Entry-Deterring Agency*

Simon Loertscher[†] Andras Niedermayer[‡]

December 19, 2017

Abstract

We provide a model in which an intermediary can choose between wholesale or agency. The possibility that buyers and sellers transact directly limits his market power and, thus, creates incentives for him to deter the emergence of bilateral exchanges. In equilibrium, the intermediary chooses agency and thereby pre-empts the emergence of a competing bilateral exchange if the matching technology of the competing exchange is sufficiently efficient. Whenever agency is chosen in equilibrium, consumer surplus and social welfare decrease and both listing and transaction prices increase compared to wholesale. The predictions of our model are broadly consistent with empirical evidence.

Keywords: Entry deterrence, brokerage, endogenous market structure, platforms, competing exchanges.
JEL-Classification: C72, D41, D43, L13.

---

*This paper has benefited from comments by and discussions with Luis Cabral, Renato Gomes, Stephen King, Stephan Lauermann, Joao Montez, Martin Peitz, Michael Riordan, Larry Samuelson, Artyom Shneyerov and seminar audiences at the University of Mannheim. We also thank participants of AETW 2013 at UQ, EARIE 2013 in Evora, SAET 2013 in Paris, the CLEEN Workshop in 2013, IIOC 2013 in Boston, the MaCCI Annual Conference in 2013 in Mannheim, the SFB TR 15 meeting 2013 in Bonn, the 2015 ATE Symposium in Auckland, the Multi-Sided Platforms Workshop in 2015 in Singapore, and the 2015 Workshop on Consumer Search in Bad Homburg for helpful comments. The second author gratefully acknowledges funding by the Deutsche Forschungsgesellschaft through SFB TR 15 and project PE 813/2-2. A previous version of the paper circulated under the title "Predatory Platforms".

[†]Economics Department, University of Melbourne, Level 4, FBE Building, 111 Barry Street, Parkville 3010, Victoria, Australia. Email: simonl@unimelb.edu.au

[‡]Université Paris-Dauphine, PSL Research University, LEDa, 75016 Paris, France. Email: afniedermayer@gmail.com.

1 Introduction

Merchants, middlemen, and market makers have always assumed important economic roles. For just as long, they have been viewed with suspicion for doing so.[1] Traditional examples of these intermediaries are real estate brokers, employment agencies, and used car dealers. Following the entry into the 21st century, the emergence of the Internet and e-commerce has witnessed the creation of intermediaries such as Amazon, eBay, Apple with its i-products and services, AirBnB, Expedia, Uber, and Booking.com. Resonating with traditional suspicions, these Internet platforms are viewed as controversial by policy makers and the public, with a major concern being that they extract excessive rents from the two sides of the market.[2,3]

Unlike monopoly producers, an intermediary faces the competitive threat of

[1] See, for example, Karl Marx (Capital, Vol.1, Chapter 29, p.744), who wrote that "it is evident here how in all spheres of social life the lion's share falls to the middleman. In the economic domain, e.g., financiers, stock-exchange speculators, merchants, shopkeepers skim the cream...".

[2] In Europe, online hotel booking sites Booking.com and Expedia have come under the scrutiny of antitrust authorities, in the wake of which these platforms have abandoned their "favorite nation" clauses. The Apple e-books case is another prominent example of an intermediary – Apple as the provider of the iPad, contracting with publishing companies – alleged of and fined for aiming to foreclose an alternative exchange. Most recently, it has been argued that Uber is trying to "corner" capital required to subsidize drivers in emerging markets like India and China; for example, see http://www.nytimes.com/2016/06/21/business/dealbook/why-uber-keeps-raising-billions.html?_r=0

[3] While new Internet intermediaries have been particularly prominent, even the modern-day debate on excessive rent extraction by intermediaries goes back much longer, through the investigations of real-estate brokers by the Department of Justice and the conviction of Sotheby's and Christie's for collusion in the 2000s. Competitive concerns were also at the source of investigations of the US Department of Justice on real-estate brokerage in 1983 and 2007 (see, e.g., DOJ, 2007) and allegations of and convictions for collusion by the auction houses Sotheby's and Christie's (see, e.g., Ashenfelter and Graddy, 2005). The International Labor Organization's call for a ban of private fee-charging employment agencies, to the referendum on banning private labor market intermediaries in Washington State in 1914. The International Labor Organization of the United Nations passed a convention in 1949 that banned private fee-charging employment agencies, to be revoked by a second convention by the International Labor Organization only as late as 1997. See conventions C96 and C181 of the International Labor Organization, C96 Fee-Charging Employment Agencies Convention (Revised), 1949, http://www.ilo.org/ilolex/cgi-lex/convde.pl?C096, C181 Private Employment Agencies Convention, 1997, http://www.ilo.org/ilolex/cgi-lex/convde.pl?C181. Similar developments were also present in the US: in 1914 a referendum in Washington State banned private labor market intermediaries, a law that was later overruled by the US Supreme Court. See Adams v. Tanner, 244 U.S. 590 (1917) and a description of the controversy in Foner (1965, p. 177-185).

being circumvented by its customers even if it is a monopoly: in principle, buyers and sellers can always trade directly with each other. From a consumer surplus perspective, this competitive threat is welcome because, all else being equal, it reduces the intermediary's market power.[4] However, it also provides the intermediary with incentives to (potentially inefficiently) deter the emergence of a competing bilateral exchange.

In this paper, we focus on two widely used models of intermediation, called *wholesale* and *agency*, which feature prominently in recent antitrust debates and cases (see, e.g., Johnson, 2013, 2017). We show that a monopoly intermediary optimally uses agency if and only if the competitive threat is sufficiently strong. It does so to pre-empt the emergence of a bilateral exchange and thereby reduces both social surplus and consumer surplus while increasing the final prices buyers face and pay.

A sketch of our model and the forces at work is as follows. We assume a continuum of buyers and sellers with heterogeneous valuations and costs for a homogeneous good. Buyers have unit demand and sellers have unit capacities. In the absence of a bilateral exchange, wholesale with appropriately chosen prices for buyers and sellers yields higher profits for the intermediary than agency. However, because it induces a spread between buyers' and sellers' prices, traders with values and costs in between these prices cannot trade with the intermediary but can, to their mutual benefit, trade with each other. Wholesale is thus associated with an active bilateral exchange that improves the outside option of not trading with the intermediary even for traders who in equilibrium join the intermediary. Consequently, the bilateral exchange puts pressure on the spread the intermediary can charge, and the more so, the more efficient the bilateral exchange. In contrast, under agency the intermediary does not take a position but rather has the sellers set prices, randomly matching them to buyers and charging a percentage fee on the price whenever a transaction occurs. Because sellers are heterogeneous with respect to their costs, agency is associated with a non-degenerate distribution of transaction prices, and more importantly, with overlapping supports in the valuations and costs of buyers and sellers who join the intermediary under agency. Because of this overlap, no buyers and sellers are left who could trade to their mutual benefit in a bilateral exchange. Consequently, the intermediary can use

[4] See, for example, Gehrig (1993).

agency to deter the emergence of the bilateral exchange.

Whether such entry deterrence via agency is profitable depends on additional details of the setup. We show that for a variety of settings it is profitable, provided the bilateral exchange's matching technology is sufficiently efficient. We also show that whenever entry deterrence pays off, social surplus and consumer surplus (defined as the aggregate of buyers' and sellers' surplus) decrease, and buyers' prices increase.

A policy implication of our results is that banning the agency model can in some cases be welfare improving. This has indeed been the Department of Justice's decision in the case against Apple's pricing of e-books. In other cases, banning the agency model outright may not be feasible: e.g. for real estate brokerage, a wholesale model may not be practical or feasible. An alternative solution is to require flat fees paid independent of whether there is a transaction, which can be shown to be equivalent to a wholesale model of intermediation.[5] The Department of Justice's 2007 settlement with the Realtor's Association was a milder form of this alternative: while the agency model was not banned, the Realtor's Association was required to abandon practices that discriminated against flat fee agents.

To the best of our knowledge, ours is the first paper to analyze an intermediary's incentives and options to deter entry by an alternative exchange. The emergence of the Internet and e-commerce has led to new intermediaries and to an upsurge of research on two-sided markets. Starting with the pioneering work by Caillaud and Jullien (2001, 2003), Rochet and Tirole (2003, 2006), Anderson and Coate (2005) and Armstrong (2006), this literature has primarily focused on the monopoly platform or competition between platforms, abstracting from the exact mechanisms the platforms employ to generate surplus, from traders' options to circumvent the platform, and the platform's incentives to prevent them from so doing.[6] Two notable exceptions that explicitly analyze the

[5]The equivalence between flat participation fees and the wholesale model have been shown e.g. in Niedermayer and Shneyerov (2014) for arbitrary distributions of traders' types. See also (Belleflamme and Peitz, 2015, Chapter 22) for equivalence between the wholesale model and a flat *transaction* fee (i.e. a fee that is only paid in case of a transaction occurring), however, this equivalence only holds for uniform distributions.

