In the case where $k_I > k_V$, we have underinvestment, as $c_I'(k_I) = r$, $c_V'(k_V) = r/2$. In the case where $k_V > k_I$, we have overinvestment as $c_I'(k_I) = a + (3r/2)$, $c_V'(k_V) = a + r$. The reason for overinvestment is that absent congestion, I's private return exceeds the public return from each additional unit of investment, as it gets to charge Q $r/2$ for the additional investment (which correspondingly lowers V's profit by $r/2$ due to competition). However, the net effect on social welfare is positive as long as the condition in equation 4 is fulfilled.

The problem could also be viewed as a social planner's problem who chooses k_V and k_I to maximize social surplus. Here, it is necessary to explicitly state a side-condition that the total sum of demand equals the maximum capacity from the combined networks of V and I. This planner problem can be written as

$$\max_{k_I, k_V} S = (k_I + k_V)(a + r) - c_I(k_I) - c_V(k_V) \tag{27}$$

$$\text{subject to} \sum_{i=V,Q} \sum_{h=c,m} d_{i,h} = k_I + k_V. \tag{28}$$

and it is clearly also maximized where $c_I'(k_I) = c_V'(k_V) = a + r$. The side-condition in equation 28 is also fulfilled in the Nash equilibrium as both Q and V find it optimal to sign the contract with I. This side-condition rules out inefficient configurations for the social planner's problem in which V would not have access to k_I and the total quantity of digital services served downstream would be lower than in the Nash-equilibrium.

The intuition for the result on social welfare comes from the comparison of two competing forces: when comparing the equilibrium where $k_I > k_V$ to the equilibrium where $k_V \geq k_I$, we are moving from a situation in which both k_I and k_V are below the social optimum to a situation where k_I is above and k_V is exactly at the social optimum. Whether social welfare increases then simply depends on whether the increase in k_I is "too much" relative to the increase in k_V.

1.5 Applications

1.5.1 Contestability

How does vertical integration into Internet infrastructure by V impact the "contestability" of the downstream market? Contestability is a policy objective in the regulation of digital markets. The European Commission implicitly defines contestability as the absence of "very high barriers to entry or exit, including high investment costs, which cannot, or not easily, be recuperated in case of exit, and the absence of, or reduced access to, some key inputs in the digital economy, such as data", also referring to the network effects implied by access to more data.[16]

The justification for pursuing contestability is that in its absence, platform markets in particular are prone to "tipping", resulting in entrenched market power, leading to higher costs for consumers and less innovation.[17] A straightforward way to approach this question in the context of this model is to look at how the market shares of V and Q (both on the advertising side and the consumer side) behave between the two equilibria.

Market shares in an antitrust context are typically measured in terms of revenue (rather than profit). Revenue on the advertising side is proportional to the number of connected users for both V and Q. Revenue on the consumer side is nil for Q due to Bertrand competition and $k_V a$ for V due to market power for digital services on the monopolistic segment.

[16]The DMA discusses contestability repeatedly without giving a simple definition. Its meaning is implicitly explained most clearly in recitals 3 and 13 of the preamble.

[17]For example, the preamble of the EU's DMA asserts that "specific features of core platform services make them prone to *tipping*: once a service provider has obtained a certain advantage over rivals or potential challengers in terms of scale or intermediation power, its position may become unassailable and the situation may evolve to the point that it is likely to become durable and entrenched in the near future." (emphasis added), DMA, preamble, paragraph 25.

We denote the market shares of firm j as MS_j. It is measured in terms of revenue in the digital services market across both sides. It can be simplified by dividing by r and substituting infrastructure levels for demand. Now, market share is a function of infrastructure levels k_V, k_I as well as the ratio of the value generated by digital services and advertisement (a/r).

$$MS_Q = \frac{r(k_I/2)}{r(k_I + k_V) + ak_V} = \frac{(k_I/2)}{(k_I + k_V) + (a/r)k_V} \tag{29}$$

$$MS_V = 1 - MS_Q \tag{30}$$

We analyze the change in MS_Q in the case where two equilibria are possible. Denoting the equilibrium where $k_V \geq k_I$ ($k_V < k_I$) with superscript ** (*) we find:

Proposition 3: $MS_Q^{**} - MS_Q^{*}$ is negative and decreasing in r, β_I.

Proof: In the Appendix.

For the cases when we find that social welfare increases at the higher levels of vertical integration that come with the $k_V > k_I$-case, the model illustrates a trade-off: the dynamics of Internet infrastructure and vertical integration that can lead to increased efficiency and greater social welfare come at the cost of lower market share for non-integrated rival platforms. Pursuing the EU's policy goal of contestability in digital markets comes at a cost in terms of efficiency. It is at odds with the equilibrium that arises when V's infrastructure dominates.

High market shares are not harmful to consumers per se but often serve as shortcuts to market power analysis by competition authorities. Under some conditions, such as incomplete capital markets or network effects, high market shares can enhance market power, however, which may justify concerns about the market structure of the downstream market.

1.5.2 Merger: I buys Q

What happens when I is not a pure upstream player but can integrate downstream as well and provide digital services? While there is no notable example of a CDN operator offering consumer-facing services, such as media entertainment, there are traditional Internet infrastructure firms active in downstream consumer markets. For example, Deutsche Telekom operates on multiple levels of Internet infrastructure for business customers (including data centers and cloud computing) but also has media offerings including sports and streaming for final consumers.[18]

To remain as close as possible to the base model, we consider a merger between I and Q, forming a new merged entity, M. Now, the two firms V and M are symmetric except for the potentially different marginal cost variables β_M, β_V. In the first stage, M and V invest in infrastructure k_M, k_V. At the second stage M makes a take-it-or-leave-it-offer to V for access to k_M. Both firms anticipate downstream competition at the third stage. We analyze M's decision to rent its infrastructure to V as well as M's and V's infrastructure investment decisions.

In the base model, a firm never has an incentive to provide less digital services than its available infrastructure allows (as infrastructure costs are sunk at the phase of downstream competition and digital services are provided at zero marginal cost). Now, however, if M offers more digital services downstream while it has the smaller network, it provides additional competition to its rival V which may decrease V's ability to pay high transfers for network access. The equilibrium depends on whether M can commit to a level of downstream services q_M alongside the transfer t_V at the second stage.[19]

[18] See Deutsche Telekom website, for business customers, for media and entertainment [both in German], last accessed 31.10.2022.

[19] While the base model allowed no commitment of I towards V not to renegotiate transfers with Q, it seems

This is described in Proposition 4 which comes in two parts, 4a) describes the case without commitment by M for the quantity q_M, while 4b) describes the case with commitment:

Proposition 4a: If M cannot commit to a level of q_M at the second stage, $q_M = k_M$ in equilibrium. Furthermore,

$$\left.\begin{array}{ll} c'_M(k_M) & = r \\ c'_V(k_V) & = r/2 \\ t_V & = (k_M - k_V)(r/2) + k_V(a+r) \end{array}\right\} \text{ if } \beta_V > \beta_M + r/2$$

and

$$\left.\begin{array}{ll} c'_M(k_M) & = (3r/2) + a \\ c'_V(k_V) & = a + r \\ t_V & = k_M(a+r) \end{array}\right\} \text{ if } \beta_V < \beta_M + r/2$$

Proposition 4b: If M can commit to a level of q_M at the second stage, M offers

$$\left.\begin{array}{ll} t_V & = k_M(a+r) \\ q_M & = k_M \end{array}\right\} \text{ if } \beta_V < \beta_M + r/2$$

at the second stage and first-stage investment is given by $c'(k_M) = a + (3r/2)$ and $c'(k_V) = a + r$.

$$\left.\begin{array}{ll} t_V & = k_M(a+r) \\ q_M & = 0 \end{array}\right\} \text{ if } \beta_V > \beta_M + r/2$$

at the second stage and first-stage investment is given by $c'(k_M) = a + r$ and $c'(k_V) = a + r$. In both cases, if $\beta_V = \beta_M + r/2$, either one of these sets of investment levels and transfers can arise in equilibrium.

Proof: In the Appendix.

The first part of the proposition tells us that without commitment to a quantity level, we obtain the same result as in the base model. As M will always find it attractive at the third stage to offer digital services on its infrastructure and V anticipates this at the second stage, we have similar equilibrium profits and transfers and the condition on the difference in marginal cost that determines the equilibrium outcome is identical.

The second part of the proposition tells us that when the newly merged entity can commit in advance to a downstream strategy, it may find it useful to set $q_M = 0$, or shut down its digital services operations and become a pure infrastructure provider. This is the case when its infrastructure is larger than that of the vertically integrated firm. The intuition is that in this case, it can create a downstream monopoly and participate in its profits by charging monopoly prices for access to k_M. As now the size of the competitive segment is $\emptyset$ and the size of the monopolistic segment is $k_M + k_V$, consumer welfare is 0 and lower than in the base model when $\beta_V > \beta_I + r/2$.

This result does not have novel implications per se, competition authorities already investigate vertical mergers for the ability and incentive of a merged entity to divert business either upstream or downstream. While harmful to consumers, M shutting down its downstream business maximizes social welfare and increases investment (as marginal investment fulfills the conditions of

more plausible that an integrated entity could credibly commit not to compete with its own business customers which is why we check both cases.

equation 26). The proof also involves checking that the foregone downstream revenue is smaller than the additional transfer fee but this is implied by the condition $\beta_V > \beta_M + r/2$.

1.5.3 Efficient side-payments

One question of current interest is the contribution of digital content firms to telecommunication networks.[20] We observe side payments between some large content platforms, such as Netflix, and traffic carriers such as Comcast.[21] This model can explain why such side payments can occur when I's network is only used to a limited extent by Q. There are several reasons why smaller competitors may be limited in their ability to compete with a large digital platform even in the presence of sufficient infrastructure: intellectual property (patents, technology or media content), network effects, brand attraction, or internal technical ability. For example, the lack of exclusive content may limit a video-streaming platform's attractiveness to consumers.

We think of these various limitations as some exogenous limit to the demand served by Q. If Q cannot serve more than some amount $\bar{q}_Q$ of services and V and I can agree, before $t = 1$ of the base game, that V will pay a transfer t'_V (in addition to t_V) to I conditional on choosing a certain value k_I. To make the model interesting, we only consider cases where $\bar{q}_Q$ is binding in equilibrium.

Proposition 5: If Q is capacity-constrained and conditional side-payments are possible and $\beta_V > \beta_I + r/2$, I's and Q's incentives to invest align. Investment levels are given by $c'_V(k_V) = c'_I(k_I) = a + r$.

Proof: When Q is capacity-constrained so that it can serve only a portion of the demand that is smaller than the size of I's network, every unit $k_I > \bar{q}_Q$ can be rented out only to V. V generates a profit of $(a+r)(k_V + k_I - \bar{q}_Q) + r/2\bar{q}_Q - k_V(k_V) - t'_V - t_V$. In equilibrium, I and V can jointly maximize their surplus by setting a value of k_I such that $c'_I(k_I) = a + r$. Call this value k_I^{**} and k_I^* the equilibrium level of I's investment. Then this can be implemented and is incentive compatible when V offers any value of t'_V such that $(k_I^{**} - k_I^*)(a + r) > t'_V > c_I(k_I^{**}) - c_I(k_I^*)$ if and only if $k_I = k_I^{**}$. This value exists and is positive from the fact that $c_I(k_I)$ is increasing and convex and its marginal value is $a + r$ only at $c_I(k_I^{**})$ and lower at any lower level. □

This result illustrates that in the base model, competition from Q serves as a friction that prevents I and V from maximizing their joint surplus. Absent this constraint, the value k_I that maximizes total surplus is simply the one that equates marginal cost and revenue. An enforceable contract on the choice of k_I is a simple mechanism to implement the efficient investment levels k_V, k_I.

1.5.4 Net neutrality

Net neutrality is the absence of termination charges, i.e., content providers do not pay those Internet service providers (ISP) into whose network their traffic is delivered. According to other definitions, net neutrality does not allow ISP to offer a "fast lane" in exchange for payment.

The model only translates to this setup approximately as I does not represent an ISP operating last-mile-networks (we generally do not observe competition at this stage due to fixed costs that imply a natural monopoly). However, there is some discussion as to whether certain infrastructure operators should follow the same rules as ISP. For example, in the US CDN managed to be exempted from net neutrality rules by lobbying.[22]

[20] See Reuters, EU's Vestager assessing if tech giants should share telecoms network costs (2022), last accessed 01.05.2023.

[21] See Wall Street Journal, Netflix to Pay Comcast for Smoother Streaming (2014), last accessed 01.05.2023.

[22] See the FCC's NPRM in 2014 which introduces their thinking on CDN. Protecting and Promoting the Open Internet NPRM (2014). Akamai and other CDNs lobbied to FCC to amend the Open Internet Order which would exempt them from regulation See filings received by Akamai between 02.02.2014 and 31.03.2015. This was

Consider a net neutrality scenario in which I is constrained to charge a single price $t_Q = t_V = t$. Then, access to this infrastructure can be described as a price posted by the upstream industry instead of the bilateral bargaining of the base model. More precisely, after investment decisions in k_V, k_I have been made, I posts a price under which platforms can purchase non-rival access to k_I. V and Q decide simultaneously whether to pay the price, and finally downstream competition takes place. Here, we do not consider renegotiation of the transfer t (as there is only a single profit-maximizing price, renegotiation makes no difference in this case).

Proposition 6: Under net neutrality, I charges $t = k_I(a + r)$ and chooses $c'(k_I) = a + r$. In equilibrium, Q chooses not to pay this price and $q_Q = \{\emptyset\}$, $q_V = k_V + k_Q$. $p_{V,m} = a$, $d_{V,m} = q_Q$.

Proof: In the Appendix.

Net neutrality for infrastructure providers thus harms downstream competition as Q is essentially excluded from the market and consumers pay higher prices due to the absence of a competitive segment. The intuition behind this result is that I has no interest in downstream competition as it reduces industry profit. There is no equilibrium where Q accepts the offer and V rejects because in this case, V's best response would be to invest at least up to the point where $c'_V(k_V) = r/2$. However, due to the presence of competition, Q cannot operate profitably when $t = k_I(a+r)$. The resulting outcome maximizes social welfare as the resulting equilibrium levels of investment fulfill the conditions of equation 26.

Net neutrality in this model unravels the assumption that I cannot commit not to renegotiate and sell to a second firm, limiting its power to create a downstream monopoly. Other papers show similarly how the effects of net neutrality hinge on an infrastructure provider's ability to commit to its pricing (see Footnote 7 of Schuett, 2010) or alternatively on pricing commitment by the downstream industry (Greenstein, Peitz, et al., 2016).

1.6 Robustness

1.6.1 Congestion

So far, we have abstracted from the possibility of congestion. Each unit of infrastructure could be used as a non-rival input in the provision of digital services by both V and I. This most accurately describes investments that enhance capabilities rather than capacity. For example, in a network that does not suffer from congestion, additional investment may represent new network nodes that reduce the average (physical) distance that data packages travel between digital services firms and consumers. In such a case, it may be reasonable to abstract away from congestion and to focus on the role of the network in increasing demand through enhanced quality.

In reality, there are few examples of infrastructure that are truly non-rival, and the aspect of congestion may just be more or less important. To reflect this, we consider a robustness check in which returns on the competitive segment of the market are decreased. We introduce a *congestion factor* $1/2 \leq \phi < 1$. We consider adjusted demand functions on the competitive segment:

$$d'_{V,c} = \phi d_{V,c} \quad \text{and} \quad d'_{Q,c} = \phi d_{Q,c}. \tag{31}$$

The closer ϕ is to 1, the smaller is the role played by congestion and for $\lim_{\phi \to 1}$, we are back in the base model. As ϕ moves closer to 1/2, congestion reduces the returns on the competitive segment.

reflected in the final order. FCC Releases Open Internet Order (2015). All links last accessed 17.01.2023.

Simultaneous use of infrastructure (i.e., on the competitive but not on the monopolistic segment), decreases the quantity of digital services the lower is ϕ. For example, the time that videos buffer before playing increases or the time that websites take to load and consumers switch away.[23] As firms compete Bertrand-style on the competitive segment, this results in lower advertisement revenues.

Proposition 7: Under congestion, investment is given by

$$\begin{cases} k_V = \left(\frac{(r\phi/2)-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_V = \left(\frac{a+r-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \quad \text{and} \quad \begin{cases} k_I = \left(\frac{r\phi-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_I = \left(\frac{a+(1+\phi/2)r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{32}$$

Proof: In the Appendix.

The overall impact of congestion on the model is low. Investment in the case $k_I > k_V$ is decreased overall for both V and I. The part of the equations for k_I, k_V related to revenues is multiplied by ϕ. When I has the larger network, the transfer it gets from V reflects that marginal investment by V allows V not only to sell more services but also to escape the “congestion tax” (the term $k_V(1-\phi/2)r$ in equation 65 is greater than the corresponding term $k_V(r/2)$ in the base model, equation 11). Interestingly, in the case where $k_V \geq k_I$, investment by V is unaffected and investment by I is reduced to a lesser degree: only the part of the expression that pertains to the transfer paid by Q is multiplied by ϕ in equation 63. The intuition is that the transfer t_V is solely determined by the returns of V on the monopolistic segment which are unaffected by congestion.

1.6.2 Nash-in-Nash bargaining

In the preceding analysis, we have not focused on surplus division between I on the one hand and V and Q on the other hand. In the base model, I proposes a take-it-or-leave-it offer that allows it to extract the entire surplus from the rental of k_I, only being disciplined by future competition from itself, due to the lack of commitment. Focusing on revenues instead of profits, for example in the analysis of market shares in Section 1.5.2, has allowed us to side-step this issue.

This assumption is less credible the more we think of the upstream industry as competitive. Nevertheless, intuitively it is not clear that limiting I's ability to extract surplus from the transaction should fundamentally change the model, rather than modifying the marginal-revenue conditions that determine equilibrium investment. Having introduced a source of interdependence of bargaining outcomes in the form of congestion, it seems desirable, however, to study the effect of an alternative surplus-division rule on the outcome of the model.

The Nash-in-Nash bargaining framework seems to be the most reasonable alternative surplus division rule for our model. The model satisfies the conditions of weak conditional decreasing marginal contributions, feasibility, and gains from trade posed by Collard-Wexler et al. (2019). Alternative frameworks to intensify competition upstream, such as entry from a competitive fringe or adding one or several additional upstream firms, would require additional assumptions on the timing of investment and combinatorics for the now exponentially increased number of constellations in which V and Q can have access to different subsets of networks. These

[23] A leaked e-mail by Mark Zuckerberg about Facebook from February 14, 2008 is instructive about the relationship between transmission quality and the quantity of services consumed: “We have a lot of stats that show that usage of the site is basically tied to how fast the site is. The faster we make the site, the more activity we see. I believe the latest data I saw was that if we made the site 100ms faster we'd have about 3% more activity and if we made the site a second faster we'd have about 20% more activity. That's a really big deal. What it means is that even if users don't consciously notice the speed, it's subconsciously making them do fewer pager views and less activity.” The Zuckerberg Files, “Six4Three v. Facebook sealed exhibits”, last accessed 27.02.2023.

assumptions may drive results in addition to any potential effect from alternative modes of surplus division, making the model less tractable.

Again we consider the case with congestion, in which demand on the competitive segment is modified by a congestion parameter $1/2 \leq \phi < 1$. In addition, we define bargaining weights $0 < \delta_j < 1$, $j = V, Q$ which represent the share of the surplus that firm I can extract in the negotiation with V and Q, respectively. The impact of these changes on the model is that we now assume agreements in period 2 to occur immediately and transfers t_Q, t_V to be given by the following expressions:

$$t_Q = \delta_Q \phi k_I (r/2) \tag{33}$$

$$t_V = \begin{cases} \delta_V (k_I \frac{r\phi}{2} + k_V (a + r(1 - \frac{\phi}{2}))) & \text{if } k_I > k_V \\ \delta_V k_I (a + r) & \text{if } k_V \geq k_I \end{cases} \tag{34}$$

Proposition 8: In the case with congestion and Nash-in-Nash bargaining, equilibrium investments are given by

$$\begin{cases} k_V = \left(\frac{(1-\delta_V)a + r(1-\delta_V(1+(\phi/2))) - \beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} \\ k_V = \left(\frac{a+r-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} \end{cases} \quad \begin{cases} k_I = \left(\frac{(\delta_V+\delta_Q)\frac{r\phi}{2} - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_I = \left(\frac{\delta_V(a+r) + (\delta_Q \phi r/2) - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{35}$$

Proof: In the Appendix.

V's marginal investment when $k_V \geq k_I$ does not depend on δ_V. When $k_I > k_V$, however, V's investment now also depends on a as I no longer extracts the whole surplus on this margin. The new expressions reinforce the intuition about the cases in which V's profit drives investment by V and I, respectively. Unsurprisingly, I is now left with lower investment incentives although cases in which I over-invests are still possible (but ruled out in cases where $\delta_V + \frac{\delta_Q \phi}{2} \leq 1$).

Overall, modeling surplus division via Nash-in-Nash bargaining makes the model richer as it highlights the cases in which individual parameters do (a for V in the case $k_I > k_V$) or don't (δ_V for V's investment in the case $k_V \geq k_I$) matter. It does not fundamentally change the model however, reassuring us that the upstream monopoly is a relatively innocent assumption.

1.6.3 Product differentiation

Finally, we are interested how the introduction of product differentiation impacts the model. Until now, V and Q have offered undifferentiated services in the downstream market. On the competitive segment, consumers view their offers as perfect substitutes. Now, we consider that consumers have an innate preference for the products offered by one firm or the other. This may be due to personal preference, for example, as a result of branding, or differentiated product offering, such as exclusive content in the case of video-streaming platforms.

We follow the well-known framework of Shubik and Levitan (1980) for a differentiated goods model. This framework has the advantage that the total market size is unaffected by the number of products or their degree of substitutability. We have already considered a modification of total market size through congestion which shrinks the total amount of services sold on the competitive segment of the downstream market. We maintain both congestion and Nash-in-Nash bargaining as in the previous two subsections.

Denoting the degree of substitutability $\mu \in [0, \infty]$ for $n = 2$ products (digital services by firms V and Q, respectively), following the notation of Motta (2004) with a as the demand scaling parameter, we write indirect demand for $i = V, Q$ as

$$p_{i,c} = a - \frac{1}{1+\mu}\left(nq_i + \mu\sum_{j=1}^{n} q_j\right), \tag{36}$$

and direct demand as

$$d_{i,c} = \frac{1}{n}\left[a - p_i(1+\mu) + \frac{\mu}{n}\sum_{j=1}^{n} p_j\right]. \tag{37}$$

We leave demand on the monopolistic segment unaffected. On the competitive segment, price competition will now not generally result in marginal-cost pricing because a firm that increases its price above marginal cost will still face positive demand. However, due to the presence of advertisement revenues r, it is not clear that this price increase will be profitable. It could be that the marginal loss of advertisement revenue, which is linear in $q_{i,c}$, outweighs the marginal gain from a price increase for any price above marginal cost.

Under price competition for two firms and interpreting r as a negative marginal cost (as in Gans, 2022), equilibrium prices are

$$p_{i,c} = \frac{2a - r(2+\mu)}{4+\mu} \tag{38}$$

First, note that if advertisement revenues are too large, firms will indeed lower their price below marginal cost (0). As the model only allows non-negative prices, to rule out this case, we impose

$$r < \frac{2a}{2+\mu}. \tag{39}$$

Proposition 9: Equilibrium investments are now given by

$$\begin{cases} k_I &= \left(\frac{(\delta_Q+\delta_V)\frac{\phi}{2}\left(\frac{2a+2r}{4+\mu}\right)-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} \\ k_I &= \left(\frac{\delta_Q\frac{\phi(a+r)}{4+\mu}+\delta_V\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right)}{\alpha}\right) \end{cases} \begin{cases} k_V &= \left(\frac{a+r-\delta_V\left(\frac{4+(a+r)(\mu-2\phi)}{4+\mu}\right)-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} \text{ if } k_I > k_V \\ k_V &= \left(\frac{a+r-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} \text{ if } k_V \geq k_I \end{cases} \tag{40}$$

The $k_I > k_V$-equilibrium arises if

$$\beta_I \geq \beta_V + \delta_Q\left(\frac{\phi(a+r)}{4+\mu}\right) + \delta_V\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right) - (a+r) \tag{41}$$

and the $k_V \geq k_I$-equilibrium arises otherwise.

Proof: In the Appendix.

The resulting expressions for equilibrium investment levels and the parameter thresholds that give rise to the different equilibria are now more complicated but have some interesting properties. The new parameter μ does not merely shift all outcomes. Product differentiation always increases the share of the surplus that I can capture when $k_I > k_V$. In the case $k_V \geq k_I$, it also increases t_Q but decreases (increases) the share captured by I from V if $a + r >$ (<) 1. Notably, investment by V is still at the efficient level if $k_V \geq k_I$. When $k_I > k_V$, the deviation from the efficient level of k_V is governed by the share of surplus captured by V and only this part is affected by congestion (the intuition for which was given in the previous section) and product differentiation.

1.6.4 Discussion

The key prediction of the model is a change in the relationship between the traditional carriers of data on the Internet and vertically integrated platform companies. Depending on the role that the upstream firm plays for the vertically integrated firm's value creation, we observe a drastic, non-continuous change in investment incentives. I argue that this mechanism points to issues beyond digital services. Collaborations between digital platform companies and other industries, such as the automotive industry, have led to fears about the future focus of value creation. This model shows that the shift in outside options through vertical integration can have a disruptive impact on an industry.

This model is intended as a first attempt to analyze the questions posed by the vertical integration of digital platforms through proprietary infrastructures. Both competition in downstream markets as well as Internet infrastructure are complex and technical issues, and sector-specific regulation differs between Europe, North America, and other regions of the world. As such, it is not the purpose of this model to predict exactly the behavior and contracts that will arise in any specific geographic or product market. Instead, the model illustrates key features of proprietary Internet infrastructure: the potential efficiency and new goods and services provided by big tech investment, but also the interaction in the market place with smaller players which can be harmed by well-intended regulation or marginalized through increased efficiency.

The setup is standard except for the way that upstream investment determines the size of the downstream market. Intuitively, this one-dimensional measure of consumer demand can be read either as the intensive margin of demand (existing customers demanding additional services as bandwidth increases) or the extensive margin of demand (new customers are won as networks improve and bandwidth increases). For a more stringent exposition, I have focused only on the first interpretation in this paper but note that investment can also affect the extensive margin, especially for new services.

This is a stylized model of a complex environment. This model illustrates the incentives for a large digital services company to invest in proprietary Internet infrastructure. Even when applying this model to an example as specific as video-on-demand streaming, we are still folding several kinds of differentiated products downstream and different kinds of infrastructure upstream into a simple framework that relates the amount of the available infrastructure to the effective quantity of services that can be provided.

The issue of downstream foreclosure (I refusing to deal with a downstream firm) is sidestepped by limiting I's ability to commit to a deal for monopoly. While this is partially motivated by the existence of upstream competition (and absent upstream competition, the essential facilities doctrine in competition law can limit anticompetitive foreclosure), a more economic formulation could point out that two forces in the model drive the incentive for foreclosure: congestion makes downstream foreclosure more attractive, as letting a second firm compete downstream limits the revenue I can extract from the first firm, while product differentiation points in the exact opposite direction: allowing a second firm in expands the downstream market. For the sake of exposition, I chose to move both of these effects into robustness checks but it seems worth pointing out that foreclosure can also be controlled through appropriate assumptions on the relative strength of these forces.

What is k? We are looking at proprietary networks that consist of many different parts. For the purpose of this model, we are not interested, for example, what share of investment goes towards data centers vs. submarine cables. Instead, we are interested in the service improvement that can be purchased at a given price. Therefore, infrastructure k can be thought of as a measure of real quality gain from a given level of investment. This only requires downstream demand to be at least somewhat elastic to the quality improvement induced by investment which seems a reasonable assumption (see also Footnote 23).

The 1:1 ratio between investment and downstream capacity is a free variable, as even relatively inelastic consumer demand can be expressed in the slope of $c_i(k_i)$. A steep investment cost function means that it is very costly to expand the market through further investment, and a low elasticity of demand with respect to quality improvements implies that large investments are needed to expand downstream demand. The main assumptions on investment in the model are therefore that investment costs are convex (increasing demand through additional infrastructure becomes more expensive as efficient and low-cost investment opportunities are realized) and potentially different for I and V.

The model is balanced on a "knive's edge" as for given β_I there is only one value of β_V that gives rise to the two types of equilibria that we compare in our analysis. The key mechanism, also throughout the robustness checks, is the change in the outside option. A more elegant but much more complicated approach would assume that the own-price elasticity of demand differs along the dimension of q_i independent of whether a particular level of infrastructure is reached by one firm or many, and potentially decreases at higher levels of q_i. This would also generate higher markups in the area where the vertically integrated firm does not face competition. This kind of demand could correspond to markets where infrastructure supports innovative and novel services that have fewer alternatives than basic ones.

This alternative formulation might be useful to model the entire ecosystem of a large digital services company, where the range of services offered through increasingly sophisticated infrastructure could range from e-mail, over online search, and real-time virtual realities (a "Metaverse") to future technologies. Such a demand structure would, however, assume the existence of a wider market with substitutes of varying qualities for all kinds of services offered by such a firm. For a tighter exposition and a more specific example, I choose to concentrate on a market in which the consumers' choice of options is always well-defined, as each segment of demand represents a product that is offered either by a monopolist or a Bertrand oligopolist and an outside option of 0.

While it seems natural to think about vertical separation, a counterpart to the forward integration considered in Section 1.5.2, it is not a case that I analyze. The technical reason is that with two independent upstream networks that both downstream players could access, the number of possible permutations of network configurations increases from 4 to 16 (as each downstream firm can have access to either no network, network one, network two, or both networks). Not only does this make the model a lot less tractable but it also requires additional assumptions on the timing and commitment regarding the access of downstream firms to particular networks. More practically, vertical separation in this context seems a less credible policy than in the natural monopoly applications where it normally arises (railroads, electricity networks, other utilities). It seems harder to justify punishing a firm for improving its IT. Moreover, since the defining characteristic of proprietary networks is their specialization to the business and demand of a particular firm, the value of V's infrastructure would be likely diminished or completely lost upon separation.

The most promising avenue for future research appears to be innovation in proprietary networks: While the motivation to study proprietary networks is partly also the ability of platforms to steer innovation in their ecosystems, innovation is not an explicit feature of the model. In part, the increased demand as a result of increased investment can be understood as demand for innovative services that only become feasible with increased infrastructure. However, a practical concern around proprietary networks owned by big tech firms (e.g., deployment of software products through the cloud services of firms such as Microsoft, Amazon, Google) is that innovation might shift into their "walled gardens", decreasing overall benefits from innovation by limiting spillover and network effects.

1.7 Conclusion

The Internet has affected the global economy on many levels. It has enabled some platform businesses to grow to spectacular size in a short amount of time. Understanding the economics underpinning its infrastructure is key to successful economic policy and regulation. In particular, an effects-based assessment of regulation, potential anti-competitive conduct and merger review needs economic guidance. This paper illustrates the economic effects of the increasing vertical integration into Internet infrastructure by digital platform companies.

The model illustrates investment incentives for Internet infrastructure that impacts competition in a digital market downstream. I show that investment incentives increase both for a pure upstream player and a vertically integrated firm when the latter owns more infrastructure. The intuition is that in this case, the vertically integrated firm has market power over the additional demand generated by its investment even absent the additional infrastructure from the infrastructure firm.

Marginal investment by the infrastructure firm increases the total surplus to be shared between it and the vertically integrated firm. As each unit of upstream investment results in a constant increase in revenue for the vertically integrated firm, the infrastructure of the infrastructure firm becomes fully commoditized. An important exception is the case when the infrastructure firm has lower investment costs and fringe players cannot compete with the vertically integrated firm over the whole range of services, for example because of a lack of technical ability, patents or exclusive content. In this case, the vertically integrated firm may prefer to subsidize investment by the infrastructure firm directly. At least in this model, there is no market failure and no need for intervention for this to happen.

As a consequence, this model explains different aspects of the rise of private, proprietary infrastructure. I predict that this additional investment is socially desirable but might increase the market shares of the largest companies. High market shares are not problematic per se but often serve as shortcuts to market power analysis by competition authorities. Under specific circumstances, such as incomplete capital markets, present strength can beget future strength, however, justifying a concern about the market structure of the downstream market.

Expanding net neutrality, a pet policy of some Internet activists, to the infrastructure firms currently not covered, may lead to exclusion of rivals and harm to consumers. Entry downstream by the upstream firm may also have the surprising effect of reducing the amount of downstream services whenever the newly vertically integrated firm prefers infrastructure separation, yielding the inefficient outcome that does not occur in the equilibrium of the base model.

Caution is advised before drawing policy conclusions from a literal reading of the model. It is understood that this model is not a full simulation of any particular downstream market with its generic features such as paid-for premium services and advertisement revenues. Nor does it necessarily describe the market structure for the upstream industry. Instead I purposefully aggregate infrastructure investment into a black box variable to study the effect of quality-improving infrastructure investment. Intervention in any particular market would need to carefully evaluate the sources of revenue and business models of firms in the downstream market and to identify the most important components of Internet infrastructure related to an industry.

Nevertheless, this paper can help policy makers and enforcers ask the right questions both for a competitive analysis of a digital market and for a forward-looking market investigation: First, it describes how even efficient and increasing investment in proprietary and public networks can enhance the unequal footing on which vertically integrated firms and smaller rivals compete. Second, it illustrates the kingmaker role of third-party infrastructure providers, especially when they become active downstream themselves.

Finally, the model points at questions beyond digital services. Large technology firms have begun vertical integration in other fields, including automotive, where questions about the future focus of value creation have also been asked. The model allows for many rich expansions as discussed above. In addition, the analysis can be expanded by appropriate data to test model predictions empirically.

Acknowledgments

Funding by the German Academic Exchange Service, the International Center for Law and Economics, and Lear is gratefully acknowledged. The author thanks Giacomo Calzolari, David K. Levine, Zeinab Aboutalebi, Pierre Regibeau, Volker Stocker, Sean Ennis, Patrick Rey as well as participants at the DG COMP Chief Economist Team internal seminar, CLEEN 2022, ITS 2022, CRESSE 2022, ICLE Market Structure Roundtable 2022, Lear Competition Festival 2022, the 31st Conference on Postal and delivery economics, EARIE 2023 and at EUI for helpful comments and discussion. All errors and views are my own.

References

Armstrong, Mark (2006). “Competition in two-sided markets”. en. In: *The RAND Journal of Economics* 37.3, pp. 668–691. DOI: 10.1111/j.1756-2171.2006.tb00037.x.

Armstrong, Mark and Julian Wright (2007). “Two-sided Markets, Competitive Bottlenecks and Exclusive Contracts”. en. In: *Economic Theory* 32.2, pp. 353–380. DOI: 10.1007/s00199-006-0114-6.

Arnold, Todd (2020). “Understanding Cloud Network Performance”. PhD Thesis. Columbia University.

Arnold, Todd et al. (2020). “Cloud provider connectivity in the flat internet”. In: *Proceedings of the ACM Internet Measurement Conference*, pp. 230–246.

Avenali, Alessandro, Giorgio Matteucci, and Pierfrancesco Reverberi (2014). “Broadband investment and welfare under functional and ownership separation”. en. In: *Information Economics and Policy* 28, pp. 70–82. DOI: 10.1016/j.infoecopol.2014.07.003.

