Contractual Data and Market Power

Anil Savio Kavuri and Alistair Milne*

30th December 2023

Abstract

A growing literature on the economics of data has focused on ownership and exchange of consumer data. This paper provides an industrial organisation perspective, discussing contractual data as an input to production of goods and services whenever supply involves either changes in asset ownership (property rights) or agreed coordinated performance across multiple businesses (contractual rights). This 'wholesale' data appears in many contexts, both financial (payments, securities clearing and settlement, trade finance) and non-financial (logistics and supply chains, registers of land and vehicle ownership, flight bookings and patent registration). Arrangements for control of this wholesale business data affects competition and market power, captured here in a model of vertical separation, with an 'upstream' market with connection (or switching costs) for data control and a 'downstream' market with a Salop Circle capturing imperfect competition in the supply of goods and services. This highlights the role of upstream 'data validators', the economic importance of data governance and data control in these wholesale data markets, and the potential for reducing connection and switching costs, and hence for promoting greater competition and improved economic outcomes, through standardisation and overcoming data fragmentation. [190 words]

Keywords: Asset Ownership, Blockchain, Clearing and Settlement, Contracts, Contractual Data,Connection Costs, Data Control, Data Validators, Economics of Contract, Property Ownership, Social Optimum, Supply Chains, Switching costs, Salop Circle, Financial Assets, International Payments, International Logistics, Retail Payments, Vertical Separation

JEL Codes: D43, L13, L14, L15, L22

*Kavuri: Research Associate at Loughborough University and at Australian National University and Honorary Associate at University of Cambridge. a.s.kavuri@lboro.ac.uk; Milne: Professor at Loughborough University. a.k.l.milne@lboro.ac.uk. We thank Antonio Russo, Luke Garrod and Christopher Wilson and an audience at the 2023 Markets, Contracts and Organizations conference at Australian National University, Canberra, for valuable comments.

1 Introduction

The continuing rise of online and digital economic transactions has motivated a growing literature on the economics of data. The issues addressed in this literature have included: property rights (Posner and Weyl 2019, Jones and Tonetti 2020, Acemoglu et al. 2019); privacy (Varian 2018, Acemoglu et al. 2019) and market failure (OECD 2016, Morton-Scott et al. 2019, Farboodi et al. 2019, Veldkamp and Chung 2019, Acemoglu et al. 2019, Jones and Tonetti 2020). This literature is though imbalanced , largely focused on consumer-to-business (C2B) and consumer-to-government (C2G) data exchange; i.e. what data in retail goods and services. It has paid relatively little attention to business-to-business (B2B) data exchange in the supply chains, services and infrastructures supporting the provision of final goods and services.

This paper explores a hitherto little examined issue in B2B data exchange: how do the arrangements for the writing/updating of 'wholesale' data – i.e. shared data used as inputs to the production of contract-based outputs – affect competition and market power? The contribution is exploring how the institutional and technical arrangements for operations with this wholesale data impact on the level of competition and efficiency of pricing. This highlights under-researched issues in the economics literature: what are the arrangements for validation and control of ownership and contract related data? how do these affect competition and market power? and to what extent are and should businesses be competing in the provision of data based services and collaborating on data standards and mechanisms of data exchange?

There are commonalities between B2B and C2B data exchange. What are the responsibilities for protecting data relating to business counterparties? To what extent should businesses be able or required to pass on data to other businesses. How are the rights over that data to be protected? These other data rights issues are put to one side, in order to focus on the implications of arrangements for validation and control of contractual and ownership data for market power.

The analysis is developed as follows. Section 2 is conceptual, presenting the key economic intuition. Contractual data here refers to both the digital representation of contracts subsequent contract related data. The processing of this contractual is an input to the provision of many business-to-business and business-to-consumer services. This in turn suggests that there can be an economically relevant vertical separation between the validation of contractual data and the provision of goods and services using this data; that there is a parallel with other vertically separated markets such as those for household utilities; and that effective competition can be promoted by interventions focused on reducing barriers to the validation and processing of data in ownership records and contractual commitments.

Sections 3 and 4 then present some formal modelling, comparing market outcomes for the validation and processing of contractual data with that chosen by a social planner. The upstream market is for data validation, providing access to and updating of underlying contractual data. Since computerised data operations are typified by large fixed costs and low variable cost, this modelling assumes that the only economically relevant upstream costs are connection and switching costs: the connection costs of first accessing data validation services and the switching costs of moving from one data validator to another. Imperfectly competitive provision in the downstream services employing contractual data as an input is captured through location of downstream firms and consumers on a Salop circle.

The simplest version of this modelling (the baseline of Section 3) assumes vertical separation with a choice of upstream data validators. The magnitude of the fixed cost of connection and switching determines the number of providers of downstream services – the lower the cost then the more firms are able to profitably enter the downstream market and the less their market power. This baseline inherits the limitations of the standard Salop circular city model, assuming price inelastic consumption of the downstream good and the resulting possibility of excess entry as a market outcome with excessive resources devoted to connection relative to the consequent reduction of transport costs.

Variations of this model (developed in Section 4) allow for price elastic consumption and take account of the varying institutional arrangements in different data contract based industries. Downstream firms may own data validators. A special case of is where there is only one data validator. This may be owned by one or several downstream firms. In this situation switching costs are no longer relevant, but the joint ownership giving of the data validator is then a barrier to downstream entry. A further possibility is a separated monopolicy data validator, either a for profit entity, a not-for profit or a state or local government owned entity.

The concluding Sections 5 and 6 discuss contractual data and market power, in the light of this economic intuition and formal modelling, across a number of economically significant financial and non-financial markets. Section 5 considers those markets, which include most financial services, with existing market-wide institutional arrangements for data validation. These examples are primarily financial (asset management, capital market operations, payments) but there are also important non-financial (property ownership and transfers, flight bookings and patent registration). As well as the two vertical layers ('upstream'/ 'downstream'), there are sometimes additional 'midstream' services linking upstream and downstream. The upstream services of data validation and control can be provided by specialised firms operating independently of the downstream customer facing firms. In other cases downstream firms are responsible for upstream data validation and control.

Section 6 discusses the potential gains of developing market wide arrangements for the validation of contractual data in supply chains from a transaction costs perspective. Here arrangements for data validation are typically fragmented, developed on a relationship specific basis for bilateral transactions. Arguably, adoption of market-wide institutional arrangements for data validation and control could substantially lower costs of production and, also, through greater transparency support access to complementary third party financing and insurance, facilitate supply chain resilience and help with the traceability of inputs and the promotion of ESG goals. However the analysis of contractual data and market power develop here indicates that the required change in economic organisation will be a substantial strategic and policy challenge: requiring industry wide consensus or regulatory intervention to overcome the challenges of co-ordination and resistance from incumbents whose market power is threatened by standardised data validation.

While none of this analysis empirically quantifies the economic impact of arrangements for wholesale data control on competition and market power, some general qualitative conclusions can be drawn. Vertical separation, with competition in the provision of data validation services, is favourable for encouraging entry and competition in downstream provision. Vertical integration of data validation does not necessarily prevent downstream competition, but avoiding excessive downstream market power then requires that the upstream market for data control is open to new entry, not restricted to incumbent firms.

Whether data validation is vertically separated or vertically integrated, data standardisation is of central economic importance: both directly reducing costs and lowering barriers to entry. This also highlights the challenge of achieving appropriate economic governance of collective arrangements in markets characterised by reliance on contractual data. The development and adoption of common data standards, in order to reduce costs of data validation and control, may not be embraced by incumbents both because of the difficulties of achieving a consensus on co-ordinated change and also resistence from downstream firms who lose market power as a consequence of upstream standardisation.

This analysis of the economic role of data validation also suggests reasons for scepticism about the currently widespread claims for economic benefits from the wideapread adoption of the 'blockchains' (or distributed ledgers). These technologies have attracted attention from their application in cryptocurrency transactions and decentralised finance, where they are employed on a 'permissionless' basis, validating data without requiring any institutional involvement. This is in effect a maximal form of standardisation with upstream data record keeping entirely automated and data validation conducted by the consumers of downstream services. Concerns about market power are limited to the possibility of large better informed market participants ('whales') being able to manipulate market prices to their advantage. It is unclear that such an approach to data validation is more widely applicable outside of this institution free activities. The need to validate data means that applications of these technologies to other goods and services must be on a 'permissioned' basis with blockchain data jointly controlled by a group of institutional data validators. This institutional control in turn means that, even when these new technologies offers substantial efficiency gains in the use of contractual data as an input to production, concern remains about the market power of data validators and the economic incentives for adoption by incumbent firms.

All this leaves many questions for further investigation:

- The economics of data remains an immature sub-field of economic scholarship. A basic but as yet unfinished task is establishing a consensus characterisation of the economic role of the recording, exchange and use of data. See Carrière-Swallow and Haksar (2019) for a review of the existing literature. They distinguish the economic functions of data collection and data processing, discuss the economics of data privacy, and highlight the critical role data access and of control of data access; but they do not discuss data validation.

- The vertical separation explored here parallels that extensively discussed in the economics of utilities regulation, where ownership of and access to 'upstream' infrastructure is a substantial competition concern (e.g., Armstrong and Sappington 2006, Armstrong, Cowan and Vickers 1994). If upstream data validation and control is indeed a major concern for market power, then what are the implications for competition law and market interventions by competition authorities?

- Switching costs (Farrell and Klemperer 2007) result in customer lock-in, discouraging customers from changing suppliers and giving vendors ex-post market power. A range of work on vertical separation addresses imperfect competition in upstream and downstream markets (e.g.,Buehler and Schmutzler (2008), Reisinger and Schnitzer (2012) and Ordover, Saloner and Salop (1990)). Detailed empirical investigation to quantify the impact of arrangements for data validation and control on individual markets is warranted, in those markets discussed here and others.

- The analysis also relates to the body of research on two-sided platforms (e.g. Rochet and Tirole 2003, Gilbert 2021), in that platform services are often based on data access and arrangements for data validation and control. To what extent the issues of data standardisation and governance highlighted a determinant of platform access and the extent of multi-homing? Platforms are also data validators, so to what extent might data validation and control be separated from the search and communication services? Could amended arrangements for data validation and control address concerns about platform market power?

- Contractual arrangements are central to the literature on corporate organisation and contract design (Grossman and Hart 1986; Hart and Moore 1990). The analysis of Sections 2-5 assumes existing institutional arrangements. If institutional arrangements and the recording and processing of contractual data can be restructured, then how should this be done to make best use of technologies for validation and control of contractual data, so as to reduce transaction costs and promote the efficient resolution of contractual default? What are the implications for the economic role of the firm and for the economically efficient organisation of markets?

- There has been little work on the industrial organisation of permissioned blockchains, critical to understanding the application of blockchain technologies outside of crypto-currency transactions and decentralised finance. Our literature search uncovered only one such paper, Iyengar et al. (2023), modelling the possibility of adoption failure of permissioned blockchain because producers are unable to gain competitive advantage from adoption, even though consumers benefit. There is a need for further theoretical and empirical research on data validation using permissioned blockchains and its implications for both market power and incentives for adoption.