[6]That circumventing the intermediary is an option will not come as a surprise to any academic who has ever hired a research assistant from a developing country using an online intermediary.

platform's trading mechanism are Gomes (2014) and Niedermayer and Shneyerov (2014). Competition between wholesale intermediaries and alternative exchanges fares prominently in the works of Rubinstein and Wolinsky (1987), Stahl (1988), Gehrig (1993), Spulber (1996), Bloch and Ryder (2000), Rust and Hall (2003), Loertscher (2007) and Neeman and Vulkan (2010), which, however, do not address intermediaries' incentives and options to drive out the competing exchanges.

Our paper builds on the existing literature. Participation fees, which are equivalent to the wholesale model, have been analyzed by Niedermayer and Shneyerov (2014), while Loertscher and Niedermayer (2017a,b) analyze brokers' fee structures from an optimal pricing perspective for the agency model. These papers assume that there is no competing bilateral exchange.[7]

Our paper also shares features with the Industrial Organization literature on predation, in particular with the strand of literature where predatory pricing is made credible, without invoking differential access to financing, by the presence of learning-by-doing as analyzed by Cabral and Riordan (1994, 1997), Besanko, Doraszelski, Kryukov, and Satterthwaite (2010), and Besanko, Doraszelski, and Kryukov (2014). In this strand of literature, the cumulative effects of learning-by-doing render predatory pricing credible. In our model it is the threat of the emergence of an alternative exchange between buyers and sellers that makes agency credible for deterring entry. However, the incentives and methods used for entry deterrence in the literature mentioned above and the present paper are different and complementary. A novel feature of our model is that a strategic player, the intermediary, may predate the emergence of an alternative exchange who is not a player.

Recently, Edelman and Wright (2015) developed a model that exhibits excessive intermediation in that more agents are active than would be under Walrasian conditions, but the mechanisms in the two papers are quite different. In their paper, it is price coherence that induces inefficient outcomes whereas in our model it is the opposite – the freedom of sellers to set prices and the non-degenerate price distributions this induces. The two papers thus complement each other. Studying credit card fees, Bourguignon, Gomes, and Tirole (2014) analyze alternative regulatory interventions in a setting with one-sided incomplete informa-

[7]Equivalently, these papers can be viewed as having a bilateral exchange, for which the payoffs are exogenously given and cannot be influenced by the intermediary.

tion. Biglaiser and Li (2016) study a moral hazard model in which the presence of a middleman who uses (in our terminology) wholesale can increase or decrease welfare. In contrast, we study an adverse selection model in which the intermediary is always present but in equilibrium chooses an inefficient mechanism to reduce traders' outside options. While the themes of our paper and theirs are similar, the setups, mechanisms, and outcomes are thus quite different, which renders the papers complementary. Our analysis has a similar flavor as Johnson (2017) because both analyze agency versus wholesale. However, the setup and the question are very different: our focus is on the entry deterrence effect of the agency model, which is an effect that cannot occur in Johnson (2017) because in his setup it is, by assumption, impossible for the supplier to circumvent the retailer.

Our paper also contributes to the literature on intermediation and specifically on real-estate brokerage such as the papers by Yavas (1996), Hsieh and Moretti (2003), Rutherford, Springer, and Yavas (2005), Levitt and Syverson (2008), Hendel, Nevo, and Ortalo-Magné (2009) and Loertscher and Niedermayer (2017a,b). The agency model is difficult to reconcile with a principal-agent perspective, but consistent with an optimal pricing perspective.[8] In the companion paper (Loertscher and Niedermayer, 2017b), we also take the theory to the data and estimate demand and supply under the assumption that payoffs from a bilateral exchange market are exogenous (or that no bilateral exchange market exists). Yavas (1996) has investigated whether real-estate brokers have an incentive to induce inefficiently many trades by matching buyers and sellers "horizontally" rather than "vertically" (i.e. high-cost sellers to high-value buyers rather than inducing trades among the most efficient pairs) because brokers earn commissions that depend on the price and not on the surplus trades generate. While the specifics of the models and mechanisms are different, we obtain a similar result: that under agency horizontal matchings occur with positive probability. This contrasts with the wholesale model, which induces trade among the most efficient buyers and sellers with probability 1.

In a wider sense, this paper also relates to Fudenberg and Tirole (2000)'s results of entry deterrence in a dynamic market with network externalities.[9]

[8] For a discussion, see Loertscher and Niedermayer (2017a,b).

[9] The key difference is that in Fudenberg and Tirole (2000) the focus is on initial low pricing by a monopolist in the presence of network externalities, which results in entry deterrence in

The remainder of this paper is organized as follows. Section 2 describes the setup. In Section 3, we derive the conditions under which entry deterrence via agency is profitable, and the effects of profitable entry deterrence on social surplus, consumer surplus, and prices. Section 4 contains discussion of extensions and policy implications, and Section 5 concludes. The Appendix contains omitted proofs and an alternative model with binary types, which lends itself to a dynamic random matching extension.

2 Setup

We consider a one-period model with a continuum of buyers and a continuum of sellers, each with mass 1. Buyers have unit demand and sellers have unit capacities. All agents are risk-neutral, have quasilinear preferences, and outside options with value 0. Buyers draw their valuations v independently from the *(generalized) Pareto distribution* $F(v)$ and sellers draw their costs c independently from the generalized Pareto distribution $G(c)$ with supports $[0, 1]$ and elasticity parameter σ:

$$F(v) = 1 - (1 - v)^{\sigma} \quad \text{and} \quad G(c) = c^{\sigma},$$

whose densities are $f(v) = \sigma(1 - v)^{\sigma-1}$ and $g(c) = \sigma c^{\sigma-1}$, respectively. Generalized Pareto distributions are analytically convenient because they permit closed-form solutions and yet allow comparative statics with respect to the elasticities of supply and demand. The elasticity of supply is $\eta_s(c) = cg(c)/G(c) = \sigma$ and the elasticity of demand $\eta_d(v) = (1 - v)f(v)/(1 - F(v)) = \sigma$.[10] We assume that the elasticities of demand and supply are weakly larger than 1; that is, $\sigma \geq 1$. We defer discussing the relevance of this assumption to the section after the equilibrium analysis. By symmetry of F and G, the Walrasian price satisfying $1 - F(p) = G(p)$ is $p = 1/2$. Observe that F and G are uniform distributions for $\sigma = 1$.

later periods due to an installed base. In our paper, the focus is not the level of pricing, but the type of pricing: wholesale versus agency.

[10] We are simplifying terminology here. Technically speaking, F is a finite support generalized Pareto distribution and G is a mirrored finite support generalized Pareto distribution. While the elasticity of supply is the standard definition, the elasticity of demand is mostly defined as $vf(v)/(1 - F(v))$, but for dealing with markets with intermediaries, the above definition is more convenient.

There is an intermediary that is endowed with a technology that allows a higher matching probability for buyers and sellers than if buyers and seller were to search for each other bilaterally.

Our analysis is motivated by the prevalence of two different trading modes of platforms used in the real world. These are sometimes called **wholesale** and **agency**.[11] Under *wholesale*, the intermediary takes a position and stands ready to trade with buyers and sellers. In contrast, under *agency* the intermediary does not take a position but rather has the seller set a price, on which it charges a percentage fee $b \in [0, 1]$.[12] Following the two-sided markets literature, we assume that the intermediary determines the trading mode and its prices and fees, taking as given the elasticity of supply and demand. Accordingly, we model wholesale as posting two prices, a price $p_B \in [0, 1]$ for buyers and a price $p_S \in [0, 1]$ for sellers. At these prices, the intermediary stands ready to sell and buy, respectively. When the intermediary chooses agency, we assume that he sets a percentage fee $b \in [0, 1]$, which the seller has to pay if and only if a transaction occurs at the price the seller sets. Under agency, buyers and sellers are matched one-to-one at random. Traders on the long side (if there is a long side) are randomly rationed.

Buyers and sellers can trade via this intermediary (where they are matched with probability 1), choose to remain inactive, or try to trade in a bilateral matching market in which the matching probability is $\lambda \in [0, 1]$ if the mass of buyers and sellers is the same. In the main model, we assume *fixed-price bargaining* in the non-intermediated random matching market. That is, a buyer and seller who are matched either trade at the Walrasian price of $p = 1/2$ or do not trade at all.[13] Consequently, the expected payoffs of a buyer with value v and of a seller with cost c from participating in the bilateral matching market,

[11] This language was used by the court in *United States of America v Apple*. It has also found its way into the economics literature; see, for example, Johnson (2013) or De Los Santos and Wildenbeest (2016).

[12] What is called agency here is referred to as fee-setting in Loertscher and Niedermayer (2017a,b).

[13] Alternatively, and in many ways equivalently, one could stipulate that, in the absence of an intermediary that uses wholesale, fixed-price bargaining occurs at the Walrasian price of $1/2$ while in the presence of a wholesale intermediary who trades at prices p_B and p_S fixed price bargaining occurs at the price $(p_B + p_S)/2$.

denoted respectively $V_B(v)$ and $V_S(c)$, are

$$V_B(v) = \lambda \max\left\{v - \frac{1}{2}, 0\right\} \quad \text{and} \quad V_S(c) = \lambda \max\left\{\frac{1}{2} - c, 0\right\}.$$

Observe that for $\lambda < 1$, $V_B'(v)$ and $V_S'(c)$ are strictly less than 1 everywhere where they are defined.[14] In Section 4.3, we show that the insights from the fixed-price bargaining model extend to alternative specifications of bargaining, which shows that the assumption of fixed-price bargaining used in the main part of this paper is indeed only for expositional simplicity.