Balakrishnan, Hari et al. (2021). “Revitalizing the public internet by making it extensible”. en. In: *ACM SIGCOMM Computer Communication Review* 51.2, pp. 18–24. DOI: 10.1145/3464994.3464998.

Buehler, Stefan, Armin Schmutzler, and Men-Andri Benz (2004). “Infrastructure quality in deregulated industries: is there an underinvestment problem?” en. In: *International Journal of Industrial Organization* 22.2, pp. 253–267. DOI: 10.1016/j.ijindorg.2003.07.004.

Bulow, Jeremy I. (1982). “Durable-Goods Monopolists”. en. In: *Journal of Political Economy* 90.2, pp. 314–332. DOI: 10.1086/261058.

Carroni, Elias, Leonardo Madio, and Shiva Shekhar (2018). “Superstars in Two-Sided Markets: Exclusives or Not?” en. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3243777.

Chaturvedi, Rakesh, Souvik Dutta, and Kiriti Kanjilal (2021). “An economic model of the last-mile internet”. en. In: *Journal of Economic Behavior & Organization* 191, pp. 620–638. DOI: 10.1016/j.jebo.2021.09.025.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S. Lee (2019). "“Nash-in-Nash” Bargaining: A Microfoundation for Applied Work”. en. In: *Journal of Political Economy* 127.1, pp. 163–195. DOI: 10.1086/700729.

D’Annunzio, Anna (2017). “Vertical integration in the TV market: Exclusive provision and program quality”. en. In: *International Journal of Industrial Organization* 53, pp. 114–144. DOI: 10.1016/j.ijindorg.2017.05.002.

Easley, Robert F., Hong Guo, and Jan Krämer (2018). “Research Commentary—From Net Neutrality to Data Neutrality: A Techno-Economic Framework and Research Agenda”. en. In: *Information Systems Research* 29.2, pp. 253–272. DOI: 10.1287/isre.2017.0740.

Economides, Nicholas and Joacim Tåg (2012). “Network neutrality on the Internet: A two-sided market analysis”. en. In: *Information Economics and Policy* 24.2, pp. 91–104. DOI: 10.1016/j.infoecopol.2012.01.001.

Gans, Joshua S. (2022). "The Specialness of Zero". en. In: *The Journal of Law and Economics* 65.1, pp. 157–176. DOI: 10.1086/714971.

Greenstein, Shane (2020). "The Basic Economics of Internet Infrastructure". en. In: *Journal of Economic Perspectives* 34.2, pp. 192–214. DOI: 10.1257/jep.34.2.192.

Greenstein, Shane and Tommy Pan Fang (2020). "Where the Cloud Rests: The Economic Geography of Data Centers". Harvard Business School.

Greenstein, Shane, Martin Peitz, and Tommaso Valletti (2016). "Net Neutrality: A Fast Lane to Understanding the Trade-offs". en. In: *Journal of Economic Perspectives* 30.2, pp. 127–150. DOI: 10.1257/jep.30.2.127.

Hagiu, Andrei and Robin S. Lee (2011). "Exclusivity and Control". en. In: *Journal of Economics & Management Strategy* 20.3, pp. 679–708. DOI: 10.1111/j.1530-9134.2011.00302.x.

Lee, Robin S. (2013). "Vertical Integration and Exclusivity in Platform and Two-Sided Markets". en. In: *American Economic Review* 103.7, pp. 2960–3000. DOI: 10.1257/aer.103.7.2960.

Lehr et al. (2019). "Whither the Public Internet?" In: *Journal of Information Policy* 9, pp. 1–42. DOI: 10.5325/jinfopoli.9.2019.0001.

McAfee, R. Preston and Marius Schwartz (1994). "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity". In: *The American Economic Review* 84.1, pp. 210–230. DOI: https://www.jstor.org/stable/2117979.

Motamedi, Reza et al. (2019). "On Mapping the Interconnections in Today's Internet". In: *IEEE/ACM Transactions on Networking* 27.5. Conference Name: IEEE/ACM Transactions on Networking, pp. 2056–2070. DOI: 10.1109/TNET.2019.2940369.

Motta, Massimo (2004). *Competition Policy: Theory and Practice*. Cambridge University Press.

Petit, Nicolas (2020). *Big Tech and the Digital Economy: The Moligopoly Scenario*. Oxford University Press.

Schuett, Florian (2010). "Network Neutrality: A Survey of the Economic Literature". en. In: *Review of Network Economics* 9.2. Publisher: De Gruyter. DOI: 10.2202/1446-9022.1224.

Sermpezis, Pavlos, George Nomikos, and Xenofontas Dimitropoulos (2017). "Re-mapping the Internet: Bring the IXPs into Play". In: *arXiv:1706.07323 [cs]*. arXiv: 1706.07323.

Shubik, M. and R. Levitan (1980). *Market structure and innovation*. John Wiley & Sons.

Stocker, Volker, Günter Knieps, and Christoph Dietzel (2021). "The Rise and Evolution of Clouds and Private Networks Internet Interconnection, Ecosystem Fragmentation". In: *SSRN Electronic Journal*. DOI: https://dx.doi.org/10.2139/ssrn.3910108.

Stocker, Volker, Georgios Smaragdakis, and William Lehr (2020). "The State of Network Neutrality Regulation". In: *ACM SIGCOMM Computer Communication Review* 50.1, pp. 45–59.

Wilson, Kyle, Mo Xiao, and Peter F. Orazem (2021). "Entry threat, entry delay, and Internet speed: The timing of the U.S. broadband rollout". en. In: *Journal of Economics & Management Strategy* 30.1, pp. 3–44. DOI: 10.1111/jems.12407.

2 Does buyer discretion facilitate home bias in procurement? Cross-border procurement of medical supplies under Covid-19

Abstract

Public procurement markets are often national despite a general agreement against national preferencing. I exploit shocks occurring during the Covid-19 pandemic to two important factors, *crisis urgency*, measured through local infection rates, and increased *buyer discretion*, to study home bias in public procurement. Two causal difference-in-difference analyses on novel data for medical supplies in Europe show that home bias is not inevitable. An increase in local infection rates by one standard deviation locally increases the share of cross-border procurement by 19.3 percentage points over a baseline of 1.5 percent. Also, deregulation that allowed for buyer discretion caused cross-border procurement to increase by more than 35 percentage points. A simple theoretical model systematizes these findings.

Keywords: public procurement, home bias, regulation, difference-in-differences, Covid-19
JEL: H12, H57, L51

2.1 Introduction

Public sector procurement accounts for a large share of the economy, at 15-20% of gross domestic product in the European Union (EU). Economists have identified sources of misallocation that result in inefficiencies, for example, due to favoritism of public buyers towards firms from their own country (home bias). Home bias can emerge when buyers care not only about the purchase itself, but also about secondary goals, such as boosting local jobs or pursuing political goals. In the EU, persistent home bias also undermines the policy goal of completing the "Single Market". This is in spite of regulation aimed at promoting competition and a level playing field. Instead, most contracts are awarded nationally, fragmenting public procurement markets along national borders.

In this paper, I study home bias using the impact of the early Covid-19 pandemic. Two factors have large effects on home bias in procurement in an emergency: crisis urgency and buyer discretion. These effects are not explained by potential supply-side limitations that may have pushed buyers to purchase abroad. Using the natural experiment that the pandemic poses for certain regions and certain product categories, I draw lessons for procurement beyond the context of the pandemic and argue that the effective lever to reduce home bias is in buyer incentives. However, given that the costs of the pandemic are estimated anywhere between 12.5 trillion USD globally through 2024 (by the IMF)[24] to as much as 16 trillion USD for the US alone (Cutler and Summers, 2020), also avoiding misallocation in the context of emergencies is an important application of these results.

In 2020, the Covid-19 pandemic shocked the global economy on many levels, impacting international supply chains and firm activities. Changing market conditions overturned the conventional wisdom of the economics of procurement. As the crisis intensified, rules which prioritize transparency and competition were abandoned. In some places, public authorities found themselves competing against private sector buyers and each other. In the EU, the European Commission eventually dropped publication and transparency requirements and gave buyers full discretion in their purchasing decisions for certain goods. The focus of the analysis is on this deregulation, together with the effect of crisis urgency.

Potential misallocation in the procurement for medical supplies has come under scrutiny and has garnered high levels of public attention. Anti-corruption activists caution against the risk that vested private interests capture public resources and distort decisions as procurement is

[24] Reuters, 20.01.2022: https://www.reuters.com/business/imf-sees-cost-covid-pandemic-rising-beyond-125-trillion-estimate-2022-01-20/

highly vulnerable to corruption.[25] Media reports uncover problems ranging from low-quality products in Austria[26] and Switzerland[27], over failures to organize distribution[28] and payment of deliveries[29] to even major irregularities in contract awards due to political influence-taking in Germany[30] and Finland[31]. Examples from other countries abound, making it clear that misallocation and inefficiencies did not fully disappear in this crisis situation.

This article contributes to the study of home bias and cross-border procurement: I create a novel data set of procurement contracts for medical supplies. I analyze tender documents published between 2018 and 2020 in Tenders Electronic Daily (TED), an online register of European procurement contracts. The descriptive statistics show that a temporary and large surge in cross-border procurement during the first wave of infections coincides with a period of few competitive tenders and many direct awards. This surge of direct cross-border awards already suggests the presence of a phenomenon other than mutual unawareness of international buyers and sellers that explains home bias.

Among the potential reasons for home bias, the shock provided by the pandemic illustrates two in particular, crisis urgency and buyer discretion. We use two distinct analyses: For crisis urgency, I study the effect of local infection rates on cross-border procurement on a sample of Covid-19-related medical products. This analysis finds that an increase in infection rates by one standard deviation increases the likelihood of a cross-border award for medical supplies by more than 19 percentage points (over a pre-pandemic baseline of 1.5 percent). A country-by-country leave-one-out analysis finds a still sizeable lower bound on the effect size of 9.3 percentage points.

Then, I analyze the lifting of restrictions on buyers purchasing these Covid-19-related medical supplies. Under the new rules, buyers purchased directly from sellers rather than posting calls for tender. To estimate the effect of this regulation change, I draw upon additional data of procurement contracts of medical supplies which were not affected by these new rules but which are still similar in product group and therefore, presumably, similar in their supply characteristics. I compare the share of cross-border procurement for these supplies with medical supplies from related product categories that were unaffected by deregulation using additional data. Following this increase in buyer discretion, the probability of a cross-border award increased by 35.7 percentage points for affected products and services (with a lower bound of 16 percentage points in a country-by-country leave-one-out analysis).

Informed by a simple model of monitoring, I suggest that the observed changes in buyer behavior are caused by the changed incentives and informational advantages of the buyers. Importantly, the empirical results persist in an analysis on a subsample of contracts where there is domestic spare capacity, conservatively approximated as the extensive margin of supply (additional domestic firms standing ready to compete for a contract but not winning it). Therefore, a lack of domestic production capacity alone does not explain the surge in cross-border awards during the pandemic. At the same time, this implies that the low share of cross-border awards outside of the emergency is not driven by a lack of competitive foreign sellers, for example due to potential transportation cost differentials or frictions related to regulatory, language, or cultural barriers.

[25] Transparency International, 29.03.2021: https://www.transparency.org/en/blog/g20-italy-covid-19-recovery-corruption-priorities

[26] Der Standard, 09.12.2020: https://www.derstandard.de/consent/tcf/story/2000122115865/mangelhafte-ffp2-masken-werden-zurueckgeholt-und-beschaffung-rechtlich-geprueft [in German]

[27] blue News, 08.03.2021: https://www.bluewin.ch/de/news/international/armee-tauscht-schutzmasken-der-firma-emix-um-616790.html [in German]

[28] Osnabrücker Zeitung, 05.03.2021: https://www.noz.de/lokales/osnabrueck/artikel/2246556/osnabrueck-loest-corona-materiallager-auf [in German]

[29] Die Zeit 13.06.2020: https://www.zeit.de/politik/deutschland/2020-06/mundschutz-atemschutzmasken-lieferproblem-zahlungsverzug-coronavirus [in German]

[30] Deutsche Welle, 08.03.2021: https://www.dw.com/en/german-lawmaker-resigns-over-face-mask-scandal/a-56798497

[31] Politiikka, 09.04.2020: https://www.hs.fi/politiikka/art-2000006469197.html [in Finnish]

I conclude that home bias can subside temporarily in an emergency. The empirical findings then vindicate the crisis response of lifting restrictions. However, a return to pre-pandemic levels of cross-border procurement in late 2020 shows that both effects were just transitory. Overcoming home bias in procurement therefore requires more permanent policies. If regulation to limit buyer discretion has the straightforward effect of lowering this monitoring cost, one would expect regulation to decrease misallocation. However, the above-mentioned scrutiny and public attention on procurement of medical supplies during the pandemic might have worked in the opposite direction by disciplining buyer behavior.

The rest of the chapter is organized as follows: In Section 2.2, I review the literature on misallocation in procurement. In Section 2.3, I present the data set. I outline the empirical strategy and regression results in Section 2.4. Section 2.5 concludes. A model of monitoring in Appendix B.4 supports the findings.

2.2 Literature

This chapter relates most closely to the empirical literature on the effect of procurement design, in particular buyer discretion, on various outcomes of the procurement process. Coviello et al. (2017) study procurement in Italy with 2000-2005 data and find that increased buyer discretion on average improves the functioning of the procurement system. In contrast, Baltrunaite et al. (2021) find that an increase in bureaucrat discretion in Italy increased the rents of politically well-connected and less efficient firms in data from 2009-2013. This paper is the first to study the economic channels of crisis urgency and increase of buyer discretion on cross-border awards. This analysis is not limited to contracts around specific thresholds and contributes a broad, European overview.

An older strand of the literature studies how misallocation in procurement can occur due to favoritism in auctions (McAfee and McMillan, 1989; Laffont and Tirole, 1991), outright corruption (Burguet, 2017), or “buy national” policies (Cernat and Kutlina-Dimitrova, 2015). Allocation of procurement contracts in the EU is heavily skewed towards domestic firms (Vagstad, 1995), indicating strong discrimination against foreign firms. Based on an empirical analysis of procurement in 29 OECD countries, Hessami (2014) documents misallocation in procurement due to political corruption and rent-seeking behavior in OECD countries.

I call this skew towards domestic firms “home bias”, which need not arise from corruption, but which can also be a deliberate policy choice.[32] Procurement processes can encourage home bias to promote small-and-medium-sized firms, jobs, growth, or innovation. Dynamic considerations can justify home bias economically (Barbosa and Boyer, 2021). Laffont and Tirole (1991) suggest that the share of cross-border awards can serve as a simple screen for favoritism. They specifically analyze collusion of domestic buyers and firms against the government and conclude that under some conditions a “law of large numbers” should hold for awards to the disfavored group. I study this outcome empirically in the context of European medical supplies.

In spite of its richness and accessibility, few articles have used procurement data published on TED. Prier et al. (2018) describe a consolidated data set published by TED for the years 2009-2015. La Cour and Ølykke (2018) find that data completeness differs depending on the country submitting tender information. La Cour and Milhøj (2013) attempt a mainly statistical exploration of the data on Danish contracts. Similar to this paper, Kutlina-Dimitrova and Lakatos (2016) use data for cross-border contract awards in Europe and argue that strong product market regulation may act as a hidden anticompetitive barrier. Carboni et al. (2018) also discuss

[32] The European Commission claims that “the public sector can use procurement to boost jobs, growth and investment, and to create an economy that is more innovative, resource and energy efficient, and socially-inclusive” (https://ec.europa.eu/growth/single-market/public-procurement_en). Loader (2007) writes that “UK government and small firms believe that public sector procurement is a good way of helping to support small business”. See also Loader (2016) on favoring small and medium-sized firms in procurement.

empirical methods to study discrimination of foreign firms in procurement.[33] I discuss the issues raised by previous scholars with regards to the present data set for sample selection and identification. Although missing values are common in the data set, they seem uncorrelated with the award of contracts to foreign or domestic firms. This is the first paper to my knowledge that uses TED data to investigate procurement in the context of the Covid-19 pandemic and creates an original data set from individual contract award notices, rather than using the consolidated data sets which are only published with a lag of several years.

So far, the effect of Covid-19 on procurement has been studied by legal scholars (Lalliot and Yukins, 2020; Sanchez-Graells, 2020) while Hoekman et al. (2021) focus on the trade dimension and the implications for public-private partnerships (PPP) in the global medical industry (on PPP, see Casady and Baxter, 2022; Vecchi et al., 2020). These early responses to the pandemic lay out the policy challenges. By studying empirically the outcomes of procurement for medical supplies during the Covid-19 pandemic, this paper helps to quantify these concerns and to draw lessons for the future.

2.3 Data

This section provides a brief summary of each each data set. Additional details can be found in Appendix B.1.

2.3.1 Data sets used

Procurement data: Contract award notices (CAN) for public procurement 2018-2020 in Europe come from "TED: tenders electronic daily, Supplement to the Official Journal of the EU" (TED).[34] According to its website, "TED publishes 746 thousand procurement award notices a year, including 235 thousand calls for tenders which are worth approximately €545 billion." Each CAN describes the outcome of a tender, including information on the buyer (name, type and location of the authority), the object (total value of the procurement, product category), and, where applicable, the division of the contract into individual lots.

For each lot, the CAN reports the number of companies that bid for the contract (bidders), as well as the number of bidders that are foreign, non-EU, and small-and-medium-sized firms. The bidder winning the lot, the contractor, is listed with its location as well as an initial estimate and final reported value of the award.

The unit of observation is a separate contract award to a contractor. I study the cross-section of awards as there are too few repeat observations for firms and buyers to create a panel. A contract award is defined by a contract date and a contractor and may represent an entire contract or just a contract lot.

Covid-19 infection rates: Regional infection rates come from the European Centre for Disease Prevention and Control (ECDC).[35] These are reported as average cases per 100,000 inhabitants over a 14-day period. Regions are listed by NUTS2-code.

Population data: Population data at national and regional (by NUTS2-code) levels from Eurostat are used for a robustness check to compute average national infection rates (weighted by population) excluding individual regions.

Exchange rates: Non-Euro currencies are converted to Euro using data from the European Central Bank (supplemented in a few cases by online sources described in Appendix B.1).

[33] This is also the subject of a study commissioned by the European Commission using TED data (Ramboll/HTW Chur, 2011).

[34] https://ted.europa.eu

[35] https://qap.ecdc.europa.eu/public/extensions/COVID-19/COVID-19.html

2.3.2 Sample selection

Entries with incomplete data or token values are removed. Especially the largest countries often do not report the total values of procurement contracts or only report token values (such as 1 EUR) which are treated as missing. Of the 295 documents for medical supplies where no total value is reported (out of a total of 8,054 CAN that remain after dropping failed tenders from the original 9,322 CAN), 75 percent have buyers from Germany, France, and the UK. This is consistent with differences in reporting described by previous researchers (see Section 2.2).

The sample is further restricted to contract awards by national governments and public authorities (excluding EU bodies) from 2018 - 2020 with contract values reported. Thus, I obtain a regression data set of 67,638 observations of individual awards for medical supplies.[36] These contracts were designated by TED as related to the Covid-19 pandemic and are the subject of a comparison of cross-border procurement across regions by their infection rates. For the analysis of the deregulation, I identify another 252,575 observations which are used as the control group in Section 2.4.3. These products are identified from belonging to the same larger product groups and are mostly medical in nature as well. They were, in contrast with the previous group, not identified as Covid-relevant. For example, the former data set includes "antiseptics and disinfectants" which are Covid-relevant while, based on the CPV hierarchy level, "antipsoriatics" are not Covid-related and enter the control group. Additional examples are listed in Table 13.

2.3.3 Descriptive statistics

Table 1 presents summary statistics for Covid-19-related tenders. Further descriptive statistics, including on the control group data, are in Appendix B.2. The most important findings of this section are that the value of cross-border awards increases from 1.5% on average pre-pandemic to more than 50% between April and June 2020 and that a majority of tenders in April and May 2020 were direct awards to foreign sellers.

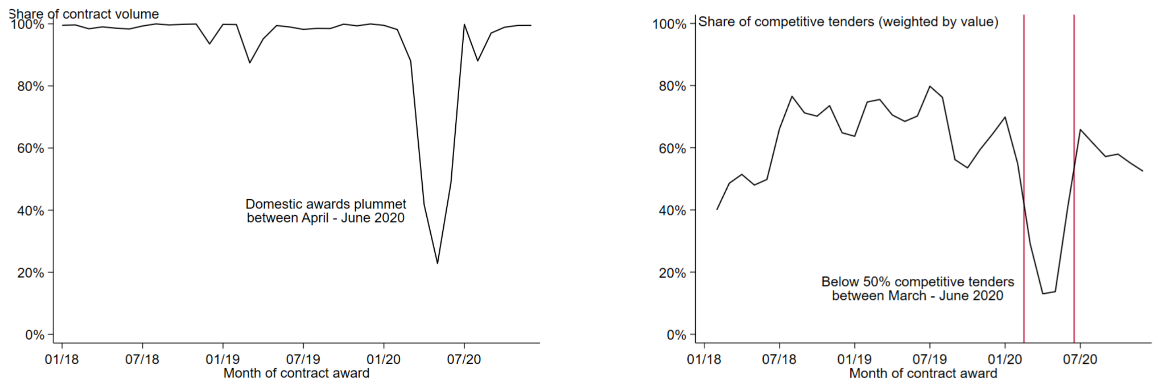
Table 1: Descriptive statistics of tenders for Covid-19-related product groups, n = 67,638

	Mean	Std. dev.	Minimum	Median	Maximum
Lot value (excluding VAT) in EUR	137,115.50	3,640,437.11	0.02	1,999.87	420,499,749
Indicator: domestic award	0.99	0.08	0	1	1
14-day average infection rate per 100 inhabitants	0.10	0.08	0	0	2.35
Infection rate at seller location	0.03	0.12	0	0	2.35
Total number of bidders	5.00	10.52	1	2	350
Share of foreign bidders	0.00	0.06	0	0	1

Spending patterns: The total spending on Covid-19-related medical supplies by country is summarized in Table 15. It is apparent that the amount procured by a country is not just a function of the size of its economy or population.[37] The five countries that procured the greatest amount during the pandemic are the United Kingdom (UK), Ireland, Germany, Norway and Romania. In both tables, the UK represents a large amount of spending. This is due to some

[36] The number of observations exceeds the number of CAN because one CAN may include several individual lots.

[37] An economic explanation is that countries with a well-funded healthcare system may already possess large stocks of commodities such as disinfectant or PPE and need to procure less in an emergency than countries with an under-funded healthcare system. However, I do not observe such stocks and in any case this would only represent a part of the spending. The overall decisions to procure medical supplies also depend on the incentives of buyers (e.g. differences in the funding of the healthcare system between countries, Beveridge vs. Bismarck systems) and differences in capacity planning for goods such as intensive care units (ICU) or medical labs (the data set includes construction services for both ICU and medical labs).



(a) Share of contract value awarded to domestic firms (b) Share of competitive tenders (moving average)

exceptionally large contracts for goods from China and Hong Kong. Most contracts are worth between 1,000 and 10,000 EUR (Figure 9). The ten largest individual lots are reported in Table 16. As a result of the presence of outliers, and to prevent one country from impacting the estimation by too much, we include a robustness check by leaving one country out at a time.

Finally, there are purchases completely outside the EU procurement system that do not appear in the data set. For example, the German "open house" purchase of face masks in March 2020 accepted face masks from all sellers who committed to a minimum delivery amount. According to media reports, circa a billion face masks at a unit price of 4.50 EUR were ordered.[38] While the contract notice is posted on TED, no CAN exist. In contrast with traditional contract awards, all firms that fulfilled certain conditions could deliver goods at the posted price.[39]

Net change in cross-border awards: The outcome of interest is the share of contracts that were awarded cross-border. This is an indicator variable that takes the value 1 if buyer and seller are located in the same country and 0 if not. At the start of the pandemic, the value of cross-border awards increases drastically.[40] Figure 4a shows the average monthly share of domestic procurement. The overall level of domestic procurement before the pandemic is consistent with previous studies on procurement in Europe by Vagstad (1995).[41]

Figure 4a shows that in a sudden reversal, the volume of domestic procurement dips in April 2020. Less than 50% of contract volume were awarded to domestic companies between April and June 2020, reaching a low of 22.8% in May 2020. Purchases in several European countries, including France, the UK, and Italy, contribute to this dip. Countries that received a large number of cross-border orders include large non-European countries, including the US and China, and within Europe Switzerland and Austria. Overall, the value of domestic contract awards are 98.5% of the total in the pre-pandemic period from January 2018 to January 2020 but only 60.3% in the pandemic period from February to December 2020.

Competition and direct awards: Competition for contract awards decreased at the start of the pandemic. Figure 4b describes the share of the total value that buyers awarded competitively in each month. I define competitive contract awards as having more than one bidder and not being categorized as "Contract awards without prior publication" or "Negotiated without a prior call for competition". Buyers awarded over 90% of contract volume non-competitively in April and May 2020. Direct awards to foreign sellers represent 56.5% and 77.1% of the total observed contract volume in these two months. While the presence of several bidders does not guarantee

[38] See Tagesschau, 16.05.2021: https://www.tagesschau.de/investigativ/ndr-wdr/open-house-101.html [in German]

[39] See https://ted.europa.eu/udl?uri=TED:NOTICE:147548-2020:TEXT:EN:HTML

[40] I do not observe local subsidiaries of foreign firms or foreign value-added to contracts awarded to domestic firms, e.g., through traded inputs or sub-contracting. Accounting for these alternative channels would result in a larger value share to foreign companies. See e.g. Cernat and Kutlina-Dimitrova (2015).

[41] Before the pandemic, buyers awarded over 90% of total contract value domestically. Over 99% of all contracts are awarded domestically, which implies that cross-border awards are of higher value on average.

a competitive auction, it is noteworthy that buyers placed many contracts directly with foreign sellers. This suggests that search frictions, such as a lack of awareness of buyers about foreign sellers or of firms about foreign procurement tenders, are unlikely to explain home bias before the pandemic.

Explanatory variables: One of the main explanatory variables is the average 14-day infection rate per 100 inhabitants. The overall shape of the infection rate is plotted in Figure 20.[42] This figure shows the average infection rate per month, weighting regions by their monthly contract value. This curve tracks closely the total European infection rate reported by the ECDC with a first wave in early 2020 and a much higher second wave in late 2020, as plotted in Figure 21. The ECDC figure contains weekly data separate by country. Although there are some differences, such as a slightly steeper drop in infection rates towards the end of the sample period, this comparison suggests that the weight of different European regions in the data set follows the general trend of the pandemic.

The variable "Regulation change" is a dummy that takes the value 1 for contracts signed after April 1, 2020, the publication date of the "European Commission guidelines on procurement" and 0 before. It always takes the value 0 in the control group. The purpose of the change in regulation was to increase buyer discretion. I test the hypothesis that the publication of these guidelines, as buyers had more discretion in awarding contracts, changed the share of cross-border procurement.

Further discussion on missing data and potential sample selection is in Appendix B.2.1. Procurement documents are sometimes incomplete, leading to missing values. These values are most likely not missing at random, yet a regression analysis shows that they likely just result in a downward bias of regression estimates, leaving us with a lower bound on possible effect sizes. Also, contracts below an administrative notification threshold don't have to be reported. Contracts that should normally fall above a notification threshold can be split or shaded intentionally to fall below the threshold and go unreported. While I cannot exclude the possibility that this also affects tenders for medical supplies, they likely do not impact the estimates. Only systematic differences in contract size between times of high cross-border awards and other periods would be of concern. The main period of high cross-border awards, April to June 2020, does not appear to be anomalous from visual inspection. There is also no significant seasonal variation in the main outcome or independent variables.

2.3.4 Summary of the data

Purchasing patterns across countries and product-groups are heterogeneous and lot values are highly dispersed. Beyond the documented differences in reporting, there are other unobserved differences between countries. For example, a country with a more technocratic public sector may have a higher base-line of cross-border awards and recalibrate less due to the pandemic.[43] This motivates a leave-one-out robustness check by country. Descriptive statistics show that the net effect of the pandemic and deregulation was a temporary surge in cross-border procurement. Direct awards to foreign sellers represent a high share of the total observed contract volume, especially in April and May 2020. While the presence of many bidders does not guarantee a competitive auction, it is noteworthy that buyers placed most contracts directly with foreign sellers in these months. This suggests that there is no lack of mutual awareness of buyers and sellers across borders driving home bias in normal times.

2.4 Estimating the impact of the emergency and buyer discretion

I estimate economically large effects of crisis urgency (measured through local infection rates) as well as increased buyer discretion (measured through a change in regulation) on the likelihood of cross-border procurement. First, I analyze medical supplies that are designated as Covid-19

[42] Scaled down by a factor of 1,000 from the original data for readable regression coefficients.

[43] I thank a referee for this suggestion.

related. Using local infection rates as a treatment for a staggered difference-in-difference analysis, there is an increase in the likelihood of a cross-border award by 19.3 percentage points following a one-standard deviation increase in infection rates.

The contracts in the first analysis were all subject to deregulation that increased buyer discretion. For a second analysis, I draw upon additional data for similar products from the same product categories but for which regulations were not lifted. The likelihood of a cross-border award increased by 35.7 percentage points for products and services for which the regulation was lifted relative to the control group. However, each analysis finds a large increase in cross-border procurement compared to the pre-pandemic baseline when only 1.5 percent of the total value was awarded internationally. Even the lower bound effect sizes in a "leave-one-out" analysis remain economically very large, at 9.3 and 16 percentage points, respectively.[44]

2.4.1 Empirical strategy

Ideal experiment: What would be the ideal experiment to causally identify the effect of increased *buyer discretion* and *crisis urgency*? For the former effect, consider a counterfactual Europe with identical infection rates but without the lifting of regulation. That setting would allow a direct causal interpretation of the regulation effect. It is apparent that the ideal experiment does not exist: non-European procurement markets, such as in Asia or the United States enacted their own regulatory responses to the pandemic. Both the pandemic and policy responses evolved at different speeds across the world. Alternatively, with complete information on the bidding history of all firms, one could analyze firms' likelihood of winning bids (conditional on their bid and other bids) around the regulation change. This data is not available, as individual bids are not released, nor is the identity of non-winning bidders known.

Likewise, a naive regression on infection rates would admit a causal interpretation only if infection rates were randomly assigned. This is clearly not the case: infection rates are endogenous, as they depend on previous infection rates, infection rates in neighboring regions, and they are correlated with factors such as the efficiency of the public sector or existing trade links. Changes in infection rates might not be random, as they might be influenced, e.g., government quality which might impact the likelihood of cross-border procurement and influence infection rates.

Crisis urgency: Nonetheless, our setting provides a unique chance to identify the effects of *crisis urgency* and *buyer discretion.* The first analysis exploits the fact that the pandemic intensified in different locations at different times. Using only data set on Covid-19-related medical supplies, local infection rates act as a treatment in a staggered difference-in-differences setup, where higher levels of treatment imply greater volumes of cross-border procurement. To obtain an unbiased estimate of this treatment effect, we have to assume i) parallel trends and ii) the linearity of the infection effect. While we cannot test whether infection rates would have developed in parallel in different locations, we do compare regions that were treated earlier and later further below.

As we have data on multiple locations and time periods, our data lends itself to a two-way fixed effects analysis with staggered treatment (Callaway and Sant'Anna, 2020; Athey and Imbens, 2021). A two-way fixed effects analysis generalises the canonical two-period difference-in-differences approach to multiple periods. The characteristic of the staggered treatment is that units are treated at different times. The unit that is treated is not the contract award (our unit of observation) itself, but the geographic region for which the contract is specified. The control group for treated units (regions) consists both of never-treated units and potentially also units that have been treated in the past. This setting also features varying treatment intensity as the treatment variable, infection rates, vary across locations and periods. Wooldridge (2021) shows that also with a staggered treatment, the two-way fixed effects estimator is appropriate to account for treatment intensity, covariates and interactions.

[44] I emphasize that the two analyses have different treatment and control groups and effects are not like-for-like or additive.

I estimate the following baseline equation which permits an interpretation of the effect of infection rates on cross-border awards.

$$y_{it} = \alpha_i + \lambda_t + \beta I_{it} + \vec{\beta} X_{it} + \epsilon_{it} \tag{42}$$

α_i denotes a contract's performance location, λ_t is the week. Recall that the indicator variable y_{it} takes the value 1 if the contractor is located in the same country as the buyer and 0 otherwise. I_{it} is the infection rate at the buyer's location for observation i.[45] X_i is a vector of control variables, the infection rate at the seller location, the share of foreign bidders, the total number of bidders, and dummies for country and product-group.

Buyer discretion: By drawing upon additional data, the events surrounding the pandemic allow us also to study the effect of a deregulation act. We collect additional data on contracts for products which are in adjacent product groups relative to the previously analyzed ones. These contracts fall outside the classification as "Covid-19 related". Between these contracts and the previously analyzed ones, a deregulation action on April 1st, 2020 provides a treatment that only affects the latter group but not the former, which acts as a control group of products with similar supply characteristics for which the procurement rules were unchanged. Identification of the effect of deregulation requires parallel trends between the control and treatment group. We show that trends in procurement for these products are indeed comparable to the treatment group.

There are procedural restraints (open and competitive tenders, transparent award criteria, review by external authorities) which were lifted in an announcement by the European Commission on April 1st, 2020, two months after the start of the pandemic.[46] After that date, buyers could award contracts directly and as fast as possible. The "Guidance from the European Commission on using the public procurement framework in the emergency situation related to the COVID-19 crisis" (2020/C 108 I/01) states:

> "for a situation such as the current COVID-19 crisis which presents an extreme and unforeseeable urgency, the EU directives do not contain procedural constraints. [...] [P]ublic buyers may negotiate directly with potential contractor(s) and there are no publication requirements, no time limits, no minimum number of candidates to be consulted, or other procedural requirements. No procedural steps are regulated at EU level. In practice, this means that authorities can act as quickly as is technically/physically feasible – and the procedure may constitute a de facto direct award only subject to physical/technical constraints related to the actual availability and speed of delivery."

Generally, this announcement applied only to tenders for certain goods and services that were seen as directly related to the pandemic, not to procurement overall. This communication from the European Commission was not a change in hard law. It merely clarified how to use the procurement framework. However, notes in procurement contracts refer explicitly to the use of expedited rules under the conditions of the pandemic.[47]

---

[45] The 14-day average rate of infection per 100,000 inhabitants published by the ECDC is scaled up by a factor of 1000 to improve readability of the coefficient estimates.

[46] The WHO declared the novel Coronavirus outbreak a public health emergency of international concern (PHEIC), WHO's highest level of alarm, on January 30, 2020, see https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline.

[47] One example is contract award notice 600222-2020, a British purchase of PPE from Austria, concluded on May 30, 2020. It says "In March the NHS experienced severe shortages of PPE [...]. In these circumstances, a procurement following the usual timescales under the PCR 2015, including accelerated options, was impossible. PPE manufacturers and supply chains were under immediate and unprecedented global pressure to provide products. A delay in engaging with the market by running a usual procurement process ran the risk of failing to acquire the necessary stock of PPE equipment and presenting a significant risk to life. The Department is content the tests permitting use of the negotiated procedure without prior publication (Regulation 32(2)(c)) are met: 1)

I estimate the following equation:

$$y_{it} = \tau_i + \lambda_t + \beta D_{it} + \vec{\beta} X_{it} + \epsilon_{it} \tag{43}$$

Compared to equation 42, τ_i now denotes whether a contract awards belongs to the treatment or control group, the set of control variables does not include product group dummies (to avoid co-linearity with treatment group status), and D_{it} takes the value 1 for observations in the treatment group after the treatment date (April 1st, 2020). Again, a leave-one-out analysis by country allows to investigate the importance of outliers.