- Data technologies , including AI and 'big data', are central to current policy discussions of innovation and productivity (see e.g. Babina et al. (2022),Beraja, Yang and Yuchtman (2023)). What is the relevance of data validation and control to the relationship between competition and innovation in data technologies? Is the argument that emerges here for vertical separation and standardisation of this data, in order to promote downstream competition, reinforced by a resulting strengthening of incentives for innovation? Or is there a trade-off between upstream standardisation promoting downstream competition and downstream innovation?

2 Intuition

The 'wholesale' data investigated in this paper emerges whenever any form of economic output (where the customers could be businesses or households) that depend on first creating agreements (contracts), whether for exchange of assets or for supply of goods or services, and then on co-operative fulfilment of those contracts by two or more firms. Examples of such contract-based outputs include: retail and wholesale payments (including card and bank payments, both domestically and internationally); securities trading and settlement; trade finance; logistics; construction supply chains; property ownership registration and transfer; and flight bookings. The common feature is that ownership and contractual obligations are or can be represented in shared databases which are then used as input to production of goods and services. This in turn requires the services of data validators to validate and update data records, the operations that we refer to as 'wholesale' data control.

Exploration of some of these examples illustrates this concept of wholesale data and the economic role of data validators:

- An familiar example is house purchase. House purchase and sale requires access to and updating of records of title to land and property. Historically these records were either paper documents (deeds of title) or a written entry in a register. Nowadays the title is represented as a record in a computer database, either a digital registry in which an entry itself represents legal title, or a digital system for confirming the validity of paper title documents. The 'data validators' are lawyers, working for the house purchaser and the house seller, responsible for ensuring that the transfer of title is legally valid, that the required exchange of money to complete the transaction takes place as agreed, and after this for updating paper and digital records. Their economic role is not placing new entries in the database (that final task can be carried out by a separate data base operator our through automated software acting on a request by the data validators); it is ensuring that the transaction is legally valid and completed according to the contract of purchase and instructing the consequent changes in recorded title.

- A second example is that of bank or card payments. Here there are two providers, a bank or a non-bank account providers, who must operate in conjunction to successfully deliver the payment service, i.e. the transfer of funds from an account held with one provider to an account held with another. These account providers are the data validators: they must ensure that there is a sufficient account balance or line of credit for the payer (the sender of funds) to make the payment; apply checks to satisfy themselves that the payment instruction is genuine and that there is an identified account for the payer to receive funds; and finally settle the payment with a matching transfer of a settlement asset, nowadays almost always central bank reserves, between the two account providers. There are thus three different sets of data records that must be updated: those of the sending institution, those of the receiving institution and finally those of the central bank providing the reserve accounts for settlement.

- A third somewhat different example is that of supply chains, particularly the case of small companies supplying larger corporate clients e.g.: consumer products for large retailers such as Walmart or Carrefour; components to large manufacturing companies such as the global car manufacturers Ford or Toyota; or sub-contractors to a large construction company such as Bechtel or Balfour Beatty. Unlike the case of the house purchase or bank payment transfers there is usually no standardised system for maintaining the associated contractual data records. The commercial arrangements are bilateral and conducted using record systems established by the purchaser. Contractual data control and data validation is though economically relevant because of the involvement of other parties in the supply chain: for financing production and then for transporting, warehousing and delivery. The absence of validated data on the bilateral supply contract and its delivery interferes with the supply of these supporting services. The consequence is inefficient use of data in the process of production and supply. The creation of wholesale markets for data with data validation can potentially address this issue but this development may be blocked by coordination and incentive problems.

There are many different possibilities for the economic organisation of such 'wholesale' data control. One possibility – captured in the baseline model of the following section – is of competition between validators to provide access to wholesale data control, allowing entry to the downstream market but entry limited by the costs of accessing these data validation services. Another possibility – raising concerns about heightened barriers to entry the downstream market – is for the wholesale data control to be provided by a single operator owned by the incumbent downstream firms. Another possibilty is wholesale data control to be offered by a single operator operating on a commercial basis – with the potential consequence of downstream monopoly. Yet another possibility is an operator owned by one or a subset of downstream firms.

The perspective of this paper also provides insight on the large volume of recent work claiming transformative commercial benefits from the adoption of blockchain (or distributed ledger) technologies underpinning cryptocurrencies such as Bitcoin and also so called 'decentralised' finance. The areas of application explored here - in both financial transactions and supply chains - are largely the same as the proposed business and economic applications of the blockchains. The original blockchains were *permissionless* meaning that they operate without institutional involvement and employ open source software available to anyone with internet access. Permissionless blockchain can be interpreted as the ultimate standardisation of wholesale data records, removing entirely the need for the data validators, highlighted in our analysis, or other intermediaries and hence fully removing barriers to entry in downstream markets which rely on data as an input to production.

The achievement of such blockchain based standardisation is however limited by the need for data validation. Validation is trivial in the special case of permissionless blockchains because all that is required to initiate a valid transaction is possession of the relevant private cryptographic keys. Validation is no longer trivial when data is used as an input to production, including all the markets discussed in this paper. Decentralised finance has demonstrated the possibility of using software based 'oracles', replacing the role of data validators, capturing and validating internet data on for example weather events or movements in financial market prices. Such publicly available data is though only a small subset of all the contractual data used as inputs in production. So, while standardisation can reduce connection costs and promote downstream competition, in most contexts institutional data validation will still be necessary. This in turn rules out the use of 'permissionless' blockchains requiring instead more conventional 'permissioned' data architectures.

3 A baseline model

This section presents a baseline model of a vertically separated market for data validation and control. Firms providing consumer services in a downstream market require access to the upstream market for the updating and control of data. Since data operations are largely based on fixed costs of establishing systems for data processing with low or negligible variable costs, we characterise the upstream market as one with switching or connection costs but otherwise with zero marginal costs. The downstream services market is imperfectly competitive, with consumers preferences amongst different suppliers characterised by location around a circular product space. For simplicity, we assume that all production and consumption takes place in a single period $t = 1$.

3.1 Assumptions

Data operators, indexed by $j = [1, 2 \cdots J]$, compete in the upstream market; and N firms, indexed by $n = [1 \cdots N]$, are in imperfect competition in the downstream market. We capture this imperfect competition by employing the standard Salop circular city model, assuming that the competitive position of each firm is represented by its location on a circle. The assumptions are as follows:

1. There are multiple data validators who operate without cost. In the baseline model we assume these operate as profit maximiser, each setting a price for connection consistent with the price set by other data validators. It then turns out that the number of operators (as long as there is more than 1) is immaterial.

2. Initially, before decisions on entry or exit are made, there are N_0 downstream firms connected to these data validators and an unlimited fringe of potentially new firms who can pay connection costs of S to connect to an existing operator. Firms can also switch data validators, paying the same cost S as a switching cost to move from their current operator to a new operator. We adopt the convention that the set of downstream firms given by $n \in [1 \cdots \tilde{n}]$ is initially connected to data validator $j = 1.$[1].

3. Subsequently, after decisions on entry or exit are made, the equilibrium number of downstream firms N may higher, lower or equal to N_0. As we discuss, the outcome that maximises social welfare depends on N_0, since entry $N > N_0$ involves additional new costs, whereas for exit $N < N_0$ the past costs of connection are sunk costs that can be disregarded when assessing social welfare.

4. Consumer preferences are captured by the location of consumers, with a total mass of 1, evenly distributed around a circle, at positions indexed by $\theta = [0 \cdots 1]$.[2]

5. Downstream firms have identical costs and so locate themselves at equidistantly around the circle. Firm n is located on this circle at position $\theta_n \in [0, 1]$ given by:

$$\theta_n = \frac{n-1}{N}$$

6. Transport costs, paid by the consumers for purchasing from a firm n located at θ_n, depend on a parameter t and are a linear function of the distance moving around the diameter of the circle between the location of the consumer and the location of the firm $\theta - \frac{n-1}{N}$. The consumers purchase one unit from their preferred firm n at a price P_n. So total costs, price and transport, for purchase from n are given by

$$c(\theta) = P_n + t|\theta - \frac{n-1}{N}|$$

7. The profits of downstream and upstream firms, connection costs and data access fees and the consumer expenditure on transport and on purchase of downstream goods, are all measured relative to an outside good of unit price. Consumers have lexicographic preferences: they must purchase one unit of the downstream good to obtain a non-zero utility and thereafter utility is increasing in the consumption of the outside good.

The time line for firm decisions, shown below in Figure 1, is as follows. First in a pre-production period $t = 0$ each data validators sets the two prices it will charge in t for use of its services: one for existing customers (firms already connected to the operator who have no connection costs); the other new customers joining it either as new entrants to the downstream market or switching from other data validators. At the beginning of the operating period $t = 1$ the downstream firms decide

1. since the data validators costs are identical there is no loss of generality in ordering the data validators in this way

2. Geometrically, in Cartesian co-ordinates, we may assume that a position on the circle is represented by the point ($x = cos(2\pi\theta)$,$y = -sin(2\pi\theta)$) and so consumer θ is located on the circle at angle $2\pi\theta$ measured clockwise relative the positive horizontal axis.

- on whether to enter the market (if they are initially outside) or exit the market (if they are initially inside)
- if they enter or stay in the market, which data validator they will then use for data control; and
- their location on the Salop circle characterising the downstream market and the price they charge to consumers.

Production and consumption then take place and profits are earned during period $t = 1$.

The environment is illustrated in Figure 2. The figure indicates the positions of firm n and firm $n+1$. Consumer θ is located between the two firms and chooses one to use for their services.