The timing of the game is as follows. In Stage 1, the intermediary chooses the mode of operation (wholesale or agency) and, conditional on that, the optimal prices and the optimal percentage fee, respectively. The mode of operation and the prices or fee are observed by all agents. In Stage 2, all buyers and sellers simultaneously decide whether to join the intermediary, the non-intermediated random matching market, or to remain inactive. These choices are mutually exclusive. We assume that agents who trade with probability 0 at a given market (i.e. at the intermediary's exchange or in the matching market) do not enter this market.[15] We defer further discussions of equilibrium selection to the next section.

3 Equilibrium

We now analyze the model laid out above. We first derive the equilibrium outcome and then analyze effects on social surplus and consumer surplus.

3.1 Wholesale

Under wholesale the profit-maximizing intermediary operates at a positive bid-ask spread by posting a buyers' price p_B that exceeds the sellers' price p_S. Consequently, buyers and sellers with values and costs in between p_S and p_B cannot trade via the intermediary but may benefit from trading with each other. This

[14]These expressions are exact when the mass of buyers and sellers in the matching market are the same. If there are, say, more buyers than sellers in the bilateral exchange, $V_S(c)$ is as displayed while $V_B(V)$ decreases in proportion to excess demand. Obviously, this does not affect the conclusion that the slopes are less than 1.

[15]This can be justified as the limit of a slightly richer model in which a positive fixed cost of market participation goes to 0.

opens the scope for a bilateral exchange with an expected payoff of $V_B(v)$ for a buyer and $V_S(c)$ for a seller. Of course, the option of participating in the bilateral exchange not only provides buyers with values $v \in (p, p_B]$ and sellers with costs $c \in [p_S, p)$ with a positive payoff, where p is the Walrasian price, but also gives a positive outside option to those more efficient traders who, in equilibrium, trade via the intermediary. Therefore, the willingness to pay for trading via the intermediary of buyers with value v and the reservation price for selling to the intermediary for sellers with costs c are $v - V_B(v) = (1-\lambda)v + \lambda/2$ and $c + V_S(c) = (1-\lambda)c + \lambda/2$ for $v \geq p$ and $c \leq p$.

If the intermediary sets the price p_B for buyers, the buyer with value $\overline{v}$ such that $\overline{v} - p_B = V_B(\overline{v})$ will be indifferent between trading with the intermediary and participating in the bilateral exchange. Because for any $\lambda < 1$ the slope of $V_B(v)$ is less than 1, all buyers with higher values will strictly prefer trading with the intermediary while buyers with values $v \in (p, \overline{v})$ will be strictly better off in the bilateral exchange. That is, single-crossing holds. Solving $\overline{v} - p_B = V_B(\overline{v})$ for $\overline{v}$ yields $\overline{v} = (p_B - \lambda/2)/(1-\lambda)$. Analogously, given a price p_S set by the intermediary for sellers, the seller with cost $\underline{c}$ satisfying $p_S - \underline{c} = V_S(\underline{c})$ will be indifferent between trading with the intermediary and participating in the bilateral exchange. Because of single-crossing, all sellers with lower costs will strictly prefer trading with the intermediary while all sellers with costs $c \in (\underline{c}, p)$ are better off in the bilateral matching market. Solving $p_S - \underline{c} = V_S(\underline{c})$ for $\underline{c}$ gives $\underline{c} = (p_S - \lambda/2)/(1-\lambda)$. Consequently, the intermediary's profit under wholesale with prices p_B and p_S is $(p_B - p_S)\min\{1 - F(\overline{v}), G(\underline{c})\}$, where $\min\{1 - F(\overline{v}), G(\underline{c})\}$ is the quantity he trades. Because there is no point choosing p_B and p_S to induce an imbalance in demand and supply, the intermediary's price will be such that $1 - F(\overline{v}) = G(\underline{c})$, yielding $p_B = 1 - p_S$. Thus, the intermediary's profit maximization problem under wholesale can be expressed as setting $p_B = 1 - p_S$ and choosing p_S to maximize

$$(1 - 2p_S)G\left(\frac{p_S - \lambda/2}{1-\lambda}\right) = (1 - 2p_S)\left(\frac{p_S - \lambda/2}{1-\lambda}\right)^{\sigma}, \tag{1}$$

where the equality follows from our assumption of generalized Pareto distributions. It is straightforward to verify that the optimal prices, denoted $p_S^*(\sigma, \lambda)$ and $p_B^*(\sigma, \lambda)$, satisfy

$$p_S^*(\sigma, \lambda) = \frac{\sigma + \lambda}{2(1+\sigma)} \quad \text{and} \quad p_B^*(\sigma, \lambda) = 1 - \frac{\sigma + \lambda}{2(1+\sigma)}.$$

Plugging these values into (1), one sees that the intermediary's maximized profit under wholesale is

$$\Pi_W^*(\sigma,\lambda) = (1-\lambda)\left(\frac{\sigma}{2}\right)^{\sigma}\left(\frac{1}{1+\sigma}\right)^{1+\sigma}.$$

For example, for uniform distributions, we have $\Pi_W^*(1,\lambda) = (1-\lambda)/8$. For any σ, $\Pi_W^*(\sigma,\lambda)$ monotonically decreases in λ and satisfies $\Pi_W^*(\sigma,1) = 0$. This is the downward pressure on profits under wholesale mentioned in the Introduction. The observation that

$$\Pi_W^*(\sigma,\lambda) = (1-\lambda)\Pi_W^*(\sigma,0) \tag{2}$$

will prove useful below. The special case $\lambda = 0$ can be interpreted as no bilateral market existing, since the matching probability λ is so low that no-one finds a trading partner there.

3.2 Agency

Under agency, the intermediary sets a percentage fee b and buyers and sellers joining the intermediary will be randomly matched. Assume temporarily that all buyers and sellers join the intermediary.

Observing b and his own cost c, the seller sets the price p to maximize

$$((1-b)p - c)(1-F(p)). \tag{3}$$

The maximizer is

$$p(b,c) = \frac{1}{1+\sigma} + \frac{\sigma}{1+\sigma}\frac{c}{1-b}.$$

Notice that the lowest price that will be set is $\underline{p} := p(b,0) = 1/(1+\sigma)$, which is independent of b, while the highest price is 1, which will be set by the seller with cost $1-b$. Consequently, buyers with values below $\underline{p}$ and sellers with costs above $1-b$ can never trade in equilibrium under agency. Because we assume that traders who do not benefit from a given exchange do not enter it, these agents (i.e. buyers with values below $\underline{p}$ and sellers with costs above $1-b$) will not join the intermediary under agency. However, this does not change the fact that the seller with cost $c \in [0, 1-b]$ optimally sets the price $p(b,c)$ because $p(b,c)$ is obviously also the maximizer of $((1-b)p-c)(1-F(p))/(1-F(\underline{p}))$, where $(1-F(p))/(1-F(\underline{p}))$ is the probability of selling at a price $p \geq \underline{p}$ conditional on buyers with values below $\underline{p}$ not being active: The truncation $v \geq \underline{p}$ only yields to a multiplication of the seller's profit function provided in (3) by a constant.

Given this, the intermediary's profit under agency with fee b is

$$\Pi_A(b) = b \min\{1 - F(\underline{p}), G(1-b)\} \frac{\int_0^{1-b} (p(b,c)(1 - F(p(b,c))))dG(c)}{(1 - F(\underline{p}))G(1-b)},$$

where $f(v)/(1 - F(\underline{p}))$ and $g(c)/G(1-b)$ are the conditional densities of active traders, integrated out in the case of buyers, and $\min\{1 - F(\underline{p}), G(1-b)\}$ is the mass of matches. As shown in the proof of the following proposition, $\Pi_A(b)$ is maximized at $b = 1/(1+\sigma)$.

Observe also that with $b = 1/(1+\sigma)$, we have $\underline{p} \leq 1 - b$ for any $\sigma \geq 1$. That is, if all buyers and sellers who can profitably trade via the intermediary under agency join the intermediary, there are no agents left who can trade to their mutual benefit in the random matching market. This brings us to the key for why agency can deter the emergence of a competing exchange: By setting $b = 1/(1+\sigma)$, the intermediary can pre-empt (or predate) the emergence of a bilateral random matching market. The following proposition, whose second part we have just established, summarizes the equilibrium outcome under agency. We prove the first part in the Appendix.

Proposition 1. *The optimal fee under agency is $b = 1/(1+\sigma)$, yielding a intermediary profit of*

$$\Pi_A^*(\sigma) = \left(\frac{\sigma}{1+\sigma}\right)^\sigma B(1+\sigma, 1+\sigma), \tag{4}$$

where $B(x,y) := \int_0^1 t^{x-1}(1-t)^{y-1}dt$ is the Beta function. Moreover, agency pre-empts the emergence of the bilateral random matching market

For example, for uniform distributions we have $\Pi_A^*(1) = 1/12$. This is evidently larger than $\Pi_W^*(1,\lambda) = (1-\lambda)/8$ if and only if $\lambda > 1/3$, so that for uniform distributions one would expect agency to be the equilibrium choice of the intermediary if and only if $\lambda > 1/3$. Because it pre-empts the emergence of the bilateral exchange, $\Pi_A^*(\sigma)$ is independent of the matching technology parameter λ.