2.4.2 Emergency effect

The first hypothesis is that greater levels of emergency lead to a greater share of cross-border awards. Measuring the degree of emergency implied by the pandemic through local infection rates, I investigate the impact of emergency on cross-border awards. If the absence of cross-border awards is associated with misallocation, cross-border procurement should increase when *crisis urgency* is greater. Consistent with the descriptive statistics from Section 2.3.3, any effect of infection rates on cross-border procurement is, if anything, transitory. The overall plummet in cross-border procurement coincides with the first wave of Covid-19 infections, but not the higher, second wave. This supports the interpretation of infection rates as a measure of *crisis urgency* which presumably played a bigger role when the world was initially confronted with the pandemic than in later stages of the pandemic.

The present analysis differs from a standard setup of the staggered difference-in-difference estimator first because of the binary outcome variable. Both linear probability models and non-linear alternatives such as probit are available and have been used in these cases (Finkelstein, 2002) although the interpretation of the treatment effect in the non-linear model is not straightforward (Puhani, 2012). I do not repeat the exercise with a non-linear model as the focus is on the demonstration of an effect of considerable economic size, rather than a precise estimation of magnitude.

Difference-in-differences diagnostics: First, I investigate the standard "parallel trends" assumption and potential spillover in treatment or outcomes. A standard approach to demonstrate parallel trends is to plot control and treatment group around the threshold and visually confirm that their trends are approximately parallel. Here, this is difficult due to the large number of regions with distinct infection profiles. Therefore, I aggregate regions into two groups depending on the date when infections were first recorded. Figures 22 and 23 in the Appendix compare the share of domestic purchases and average infection rates for these two aggregated groups. The former figure shows differences with two visible spikes in cross-border awards for the "early" group and a spike in-between for the "late" group. Looking at the flat section of the graphs in Figure 23 until the beginning of April, it is also clear that both groups do not experience a fundamentally different development in the infection numbers for the first few weeks. This strengthens the case for areas that were infected a few weeks later as a control group for those that were infected earlier.

Estimation results: The outcome of interest is the share of the value awarded abroad, so each observation is weighted by the value of the award. This avoids distortion from potentially arbitrary divisions of contracts into more lots. I use White's heteroskedasticity-robust standard

The purchasing of PPE was identified as strictly necessary to meet anticipated demand. 2) It is responding to Covid-19 immediately because of public health risks presenting a genuine emergency. 3) The events that led to the need for extreme urgency were unforeseeable: the Commission itself confirmed: 'The current coronavirus crisis presents an extreme and unforeseeable urgency – precisely for such a situation our European rules enable public buyers to buy within a matter of days, even hours, if necessary.' 4) There was no time to run an accelerated procurement under the open, restricted or competitive procedures with negotiation that would secure products within the required timescales. 5) The situation is not attributable to the contracting authority: It has not done anything to cause or contribute to the need for extreme urgency."

errors to account for the well-known fact that linear regressions in a binary outcome framework have heteroskedastic residuals. The results are summarized in Table 2.

Table 2: Difference-in-difference analysis with staggered treatment through infection rates

	All obs.	No UK
Dep. var.: Contract awarded to domestic company		
ATET		
14-day average infection rate	-1.376***	-0.463***
	(0.318)	(0.127)
Controls		
Infection rate at seller location	2.242***	0.529***
	(0.400)	(0.183)
Total number of bidders	0.002***	0.000
	(0.001)	(0.000)
Share of foreign bidders	-0.363***	-0.428***
	(0.100)	(0.061)
Standard error of infection rate	0.14	0.20
1-std.dev. increase	-0.193	-0.093
Dummies	yes	yes
N	67,638	67,387

Robust standard errors in parentheses, *** p<0.01. Infection rates are 14-day moving average per 100 inhabitants at the NUTS-region reported as performance location (buyer infection rate) or as location of the contractor (seller infection rate). Share of foreign bidders computed as number of foreign bidders from EU and non-EU countries divided by total number of bidders.

The estimated average treatment effect on the treated (ATET) is for a unit increase of the 14-day average infection rate per 100 inhabitants. It is best interpreted relative to observed standard deviations during the pandemic. I compute the standard deviation of the infection rate for observations with non-zero infection rates.[48] Column one includes all observations for medical supplies where the lot value is not missing. In the whole sample, the effect of a one-standard deviation increase in infection rates is an increase of 19.3 percentage points in the likelihood of a cross-border award.

Due to heterogeneity between countries and the presence of outliers, I conduct a "leave-one-out" analysis in which I exclude one country at a time. For all of these tests, the effect remains statistically significant and when leaving out Czech Republic, it even increases by about 50% relative to the full sample. Unsurprisingly in light of the previously documented outliers from the UK, leaving out the UK diminishes the coefficient the most. Even this case, however, I find that a one-standard deviation increase in infections increases the likelihood of a cross-border awards by 9.3 percentage points (see column 2 of Table 2). Given that cross-border contracts only represent 1.5 percent of value pre-pandemic (see Figure 4a), even an increase by just over 9 percentage points is economically large.

Spillover might occur if the treatment status of some units influences the treatment of other units. In the context of a pandemic, this is clearly a concern as the treatment literally spreads and may be transmitted by people moving between regions. However, whether this spread of infections also influences procurement decisions is less clear, as buyers should react to local infection rates independent of whether local infections arise because of movement between regions

[48] Otherwise the standard deviation of the infection rate would be artificially diminished by pre-pandemic observations.

or infections within a region. More importantly, are buyers influenced by information from other regions when making purchasing decisions? In other words, do infections in other regions provide additional information beyond what local (lagged) infection rates predict?

To test this, for each region I test for Granger-causality of infection rates against own lags and lags of other regions in the same country (see Appendix B.2.2). Allowing for up to six lags of the weekly updated infection data, I fail to reject the Null-hypothesis that the infection rates in one region are not Granger-caused by infection rates in other regions (see Table 18).[49]

I also test for spillover in outcomes (Table 19 in Appendix B.2.2): Does treatment status by one region impact outcomes in another region? Using data on population and infection rates, I compute for every region and every week the national infection rate excluding that region. In other words, it is a national measure of infection rates that excludes the contract location. Note that this measure is strongly correlated with local infection rates: this tells us that most regions at most times did not diverge too far from national trends in the infection rate. Still, even in the small subset of observations (842 observations representing ca. 45 million Euros worth of contracts) where local infections were below the overall median infection rate while national rates were above, there is no significant impact of the national infection rate on outcomes.

2.4.3 Suspension of regulation

As described in Section 2.4.1, the European Commission published guidance on using the procurement framework in which the normally very restrictive rules were widely suspended. Did the suspension of these rules lead to additional cross-border procurement? These suspensions increased buyer discretion, which previous scholars identified as a potential source of better procurement in some contexts and as a potential source of misallocation in others (Section 2.2). This hypothesis follows also from the theoretical model (Appendix B.4).

The European Commission has confirmed that these rules were indeed applied restrictively to Covid-related tenders. I leverage the TED classification of Covid-related tenders to identify those products that fell under these rules. The control group includes all product groups that are at the same CPV hierarchy level as the Covid-related tenders. This yields a selection of closely related products.

For example, from the category "beds for medical use" (33192100), the product group "hospital beds" (33192120) is Covid-related and all contract award notices 2018-2020 are part of the treatment group. Contract award notices for the remaining products in the category "beds for medical use", which are "orthopaedic beds" (33192110), "motorised beds" (33192130), "psychiatric couches" (33192140), "therapy beds" (33192150), and "stretchers" (33192160), enter the control group. A complete list of relevant product groups is provided in Table 13.

The main assumption is that outcomes for products in the control group would have a similar potential outcome to the products in the treatment group. This could mean that for a contractor, there is a high degree of supply-side substitutability between, say, "intensive-care unit construction work" (treatment group) and "diagnostic screening room construction work" (control group). It could also mean that, even if different firms produce "protective gear" (treatment group) on the one hand and "work gloves" and "safety visors" (control group) on the other hand, changes in market conditions such as increasing input costs affect these firms similarly.

Difference-in-differences diagnostics: The control group is based on product similarity in the product classification system.[50]

[49] I flexibly allow for up to nine lags and then choose the model based on Bayesian Information Criterion. Allowing for longer or shorter lag structures, e.g. six weeks or twelve weeks of past information, yields the same result.

[50] A similar method has been used by Eurostat, the statistical agency of the EU, to analyze trends in international trade related to Covid. Based on the trade statistics in their Comext data base, they compare Covid-19 specific product groups with other, similar product groups which are selected by chapters of the "Harmonised System" product classification. See Eurostat, Archive:EU trade in COVID-19 related products (31.03.2021), last accessed 01.05.2023.

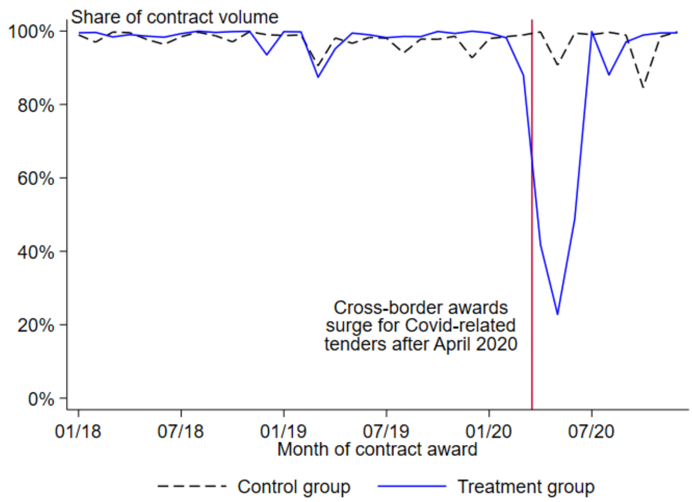


Figure 5: Average share of cross-border awards in the treatment group with greater buyer discretion and the control group

Figure 5 shows the average monthly rates of cross-border procurement for the products in the treatment and control group. This is unsurprising given the soft nature of the suspension of buyer discretion: instead of representing a change in "hard law", it is a clarification on the interpretation of existing rules. Nonetheless, a clear impact of the regulation on the impacted goods, but not similar goods can be seen. At the same time, the graph shows broadly parallel trends pre-April 2020, save for a handful of spikes in individual months.

For most product groups, even in the control group the average monthly spending during the pandemic exceeds spending before the pandemic by a large amount. Only in two product groups (albeit including the largest by spending, "various medicinal products" alongside "antineoplastic agents"), spending is slightly lower during the pandemic. For most product groups both in the control and treatment group, average spending during the pandemic exceeds average spending before the pandemic, sometimes by an order of magnitude. This suggests that differences in cross-border procurement between goods in the treatment and control group are likely not driven by an uneven spike in demand that would only affect the treatment group.

Estimation results: Estimation results are reported in Table 43. The baseline result on the full sample indicates that the increase in buyer discretion increased the share of cross-border awards for Covid-19 related tenders by circa 36 percentage points. I treat the smallest absolute regression coefficient in the leave-one-out analysis as a lower bound of this effect. This value is only 16 percentage points when the UK is removed, reflecting large cross-border contracts from the UK. The overall results are unsurprising, given the size of the drop in cross-border procurement, as well as the visually clear difference between the treatment group and the control group in Figure 5.

2.4.4 Robustness

Potential challenges to the empirical strategy include the impact of the pandemic on the manufacturers of medical supplies, but also the measurement of buyer discretion and emergency. The pandemic has lead to disruptions for manufacturers as well as for distribution and international supply chains which could impact inference on awarding practice. For example, a decrease in contract awards to foreign companies could be related to border closures or export restrictions on medical supplies that were applied in 2020.

All companies may face higher costs or lower production capacity due to sick workers, new safety measures, or general uncertainty arising from the circumstances of the pandemic and

Table 3: Difference-in-difference analysis with targeted deregulation

	All obs.	No UK
Dep. var.: Contract awarded to domestic company		
ATET		
Regulation change	-0.357***	-0.160***
	(0.055)	(0.061)
Controls		
Infection rate at seller location	0.394***	0.058*
	(0.121)	(0.033)
Total number of bidders	0.001	0.001***
	(0.000)	(0.000)
Share of foreign bidders	-0.030	-0.504***
	(0.036)	(0.075)
dummies	yes	yes
N	320,213	319,173

Robust standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01. Infection rates are 14-day moving average per 100 inhabitants at the location of the contractor (seller infection rate). Share of foreign bidders computed as number of foreign bidders from EU and non-EU countries divided by total number of bidders.

therefore decide to pursue fewer contracts or only bid on nearby (domestic) contracts. Depending on a firm's industry, production technology and location, rising infection rates can imply

- decreased production capacity, including factory shutdowns, due to worker sickness;
- decreased labor productivity, or increased production cost, due to increased safety standards (e.g. encumbering safety clothing, regular disinfection protocols, maintaining physical distance between workers);
- lack of physical access or uncertain access to raw materials due to interrupted supply chains;
- lack of physical access or uncertain access to buyers due to border shutdowns and export restrictions.

To rule out that the observed contracting practices are driven by these supply-side factors, company size would be available as a control. However, firm identity and size are not reported, only the total number of bidders is observed, and the numbers of bidders from other EU and non-EU countries. The control variables in the main regression, the number of firms bidding and the share of foreign bidders, already account for some of these supply-side effects. I also consider effects that would impact foreign awards even after controlling for participation.

Capacity constraints on the supply-side: A buyer might turn to foreign sellers because domestic industries cannot supply the short-term need for all goods and services. Then, cross-border awards should increase during the pandemic even absent any home bias. However, even applying a conservative definition of spare capacity does not change the findings of the main analysis.

Excess domestic capacity is present when there is a larger number of domestic bidders than domestic winners. For each contract, we subtract from the largest number of domestic bidders of any given lot the number of distinct domestic winners (counting over firm names after cleaning inconsistently spelled ones). This method likely underestimates domestic spare capacities because it only counts entire domestic companies that would stand ready to serve a contract, the extensive margin of domestic supply, while leaving unobserved intra-firm spare capacity, the intensive margin. This measure is also conservative because bidders are anonymous, so if different domestic firms bid for different lots on a contract, it is not possible to distinguish between them: the lot with the largest number of domestic firms is still just a lower bound of distinct domestic firms ready to supply a certain product at the time of contract award.

A potential challenge for the analysis of infection rates as a measure of *crisis urgency* could be that the outcome variable impacts infection rates. That is, less misallocation and better procurement might, for some reason, directly result in lower infection rates. However, exploiting variation in infection rates within country-month-product group brackets means that any impact of procurement on infection rates would have to occur within the same month to create this kind of endogeneity. The effects of emergency and the regulation are robust to an alternative approach to measuring emergency that specifically circumvents such potential feedback effects: the infection rate in other regions of the same country (computed via population statistics and national infection rates, see Table 20). However, these measures result in a noisier estimation than the main estimation, so I only consider them as robustness checks.

The results are reported in Tables 21 and 22. The difference-in-differences analysis by regional infection rates has now 23,443 observations (34.7% of the full sample). The effect of the infection rate remains significant and the effect size slightly increases in absolute terms to -19.7 percentage points for a one-standard deviation increase in infections. For the analysis on the regulation effect using a control group of additional products, I use the full control group. Here, the effect size is lower, at -8.8 percentage points post regulation. The coefficients are still significant but less so than in the main analysis. The leave-one-out robustness check does not change the results much, except that is pushes up the standard error for the regulation effect so that it is no longer significant at the 5% level. Overall, the analysis shows that a lack of domestic spare capacity is not a sufficient explanation for the surge in cross-border procurement. Large effects, especially from rising infection rates remain for contracts on which domestic suppliers bid.

Another supply-side concern could arise from changes in the pool of sellers. For example, the increase in domestically awarded contracts could be due to buyers using their increased discretion to purchase via resellers or trading companies. These companies may not qualify for procurement contracts for medical supplies under normal circumstances but might have privileged access to foreign manufacturers. Anecdotally, the German government called upon companies with trading links to China to purchase additional face masks around Easter 2020. While the data is not sufficiently rich to describe composition differences in their industries, I check that companies which have as name components “trade”, “logistic” or “distribution” are not more common among contractors during the pandemic than before and make up only a small portion of total sales, suggesting that this is not a large concern.

A further robustness check is to ignore the lot value and instead count the number of lots awarded domestically or cross-broder. The results do not replicate when ignoring differences in value between individual lots. In this case there is no effect of the infection rate and a negative effect but very small effect of the regulation change when looking at all contracts, including those with missing lot values (see Table 23). This is unsurprising, as this check is very much driven by two countries whose contract award notices are typically split into many, sometimes very small lots: Romania, representing 63% of all contract awards, and Poland, representing 20% of the absolute number of lots.[51]

Additionally, in a subset of countries with a series of failed procurement tenders during the pandemic, there is no evidence of any impact of failed procurement on contract awards (Table 24). I also test whether correlation between procurement contracts within countries, for example due to similarity in regulatory and legal environments or similarities in training, doctrine, and

[51] EU regulation suggests that contracts may be aggregated to achieve cost savings, while a division into smaller lots might be motivated by a desire to allow smaller companies to participate, for example in Directive 2014/24/EU, preamble, paragraph 59. However, these competing forces determining contract size and lot division are outside the scope of the paper.

perspective among public sector buyers, plays a role. I re-estimate the difference-in-differences analyses with country-clustered standard errors, as well as standard errors that are clustered by contract. Both types of clustered standard errors are moderately larger than White-robust standard errors (see Tables 25 and 26), but the estimates remain significant at conventional significance levels. Therefore, these robustness checks do not change the conclusions regarding the impact of crisis urgency and buyer discretion.

As mentioned above, differences across countries, for example differences in how procurement of medical supplies interacts with different health systems or in the quality of different public sectors, could impact results. For example, some countries might face lower base-lines of misallocation to begin with and consequently experience lower effects of the pandemic. While it is not possible to control for every conceivable difference between countries, in line with our goal of demonstrating the importance of both *crisis urgency* and increased *buyer discretion*, we account for these differences by including country-fixed effects, capturing different baselines, and with the "leave-one-out" analysis which tells us the lowest absolute effect size found among subsets in which one country is left out. These still continue to find economically large and statistically significant effects.

Our estimation of these effects requires that trends in procurement awards would have been parallel across different regions independent of when they were treated with infections in the first analysis. The best argument for this is perhaps the long-lasting and consistently high share of domestic procurement for the product categories of interest which do not suggest that anything would have changed absent the pandemic. For the second analysis, a similar argument holds that we see persistently high shares of domestic awards throughout the pre-pandemic period for all goods, and substantial cross-border procurement only after the deregulation well into the pandemic.

2.4.5 Discussion

The overall surge in cross-border awards is driven by two channels: both *crisis urgency* and increased *buyer discretion* lead to an increase in cross-border awards. This does not exclude the possibility that other channels may have an impact, too. At least the result on buyer discretion is surprising as buyer discretion may in principle also contribute to misallocation.

The effects are statistically significant, even when accounting for across-country heterogeneity, and economically large: the increase in the likelihood of a cross-border award following a one-standard deviation increase in infection rates is over 10 times larger than the baseline rate, and over 20 times in the case of the deregulation. Interpreting this in light of the descriptive statistics and charts, the large drop in value awarded to domestic companies coincides with the first, lower wave of infections, and does not repeat itself when infection rates rise much higher during a second wave in late 2020. This suggests that any mitigating effect of urgency on home bias was at best temporary.

While the absence of cross-border procurement is not direct evidence of misallocation, there are good reasons to presume that low shares of cross-border procurement indicate misallocation. Firstly, large variations in cross-border procurement as a result of a change in buyer discretion is inconsistent with a hypothesis of equal treatment. Secondly, a pure cost-based explanation could not explain the drastic surge in cross-border procurement from almost nothing pre-pandemic to more than 50 percent in some months of 2020. As argued above, the results are not driven by a lack of competitive foreign sellers. The magnitude of the observed effects is economically large relative to potential exclusionary effects of, say, fixed costs that might arise from translation of tender documents.

I rationalize these findings using a game of monitoring that is presented in Appendix B.4 to explain especially the puzzling effect of buyer discretion. In this game, a buyer chooses either selects the unconditionally best supplier or a favored supplier and the government can choose to make a costly review of that purchase. Emergency appears as the differential in the payoff that the government receives when a purchase is made either from the best supplier or the favored supplier. It is plausible that this differential is higher in an emergency due to the health threat posed by the pandemic where "procuring well" (choosing a supplier that aligns with tender requirements) is of the essence.

Buyer discretion enters the model through the cost of monitoring. When regulation on buyer discretion decreases these costs and makes the collusive outcome less likely, the lifting of such regulation should result in more misallocation. There is evidence to the contrary in the empirical analysis. The conclusion within the proposed model is that in the context of the pandemic, the lifting of regulation does not increase the cost of monitoring collusion. Indeed, as the media reports cited in the introduction show, misallocation and collusive practices were under great scrutiny. While normally, procurement outcomes rarely make headlines, in the pandemic there was great public interest in the procurement of medical supplies. Therefore, in spite of the known instances of misallocation, overall buyers were put under more pressure to avoid misallocation, potentially reducing the cost of uncovering and fighting undesirable outcomes. This can explain the observed surge in cross-border procurement.

2.5 Conclusion

A novel data set of procurement tenders in Europe allows us to study the natural experiment provided by the onset of the Covid-19 pandemic. A temporary surge in cross-border awards in 2020 was driven by suspended rules on buyer discretion and an emergency effect of the pandemic. The results of the empirical analysis suggest that *buyer discretion* and *crisis urgency* are important channels of procurement. Both channels exert an economically large impact on cross-border procurement, with effect sizes an order of magnitude greater than the baseline rate.

Some open questions merit follow-up investigation. While this paper investigates the channels that affect home bias, the data is silent on contract performance indicators, such as cost overruns, delays, or terminations. Additional data could directly demonstrate the effect of misallocation both for economic and health outcomes. The paper also does not account explicitly for the role of trade policy instruments. An interesting subject for follow-up research would also be a network analysis of buyer-seller decisions following the pandemic to study the long-term effects on buyer-seller relationships in Europe.

Beyond the context of the pandemic, we see that buyers can rapidly adjust their behavior in emergencies. In spite of decades of low shares of cross-border procurement, contradicting the EU policy goal of completing the "Single Market", there was a surge of cross-border procurement at the onset of the pandemic. Both the local extent of emergency, measured through infection rates, and increasing buyers' discretion, had contributed to this surge. This finding can be explained through a model in which the costs of monitoring buyers and the extent of their informational advantage impact misallocation. While the monitoring cost decreased due to heightened media attention, buyers' informational advantage increased as procuring well for medical supplies is of heightened interest during a medical emergency.

Also, in normal times as well as during emergencies, buyer incentives and regulatory constraints are important to foster EU Common Market policy. The prevalence of direct awards and non-competitive tenders in spring 2020 suggests that buyers and sellers are aware of each other. Search frictions, such as mutual unawareness of buyers and sellers, are then likely not an important cause of the low volume of cross-border contracts absent the pandemic. To the contrary, buyers seem capable of selecting foreign sellers if they wish to do so. This helps to rule out policies that seem well-intended but are likely ineffective to remedy home bias. For example, making it easier for sellers to enter their bids into foreign procurement tenders by reducing language barriers is likely not effective against home bias. This does not mean that existing regulation in this direction, such as Europe-wide publication requirements for large tenders, are not helpful for market integration. They might have built up existing mutual awareness of buyers and sellers that enabled cross-country procurement in the pandemic.

For policy-evaluation, the analysis suggest that lifting the regulations was successful in creating a more integrated response to the pandemic that allowed the most affected regions to draw on the Single Market, in spite of the transitory nature of the effect. Economic and social policies during the pandemic have been critically examined and found lacking in some regards. In the context of procurement policy, this is likely unjustified. The main caveat is that the relationship between procurement outcomes and medical outcomes is unobserved. Yet, the procurement system at large proved capable of adjusting to the pandemic with an unprecedented increase in cross-border purchases.

Acknowledgments

I thank Giacomo Calzolari, David K. Levine, Stephen Davies, Laurens Vandercruysse and Joosua Virtanen and two anonymous referees for comments, as well as participants at CLEEN 2021, CRESSE 2021 and working groups at EUI. The usual disclaimer applies. Funding by the German Academic Exchange Service is gratefully acknowledged.

References

Athey, Susan and Guido W. Imbens (2021). "Design-based analysis in Difference-In-Differences settings with staggered adoption". en. In: *Journal of Econometrics*. DOI: 10.1016/j.jeconom.2020.10.012.

Baltrunaite, Audinga et al. (2021). "Discretion and Supplier Selection in Public Procurement". In: *The Journal of Law, Economics, and Organization* 37.1, pp. 134–166. DOI: 10.1093/jleo/ewaa009.

Barbosa, Klenio and Pierre C. Boyer (2021). "Discrimination in Dynamic Procurement Design with Learning-by-doing". en. In: *International Journal of Industrial Organization* 77, p. 102754. DOI: 10.1016/j.ijindorg.2021.102754.

Burguet, Roberto (2017). "Procurement Design with Corruption". en. In: *American Economic Journal: Microeconomics* 9.2, pp. 315–341. DOI: 10.1257/mic.20150105.

Callaway, Brantly and Pedro H. C. Sant'Anna (2020). "Difference-in-Differences with multiple time periods". en. In: *Journal of Econometrics*. DOI: 10.1016/j.jeconom.2020.12.001.

Carboni, Chiara, Elisabetta Iossa, and Gianpiero Mattera (2018). "Barriers towards foreign firms in international public procurement markets: a review". en. In: *Economia e Politica Industriale* 45.1, pp. 85–107. DOI: 10.1007/s40812-017-0089-2.

Casady, Carter B. and David Baxter (2022). "Procuring healthcare public-private partnerships (PPPs) through unsolicited proposals during the COVID-19 pandemic". en. In: *Journal of Public Procurement* 22.1, pp. 6–16. DOI: 10.1108/JOPP-07-2020-0061.

Cernat, Lucian and Zornitsa Kutlina-Dimitrova (2015). *International Public Procurement: From Scant Facts to Hard Data*. en. Robert Schuman Centre for Advanced Studies Research Paper. DOI: 10.2139/ssrn.2682582.

Coviello, Decio, Andrea Guglielmo, and Giancarlo Spagnolo (2017). "The Effect of Discretion on Procurement Performance". In: *Management Science* 64.2. Publisher: INFORMS, pp. 715–738. DOI: 10.1287/mnsc.2016.2628.

Cutler, David M. and Lawrence H. Summers (2020). "The COVID-19 Pandemic and the $16 Trillion Virus". en. In: *JAMA* 324.15, p. 1495. DOI: 10.1001/jama.2020.19759.

Finkelstein, Amy (2002). "The effect of tax subsidies to employer-provided supplementary health insurance: evidence from Canada". en. In: *Journal of Public Economics* 84.3, pp. 305–339. DOI: 10.1016/S0047-2727(00)00155-9.

Hessami, Zohal (2014). "Political corruption, public procurement, and budget composition: Theory and evidence from OECD countries". en. In: *European Journal of Political Economy* 34, pp. 372–389. DOI: 10.1016/j.ejpoleco.2014.02.005.

Hoekman, Bernard et al. (2021). "COVID-19, public procurement regimes and trade policy". en. In: *The World Economy*. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/twec.13118. DOI: https://doi.org/10.1111/twec.13118.

Kutlina-Dimitrova, Zornitsa and Csilla Lakatos (2016). "Determinants of direct cross-border public procurement in EU Member States". en. In: *Review of World Economics* 152.3, pp. 501–528. DOI: 10.1007/s10290-016-0251-3.

La Cour, Lisbeth and Anders Milhøj (2013). "An explorative study of the contracts of the Tender Electronic Daily (TED)". In: *Symposion i anvendt statistik: 28.-29. januar 2013*. Copenhagen, pp. 178–194.

La Cour, Lisbeth and Grith Skovgaard Ølykke (2018). "Public Procurement in the EU: Another Explorative Study". In: *Symposium i anvendt statistik, Vol. 25*. Copenhagen.

Laffont, Jean-Jacques and Jean Tirole (1991). "Auction design and favoritism". In: *International Journal of Industrial Organization* 9.1, pp. 9–42.

Lalliot, Laurence Folliot and Christopher R Yukins (2020). "COVID-19: Lessons learned in public procurement. Time for a new normal?" en. In: *Concurrences* 3, pp. 46–58.

Loader, Kim (2007). "The Challenge of Competitive Procurement: Value for Money Versus Small Business Support". In: *Public Money & Management* 27.5. Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1111/j.1467-9302.2007.00601.x, pp. 307–314. DOI: 10.1111/j.1467-9302.2007.00601.x.

— (2016). "Is local authority procurement supporting SMEs? An analysis of practice in English local authorities". In: *Local Government Studies* 42.3. Publisher: Routledge _eprint: https://doi.org/10.1080/03003930.2016.1157068, pp. 464–484. DOI: 10.1080/03003930.2016.1157068.

McAfee, R. Preston and John McMillan (1989). "Government procurement and international trade". en. In: *Journal of International Economics* 26.3-4, pp. 291–308. DOI: 10.1016/0022-1996(89)90005-6.

Prier, Eric, Palina Prysmakova, and Clifford P. McCue (2018). "Analysing the European Union's Tenders Electronic Daily: possibilities and pitfalls". In: *International Journal of Procurement Management* 11.6. Publisher: Inderscience Publishers, pp. 722–747. DOI: 10.1504/IJPM.2018.095655.

Puhani, Patrick A. (2012). "The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models". en. In: *Economics Letters* 115.1, pp. 85–87. DOI: 10.1016/j.econlet.2011.11.025.

Ramboll/HTW Chur (2011). *Cross-border procurement above EU thresholds*. en. Tech. rep. KM-31-13-707-EN-N. Publisher: Publications Office of the European Union. Copenhagen: DG Internal Market and Services, URL: http://op.europa.eu/en/publication-detail/-/publication/0e081ac5-8929-458d-b078-a20676009324.

Sanchez-Graells, Albert (2020). "Procurement in the Time of COVID-19". In: *Northern Ireland Legal Quarterly* 71.1, pp. 81–87.

Vagstad, Steinar (1995). "Promoting fair competition in public procurement". en. In: *Journal of Public Economics* 58.2, pp. 283–307. DOI: 10.1016/0047-2727(94)01472-Z.

Vecchi, Veronica, Niccolò Cusumano, and Eric J. Boyer (2020). "Medical Supply Acquisition in Italy and the United States in the Era of COVID-19: The Case for Strategic Procurement and Public–Private Partnerships:" en. In: *The American Review of Public Administration* 50.6-7. Publisher: SAGE PublicationsSage CA: Los Angeles, CA, pp. 642–649. DOI: 10.1177/0275074020942061.

Wooldridge, Jeffrey M. (2021). "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators". en. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3906345.

3 Pricing algorithms out of the box: a study of the repricing industry

with Giacomo Calzolari[52]
European University Institute
CEPR

Abstract

Businesses, including sellers on online marketplaces, increasingly adopt repricing algorithms, i.e., algorithms designed to set the prices autonomously. Potential consequences in digital markets, such as high prices, have drawn much attention, but there is virtually no study of the actual supply-side of the repricing industry. Relying on suppliers' descriptions, we document and classify repricers, ranging from mechanic fixed-rule algorithms to sophisticated self-learning AI algorithms and discuss product features, fees, and associated services. Although the analysis is descriptive, we offer a first, broad overview of this new and important industry, potentially affecting millions of consumers in online marketplaces.
Keywords: pricing algorithms, artificial intelligence, repricer, marketplaces, online commerce
JEL: D21, D43, D83, L12, L13

3.1 Introduction

While shopping online, buyers encounter prices increasingly determined by software algorithms. To automate pricing, sellers in online marketplaces rely on these repricing algorithms, called "repricers" or "dynamic pricing algorithms", to automate pricing. Such repricing algorithms can use significant amounts of information at a high pace. For example, they may feed in the prices posted by competitors for similar goods, as well as other factors such as competitors' stock levels and customers' reactions and quickly adjust prices. Faced with vast and differentiated online markets, repricing algorithms promise to use large quantities of data to automate sellers' profit maximization over many different items and possibly increase market efficiency. Since even tiny sellers in online mass markets currently have access to repricing algorithms, as sometimes relatively inexpensive and turn-key services, the overall impact can be considerable.

On the flip side, repricing algorithms may open the door to new forms of market abuse. For example, already as early as 2015, the UK Competition and Market Authority fined online sellers of posters for using a repricing software to fix prices on Amazon Marketplace and a similar case was litigated in the US by the US Department of Justice.[53] In *Eturas* travel agents used algorithms to fix prices on an online booking platform. In particular, the online booking system informed travel agents that the pricing algorithm was capping price discount at 3% maximum and fined the platform and the agents.[54] In 2018, the European Commission as well as the Competition and Markets Authority found in several cases that repricing algorithms could be used to enforce the (forbidden) resale prince maintenance, as algorithms allow to promptly and automatically detect deviation from agreed prices.[55]

[52]Both authors contributed equally to the writing of the article.

[53]See the decision of the Competition and Markets Authority, Online sales posters and frames, Case 50223, 12 August 2016 and United States of America v David Topkins, Plea Agreement, Department of Justice, Antitrust Division, No. CR 15-00201 WHO. Another early case in the airline industry was United States v Airline Tariff Publ'g Co., 836 F. Supp. 9 (D.D.C. 1993).

[54]Case C-74/14 Eturas' UAB and Others v Lietuvos Respublikos konkurencijos taryba EU:C:2016:42, Judgment of the Court, 21 Jan 2016.

[55]Cases AT.40465 (Asus), AT.40469 (Denon and Marantz), AT.40181 (Philips) and AT.40182 (Pioneer), decisions of 24 July 2018; decision of the Competition and Markets Authority, Digital piano, and digital keyboard sector: anti-competitive practices 50565-2, 8 October 2019.

A nascent economic literature has addressed these developments with theoretical and empirical research and studies the novel impact of repricing algorithms on competition in digital marketplaces. Calvano et al. (2020), Klein (2021) and Johnson et al. (2020) use simulations and show that self-learning algorithms, that is, algorithms (in the class of reinforcement-learning) that learn autonomously, can also learn collusive outcomes with supra-competitive prices and mechanisms to punish competitive deviations. A (scant) empirical literature shows that, depending on market structure, the adoption of pricing algorithms may increase price-cost margins, e.g., in the German retail gasoline market (Assad et al., 2020), or online marketplaces (Wieting and Sapi, 2021). With a theoretical analysis, Brown and MacKay (2021) show that simple pricing algorithms that react quickly to a rival's price change can increase price levels and price dispersion.

Notwithstanding these interesting early results, the research on repricing algorithms is hampered by limited knowledge about the actual functionality of repricing algorithms and their deployment. Algorithmic repricing companies are often opaque about their products to protect their intellectual property and proprietary software. Furthermore, the repricing industry is growing quickly and remains largely uncharted territory. For example, we do not know how many options sellers have, e.g., for repricing mass-market products and in terms of reaction speed. Nor do we know how competitive the repricing market is and which contractual agreements it uses to attract clients.

For these reasons, early research on repricing algorithms has relied so far on indirect assessments of the properties of repricing software. For example, early discussions conjectured features of such software (Schwalbe, 2018), or relied on certain classes of practical algorithms in simulated environment (Calvano et al., 2020). The empirical research identified the use of algorithms for pricing, but with no information on their characteristics, as in Assad et al. (2020) for the gasoline market, Chen et al. (2016) on Amazon, and Bol.com, another marketplace (Wieting and Sapi, 2021). However, the lack of information about the characteristics of the available algorithms, and the range of offers, is problematic. The details of algorithms turn out to be essential for their implications and also for policy assessment.