Figure 1: Timeline

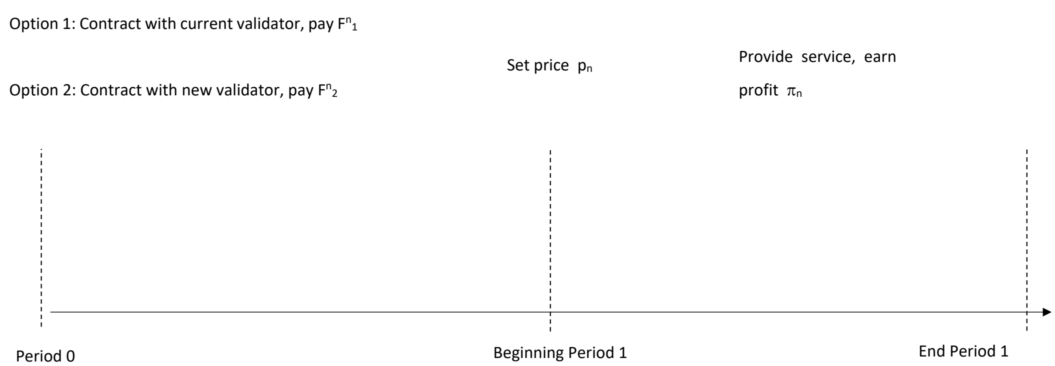
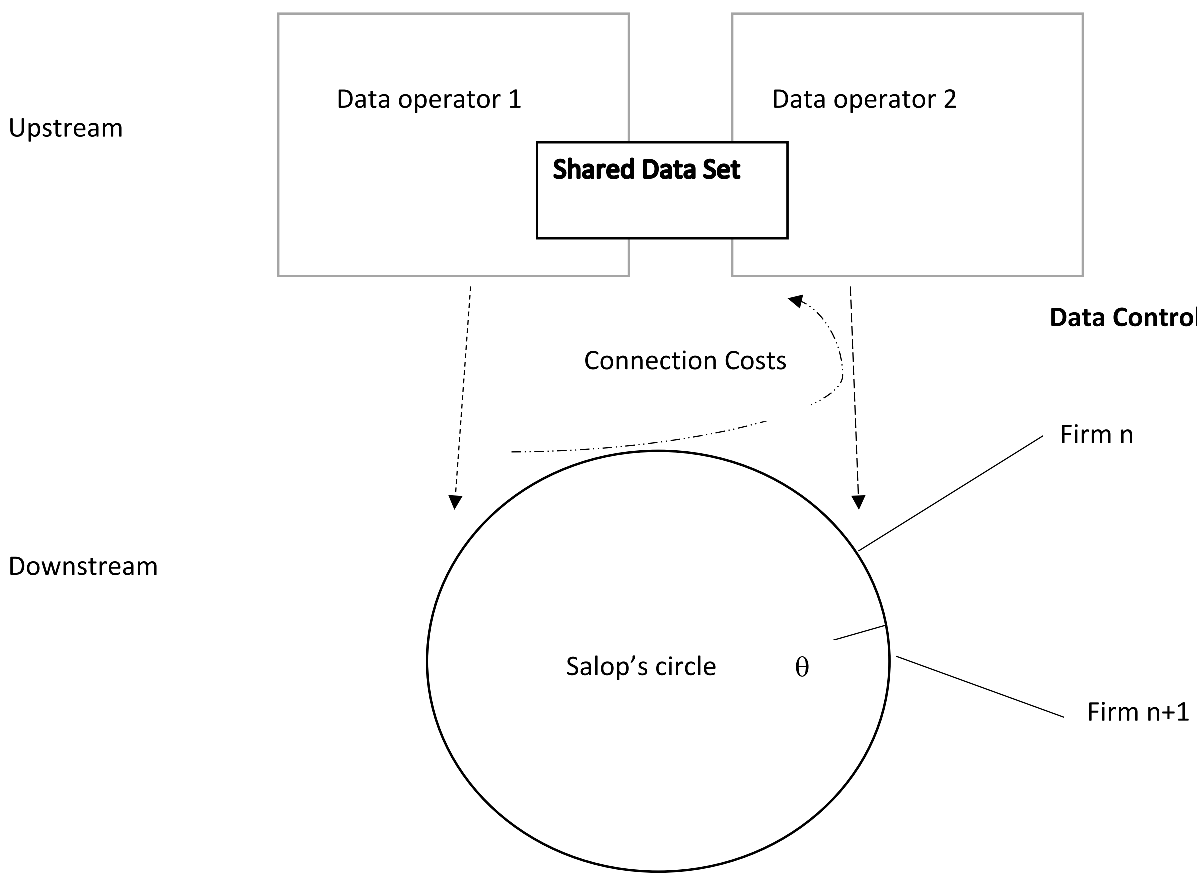


Figure 2: Data Validation



3.2 Baseline solution

Solution of this baseline model is stated in the following three propositions (proofs of all propositions are provided in the appendix):

Proposition 1. *Let the fee charged for services by a data validator to an already connected downstream firm (in the case data validator $j = 1$ this to firms $n \leq \tilde{n}$, but the ordering is not material) be F^n and the fee charged to a new previously unconnected firm be $\bar{F}^n$ In equilibrium the following holds:*

$$F^n = S$$
$$\bar{F}^n = 0$$

This result holds for any number $J \geq 2$ of data validators in the upstream market. It arises because any existing data validator can capture value of S from existing connected customers (so operator $j = 1$ obtains revenue and profits of $\tilde{n} \times S$), but to capture a new customer (either a new entrant to the downstream market or a firm switching from another data validator) must charge a zero price. As a result all firms those initially connected with an upstream firm and those entering to connect for the first time, pay the same cost S for data access.

Proposition 2. *With free entry and exit:*

- *N is a function of t and S, and can be expressed as*

$$\bar{N}\left(\frac{t}{S}\right) = \max 1, \sup\left\{z \in \mathbb{Z} : z \leq \sqrt{\frac{t}{S}}\right\}$$

- *If there is downstream competition, i.e. $\bar{N} > 1$, all downstream firms charge the same price (P) and earn the same profit (π) given by:*

$$P = \frac{t}{\bar{N}}$$
$$\pi = \frac{t}{\bar{N}^2} - S$$

but if there is a downstream monopoly, i.e. $\bar{N} = 1$, then the price charged by and profit of the downstream monopoly are unbounded.

The price charged is the same for all downstream firms and depends on transport costs t and the total number of firms N in the market. Profit of each firm then further depends on market share ($\frac{1}{N}$) and the costs of data access S. This then determines the number of firms in the market, as the given largest integer, because firm $n = N + 1$ cannot profitably either enter or remain in the market. Doing so reduces revenues for all firms below the cost of data access S.

3.3 Social planner's solution

This subsection examines the departure between the baseline equilibrium and optimal outcome determined by a social planner that intervenes in market outcomes with the objective of maximising the sum of consumer and producer surplus. The use of the standard 'circular city' to model downstream imperfect competition means that consumption levels are exogenous. The only decision of the social planner is the determination of the number of competing firms N based on the trade off between connection costs (increased by new entry) and transport costs (reduced by new entry). The optimal choice is stated in the following proposition:

Proposition 3. *In the model of this section the social welfare, the sum of consumer and producer surplus, is given by:*

$$\Omega(N) = -S\max(0, N - N_0) - \frac{t}{4N} \tag{1}$$

and the socially optimal number of firms is an integer N^*, that depends upon N_0, t and S. Let $\tilde{N}$ be the largest integer such that $\tilde{N} \leq \frac{1}{2}\sqrt{\frac{t}{S}}$ and so can be expressed as:

$$\tilde{N}\left(\frac{t}{S}\right) = \sup\left\{z \in \mathbb{Z} : z \leq \frac{1}{2}\sqrt{\frac{t}{S}}\right\}$$

Then N^* is given by:

$$N^* = \begin{cases} N_0 & \text{if } N_0 \geq \frac{1}{2}\sqrt{\frac{t}{S}} \\ \tilde{N} & \text{if } N_0 < \frac{1}{2}\sqrt{\frac{t}{S}} \text{ and } \frac{1}{2}\sqrt{\frac{t}{S}} - \tilde{N} \leq \frac{1}{2}\left(1 + \sqrt{\frac{t}{S}} - \sqrt{1 + \frac{t}{S}}\right) \\ \tilde{N} + 1 & \text{otherwise} \end{cases}$$

Social welfare depends on the sum of total connection costs for new firms entering the downstream market ($N_0 - N$) and of transport costs paid by consumers, but it is independent of both downstream prices (with output fixed these are simply a transfer between consumers and producers) and of connection costs paid by established firms (a transfer between downstream and upstream firms). Exit of firms from the downstream market is always socially costly, because the sunk cost of entry S is already spent, and their exit increases the burden of transport costs on consumers downstream; hence $N \geq N_0$; but new entry is only socially worthwhile when the resulting reduction in transport costs is larger than the fixed costs $S(N - N_0)$ of this entry.The solution for N however must be an integer, one of the two nearest neighbour integers to $N = \frac{1}{2}\sqrt{\frac{t}{S}}$ otherwise the smallest integer greater than $\hat{N}$.

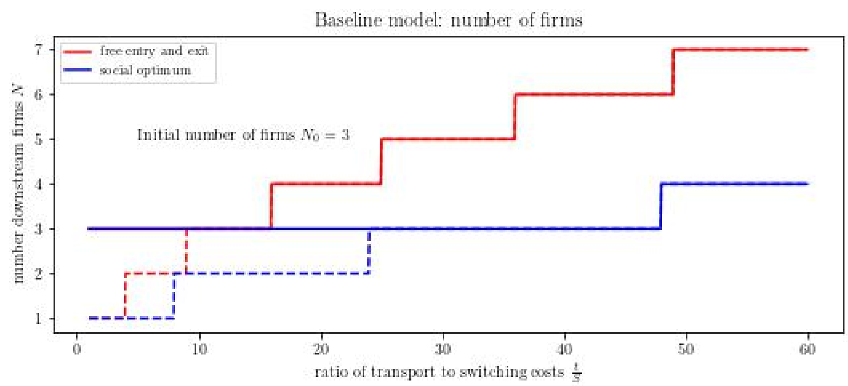
Comparing **Proposition 3** with **Proposition 2** highlights a standard result, the emergence of 'excess entry'.[3] Excess entry arises in our baseline model with fixed output in the downstream market, whenever N_0 is relatively low, because the profitability of a firm entering the downstream market is inversely proportional to N^2 (prices ($\frac{t}{N}$) $\times$ market share ($\frac{1}{N}$)) less the costs of connection; whereas social welfare is inversely proportional to N (total

3. a possibility in the circular city model first noted in the Vickrey (1964) textbook discussion of spatial models of competition, section Nine pgs. 323-334. This the so called excess entry theorem; conditions for which are explored in inter alia (Salop (1979), Matsumura and Okamura (2006a, 2006b) and Gu and Wenzel (2009))

transport costs $\frac{t}{4N}$). The relatively lower sensitivity of total transport costs to Nmeans that the number of firms achieving optimal social welfare is less than that supported by the zero-profit condition determining entry. The integer adjustment does not substantially change this outcome.

Figure 3 illustrates how the number of firms, both those determined by free entry and by the social planner, vary with $\frac{t}{S}$.

Figure 3: Number of Firms

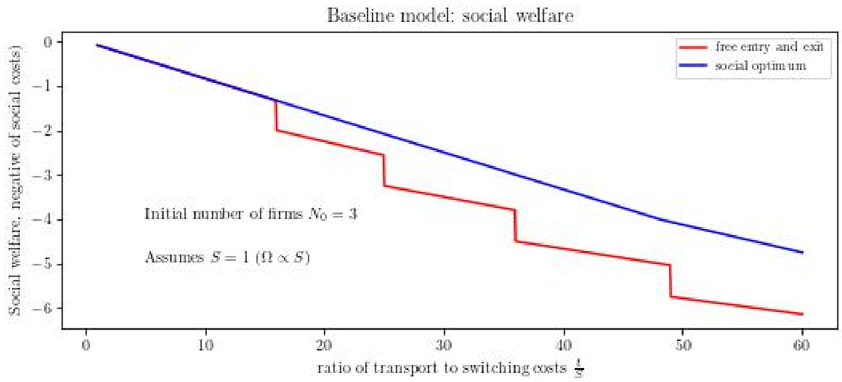


Ignoring the need for N to take integer values and assuming $N_0 = 0$ i.e. ignoring the possibility of firms being forced out of the downstream market, allows a simple welfare comparison between the social planner's solution (when $N = N^s = \frac{1}{2}\sqrt{\frac{t}{S}}$) and the market outcome (when $N = N^m = \sqrt{\frac{t}{S}}$). In this case the gain of social welfare from the social optimum compared to the market outcome, is given by:

$$\begin{aligned}
\text{welfare loss} &= -SN^s - \frac{t}{4N^s} + SN^m + \frac{t}{4N^m} \\
&= S(N^m - N^s) + \frac{1}{4}(\frac{t}{N^m} - \frac{t}{N^s}) \\
&= S\frac{1}{2}\sqrt{\frac{t}{S}} + \frac{1}{4}(\sqrt{tS} - 2\sqrt{tS}) \\
&= \frac{1}{2}\sqrt{tS} - \frac{1}{4}\sqrt{tS} \\
&= \frac{1}{4}\sqrt{tS}
\end{aligned}$$

This is an illustration of the established finding of "excess entry" in the standard Salop circular city model: the higher costs of market determined entry exceeding the consequent reduction in transport costs. Allowing for the integer restriction on N and for firms already in the downstream market modifies but does not entirely overturn this finding. This is illustrated by Figure 4.