3.3 Equilibrium Entry Deterrence

At $\lambda = 0$, wholesale is more profitable than agency. This is true because absent competing exchange (i.e., $\lambda = 0$) wholesale with the prices $p_B^*(\sigma, 0) =$

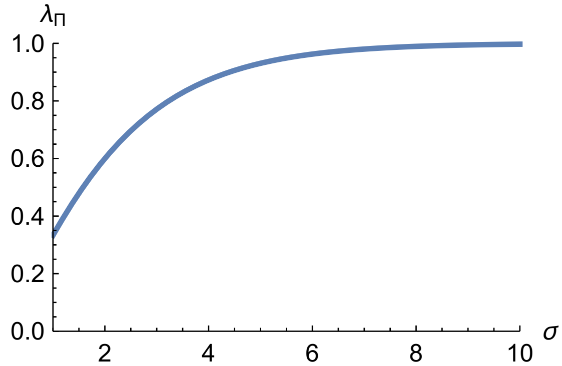


Figure 1: Agency is more profitable than wholesale if and only if $\lambda > \lambda_\Pi(\sigma)$.

$1 - \sigma/((2(1+\sigma))$ and $p_S^*(\sigma, 0) = \sigma/(2(1+\sigma))$ is the profit-maximizing mechanism for the intermediary facing a continuum of buyers and sellers as we show in Appendix B. Because $\Pi_W^*(\sigma, \lambda)$ monotonically decreases in λ and satisfies $\Pi_W^*(\sigma, 1) = 0$ while $\Pi_A^*(\sigma)$ is positive and independent of λ, it follows that for any given $\sigma \geq 1$, there is a unique $\lambda_\Pi(\sigma) \in (0, 1)$ such that $\Pi_W^*(\sigma, \lambda_\Pi(\sigma)) = \Pi_A^*(\sigma)$. Combining this with (2), we get

$$\lambda_\Pi(\sigma) = 1 - \frac{\Pi_A^*(\sigma)}{\Pi_W^*(\sigma, 0)} = 1 - 2^\sigma(1+\sigma)\sigma B(1+\sigma, 1+\sigma).$$

Consequently, agency is more profitable for the intermediary, that is, $\Pi_A^*(\sigma) > \Pi_W^*(\sigma, \lambda)$ if and only if $\lambda > \lambda_\Pi(\sigma)$. Figure 1 plots $\lambda_\Pi(\sigma)$. For example, for uniform distributions we have $\lambda_\Pi(1) = 1/3$ as noted after Proposition 1.

Proposition 2. *In equilibrium, the intermediary chooses wholesale for $\lambda < \lambda_\Pi(\sigma)$ and agency for $\lambda > \lambda_\Pi(\sigma)$, where $\lambda_\Pi(\sigma) \in (0, 1)$ for $\sigma \in [1, \infty)$.*

Because $\lambda_\Pi(\sigma)$ is an increasing function of σ, this means that entry-deterring agency is less of a concern in more elastic environments in the sense that it occurs for a smaller subset of the parameter space. More elastic demand and supply harm the intermediary under both models as the ratio of profit over Walrasian surplus decreases in σ for both agency and wholesale. However, the ratio decreases much faster under agency and goes to 0 as σ goes to infinity while the ratio is bounded by $2/e \approx 0.73$ for wholesale.

Other equilibria under agency exist. For example, equilibria with less buyer entry can be supported. If all buyers expect the lowest price to be some $\underline{p}' > \underline{p}$, no buyer with value below $\underline{p}'$ will enter, and no seller will optimally set a price below $\underline{p}'$. This multiplicity of equilibria is inevitable in models with two-sided markets. However, any such equilibrium requires coordinated beliefs by agents on both sides of the market and buyers with values $v \in (\underline{p}, \underline{p}')$ to abstain even though there is an equilibrium in which they would expect a positive payoff from joining the intermediary. We therefore focus on equilibria in which all buyers with values above $\underline{p}$ (and all sellers with costs below $1-b$) join the intermediary.

Under wholesale, there is also an equilibrium in which no agent joins the bilateral exchange. However, traders with values and costs between p_S and p_B obtain a payoff of 0 in this equilibrium whereas they would obtain a positive expected surplus if they all participated in the bilateral exchange. Consequently, the equilibrium without an active bilateral exchange hinges on a failure to coordinate by these traders, which is arguably not all that compelling. The same is not true for the entry-deterring equilibrium under agency: The traders who join the intermediary obtain a payoff that exceeds the payoff of not being active at all.[16]

3.4 Social and Consumer Surplus Effects

Social surplus is defined as the equally weighted gains from trade of all buyers and sellers and of the profit of the intermediary. Under wholesale, social surplus, denoted $S_W(\sigma, \lambda)$, is

$$S_W(\sigma, \lambda) = \frac{\lambda(1+\sigma)^{1+\sigma} + (1-\lambda)\sigma^\sigma(1+2\sigma)}{2^\sigma(1+\sigma)^{2+\sigma}}.$$

For example, for uniform distributions, we have $S_W(1, \lambda) = (3+\lambda)/16$. In contrast, social surplus under the agency model, denoted $S_A(\sigma)$, is

$$S_A(\sigma) = \sigma^\sigma \frac{1+3\sigma}{(1+\sigma)^{1+\sigma}} B(1+\sigma, 1+\sigma).$$

[16]Whether they could jointly benefit depends in intricate ways on the assumptions one makes about how such joint deviations would occur. For example, it is not the case that all agents are necessarily worse off in the entry-deterring equilibrium under agency than if all traders who can benefit from joining the bilateral exchange joined it. For example, with uniform distributions sellers with cost 0 get a payoff of 1/4 under agency and a payoff $V_S(0) = \lambda/2$ from participating in the bilateral exchange. For $\lambda \in (1/3, 1/2)$, agency is thus an equilibrium outcome and this type of seller prefers agency with no bilateral exchange to a "pure" bilateral exchange in which all buyers with $v \geq 1/2$ and all sellers with $c \leq 1/2$ participate.

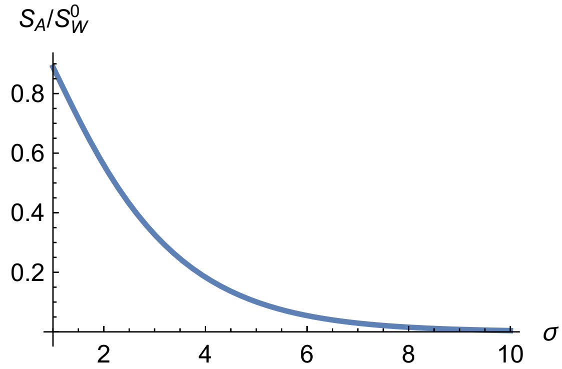


Figure 2: Ratio of social surplus under agency and wholesale for $\lambda = 0$.

For uniform distributions, we have $S_A(1) = 1/6$.

Social surplus under wholesale increases with λ and is therefore at its minimum at $\lambda = 0$. Figure 2 plots the ratio $S_A(\sigma)/S_W(\sigma, 0)$ for $\sigma \in [1, 10]$. It can be seen from the figure that this ratio is less than 1 for all values of σ and decreases with σ. This implies that the wholesale model generates higher social surplus than the agency model for all values of λ. Formally, the ratio $S_A(\sigma)/S_W(\sigma, 0)$ is given by

$$\frac{S_A(\sigma)}{S_W(\sigma, 0)} = \frac{2^\sigma(1+\sigma)(1+3\sigma)\Gamma(1+\sigma)^2}{(1+2\sigma)\Gamma(2+2\sigma)}, \tag{5}$$

where $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$ is the Gamma function. In the proof of the following proposition, we show that for all all $\sigma \geq 1$, $S_A(\sigma)/S_W(\sigma, 0)$ is less than 8/9.

Proposition 3. *For all $\sigma \geq 1$, social surplus under wholesale exceeds social surplus under agency; that is,*

$$S_W(\sigma, \lambda) > S_A(\sigma).$$

Proposition 3 implies the following corollary.