This paper contributes to the debate about repricing algorithms with a systematic overview of these algorithms' supply, fees, market positioning, and claimed features. Underlying this paper is a descriptive study of 185 repricing algorithms by 130 firms, spanning different geographies and use cases. While we do not have the opportunity to investigate the actual underlying software of these proprietary algorithms, we study claimed features and pricing schemes. The analysis allows identifying new and valuable information that could help understand this market and its segments and characterize the capabilities of today's software.

Our focus is on third-party repricing algorithms that are available for purchase online rather than bespoke repricing algorithms developed by large and sophisticated software companies. The reasons are twofold: first, we only observe the repricing algorithms offered in the market but not those designed and deployed "in-house". Second, we are interested in the extent to which mass markets populated by many small sellers could already be affected by algorithmic pricing. We find a wide variety of products that fall under the broad category of "price optimization software", ranging from off-the-shelf, third-party Amazon repricing algorithms for various goods to decision support software for specialized industries.

These products vary in how they are deployed, arrive at their price recommendations, and how expensive and sophisticated they are. All repricing algorithms that we study have in common that they are readily available for purchase and allow instant or at least swift repricing (typically with no more than a claimed few minutes delay) of hundreds, if not thousands of products without human intervention. However, there is a wide range of (claimed and actual) characteristics. When and why a repricing algorithm decides to adjust prices can be different from case to case. While many off-the-shelf products offer to reprice products only based on fixed, predetermined rules, other products claim to run independent "A/B-tests" (i.e., testing consumers' and rivals' reactions to variants of selling policies) or to learn about competitive parameters such as own-price elasticity, rivals' strategies, and demand.

While the current interest in big data analytics and artificial intelligence (AI) makes the latter cases the focus of much conjecture about algorithmic collusion, even less sophisticated products can impact competition. Ezrachi and Stucke (2017), for example, discuss the justification strategies used by cartelists and assert that even an unsophisticated computer program increases the mental distance to any wrongdoing, making collusion easier. Brown and MacKay (2021) and Wieting and Sapi (2021) show that some of the observed algorithmic pricing behavior can be generated by simple pricing rules.

While our analysis is a snapshot of a rapidly evolving industry and is necessarily limited by what information firms share, we have already found a deep market for online stores. Sellers have a choice of hundreds of repricing algorithms that can be as cheap as a few dozen Euros per month. This is irrespective of whether sellers rely on marketplaces such as Amazon and eBay or own shops built on shop software such as Shopify or Magenta. As a result, we expect that algorithmic repricing might already be commonplace, especially in markets where the Internet is an important sales channel. At the same time, even firms without specific technical expertise can purchase high-end repricing algorithms that claim to make use of self-learning AI and features such as A/B-testing or demand elasticity estimation.

Researchers and authorities should therefore take note of algorithmic repricers as an important market feature especially for products for which the Internet is the main distribution channel. It is highly likely that many of these products are already priced by algorithms.[56] Where they aren't, adoption can potentially happen quickly due to the wide availability of affordable repricers. Nonetheless, there are price differences between the more and less sophisticated repricers that might put the most powerful repricing technology out of reach for small sellers, placing them at a disadvantage on marketplaces. To our knowledge, the literature has so far ignored differences in the fees and sophistication of algorithms that different sellers may use. This difference could be weighed against the possibility of free riding: smaller sellers might use simple rules-based algorithms to piggyback on demand-driven price changes implemented by a rival's sophisticated algorithm.

We suggest that algorithms are not just of interest because of the risk of high prices. They also give an insight into the objectives of the firm, due to the availability of strategies other than profit maximization. Finally, researchers and authorities might even be able to leverage the ability of repricers to identify competing products for their own purposes.

The paper is organized as follows. Section 3.2 provides an introduction to the class of algorithms that we call "repricer" and are the subject of this paper. Section 3.3 contains a general description of algorithms and repricers, an explanation of the methodology under which the data was collected. We present summary statistics in Section 3.4. Section 3.5 contains a discussion of categorizations and insights on algorithmic repricers based on our data. Section 3.6 concludes.

3.2 Introduction to algorithms and repricing

In its broadest definition, an algorithm is a computational procedure to solve a mathematical problem in a finite number of steps. Companies working in digital markets (and others) use increasingly sophisticated algorithms in everyday products and services, for example, to find more relevant results on search engines or to curate feeds on content platforms.

A repricer is an algorithm that recommends and sometimes implements price changes for products usually sold online. For example, a seller of electronic gadgets on a digital marketplace such as Amazon may use a repricer to continuously update the prices and inventory. The repricer may be able to change a large number of prices simultaneously and take into account a variety

[56] In a study of 2015, Chen et al. (2016) show that roughly 30% of prices in US Amazon Marketplace were most likely set by repricers, a figure that we expect to have considerably increased since then.

of data that surpasses the monitoring and analytical capabilities of a human decision-maker. For example, a repricing algorithm may observe a large number of rival products simultaneously and increase the prices of products for which competitors have run out of stock while decreasing the price for goods for which competitors are also reducing the price. A repricer might identify additional data helpful to predict demand and suggest prices according to its objectives.

One main distinction between "fixed-rule" algorithms and "self-learning" ones can be drawn. Fixed-rule repricing means the algorithm adjusts prices regularly based on predetermined and user-provided rules. Such a rule can be, for example, to always undercut the cheapest competitor by a small amount. Alternatively, it could be to adjust the price when it deviates too far from the median of a comparison group. Despite their seemingly simple, deterministic nature, these algorithms can potentially include many data and an extensive set of "if-then" rules that can condition the behavior of the repricing on multiple and sophisticated dimensions.

For example, a rules-based repricer might automatically identify relevant competitor products based on matching images and product descriptions. The user might then manually modify the selection of competitors, e.g., to tell the algorithm to ignore a particular competitor that charges very low prices. Other standard inputs for such an algorithm may include maximum and minimum prices or price-cost margins. Different rules-based algorithms are designed to continuously undercut rivals by some predetermined tick, and at regular intervals, they automatically "reset" the price to a high level. With these fixed rules in place, the algorithm monitors changes in the input data that may trigger a repricing event depending on the pre-specified rules. For example, a drop in a competitor's price may trigger a price decrease. An example is provided in Figure 25 in the Appendix. This screenshot from the repricing algorithm Prisync allows users to chain simple conditions into rules such as "I would like to be 5.00 EUR higher than the cheapest of all my competitors as long as I respect min margin/markup as Cost+5.00 EUR."

Fixed-rules algorithms have been used for repricing for several decades, for example, in the hotel and airline industries. Although these fixed-rule repricing can be pretty sophisticated, it should be understood that their behavior is entirely pre-designed and built into the algorithm. By contrast, self-learning algorithms are equipped with fewer rules and are designed to learn how to respond to market conditions autonomously. The ability to learn from market conditions allows these algorithms to adapt and change their actual behavior that may thus be different from its behavior under the same conditions observed at other points in time. Simplifying, the programmer of these algorithms indicates the possible goals, such as sales/stock turnover maximization or profit maximization, between which the seller can choose and change, depending on business strategy and current position in the marketplace. An example is provided in Figure 26 in the Appendix for the company Bqool, which allows users the choice between repricing strategies that vary between turnover maximization and per-unit profit maximization.

The learning method that transforms what the algorithm monitors at any point in time into some pricing decision is also specified, taking into account the goal set by the seller. Finally, these algorithms can be designed to actively experiment, sometimes using certain price levels just to determine how consumers and rivals react. These algorithms are considered more modern with respect to the fixed-rules algorithms and are an application of AI and Machine Learning in particular.

With self-learning, different kinds of algorithms mainly differ in three dimensions. First, if they learn, for example, whether they consider intertemporal trade-offs (e.g., considering profits over a long time horizon). Second, to what extent do they deal with experimentation while learning. Third, if they are model-free, thus abstracting entirely from the market environment, or if they instead contain pre-coded conditions and information, such as, for example, that price demand tends to increase when decreasing the price (information that could also be itself updated with learning), or that the algorithm can only consider using certain price levels.

Certain repricing companies claim to use these types of more modern algorithms that learn from market conditions and continuously improve over time.[57] Others instead focus on more mundane fixed-rule repricers. Clearly, these two types of algorithms have pros and cons. In particular, fixed-rule repricers lack the flexibility to adapt to unforeseen changes in market conditions. On the other hand, the flexibility with self-learning repricers may back-fire as learning and experimentation while the algorithm operates in the markets can be expensive for sellers.[58]

However, as seen from this brief discussion, repricers of either type can increase economic efficiency by promptly reacting to changing market conditions such as consumers' demand and possibly intensifying competition. They can also feed information into the pricing of products more systematically and comprehensively than human decision-makers. As shown in the introduction, there are also potential risks for consumers in spreading repricing algorithms. The most salient risk is the fear that repricing algorithms might allow directly or indirectly help sellers coordinate to raise prices, resulting in less competitive online marketplaces.

3.3 Data and methodology

In this section, we describe our approach to collecting the data, the method for identifying relevant repricers, and the variables we collect. We then present descriptive statistics of our data. As a first step, we identify many potential repricing tools for online markets. We spot companies using conventional online search for "repricing" and "dynamic pricing" tools and complement this with a market overview of so-called "Pricing Optimization Software" by Capterra, a firm that compares business software. Our search follows the industry's convention for naming these products. We expand our search through comparisons of pricing software by market research firms G2[59] and Gartner.[60]

We choose this method to find the largest possible number of mass-market third-party repricing software available for sale. A limitation of this method is that we do not observe pricing software developed by companies in-house or fully bespoke software (although we do observe some firms offering customized products). We take this approach because our focus is on the actual offer rather than the precise effects of algorithmic pricing software, especially for off-the-rack use in the mass market of online retail. To this end, we analyze pricing algorithms advertised for marketplaces rather than their behavior, as is common in the literature.

In a second step, we narrow the sample of pricing software to repricing tools capable of automatically changing sellers' prices. We thus exclude software that is purely focused on monitoring competitors' prices, cost-control, price classification (for example, to create price catalogues or to keep track of discounts and exchange rates) or making price suggestions without being capable of actively changing prices. With these two steps, we identify 185 distinct pricing tools by 130 different firms.

In a third and final step, we collect publicly available information from the websites of the repricing companies and by signing up for trial accounts and obtaining additional information where possible. For each firm in our sample, we analyze their range of repricing software and create one observation for each feature-distinct software product. We consider two products of the same firm feature-distinct if they offer a different set of features, for example, additional variables that are taken into account or a significantly higher repricing frequency. By contrast, we do not consider two products feature-distinct if different versions only differ in scale, for example, by enabling a retailer to connect to a greater number of marketplaces or to reprice

[57] One example is the algorithmic repricer by BQool that claims to have these features in contrast to a rules-based repricer. We will discuss this further in the next sections.

[58] In fact, often these types of algorithms are at least in part pre-trained offline, that is, in virtual environments that emulate the markets.

[59] https://research.g2.com/

[60] Gartner, "Market Guide for Retail Unified Price, Promotion and Markdown Optimization Applications", Robert Hetu, 6 January 2021. https://www.gartner.com/en/documents/3995226/market-guide-for-retail-unified-price-promotion-and-mark

a greater number of stock-keeping units (SKU).[61] This decision is motivated by our focus on studying the capabilities and varieties of repricing algorithms in the market. To account for tiered software that is not feature-distinct, we do, however, note the pricing structure (e.g., tiered by number of SKU, by revenue) and compute fees, where possible, for hypothetical sellers of different sizes.[62]

Finally, we analyze each website according to qualitative and quantitative criteria, which we define from the terminology and features we find in common between several sellers. We record the features that we consider most helpful and interesting in classifying and distinguishing websites. Our criteria reflect that repricing algorithms are heterogeneous in terms of their sophistication, transparency, pricing, features, scope, and breadth, ranging from industry-specific expert systems to all-purpose pricing software.

Apart from basic information about the firm selling the repricing software, such as the URL of their website and country, we collect information on the marketplaces where this software can be used, with additional information on which national Amazon marketplaces are mentioned, if any. We also identify if a firm offers pricing-related consulting (possibly with a more strategic rather than operational focus) with human experts as an add-on to their repricing service. We collect information on whether a firm makes certain specific claims about its product: whether it learns demand, uses AI, is based on game theory, is based on economics, is self-learning, uses dynamic pricing, and whether it uses price discrimination (or segmentation). We also note if the product is advertised only to B2C channels, for example, retailers with webshops, or B2B channels, helping manufacturers or wholesalers selling to retailers. We look for evidence of facilitating switching from rival services.

For the fees of the different tariffs of the repricers, we consider the following: the components of the overall service fee (e.g., setup fee, revenue thresholds, per number of SKU), the currency, the costs of contracting for one month for hypothetical small, medium, or large shops and whether a free trial exists. For this purpose, we define representative small, medium, and large shops, which we explain in detail further below. Practically all products in our sample are subscription services, often as a “Software as a Service” (SAAS) model, where customers purchase the use of the repricing algorithm for a specific period. Therefore, we also check for discounts for long-term contracts and, if such a discount exists, the shortest and most extended periods for which a buyer can contract and the implied discount.

We also analyze the following features of repricing algorithms: the declared number of price changes per hour, the maximum number of products that an algorithm can handle, whether the algorithm can focus on or exclude specific competitors, which pricing rules are used, whether an algorithm specifically targets or claims to win the Amazon Buy Box or similar (where applicable), and whether it keeps track of rivals’ stocks.[63] Finally, we record if the product comes with analytics and, if so, which form they take.

We also collect geographical information. To this end, we first look for information on the website about the company’s home country. If we do not find any further information, we assign a country based on office location. When several offices are listed, we assign a country based on the location of the first office listed.[64]

[61] SKU are code numbers that identify single inventory items at the lowest level of granularity for inventory control, for example selling or purchasing.

[62] To avoid confusion, we indicate with the fee the payment charged by the software firm for the repricing services and the price that the algorithm automatically determines for the seller.

[63] The Amazon Buy Box is the clickable white or yellow box, generally located at the top right part of the web page, that buyers may click to initiate their purchase of a particular product. It is estimated that more than 80% of Amazon transactions refer to products associated with the Buy Box. For this reason, competing sellers of the same product try to win it. Although the precise rule used by Amazon to temporarily allocate the Buy Box to a seller is not precisely known, a low price and high-ranked reviews are among the most relevant elements, although certainly not the only ones.

[64] Typically, nothing precludes that a software company based in one country deals with customers in other countries. Our focus on the English languages allows us to identify companies that operate in many countries.

To collect all this information, we rely on the repricers' websites and access the publicly available product demo where possible. Notice that our data set then records whether a product is advertised to have a specific feature rather than whether it is really present. In particular, the binary variables we use have to be read accordingly, e.g., a certain number of repricers are advertised to track rivals' stock, but other repricing algorithms may feature the same property without this being listed as an explicit product feature. At the same time, firms' claims about product features may be misleading or exaggerated. Finally, some omissions concern product features that lend themselves to advertising, so we would expect them to be mentioned if they apply, for example, if a repricing algorithm learns demand, while others may be considered limitations, and it is not clear why some firms do mention them at all.

Table 28 in the Appendix lists and describes all the variables that form our analysis.

3.4 Descriptive analysis of repricers

A first interesting observation is the distribution of locations for the 130 repricing firms, Table 4. About a third of the 185 repricing algorithms they offer is associated with companies located in the United States (32.4%), with the United Kingdom (11.9%), France (7.6%), and Germany (7.6%) the next most common locations. Overall, 16 repricing algorithms (8.6%) are associated with the Asia-Pacific region (with entries from Australia, India, New Zealand, Singapore and Taiwan) and 2 (1%) with South America. Clearly, this overview shows some gaps, most notably in China which we expect to have a domestic industry. Our English-language search is also more likely to miss firms that do not offer websites in English.[65] We expect our results, therefore, to be most applicable to Europe and North America.

However, it is not clear that the country where a repricing firm is located restricts its ability to sell services globally (apart from reasons unrelated to the market, such as international sanctions). We do not find any statements that indicate geoblocking or other geographical restrictions. The accessibility of online marketplaces is another, possibly more relevant, constraint, such as the different national Amazon marketplaces that serve different geographies (e.g., www.amazon.it, www.amazon.de). Sometimes, a repricing company with its main office in the US has relocated there after an initial phase in another country. This is consistent with the fact that many startups, especially in information-technology-related areas, find better access to financing in the US.

Of 185 repricing algorithms, 69 (37%) list specific marketplaces, such as Amazon and eBay, for which they are suitable. The remainder either make no explicit statement on the marketplaces they integrate with or offer connection via API to existing webshops. We list those as "platform-independent". Of those 69 repricing algorithms that name compatible marketplaces, 41 (59% of 69) list more than one marketplace. The ability of repricers to operate in several marketplaces raises possible concerns related to theories of strengthening collusion through multi-market contact, which in this case could take place on different platforms and thus be more challenging to identify. It would be interesting to see if algorithms learn from other marketplaces simultaneously. For example, in an across-platforms coordination scheme, repricing algorithms may react to price cuts by a competitor on one platform with price cuts across several platforms to inflict harsher punishments on competitors that deviate from high prices.

The most commonly mentioned marketplace is Amazon (57 mentions, 83% of 69), followed by eBay (29, 42%), and Walmart (9, 13%). Some repricers mention price comparison websites, such as Google Shopping (6, 9%). 16 other marketplaces were mentioned five times or less. Some of these less frequently mentioned marketplaces are likely specific to the main region or country where the repricing firm is active, such as Mercado Libre (South America), Fnac Darty (France), Kaufland.de (Germany). The full list is reported in Table 5.

[65]Some websites we covered are not available in English or only translate a limited part of their website to English, for example, https://by.prexus.co/

Table 4: Number of repricing firms by country of origin

Country	Frequency	Country	Frequency
United States	37	Bulgaria	1
United Kingdom	15	Canada	1
France	12	Colombia	1
Germany	12	Czech Republic	1
Netherlands	8	Denmark	1
Spain	7	Finland	1
Poland	5	Luxembourg	1
India	4	New Zealand	1
Israel	4	Portugal	1
Italy	4	Russia	1
Australia	3	Serbia	1
Switzerland	2	Singapore	1
Turkey	2	Sweden	1
Brazil	1	Taiwan	1
		Total	130

Table 5: Number of compatible repricing algorithms for different marketplaces

Marketplace	Count of repricing algorithms	Marketplace	Count of repricing algorithms
Amazon	57	Etsy	2
Ebay	29	Bonanza	2
Walmart	9	Wish	2
Google Shopping	6	Mercadolibre	2
Cdiscount	5	Facebook	2
Cdiscount	5	Instagram	2
Fnac	4	Tesco	1
Darty	3	Rakuten	1
Manomano	3	Kaufland.de	1
Rueducommerce	3	Idealo	1

In addition to the observation that most repricing algorithms are geared towards sellers on Amazon, it is worth mentioning that Amazon offers a free repricing service as well.[66] Amazon claims that its tool can "adjust your prices quickly and automatically against your competition" to increase sellers' chances of winning the Buy Box, which Amazon also calls the "Featured Offer", the first offer that is shown on an Amazon product site. The main listed features of Amazon's repricing algorithm are the options to select a "pre-configured 'competitive price rule' or customize your strategy, choosing whether to compete based on feedback rating, fulfilment channel, and other options". Amazon's repricer is based upon simple, fixed rules that allow sellers to start and stop repricing for designated products and to set price limits.

Algorithmic repricers also advertise being compatible with common enterprise software packages, including marketplace software, and customer relationship management (CRM)[67] software, or enterprise resource planning (ERP)[68] software. Some repricers advertise compatibility with industry-specific revenue management software solutions, such as Marketron (media), Wideorbit (premium ads) or Sqills (transportation). In total, 43 of 185 repricing algorithms mention inte-

[66] https://sell.amazon.com/tools/automate-pricing

[67] CRM encompasses a firm's practices in dealing with its clients, including prediction and analysis of customer trends, sales, after-sales and support.

[68] Companies use ERP to integrate functions such as inventory planning, sales, marketing, human resources, and finance into a unified and comprehensive system.

gration with such tools (23%). The most commonly named ones are tools that allow sellers to create their online shops, such as Shopify (13 mentions, 30% of 43) and Magento (6, 14%). 35 other software packages are mentioned 5 times or less, with some examples including Tableau (data visualization), LS Retail (retail point-of-sale software), or Microsoft Dynamics (a business software suite), each of which is mentioned once. The complete list is reported in Table 29 in the Appendix.

3.5 Analysis

In this section, which is the core of the paper, we analyse our sample of repricing algorithms. We investigate, in turn, the information related to the different types of companies that offer algorithmic repricing in Section 3.5.1, their fees in Section 3.5.2, learning and AI in Section 3.5.3, repricing speed in Section 3.5.4 and the relevant economic implications for economists in Section 3.5.5.

3.5.1 Analysis of algorithmic repricing companies

Repricing software is offered by various companies that can generally be ordered according to the breadth of their product portfolio. We propose the following four characterisations of typical repricing companies: a) startups that only offer repricing algorithms, b) pricing consultants that extend pricing consulting services with an algorithmic repricer, c) price management and observation companies that offer algorithmic pricing software either as an extension of their existing software, as a bundle, or as a standalone product, and on the broad end of the spectrum d) sellers of corporate software that offer repricers as one of many tools in their product portfolio.

Some repricing software companies are also specialised in terms of the industry that they target. This dimension is prevalent for the b)-type of firms, which might have some long-standing experience with a specific sector (e.g., Heating, Ventilation, Air Conditioning (HVAC), government procurement, transport, energy contracts) and leverage this knowledge into a repricing software. In contrast, especially type-a) but also type-c) firms typically stress the universality of their algorithm to generate prices based on competitors' prices and stocks without regard to the industry. Other remarkable variations on repricing offers include i) promotional planning (pricing discounts) and ii) analysis of "price-image", i.e., influencing customer perception about the general price level of a seller.

In addition to this customer-specific technical expertise, several repricers are advertised alongside additional offers for pricing expertise. In particular, 49 (26 %) of repricing algorithms are promoted jointly with some pricing consulting services, typically purchased as an optional add-on. These can take the form of strategy training or workshops with customer sales teams or individualised algorithms, where the pricing expertise is either provided by the seller of the algorithmic pricing software or in collaboration with a third-party consultant.

As for the specific sellers' sector, we identify ten different specialisations among companies that offer repricing algorithms for specific industries: i) pricing tickets for the entertainment industry, ii) pricing for the hospitality industry (hotels, short-term rentals), iii) power purchasing agreements (and energy markets in general), iv) transport and logistics, v) HVAC, electric and plumbing, vi) tyre manufacturers, vii) travel ("capacity constrained industries"), viii) revenue management for airlines, ix) corporate banking, x) telecommunications. We were particularly surprised to find examples for viii) as we would have expected this ability typically to be in-house with large airlines.

While some repricing firms advertise almost identically to B2C and B2B companies, some offers are geared explicitly towards B2B offerings or firms specialised in government contracts. While we do not consider them as repricing software, we believe these adjacent pricing products worth mentioning. Some specialise in generating quotes or supporting sales teams on request rather than continuously changing a publicly advertised price. Specific features for the benefit of

B2B companies include preparation of compliance-relevant paperwork for procurement contracts, management of complex rebate and discount schemes, and electronic management of complex pricing catalogues (e.g., in the case of an automotive company that has an extensive catalogue of product specifications, spare parts, pricing and discount schemes, sometimes across several currencies).

Repricing algorithms can be classified by use-case or by functionality. Use-cases generally fall into two categories: repricers for third-party marketplaces and shops (Amazon, eBay, Walmart) and repricers for sellers' own shops. The former kind seems to be more standardised: these repricers typically offer detailed information about prices and product features. By contrast, solutions for sellers' websites are more likely to be bespoke, with individual and non-public prices and sales through a personal demo.

3.5.2 Repricing fees

We find that rather than selling software as a one-time purchase (like a video game or a piece of office software on CD in the early 2000s might have been sold), repricing algorithms are typically sold in a subscription- and licensing model. Customers buy access to pricing software through monthly or annual subscriptions. Only half of all repricers come with sticker prices posted. The other half require contacting the seller to get an individual quote. We would expect sellers of algorithmic repricing software also to use sophisticated algorithms or some data-driven approach to choose the fees for their own services. In Table 6 we tabulate the main fee models we observe. In 93 cases (50%), no information about the fees is publicly posted. Another 64 (35%) repricing algorithms are offered at a fixed fee for a fixed limit to the functionality (typically including a maximum number of SKU or marketplaces). The remaining ones offer flexible fees that scale with the number of items, revenue or some mixed model, for example charging separate setup fees, thus inducing a pricing structure featuring a two-part tariff.

Table 6: Frequency of repricing fee rules

Main determinant of repricing fee	Number of repricers	% of total
Individual quotes	93	50
Fixed fee	64	35
Number of items	12	6
Mixed rule	6	3
Other	6	3
Total revenue	3	3
Total revenue plus fixed fee	1	1

Percentage relative to total of 185 repricers, difference to 100% from rounding.

Based on the information we collected from the websites, we have computed the fees for three hypothetical shops based on the following characteristics[69]:

1. Small shop: 10 items, 100 sales, 10,000 € monthly turnover, 1 national Amazon marketplace or sales platform

2. Medium-sized shop: 500 items, 5,000 sales, 500,000 € monthly turnover, 7 national Amazon marketplaces or sales platforms

3. Large shop: 15,000 items, 150,000 sales, 5 mio. € monthly turnover, 10 national Amazon marketplaces or sales platforms

[69] These numbers were also informed by personal communication with actual sellers.

We measure the fee for a one-month subscription, including one-time setup costs, separate by currency.

Table 7 presents the distribution of fees for these three hypothetical shops. We distinguish between offers in USD and Euros. There are only three products with fees in other currencies (two in GBP, one in BRL). We observe 50 repricers that list fees in "EUR" and 31 in "USD" (the websites were accessed from Eurozone countries). So, only 44% of repricers (84) come with a publicly posted fee. The overall range of fees is extensive. While several offers cost no more than a hundred Euros or US-Dollars to a small shop per month, the most expensive offers can be several times more costly. In this context, it would be interesting to learn whether this fee disparity is driven by differences in value or uncertainty about the value delivered by repricers. The wide range of fees, as well as the proliferation of trial versions and live demonstrations, shows that convincing customers of the value-added of a repricer is a major challenge.

Nonetheless, another takeaway from the analysis of the general price level is that algorithmic repricing is quite affordable. Sellers of different sizes have low-cost options to price online stores with repricing algorithms. Therefore, we should expect many sellers, even relatively small ones, to adopt pricing algorithms quickly if they haven't already done so.

Table 7: Repricing algorithm costs

	Currency	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Small shop	EUR	20.00	81.50	197.00	402.92	449.00	2490.00
	USD	0.00	20.50	75.00	755.09	124.50	14000.00
Medium shop	EUR	25.00	121.25	274.00	501.76	649.00	2490.00
	USD	4.00	28.21	99.00	911.00	186.50	14000.00
Large shop	EUR	59.00	144.50	304.00	627.57	860.00	2490.00
	USD	8.00	97.50	149.00	1478.61	772.50	14000.00

The minimum of 0 is due to one firm, Priceva, which offers a permanent free tier for their repricer that is limited to 20 products.

In Figure 6 we present a boxplot of long-term discounts. So far, to put fees on a comparable scale, we have compared fees for the shortest possible subscription, typically one month. As for many other subscription products, such as newspapers or gyms, many firms offer discounts for longer subscription periods. We find 38 repricers (21 %) that offer long-term discounts, as opposed to 65 that don't (35%) and 82 (44%) for which no information is available. We report the distribution for the 103 entries that are not missing to give readers a clearer idea of the average fees for repricers that sellers would pay in a regular market environment where discounts play some role. The discount is for an annual subscription relative to a one-month contract, with few exceptions. The descriptive statistics are in Table 8. Long-term discounts are offered by a sizeable minority of repricers (and possibly some of those for which information on fees is missing) and lowers the final fee to sellers by typically between 5 and 10% (up to 40% in one case).

Table 8: Summary statistics of long-term discounts

Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
0%	0%	0%	6%	10%	40%

3.5.3 Learning and AI

Given the various approaches, a useful distinction is the one introduced above on fixed-rules and self-learning repricers, although a diverse spectrum of more and less sophisticated repricing tools shows up. As an example, in one case, a company mentions that they "develop mathematical and

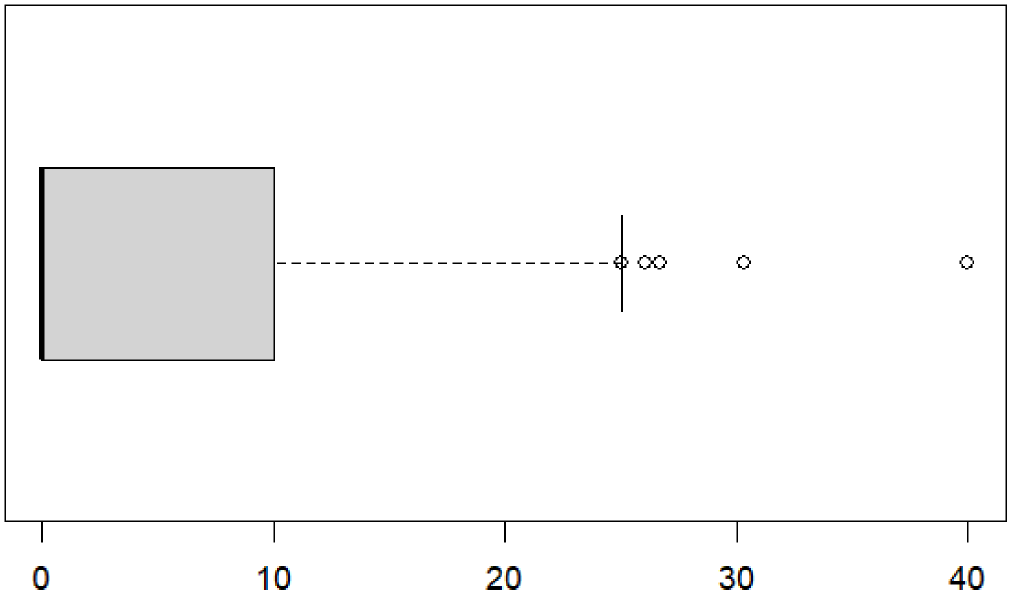


Figure 6: Distribution of long-term discounts in percent

computational methodologies based on statistics, physics and mechanics, as well as time series analysis and operational research to understand and predict events and behaviors."[70] Another company offered "14 different time series methods, including recurrent neural networks, deep learning, probabilistic, ARIMA and other models to choose from."[71] Even though we could not effectively verify these features, as explained, the presence of these precise details seems to indicate that they should be genuine features, as it would be relatively simple to verify if that is not the case.

Many tools offer repricing without claims of using AI. Instead, they typically allow the user to set rules (e.g., minimum and maximum prices for products, desired position towards certain other products or sellers). Notably, while most repricers stress their ability to maximize profit, repricers are also advertised to be able to pursue other goals, such as maximizing revenue or achieving targets in terms of inventory management/turnover speed. This possibility could indicate that repricers are used as tools in more complex organizational settings, perhaps with delegated management and managers purposefully incentivized to maximise revenue or some target other than profit.

We are particularly interested in the technical features of repricing algorithms, precisely how many of them claim to include AI or claim to be self-learning. Out of 185 repricing algorithms, 101 (55%) explicitly claim to use AI, but only half as many indicate in some sense that their algorithms are self-learning (50 mentions, 27%), that is AI that trains itself and determines the pricing strategy autonomously, using unlabeled data. Admittedly, though, we do not know whether marketers that advertise their algorithms as "self-learning" apply this definition, and we reached the view that some of these claims are far-fetched. We find it most helpful to distinguish

[70]https://premoneo.com/en/solutions/
[71]https://www.remi.ai/demandforecasting

between more and less sophisticated repricers. We classify all algorithms that claim to either use AI or to be self-learning as more sophisticated and the remainder as less sophisticated (keeping in mind the previously mentioned limitations). One main purpose of this distinction is to compare prices between these two groups. Just over half (106, or 57%) of repricers are more sophisticated by this definition.

The number of algorithms that claim to learn demand, or demand elasticities, falls in-between those claiming to use AI and those claiming to be self-learning, at 74 (40%). Only 10 repricing algorithms (5%) are claimed to be built using game theory and 15 (8%) explicitly reference economics. 127 repricing algorithms (69%) are described as pricing dynamically and 158 (85%) offer additional analytics. About half of all algorithms (92, 50%) can target specific competitors (i.e., condition pricing rules on their prices or exclude specific firms), but only 58 (31%) actively track rivals' stocks (for example, to increase prices when rivals run out of stock).[72]

We also look at the subset of the most expensive repricers. This analysis is limited by a large number of missing values for the fee structure of repricers because most firms do not report their fee structure but instead rely on individual offers. We tabulate the technical features separately for those 34 repricers with above-median prices for the hypothetical large shop. The frequency of different technical claims for the full sample as well as the 34 most expensive repricers is shown in Table 9.

Overall, no clear picture emerges. One might expect technically sophisticated features such as self-learning and AI to be more prevalent in this sample but they are either less frequent (AI use, demand learning) or only slightly more prevalent (27% of all repricers claim to be self-learning, but only 32% of the most expensive ones). Due to the low sample size and many missing values for fees, we refrain from interpreting too much (or conducting regression analysis) as the results would likely be misleading. On the hand we will come back below looking at this point from another perspective, that is identifying algorithms that incorporate sophisticated features, finding in that case significant differences in the fees.

Rather than comparing the features of the most and least expensive repricers, we can also compare repricer fees by category. Using our proposed classification of self-learning and AI-based repricers, as opposed to "simpler", fixed-rule repricers, we find that the latter tend to be much cheaper. For example, the hypothetical large shop would face a monthly mean price of 1967.6 Euros for a "self-learning" repricer versus 338.68 Euros for one of the remaining repricers (median: 617.9 Euros and 195 Euros, respectively). Sophisticated repricers end up costing ca. 6 times more than the simple ones. Similar figures emerge for medium-sized and small shops, with self-learning and AI-powered repricers costing ca. 4-5 times more than other repricers. While these prices might seem small relative to the turnover of large sellers, the difference in fees between the most and least sophisticated algorithms may prevent smaller sellers from adopting the most sophisticated technology. This can put smaller sellers at a disadvantage on marketplaces where they compete with larger sellers for which even a sophisticated repricer is but a small expense. This disadvantage could be weighed against the possibility of free riding: smaller sellers might use simple rules-based algorithms to piggyback on demand-driven price changes implemented by a rival's sophisticated algorithm. This could limit the attractiveness of investing into a sophisticated repricing algorithm in the first place.

Regarding the use of AI to choose prices, repricers vary drastically. Most companies that claim to use AI or algorithms for pricing do not provide detailed information about how their models are trained, what kind of model is used, or what data is used to train the repricers. When information about the algorithm's training is provided, it typically falls into one of two categories of learning: either the repricing company uses historical sales data of the seller to train its algorithm, or it conducts its own A/B tests. In both cases, the repricing companies often claim to learn demand elasticities to inform their models.