Figure 4: Number of Firms



This figure assumes the normalisation $S = 1$ (for given $\frac{t}{S}$, the social welfare is proportion al to S). Social welfare is the negative of entry and transport costs and so always less than zero. When when $N_0 = 0$ (no firms initially in the downstream market) then excess entry arises for most values of $\frac{t}{S}$. When $N_0 > \frac{1}{2}\sqrt{\frac{t}{S}}$ then excess entry arises only for larger values of $\sqrt{\frac{t}{S}}$.

The welfare comparison is more substantially changed when allowing for price elasticity of demand, as in the following section, where the opportunity to limit entry and exploit market power in the downstream market strengthens the market incentives for reducing the number of firms in the downstream market.

4 Ownership and governance

This section develops a number of extensions of the baseline model, in order to explore the implications of some of the many possible arrangements for ownership and governance of upstream data validators. We begin first by relaxing the standard assumption of the circular city model, as employed in our baseline model, that demand is price inelastic. Having relaxed this assumption, we explore the difference this makes to the social optimum, comparing to the outcome in the baseline model. We then vary the model to capture variety of possible arrangements for ownership and governance of the upstream data validators: a profit seeking data control monopoly, vertical integration of a data controller with a downstream firm, collective ownership of the data controller by the downstream firms.

4.1 Price elasticity of demand

To introduce price elasticity we follow Gu and Wenzel (2012) by assuming that consumption of downstream goods is a function of the price at which they are sold but independent of the distance between the firm supplying the good and consumer and the consequent transport costs.[4]

4. Microfoundations for this assumption can be provided by assuming that the transport costs are paid through foregoing consumption of the alternative numeraire good that is in fully elastic supply with utility a separable function of the two goods, non-linear in the downstream consumption and linear in the numeraire.

We adopt the convention of using upper case (P, Q) to represent the absolute value of prices and quantities and lower case to represent the natural logarithm of prices and quantities (($p = \ln P$, $q = \ln Q$ with the inverse functions ($P = \exp p$, $Q = \exp q$). This convention is convenient for exploring specific demand functions such as $q = \alpha_0 - \alpha_1 p$ (constant elasticity of demand) or $q = \max\left(\alpha_0 - \frac{\alpha_1{}^2}{4\alpha_2}, \alpha_0 - \alpha_1 p + \alpha_2 p^2\right)$ (elasticity of demand falling to minimum of 0 as price rises towards a value of $p = \frac{\alpha_1}{\alpha_2}$).

Assumption 1. *Consumption of the good provided by firm n is represented by the function $Q_n = Q(P_n) \geq 0$ of the firm's price, with $Q'() < 0$ and with a normalisation of the units of consumption and of price so that $Q(1) = 1$ (the equivalent of the normalisation in the baseline, where demand is completely inelastic, of $Q = 1$).*

As long as the equilibrium is symmetric the n subscript is not needed for statement of the equilibrium outcome and we can then refer simply to $Q = Q(P)$ or, restated in natural logarithms, $q = q(p)$. The price elasticity of demand is a function of P and can be expressed as:

$$\epsilon(P) = -\frac{Q^{'}(P)P}{Q} = -q^{'}(p) > 0 \tag{2}$$

On occasion we make the following further assumption:

Assumption 2. *$\epsilon'(P) > 0$ and there is a level of price $P = \bar{P}$ for which $\epsilon(\bar{P}) = 1$.*

i.e. the price elasticity of demand is monotonically increasing with price; with $\epsilon(0) < 1$ and with $\epsilon(P) \lesseqgtr 1 \iff P \lesseqgtr \bar{P}$ for some $\bar{P}$. The role of this additional assumption is to ensure the existence of a unique equilibrium.

4.2 Social planner's solution

As with the baseline a social planner chooses the number of downstream firms with access to upstream data control services; now however the social planner also chooses the price at which they sell to consumers, again in order to maximise the sum of producer and consumer surplus. Because of the assumed separability of consumption and transport, these two decisions are independent. Thus proposition 3. continues to apply and the socially optimal outcome, taking account of both decisions, can be summarised in the following proposition:

Proposition 4. *The socially optimal price under the assumptions of this section is $P = P^* = 0$. While the socially optimal number of firms is given by N^* as stated above in Proposition 3. This yields a total social surplus, the sum of consumer and producer surplus, of:*

$$\Omega(N^*) = \int_{\phi=0}^{1} Q(\phi)\mathrm{d}\phi - S\max\{N^* - N_0, 0\} - \frac{t}{4N^{*2}}$$

Because marginal costs of production are zero, the socially efficient price is also zero. The social surplus now takes acccount of the surplus then achieved by consumers from consumption (the integral of consumption, using the variable of integration ϕ, from $P = 0$ ($p = -\infty$) to an arbitrary normalisation reference price of $P = 1$ ($p = 0$). This is in addition to transport and connection costs which are unchanged from the baseline model.

See Gu and Wenzel (2012) for details and discussion.

4.3 Market solution with elastic demand

We now consider the market solution with elastic demand, employing the further **Assumption 2.** Instead of proposition 2 we now have:

Proposition 5. *Assuming that individual consumption of the preferred downstream good is represented by $Q = Q(P)$ with $Q'(P) < 0$, that $N_0 \geq 1$, that the elasticity of demand $\epsilon(P)$ is characterised by Assumption 2, with a unique price $\bar{P}$ for which $\epsilon(P) = 1$, and $S \leq \bar{P}Q(\bar{P})$ then:*

- *there is a unique pair of rational numbers $P = \tilde{P}$ and $N = \tilde{N}$ (N is not necessarily integer) that jointly solve the equations:*

$$\frac{P^2 Q(P)}{1 - \epsilon(P)} = tS$$

$$P = \frac{t(1 - \epsilon(P))}{NQ(P)}$$

- *N, a function of t, S, $\tilde{P}$, is the largest integer such that:*

$$N \leq \sqrt{\frac{t}{S}(1 - \epsilon(\tilde{P}))Q(\tilde{P})}$$

and so can be expressed as

$$N\left(\frac{t}{S}\right) = \sup\left\{z \in \mathbb{Z} : z \leq \sqrt{\frac{t}{S}(1 - \epsilon(\tilde{P}))Q(\tilde{P})}\right\}$$

- *the downstream price P is the solution (implicit for $N \geq 2$) of:*

$$P = \begin{cases} \frac{t(1-\epsilon(P))}{NQ(P)} < \tilde{P} & \text{if } N \geq 2 \\ \bar{P} & \text{if } N = 1 \end{cases}$$

- *downstream profit π are non-negative and given by:*

$$\pi = \begin{cases} \frac{t(1-\epsilon(P))}{N^2} - S < \tilde{P}Q(\tilde{P}) - S & \text{if } N \geq 2 \\ \bar{P}Q(\bar{P}) - S & \text{if } N = 1 \end{cases}$$

The number of firms N in the downstream market and the price they can charge P are jointly determined, just as in the baseline of Section 3.2 . The formal statement of solution is though a little more complicated because the equilibrium price P_n is no longer inversely proportional to N (which allowed us to substitute out for P_n in the baseline and solve directly for N as a function of t and S). Instead P_n is an implicit function of N that depends also on $\epsilon(P)$. Assumption 2. ensures that price and profitability are both decreasing functions of N and there is a unique N (treated as a continuous number) and associated price $P = \tilde{P}$ at which downstream firms make a zero profit. This this means that there is a unique N, as stated in the proposition, for which downstream operations are profitable when there are N downstream firms but unprofitable for $N+1$ downstream firms, determining N when there is competition in upstream provision of data control.

Downstream pricing and profitability then reflects one of two possible equilibrium outcomes: (i) when the product tS is relatively low, then there are two or more downstream firms and price setting is constrained by both elasticity of demand and by the transport costs of the marginal consumer (in equilibirium equi-distant between two neighbouring firms) and price is less than the downstream monopoly price $P = \bar{P}$; or,(ii) when the product tS is relatively high, only one downstream firm, i.e., a monopoly when the transport costs no longer constrains pricing and price is set at the monopoly price of $P = \bar{P}$. Assumption 2 ensures that the equilibrium price P is bounded even in the case of monopoly pricing.[5]

4.4 A downstream firm owns one of the competing upstream firms

As one variation we consider one of the upstream firms being owned by a downstream firm.

Proposition 6. *Let data validator* 1 *in the upstream market be a subsidiary of firm* 1 *in downstream market, if data validator* 2 *does not compete downstream the equilibrium conditions are unchanged.*

The proposition can be best understood by a simple explanation. If there is no competition by the data validator in the downstream market firms can still enter and establish relationships with data validator 2.

4.5 Profit maximising monopoly upstream

A profit maximising monopoly upstream can set service fees higher than in Proposition 1 using these high fees to extract all profit from downstream firms. The outcome can be summarised as follows:

Proposition 7. *Suppose that the upstream data control is provided by profit-maximising monopoly data validator, and that the elasticity of demand is characterised by Assumption 2, then the outcome is*

- *Only one firm remains in the downstream market i.e.,* $N = 1$
- *The downstream price is the monopoly price of Assumption 2,* $P = \bar{P}$
- *The profit of the one downstream firm is* $\pi = 0$
- *the service fee* F *for the one remaining firm equals its revenue less connection charge, and so depends on the number of firms* N_0 *initially connected:*

$$F = \begin{cases} \bar{P}Q(\bar{P}) & \text{if } N_0 \geq 1 \\ \bar{P}Q(\bar{P}) - S & \text{if } N_0 = 0 \end{cases}$$

The economic intuition in this case is straightforward. The upstream monopolist is able to extract all revenue, less any connection charge, from the downstream firms; and total revenue less any connection charge is maximised when there is also a downstream monopoly.

5. If there is no solution $P = \bar{P}$ with $\epsilon(\bar{P}) = 1$, then there is no economically sensible solution, a monopoly downstream firm is entirely unconstrained in its pricing, and is able to set an infinitely high price.

4.6 Monopoly upstream data validator owned by subset of downstream firms

Proposition 8. *Suppose that the upstream data control is provided by monopoly data validator, owned by a subset N_s of the N_0 firms initially in the market, with both the N_s the remaining $N_0 - N_s$ downstream firms connected to the monopoly data validator; that the elasticity of demand is characterised by Assumption 2; and that $1 < N_s < N_0$. Then the market outcome is:*

- *The monopoly data validator makes no charge its owners; and sets a sufficiently high charge of F to $N_0 - N_s$ remaining firms that they all exit the market; as a result $N = N_s$*
- *The market price P in the downstream market N is determined implicitly by the unique solution of:*

$$P = \frac{t}{N_s} \frac{1 - \epsilon(P)}{Q(P)}$$

- *The net profit of each of the N_s firms that own the validator, net of charges and distribution of upstream validator firm profits, is given by:*

$$\pi_s = \frac{t(1 - \epsilon(P))}{N_s^2}$$

The economic intuition is again straightforward. The N_s firms are able to act as a joint monopoly in determining the number of downstream firms, even though they then subsequently compete with each other downstream; their profits (which equal total revenues since they are already connected) are increased by reducing the number of downstream firms as far as possible (Assumption 2 ensures aggregate revenue is a monotonically decreasing function of the number of firms).