Corollary 1. *Agency decreases social welfare.*

Another question of interest concerns effects on consumer surplus, which are at the center of attention for many antitrust authorities around the globe. Defining consumer surplus as social surplus minus the intermediary's profit, consumer surplus under agency, denoted $CS_A(\sigma)$, is independent of λ and given as $CS_A(\sigma) = S_A(\sigma) - \Pi_A^*(\sigma)$. Under wholesale, consumer surplus, denoted

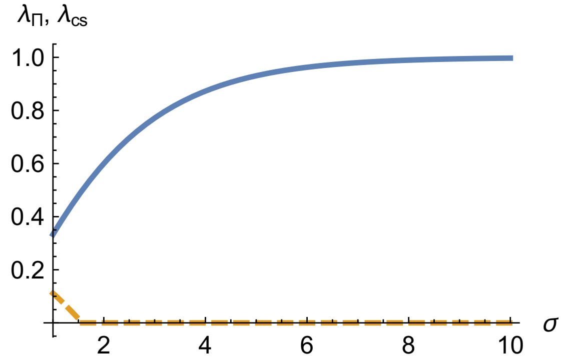


Figure 3: The critical values $\lambda_\Pi(\sigma)$ (solid) and $\lambda_{CS}(\sigma)$ (dashed) as functions of σ. For $\lambda > \lambda_\Pi(\sigma)$, $\Pi^*_A(\sigma) > \Pi^*_W(\sigma, \lambda)$. For $\lambda > \lambda_{CS}(\sigma)$, $CS_W(\sigma, \lambda) > CS_A(\sigma)$. For large values of σ the critical value λ_{CS} is negative.

$CS_W(\sigma, \lambda)$, strictly increases in λ because λ does not affect social surplus generated by traders who join the intermediary and strictly increases social surplus generated in the matching market. For example, for the uniform case we have

$$CS_W(1, \lambda) = \frac{3+\lambda}{16} - \frac{1-\lambda}{8} = \frac{1+3\lambda}{16} \quad \text{and} \quad CS_A(1) = \frac{1}{6} - \frac{1}{12} = \frac{1}{12}.$$

Let $\lambda_{CS}(\sigma)$ be the smallest value of $\lambda \in [0,1]$ such that $CS_W \geq CS_A$.[17] For the uniform example, we get $\lambda_{CS}(1) = 1/9$. Because $CS_W(\sigma, \lambda)$ increases in λ while $CS_A(\sigma)$ is independent of λ, it follows that consumer surplus under wholesale exceeds consumer surplus under agency if and only if $\lambda > \lambda_{CS}(\sigma)$.

Figure 3 plots $\lambda_\Pi(\sigma)$ and $\lambda_{CS}(\sigma)$ as functions of σ. Because $\lambda_\Pi(\sigma) > \lambda_{CS}(\sigma)$ for all values of σ, it follows that consumer surplus under agency is always lower than consumer surplus under wholesale whenever the intermediary's profits under agency are higher than under wholesale. Because consumer surplus is social surplus minus profits, Proposition 3 has another corollary.

Corollary 2. *Consumer surplus under the wholesale model is higher than under the agency model if the intermediary's profits under agency exceed those under wholesale.*

Put differently, Corollary 2 says that entry deterrence decreases consumer surplus.

[17]Formally, λ_{CS} is defined as $\lambda_{CS} := \inf_{\lambda \in [0,1]}\{\lambda | CS_W - CS_A \geq 0\}$. For sufficiently large values of σ there will be no $\lambda \geq 0$ such that CS_W does not exceed CS_A, which explains the somewhat complicated definition.

3.5 Price-Increasing Agency

An important issue in the debates on agency versus wholesale are their effects on prices. For example, in the Apple-e-books case, agency was associated with higher prices faced by buyers. As we show next, our model also sheds light on these questions.

Under agency one needs to distinguish between the average listing price, which we denote by $p^L(\sigma)$, and the average transaction price, which we denote $p^T(\sigma)$. The listing price of a seller with cost c who faces a fee b is $p(b, c)$. Because higher listing prices will garner a smaller market share, the average transaction price under agency will be below the average listing price. Depending on the empirical application, either $p^L(\sigma)$ or $p^T(\sigma)$ (or the distributions of transaction prices) are typically observed. Observing posted prices on online platforms corresponds to observing listing prices. In real-estate research, typically transaction prices are observed and analyzed. The average (or expected) listing price under agency $p^L(\sigma)$ given a percentage fee b is

$$p^L(\sigma) = \frac{\int_0^{1-b} p(b,c)dG(c)}{G(1-b)}|_{b=1/(1+\sigma)} = 1 - \frac{\sigma}{(1+\sigma)^2}. \tag{6}$$

where the second equality follows by plugging in the optimal fee $b = 1/(1 + \sigma)$. Of course, $p^L(\sigma)$ is independent of λ. Observe also that $p^L(\sigma)$ is monotonically increasing in σ and equal to $3/4$ for $\sigma = 1$ and converging to 1 as σ goes to infinity. While higher values of σ imply more elastic demand, and thus all else equal lower prices, this price-decreasing effect is dwarfed by the price-increasing effect that stems from the distributional shift on the supply side toward high-cost sellers as σ increases.

Under wholesale, there is no difference between the listing and the transaction price because all transactions occur at the uniform listing price. This price is

$$p_B^*(\sigma, \lambda) = 1 - (\sigma + \lambda)/(2(1 + \sigma)),$$

which is decreasing in σ and λ. Thus, $p_B^*(\sigma, \lambda)$ is maximized over λ at $\lambda = 0$, in which case it is $p_B^*(\sigma, 0) = 1 - \sigma/(2(1 + \sigma))$. Therefore, $p_B^*(1, 0) = p^L(1)$. For all other values of λ and σ with $\lambda \in [0, 1]$ and $\sigma \geq 1$, we have $p^L(\sigma) > p_B^*(\sigma, \lambda)$.

Proposition 4. *The average listing price under agency, p^L, is equal to the listing price under wholesale, $p_B(\sigma, \lambda)$, if and only if $\lambda = 0$ and $\sigma = 1$. For all other* *parameter values, the listing price under agency is larger than the listing price under wholesale, i.e.,* $p^L(\sigma) > p_B^*(\sigma, \lambda)$.

Thus, with regards to listing prices our model implies that agency leads to price increases.

As mentioned, in some industries analysts may have access to transaction prices. This is, for example, the case for the real-estate data set compiled and analyzed by Genesove and Mayer (2001). Under wholesale, the listing price $p_B^*(\sigma, \lambda)$ is also the transaction price. Under agency, the average transaction price $p^T(\sigma)$ is given as[18]

$$p^T(\sigma) = \frac{(1+\sigma)\Gamma(1/2+\sigma)}{2\Gamma(3/2+\sigma)}.$$

Defining $\lambda_P(\sigma)$ as the unique number such that $p^T(\sigma) = p_B^*(\sigma, \lambda_P(\sigma))$, one can show numerically that for all $\sigma \geq 1$

$$\lambda_P(\sigma) \leq \lambda_\Pi(\sigma),$$

with strict inequality for $\sigma > 1$, where $\lambda_\Pi(\sigma)$ is the unique number, defined above, such that $\Pi_A^*(\sigma) = \Pi_W^*(\sigma, \lambda_\Pi(\sigma))$. This is illustrated in Figure 4 and implies that agency increases average transaction prices whenever it occurs in equilibrium. Put differently, agency is profitable only if it increases average transaction prices.

Summarizing, we have the following: If agency is observed empirically if and only if it is more profitable than wholesale, then agency is associated with higher buyer prices on average, regardless of whether these are measured as listing prices or as transaction prices (weighted by the probability of sales). This is in line

[18] To see this, notice that

$$p^T(\sigma) = \frac{\int_0^{\sigma/(1+\sigma)} p(1/(1+\sigma), c)(1 - F(p(1/(1+\sigma), c)))dG(c)}{G(\sigma/(1+\sigma))(1 - F(1/(1+\sigma)))\mu^T},$$

where μ^T is the share of matches that lead to a transaction given as

$$\mu^T = \frac{\int_0^{\sigma/(1+\sigma)} (1 - F(p(1/(1+\sigma), c)))dG(c)}{G(\sigma/(1+\sigma))(1 - F(1/(1+\sigma)))}.$$

Plugging in the functions for F and G and simplifying yields the result.

With a percentage fee b, there is an alternative way to define average transaction prices, which would just be $1/b$ times Π_A because the intermediary's profit is just the percentage b of the average transaction price occurring on its intermediary. However, this is not average over the transactions actually occurring, and thus differs from what an econometrician would observe who only has data on transactions that actually occur.

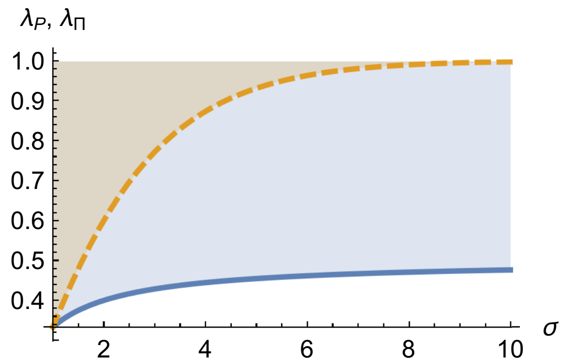


Figure 4: Agency increases average transaction prices for all $\lambda > \lambda_P(\sigma)$ (shaded area between the solid and the dashed line) and profits for all $\lambda > \lambda_\Pi(\sigma)$ (shaded area above the dashed line).

with recent empirical evidence; see, for example, De Los Santos and Wildenbeest (2016), who find that the change from agency to wholesale led to a decrease of prices for e-books of 18 percent on average.

4 Discussion

We now discuss a number of aspects of our analysis and its implications.