Other firms claim to "estimate demand models". It is questionable whether the use of these

[72] See below for an analysis of correlations between these features.

expressions in the self-marketing of these products is equivalent to the use of these terms in the economics literature (e.g., in empirical Industrial Organization), if for no other reason than the lack of detailed explanation (e.g., which kinds of demand model are considered by the algorithm and how the applicable demand model is chosen). Other approaches to learning demand are mentioned as well. These include conjoint analysis[73] and the use of historic data. An example of the latter case is when a repricing company advertises the size of its proprietary data set rather than emphasizing the use of client data.

Looking at the correlation between these different features in Table 10, we see claims of demand-learning, AI, and self-learning to appear together often. Similarly, claims to track rivals' stock and to target specific rivals often appear together, unsurprisingly. Interestingly, a few features have a (small) negative correlation, suggesting that they appear together less frequently. For example, game theory/dynamic pricing and economics/analytics are two such pairs, as are demand/target competitors and AI/analytics. Self-learning is negatively (albeit only slightly) correlated with tracking rivals' stock and targeting competitors, as are analytics and tracking rivals' stock.

The correlations should not be misunderstood as being quantitatively interesting in and of themselves. Instead, they help us identify those pairs of features or terms that often appear together as relevant features offered in the market, thus allowing to classify products. We can then identify some products for which both of these terms appear and understand why they do so, thereby guiding an essentially qualitative examination of the sample. Understanding *when* these terms related to the theoretical and technical sophistication of the algorithm appear together is often simple, understanding *why* is harder. Most commonly, these terms appear together in white papers, blogs, or additional resources. We suggest that some repricing firms differentiate themselves by discussing the technical background of their repricing algorithms as a way to signal quality.

These correlations and a more qualitative assessment of the studied websites allow a rough distinction between two broad categories for repricers. We find a cluster of AI-first, technology-driven repricing algorithms that, in the very extreme, are very narrow products that seek to integrate with existing revenue-management software. These products may sometimes forego more basic features such as data analytics, as the requirement of input data implies that a customer must have revenue management and analytics already in place. In terms of our analysis, these products are advertised as using "self-learning" and "artificial intelligence", sometimes powered by "game theory" or "economics".

A different group of products includes broader price management systems that sometimes offer repricing as an add-on on top of a range of more basic analytics or competitor monitoring features. This includes the family of rules-based repricing algorithms that do not claim to learn competitors' behavior. They instead focus on identifying relevant competitors and then allow the implementation of mechanical rules (e.g., always undercut the lowest rival offer, always track the median price of a comparison group, subject to a specific price corridor, minimum margins etc.)

3.5.4 Repricing speed

One of the main attractions of repricers is their ability to change prices across a large inventory faster and more frequently than humans can. A total of 50 repricers made some statement about repricing speed. Interpreting a rather vague but common claim of "instant repricing" as up to 3600 repricing events per hour (one per second), we find a range between 0.042 (once per day) and 3600 (once per second). Table 11 reports some summary statistics on these claims. Almost half

[73]In conjoint analysis, a group of respondents is shown a sequence of products pairs with different characteristics. Respondents have to communicate a preference between two shown products. By repeating this procedure and systematically varying the product characteristics, conjoint analysis helps to identify hedonic preferences, i.e., preferences over product features.

Table 9: Frequency table of technical features of repricing algorithms

	All repricers		Most expensive repricers	
Feature	Yes (%)	No (%)	Yes (%)	No (%)
Learns demand	111 (60%)	74 (40%)	9 (26%)	25 (74%)
Uses AI	103 (56%)	82 (44%)	15 (44%)	19 (56%)
Uses game theory	10 (5%)	175 (95%)	1 (3%)	33 (97%)
Uses economics	15 (8%)	170 (92%)	1 (3%)	33 (97%)
Is self-learning	50 (27%)	135 (73%)	11 (32%)	23 (68%)
Uses dynamic pricing	127 (69%)	58 (31%)	27 (79%)	7 (21%)
Offers analytics	158 (85%)	27 (15%)	31 (91%)	3 (9%)
Targets specific competitors	92 (50%)	93 (50%)	22 (65%)	12 (35%)
Tracks rivals' stock	58 (31%)	127 (69%)	15 (44%)	19 (56%)

Based on repricing firms' claims, e.g., 50 out of 185 repricers are claimed to be "self-learning", 135 make no such claim. "Most expensive" repricers selected as 34 repricers with above median fee for the large hypothetical shops when the fee was available.

Table 10: Correlation matrix between technical features

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Learns demand [1]	1.00	0.68	0.24	0.12	0.45	0.10	0.03	-0.06	0.04
Uses AI [2]	0.68	1.00	0.17	0.19	0.47	0.10	-0.03	0.15	0.06
Uses game theory [3]	0.24	0.17	1.00	0.28	0.07	-0.10	0.10	0.19	0.25
Uses economics [4]	0.12	0.19	0.28	1.00	0.09	0.12	-0.10	0.10	0.18
Is self-learning [5]	0.45	0.47	0.07	0.09	1.00	0.15	0.11	0.00	-0.02
Uses dynamic pricing [6]	0.10	0.10	-0.10	0.12	0.15	1.00	0.12	0.11	0.13
Offers analytics [7]	0.03	-0.03	0.10	-0.10	0.11	0.12	1.00	0.07	-0.02
Targets specific competitors [8]	-0.06	0.15	0.19	0.10	0.00	0.11	0.07	1.00	0.59
Tracks rivals' stock [9]	0.04	0.06	0.25	0.18	-0.02	0.13	-0.02	0.59	1.00

Column headers replaced by index numbers to save space.

of the 50 repricers (21, or 42%) that give any information about repricing speed claim "instant repricing," and only three repricers (6%) reprice only once a day.[74]

Table 11: Summary statistics for repricing frequency

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.04	1.00	30.00	1553.55	3600.00	3600.00

Sample refers to 50 repricers out of the full sample which report information about repricing speed. Repricing speed measured as frequency of repricing per hour.

Repricers that focus on Amazon emphasize the importance of winning the Buy Box, noting that 80-90% of sales happen there. They may hint at the opacity of the BuyBox algorithm to induce urgency for the reader. They will also emphasize (more than other repricers) the speed at which they reprice, some of them claiming that Amazon limits 3rd-party tools to up to 30 repricing events per hour, while others claim not to suffer from any of these constraints. From the authors' experience with Amazon, Amazon limits the frequency of price changes, so even if a repricer claims to adjust prices frequently, a marketplace may still set binding and relevant constraints on the frequency of price changes, which some sellers may not know.

[74] Those three algorithms are related to only two firms, one of which focuses on e-commerce, and one on hotel room pricing.

3.5.5 General implications

The use of AI is not only limited to learning and applying pricing strategies but very often also for discovering related and competitively relevant products. It is unclear whether this only refers to recognising similar product names, descriptions, and images or if this also includes using AI to find products with high price correlations or other economic features (finding products that pose a competitive constraint). In the latter case, this could be an interesting and novel case of AI-driven market definition. This use of pricing algorithms could be potentially interesting for antitrust authorities. However, the way this feature is typically described, the former seems more likely. Product discovery and market observation are closely linked to algorithmic repricing because competitors' prices are typically a necessary input variable to the suggested price. Unsurprisingly, many repricing software come with analytics and business intelligence tools that allow competitive monitoring. This feature is most common for the c)-type of firms.

In particular for repricing software that are advertised to "brands", "manufacturers" or any other kind of B2B business, repricing is commonly offered alongside tools to monitor compliance with retail price maintenance, i.e., observing whether merchants are undercutting the manufacturer on its own product. The common industry jargon is "MAP violations" (MAP = minimum advertised price). The legality of this practice and compliance in regions where retail price maintenance is banned is an open question.

The price monitoring feature for discovering competitively relevant products is potentially of great practical relevance to research economists and practitioners, such as competition authorities. In terms of market definition, an AI-based identification of competitively relevant products, mainly when it includes the estimation of a complete demand system of cross-price elasticities, could provide a more accurate and systematic approach to product market definition. Conceptually, product market definition in competition economics relies on identifying all products that provide a competitive constraint to each other and are jointly worth monopolising. Practically, the need to limit quantitative investigations of product market definition (for example, to respect the resource constraints of an agency) requires a selection of "plausible" candidate products that could form part of a relevant market. AI-powered product market definition could also impact empirical research in industrial organisation.

Few algorithmic repricers mention the economic underpinnings of their algorithms. Economic theory is rarely mentioned in the description of repricers, game theory is sometimes mentioned. To the extent that game theory is discussed as a consideration in the design of the repricer, this often refers to the avoidance of price wars, comparing pricing to either a prisoner's dilemma or some variation on the Cournot or Bertrand game. However, economic theory informs some of the strategies implemented by repricing algorithms or the reasons for their adoption. Similarly to how the Bertrand pricing game in the managerial/MBA education is framed in terms of "escaping the Bertrand trap", some repricers mention how they help avoid the "Bertrand trap".

Other strategies and pricing rules involve cyclical prices, defined by undercutting other sellers to capture the Buy Box and resetting to a higher price once a predefined minimum price is reached. One repricing firm calls this an "oscillation" or "oscillate" strategy.[75] Another way to think of such a strategy is "tit-for-tat" with forgiveness from the prisoner's dilemma literature. Under this interpretation, the seller using the repricer retaliates against price cuts with own price cuts but "forgives" by resetting to high prices to avoid permanent price wars.

Finally, the data sheds some light on the potential for price discrimination in the form of "personalized pricing". While theoretically well understood by economists, personalized pricing has been rarely observed in practice. With the advent of data-driven, real-time pricing, however, there have been increasing predictions that personalized pricing will become more common (Goldfarb and Tucker, 2019). In our data, 59 repricers (32%) claim to offer either personalized or

[75] Bqool mentions this strategy on a website, `https://www.bqool.com/products/rule-based-repricing-central/`, and provides a brief characterization in a video, last retrieved from YouTube on 27.05.2022, `https://www.youtube.com/watch?v=zuyBGieuj_E`.

segmented pricing. However, it is clear that repricing companies try to disassociate themselves from the negatively charged term "price discrimination".[76] In our survey of real user interfaces, we did not find any obvious examples of price discrimination or segmentation such as "charge 5% extra for users on an iOS device". Further research documenting and describing personalized pricing is therefore needed.

3.6 Conclusion

This paper presents a snapshot of the relevant and nascent repricing industry that offers software algorithms for automating pricing decisions in online markets. We offer a categorization of the players in this industry and a classification of repricing algorithms. A useful distinction can be made between those self-learning repricing algorithms and those that instead use fixed rules. This paper shows the main advertised features of 185 repricing algorithm and their prices. Although this approach does not allow for precise verification of what these repricers are effectively capable of, it still offers an exciting and, to our knowledge, unique view of this industry.

Regulators, researchers, and consumers should note that affordable, off-the-rack pricing algorithms are widely available online. In markets where the Internet is a major distribution channel, it is therefore not unlikely that many sellers are already actively making use of algorithmic repricing or have the possibility to adopt this technology soon. At the same time, it shows that increasingly sophisticated AI capabilities are available for sale. Our findings vindicate the urgency with which economists have advanced the literature on algorithmic pricing in the past years. Yet we also hint at significant differentiation in the fees and sophistication of repricers. This difference is consequential because it can affects the level of sophistication and effectiveness of pricing strategies that small and large seller can afford, with material consequences for the structure of online retail markets and consumers' surplus.

Of particular interest for economists should be that pricing algorithms are not exclusively advertised as tools for profit maximization but also to optimize other company KPIs, such as revenue or inventory turnover speed. This observation hints at potential future research of algorithmic pricing in delegated management scenarios and institutional setups where managers are instructed to pursue other goals than just profit maximization.

Competition authorities might be alerted to the use of repricers to maintain vertical restraints, such as minimum advertised prices. At the same time, some capabilities of repricers, namely the identification of relevant competitor products, could be seen as helpful in market definition and point towards a more sophisticated method of market definition in competition policy. Tools that can identify pricing constraints, substitutes and complements from an unfiltered, potentially large universe of products could help better implement existing approaches to market definition. We leave for future research the analysis of the practical features of the algorithms on offer and their implications.

Acknowledgments

We are grateful to Sara Guidi for her help in collecting the data and thank for comments Maximilian Schaefer, Vito Bramante and Emilio Calvano.

References

Assad, Stephanie et al. (2020). "Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market". In: *CESifo Working Paper No. 8521*.

[76]For example, this article discusses price discrimination alongside a wider scope for personalized dynamic pricing which includes price matching in response to consumers discovering cheaper competitors, https://competera.net/resources/articles/dynamic-pricing-personalization.

Brown, Zach Y. and Alexander MacKay (2021). *Competition in Pricing Algorithms*. Working Paper 28860. Series: Working Paper Series. National Bureau of Economic Research. DOI: 10.3386/w28860.

Calvano, Emilio et al. (2020). “Artificial Intelligence, Algorithmic Pricing, and Collusion”. en. In: *American Economic Review* 110.10, pp. 3267–3297. DOI: 10.1257/aer.20190623.

Chen, Le, Alan Mislove, and Christo Wilson (2016). “An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace”. en. In: *Proceedings of the 25th International Conference on World Wide Web*. Montréal Québec Canada: International World Wide Web Conferences Steering Committee, pp. 1339–1349. DOI: 10.1145/2872427.2883089.

Ezrachi, Ariel and Maurice E. Stucke (2017). *Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy*. en. Harvard University Press.

Goldfarb, Avi and Catherine Tucker (2019). “Digital Economics”. en. In: *Journal of Economic Literature* 57.1, pp. 3–43. DOI: 10.1257/jel.20171452.

Johnson, Justin, Andrew Rhodes, and Matthijs R. Wildenbeest (2020). “Platform design when sellers use pricing algorithms”. In: *Available at SSRN 3753903*.

Klein, Timo (2021). “Autonomous algorithmic collusion: Q-learning under sequential pricing”. en. In: *The RAND Journal of Economics* 52.3, pp. 538–558. DOI: 10.1111/1756-2171.12383.

Schwalbe, Ulrich (2018). “Algorithms, machine learning, and collusion”. In: *Journal of Competition Law & Economics* 14.4, pp. 568–607. DOI: 10.1093/joclec/nhz004.

Wieting, Marcel and Geza Sapi (2021). *Algorithms in the Marketplace: An Empirical Analysis of Automated Pricing in E-Commerce*. en. SSRN Scholarly Paper ID 3945137. Rochester, NY: Social Science Research Network.

4 Algorithmic pricing - a literature review

Abstract Economists have studied the use of pricing algorithms to understand whether these algorithms learn to collude. In doing so, they have ignored decades of research into algorithmic learning. Still, the economics literature on algorithmic pricing has developed a better understanding of the circumstances under which algorithmic pricing leads to higher prices, replacing earlier concerns about rampant algorithmic collusion. I summarize findings from the economics literature across computational, experimental, and empirical methods as well as adjacent fields. I argue that a lack of understanding of buyer responses to algorithmic pricing cycles and endogenous adoption of algorithmic pricing are the main gaps in the literature.

Keywords: algorithmic pricing, repeated games, collusion, learning in games
JEL: C62, C73, K21, L41, L81

4.1 Introduction

A rapidly growing literature has investigated the role of pricing algorithms on prices, particularly on online marketplaces. This practice is known as "revenue management" among airlines and in the hospitality industry. In retail, sellers increasingly let algorithms determine prices for their goods. This has sparked interest both from a theoretical and a policy perspective. Through the lens of algorithmic pricing, economists hope to better understand competition, for example, what strategies algorithms learn and pursue. Meanwhile, competition lawyers and antitrust authorities are concerned about the potential of algorithms to coordinate on setting high prices or to exhibit other potentially undesirable behaviors in the marketplace, such as discrimination.

As a result, in recent years economists and other researchers have produced a body of literature that has approached these questions from different angles. Some papers take a descriptive approach to the phenomenon of algorithmic repricing or the industry supplying algorithmic repricers, while others make predictions about the behavior of algorithms based on simulations or economic theory. Finally, an emerging empirical literature studies the effect of algorithms on prices and documents the effect of algorithms on pricing patterns.

This paper organizes the literature, from the initial engagement of economists with the existing legal and computer-science literature on pricing algorithms, to the methods and questions contributed by economists. Its purpose is to help academic economists to understand frontier research across methodological approaches on this research area and to identify new research questions. I do not focus exclusively on collusion but I discuss the determinants of algorithmic collusion found in the literature. I provide an overview of data sets and empirical studies and their findings.

Pricing algorithms have been touched upon in several surveys of the relationship between artificial intelligence/machine learning and economics/economists (Varian, 2019; Athey and Luca, 2019; Abrardi et al., 2022). However, other dedicated literature reviews of algorithmic repricers precede the emergence of the economics literature that contributes substantive analyses of algorithmic behavior (Pohlmann and Schütte, 2018; Van Uytsel, 2018).

The initial concerns that algorithms would always find collusive strategies with ease have largely not been confirmed empirically. Instead, even papers that do find evidence for higher prices as a result of algorithmic pricing do so only under specific circumstances. It is already clear that the effect of algorithmic pricing depends on market circumstances, including both the market structures, how many competitors adopt pricing algorithms and product characteristics. However, there is little general understanding about the market features that enable harmful outcomes of algorithmic pricing, so the question remains which insights generalize into new contexts.

Furthermore, consumers have an increasing number of tools to automatically monitor price cycles on online marketplaces such as Amazon. So even when pricing algorithms lead to higher prices on average, it is not clear that consumers also purchase at this higher, average price, rather than at a low-point in the pricing cycles that have been documented by the empirical literature.

In light of these challenges, the number of empirical studies that quantify the effect of algorithms on prices is relatively small. I argue that research on general price levels is of limited insight without considering the reaction of sophisticated customers. Therefore, the cliche conclusion that "more research is needed" applies here. At the same time, theoretical researchers interested in the use of pricing algorithms should explain which novel questions are posed by algorithms that have not already been covered in the traditional literature on dynamic games.

4.2 An overview of the algorithmic pricing literature

The algorithmic pricing literature is conveniently summarized with a joint methodological and chronological approach. First, there is literature preceding the debate on algorithmic collusion. These researchers and their successors are either interested in the features of dynamic pricing algorithms from a computer science perspective or they are looking for new computational approaches to simulate economics models (Section 4.2.1).

This is followed by an emerging literature on algorithmic collusion (Section 4.2.2). This early literature on collusion is partly speculative, i.e., without substantive analysis of algorithmic behavior and learning, especially from the legal perspective. I also include in this section papers that have critically refined this literature by comparing the market structure that allow tacit collusion to arise with the settings in which we encounter algorithmic pricing.

The third and current wave of (economics) literature on pricing algorithms is usefully distinguished by their methodology, including simulations and computational methods (Section 4.2.4), experimental research (Section 4.2.5), and empirical papers (Section 4.2.6).

4.2.1 Algorithmic pricing before and besides the debate on collusion

Pricing algorithms appear in the economics literature at least a decade before the policy discussion of algorithmic collusion by at least a decade. Already Tesfatsion (2006) describes algorithmic pricing rules for agents in computational representations of economies. This agent-based-modeling (ABM) approach studies emergent behavior of economic agents. It does not compare pricing outcomes of algorithmic and conventional, "human" pricing. While these papers do not focus on competition policy, they incorporate aspects that have received relatively scant attention in the industrial organization literature, such as buyer-seller interaction and bounded rationality.

Boer (2015) surveys the earlier literature on dynamic pricing, including price formation and statistical learning, covering the literature in computer science, management, operations research, and economics. However, the discussion of the economics literature and competition is limited to general questions of price formation and less on how the use of pricing algorithms by economic agents impacts price formation in markets.

Studying electricity markets, Barazza and Strachan (2020) employ pricing algorithms that account for bounded rationality by producers who algorithmically determine production. Marzband et al. (2017) simulate electricity markets with algorithms that govern the generation of electricity where consumers are sometimes also producers, and prices follow from generation decisions through peak-load pricing. Filatova et al. (2009) demonstrate the use of a simple pricing algorithm to study price formation on land markets using ABM. Here, buyers and sellers adjust their prices starting from their reservation prices by a gradient that is determined by the relative size of the buyer and seller population.

Management scholars have provided alternative perspectives on pricing algorithms: Gerlick and Liozu (2019) suggest that ethical considerations including fairness, (avoiding) deception, and social justice, as well as legal considerations regarding data privacy and non-discrimination are more important than antitrust concerns to firms that deploy pricing algorithms.

Research priorities in adjacent literatures highlight the importance of carefully posed research questions. In particular, the study of algorithms per se is not of economic interest but the implications for consumers and firms are. The following sections focus on the industrial organization literature, which has largely ignored interaction between algorithmic buyers and algorithmic sellers as well as bounded rationality. Some of the examples above, even if mechanical or methodologically simplistic, have started to incorporate these aspects. I argue that this is a gap in the industrial organization literature on algorithmic pricing.

4.2.2 Speculative and legal literature

Algorithms have appeared in an antitrust context at least since a 2015 US case against price fixing for posters on Amazon.[77] The use of algorithms as a tool for collusion is not identical with algorithmic collusion, however. In the above case, algorithms did not differ from the use of other technology to better communicate and coordinate.[78] Ezrachi and Stucke (2017) point out that algorithms can facilitate collusive practices that are not fundamentally different from traditional collusion, for example, if the use of an algorithm lowers a perpetrator's sense of wrongdoing, or because it facilitates a hub-and-spoke cartel where the supplier of the algorithm serves as ring-leader.

Early contributions by competition lawyers are mostly speculative, and divided on whether we should expect algorithmic collusion. Some legal scholars assume algorithmic collusion to be easy, sometimes calling for entirely new legislation and enforcement tools.[79] Petit (2017) cautions against presumptions of algorithmic collusion and emphasizes the need for empirical analysis. Pohlmann and Schütte (2018) and Van Uytsel (2018) summarize this literature. The latter also lists suggestions to deal with algorithmic pricing: algorithmic audit, reduced price transparency, and (at the time of the EU's General Data Protection Regulation) enhanced consumer privacy.

Economists tend to be skeptical towards the prospect of algorithmic collusion. Schwalbe (2018) contrasts the limited capability of machine-learning based algorithms at the time with the high hurdles that the economics literature has identified for successful collusion. While algorithmic collusion is possible, algorithms would need to be able to communicate, and even if this resulted in reduced price competition, other competitive parameters might remain. Miklós-Thal and Tucker (2019) explain that the superior ability of algorithms to detect market features might exactly counteract the traditional, so-called "plus-factors" of tacit collusion.

Leisten (2022) investigates the question of human-algorithmic interaction theoretically. He emphasizes the role of algorithms as enhancing both prediction and commitment and that only one of the aspects needs to be retained in order to obtain supracompetitive prices. In a setting where competing pricing algorithms collude, he finds that even when managers can override algorithmic decisions, prices remain supra-competitive. Cartea et al. (2022) represent many learning algorithms as stochastic equations and find that they converge to the Nash equilibrium of the stage game but that competition might be impeded by tick size.

Harrington (2018) asks whether algorithmic collusion made tacit collusion so much easier that a shifting of the legal standard is warranted. Typically, algorithmic collusion scenarios focus on tacit collusion, that is, firms succeeding to sustain high prices without explicit communication. From an economic welfare point of view, the outcome of tacit collusion is not materially different to that of a cartel that maintains high prices by explicit communication. The legal treatment is

[77]FTC (2015): Former E-Commerce Executive Charged with Price Fixing in the Antitrust Division's First Online Marketplace Prosecution, `https://www.justice.gov/opa/pr/former-e-commerce-executive-charged-price-fixing-antitrust-divisions-first-online-marketplace`.

[78]Commenting on its use as an indirect communication tool, the FTC chairman reportedly quipped that "algorithm" in this case may just as well have been "a guy named Bob" (Ohlhausen, 2017).

[79]Mehra (2014) claims that "robo-sellers will increase the risk that oligopolists will coordinate prices above the competitive level" and proposes solutions that the author finds "quite difficult to reconcile with current antitrust law".

different, as tacit collusion is not illegal in most jurisdictions. The author suggests adjusting the legal standard to make collusion between "autonomous artificial agents" illegal.

4.2.3 Theoretical literature

The literature on algorithmic collusion can be seen as a strand of the literature on repeated games.[80] The Folk theorem states that a large number of equilibria can be supported in repeated play with patient players (Fudenberg and Maskin, 1986; Abreu, Pearce, et al., 1990; Abreu, Dutta, et al., 1994; Fudenberg, Levine, et al., 2008). These equilibria include outcomes that could be called "competitive" or "collusive" equilibria in some settings[81] as well as any number of intermediate outcomes (Skrzypacz and Hopenhayn, 2004).

Starting from the study of finite automata (Rubinstein, 1986; Abreu and Rubinstein, 1988), the repeated games literature later focused on Turing machines (Binmore, 1987). These are relevant as every (pricing) algorithm is a finite sequence of instructions to solve a calculation. The parallel between the computerized terminology adopted in the literature 40 years ago and today's economic analysis of computer algorithms is remarkable. Aumann and Sorin (1989) and Anderlini and Sabourian (1995) show that players will converge to their pareto-optimal equilibria in common-interest games. Cooperation emerges even in an evolutionary setting under weak assumptions. It is not even required that the more successful algorithms are adopted more frequently (Anderlini and Sabourian, 1996).

The reinforcement algorithms considered below learn probabilistically, comparable to hidden Markov models. This literature is summarized in Marks (1990) and Almanasra et al. (2013) but more recent papers also have applications to algorithmic pricing. For example, Dal Bó and Fréchette (2019) experiment with repeated PD under perfect monitoring and discuss the recovery of strategy from observed actions, a setting that applies to algorithms monitoring competitors' prices.

Without narrowing down the context that would permit a game-theoretic analysis, for economists who have studied repeated games, it would not be surprising to find that real-life computer algorithms would learn to collude. How can this be reconciled with the skepticism of some economists against the early papers on algorithmic collusion (previous section)? One possible explanation is that the real-life equivalent to the environment studied by IO economists is usually not well defined: It is not clear whether in real marketplaces algorithms compete against agents with fixed strategies, learning agents, the same agents or random agents from a large population. A second criticism, discussed in more detail in the context of the following section, comes from papers such as Eschenbaum et al. (2022) that point out how differences between the training environment of an algorithm and a real marketplace can matter. In game theoretic analysis, these are usually the same thing: the agent learns in the game which provides a sufficiently rich description of the setting in which she operates.

A complete overview of repeated games is beyond the scope of this article. However, when restrictive features of algorithms are assumed (such as a memory of specific length), models of algorithmic collusion are essentially looking for the equilibria that emerge under specific learning protocols and restrictions. Researchers investigating algorithmic collusion should therefore explain whether the restrictions that allow equilibrium selection should be taken as exogenous and fixed (for example due to technical reasons) and if not, whether any equilibria that emerge are robust to perturbations (for example, the introduction of new algorithms with a larger state-space, faster reaction etc.)

[80] Calzolari and Hanspach (2022) document algorithmic repricers even for pricing one-off contracts and procurement tenders that do not seem to consider repeated interactions. The majority of the literature focuses on sellers that meet repeatedly.

[81] In the stage game of a Bertrand pricing game with two firms, the competitive outcome would be equivalent to the Nash equilibrium with low prices while the collusive outcome would correspond to high prices, possibly monopoly prices, which leave both firms better off but do not form an equilibrium.

4.2.4 Simulations and computational approaches

Companies typically do not reveal their pricing strategies for competitive reasons.[82] For the most part, researchers have therefore either studied simple algorithms in simulated environments, or made inference from observable characteristics of pricing data (such as the frequency of price changes) to make an educated guess about the presence of pricing algorithms in observational data. This section looks at cases of the former, papers that study the behavior of certain algorithms in stylized oligopoly models.

Several proof-of-concept algorithms arrive at collusive outcomes in these environments. Calvano, Calzolari, Denicolò, et al. (2020) study Q-learning algorithms (a class of reinforcement-learning algorithm[83]) in a repeated Bertrand-setting with simultaneous updating while Klein (2021) studies a similar problem with sequential updating. Both papers find that the algorithmic agents set prices above the "competitive price" associated with the Nash equilibrium of the stage game.[84] Importantly, these high prices are supported by a collusive strategy that punishes deviations from the collusive price by a price cut, followed by a gradual readjustment to higher prices.

Hansen et al. (2021) show in a multi-armed bandit framework that algorithms learn to set supra-competitive prices when rivals' prices are not observed. These algorithms run concurrent experiments leading to biased estimates of own-price elasticities and high prices. Sánchez-Cartas and Katsamakas (2022) simulate algorithmic price setting for platform companies and compare outcomes for Q-learning, price matching, and evolutionary algorithms. They find different behavior of different algorithms, advising against blanket policy descriptions. Banchio and Skrzypacz (2022) find that simulated Q-learning agents can achieve collusive prices in a bidding context resembling online advertisement.

Similarly, Kastius and Schlosser (2022) study the behavior of two RL algorithms (Deep Q-Networks and Soft Actor Critic) in several stylized models of imperfect competition but only describe high-price outcomes for certain pre-specified rival strategies without searching for equilibrium strategies with multi-agent learning. Kastius and Schlosser (2022) also list several papers that study Q-learning algorithms in different market environments. Kutschinski et al. (2003) compare different RL algorithms, noting the high cost of training Q-learning algorithms. When the state-space is high-dimensional, it increases the number of cases that need to be trained for exponentially.

Abada and Lambin (2023) train Q-learning algorithms in the particular context of electricity markets, or more precisely battery charging in a setup modeled on Tesla's "autobidder" platform. They show numerically that supracompetitive prices arise due to a failure to fully learn competitive strategies, rather than collusion. They discuss policy remedies and challenges for regulators specific to their application. While enforcing local learning (at the level of the user, rather than at the level of the aggregator) works well in a simulation, it is doubtful whether it is cost-efficient in a real setting. Intervention by a regulated market participant does not improve welfare in the simulation.

Do algorithms *learn to collude* or do they *fail to learn to compete*? Calvano, Calzolari, Denicolò, et al. (2020) emphasize that economists speak of collusion if high prices are underpinned by retaliation following a deviation from collusion. Just observing high prices can indicate algorithms not learning to compete, or optimizing for the wrong problem. Xie and J. Chen (2004)

[82] "The precise choice of algorithms is one of the most closely held secrets in Silicon Valley and elsewhere." (Abada and Lambin, 2023)

[83] Reinforcement learning (RL) is a machine-learning technique in which an agent seeks to maximize some objective function. The agent selects actions based on observable states and receives rewards and punishments in accordance with his actions in a given state. The agent updates his preferred actions through trial and error. Q-learning is a popular, model-free implementation of RL that starts with initially random beliefs over the reward-punishment matrix which it slowly updates through (gradually decreasing) experimentation.

[84] Calvano, Calzolari, Denicoló, et al. (2021) show that Q-learning algorithms can also learn collusive strategies in the Green and Porter (1984)-environment of imperfect monitoring.

propose a "Nash test" in which, first, agents learn a set of converging strategies. Then, holding all but one agents' learned strategies fixed, a perturbation is introduced to see if the remaining agents' strategies converges to their initial strategies, generating a simulation equivalent to a best response.

Eschenbaum et al. (2022) question whether collusive strategies generalize from training environments to real markets. Differences in context make it difficult to train algorithms for robust collusion unless firms coordinate on *algorithm design*.[85] However, as Calzolari and Hanspach (2022) document, some firms offering pricing algorithms claim to train in real-life environments. It is unclear whether there is any comfort in the idea that too simplistic training environments might hamper algorithmic collusion.

Overall, the simulation literature confirms that pricing algorithms can learn collusive strategies that result in supracompetitive prices in controlled setups. Simple and robust strategies of reward-and-punishment that underlie these strategies can occur as emergent properties in multi-agent learning. These strategies resemble equilibrium strategies that are familiar to economists in the context of tacit collusion. Further testing is needed, however, how these algorithms perform in real-life environments. It is not a-priori clear when the added complexity of a real marketplace effectively limits algorithmic collusion.

4.2.5 The experimental literature

Lab experiments are an established tool to test theories and to generate data in a controlled environment. Human test subjects have been found to quickly converge on the theoretically predicted Nash equilibria in a plethora of circumstances, including collusion (Levine, 2012). Dal Bó and Fréchette (2011) point out the sharp conditions under which cooperation arises in infinitely repeated games. Given the existing literature on collusion in a lab setting, it seems useful to compare the existing results with experiments that add pricing algorithms, and knowledge about their presence, as a treatment.

Human-machine interaction has been recognized as its very own source of friction in a market environment (Crandall et al., 2018; Fu et al., 2022). In this field, phenomena such as algorithm aversion appear, some of which have also been of interest to the psychology literature (Dietvorst et al., 2015).

Werner (2021) uses human participants as a benchmark against which to test pricing algorithms and finds positive effects of algorithms on prices, in particular when adoption is widespread. Normann and Sternberg (2023) let algorithms and human experiment participants compete and find that "firms employing an algorithm earn significantly less profit than their rivals. (Un)certainty about the actual presence of an algorithm does not significantly affect collusion, although humans do seem to perceive algorithms as more disruptive."

This is consistent with some results of the empirical literature described in the following section that emphasizes the importance of several, if not all, firms in a market to adopt an algorithm to obtain higher prices.

4.2.6 The empirical literature

A major reason to study pricing algorithms is uncertainty about their effect on prices. A limited number of empirical studies have been conducted to analyze the effect of pricing algorithms on retail prices in different industries and locations. Table 12 summarizes some current papers that include substantial empirical analysis of pricing algorithms, reporting the data source, location, time period, and industry studied for the purpose of pointing to available data.

Several authors study pricing patterns in scraped online data. Early papers focus on describing stylized patterns of algorithmic pricing and detection of algorithmic sellers. L. Chen et al.

[85] Algorithm design refers to the way the pricing algorithms are trained, in particular restricting the information that their algorithms can operate on to avoid overfitting to rivals' strategies.

(2016) study pricing data from Amazon. They find that algorithmic sellers typically win the Buy Box, closely matching the lowest price. Brown and MacKay (2021) collect data on pricing by online pharmacies and retailers selling over-the-counter medicine. They propose that algorithmic sellers change prices faster than others, and generally charge lower prices, reacting faster to competitor price changes. The authors use their findings to motivate a theoretical model of asymmetric pricing speeds.

The identification of algorithmic sellers is a problem that is common to all empirical papers in this field. To the best of my knowledge, there exists no data set of prices set by competing sellers that includes observed data on the use of pricing algorithms. All papers mentioned in this section infer algorithm usage from some observable characteristic of seller behavior. Brown and MacKay (2021) do not make inference about which sellers might use algorithms but simply describe general trends in the observed pricing of online retailers, which they then argue are consistent with the use of pricing algorithms.

Most of these approaches follow L. Chen et al. (2016) who define two main criteria to infer that a seller uses pricing algorithms: first, a close correlation of a seller's price series against a meaningful benchmark, such as the lowest- or second-lowest price; second, a high frequency of price changes. The cutoff values for both criteria are chosen empirically, that is, eyeballed from the "kink" or "knee" in the empirical distribution. Assad et al. (2020) study the adoption of algorithms by retail gasoline stations similarly, focusing on structural breaks in observable features of the price series. These features are (i) number of price changes, (ii) average size of price changes, and (iii) rival response time.

Some papers identify a causal effect of the adoption of pricing algorithms on market prices. Assad et al. (2020) use public data on German gas station prices. Gas stations that adopt pricing algorithms increase their margins in non-monopoly markets. Adoption has no impact on margins in monopoly markets. Margins in duopoly market increase only when both firms adopt algorithms. Musolff (2022) studies pricing cycles on the Amazon marketplace using proprietary data. He finds that the adoption of pricing algorithms led to an immediate decline in prices. However, the resulting cycling strategies induce competitors to raise prices such that long-term price increases are possible.