Many variations of this set up are possible. If mergers or side payments take place then the number of firms may fall below N_s and the outcome tend towards a downstream monopoly. The industry owned data validator could apply charges based on the quantity of downstream services, limiting downstream price competition and raising revenue and profits. Also, it is conceivable that individual firms could offer other otherwise excluded firms 'agency' access to data validation services, connecting them indirectly through their own systems. Such extensions of Proposition 8. are left for further work.

5 Applications

This section applies the framework developed in this paper to a variety of different contract based markets. These examples are predominantly financial: asset management; capital market trading and settlement; and payments (card and bank payments made both domestically and internationally); but there are also non-financial examples: property transactions, flight bookings and patent registration.

These examples illustrate substantial variation in the arrangements for validation of contractual data and governance of data operations. The relatively simple baseline of Section 3, with a separation of data validators and downstream providers of goods and services, is the exception rather than the rule. Reality, unsurprisingly, is considerably more complex than the formal modelling developed here. Still, the analytical and modelling framework developed in this paper is able to offer a number of insights into contractual data, data validation and market power in these various markets.

5.1 Financial services

- Asset management. One financial service, whose structural arrangements correspond fairly closely to our baseline model with seperation of upstream data validation, is asset management, responsible for managing around $100tn of financial assets globally on behalf of institutional and retail clients.[6]. The downstream service of asset managers are provided by many institutions, ranging from large global players such as Blackrock, Vanguard and Fidelity to much smaller firms. Asset managers hold their assets with custodian banks, such as JP Morgan Securities Services, State Street, Bank of New York Mellon and BNP Paribas. These custodians are the upstream data validators, responsible for both updating records of ownership and for so called 'corporate actions' such as payment of coupons and dividends, the withholding of tax and communication between securities issuers and investors.

  The framework of this paper offers some insights into this market. The vertical separation of upstream data validation and control from the provision of downstream customer services, is supportive of downstream entry and competition. Such separation is found in asset management and the large number of firms competing downstream is consistent with the modelling framework provided here. This is not to say that asset managers have no market power; but to the extent that they do this arises for other reasons, for example the limited knowledge and understanding of financial products amongst retail consumers preventing them 'shopping around'. At the same time the framework fo this paper highlights the major role of standardisation, for reducing costs and delivering better value to customers. Here the very size of the asset management market is a barrier to coordinated change in data technologies. Intervention to promote data standardisation can be one way of addressing concerns about high asset management fees, especially for retail customers.

- Financial market trading. Closely related to asset management is financial market trading and associated post-trade clearing and settlement.[7] Here, asset managers are now the downstream clients along with hedge funds, high frequency traders and retail intermediaries, such as spread betting firms and investment brokers, providing indirect market access to retail clients. Once again there is vertical separation, with trade execution to these downstream clients provided by brokers in equities and dealers in bond, OTC derivatives and foreign exchange. This market is much more complicated than in asset management. There is additional upstream horizontal separation, with post trade clearing and settlement is provided by a range of further data validators,

6. McIntyre et al. (2023)

7. The complexities of this market make it difficult to quantify revenues and margins and assess welfare losses available from reducing processing costs and promoting competition; Cunliffe (2022) refers to industry estimates of post-trade global processing costs of $40-$45bn, much of which is associated with lack of interoperability between systems and resulting need for extensive manual processing; to which must be added high costs of liquidity management when assets are trapped or held in different locations and of the impact of fragmentation of trading on bid-ask spreads and market impact of trades. A recent industry report Choudhury et al. (2023) estimates $15-20bn in annual operational cost savings, from adoption of standradised industry wide data validation through employing distributed ledger technologies, along with further benefits from efficiencies of employment of collateral and widening access to capital markets.

including central counterparties (for trade guarantees and netting out of offsetting obligations), commercial banks (for final payment) and custodian banks (for final transfer of ownership). Arrangements are even more complicated in cross-border financial trading to accommodate clients operating in different time zones and under different legal jurisdictions. All these processes are subject to a web of supervisory oversight and financial rules and regulations, in order to contain prudential and systemic financial risk, promote market integrity and protect customers.

The framework of this paper again offers insight into reducing costs and achieving effective competition in financial market trading and post-trade operations. In contrast to asset management, the complexity of operations in financial trading make it far from clear that the current arrangements for data validation support sufficient competition in trade execution. As in asset management there is also, prima-facia, a case that standardisation could support substantial operational cost savings and lower barriers to entry; but as in asset management it is difficult to achieve the necessary co-ordinated adoption and the consequent reduction in barriers to downstream entry may lead to substantial resistance from downstream providers.

- Domestic retail payments. In domestic bank retail payments there is vertical integration, with banks combining the downstream role of providing transaction services and the upstream role of 'data validator'. They are responsible for updating records of their customers account ownership claims according to the requirements of various domestic payment scheme, scheduling settlement between participating payment institutions and setting the technical standards employed in payments.[8] These payment schemes are operated as utilities, with bank membership and often historically bank owned though other governance and ownership arrangements are also common (for example the US Federal Reserve banks own and operates payment schemes such as FedWire, the Fed's ACH and the newly established FedNow).

  Complementing these bank payment schemes, in online and 'point of sale' (i.e. face to face) consumer payments, are the various debit and credit card payment schemes. The largest of these globally are Visa and Mastercard. Here the schemes are the data validators (though final payments to merchants must be made as a domestic bank retail payment), along with the provision of further services such as fraud control, guarantees for consumers and reversal of payments for returned goods. Historically these card schemes were bank co-operatives but now the major card schemes are now for-profit public companies.

  What insights does the analytical framework of this paper offer into these arrangements? Until two decades ago access to the payment schemes, determining which firms were able to validate payments transactions, was limited to regulated banks. This effectively limiting entry as in Propositions 8 above.[9] This has attracted the attention of policy makers. The economics literature and competition authorities have paid considerable attention to card schemes, focusing on regulation of 'interchange fees' paid to card issuers.[10]. More recently the focus of policy makers has shifted to bank

8. Each jurisdiction has several such schemes for different categories of payment, for example the US ACH schemes, the various national schemes falling under EU SEPA or Single European Payments Area and the UK BACS and Faster Payments schemes.

9. The case of the UK is described in Chapter 3 of Cruickshank (2000).

10. Rochet and Wright (2010) and Verdier (2011)

payment schemes, notably with measures to reduce barriers to entry for non-bank payment service providers, arising from bank ownership and governance of domestic bank payment schemes, for example through the EU Second Payments Service Directive PSD2;[11] and from initiatives to develop and promote greater standardisation, aligned with the ISO20022 XML based messaging standards, and faster payment schemes in which domestic bank payments can be credited to recipients within minutes of initial instruction.[12] This shift in policy, emphasising the role of standards and governance and reducing barriers to entry. In terms of the framewofk of this paper, the earlier intervention of regulators in card interchange fees was undertaken through the powers of competition authorities to intervene in collective price setting arrangements; but interchange fees are paid between co-operating providers of downstream services, not by downstream consumers. This suggests that the shift since then to focusing on standard setting and removing barriers to entry to upstream data validation is a more effective approach to promoting competition in retail payments.

- International payments. Similar issues of data control arise in international payments. A major difference from domestic retail payments is the need for the updating of data in two jurisdictions: those of the payment sender and of the payment recipient. As a consequence the downstream provider of the payment service relies on access, from a correspondent bank, to upstream data validation services in a foreign jurisdiction. This is a case if multiple vertical separation with an additional 'midstream' service. The downstream, the final customers are the account holders. Midstream are the account providers, the data validators for domestic retail payments. These account providers in turn rely on the further data validation services of the correspondent bank who connects the account provider from one jurisdiction to the retail payment systems in the other jurisdiction.

  The high costs and often slow and uncertain processing of cross border payments have become a recent focus of attention of the G20 group of nations and, working through the Financial Stability Board, of central banks worldwide.[13] The analytical framework of this paper, viewing correspondent banks as the data validators in cross border payments, offers some insight. One reason for high costs are substantial barriers to entry in correspondent banking, though in this case the barriers are to an important extent the consequence of global policy measures, through raising costs of validation through stringent anti money laundering (AML) and countering financing of terrorism (CFT) regulations and substantial fines for non-compliance. Efforts at standardisation, especially standardisation of identity validation, to restore the commercial viability of correspondent banking are a promising way of addressing the concerns about costs and inefficiency in cross-border payments.

5.2 Non-financial data validation

The remainder of this section briefly discusses three examples of validation of non-financial contractual data, again illustrating the variety of possible outcomes.

11. For evidence on the competition impact of PSD2 see Polasik et al. (2020)
12. See CPMI (2021)
13. See He, Milne and Zachariadis (2022) for summary of this policy agenda as of early 2022 and of the potential for using central bank digital currencies to reduce costs and inefficiencies in cross border payments.

- Property ownership registration and ownership transfer. Registration of ownership of land and physical assets such as vehicles is typically organised through a data record system provided by state or national government.

  These cases of a state owned registry system can be viewed as that of a monopoly data validator seeking to maximise social welfare. There can though be barriers to access, imposed for example because of security concerns, that in turn limit competition in the downstream market.[14]

- Flight and other travel bookings. Passenger flight bookings require co-operation between airlines in order to deliver a complete journey provided by multiple providers. Here the required contractual data validation is provided by global distribution systems (GDS), the platforms provided by one of the three global providers: Amadeus, Sabre and Travelport (which operates Galileo, Worldspan and other GDS systems). These platforms did not originate as independent data validators. They were founded in the 1906s and 1970s under airline ownership in order to apply the then mainframe computing technologies to automation of passenger reservations. Over time they developed as two sided platforms facilitating search by travel agents to locate and agree airline reservations and then, by the late 1990s, for online reservations via internet platforms instead of travel agents.

  Two sided platforms are not data validators, as long as they are simply bringing together two sides of a market. In the case of GDS platforms they do more than simply connecting airline booking providers with airlines, they also take on a data validation role when connecting flights are provided by different airlines through codesharing arrangements. This vertical separation of data validation does not however appear to substantially limit downstream competition, because the airlines participating in codesharing agreements, most commonly amongst members of main airline alliances, are able to distribute their available flights through more than one GDS (i.e. there is multihoming on the carrier's side of the platforms).

- Patent records are also organised through national record systems. In the UK a patent is granted by the UK Intellectually Property Office. In the US, the United States Patents and Trademark Office (USPTO) is the national patent office and trademark registration authority. In European for a patent one needs to apply to the national office of each country.

  The case of patents is somewhat different from our other examples, because patent record systems are set up in order to provide owners of patents with market power, ensuring that competitors cannot exploit their intellectual property and reduce commercial returns to the creation of new intellectual property.

14. The registration of pedigree dogs is a related example. These ownership records are of less economic significance than land or vehicles, but they illustrate how data validators such as The Kennel Club in the UK or The American Kennel Club, even though they operate as not-for-profits appear to enjoy market power that allows them to extract rent from dog breeders operating downstream, resulting in healthy operating margins.