4.1 Price Elasticity and Equilibrium Entry Deterrence

The price set by a seller under agency is $p(1/(1+\sigma), c) = c + 1/(1+\sigma)$. Buyers with values greater than $p(1/(1+\sigma), 0) = 1/(1+\sigma)$ and sellers with costs less than $\sigma/(1+\sigma)$ trade with positive probability under agency. Because the function $p(1/(1+\sigma), c)$ decreases in σ, buyers with lower values and sellers with higher costs will be active when the elasticity increases. Nevertheless, the probability that trade actually occurs decreases in σ, the intuition being that as σ increases, there are more low valuation buyers and more high cost sellers. This more than offsets the less aggressive optimal pricing. Increases in σ decrease the profit under agency like they decrease the profit under wholesale, but eventually $\Pi^*_A(\sigma)$ decreases faster in σ than $\Pi^*_W(\sigma, \lambda)$. That is $\Pi^*_A(\sigma)/\Pi^*_W(\sigma, \lambda)$ goes to 0 as σ goes to infinity.[19]

[19]Interestingly, the ratio $\Pi^*_A(\sigma)/\Pi^*_W(\sigma, \lambda)$ is increasing in σ for σ close to 1.

For $\sigma = 15$, we have $b \approx 0.06$, which is the empirically observed fee used by real-estate brokers in the U.S. (see e.g., Hsieh and Moretti (2003)). Calibrating the supply side in this way (and assuming that buyers draw their types from a symmetric distribution), we get $\Pi_A^*(15) \approx 0.00345$ and $\Pi_W^*(15,0) \approx 0.0293$. Agency is thus more profitable than wholesale when $\lambda > \lambda_\Pi(15) = 1 - \Pi_A^*(15)/\Pi_W^*(15,0) \approx 0.882$.

4.2 Policy Implications

Policy-makers and antitrust authorities concerned with the competitiveness of a brokerage market that uses agency may quite naturally look at the level of the fees. Intuition suggests that lowering fees will enhance consumer surplus. However, in our model the anticompetitive effects of, say, percentages fees do not stem so much from their level as from their very nature: if a given percentage fee is used to deter the emergence of a competitive exchange, then so will any lower percentage fee. Therefore, if agency is used for the purpose of entry deterrence, standard regulatory approaches will not be effective. A policy implication of our results is thus that in some cases, banning the agency model can be welfare enhancing. This has indeed been the decision of the Department of Justice in the case of Apple's iTunes e-books.

For some markets, such as real-estate brokerage, the rather drastic intervention of banning the agency model and requiring the wholesale model may not be feasible for a variety of reasons. For one thing, requiring real-estate brokers to buy and sell houses, like retailers buy and sell all sorts of storable goods, is not going to be feasible, simply because of the liquidity constraints these brokers face. However, a more moderate policy intervention that might pre-empt entry-deterring behavior by brokers would be to require brokers to use flat fee mechanisms; that is, a fixed fee, which is paid when listing the property and which is paid irrespective of whether the property was sold. It can be shown that a flat fee mechanism is equivalent to posted prices (i.e., the wholesale model) in our setup. Suppose, for example, that, like in Niedermayer and Shneyerov (2014), the intermediary charges upfront *participation fees* τ_B from buyers and τ_S from buyers and then allows buyers and sellers to randomly match and bargain by a random proposer game in which the seller makes an offer with probability α and the buyer with probability $1 - \alpha$. In equilibrium, any seller that makes an offer offers p_B^0; any buyer that proposes offers p_S^0; and only buyers with $v \geq p_B^0$ and sellers with $c \leq p_S^0$ enter. The optimal fixed participation fees $\tau_S = \alpha(p_B^0 - p_S^0)$ and $\tau_B = (1-\alpha)(p_B^0 - p_S^0)$ extract the same rents as wholesale and provide the same incentives, so that the outcome is equivalent to wholesale.

The agreement of real-estate brokerage associations with the Department of Justice to stop practices which have been seen as discriminatory toward flat-fee brokers can be seen as a step in this direction (see DOJ, 2007).[20]

4.3 Alternative Bargaining Protocols

In this subsection, we show that entry deterrence is an equilibrium outcome for the alternative widely used bargaining protocols of take-it-or-leave-it offers, Nash bargaining, and double-auctions. As is well known, deriving equilibrium predictions is notoriously difficult for bargaining under incomplete information with general distributions. Therefore, we now assume uniform distributions; that is, we set $\sigma = 1$.

We briefly describe the standard bargaining protocols we now consider. With *random proposer take-it-or-leave-it-offers*, the buyer makes an offer p_b with probability α, which the seller can either accept or reject. With probability $1-\alpha$, the seller makes the take-it-or-leave-it-offer p_s. Under *generalized Nash bargaining*, the buyer with value v gets α of the joint surplus $\max\{v-c, 0\}$ when matched to a seller with cost c while the seller gets $1-\alpha$ times this surplus. In the $k = 1/2$ *double-auction*, a buyer and a seller submit bids p_b and p_s simultaneously and trade at the price $p = (p_b + p_s)/2$ if and only if $p_b \geq p_s$; for examples, see Chatterjee and Samuelson (1983); Rustichini, Satterthwaite, and Williams (1994), and Satterthwaite and Williams (2002). With buyers' values and sellers' costs uniformly distributed on $[\underline{v}, \overline{v}]$ and $[\underline{c}, \overline{c}]$, respectively, the $k = 1/2$ double-auction has a Bayes Nash equilibrium in which the buyer with value $v \in [\underline{v}, \overline{v}]$ bids

$$p_b(v) = \frac{1}{4}\underline{c} + \frac{1}{12}\overline{v} + \frac{2}{3}v$$

and the seller with cost $c \in [\underline{c}, \overline{c}]$ bids

$$p_s(c) = \frac{1}{4}\overline{v} + \frac{1}{12}\underline{c} + \frac{2}{3}c,$$

[20] The allegations were that flat-fee brokers were put at a disadvantage in the Multiple Listing Services, e.g., in cases where a buyer's broker and a seller's broker would usually cooperate, cooperation was refused for flat-fee brokers.

inducing trade whenever $v \geq c + (\overline{v} - \underline{c})/4$.[21]

We defer the technical derivation of the following proposition to the appendix.

Proposition 5. *Under each of the three standard bargaining protocols, (i) entry deterrence occurs in equilibrium for λ sufficiently large, and (ii) entry deterrence, when it occurs, is harmful for social welfare and for consumer surplus.*

4.4 Self-Enforcing and Optimal Mechanisms

It is remarkable that a market structure that at face value looks very competitive consisting of a large number of active brokers each of whom has a small market share may have predatory effects. This is possibly even more remarkable because it is an optimal mechanism given that with the matching technology agency is *self-enforcing* in the following sense. Keeping fixed the random matching technology across brokers, the fee $b = 1/(1+\sigma)$ is optimal given the distributions $F(v)$ and $G(c)$ on $[0, 1]$ and optimal when these distributions are truncated to the sets of buyer- and seller-types that can trade with positive probability under agency.[22]

Wholesale is an optimal mechanism for a large intermediary facing a deep market with a continuum of buyers and sellers when there is no random matching market in the sense that there is no other incentive compatible and individually rational mechanism that generates higher profits. We show this in Appendix B. In the same sense, agency is an optimal mechanism for a broker in a thin market where the broker is randomly matched to a buyer-seller pair. This is an implication of our results in Loertscher and Niedermayer (2017a). An open question is to derive the optimal mechanism for an intermediary who faces an active competing exchange. Theoretically, this is challenging because it corresponds to a mechanism design problem with endogenous and type dependent participation constraints. Even *exogenous* type dependent participation constraints are known to be difficult (see e.g., Jullien, 2000). From a practical perspective, it requires being specific about how payoffs in the random matching market are determined. What this paper has shown is that the wholesale model, which is optimal in the absence of random matching markets, will no longer be optimal when there is the

[21] See, for example, Chatterjee and Samuelson (1983) and Myerson and Satterthwaite (1983), who also observe that this is a second-best mechanism.

[22] See also Tirole (2016) for the importance of time consistency and selection effects for optimal mechanisms.

threat of an active random matching market.

4.5 Equilibrium Selection

Multiplicity of equilibria is inevitable in almost any model of two-sided markets with endogenous participation simply because not going to an exchange if no one else does is typically a best response. Not surprisingly, therefore, our model also exhibits multiple equilibria. For example, if one assumes that under wholesale the intermediary is not committed to buying from sellers willing to sell regardless of demand, it is always part of an equilibrium strategy profile that no buyer and no seller joins the intermediary under wholesale. This equilibrium is arguably not very compelling as small perturbations to the mechanism (such as a commitment to buy) would eliminate it.[23] Similarly, there is always an equilibrium in which no agent goes to the random matching market. This is so even when there is no intermediary, but again does not strike us as particularly plausible.

Our theory of entry deterring agency, then, rests on the following assumptions. Given posted prices that leave a positive spread, the buyers and sellers with values and costs inside the price gap $[p_S, p_B]$ would realize that there are gains from joining the random matching market. Consequently, assuming a positive bid-ask spread $p_B - p_S$, our assumption is that the random matching market is active whenever $\lambda > 0$. For any given posted prices, however, any buyers and sellers who derive greater utility from trading with the intermediary at these prices than they expect from participating in the random matching market trade with the intermediary, where expectations are taken under the assumptions that all traders behave in the same way. Given agency, we maintain the same assumption: For the announced fee b all those buyers and sellers join the intermediated market who expect greater gains from doing so than when joining the random matching market, again taking expectations under the assumption that all other traders behave in the same way. (Of course, the twist now is that the expected payoff from random matching market participation is 0 for all agents if all agents behave and form expectations in this way.)