Holt et al. (2022) compare detection algorithms for pricing cycles on gas station prices for Australia and Germany, overlapping with the German data used by Assad et al. (2020) They test for the presence of Edgeworth-cycles. Edgeworth-cycles are one model that generates the pricing cycles observed in the study of algorithmic pricing, although Hortacsu et al. (2021) and Musolff (2022) find that pricing cycles in their data are better explained by delegated strategies.

Hortacsu et al. (2021) study dynamic pricing competition in a setting characterized by capacity constraints and sales deadlines. Using a theoretical model and proprietary 2019 data from a US airline, they find that airlines' use of ticket "fare buckets" as a pricing heuristic can improve welfare over a benchmark with repricing in every period. The authors relate this finding to the discussion around strategic selection of pricing algorithms by firms.

The data used in these papers is summarized in Table 12. I summarize the most important questions tackled in this literature as follows:

1. How many sellers use pricing algorithms?
2. What is the effect of pricing algorithms on pricing patterns?
3. What is the effect of pricing algorithms on average prices?
4. How does competition affect the impact of pricing algorithms?

Question 4 differs from from how the question is often phrased. More commonly, researchers ask "what is the effect of algorithms/AI/machine learning on competition?" However, one finding that is common to several papers that we have already discussed above is that the adoption of

Authors	Industry	Time period	Data source	Key results
L. Chen et al. (2016)	Online retail	2014 - 2015	Scraped public data	Buy box non-price determinants, identification of dynamic pricing
Assad et al. (2020)	Gasoline	2016 - 2019	Public data set	Increased prices in duopolies wide algorithm adoption matters, market structure matters
Brown and MacKay (2021)	Medicine	2018 - 2019	Scraped public data	Data-driven model explains high and dispersed prices, greater effect of mergers on prices
Hortacsu et al. (2021)	Airlines	2019	Proprietary data	Organizations constrain pricing
Aparicio et al. (2021)	Online grocery retail	2006 - 2017	Scraped public data	Greater online price dispersion for multi-channel retailers
Musolff (2022)	Online retail	2018 - 2020	Proprietary data	Algorithms facilitate tacit collusion
Holt et al. (2022)	Gasoline	2001 - 2020	Public data	Screening methods for price cycles impact proper screens for collusion

Table 12: Overview of data sets used by empirical papers studying pricing algorithms

pricing algorithms yields differential impacts depending on the market structure. By contrast, little research has been done to identify an effect of the adoption of pricing algorithms on market structure (in particular, entry and exit).

Varian (2019) shows that much that is discussed regarding algorithmic collusion is very similar indeed to the older literature on equilibria in repeated games. The question of algorithmic collusion can be boiled down to whether pricing algorithms converge to collusive equilibria in repeated games. All markets that have been studied so far have existed before the introduction of pricing algorithms and the empirical approach has therefore been a before-after comparison.[86] The question at stake is really whether adding pricing algorithms to a market environment allows coordination on different (in particular higher-price) equilibria.

Pricing algorithms may impact average prices through channels other than collusion. By drawing on more and new data, algorithms may alleviate human biases and systematic errors in pricing (or introduce new systematic errors). If we assume that algorithms are better than humans in choosing profit-maximizing prices, the impact of algorithms depends on the previous bias of the human decision maker. Algorithms then merely correct prices to the profit-maximizing level. If humans initially set prices excessively high, pricing algorithms can lead to lower prices.[87] By the same mechanism algorithms could also result in price increases if the human-set prices are downward-biased.

Most regression analyses default to analyzing average prices. If prices are, for example, higher on average following the adoption of pricing algorithms, but also more volatile, sophisticated consumers will purchase when prices are low (Acquisti and Varian, 2005). Higher average prices are not a proof of consumer harm if algorithmic pricing allows (some) consumers to purchase at lower prices than before.[88] Therefore, documenting pricing patterns together with average

[86] To my knowledge, there are no studies of markets that emerged together with pricing algorithms or markets from which pricing algorithms were removed.

[87] One possible human bias could be managerial overconfidence, which is documented in the managerial literature regarding demand (Montgomery and Bradlow, 1999), new product introduction (Simon and Shrader, 2012, Markovitch et al., 2015, Feiler, 2021) and other corporate decision variables (Malmendier and Tate, 2015). Overconfidence, for example in product demand or quality can result in excessive prices. Kahneman (2011) recognizes overconfidence as "*the most significant of the cognitive biases.*"

[88] Consumers that search for flights on price comparison engines sometimes face price ranges that inform them of average prices and the range of prices on their desired route, for example when using Google Flights. Websites such as Keepa or CamelCamelCamel track prices on Amazon and can send automated alerts to consumers. Importantly, both of these services offer simple and intuitive alerts that do not require any technical expertise beyond using a web browser.

prices is important to quantify potential consumer harm. When pricing algorithms lead to higher prices, consumer harm is more likely if prices are stable than if prices remain variable (Huber and Imhof, 2019).

Finally, there is scant research on the endogenous features of sellers that lead to the adoption of algorithms and their impact on market structure. In Assad et al. (2020), brand size is the only determinant of the brand decision to adopt pricing algorithms. In a setting where individual gas stations belong to a larger chain making strategic decisions, a brand-level decision to adopt pricing algorithms (or to encourage franchisees to do so), can increase adoption of algorithms by individual retail outlets. Brand adoption therefore serves as an instrument to deal with endogeneity in station-level adoption of pricing algorithms.

4.3 Conclusion

In surveying the literature on algorithmic pricing, we have gathered some facts on algorithmic pricing:

Algorithmic pricing does not always lead to coordinated higher prices. In particular, early speculation that warned of ubiquitous algorithmic collusion has not been confirmed. The empirical literature has accepted, however, that the adoption of pricing algorithms leads to changes in pricing behavior, in particular frequent price changes and, in some markets, recurring price cycles.

In controlled environments, simple algorithms are capable of learning sophisticated reward-and-punishment strategies. In theory, these strategies are capable of supporting collusive behavior. Empirically, higher prices seem to be a more likely outcome in more concentrated markets and with universal adoption of pricing algorithms.

However, higher average prices do not necessarily imply harm to consumers if prices also cycle. Sophisticated consumers may still be able to purchase at equal or even lower prices compared to a situation without algorithmic pricing. If pricing algorithms indeed learn market features with a high degree of precision, the resulting pricing patterns may also be explained as a form of (potentially efficiency-enhancing) price discrimination, rather than algorithmic collusion.

Therefore, in spite of the great attention afforded to the risk of algorithmic collusion, there are open questions before we can accurately describe the effect of pricing algorithms in markets. The ability of algorithms to draw on superior data and to outperform human decision makers remains speculative. Also, the effect of algorithmic adoption on market structure remains unknown. Finally, we can say little about the effect of algorithmic pricing on consumers without research on strategic reactions of buyers to this new phenomenon.

Acknowledgments

Funding by the German Academic Exchange Service is gratefully acknowledged. Giacomo Calzolari and David Levine provided helpful comments. All errors and views are my own.

References

Abada, Ibrahim and Xavier Lambin (2023). “Artificial Intelligence: Can Seemingly Collusive Outcomes Be Avoided?” en. In: *Management Science*, mnsc.2022.4623. DOI: 10.1287/mnsc.2022.4623.

Abrardi, Laura, Carlo Cambini, and Laura Rondi (2022). “Artificial intelligence, firms and consumer behavior: A survey”. en. In: *Journal of Economic Surveys* 36.4, pp. 969–991. DOI: 10.1111/joes.12455.

Abreu, Dilip, Prajit K. Dutta, and Lones Smith (1994). “The Folk Theorem for Repeated Games: A Neu Condition”. In: *Econometrica* 62.4, p. 939. DOI: 10.2307/2951739.

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990). “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring”. In: *Econometrica* 58.5, p. 1041. DOI: 10.2307/2938299.

Abreu, Dilip and Ariel Rubinstein (1988). “The Structure of Nash Equilibrium in Repeated Games with Finite Automata”. In: *Econometrica* 56.6, p. 1259. DOI: 10.2307/1913097.

Acquisti, Alessandro and Hal Varian (2005). “Conditioning Prices on Purchase History”. In: *Marketing Science* 24.3, pp. 367–381.

Almanasra, Sally, Khaled Suwais, and Muhammad Rafie (2013). “The Applications of Automata in Game Theory”. In: *Intelligent Technologies and Techniques for Pervasive Computing.* 10.4018/978-1-4666-4038-2.ch011. IGI Global.

Anderlini, Luca and Hamid Sabourian (1995). “Cooperation and Effective Computability”. In: *Econometrica* 63.6, p. 1337. DOI: 10.2307/2171773.

— (1996). “The evolution of algorithmic learning rules: a global stability result.”

Aparicio, Diego, Zachary Metzman, and Roberto Rigobon (2021). *The Pricing Strategies of Online Grocery Retailers.* en. Tech. rep. w28639. Cambridge, MA: National Bureau of Economic Research, w28639. DOI: 10.3386/w28639.

Assad, Stephanie et al. (2020). “Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market”. In: *CESifo Working Paper No. 8521.*

Athey, Susan and Michael Luca (2019). “Economists (and Economics) in Tech Companies”. en. In: *Journal of Economic Perspectives* 33.1, pp. 209–230. DOI: 10.1257/jep.33.1.209.

Aumann, Robert J and Sylvain Sorin (1989). “Cooperation and bounded recall”. en. In: *Games and Economic Behavior* 1.1, pp. 5–39. DOI: 10.1016/0899-8256(89)90003-1.

Banchio, Martino and Andrzej Skrzypacz (2022). “Artificial Intelligence and Auction Design”. en. In: *Proceedings of the 23rd ACM Conference on Economics and Computation.* Boulder CO USA: ACM, pp. 30–31. DOI: 10.1145/3490486.3538244.

Barazza, Elsa and Neil Strachan (2020). “The co-evolution of climate policy and investments in electricity markets: Simulating agent dynamics in UK, German and Italian electricity sectors”. en. In: *Energy Research & Social Science* 65, p. 101458. DOI: 10.1016/j.erss.2020.101458.

Binmore, Ken (1987). “Modeling Rational Players: *Part I*”. en. In: *Economics and Philosophy* 3.2, pp. 179–214. DOI: 10.1017/S0266267100002893.

Boer, Arnoud V. den (2015). “Dynamic pricing and learning: Historical origins, current research, and new directions”. en. In: *Surveys in Operations Research and Management Science* 20.1, pp. 1–18. DOI: 10.1016/j.sorms.2015.03.001.

Brown, Zach Y. and Alexander MacKay (2021). *Competition in Pricing Algorithms.* Working Paper 28860. Series: Working Paper Series. National Bureau of Economic Research. DOI: 10.3386/w28860.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicoló, et al. (2021). “Algorithmic collusion with imperfect monitoring”. en. In: *International Journal of Industrial Organization* 79, p. 102712. DOI: 10.1016/j.ijindorg.2021.102712.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, et al. (2020). “Artificial Intelligence, Algorithmic Pricing, and Collusion”. en. In: *American Economic Review* 110.10, pp. 3267–3297. DOI: 10.1257/aer.20190623.

Calzolari, Giacomo and Philip Hanspach (2022). *Pricing Algorithms Out of the Box: A Study of the Algorithmic Repricing Industry.*

Cartea, Álvaro, Patrick Chang, and José Penalva (2022). “Algorithmic Collusion in Electronic Markets: The Impact of Tick Size”. en. In: *SSRN Electronic Journal.* DOI: 10.2139/ssrn.4105954.

Chen, Le, Alan Mislove, and Christo Wilson (2016). “An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace”. en. In: *Proceedings of the 25th International Conference on World Wide Web.* Montréal Québec Canada: International World Wide Web Conferences Steering Committee, pp. 1339–1349. DOI: 10.1145/2872427.2883089.

Crandall, Jacob W. et al. (2018). "Cooperating with machines". en. In: *Nature Communications* 9.1, p. 233. DOI: 10.1038/s41467-017-02597-8.

Dal Bó, Pedro and Guillaume R. Fréchette (2011). "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence". en. In: *American Economic Review* 101.1, pp. 411–429. DOI: 10.1257/aer.101.1.411.

— (2019). "Strategy Choice in the Infinitely Repeated Prisoner's Dilemma". en. In: *American Economic Review* 109.11, pp. 3929–3952. DOI: 10.1257/aer.20181480.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey (2015). "Algorithm aversion: People erroneously avoid algorithms after seeing them err." en. In: *Journal of Experimental Psychology: General* 144.1, pp. 114–126. DOI: 10.1037/xge0000033.

Eschenbaum, Nicolas, Filip Mellgren, and Philipp Zahn (2022). *Robust Algorithmic Collusion.* arXiv:2201.00345 [econ, q-fin]. DOI: 10.48550/arXiv.2201.00345.

Ezrachi, Ariel and Maurice E. Stucke (2017). *Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy.* en. Harvard University Press.

Feiler (2021). *From Noise to Bias: Overconfidence in New Product Forecasting | Management Science.*

Filatova, Tatiana, Dawn Parker, and Anne van der Veen (2009). "Agent-Based Urban Land Markets". en. In: *Journal of Artificial Societies and Social Simulation* 12.13, p. 31.

Fu, Runshan, Ginger Zhe Jin, and Meng Liu (2022). *Human-Algorithm Interactions: Evidence from Zillow.com.* en. Tech. rep. w29880. Cambridge, MA: National Bureau of Economic Research, w29880. DOI: 10.3386/w29880.

Fudenberg, Drew, David Levine, and Eric Maskin (2008). "The Folk Theorem with Imperfect Public Information". en. In: *A Long-Run Collaboration on Long-Run Games.* WORLD SCIENTIFIC, pp. 231–273. DOI: 10.1142/9789812818478_0012.

Fudenberg, Drew and Eric Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information". In: *Econometrica* 54.3, p. 533. DOI: 10.2307/1911307.

Gerlick, Joshua and Stephan M. Liozu (2019). "A Conceptual Framework of Ethical Considerations and Legal Constraints in the Algorithm-Driven Pricing Function". en. In: *SSRN Electronic Journal.* DOI: 10.2139/ssrn.3454123.

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

Hansen, Karsten T., Kanishka Misra, and Mallesh M. Pai (2021). "Frontiers: Algorithmic Collusion: Supra-competitive Prices via Independent Algorithms". en. In: *Marketing Science* 40.1, pp. 1–12. DOI: 10.1287/mksc.2020.1276.

Harrington, Joseph E (2018). "Developing Competition Law For Collusion by Autonomous Artificial Agents". en. In: *Journal of Competition Law & Economics* 14.3, pp. 331–363. DOI: 10.1093/joclec/nhy016.

Holt, Timothy, Mitsuru Igami, and Simon Scheidegger (2022). *Detecting Edgeworth Cycles.* Number: arXiv:2111.03434 arXiv:2111.03434 [physics].

Hortacsu, Ali et al. (2021). "Organizational Structure and Pricing: Evidence from a Large U.S. Airline". In.

Huber, Martin and David Imhof (2019). "Machine learning with screens for detecting bid-rigging cartels". en. In: *International Journal of Industrial Organization* 65, pp. 277–301. DOI: 10.1016/j.ijindorg.2019.04.002.

Kahneman, Daniel (2011). *Thinking, fast and slow.* Macmillan.

Kastius, Alexander and Rainer Schlosser (2022). "Dynamic pricing under competition using reinforcement learning". en. In: *Journal of Revenue and Pricing Management* 21.1, pp. 50–63. DOI: 10.1057/s41272-021-00285-3.

Klein, Timo (2021). "Autonomous algorithmic collusion: Q-learning under sequential pricing". en. In: *The RAND Journal of Economics* 52.3, pp. 538–558. DOI: 10.1111/1756-2171.12383.

Kutschinski, Erich, Thomas Uthmann, and Daniel Polani (2003). "Learning competitive pricing strategies by multi-agent reinforcement learning". en. In: *Journal of Economic Dynamics and Control* 27.11-12, pp. 2207–2218. DOI: 10.1016/S0165-1889(02)00122-7.

Leisten, Matthew (2022). "Algorithmic Competition, with Humans". en. In: *working paper*.

Levine, David (2012). *Is behavioral economics doomed? the ordinary versus the extraordinary.* OCLC: 849917882. Cambridge: Open Book Publishers. DOI: 10.11647/OBP.0021.

Malmendier, Ulrike and Geoffrey Tate (2015). "Behavioral CEOs: The Role of Managerial Overconfidence". en. In: *Journal of Economic Perspectives* 29.4, pp. 37–60. DOI: 10.1257/jep.29.4.37.

Markovitch, Dmitri G. et al. (2015). "Behavioral Reasons for New Product Failure: Does Overconfidence Induce Overforecasts?: Behavioral Reasons for New Product Failure". en. In: *Journal of Product Innovation Management* 32.5, pp. 825–841. DOI: 10.1111/jpim.12252.

Marks, Robert E. (1990). "Repeated games and finite automata." In: *Australian Graduate School of Management, University of New South Wales*.

Marzband, Mousa et al. (2017). "Distributed generation for economic benefit maximization through coalition formation-based game theory concept: Distributed Generation economic benefit maximization". en. In: *International Transactions on Electrical Energy Systems* 27.6, e2313. DOI: 10.1002/etep.2313.

Mehra, Salil K. (2014). "De-Humanizing Antitrust: The Rise of the Machines and the Regulation of Competition". en. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2490651.

Miklós-Thal, Jeanine and Catherine Tucker (2019). "Collusion by Algorithm: Does Better Demand Prediction Facilitate Coordination Between Sellers?" en. In: *Management Science* 65.4, pp. 1552–1561. DOI: 10.1287/mnsc.2019.3287.

Montgomery, Alan L. and Eric T. Bradlow (1999). "Why Analyst Overconfidence About the Functional Form of Demand Models Can Lead to Overpricing". In: *Marketing Science* 18.4, pp. 569–583. DOI: 10.1287/mksc.18.4.569.

Musolff (2022). "Algorithmic Pricing Facilitates Tacit Collusion". In: pp. 32–33. DOI: https://doi.org/10.1145/3490486.3538239.

Normann, Hans-Theo and Martin Sternberg (2023). "Human-algorithm interaction: Algorithmic pricing in hybrid laboratory markets". en. In: *European Economic Review* 152, p. 104347. DOI: 10.1016/j.euroecorev.2022.104347.

Ohlhausen, M. K. (2017). "Should we fear the things that go beep in the night? some initial thoughts on the intersection of antitrust law and algorithmic pricing". In: *Discurso pronunciado en la Conferencia «Antitrust in the Financial Sector» organizada por Concurrences en Nueva York. Recuperado el.* Vol. 9.

Petit, Nicolas (2017). "Antitrust and Artificial Intelligence: A Research Agenda". en. In: *Journal of European Competition Law & Practice* 8.6, pp. 361–362. DOI: 10.1093/jeclap/lpx033.

Pohlmann, Petra and David Schütte (2018). "Algorithms As Cartel Infringements A Bibliography". en. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3200004.

Rubinstein, Ariel (1986). "Finite automata play the repeated prisoner's dilemma". en. In: *Journal of Economic Theory* 39.1, pp. 83–96. DOI: 10.1016/0022-0531(86)90021-9.

Sánchez-Cartas, J. Manuel and Evangelos Katsamakas (2022). *Effects of Algorithmic Pricing on Platform Competition*. en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.4027365.

Schwalbe, Ulrich (2018). "Algorithms, machine learning, and collusion". In: *Journal of Competition Law & Economics* 14.4, pp. 568–607. DOI: 10.1093/joclec/nhz004.

Simon, Mark and Rodney C. Shrader (2012). "Entrepreneurial actions and optimistic overconfidence: The role of motivated reasoning in new product introductions". en. In: *Journal of Business Venturing* 27.3, pp. 291–309. DOI: 10.1016/j.jbusvent.2011.04.003.

Skrzypacz, Andrzej and Hugo Hopenhayn (2004). "Tacit collusion in repeated auctions". en. In: *Journal of Economic Theory* 114.1, pp. 153–169. DOI: 10.1016/S0022-0531(03)00128-5.

Tesfatsion, Leigh (2006). "Chapter 16 Agent-Based Computational Economics: A Constructive Approach to Economic Theory". en. In: *Handbook of Computational Economics*. Vol. 2. Elsevier, pp. 831–880. DOI: 10.1016/S1574-0021(05)02016-2.

Van Uytsel, Steven (2018). "Artificial Intelligence and Collusion: A Literature Overview". In: *Robotics, AI and the Future of Law*. Ed. by Marcelo Corrales, Mark Fenwick, and Nikolaus Forgó. Series Title: Perspectives in Law, Business and Innovation. Singapore: Springer Singapore, pp. 155–182. DOI: 10.1007/978-981-13-2874-9_7.

Varian, Hal (2019). "16. Artificial Intelligence, Economics, and Industrial Organization". en. In: *16. Artificial Intelligence, Economics, and Industrial Organization*. University of Chicago Press, pp. 399–422. DOI: 10.7208/9780226613475-018.

Werner, Tobias (2021). "Algorithmic and human collusion". In: *Available at SSRN 3960738*.

Xie, Ming and Jian Chen (2004). "Studies on horizontal competition among homogenous retailers through agent-based simulation". en. In: *Journal of Systems Science and Systems Engineering* 13.4, pp. 490–505. DOI: 10.1007/s11518-006-0178-7.

A Appendix to Chapter 1

A.1 Proofs

A.1.1 Proof of Theorem 2

For given β_V, β_I, we can write social welfare when $k_V \geq k_I$ as

$$\begin{aligned} S|_{k_V \geq k_I} = (a+r) & \left[\left(\frac{a + \frac{3}{2}r - \beta_I}{\alpha} \right)^{\frac{1}{\alpha-1}} + \left(\frac{a + r - \beta_V}{\alpha} \right)^{\frac{1}{\alpha-1}} \right] \\ & - \left[\left(\frac{a + \frac{3}{2}r - \beta_I}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_I \left(\left(\frac{a + \frac{3}{2}r - \beta_I}{\alpha} \right)^{\frac{1}{\alpha-1}} - k_0 \right) \right] \\ & - \left[\left(\frac{a + r - \beta_V}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_V \left(\left(\frac{a + r - \beta_V}{\alpha} \right)^{\frac{1}{\alpha-1}} - k_0 \right) \right] \end{aligned} \tag{44}$$

while if the inequality is reversed, we have

$$\begin{aligned} S|_{k_I > k_V} = (a+r) & \left[\left(\frac{r - \beta_I}{\alpha} \right)^{\frac{1}{\alpha-1}} + \left(\frac{r/2 - \beta_V}{\alpha} \right)^{\frac{1}{\alpha-1}} \right] \\ & - \left[\left(\frac{r - \beta_I}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_I \left(\left(\frac{r - \beta_I}{\alpha} \right)^{\frac{1}{\alpha-1}} - k_0 \right) \right] \\ & - \left[\left(\frac{r/2 - \beta_V}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_V \left(\left(\frac{r/2 - \beta_V}{\alpha} \right)^{\frac{1}{\alpha-1}} - k_0 \right) \right] \end{aligned} \tag{45}$$

Now, for $S|_{k_V \geq k_I} - S|_{k_I > k_V} > 0$, it must be that

$$\begin{aligned} (a+r) & \left(\left[\frac{2a + \frac{5}{2}r - \beta_I - \beta_V}{\alpha} \right]^{\frac{1}{\alpha-1}} - \left[\frac{\frac{3}{2}r - \beta_I - \beta_V}{\alpha} \right]^{\frac{1}{\alpha-1}} \right) \\ & - \left[\left(\frac{a + \frac{3}{2}r - \beta_I}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_I \left(\frac{a + \frac{3}{2}r - \beta_I}{\alpha} \right)^{\frac{1}{\alpha-1}} \right] \\ & - \left[\left(\frac{a + r - \beta_V}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_V \left(\frac{a + r - \beta_V}{\alpha} \right)^{\frac{1}{\alpha-1}} \right] \\ + \left(\frac{r - \beta_I}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_I \left(\frac{r - \beta_I}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} & + \left(\frac{\frac{r}{2} - \beta_V}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \beta_V \left(\frac{\frac{r}{2} - \beta_V}{\alpha} \right)^{\frac{1}{\alpha-1}} > 0 \end{aligned} \tag{46}$$

$$\leftrightarrow (a+r) \left[\frac{2a + r}{\alpha} \right]^{\frac{1}{\alpha-1}} - \left(\frac{2a + r}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} - \beta_I \left(\frac{a + \frac{r}{2}}{\alpha} \right)^{\frac{1}{\alpha-1}} - \beta_V \left(\frac{a + \frac{r}{2}}{\alpha} \right)^{\frac{1}{\alpha-1}} > 0 \tag{47}$$

$$\leftrightarrow (a+r)^{\alpha-1} \left[\frac{2a + r}{\alpha} \right] > \left[\frac{2a + r}{\alpha} \right]^{\alpha} + (\beta_I + \beta_V)^{\alpha-1} \left(\frac{a + \frac{r}{2}}{\alpha} \right) \tag{48}$$

which is the expression in the proposition. □

A.1.2 Proof of Proposition 3

Substitute the expressions from equations 16, 20 into

$$
\begin{aligned}
MS_Q^{**} - MS_Q^{*} &= \frac{k_I^{**}/2}{k_I^{**} + (1+a/r)k_V^{**}} - \frac{k_I^{*}/2}{k_I^{*} + (1+a/r)k_V^{*}} \qquad (49)\\
&= \frac{\frac{1}{2}\left(\frac{a+\frac{3r}{2}-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}}}{\left(\frac{a+\frac{3r}{2}-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} + \left(1+\frac{a}{r}\right)\left(\frac{a+r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}}} - \frac{\frac{1}{2}\left(\frac{r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}}}{\left(\frac{r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} + \left(1+\frac{a}{r}\right)\left(\frac{\frac{r}{2}-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}}}\\
&= \frac{1}{2}\left(\frac{a+\frac{3}{2}r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}}\left[\left(\frac{r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} + \left(1+\frac{a}{r}\right)\left(\frac{\frac{r}{2}-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}}\right]\\
&- \frac{1}{2}\left(\frac{r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}}\left[\left(\frac{a+\frac{3r}{2}-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} + \left(1+\frac{a}{r}\right)\left(\frac{a+r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}}\right]\\
&= \left(a+\frac{3}{2}r-\beta_I\right)\left(r-\beta_I+\left(1+\frac{a}{r}\right)^{\alpha-1}\left(\frac{r}{2}-\beta_V\right)\right)\\
&- (r-\beta_I)\left[a+\frac{3}{2}r-\beta_I+\left(1+\frac{a}{r}\right)^{\alpha-1}(a+r-\beta_I)\right]
\end{aligned}
$$

which, after multiplying out the brackets and some simplifications, becomes

$$
\begin{aligned}
&\left(a+\frac{3r}{a}-\beta_I\right)\left(\frac{r}{2}-\beta_V\right) - (r-\beta_I)(a+r-\beta_I) \qquad (50)\\
&= -\frac{ra}{2} - \frac{r^2}{4} + \beta_I\frac{r}{2} + (\beta_I-\beta_V)a + \beta_I - \beta_V\frac{3}{2}r + \beta_I\beta_V - \beta_I^2.
\end{aligned}
$$

As we consider the case where both equilibria can exist, we substitute $\beta_V = \beta_I - \frac{r}{2}$.

$$
\begin{aligned}
-\frac{ra}{2} - \frac{r^2}{4} + \beta_I\frac{r}{2} + (\beta_I-\beta_I+\frac{r}{2})a + \beta_I r - \frac{3}{2}\beta_I r - \frac{3}{4}r + \beta_I^2 - \beta_I\frac{r}{2} - \beta_I^2 \qquad (51)\\
= -\frac{r^2}{4} - \frac{3}{4}r - \beta_I\frac{r}{2}
\end{aligned}
$$

which is clearly negative, given that $r > 0$, $\beta_I > 0$, proving the proposition.

A.1.3 Proof of Proposition 4

First, note that in the case without commitment, at the third stage, M will always compete downstream, so $q_M = k_M$ because doing so strictly increases M's profit, at least by $k_M(r/2)$ if $k_M \leq k_V$ and $k_M(r/2) + max(0, (k_M - k_V)(a + r/2))$ otherwise, while withholding its capacity at the third stage yields no profit. Anticipating this, the highest transfer t_V that V is willing to pay is the difference between the profit with or without k_M given that the competitive segment downstream will be k_M if V agrees to rent access to M's infrastructure or k_V if it does not and $k_V \leq k_M$. But then the analysis is equivalent to the base model with k_I instead of k_M.

In the case with commitment, if $k_M > k_V$, offering $q_M = k_M$ limits M to charge $t_V \leq (k_M - k_V)\frac{r}{2} + k_V(a+r)$, or the additional profit V would make competing when M's infrastructure is larger. If M commits to $q_M = 0$, the highest transfer t_V that V would accept is $k_M(a + r)$. There are only these two candidate values of q_M as profits and the highest transfer t_V are linear in the value of q_M that M offers. Comparing profit under these two candidate commitments, we find

$$k_M(a+r) > k_M\frac{r}{2} + k_V(a+\frac{r}{2}) \tag{52}$$

which is true only if $k_M > k_V$. The proposition follows from this. $\square$

A.1.4 Proof of Proposition 6

There are two candidate prices for t, either the highest price that both Q and V are willing to pay or the highest price that a single firm is willing to pay. The highest price that both Q and V are willing to pay must be $min(k_I(r/2), (k_I + k_V)(r/2) + max(0, (k_V - k_I)(a + r/2))$, that is, Q's revenue under competition, $k_I(r/2)$. The highest revenue that V could make renting k_I is the monopoly profit $k_I(a + r)$ which is evidently higher. The highest revenue that Q could make, given k_V, is $k_I(r/2) + max(0, (k_I - k_V)(a + r/2))$. This is lower than $k_I(a + r)$ for all $k_V > 0$. Whenever V sets $k_V > 0$, only V will accept. By setting $t = k_I(a + r)$, I ensures that V agrees to rent at the proposed price. Off-path, V never has an incentive to choose $k_V = 0$ as $c'_V(0) = \beta_V$ which is smaller than the marginal revenue of $r/2$ by assumption. $\square$

A.1.5 Proof of Proposition 7

We now analyze the model and its applications under congestion. Third stage prices conditional on available infrastructure remain unchanged: as products are still perfect substitutes and marginal costs are zero, the only prices that can emerge in equilibrium are $p_{V,c} = p_{Q,c} = 0$. However, revenues on this segment are now lower. Given that both firms rent access to I's infrastructure, V's price on the monopolistic segment remains unchanged at $p_{V,m} = a$.

The resulting profits are

$$\Pi_Q = \phi k_I(r/2) - t_Q \tag{53}$$

$$\Pi_V = \phi k_I(r/2) + (a+r)k_V - t_V - c_V(k_V) \tag{54}$$

Following the analysis from the base model, resulting transfers are

$$t_Q = \phi k_I(r/2) \tag{55}$$

$$t_V = \begin{cases} k_I\frac{r\phi}{2} + k_V\left(a + r(1-\frac{\phi}{2})\right) & \text{if } k_I > k_V \\ k_I(a+r) & \text{if } k_V \geq k_I \end{cases} \tag{56}$$

Note that the expression for t_V if $k_V \geq k_I$ is identical to the base model but if $k_I > k_V$, the expression for the transfer is different and lower.[89] In particular, the contribution of k_I to the transfer that I can demand is now lower, as each additional unit of k_I that increases the competitive segment has a lower return due to congestion. At the same time, the return to investment by V is higher because an increase in k_V and therefore an expansion of V's monopolistic segment allows V not only to escape competition with Q but also to escape the congestion "tax".

[89] Which can be easily seen from solving $(k_V + k_I)\frac{1}{2} > k_I\frac{\phi}{2} + k_V(1-\frac{\phi}{2}) \leftrightarrow k_I\frac{1-\phi}{2} > k_V\frac{1-\phi}{2}$ which is true when $k_I > k_V$.