6 Data fragmentation and validation: the case of supply chains

This final section offers some reflections on the extent to which separating downstream data based production of goods and services and upstream validation of contractual data could create economic value in industries where data is currently fragmented. For concreteness the focus here is on supply chain data in production and sale of goods.[15]

This challenge of the economic organisation of contractual data validation can be related the established literature on the economics of transaction costs. A large literature has explored the insights of transaction cost economics for economic organisation: economic activity takes place through either market exchange or of administrative allocation of resources within organisations, the choice being made according to which arrangement achieves lower transaction costs.[16]

The example of supply chain data is pertinent because of substantial transaction costs of validating contractual data in market exchange between independent organisations. Most data record systems in supply chains are proprietary. Manufacturers, construction firms, retailers and other downstream firms establish contracts with individual supplying firms using proprietary systems for recording contractual information, creating purchase orders, monitoring the quality and timeliness of delivered goods and for processing invoices and payments. Thus the initial situation is data fragmentation, with each downstream firm conducting its own separate validation of contractual data. The associated transaction costs are not so large that inputs are made in house, rather than bought externally; but the adoption of digital data technologies, and the possibility of implementing them using vertically separated standardised data validation services, create opportunity for substantially reducing transaction costs and increasing economic value.[17]

Insight into this opportunity can be obtained from the framework developed in this paper. This highlights: (a) the potential for efficiency gains from vertical separation, with standardisation of data, digital storage and retrieval along with independently provided data validation; and (b) the possible resistance by incumbents to such vertical separation and standardisation if it reduces their market power and operating margins.

A prominent illustrative example is that of international goods trade with the additional complexities of both compliance with product regulations and payment of tariffs along with the provision of supporting financial and logistical services. Currently all the contractual data and associated data processing remains largely paper based requiring costly manual processing and adding substantially to the costs of international goods trade. The potential gains of digitisation through vertical separation are large. Widespread manual processing and validation currently accounts for some 3% of the value of international goods trade in the G7, a figure which has been estimated could be reduced to 0.7%.[18] These are

15. Similar issues about fragmentation of data and economic organisation arise in other contexts: for example in capital markets for the post-trade processing of over the counter derivative contracts, securities lending and sale and repurchase (repo) markets, where fragmentation of data complicates the resolution of financial distress and creates systemic risk; also in public sector procurement.

16. A literature dating back to the work of Coase (1937), key contributions are reviewed in Williamson (1981), Cuypers et al. (2021) summarise the substantial subsequent scholarship in the management journals; the literature of course recognises the possibility of other intermediate arrangements between these poles, such as joint ventures or franchising see e.g. Williamson (1991).

17. See related work on information systems, transactions costs and the governance and structure of supply chains e.g. Croom 2001; Arlbjørn, Haas and Munksgaard 2011.

18. Coriolis Technologies 2022.

the direct costs; but it is arguable that the associated reduction in indirect costs, arising because opaque data limits access to and increases the costs of trade finance, insurance and logistical services, could be an order of magnitude bigger. A case can be made that digitisation and the associated increased competition in facilitating services could lead to an overall boost to the volume of world trade and of global output of 5% or more with particular benefits for SMEs and low and middle income countries.[19].

Similar impacts can be expected from vertical separation of data validation within domestic supply chains. The gains are not limited to cost reductions. The transparency supported by standardised data validation can also be a key protection against fraud. Supplying firms are currently financially vulnerable in the event of late payment by contractors or increased risk aversion by the suppliers of trade finance (a risk that materialised in the aftermath of the global financial crisis of 2007-08). Increased transparency of supply chains through vertical separation of data validation could limit such financial exposure. It could also help purchasing firms anticipate supply shocks and improve the resilience of their supply chains. Finally it could facilitate much greater traceability, for compliance with ESG goals including the assessment of the carbon contribution of supply chains and support the increased recycling of materials through circular manufacturing.

The framework of this paper also offers insight into the barriers to unlocking these economic gains from standardised supply chain data validation. This is a governance and regulation challenge of the kind addressed at length in the transaction cost literature. The organisational choice is however more complicated than the basic one between resource allocation through market exchange or organisational decision making. The shift to vertical separation of data validation requires industry wide collaboration. Establishing a consensus for change is not easy and, even when the value created is substantial, may not take place voluntarily without national and international policy intervention. Examining and making the case for change and winning business and political support is a substantial agenda for future research and policy analysis.

19. Even larger estimates have been produced in a series of quantitative reports for the International Chamber of Commerce Harding, Haywood and Strazdins 2022; Coriolis Technologies 2022; as a result of the perceived large gains digital facilitation of trade is now a policy priority for the G20 group of nations

Appendix: Proofs of propositions

The following table summarises the notation employed in these propositions and their proofs.

Variable	Type	Definition
S	Parameter	Switching/ entry cost connecting to a data validator
t	Parameter	Transport cost on the Salop Circle (linear in distance)
N_0	Parameter	Initial number of firms in the downstream market
N	Endogenous	Final number of firms in downstream market
$n \in [1, \ldots N]$	Index	Distinguishes downstream firms
$\theta_n = 2\pi \times n \div (N-1)$	Angle	Location of firm n on Salop circle
J	Exogenous	Number of upstream data validators
$j \in [1, \ldots J]$	Index	Distinguishes upstream data validators
P_n	Endogenous	Price of downstream output of firm n
$Q(P_n)$	Exogenous	Demand for output of firm n (independent of location)
$q(p)$	Exogenous	Q() as logarithmic function, $q = \ln Q$, $p = \ln P$
$\epsilon(P) = -q'(p)$	Derived	Price elasticity of consumption
$\Omega(N; N_0, S, t)$	Endogenous	Social welfare

Proposition 1: Let firm n be initially using data validator 1, S be the connection costs for firm n and the fee charged for service from data validator j to firm n be F_j^n. In equilibrium the following holds:

$$F_1^n = S \tag{3}$$

$$F_2^n = 0 \tag{4}$$

Proof. Assume firm n is already connected to data validator 1. The choices for firm n are (i) staying with data validator 1 paying a fee F_1^n; or (ii) paying a fee F_2^n plus connection cost S to switch to operator 2. So it will switch only if $F_2^n + S < F_1^n$.

Consider the best responses of the two firms. Given F_1^n the best response of firm 2 is to set a price $F_2^n = F_1^n - S - \epsilon$ subject to the constraint $\pi_2 = F_2^n > 0$ for some small ϵ. Given F_2^n the best response for firm 1 is to set the highest price that make firm n indifferent between staying with operator 1 or moving to operator 2 i.e. $F_1^n = F_2^n + S$. Solving for the best-response Nash equilibrium yields Proposition 1. □

Proposition 2: With free entry and exit:

- **N is a function of t and S, and can be expressed as**

$$\bar{N}\left(\frac{t}{S}\right) = \max\left(1, \sup\left\{z \in \mathbb{Z} : z \leq \sqrt{\frac{t}{S}}\right\}\right)$$

- **If there is downstream competition, i.e. $\bar{N} > 1$, all downstream firms charge the same price (P) and earn the same profit (π) given by:**

$$\begin{aligned} P &= \frac{t}{\bar{N}} \\ \pi &= \frac{t}{\bar{N}^2} - S \end{aligned}$$

but if there is a downstream monopoly, i.e. $\bar{N} = 1$, then the price charged by and profit of the downstream monopoly are unbounded.

Proof. Suppose there are N downstream firms. Each downstream firm sets a price P_n. So equilibrium is a choice of the vector of prices $(P_1 \cdots P_N)$ The 'marginal' customer between downstream firm n and $n+1$ is located at ${\theta_n}^{n+1}$, which satisfies:

$$P_n + t\left(\theta_n^{n+1} - \frac{n-1}{N}\right) = P_{n+1} + t\left(\frac{n}{N} - \theta_n^{n+1}\right)$$

so

$$\theta_n^{n+1} = \frac{2n-1}{2N} + \frac{P_{n+1} - P_n}{2t}$$

The marginal customer is positioned half-way between the two downstream firms with an adjustment for price difference that depends on the transport cost t. The profit of downstream firm n is then given by:

$$\begin{aligned} \pi_n &= P_n\left(\frac{n-1}{N} - \theta_{n-1}^{n} + \theta_n^{n+1} - \frac{n-1}{N}\right) - S && (5) \\ &= P_n(\theta_n^{n+1} - \theta_{n-1}^{n}) - S && (6) \\ &= P_n\left(\frac{P_{n+1} - 2P_n + P_{n-1}}{2t} + \frac{1}{N}\right) - S && (7) \end{aligned}$$

The first order condition for profit maximisation is:

$$\frac{P_{n+1} - 4P_n + P_{n-1}}{2t} + \frac{1}{N} = 0$$

which yields:

$$P_n = \frac{P_{n+1} + P_{n-1}}{4} + \frac{t}{2N}$$

The symmetry of the downstream firms then implies that all downstream firms set the same price $p_n = p$ in equilibriums, given by:

$$P = \frac{t}{N}$$

and equilibrium profits are:

$$\pi_n = \frac{t}{N^2} - S$$

Firms will exit the market if they are unable to make a profit paying the charge S for data services from data validators. Similarly firms will enter the market if they are able to make a profit paying the connection charge S to establish data services. For a given N profits are given by:

$$\pi(N) = \frac{t}{N^2} - S$$

This is a monotonically decreasing continuous function of N. So if we solve for $\pi(N) = 0$, we obtain an upper bound for N, the number of firms that can profitably remain in the market. This leads directly to:

$$N \leq \sqrt{\frac{t}{S}}$$

and the expression for N □

Proposition 3. *In the model of this section the social welfare, the sum of consumer and producer surplus, is given by:*

$$\Omega(N) = -S\max(0, N - N_0) - \frac{t}{4N}$$

and the socially optimal number of firms is an integer N^*, that depends upon N_0, t and S. Let $\tilde{N}$ be the largest integer such that $\tilde{N} \leq \frac{1}{2}\sqrt{\frac{t}{S}}$ and so can be expressed as:

$$\tilde{N}\left(\frac{t}{S}\right) = \sup\left\{z \in \mathbb{Z} : z \leq \frac{1}{2}\sqrt{\frac{t}{S}}\right\}$$

Then N^* is given by:

$$N^* = \begin{cases} N_0 & \text{if } N_0 \geq \frac{1}{2}\sqrt{\frac{t}{S}} \\ \tilde{N} & \text{if } N_0 < \frac{1}{2}\sqrt{\frac{t}{S}} \text{ and } \frac{1}{2}\sqrt{\frac{t}{S}} - \tilde{N} \leq \frac{1}{2}\left(1 + \sqrt{\frac{t}{S}} - \sqrt{1 + \frac{t}{S}}\right) \\ \tilde{N} + 1 & \text{otherwise} \end{cases}$$

Proof. The social objective is the total profits of all firms less the costs to consumers of both transport and of consumption of the downstream good. These three components of social welfare are as follows:

- total profits: the sum of downstream profits ($p - NS$) and upstream profits ($S\min(N, N_0)$ i.e. connection charges paid by firms which are initially in the downstream market at $t = 0$ and still remain in period $t = 1$), which added together equal:

$$p - S\max(0, N - N_0)$$

  i.e. total revenue from consumers less total new connection costs.