[23]With fixed participation fees, the commitment to buy would not be available. However, following Jullien (2006), one can alternatively focus on stable equilibria with stability referring to a dynamic adjustment process where the two sides alternate in their registration choice and respond myopically to the other side's market share.

4.6 Dynamic Random Matching

In Appendix C we derive results for a variant of our model with a binary type space, in which buyers' valuations are drawn from $\{v_1, v_2\}$ and sellers' costs are drawn from $\{c_1, c_2\}$. The main results on entry deterrence go through in this alternative setting.

The binary type space has the advantage that some of the calculations become more tractable. In particular, this tractability allows us to address dynamics: consider a model in which a trader that does not trade today has the chance to trade in the future. Such traders wait in the market to try to get a future trade opportunity. In such a dynamic random matching setup, the relevant distributions of valuations and costs in the market are different from the primitive distributions for two reasons. First, traders have the option value of future trade which should be added to the seller's cost and subtracted from the buyers valuation. Second, low valuation buyers and high cost seller trade with a lower probability, wait longer until they get to trade, and hence are overrepresented in the pool of traders in the market compared to the distribution of entrants in every period.

Dealing with these two changes of the type distributions in a dynamic model is known to be a very hard problem, see Satterthwaite and Shneyerov (2007, 2008).[24] For binary type spaces, a dynamic model is more tractable, see Duffie, Garlenau, and Pedersen (2005). We analyze entry-deterring agency in a dynamic random matching with a binary type space in Appendix C and find that our results go through qualitatively. In a dynamic model, traders' concern about a lower probability of trade becomes less important, but in exchange, traders worry about a longer time on market, which has a similar effect.

4.7 Incomplete Foreclosure

We have stated our results on entry-deterring agency in the clearest possible way by assuming that entry deterrence leads to bilateral exchange shutting down completely. However, our results also hold for a less extreme version of our model: if entry deterrence does not lead to the bilateral exchange shutting down completely, but only to operating at an inefficiently low scale.

[24] It can be seen from these articles how complex the analysis of dynamic random matching with continuous type spaces is, even if one considers the limit as search frictions vanish. Away from the limit the analysis typically becomes even harder.

Consider a bilateral market that operates at an inefficiently low scale if only a small mass of buyers and sellers enter (e.g. if $\min\{1 - F(\underline{v}), G(\overline{c})\} < T$ for some threshold T), so that the probability of meeting a trading partner is λ_0. Above the threshold T, the probability of meeting a trading partner is $\lambda > \lambda_0$. Further assume that there is a small mass ϵ of both buyers and sellers who always go to the bilateral market, possibly because they are sophisticated at searching and do not need an intermediary or because they cannot afford the intermediary's fees. As long $\epsilon < T$, entry deterrence leads to the bilateral exchange operating at an inefficiently low scale, which reduces (but does not completely eliminate) the competitive threat the bilateral exchange poses to the intermediary's rent extraction.

4.8 Empirics

An empirical analysis is outside of the scope of this article. However, there are a few issues worth discussing.

One is the extent of price dispersion in search markets. This is difficult to measure empirically because differences in sales prices can be attributed to either observed heterogeneity (for real-estate, the number of bedrooms of a property, whether it has air conditioning, etc.), to unobserved heterogeneity (characteristics of a property that are observed by market participants, but not by the econometrician, e.g., whether the view from the apartment is beautiful), or to true price dispersion.

In Loertscher and Niedermayer (2017b), we structurally estimate an agency model, taking payoffs from a bilateral exchange as exogenously given. The structural estimates allow us to disentangle the three effects leading to differences in sales prices. We show how the three effects can be disentangled in Figure 5. It should be noted that not only the analysis of observed and unobserved heterogeneity poses challenges, but also its graphical representation: while sales prices are measured in hundreds of thousands of dollars, the quality-adjusted prices one obtains after correcting for observed heterogeneity and the counterfactual simulations of how much price variation one would see if there were no unobserved heterogeneity are quality-adjusted prices normalized to a mean of 1. We deal with this challenge by computing "denormalized quality-adjusted prices": we multiply the quality-adjusted prices with \$230,000, the average transaction price in the data set in Loertscher and Niedermayer (2017b).

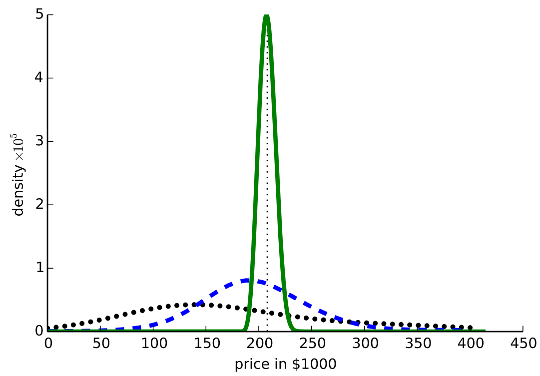


Figure 5: Empirically estimated density of prices without correcting for heterogeneity (dotted, black), correcting for observed heterogeneity (dashed, blue) and correcting for both observed and unobserved heterogeneity (solid, green).

Based on the structural estimate of true price dispersion (solid line in Figure 5), we can look at how the 6% fees charged by real-estate brokers compare to price dispersion. Figure 6 shows the distribution of "denormalized quality-adjusted" gross prices (solid) and "denormalized quality-adjusted" prices net of the 6% fee charged by real-estate brokers. It shows that the two distributions have considerable overlap.

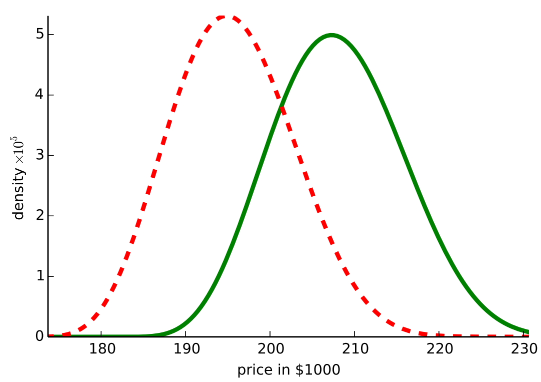


Figure 6: Empirical density of gross (solid, green) and net (dashed, red) transaction prices adjusted by both observed and unobserved heterogeneity. Based on Loertscher and Niedermayer (2017b).

An issue Loertscher and Niedermayer (2017b) do not deal with is the endogeneity of the bilateral exchange market that is the focus of this article. This is the right approach for that paper, since they analyze data from Boston in the early 1990s when the fraction of bilateral transactions was quite small.

An interesting question for future research would be to analyze data from regions and periods of time in which a bilateral exchange came into existence and started growing. One such data set is that analyzed by Hendel, Nevo, and Ortalo-Magné (2009), who consider both real-estate broker sales and For-Sale-By-Owner intermediary sales in Madison, Wisconsin, from 1998 to 2004. In this period of time, the For-Sale-By-Owner intermediary grew considerably. They find that properties sold through the For-Sale-By-Owner intermediary took considerably longer to sell than properties sold through real estate brokers, which is consistent with the assumption $\lambda < 1$; that is, that matchings through brokers are more efficient than bilateral matchings.[25] In a dynamic setting, λ determines that probability of sale in a single period and hence the time on market.

From the perspective of our model, the emergence of the For-Sale-By-Owner intermediary in Madison could be interpreted as either entry deterrence by brokers still being successful and the bilateral market still operating at an inefficiently small scale, but the efficiency at this small scale improving due to technological progress such as the Internet (i.e. λ_0 in Section 4.7 increasing). Alternatively, it could be seen as one of the few occasions in which a bilateral market managed to break free from entry deterrence.[26]

Another interesting question is the comparison to the used car market. Used car dealers typically choose the wholesale model and the fraction of bilateral transactions is known to be considerably higher for used cars than for real-estate transactions. An empirical study by Gavazza, Lizzeri, and Roketskiy (2014) that analyzes the (bilateral) used car market from a search theoretic perspective finds that there are considerable search frictions in this market, which from the perspective of our model means a low λ.[27] In a market with a low λ, intermediaries

[25] It should be noted that Hendel, Nevo, and Ortalo-Magné (2009) also find that the probability of sale is slightly higher for the For-Sale-By-Owner intermediary than for real-estate brokers; however, the difference is not statistically significant.

[26] We do not have an answer to why a bilateral exchange can break free. Possibly, enough sophisticated buyers and sellers are present in a region, who are efficient at searching for a trading partner even without an intermediary.

[27] There are different ways to model frictions in the used car market. In both Gavazza, Lizzeri, and Roketskiy's (2014) and our article these frictions are modeled as search frictions; that is, a lower probability of meeting a trading partner in a given period. An alternative is to model frictions as adverse selection: a buyer may find a seller, but there is a high probability that the

do not have an incentive to collude to foreclose the bilateral market, since it is a lesser competitive threat than in markets with a larger λ. Hence, intermediaries are more likely to choose wholesale, which is indeed what we observe for used car dealers.