The resulting first-period investments follow:

$$\max_{k_V} \Pi_V = d'_{V,c}(p_{V,c}+r) + d'_{V,m}(p_{V,m}+r) - t_V - c_V(k_V) \tag{57}$$

$$\frac{\partial \Pi_V}{\partial k_V} = a + r - c'_V(k_V) - \frac{\partial t_V}{\partial k_V} = 0 \tag{58}$$

$$\frac{\partial t_V}{\partial k_V} = \begin{cases} a + r(1-\frac{\phi}{2}) & \text{if } k_I > k_V \\ 0 & \text{if } k_V \geq k_I \end{cases} \tag{59}$$

$$\frac{\partial \Pi_V}{\partial k_V} = \begin{cases} c'_V(k_V) = \frac{r\phi}{2} \\ c'_V(k_V) = a + r \end{cases} \leftrightarrow \begin{cases} k_V = \left(\frac{r\phi - \beta_V}{2\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_V = \left(\frac{a+r-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{60}$$

$$\max_{k_I} \Pi_I = t_Q + t_V - c_I(k_I) \tag{61}$$

$$\frac{\partial \Pi_I}{\partial k_I} = \frac{\phi r}{2} + \frac{\partial t_V}{\partial k_I} - c'_I(k_I) \tag{62}$$

$$= \begin{cases} c'_I(k_I) = r\phi \\ c'_I(k_I) = a + (1+\phi/2)r \end{cases} \leftrightarrow \begin{cases} k_I = \left(\frac{r\phi - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_I = \left(\frac{a+(1+\phi/2)r-\beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{63}$$

□

A.1.6 Proof of Proposition 8

Working from the expressions for t_V, t_Q,

$$t_Q = \delta_Q \phi k_I (r/2) \tag{64}$$

$$t_V = \begin{cases} \delta_V \left(k_I \frac{r\phi}{2} + k_V \left(a + r\left(1 - \frac{\phi}{2}\right)\right)\right) & \text{if } k_I > k_V \\ \delta_V k_I (a+r) & \text{if } k_V \geq k_I \end{cases} \tag{65}$$

we work out the resulting first-period investments:

$$\frac{\partial t_V}{\partial k_V} = \begin{cases} \delta_V \left(a + r\left(1 - \frac{\phi}{2}\right)\right) & \text{if } k_I > k_V \\ 0 & \text{if } k_V \geq k_I \end{cases} \tag{66}$$

$$\frac{\partial \Pi_V}{\partial k_V} : \begin{cases} c'_V(k_V) = (1-\delta_V)a + r\left(1 - \delta_V\left(1 + \frac{\phi}{2}\right)\right) & \text{if } k_I > k_V \\ c'_V(k_V) = a + r & \text{if } k_V \geq k_I \end{cases} \tag{67}$$

$$\tag{68}$$

Using the cost function, we obtain the following levels of investment:

$$\leftrightarrow \begin{cases} k_V = \left(\frac{(1-\delta_V)a + r(1-\delta_V(1+(\phi/2))) - \beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_V = \left(\frac{a+r-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{69}$$

Similarly for firm I:

$$\frac{\partial \Pi_I}{\partial k_I} = \frac{\partial t_Q}{\partial k_I} + \frac{\partial t_V}{\partial k_I} - c_I'(k_I) \tag{70}$$

$$\frac{\partial t_Q}{\partial k_I} = \frac{\delta_Q \phi r}{2} \tag{71}$$

$$\frac{\partial \Pi_I}{\partial k_I} : \begin{cases} c_I'(k_I) = (\delta_V + \delta_Q)\frac{r\phi}{2} & \text{if } k_I > k_V \\ c_I'(k_I) = \delta_V(a+r) + \frac{\delta_Q \phi r}{2} & \text{if } k_V \geq k_I \end{cases} \tag{72}$$

$$\tag{73}$$

Using the cost function, we obtain the following levels of investment:

$$\leftrightarrow \begin{cases} k_I = \left(\frac{(\delta_V+\delta_Q)\frac{r\phi}{2} - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_I > k_V \\ k_I = \left(\frac{\delta_V(a+r) + (\delta_Q \phi r/2) - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} & \text{if } k_V \geq k_I \end{cases} \tag{74}$$

□

A.1.7 Proof of Proposition 9

V and Q charge identical prices on the competitive segment and face demand $d'_{V,c} = d'_{Q,c} = \phi k_I/2$. From this, we can compute third-period profits for given transfers and infrastructure size.

$$\begin{aligned} \Pi_Q &= \phi\frac{k_I}{2}(p_{Q,c} + r) - t_Q \\ &= \phi\frac{k_I}{2}\left(\frac{2a+2r}{4+\mu}\right) - t_Q \end{aligned} \tag{75}$$

$$\begin{aligned} \Pi_V &= \phi\frac{k_I}{2}(p_{V,c} + r) + (r+a)k_V - t_V - c_V(k_V) \\ &= \phi\frac{k_I}{2}\left(\frac{2a+2r}{4+\mu}\right) + (r+a)k_V - t_V - c_V(k_V) \end{aligned} \tag{76}$$

$$\tag{77}$$

Given these profits, we can now write transfers

$$t_Q = \delta_Q \phi \frac{k_I}{2}\left(\frac{2a+2r}{4+\mu}\right) \tag{78}$$

$$t_V = \begin{cases} \delta_V\left(\frac{2a+2r}{4+\mu}\phi\frac{k_I}{2} + k_V\frac{4+(a+r)(\mu-2\phi)}{4+\mu}\right) & \text{if } k_I > k_V \\ \delta_V k_I\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right) & \text{if } k_V \geq k_I \end{cases} \tag{79}$$

Finally, first period investment follows from the FOC:

$$\frac{\partial t_V}{\partial k_V} = \begin{cases} \delta_V\frac{4+(a+r)(\mu-2\phi)}{4+\mu} & \text{if } k_I > k_V \\ 0 & \text{if } k_V \geq k_I \end{cases} \tag{80}$$

$$\frac{\partial \Pi_V}{\partial k_V} : \begin{cases} c_V'(k_V) = r + a - \delta_V\left(\frac{4+(a+r)(\mu-2\phi)}{4+\mu}\right) & \text{if } k_I > k_V \\ c_V'(k_V) = a + r & \text{if } k_V \geq k_I \end{cases} \tag{81}$$

Similarly for firm I:

$$\frac{\partial \Pi_I}{\partial k_I} = \frac{\partial t_Q}{\partial k_I} + \frac{\partial t_V}{\partial k_I} - c_I'(k_I) \tag{82}$$

$$\frac{\partial t_Q}{\partial k_I} = \frac{\delta_Q \phi}{2}\left(\frac{2a+2r}{4+\mu}\right) \tag{83}$$

$$\frac{\partial \Pi_I}{\partial k_I} : \begin{cases} c_I'(k_I) = (\delta_Q + \delta_V)\frac{\phi}{2}\left(\frac{2a+2r}{4+\mu}\right) & \text{if } k_I > k_V \\ c_I'(k_I) = \delta_Q \frac{\phi}{2}\left(\frac{2a+2r}{4+\mu}\right) + \delta_V\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right) & \text{if } k_V \geq k_I \end{cases} \tag{84}$$

The parameter values underpinning the $k_I > k_V$ equilibrium are:

$$k_I = \left(\frac{(\delta_Q + \delta_V)\frac{\phi}{2}\left(\frac{2a+2r}{4+\mu}\right) - \beta_I}{\alpha}\right)^{\frac{1}{\alpha-1}} \tag{85}$$

$$k_V = \left(\frac{r + a - \delta_V\left(\frac{4+(a+r)(\mu-2\phi)}{4+\mu}\right) - \beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} \tag{86}$$

$$k_I > k_V \leftrightarrow \tag{87}$$

$$(\delta_Q + \delta_V)\frac{\phi}{2}\left(\frac{2a+2r}{4+\mu}\right) - \beta_I > r + a - \delta_V\left(\frac{4+(a+r)(\mu-2\phi)}{4+\mu}\right) - \beta_V \leftrightarrow \tag{88}$$

$$\beta_V > \beta_I + (\delta_Q + \delta_V)\frac{\phi}{2}\left(\frac{2a+2r}{4+\mu}\right) + r + a - \delta_V\left(\frac{4+(a+r)(\mu-2\phi)}{4+\mu}\right) \leftrightarrow \tag{89}$$

$$\beta_V > \beta_I + r + a + \delta_Q\left(\frac{\phi(a+r)}{4+\mu}\right) - \delta_V\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right) \tag{90}$$

And for the $k_V \geq k_I$ equilibrium:

$$k_I = \left(\frac{\delta_Q \frac{\phi(a+r)}{4+\mu} + \delta_V\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right)}{\alpha}\right) \tag{91}$$

$$k_V = \left(\frac{a+r-\beta_V}{\alpha}\right)^{\frac{1}{\alpha-1}} \tag{92}$$

$$k_V \geq k_I \leftrightarrow a + r - \beta_V \geq \delta_Q\left(\frac{\phi(a+r)}{4+\mu}\right) + \delta_V\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right) - \beta_I \tag{93}$$

$$\beta_I \geq \beta_V + \delta_Q\left(\frac{\phi(a+r)}{4+\mu}\right) + \delta_V\left(\frac{4+(a+r)(\mu-\phi)}{4+\mu}\right) - (a+r) \tag{94}$$

$$\tag{95}$$

□

A.2 Submarine cable ownership

In this Section I present data on submarine cable ownership by some firms that are traditionally labeled "big tech" firms. The purpose of this exercise is firstly, to motivate the model describing vertical integration by firms offering digital services, and secondly, to justify the assumption that only either very large or specialized infrastructure firms have the scale to make these investments.

Bischof et al. (2018) describe the increasing role of these firms in the submarine cable infrastructure: "The latest construction boom, however, seems to be driven by content providers, such [as] Google, Facebook, Microsoft, and Amazon. According to Telegeography's Research Director

Alan Mauldin, the amount of capacity deployed by content providers has risen 10-fold between 2013 and 2017, outpacing all other customers of international bandwidth."

I analyze data from Telegeography on submarine cables underlying the Submarine Cable Map.[90] The data is publicly available and as of September 2022 contains data on 516 submarine cables. The data set also includes location data on the cables which I do not use. Each observation of the data set includes the name of the cable, its length in kilometers, a list of its owners, a list of suppliers, and the year (and sometimes month) when the cable became or will become ready for service, ranging from 1989 to 2026.

Description

I search among the list of owners in our data set for Google, Amazon, Meta, Apple, Microsoft, Baidu, Alibaba, and Tencent. These firms are sometimes referred to with catch-all abbreviations such as GAMAM (GAFAM, before Facebook changed its name to Meta in 2021) or BAT. Neither Apple nor any of the BAT firms appear on the list of owners, but only Meta, Microsoft, Google, and Amazon Web Services. However, Alibaba does own terrestrial backbone within Asia (Corneo et al., 2021) Thus I identify cables that have one of the above-mentioned firms among its owners. This does not indicate sole ownership. Indeed, except for 7 purely Google-owned cables, all cables listed here as having "GAMAM owners" have co-owners. This is unsurprising, given that submarine cables typically include several fibre-optic cables and firms can own individual fibres.

I also search the list of owners for other firms and as a general observation, I remark that the list of owners includes mostly telecommunications firms and governments, as well as a few electricity companies, but no other firms that mainly sell digital services. While this is just one example of Internet infrastructure, it is consistent with the data previously collected in the literature (see references in this Section and Section 2.2) that describes the emergence of proprietary networks as a phenomenon driven by just a few of the largest technology firms.

Extending the range of the data of the previous paper by 7 years, I find that among submarine cables getting ready for service 2022-2024, the share of GAMAM climbs to between 20 and 27% (see Figure 7).[91] This is higher than the share of new cables owned by these first in the previous decade, which only exceeded 20% in one year (2018).

At the same time, the absolute number of new GAMAM-owned cables has quadrupled, from 1.1 new cables per year between 2010 - 2019, to 4.4 new cables for 2020 - 2024 (Figure 8). The overall increase in cables going ready for service has increased by a third during this period, from 15.7 new cables per year to 21 new cables for year. In other words, the cables added with the large technology firms as co-owners contribute more than half of the increase in the number of added cables between the the period before and after 2020. The phenomenon of these firms owning submarine cables is not a recent one, however, with the first such cable registered in 2010. Overall, the data confirms the increasing role of content firms among investors of infrastructure.

[90] https://www.submarinecablemap.com/

[91] Only 7 announcements have been made regarding cables that are ready for service in 2025 and only 1 for 2026, none of them involving any of the above-mentioned firms.

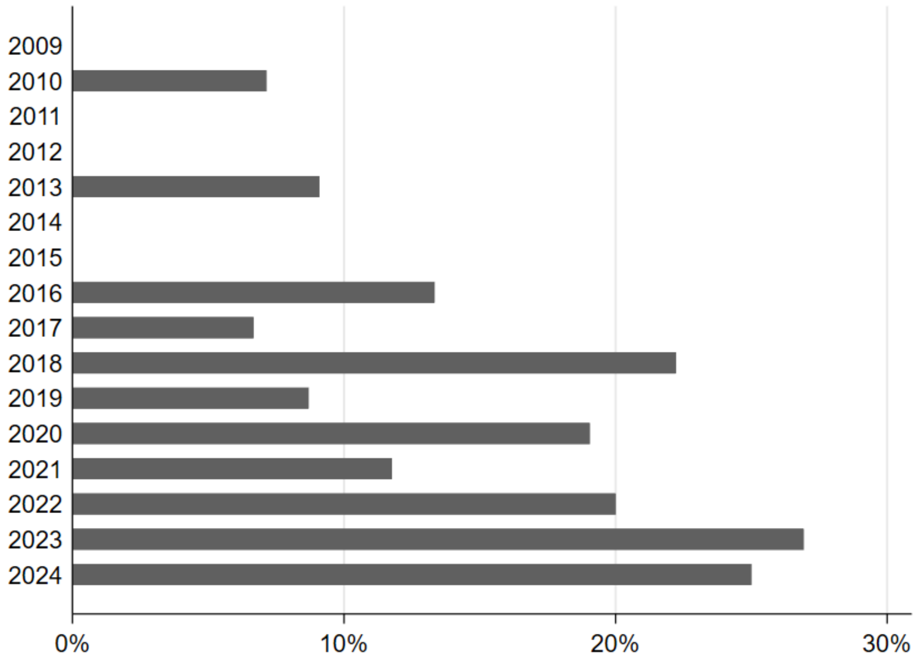


Figure 7: Share of submarine cables with GAMAM owners by ready-for-service date

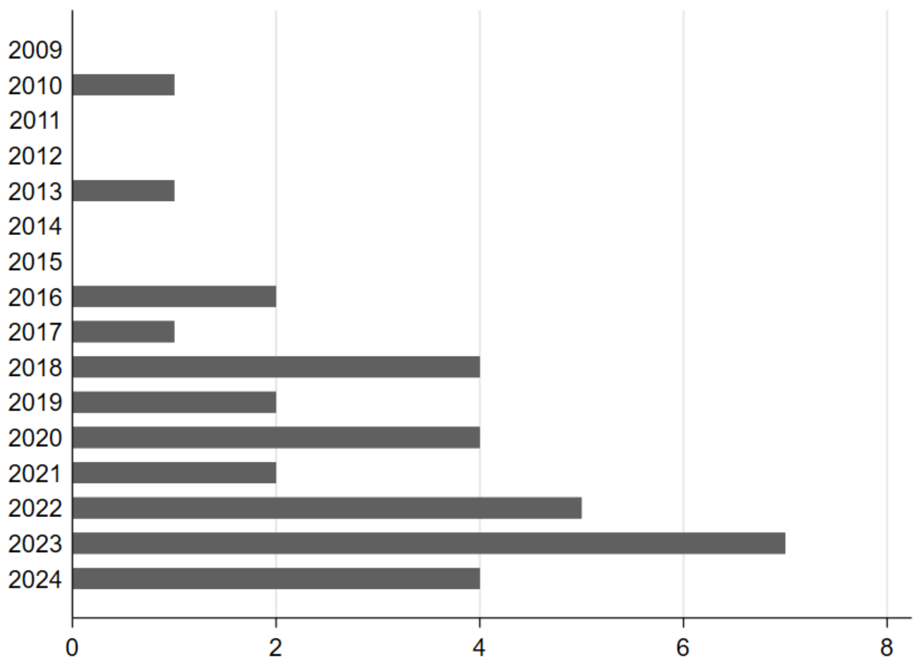


Figure 8: Sum of submarine cables with GAMAM owners by ready-for-service date

B Appendix to Chapter 2

B.1 Creation of the data set

The data set is created from individual XML files related to Contract Award Notices published on the website "TED Tenders electronic daily - Supplement to the Official Journal of the European Union".[92] The selection of relevant procurement contracts is via the pre-selected filter that is offered on a "COVID-19 dedicated page for tenders related to medical equipment needs".[93] The default filter is for contracts published between February 1, 2020 and December 31, 2020. The descriptive statistics distinguish between pre- and post-Covid contracts using this datum, which is in line with the WHO designation (see footnote 46).

I expand the default filters to include notices published as early as January 2018 to obtain a comparison period immediately preceding the pandemic. The contract date is given separately for the entire contract and its individual lots. In particular, the contract date for any individual lot, which is the unit of observation, may precede the publication date. The earliest lot in a sampled contract goes back as far as the year 2000. The filter also includes pre-selected codes from the Common Procurement Vocabulary (CPV).[94] The sample includes 461 distinct CPV codes at different levels of granularity. The first five digits of a CPV code describe the category, while three additional digits provide greater detail about the products. Different individual products are aggregated by category, using the first 5 digits of the CPV code, resulting in 311 different product categories in the data set.

I download a total of 9,233 contract award notices and voluntary ex-ante transparency notices.[95] These documents are complex and rich in features, each describing a procurement tender, giving information on the buyer (name, type and location of the authority), the object (total value of the procurement and product categories, of which only the first, main product category is observed if there are several), and possibly division of the contract into individual lots. For each lot, the document reports the number of companies that bid for the contract, as well as the number of foreign, non-EU, and small and medium-sized firms as bidders.

To create a data set at the level of individual contract awards, I process the files in two stages: first, I collect data that is common to all lots within one contract award notice, such as the date of publication, the location of the buyer, and the total value and currency of the procurement tender. In the second step, the section including the individual contract award is split at lot numbers to extract features that are specific to individual awards, such as the location and name of the contractor, the initially estimated and final value of the award, the contract date, as well as the number of bidders, further divided into bids from other EU countries, bids from outside the EU, and bids from small and medium-sized firms. These award features are then combined with the contract-specific information.

The unit of observation is a separate contract award to a contractor. Each lot that is awarded to a single company is one observation. Where several contract awards are listed under one lot number, each contract award is also counted as one observation. Entries where no contractor name or contract date exist represent failed procurement tenders. In these cases, no contract was awarded. These observations are removed from the main analysis. I do, however, construct a weekly panel by country of the total amount of failed procurement at the start of each week, which I combine with the cross-section of contract awards for a robustness check (see Appendix

[92] https://ted.europa.eu

[93] https://simap.ted.europa.eu/web/simap/covid-related-tenders

[94] "The CPV establishes a single classification system for public procurement aimed at standardising the references used by contracting authorities and entities to describe the subject of procurement contracts." https://simap.ted.europa.eu/web/simap/cpv.

[95] The latter, which only make up 0.1% of documents and 0.3% of awarded value, do not enter the regression analysis, because they contain no information about the number of bidders or foreign bidders - they all refer to directly negotiated contract awards. They are considered in Figure 4b to assess the share of non-competitive tenders.

B.3). In the remaining data set, there are 146 awards where no contract date has been listed. In these cases, contract date is approximated by publication date.

The main outcome variable is an indicator that takes the value 1 if the country of the contractor is identical to the country of the buyer. Procurement by EU agencies is dropped because the outcome variable related to domestic firms does not apply to procurement by the EU itself. Thus, I obtain a total of 125,301 observations of individual awards, some of which do not enter the main regression analysis due to lack of data on contract values.

This data is combined with monthly average exchange rates for all non-Euro currencies. Exchange rates are provided by the Statistical Data Warehouse of the European Central Bank, data set “EXR : Exchange Rates”, supplemented by individual values for the Macedonian Denar and Icelandic Kronur for six different months from “currencies.zone”, an exchange-rate information provider. All non-Euro values are then converted to Euro using the prevailing exchange rate. In a few cases, token values such as 0.01, 1, 99,999,999 or 999,999,999 appear in the data set. These are set to missing. I manually verified for the large values that these are indeed token or placeholder values, which is sometimes explicitly noted as a comment in the document or apparent from inconsistencies between total value and the value of awards.

Some contracts don’t specify a total value of the procurement at the top of the contract, but instead report a lower and upper range. Missing values for the total contract value are imputed with midpoint between the higher and lower bound of this range. To be conservative and not introduce noise (e.g., from very imprecise ranges), I only do this in cases when the upper bound is less than 10 times larger than the lower bound, so the error is not greater than an order of magnitude. In any case, the total value of the contract only enters descriptive statistics and a robustness check on the total amount of failed procurement. The main regressions use the value at the award level for which the value is always an exact amount.

Next, I add data on infection rates from the European Centre for Disease Prevention and Control (ECDC) at a national and sub-national level.[96] Infection rates are matched with the performance regions of the contracts. This is a measure of the pandemic intensity at the location of interest to the buyer. Some data cleaning is required to deal with inconsistent entries (e.g. “CZ0” appears as well as “CZ”) and to match infection rates and contract location at the same level. Infection rates are available at different levels of aggregation depending on the country. Where only infection rates for the region at the NUTS3 level is available when contract information is at the NUTS2-level, the observation is matched with the first NUTS3-region in the data. In most cases, the location is reported at a more detailed level (NUTS3), which can be easily matched with a NUTS2 code through the first two numbers following the country code.

In rare cases, the performance location is listed as a NUTS1 code, in which case it is matched to the infection rate of the first applicable NUTS2 code. When only the country is listed as location of performance, I use national data on the infection rate. I also use national data to complete entries before the 14-day sub-national average is started being reported. Dates before the first reported case in a country are assumed to have 0 infections. The procedure is repeated to match infection rates with seller location NUTS-codes and Covid-death rates for robustness checks reported in Appendix B.3. Finally, recent population data for the different NUTS codes from Eurostat is added to compute the average national infection rate excluding individual regions one at a time for a robustness check.

The data for the control group of goods closely related to the Covid-19 related medical supplies are retrieved and treated similarly. The control group comprises all product groups of the CPV that are at the same hierarchy level as the Covid-related tenders. The CPV of Covid-19 related tenders are pre-defined by TED. Table 13 shows a sample of the 969 CPV that were chosen to complement the treatment product groups:

[96] https://qap.ecdc.europa.eu/public/extensions/COVID-19/COVID-19.html#subnational-transmission-tab

Table 13: Selection of product groups by CPV for treatment and control group

Treatment	Control
45215142 Intensive-care unit construction work	45215141 Operating theatre construction work, 45215143 Diagnostic screening room construction work, 45215144 Screening rooms construction work, 45215145 Fluoroscopy room construction work, 45215146 Pathology room construction work, 45215147 Forensic room construction work, 45215148 Catheter room construction work
35113400 Protective and safety clothing	35113100 Site-safety equipment, 35113200 Nuclear, biological, chemical and radiological protection equipment, 35113300 Safety installations
18143000 Protective gear	18143100 Work gloves, 18143200 Safety visors
18424300 Disposable gloves	18424400 Mittens, 18424500 Gauntlets
33141420 Surgical gloves	33141410 Wire cutter and bistoury
33157000 Gas-therapy and respiratory devices	33151000 Radiotherapy devices and supplies, 33152000 Incubators, 33153000 Lithotripter, 33154000 Mechanotherapy devices, 33155000 Physical therapy devices, 33156000 Psychology testing devices
33192120 Hospital beds	33192110 Orthopaedic beds, 33192130 Motorised beds, 33192140 Psychiatric couches, 33192150 Therapy beds, 33192160 Stretchers, 33192200 Medical tables, 33192300 Medical furniture except beds and tables, 33192400 Dental workstations, 33192500 Test tubes, 33192600 Lifting equipment for health care sector
33195110 Respiratory monitors	33195200 Central monitoring station
33670000 Medicinal products for the respiratory system	33610000 Medicinal products for the alimentary tract and metabolism, 33620000 Medicinal products for the blood, blood-forming organs and the cardiovacular system, 33640000 Medicinal products for the genitourinary system and hormones, 33650000 General anti-infectives for systemic use, vaccines, antineoplastic and immunodulating agents, 33660000 Medicinal products for the nervous system and sensory organs, 33680000 Pharmaceutical articles, 33690000 Various medicinal products
33363600 Antiseptics and disinfectants	33363100 Antifungals for dermatological use, 33363200 Emollients and protectives, 33363300 Antipsoriatics, 33363400 Antibiotics and chemotherapeutics for dermatological use, 33363500 Corticosteroids for dermatological use and dermatological preparations, 33363700 Anti-acne preparations, 33700000 Personal care products, 33900000 Post-mortem and mortuary equipment and supplies
39330000 Disinfection equipment	39310000 Catering equipment, 39340000 Gas network equipment, 39350000 Sewerage works equipment, 39360000 Sealing equipment, 39370000 Water installation

B.2 Data description

This section reports additional summary statistics. All sums are converted to Euros, where necessary, and exclude VAT. Table 14 reports the summary statistics for the control group of the analysis of the regulation effect, that is, all those contracts that were in product groups adjacent to the Covid-19-related medical supplies for which procurement rules were lifted.

Table 14: Descriptive statistics of tenders for product groups in the control group, n = 252,575

	Mean	Std. dev.	Minimum	Median	Maximum
Lot value (excluding VAT) in EUR	162,292.00	7,076,075.20	0.07	1,380.92	1,990,000,000
Indicator: domestic award	0.99	0.07	0	1	1
14-day average infection rate per 100 inhabitants	0.02	0.10	0	0	1.87
Infection rate at seller location	0.02	0.10	0	0	1.87
Total number of bidders	4.18	7.26	1	3	350
Share of foreign bidders	0.00	0.06	0	0	1

Table 15 reports the sum of lot values for all contracts in the data set (published between 2018 and 2020, but including some contracts with earlier contract dates) and thus represents the weights of different countries in the regression analyses.

Table 15: Total value of individual contract awards published 2018-2020 by buyer country (EUR equivalent, excluding VAT)

Country	Total value	Country	Total value	Country	Total value
United Kingdom	3,825,440,159.92	Hungary	98,826,759.14	Greece	5,806,930.67
Poland	1,172,433,382.38	Slovakia	72,286,834.36	Cyprus	5,735,658.40
Ireland	771,395,491.32	Belgium	65,088,034.40	Portugal	3,481,540.76
France	767,024,944.62	Slovenia	62,360,505.66	Iceland	475,862.22
Italy	517,696,183.13	Germany	60,982,075.65	Lithuania	336,143.81
Romania	423,224,861.66	Finland	51,526,898.00	Estonia	10,190.00
Austria	402,774,108.00	Norway	43,775,788.39		
Denmark	336,247,169.13	Croatia	33,510,701.05		
Spain	215,683,821.81	Bulgaria	30,046,518.24		
Czech Republic	205,983,843.36	Netherlands	18,257,500.94		
Sweden	109,544,231.72	Latvia	6,166,053.67		

Sum of the individual contract awards by buyer country, published as "result[s] of tenders related to COVID-19" on a page with "COVID-19-related tenders" on https://ted.europa.eu/TED/.

Table 16 lists the 10 largest cross-border contract awards after February 2020. Except for one, they are all purchases from the UK. All of these purchases were under the CPV code for "Garments for biological or chemical protection" which encompasses the most common types of PPE, including face masks. These contracts range in value from almost 350 million Euros to just under 80 million Euros and most have been awarded to non-EU countries. This suggests that while cross-border procurement reached unprecedented levels in the pandemic, Covid-19 did not "integrate the Single Market", but rather pushed buyers that previously purchased domestically to an international procurement strategy.[97]

[97] Among the 50 largest lots, which includes contracts above 2.8 million Euros, there are also three French purchases of protective gear from Germany and Switzerland, one Czech purchase of medicinal products from Ireland, one Romanian purchase of medical breathing devices from Germany, three Italian purchase of medical

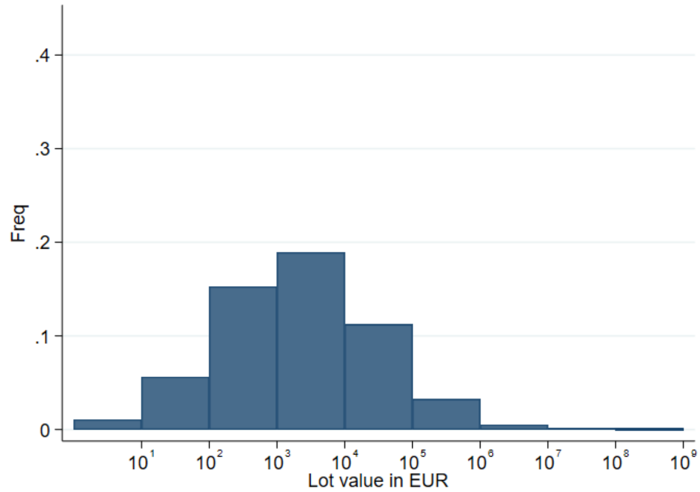


Figure 9: Distribution of lot value for contract awards, Covid-19 related tenders in EUR (log scale)

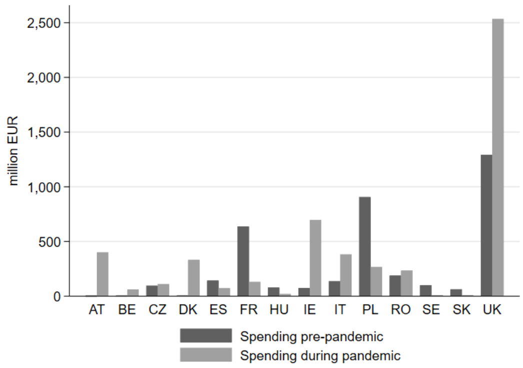


Figure 10: Spending 2018-2020 on Covid-19-related product groups, top 14 countries

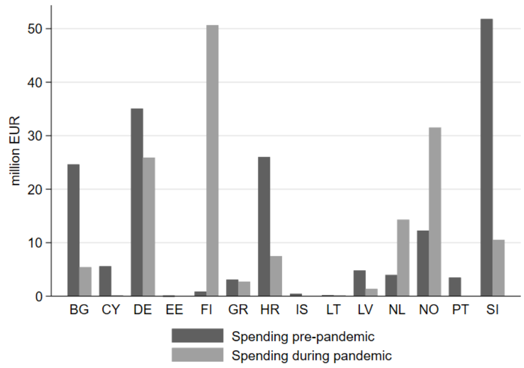


Figure 11: Spending 2018-2020 on Covid-19-related product groups, other countries

Table 16: Largest cross-border purchases

Buyer country	Seller country	Contract date	Contract value (EUR equivalent)
United Kingdom	China	20/05/2020	346,725,371
Ireland	China	03/04/2020	225,128,510
United Kingdom	Hong Kong	28/05/2020	126,582,278
United Kingdom	China	06/06/2020	110,173,256
United Kingdom	United States	04/06/2020	91,642,825
United Kingdom	Austria	30/05/2020	87,951,739
United Kingdom	Hong Kong	04/05/2020	82,553,660
United Kingdom	China	03/06/2020	81,705,908
United Kingdom	China	21/05/2020	79,801,871
United Kingdom	United States	04/06/2020	74,041,759

To compare monthly spending by country and in the 20 largest product categories, Figures 10 to 13 are reproduced for monthly values by dividing the pre-pandemic sum by 25 and the

consumables and garments from Singapore and China, while the remainder are UK purchases mostly again of garments for biological or chemical protection from mostly non-EU countries.

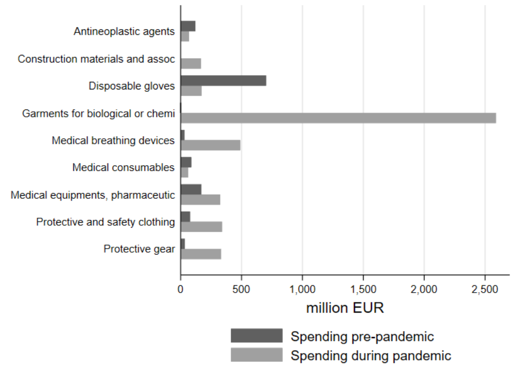


Figure 12: Spending 2018-2020 by product category, top 9 categories

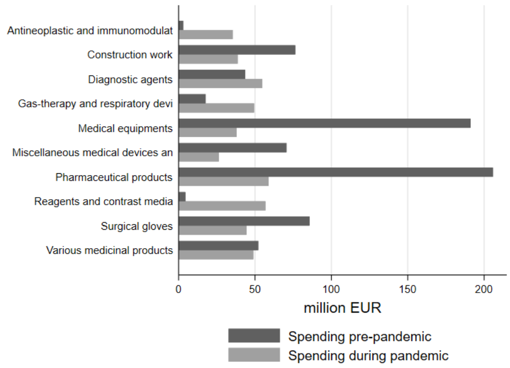


Figure 13: Spending 2018-2020 by product category, other categories

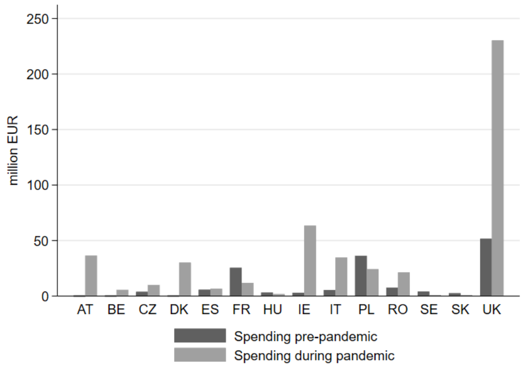


Figure 14: Average monthly spending 2018 - 2020 by country on Covid-19-related products, top 14 spenders

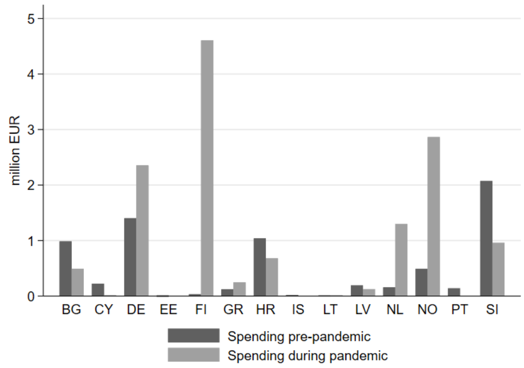


Figure 15: Average monthly spending 2018 - 2020 by country on Covid-19-related products, other spenders

sum during the pandemic by 11, reflecting the period covered in the data set. Of interest is the observation that for some countries average spending was higher in the two years preceding the pandemic than in the first 11 months of the pandemic, including Germany, Croatia, and Slovenia. This comparison also shows that monthly average spending did remain approximately constant in some product groups, including construction (related to medical work), which might reflect long-term planning for some kinds of products and services. In most of the largest product groups (Figure 16), average monthly spending during the pandemic exceeded average monthly spending before the pandemic by a wide margin.

One possible concern is that the two sets of products are not comparable due to a surge in demand for the products in the treatment group, which might not carry over to the products in the control group. If this was the case, total purchases for products in the control group would have remained constant while purchases in the treatment group spiked. Average monthly purchases by product group are reported in Figures 16 and 17. For comparison, the average monthly purchases of the 20 largest product groups in the control group are shown in Figures 18 and 19.

The following weeks are weeks where some region reports its first infection: those starting on January 20, 27, February 3, 17, 24, and March 9, 16. Based on this, observations are categorized into those that experience their first infections either early or late. For easier visualization, I classify regions that first experienced infections in the weeks starting January 20 or 27 as “early” and the rest as “late”. Figure 22 compares the evolution of the share of domestic purchases and the average infection rates.

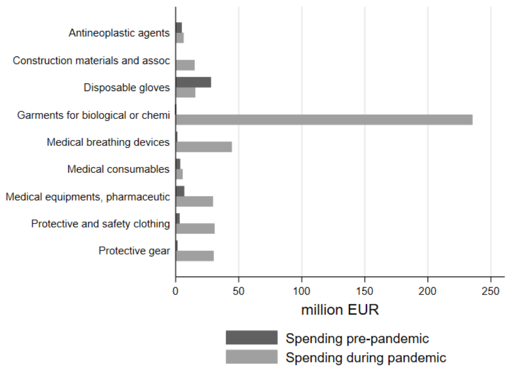


Figure 16: Average monthly spending 2018 - 2020 by Covid-19-related product category, top 9 categories

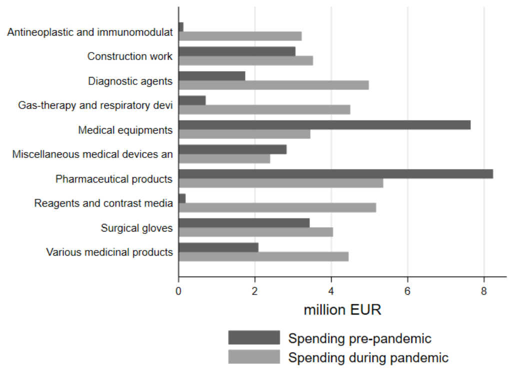


Figure 17: Average monthly spending 2018 - 2020 by Covid-19-related product category, other categories

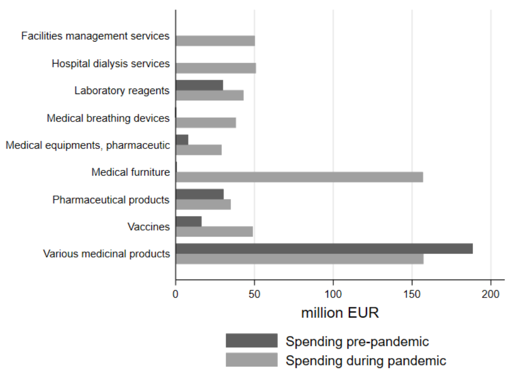


Figure 18: Average monthly spending 2018 - 2020 by product category, top 9 categories (control group products Section 2.4.3)

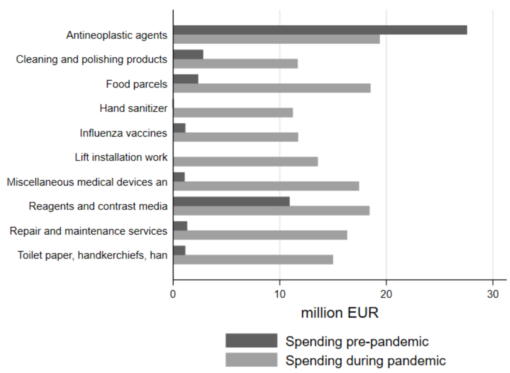


Figure 19: Average monthly spending 2018 - 2020 by product category, other categories (control group products Section 2.4.3)

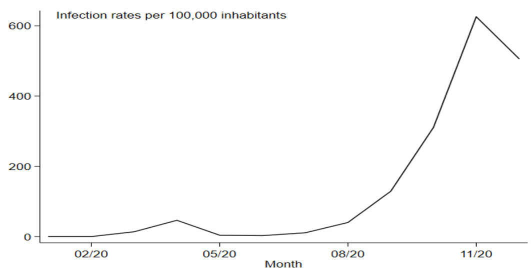


Figure 20: Monthly averages of the 14-day infection rate per 100k inhabitants in 2020 (weighted by purchase value)

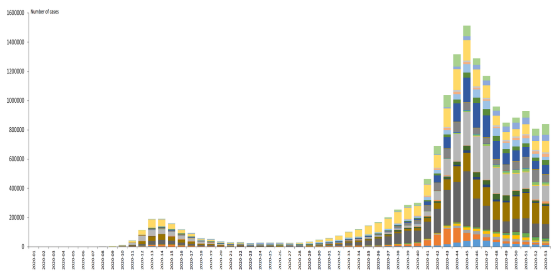


Figure 21: Weekly European infection rate according to ECDC, (c) 2021 European Centre for Disease Prevention and Control

In both figures, two red vertical lines mark the weeks beginning January 20 and 27, as a marker of the "early infection" group.