- total transport costs T_n for consumers purchasing from firm n are given by:

$$\begin{aligned} T_n &= t\left(\int_{\theta_n - \frac{1}{2N}}^{\theta_n} (\theta_n - \theta)\mathrm{d}\theta + \int_{\theta_n}^{\theta_n + \frac{1}{2N}} (\theta - \theta_n)\mathrm{d}\theta\right) \\ &= 2tN\left[\tfrac{1}{2}\theta^2\right]_{\theta=0}^{\theta=\frac{1}{2N}} \\ &= 2tN\tfrac{1}{8N^2} = \tfrac{t}{4N^2} \end{aligned}$$

  so the total transport costs of all consumers are given by:

$$N \times \frac{t}{4N^2} = \frac{t}{4N}$$

- Total consumption by all consumers equals 1, so the total consumer expenditure on the downstream good is p

Combining these three components, netting out consumer expenditure and connection charges for existing firms against revenue, total social welfare is given by:

$$\Omega(N) = p - S\max(0, N - N_0) - \frac{t}{N} - p = -S\max(0, N - N_0) - \frac{t}{4N}$$

Social welfare is the negative of total costs: connection costs for new firms joining the market and transport costs for consumers. The price paid for consumption is a transfer from consumer to producers that does not affect social welfare. This has the further implication that a downstream monopoly, even though it results in an unbounded price p, may be a socially optimal outcome.

From this we can obtain the socially optimal value of N, first considered as a continuous number and then adjusting to allow for the constraint that N must be an integer. Maximisation in continuous N is determined by the first order condition:

$$-S + \frac{t}{4N} = 0$$

which yields $N = \frac{1}{2}\sqrt{\frac{t}{S}}$. The requirement that the number of firms is an integer $N \geq N_0$ means that social welfare is maximised by either $N = N_0$ or one of the two nearest integer neighbours to the value of N satisfying this first order condition. Let $\tilde{N}$ denote the lower integer neighbour, so the higher integer neighbour is $\tilde{N}+1$. We can determine which of the two integer neighbours achieves the higher social welfare from the sign of:

$$\begin{aligned}
\Omega(\tilde{N}+1) - \Omega(\tilde{N}) &= -S - \frac{t}{4(\tilde{N}+1)} + \frac{t}{4\tilde{N}} \\
&\propto 4S(\tilde{N}+1)\tilde{N} + t\tilde{N} - t(\tilde{N}+1) \\
&\propto \tilde{N}^2 + \tilde{N} - \frac{1}{4}\frac{t}{S}
\end{aligned}$$

The same social welfare is achieved with both lower and higher integer neighbour whene $\Omega(\tilde{N}+1) = \Omega(\tilde{N})$Solving using the standard quadratic formula and taking the positive root, $\tilde{N}$ is then given by:

$$\tilde{N} = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{1}{2}\left(-1 + \sqrt{1 + \frac{t}{S}}\right)$$

If $\tilde{N}$ is greater than or equal to this value, i.e., sufficiently close to the continuous solution for N, it is the lower neighbour integer $\tilde{N}$ that achieves the highest social welfare; otherwise it is the higher neighbour integer $\tilde{N}+1$.

□

Proposition 4 The socially optimal price under the assumptions of this section is $P = P^* = 0$. While the socially optimal number of firms is given by N^* as stated above in Proposition 3. This yields a total social surplus, the sum of consumer and producer surplus, of:

$$\Omega(N^*) = \int_{\phi=0}^{1} Q(\phi)\mathrm{d}\phi - S\max\{N^* - N_0, 0\} - \frac{t}{4{N^*}^2}$$

Proof. The proof parallels that of Proposition 3. The social surplus, for given N and P must be amended by including also, alongside the transport costs and connection costs in the baseline model, the producer surplus (from providing the downstream service, before payment of any connection costs i.e. $P(Q(P))$) and consumer surplus (from consuming the downstream good rather than the alternative outside good, given by: $\int_{\phi=P}^{1} Q(\phi)\mathrm{d}\phi$). So the social planner seeks to maximise

$$\Omega(N,P) = P(Q(P)) + \int_{\phi=P}^{1} Q(\phi)\mathrm{d}\phi - S\max\{N-N_0,0\} - \frac{t}{4N^2}$$

Since this is linearly separable in the number of firms N and the price P, the optimal number of firms N^* is the same as in Proposition 3. Taking first order condition wr.t. P and equating to zero we have:

$$\frac{\partial\Omega(N,P)}{P} = Q(P) + Q'(P)P - Q(P) = Q'(P)P = 0$$

Since we assume that $Q'(P) > 0$, this implies that $P = P^* = 0$ i.e. the familiar efficiency condition that price equals marginal cost which is zero. Substituting back for P in the social objective completes the proof. □

Proposition 5. Assuming that individual consumption of the preferred downstream good is represented by $Q = Q(P)$ with $Q'(P) < 0$, that $N_0 \leq 1$, that the elasticity of demand $\epsilon(P)$ is characterised by Assumption 2, with a unique price $\bar{P}$ for which $\epsilon(P) = 1$, and $S \leq \bar{P}Q(\bar{P})$ then:

- **there is a unique pair of real numbers ($P = \tilde{P}$,$N = \tilde{N}$) solving jointly:**

$$PQ(P) = \frac{t}{N}1 - \epsilon(P) = S$$

- **the number of firms sustainable in the market N, is the largest integer such that:**

$$N \leq \sqrt{\frac{t}{S}(1-\epsilon(\tilde{P}))Q(\tilde{P})}$$

  and so can be expressed as

$$N\left(\frac{t}{S}\right) = \sup\left\{z \in \mathbb{Z} : z \leq \sqrt{\frac{t}{S}(1-\epsilon(\tilde{P}))Q(\tilde{P})}\right\}$$

- **the downstream price P is the solution (implicit for $N \geq 2$) of:**

$$P = \begin{cases} \frac{t}{N}\frac{1-\epsilon(P)}{Q(p)} < \tilde{P} & \text{if } N \geq 2 \\ \bar{P} & \text{if } N = 1 \end{cases}$$

- **downstream profit π is given by:**

$$\pi = \begin{cases} \frac{t(1-\epsilon(P))}{N^2} - S < \tilde{P}Q(\tilde{P}) - S & \text{if } N \geq 2 \\ \bar{P}Q(\bar{P}) - S & \text{if } N = 1 \end{cases} \geq 0$$

Proof. The proof is similar as that for Proposition 2. Consider first N firms, located symmetrically on the downstream circle, each charging P_n for $n = 1 \dots N$. Consumer surplus for the consumer located at $\theta > n$, less transport costs, from purchasing from the firm located at n is given by

$$\mathrm{CS}(\theta, n) = \int_{P_n}^{1} Q(\phi)\mathrm{d}\phi - t\frac{\theta - n}{N}$$

If $N \geq 2$ the marginal consumer indifferent between purchasing from the downstream at n and $n+1$ is located at $\hat{\theta}$ given by:

$$\hat{\theta} = \frac{N}{2t}\int_{P_n}^{P_{n+1}} Q(\phi)\mathrm{d}\phi + n + \frac{1}{2}$$

and total profit of the firm located at n is given by

$$\pi(P_n) = P_n Q(P_n)\left[\frac{1}{2t}\int_{P_n}^{P_{n+1}} Q(\phi)\mathrm{d}\phi + \frac{1}{2t}\int_{P_n}^{P_{n-1}} Q(\phi)\mathrm{d}\phi + \frac{1}{N}\right] - S$$

with a FOC with respect to P_n of

$$Q(P_n)(1 + \frac{P_n Q'(P_n)}{Q(P_n)})\left[\frac{1}{2t}\int_{P_n}^{P_{n+1}} Q(\phi)\mathrm{d}\phi - \frac{1}{2t}\int_{P_n-1}^{P_n} Q(\phi)\mathrm{d}\phi + \frac{1}{N}\right] - \frac{1}{t}Q^2(P_n) = 0$$

In equilibrium $P_n = P$ for $\forall n$ and this becomes (substituting using $\epsilon = \frac{P_n Q'(P_n)}{Q(P_n)}$):

$$PQ(P) = \frac{t}{N}(1 - \epsilon(P))$$

or:

$$P = \frac{t}{N}\frac{1 - \epsilon(P)}{Q(p)}$$

If on the other hand $N = 1$ i.e., there is a downstream monopoly, we have a standard monopoly pricing and with zero marginal costs the profit maximising price is given by:

$$P = \bar{P}$$

The remainder of the proof parallels that of Proposition 2. The key issue is determining whether the price equilibrium can support competition in the downstream market i.e. $N \geq 2$. Assuming $N \geq 2$, then the firm profitability is given by:

$$\pi(N) = PQ(P) - S = \frac{t}{N}(1 - \epsilon(P)) - S$$

We are interested in finding the largest integer N for which these profits are positive. $\epsilon(P)$ a function of N (unlike Proposition 2 where $\epsilon = 0$ independent of P and N) so we cannot solve explicitly for $\tilde{N}$ at which $\pi(\tilde{N}) = 0$. However, we know (from Assumption 1) that

$$\frac{\mathrm{d}P}{\mathrm{d}N} = -\frac{t}{N^2}\frac{1 - \epsilon(P)}{Q(p)}$$

and using this we can show that

$$\frac{\mathrm{d}\pi}{\mathrm{d}N} < 0$$

Since, further $\pi(N=1) > 0$ and $\lim_{N\uparrow\infty} \pi(N) < 0$, this implies that there is some value of N, $N = \tilde{N}$, treated as a continuous variable, and some associated price $P = \tilde{P}$ for which $\pi(\tilde{n}) = 0$.

The remainder of the proof follows that of Proposition 2. Given $\tilde{N}$, the number of firms that can be sustained in the market is the largest integer $N \leq \tilde{N}$; establishing the level fo pricing and profits for the two cases $N = 1$ and $N \geq 2$.

□

Proposition 6: Let data validator 1 in the upstream market be a subsidiary of firm 1 in downstream market. Data operator 2 does not compete downstream.

Proof. Let data validator 1 own firm 1 in downstream. If data validator 2 does not compete downstream, would there be a change in exit or entry conditions? No firm can always enter and establish a relationship with data validator 2. Do exit conditions change? No, it will remain the same as it still relies on t and S □

Proposition 7 Suppose that the upstream data control is provided by profit-maximising monopoly data validator, and that the elasticity of demand is characterised by Assumption 2, then the outcome is

- **Only one firm remains in the downstream market i.e., $N = 1$.**
- **The downstream price is the monopoly price of Assumption 2, $P = \bar{P}$.**
- **The profit of the one downstream firm is $\pi = 0$.**
- **the service fee F for the one remaining firm equals its revenue less connection charge, and so depends on the number of firms N_0 initially connected:**

$$F = \begin{cases} \bar{P}Q(\bar{P}) & \text{if } N_0 \geq 1 \\ \bar{P}Q(\bar{P}) - S & \text{if } N_0 = 0 \end{cases}$$

Proof. Proof is straightforward. A monopoly profit maximising upstream data validator is able to extract all rent from downstream firms, by setting a charge F equal to firm revenue net of connection costs. Thus it will charge firms that connect an amount S less than firms that are already connected at set its fees at a level at which all firms in the downstream market just break even with profits of $\pi = 0$. Total revenue extracted by the upstream monopoly is maximised when there is also a downstream monopoly, with a single firm i.e. $N = 1$. Given Assumption 2, the downstream monopoly will raise prices to $P = \bar{P}$ to maximise its own revenue.