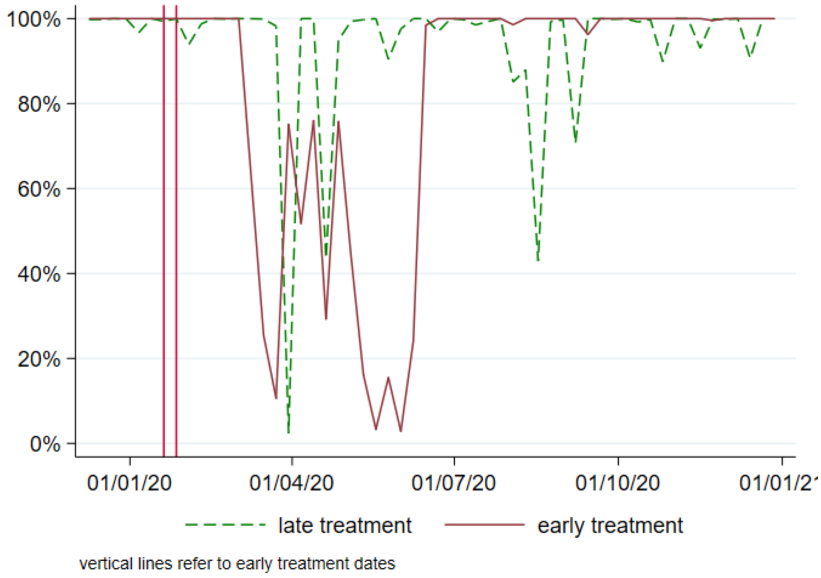


Figure 22: Domestic purchases by infection date

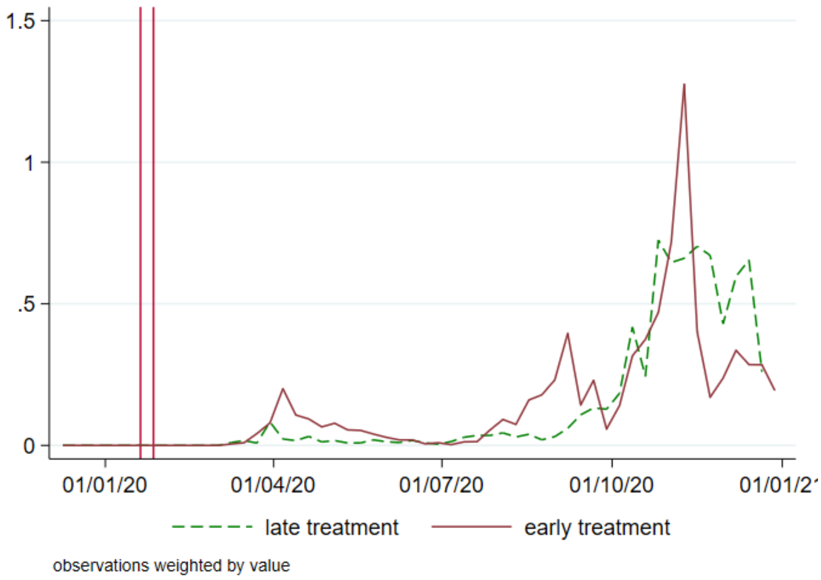


Figure 23: Average infection rate per 100 inhabitants in regions infected early/late

B.2.1 Sample selection

Procurement documents are sometimes incomplete, leading to missing values. These values are most likely not missing at random, yet a regression analysis shows that they likely just result in a downward bias of regression estimates, leaving us with a lower bound on possible effect sizes. Table 17 presents a least-squares regression of a dummy variable that takes the value 1 if the value of a contract award is missing and 0 if not. The regressor is the main dependent variable, an indicator variable for contract award to a domestic company.

Table 17: Likelihood of missing value conditional on domestic awards

Dep. var.: Award value is missing		
Domestic award	0.013	0.084***
	(0.019)	(0.019)
Controls & dummies	no	yes
N	121,120	120,962

Robust standard errors in parentheses, *** p<0.01.

Absent control variables, there is no strong correlation between domestically awarded contracts and missing information on contract value. Controlling for local infection rates for the buyer and seller, number of bidders, share of foreign bidders, as well as dummies for buyer country, product group and month-year, domestic contracts are on average circa eight percentage points more likely to not have information on award values.

Contracts below an administrative notification threshold don't have to be reported. Contracts that should normally fall above a notification threshold can be split or shaded intentionally to fall below the threshold and go unreported. While this may also affect tenders for medical supplies, they are unlikely to impact the estimates in the absence of systematic differences in contract size between times of high cross-border awards and other periods.

Figure 24 shows the share of contracts of different sizes in 2020. The categories follow the different thresholds that the EU uses to identify contracts presumed to be of cross-border interest.[98] I aggregate contracts by size and plot their relative frequency by month in 2020.[99] The main period of high cross-border awards, April to June 2020 (marked by a dashed box), is not obviously anomalous from visual inspection beyond usual fluctuations. There is also no significant seasonal variation in the main outcome or independent variables. Also, one would expect that contracts where misallocation plays a role to be more likely to go unreported. Suppose such contract awards are more likely to be concealed, for example to avoid scrutiny by oversight authorities. Then this would imply that cross-border awards are oversampled. However, since the baseline of cross-border awards is already very small (less than 1% as a simple share of all contracts), such an effect would likely be small as well.

Expanding on the discussion of missing values in Section 2.3, the presence of selection bias and unreported values would suggest that the distribution of contract award sizes was highly abnormal in this time period. Although the share of contracts in these different size categories exhibits some fluctuation year-round, the average within each category seems to be not too far off for the three months within the box and the remaining nine months. An absence of large contracts in the period of high cross-border awards would be more concerning because this might suggest that large contracts would have been deliberately not reported or shaded to avoid exceeding reporting thresholds. Visual inspection of Figure 24 does not suggest lower reporting of large contracts in the period of large cross-border awards.

B.2.2 Granger causality and spillover

This Section contains the tables referenced in Section 2.4.2 to test for the information value of infection rates in other regions and possible spillover effects. Allowing for up to six lags of the

[98]See the website of the European Commission under "Internal Market, Industry, Entrepreneurship and SMEs", https://ec.europa.eu/growth/single-market/public-procurement/rules-implementation/thresholds_en.

[99]The main value thresholds above which contracts have to be reported are 135,000 EUR for goods and services and 5,350,000 EUR for (subsidized) works contracts, which in the present context can apply to construction work. A few other thresholds may apply in special circumstances which is not identifiable from contract observables. See https://ec.europa.eu/growth/single-market/public-procurement/rules-implementation/thresholds_en

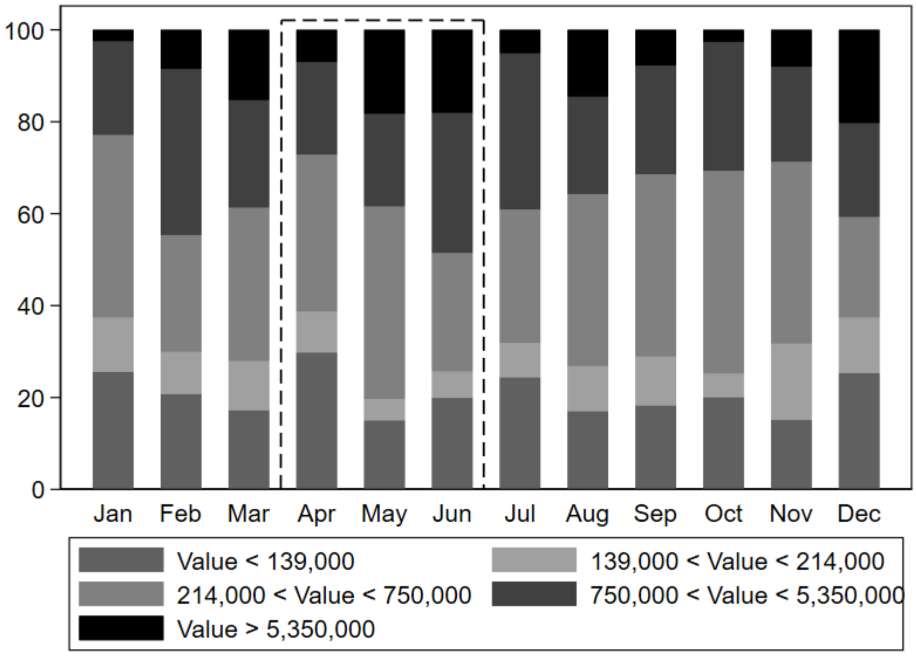


Figure 24: Contract award sizes in 2020 in EUR

weekly updated infection data, I fail to reject the Null-hypothesis that the infection rates in one region are not Granger-caused by infection rates in other regions.

Table 18: Test for Granger-causality of local infection rates

Dep. var.: Local infection rate	
Infection rate other regions	-0.017
	(0.017)
Number of periods	349
Number of units	46

Standard errors allowing for cross-sectional heteroskedasticity in parenthesis. Up to nine lags of the explanatory variable tested, six lags chosen based on Bayesian Information Criterion.

Table 19: Test for Granger-causality of domestic awards

Dep. var.: Domestic award	
Infection rate other regions	0.092
	(0.085)
Constant	0.990***
	(0.006)
R-squared	0.001
N	842

Robust standard errors in parenthesis, *** p<0.01.

B.3 Robustness

This section includes robustness checks to the regression analysis in Section 2.4.

Alternative emergency variables

The claim of the infection rates as a measure for crisis emergency, or urgency, should persist for alternative measures. In particular, one might worry about possible feedback effects from procurement on local infection rates. To alleviate such concerns, an alternative measure of crisis urgency is used to address feedback effects from procurement on infection rates. I compute the infection rate in all other regions of a country except for the one of the performance location of the contract. This helps excluding effects of local procurement outcomes on the infection rate.

This measure is computed from the absolute number of infections in each country and the region indicated in the contract award using population and the infection rate, then computing the difference between the two, and dividing by country population minus regional population to obtain the average infection rate in the country excluding the region of the observation. This measure helps us capture the common trend of infections within a country while excluding potential local feedback effects. Additionally, spillover effects are studied separately in Section 2.4.2 with the result that other regions of a country do not provide additional information about future infections over local past infections.

Using this alternative variable in the regression analysis (Table 20) yields find large and statistically significant effects for the regulation change. They are of similar quantitative size both for the full sample and the leave-one-out analysis where again leaving out the UK provides a lower bound of the effect size: In the full sample (sample excluding UK), a one-standard-deviation increase in national infections excluding the local region, results in a 25.1 (10.5) percentage points higher share of cross-border procurement.

Domestic capacity constraints

The results are robust even when accounting for buyers that had to turn to foreign sellers because of domestic capacity constraints. Consider a subsample of contracts for which the greatest number of domestic bidders for a contract (across all lots) is greater than the number of distinct domestic winners for lots on that contract. In this subsample, we identify the extensive margin of domestic spare capacity for these contracts. As discussed in Section 2.4.4, this method is likely conservative for domestic spare capacity. The results are reported in Tables 21 and 22.

The effect of a one standard deviation in the average infection rate at the buyer's location is approximately the same, at circa 20 percentage points. The effect of regulation is mitigated to circa 9 percentage points, compared to 35 percentage points in the baseline regression. Still, the effect size and statistical significance of the main variables remains high. Domestic capacity constraints were certainly important during the pandemic (affecting up to almost two-thirds

Table 20: National infection rate (excluding local infections) as alternative explanatory variable

	All countries	Excluding UK
Dep. var.: Contract awarded to domestic company		
ATET		
Infection rate other regions same country	-1.793***	-0.527***
	(0.405)	(0.141)
Controls		
Infection rate at seller location	2.050***	0.427**
	(0.398)	(0.168)
Total number of bidders	0.002***	0.000
	(0.001)	(0.000)
Share of foreign bidders	-0.361***	-0.428***
	(0.099)	(0.061)
Dummies	yes	yes
Standard error of infection rate	0.14	0.20
1-std.dev. increase	-0.251	-0.105
N	67,638	67,387

Robust standard errors in parentheses, ** p<0.05, *** p<0.01.

of contracts) but do not suffice to explain the movements in cross-border procurement in the pandemic.

Table 21: Main regression Section 2.4.2 on a subsample with domestic spare capacity

	All countries	Excluding UK
Dep. var.: Contract awarded to domestic company		
ATET		
14-day average infection rate per 100	-0.896***	-0.908***
	(0.252)	(0.258)
Controls		
Infection rate at seller location	1.212***	1.222***
	(0.405)	(0.408)
Total number of bidders	0.001**	0.001**
	(0.000)	(0.000)
Share of foreign bidders	0.057	0.051
	(0.107)	(0.110)
Dummies	yes	yes
Standard error of infection rate	0.22	0.23
1-std. dev. increase	-0.197	-0.209
N	23,443	23,294

Robust standard errors in parentheses, ** p<0.05, *** p<0.01.

Unweighted regression

I re-run the first difference-in-difference regression, but do not weight contracts by value. The results are reported in Table 23. I also use contracts for which no value is available, leading to a greater number of observations. However, a contract split into multiple small lots (observations) now carries much greater weight than a single, large contract compared to the main regression. As there is no significant variation of the outcome variable at this level (documented in Section

Table 22: Main regression Section 2.4.3 on a subsample with domestic spare capacity

	All countries	Excluding UK
Dep. var.: Contract awarded to domestic company		
ATET		
Regulation change	-0.088**	-0.089*
	(0.043)	(0.052)
Controls		
Infection rate at seller location	0.214	0.033
	(0.137)	(0.020)
Total number of bidders	-0.000	0.001***
	(0.000)	(0.000)
Share of foreign bidders	-0.147***	-0.545***
	(0.027)	(0.079)
Dummies	yes	yes
N	276,018	275,080

Robust standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

2.3.3), I do not expect to find any correlation with the explanatory variables.

Indeed, due to the great differences in award values, as well as the more significant variation of the outcome variables, the results are null. Although ignoring contract size allows us to use more observations, no meaningful analysis with regards to the outcome of interest is possible. Effects disappear (or become economically insignificant in the second column) when attempting to estimate this regression without taking into account the differences in lot size through their monetary value.

Table 23: Unweighted regression for contract award to domestic companies

	Infection rate	Regulation change
Dep. var.: Contract awarded to domestic company		
ATET		
Infection rate	-0.012	
	(0.011)	
Regulation change		-.004***
		(0.000)
Controls		
Infection rate at seller location	0.029***	0.006**
	(0.011)	(0.003)
Total number of bidders	0.000**	0.000***
	(0.000)	(0.000)
Share of foreign bidders	-0.505***	-0.597***
	(0.023)	(0.011)
Dummies	yes	yes
Standard error of infection rate	0.22	
1-std.dev. increase	-0.003	
N	120,922	540,413

Robust standard errors in parentheses, ** $p<0.05$, *** $p<0.01$.

Analysis of failed procurement contracts

Information about previously failed procurement contracts is potentially useful. For example, repeatedly failing to purchase desired goods or services may induce a buyer that discriminates against foreign sellers to reduce the extent of discrimination to avoid failed tenders in the future.

I analyze the total contract value of failed tenders (i.e., where no contractor and contract date are given) and create a weekly time series of the cumulative failed value of procurement for every country, starting in February 2020, thus restricting this analysis only to the time period that is indeed related to the Covid-19 pandemic. I add the Euro-equivalent value of total failed procurement at the time of each award and a linear time-trend as explanatory variables. The results are reported in Table 24. There is no statistically significant effect of failed previous procurement, nor does inclusion of this variable change the estimated effect drastically.

Table 24: Including the cumulative sum of failed tenders

	All countries
Dep. var.: Contract awarded to domestic company	
ATET	
Infection rate	-1.207***
	(0.300)
Controls	
Cumulative sum of failed tenders	0.000
	(0.000)
Infection rate at seller location	2.346***
	(0.432)
Total number of bidders	0.003***
	(0.001)
Share of foreign bidders	-0.614***
	(0.073)
Standard error of infection rate	0.21
1-std. dev. increase	-0.254
N	21,111

Robust standard errors in parentheses, *** $p<0.01$.

Inference based on clustered standard errors

Errors may be correlated across countries due to similarity in regulatory and legal environments and similarities in training, doctrine, and perspective among public sector buyers within each country. I re-estimate the difference-in-difference regressions in Tables 2 and 3 using clustered standard errors with clustering at the country-level. Standard errors are indeed larger with clustering, although only moderately. Only for the full-sample regression regarding infection rates are the results no longer statistically significant at the 5%-level. In the remaining regressions, the precision of the estimates remains high relative to the point estimates.

B.4 Model

This section briefly illustrates the impact of urgency and buyer discretion in a simple model of monitoring. Even though this setting is simplified, it is informative to predict the presence of misallocation in procurement. The key comparative statics include a government's cost of reviewing and punishing collusion and the importance of a procurement agent's information.

The former may represent the time and effort that needs to be expended to review a tender decision, litigate an outcome, and enforce a compensation for foregone benefits. Regulation that

Table 25: Difference-in-difference analysis for infection rates with clustered standard errors

	All countries	Excluding UK
Dep. var.: Contract awarded to domestic company		
ATET		
14-day average infection rate per 100	-1.376*	-0.463***
	(0.718)	(0.166)
Controls		
Infection rate at seller location	2.242	0.529
	(1.398)	(0.312)
Total number of bidders	0.002	0.000
	(0.001)	(0.001)
Share of foreign bidders	-0.363**	-0.428**
	(0.136)	(0.189)
Dummies	yes	yes
N	67,638	67,387

Country-clustered standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Infection rates are 14-day moving average per 100 inhabitants at the NUTS-region reported as performance location (buyer infection rate) or as location of the contractor (seller infection rate). Share of foreign bidders computed as number of foreign bidders from EU and non-EU countries divided by total number of bidders.

limits the discretion of buyers can be understood to directly reduce these costs. For example, when a regulator limits buyer discretion by mandating transparent tenders, prescribing scoring rules, and making the bidding process more transparent, he effectively reduces the cost of verifying violations of such rules.

Think of the value of the agent's information as increasing during an emergency. In particular, consider the multitude of policy objectives regarding procurement, some of which are completely unrelated to each tender (such as promoting jobs or small businesses). I assume that an increase in the urgency of a crisis makes objectives related to the tender relatively more important than secondary policy goals which may not require such information. For example, a good tender decision might require detailed information about the quality the seller, while simply spending money to secure local jobs does not. So, an increase in the urgency of a crisis should increase the importance of the bureaucrat's information. From this model, I predict that an increase in crisis urgency should result in less misallocation.

It is inspired by the mechanism-design models based upon McAfee and McMillan (1989) and Laffont and Tirole (1991). These models concern optimal auction contracts in a principal-agent-firm framework where the agent can collude with the firm (as opposed to collusion among firms). Whereas Laffont and Tirole (1991) is concerned with collusion-proof auctions and how they differ from the first-best, other models generate collusion in equilibrium (Burguet and Che, 2004; Burguet, 2017), or study dynamics and steady-state levels of corruption (Menezes and Monteiro, 2006).

Branco (1994) argues that foreign profits should not enter the utility function of the principal when the principal is a national government. In this case, favoritism may be an efficient outcome also for the principal, although it implies an international coordination problem between governments. This problem has been compared to a prisoner's dilemma where governments "want domestic protectionism and foreign liberalization, but they may prefer mutual liberalization to mutual protectionism" (Rickard and Kono, 2014). Arozamena and Weinschelbaum (2011) offer a different view and suggest that when entry into the auction is important, the principal should not favor a subset of bidders even if it cares about their payoffs, but not about the payoffs of

Table 26: Difference-in-difference analysis for the effect of deregulation with clustered standard errors

	All countries	Excluding UK
Dep. var.: Contract awarded to domestic company		
ATET		
Regulation change	-0.357***	-0.160**
	(0.062)	(0.074)
Controls		
Infection rate at seller location	0.394	0.058
	(0.319)	(0.060)
Total number of bidders	0.001*	0.001*
	(0.000)	(0.001)
Share of foreign bidders	-0.030	-0.504***
	(0.135)	(0.146)
Dummies	yes	yes
N	320,213	319,173

Country-clustered standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01. Infection rates are 14-day moving average per 100 inhabitants at the NUTS-region reported as performance location (buyer infection rate) or as location of the contractor (seller infection rate). Share of foreign bidders computed as number of foreign bidders from EU and non-EU countries divided by total number of bidders.

other bidders.

However, these models are not straightforward to take to the data as they often consider different cases, such as verifiable and unverifiable information, often with different implications. This model considers an explanation that Laffont and Tirole (1991) suggest to explain the small fraction of cross-border procurement, a fact which they suggest as a screen for misallocation, but that to my knowledge has not received much attention in this literature: rather than designing the auction himself, the principal might have to contend with reviewing procurement decisions and punishing collusion ex-post.

To be sure, Laffont and Tirole (1991) think of this as a reputation game: concerns about a reputation to “keep their mouths shut” might disincentivize firms that were unfairly treated to blow the whistle on a buyer who discriminates against them to avoid losing whatever rents they were left with. By contrast, I study a simple monitoring game between a bureaucrat who decides whether to bias his decision and a government that decides to review (and punish) without modeling firms explicitly. It suggests that the ease of monitoring the buyer and the size of his informational advantage are important channels in this context.

B.4.1 Setup

Consider a simultaneous game of complete information between a bureaucrat and a government. For a procurement tender, the bureaucrat has to decide whether to bias his decision towards a domestic supplier (cheat, C) or not (NC). I call this “cheat” for brevity, but it represents home bias as a form of misallocation in the widest sense such that not the procurement outcome is prioritized, but some other objective of the bureaucrat, say, because of regulatory capture.

If the bureaucrat plays C, this results in a low-quality product being delivered to the government from which the government obtains a payoff of $\underline{S}$, and the bureaucrat receives a private benefit, or bribe, $B > 0$. If the bureaucrat doesn't cheat, a high-quality product of value $\bar{S} > \underline{S}$ is delivered to the government and the bureaucrat receives a payoff normalized to 0. The government decides to either review the tender (R) or not (NR). If the government plays R, the government imposes a cost of k_G onto itself and k_B onto the bureaucrat. In case of review and if the bureaucrat plays C, the bureaucrat compensates the government for the foregone quality $\bar{S} - \underline{S} = \Delta S$, effectively increasing government payoff to $\bar{S}$ (before subtracting the review cost). This is summarized in Table 27 which depicts a 2x2 matrix in which the government is the row player and receives the first payoff listed in each cell and the bureaucrat is the column player and receives the second payoff listed in each cell.

Table 27: Review-and-collusion game between government and bureaucrat

		Bureaucrat	
		C	NC
Gov't	R	$\bar{S} - k_G$, $B - k_B - \Delta S$	$\bar{S} - k_G$, $-k_B$
	NR	$\underline{S}$, B	$\bar{S}$, 0

Assume that the following two conditions hold to ensure that the game is interesting: $\Delta S > k_G$ ensures that the government has an interest in playing R if the bureaucrat plays C. If k_G is too high, then the government will strictly prefer to play NR whatever the bureaucrat's strategy. The other condition is $\Delta S > B$, the difference in quality must be greater than the benefit, so the bureaucrat prefers NC when the government plays R. If this was not the case and C was the bureaucrat's dominant strategy, the government's best response would be R. These two conditions result in a well-known kind of monitoring game which has no Nash Equilibrium in pure strategies, but only in mixed strategies where the government plays R with some positive probability p and the bureaucrat plays C with some positive probability q.

These probabilities p, q are chosen such that the other player is indifferent between his strategies, so:

$$p(B - k_B - \Delta S) + (1 - p)B = p(-k_B) + (1 - p) \cdot 0 \tag{96}$$

$$q(\bar{S} - k_G) + (1 - q)(\bar{S} - k_G) = q\underline{S} + (1 - q)\bar{S} \tag{97}$$

After some simplification, this is straightforward to solve for

$$p = \frac{B}{\Delta S} \tag{98}$$

$$q = \frac{k_G}{\Delta S} \tag{99}$$

As for comparative statics, it is easy to see that the equilibrium probability p of the government playing R (reviewing a tender) increases with the size of the benefit B and decreases with the difference in the two quality levels ΔS. The equilibrium probability q of the bureaucrat playing C (favoring domestic suppliers) increases with the cost of review and enforcement k_G and decreases with the difference in the two quality levels ΔS.

B.4.2 Discussion of the results

The comparative statics for equation 99 are measurable within the empirical framework as was suggested in Section 2.4.5. The lifting of buyer discretion, which we analyze in Section 2.4, amounts to an increase in k_G. This model predicts misallocation to become more frequent when k_G increases. However, the effect found in the empirical analysis in the opposite direction. This suggests that in spite of the suspension of rules on buyer discretion, government review cost need not have increased. Perhaps the scrutiny of procurement of medical supplies was higher during the pandemic than before. The media reports cited in the introduction are an example of how procurement received unusual public attention.

ΔS stands for the difference between the high quality and the low quality product. Rather than just narrowly describing product quality in the sense of vertically differentiated products, this quality parameter S has been interpreted and described, e.g., by Laffont and Tirole (1991), as the fit of a seller with tender requirements, or more broadly as the value of the information of an agency. This motivates the focus on procurement of medical supplies. In the context of an emergency, such as the Covid-19 pandemic, the value of information on the actual quality of the seller of medical supplies is increased. In Section 2.3, I measure urgency through the local Covid-19 infection rate, but also consider alternative urgency measures. The empirical analysis confirms the model prediction that an increase in the value of information should decrease misallocation (and also lead to fewer reviews which is not observable, however).

What justifies the assumption that the quality differential is larger than the bribe? If this condition wasn't true, then collusion would be efficient in the model in the sense that it maximizes total surplus. The idea behind the second condition is that the bribing firm cannot offer a bribe that is greater than ΔS which is motivated by the notion of bilateral interim efficiency developed in Laffont and Tirole (1991). This idea was developed for the case where an agency can collude symmetrically with two firms and information about firm quality is verifiable. Then under a bilateral interim efficient auction, there are no incentive-compatible side transfers between the bureaucrat and any firm, and no announcement strategy by the agency or that firm, given the auction which is designed by the principal in the Laffont-Tirole model when taking truth-telling by the other firm as given.

Why is the bureaucrat capable of compensating for shortcomings in procurement tenders? While this question might arise of the agencies in the mechanism-design models, which are bureaucratic agencies that are merely paid to screen sellers on behalf of the principal in the spirit of Laffont and Tirole (1991), buyers in the empirical analysis typically purchase goods out of their own budgets rather than just running auctions for goods and services for which the government pays (for example, hospitals). The implications and comparative statics of the model will not change drastically when assuming imperfect recovery of foregone surplus (i.e., if the bureaucrat pays the government ΔS if caught cheating, but the government only receives a fraction $\lambda \in (0, 1)$ of that sum, q would increase to $\frac{k_g}{\lambda \Delta S}$).

C Appendix to Chapter 3

C.1 Tables

Table 28: Fields of the data set

Variable name	Description
Company name	Name of the company offering the repricer
Product name	Name of the repricer product
URL	Website of the company
Market places	Online market places that are compatible with the repricer
Country	Home country of the firm
Full Amazon integration	Repricer integrates with Amazon
Amazon marketplaces	List of compatible Amazon marketplaces
Tariff learns demand	Repricer learns demand or price elasticities
Uses AI	Repricer is advertised to use AI
Uses GT	Repricer is advertised to be designed using game theory
Uses econ	Repricer is advertised to be designed using economics
Uses self-learning	Repricer is advertised to be self-learning
Uses dynamic pricing	Repricer is advertised to use dynamic pricing
Uses price discrimination	Repricer is advertised to use price discrimination
B2B	Repricer is advertised for business-to-business/wholesale use
B2C	Repricer is advertised for business-to-customer/retail use
Facilitates switching	Firm advertises easy switching
Tariff name	Name of the repricing tariff
Tariff pricing	Mode of tariff pricing
Tariff loyalty	Tariff offers discounts for longer contract durations
Tariff shortest	Minimum contract duration
Tariff longest	Maximum listed contract duration for discounts
Tariff discount	Long-term discount relative to one month contract (or closest equivalent)
Tariff minimum	Is there a minimum cost to renting this tariff
Tariff minimum amount	What is the minimum amount of renting this tariff
Currency	Currency of prices
Tariff cost example small	Cost of hypothetical small vendor
Tariff cost example medium	Cost of hypothetical medium-sized vendor
Tariff cost example large	Cost of hypothetical large vendor
Tariff trial available	Free trial or test version of the repricer available
Tariff repricing per hour	Number of repricing events per hour
Tariff max products	Maximum number of repriced SKU
Tariff analytics yes or no	Repricer includes additional analytics
Tariff analytics notes	Self-description of analytics
Pricing consulting service	Cost of hypothetical small vendor
Tariff specific competitors	Tariff can target specific competitors
Tariff pricing rules	List of pricing rules that the repricer offers. Full list below.
Tariff buy box	Does the tariff target the Amazon Buy Box?
Tariff tracks rivals stock	Does the tariff keep track of rivals' stock and running out of stock?
Notes	Miscellaneous notes and quotes
Trial accessible	Free trial accessible without credit card information or existing online shop?

All data as of collection in October 2021. Pricing rules include optimizing for the Amazon BuyBox, competitors' prices, other suppliers' prices and stocks, combination of rules, demand factors, own stock, self-learning, time of day, price elasticity, decision tree, resetting to a maximum price, seasonality, weather, historical data, or unreported.

Table 29: Number of compatible repricing algorithms for different business software

Software	# of repricing algorithms	Software	# of repricing algorithms
Shopify	13	Microsoftdynamics	1
Magento	6	Advarics	1
Woocommerce	5	Prohandel	1
SAP	4	Roqqio	1
Bigcommerce	3	Höltl Retail Solutions	1
Prestashop	3	Intelligix	1
Salesforce	3	LS Retail	1
Shopware	2	Intersys	1
Skucloud	1	Hiltes Software GmbH	1
Lengow	1	3D Cart	1
Shoppingfeed	1	Wideorbit	1
Plentymarkets	1	Hudson MX	1
JTL	1	Marketron	1
Microtech	1	Rars2	1
Vario	1	Sqills	1
MS Dynamics AX	1	Tableau	1
MS Navision	1	NRS	1
Oxid	1	PriceFX	1
Oracle	1		

C.2 Figures

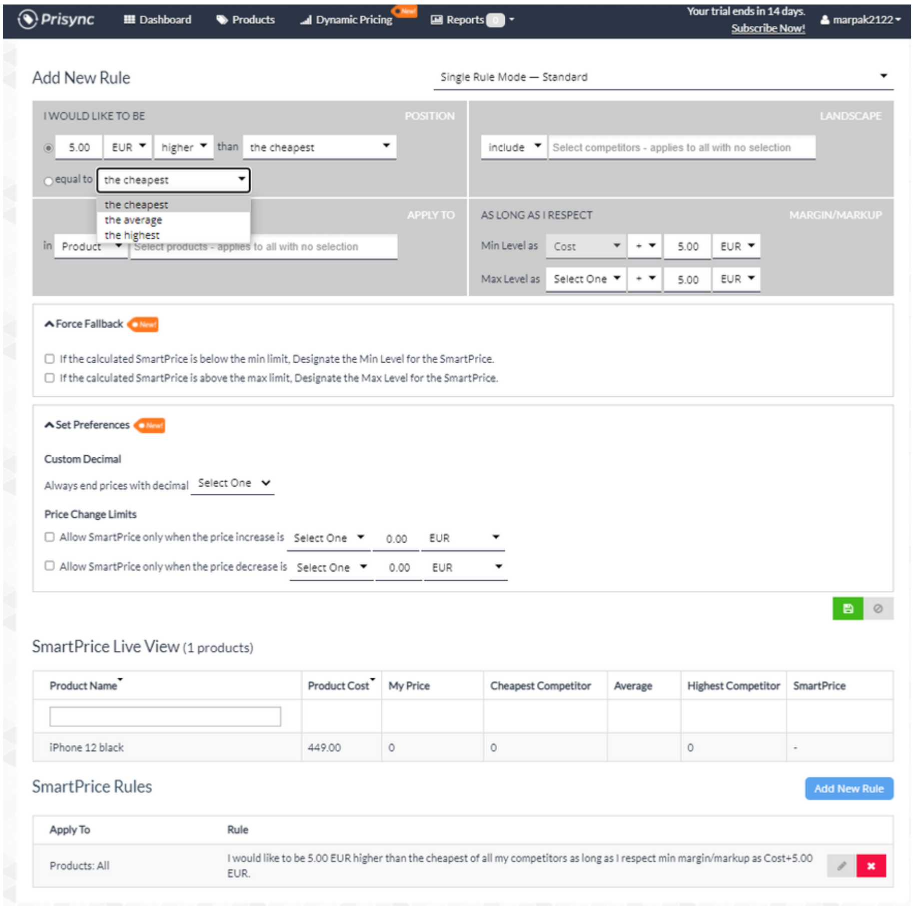


Figure 25: Prisync pricing rules, taken from https://prisync.com/

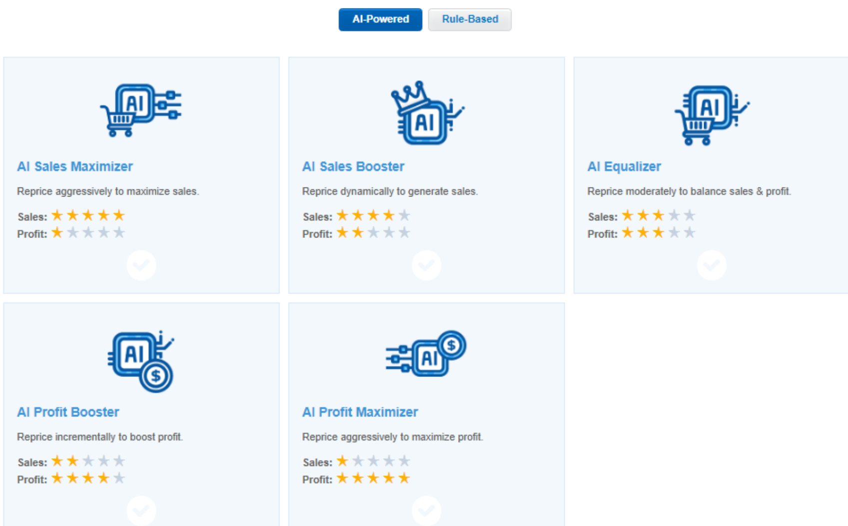


Figure 26: BQool pricing rules, taken from https://support.bqool.com