□

Proposition 8 Suppose that the upstream data control is provided by monopoly data validator, owned by a subset N_s of the N_0 firms initially in the market, with both the N_s and the remaining $N_0 - N_s$ downstream firms connected to the monopoly data validator; that the elasticity of demand is characterised by Assumption 2; and that $1 < N_s < N_0$. Then the market outcome is:

- **The monopoly data validator makes no charge its owners; and sets a sufficiently high charge of F to $N_0 - N_s$ remaining firms that they all exit the market; as a result $N = N_s$**
- **The market price P in the downstream market N is determined implicitly by the unique solution of:**

$$P = \frac{t}{N_s}\frac{1-\epsilon(P)}{Q(P)}$$

- **The net profit of each of the N_s firms that own the validator, net of charges and distribution of upstream validator firm profits, is given by:**

$$\pi_s = \frac{t(1-\epsilon(P))}{N_s^2}$$

Proof. All revenues of the monopoly data validator are distributed to its owners; so the only economically relevant charge is the connection charge F for the $N_0 - N_s$ remaining firms that stay in the market. The monopoly data validator can extract all profit from the remaining firms so it sets this price by first deciding the total number of firms on the market N, and then determining F so these additional firms make zero profits. In this way the monopoly data validator can control the number of firms N in the downstream market

The market price is depends on the number of firms in the downstream market N and given implicitly by the solution of:

$$P = \frac{t}{N}\frac{1-\epsilon(P(N))}{Q(P(N))}$$

which implies that, for any given N, the net profit of each of the N_s firms that own the validator, net of charges and distribution of upstream validator firm profits, is simply total market revenue, divided by N_s and so is given by:

$$\pi_s = \frac{P(N)Q(P(N))}{N}$$

and so

$$\frac{\partial \pi_s}{\partial N} = (\frac{Q(P(N)) + P(N)Q'(P(N))}{N}\frac{\partial P}{\partial N} - \frac{P(N)Q(P(N))}{N^2}) = \frac{Q(P(N))}{N}\left((1-\epsilon)\frac{\partial P}{\partial N} - \frac{P(N)}{N}\right) < 0$$

where the inequality is obtained because, from Assumption 2, $\frac{\partial P}{\partial N} < 0$. This in turn implies that the firms owning the monopoly data validator will seek to reduce the number of the firms in the market as far as possible, from N_0 to N_s

□

References

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian and Asuman Ozdaglar. 2019. *Too much data: Prices and inefficiencies in data markets.* Technical report. National Bureau of Economic Research.

Arlbjørn, Jan Stentoft, Henning de Haas and Kristin Balslev Munksgaard. 2011. 'Exploring supply chain innovation'. *Logistics research* 3:3–18. ISSN: 1865-035X.

Armstrong, Mark, Simon Cowan and John Vickers. 1994. *Regulatory reform: economic analysis and British experience.* Vol. 20. MIT press.

Armstrong, Mark, and David E M Sappington. 2006. 'Regulation, Competition, and Liberalization'. *Journal of Economic Literature* 44, no. 2 (June): 325–366. ISSN: 00220515. https://doi.org/10.2307/30032251. http://www.jstor.org/stable/30032251.

Babina, Tania, Anastassia Fedyk, Alex Xi He and James Hodson. 2022. 'Artificial Intelligence, Firm Growth, and Product Innovation'. *Journal of Financial Economics (JFE), Forthcoming.*

Beraja, Martin, David Y Yang and Noam Yuchtman. 2023. 'Data-intensive innovation and the State: evidence from AI firms in China'. *The Review of Economic Studies* 90 (4): 1701–1723. ISSN: 0034-6527.

Buehler, Stefan, and Armin Schmutzler. 2008. 'Intimidating competitors—Endogenous vertical integration and downstream investment in successive oligopoly'. *International Journal of Industrial Organization* 26 (1): 247–265.

Carrière-Swallow, Yan, and V Haksar. 2019. 'The Economics and Implications of Data: An Integrated Perspective'. https://www.imf.org/en/Publications/Departmental-Papers-Policy-Papers/Issues/2019/09/20/The-Economics-and-Implications-of-Data-An-Integrated-Perspective-48596.

Choudhury, Roy, Khunal Jhanji, Humza Samd, Simon Gleeson and Scott Bennett. 2023. *Impact of Distributed Ledger Technology on Global Capital Markets.* Technical report. Washington DC: Global Financial Markets Association. https://www.gfma.org/policies-resources/gfma-publishes-report-on-impact-of-dlt-in-global-capital-markets/.

Coase, Ronald. 1937. 'The theory of the firm'. *Economica* 4 (16): 386–405.

Coriolis Technologies. 2022. *G7 — CREATING A MODERN DIGITAL TRADE ECOSYSTEM CUTTING THE COST AND COMPLEXITY OF TRADE.* Technical report. International Chamber of Commerce, UK. https://www.iccgermany.de/wp-content/uploads/2021/10/Creating-a-Modern-Digital-Trade-Ecosystem-G7.pdf.

CPMI. 2021. 'Developments in retail fast payments and implications for RTGS systems'. Basel. https://www.bis.org/cpmi/publ/d201.htm.

Croom, Simon. 2001. 'Restructuring supply chains through information channel innovation'. *International Journal of Operations and Production Management* 21 (4): 504–515. ISSN: 0144-3577.

Cruickshank, Donald. 2000. *Review of Banking Services in the UK.* Technical report. London.

Cunliffe, Jon. 2022. *Innovation in post trade services - opportunities, risks and the role for the public sector.* https://www.bankofengland.co.uk/speech/2022/september/jon-cunliffe-keynote-speech-at-the-afme-operations-post-trade-technology-innovation-conference.

Cuypers, Ilya R P, Jean-François Hennart, Brian S Silverman and Gokhan Ertug. 2021. 'Transaction cost theory: Past progress, current challenges, and suggestions for the future'. *Academy of Management Annals* 15 (1): 111–150. ISSN: 1941-6520.

Farboodi, Maryam, Roxana Mihet, Thomas Philippon and Laura Veldkamp. 2019. 'Big data and firm dynamics'. In *AEA papers and proceedings,* 109:38–42.

Farrell, Joseph, and Paul Klemperer. 2007. 'Coordination and lock-in: Competition with switching costs and network effects'. *Handbook of industrial organization* 3:1967–2072.

Gilbert, Richard J. 2021. 'Separation: a cure for abuse of platform dominance?' *Information Economics and Policy* 54:100876.

Grossman, Sanford J, and Oliver D Hart. 1986. 'The costs and benefits of ownership: A theory of vertical and lateral integration'. *Journal of political economy* 94 (4): 691–719. ISSN: 0022-3808.

Gu, Yiquan, and Tobias Wenzel. 2009. 'A note on the excess entry theorem in spatial models with elastic demand'. *International Journal of Industrial Organization* 27 (5): 567–571. ISSN: 0167-7187.

———. 2012. 'Price-dependent demand in spatial models'. *The BE Journal of Economic Analysis & Policy* 12 (1).

Harding, Rebecca, Kirk Haywood and Niels Strazdins. 2022. *Quantitative Analysis of the Move to Paperless Trade.* Technical report. London: The Commonwealth Secretariat. https://thecommonwealth.org/quantitative-analysis-move-paperless-trade.

Hart, Oliver, and John Moore. 1990. 'Property Rights and the Nature of the Firm'. *Journal of political economy* 98 (6): 1119–1158. ISSN: 0022-3808.

He, Chusu, Alistair Milne and Markos Zachariadis. 2022. 'Central bank digital currencies and international payments'. London and New York. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4112382.

Iyengar, Garud, Fahad Saleh, Jay Sethuraman and Wenjun Wang. 2023. 'Economics of Permissioned Blockchain Adoption'. *Management Science* 69 (6): 3415–3436. ISSN: 0025-1909. https://doi.org/10.1287/mnsc.2022.4532. https://doi.org/10.1287/mnsc.2022.4532.

Jones, Charles I, and Christopher Tonetti. 2020. 'Nonrivalry and the Economics of Data'. *American Economic Review* 110 (9): 2819–58.

Matsumura, Toshihiro, and Makoto Okamura. 2006a. 'A note on the excess entry theorem in spatial markets'. *International Journal of Industrial Organization* 24 (5): 1071–1076. ISSN: 0167-7187.

———. 2006b. 'Equilibrium number of firms and economic welfare in a spatial price discrimination model'. *Economics Letters* 90 (3): 396–401. ISSN: 0165-1765.

McIntyre, Chris, Sultan Alsubaihin, Simon Bartletta, Philip Bianchi, Joe Carrubba, Peter Czerepak, Dean Frankle et al. 2023. *Global Asset Management 2023—21st Edition.* Technical report. Boston Consulting Group. https://web-assets.bcg.com/c8/97/bc0329a046f89c7faeef9ab6a877/bcg-global-asset-management-2023-may-2023.pdf.

Morton-Scott, Fiona, Pascal Bouvier, Ariel Ezrachi, Bruno Jullien, Roberta Katz, Gene Kimmelman, A Douglas Melamed and Jamie Morgenstern. 2019. 'Committee for the study of digital platforms: Market structure and antitrust subcommittee report'. *Chicago: Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business.*

OECD. 2016. *Maximising the economic and social value of data: Understanding the Benefits and Challenges of Enhanced Data Access.* Technical report. OECD. https://one.oecd.org/document/DSTI/CDEP(2016)4/en/pdf.

Ordover, Janusz A, Garth Saloner and Steven C Salop. 1990. 'Equilibrium vertical foreclosure'. *The American Economic Review,* 127–142.

Polasik, Michał, Agnieszka Huterska, Rehan Iftikhar and Štěpán Mikula. 2020. 'The impact of Payment Services Directive 2 on the PayTech sector development in Europe'. *Journal of Economic Behavior and Organization* 178:385–401. ISSN: 0167-2681.

Posner, Eric A, and E Glen Weyl. 2019. *Radical Markets.* Princeton University Press.

Reisinger, Markus, and Monika Schnitzer. 2012. 'Successive oligopolies with differentiated firms and endogeneous entry'. *The Journal of Industrial Economics* 60 (4): 537–577.

Rochet, Jean-Charles, and Jean Tirole. 2003. 'Platform competition in two-sided markets'. *Journal of the european economic association* 1 (4): 990–1029.

Rochet, Jean-Charles, and Julian Wright. 2010. 'Credit card interchange fees'. *Journal of Banking and Finance* 34 (8): 1788–1797. ISSN: 0378-4266.

Salop, Steven C. 1979. 'Monopolistic competition with outside goods'. *The Bell Journal of Economics,* 141–156.

Varian, Hal. 2018. 'Artificial Intelligence, Economics, and Industrial Organization'.

Veldkamp, Laura, and Cindy Chung. 2019. 'Data and the aggregate economy'. *Journal of Economic Literature.*

Verdier, Marianne. 2011. 'Interchange fees in payment card systems: A survey of the literature'. *Journal of Economic Surveys* 25 (2): 273–297. ISSN: 0950-0804.

Vickrey, William Spencer. 1964. *Microstatics.* Harcourt, Brace / World.

Williamson, Oliver E. 1981. 'The modern corporation: origins, evolution, attributes'. *Journal of economic literature* 19 (4): 1537–1568. ISSN: 0022-0515.

———. 1991. 'Comparative economic organization: The analysis of discrete structural alternatives'. *Administrative science quarterly,* 269–296. ISSN: 0001-8392.