5 Conclusions

In this paper, we present a model in which the operator of one market can successfully deter the emergence of a competing exchange. Entry deterrence is the more profitable the more efficient is the matching technology in the competing exchange whose emergence is deterred. The paper thus brings to light a relevant and novel possibility of entry deterrence. For the specifications we studied, entry deterrence is always harmful to social welfare, and when it occurs in equilibrium, to consumer surplus defined as social welfare less the intermediary's profit.

Policy-makers and antitrust authorities concerned with the competitiveness of a brokerage market that uses agency quite naturally look at the level of the fees charged. Intuition suggests that lowering fees will enhance consumer surplus. However, in our model the anticompetitive effects of, say, percentages fees do not stem so much from their level as from their very nature. If agency is used for the purpose of entry deterrence, standard regulatory approaches will not be effective as any lower fee will also deter entry. The first-order welfare gains in our model are achieved by inducing intermediaries to use the wholesale model or equivalent mechanisms.

References

Anderson, S., and S. Coate (2005): "Market Provision of Broadcasting: A Welfare Analysis," *Review of Economic Studies*, 72, 947–72.

Armstrong, M. (2006): "Competition in Two-Sided Markets," *RAND Journal of Economics*, 37(3), 668–691.

seller has a used car of inferior quality. While we have no formal model, we conjecture that our results should qualitatively go through also in the latter specification, since the basic driving force is that the intermediary has an incentive to predate if the bilateral market is too efficient.

Ashenfelter, O., and K. Graddy (2005): "Anatomy of the Rise and Fall of a Pricefixing Conspiracy: Auctions at Soheby's and Christie's," *Journal of Competition Law and Economics 1(1), 320*, 1(1), 3–20.

Belleflamme, P., and M. Peitz (2015): *Industrial organization: markets and strategies*. Cambridge University Press.

Besanko, D., U. Doraszelski, and Y. Kryukov (2014): "The economics of predation: What drives pricing when there is learning-by-doing?," *American Economic Review*, 3(3), 868 – 897.

Besanko, D., U. Doraszelski, Y. Kryukov, and M. Satterthwaite (2010): "Learning-by-Doing, Organizational Forgetting, and Industry Dynamics.," *Econometrica*, 78(2), 453 – 508.

Biglaiser, G., and F. Li (2016): "Middlemen: the good, the bad, and the ugly," Mimeo, UNC.

Bloch, F., and H. Ryder (2000): "Two-Sided Search, Marriages, and Matchmakers.," *International Economic Review*, 41(1), 93 – 115.

Bourguignon, H., R. Gomes, and J. Tirole (2014): "Shrouded Transaction Costs," Mimeo.

Cabral, L. M. B., and M. H. Riordan (1994): "The Learning Curve, Market Dominance, and Predatory Pricing.," *Econometrica*, 62(5), 1115 – 1140.

——— (1997): "The Learning Curve, Predation, Antitrust, and Welfare.," *Journal of Industrial Economics*, 45(2), 155 – 169.

Caillaud, B., and B. Jullien (2001): "Competing cybermediaries," *European Economic Review*, 45, 797–808.

——— (2003): "Chicken and Egg: Competition among Intermediation Service Providers," *RAND Journal of Economics*, 34(2), 309–328.

Chatterjee, K., and W. Samuelson (1983): "Bargaining under Asymmetric Information," *Operations Research*, 31, 835–851.

De Los Santos, B., and M. Wildenbeest (2016): "E-book pricing and vertical restraints," Mimeo, Indiana University.

DOJ (2007): "Competition in the Real Estate Brokerage Industry," *Report by the Federal Trade Commission and the U.S. Department of Justice*.

Duffie, D., N. Garlenau, and L. H. Pedersen (2005): "Over-the-Counter Markets," *Econometrica*, 73(6), 1815–1847.

Edelman, B., and J. Wright (2015): "Price Coherence and Excessive Intermediation," *Quarterly Journal of Economics*, 130, 1283 – 1328.

Foner, P. S. (1965): *History of the Labor Movement in the United States: The Industrial Workers of the World, 1905-1917*, vol. 4. New York: International Publishers.

Fudenberg, D., and J. Tirole (2000): "Pricing a Network Good to Deter Entry," *The Journal of Industrial Economics*, 48(4), 373–390.

Gavazza, A., A. Lizzeri, and N. Roketskiy (2014): "A Quantitative Analysis of the Used-Car Market," *American Economic Review*, 104(11), 3668–3700.

Gehrig, T. (1993): "Intermediation in Search Markets," *Journal of Economics & Management Strategy*, 2, 97–120.

Genesove, D., and C. J. Mayer (2001): "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116(4), 1233–1260.

Gomes, R. (2014): "Optimal Auction Design in Two-Sided Markets," *Rand Journal of Economics*, 45, 248–272.

Hendel, I., A. Nevo, and F. Ortalo-Magné (2009): "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com," *American Economic Review*, 99(5), 1878 – 98.

Hsieh, C.-T., and E. Moretti (2003): "Can Free Entry Be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry," *Journal of Political Economy*, 111(5), 1076–1122.

Johnson, J. (2013): "The Agency and Wholesale Model in Electronic Content Marks," Mimeo, Cornell University.

Johnson, J. P. (2017): "The agency model and MFN clauses," *The Review of Economic Studies*, p. rdx007.

Jullien, B. (2000): "Participation Constraints in Adverse Selection Models.," *Journal of Economic Theory*, 93(1), 1 – 47.

——— (2006): "Two-Sided Markets and Electronic Intermediaries," in *Industrial Organization and the Digital Economy*, ed. by G. Illing, and M. Peitz, chap. 10, pp. 273–302. MIT Press, Cambridge, MA.

Krishna, V. (2002): *Auction Theory*. Elsevier Science, Academic Press.

Levitt, S. D., and C. Syverson (2008): "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions," *Review of Economics and Statistics*, 90(4), 599–611.

Loertscher, S. (2007): "Horizontally Differentiated Market Makers," *Journal of Economics & Management Strategy*, 16(4), 793–825.

Loertscher, S., and A. Niedermayer (2016): "Percentage Fees in Thin Markets: An Optimal Pricing Perspective," *Working Paper*.

Loertscher, S., and A. Niedermayer (2017a): "An Optimal Pricing Theory of Transaction Fees in Thin Markets," Unsubmitted mimeo.

——— (2017b): "Percentage Fees in Thin Markets: An Empirical Perspective," Mimeo.

Myerson, R. (1981): "Optimal auction design," *Mathematics of Operations Research*, 6(1), 58–73.

Myerson, R., and M. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29(2), 265–281.

Neeman, Z., and N. Vulkan (2010): "Markets Versus Negotiations: The Predominance of Centralized Markets," *The B.E. Journal of Theoretical Economics: Advances*, 1(6).

NIEDERMAYER, A., AND A. SHNEYEROV (2014): "For-Profit Search Platforms," *International Economic Review*, 55(3), 765 – 789.

RILEY, J. G., AND W. F. SAMUELSON (1981): "Optimal Auctions.," *American Economic Review*, 71(3), 381 – 392.

ROCHET, J.-C., AND J. TIROLE (2003): "Platform competition in Two-Sided markets," *Journal of the European Economic Association*, 1, 990–1029.

——— (2006): "Two-Sided Markets: A Progress Report," *RAND Journal of Economics*, 37(3), 645–667.

RUBINSTEIN, A., AND A. WOLINSKY (1987): "Middlemen," *Quarterly Journal of Economics*, 102(3), 581–593.

RUST, J., AND G. HALL (2003): "Middlemen versus Market Makers: A Theory of Competitive Exchange," *Journal of Political Economy*, 111(2), 353–403.

RUSTICHINI, A., M. A. SATTERTHWAITE, AND S. R. WILLIAMS (1994): "Convergence to Efficiency in a Simple Market with Incomplete Information," *Econometrica*, 62(5), pp. 1041–1063.

RUTHERFORD, R., T. SPRINGER, AND A. YAVAS (2005): "Conflicts between principals and agents: evidence from residential brokerage," *Journal of Financial Economics*, 76, 627–65.

SATTERTHWAITE, M., AND A. SHNEYEROV (2007): "Dynamic Matching, Two-Sided Information, and Participation Costs: Existence and Convergence to Perfect Competition," *Econometrica*, 75(1), 155–200.

SATTERTHWAITE, M., AND A. SHNEYEROV (2008): "Convergence to Perfect Competition of a Dynamic Matching and Bargaining Market with Two-sided Incomplete Information and Exogenous Exit Rate.," *Games and Economic Behavior*, 63(2), 435–467.

SATTERTHWAITE, M. A., AND S. R. WILLIAMS (2002): "The Optimality of a Simple Market Mechanism.," *Econometrica*, 70(5), 1841 – 1863.

SHNEYEROV, A., AND A. WONG (2010): "Bilateral Matching and Bargaining with Private Information.," *Games and Economic Behavior*, 62(2), 748–762.

Spulber, D. F. (1996): "Market Making by Price-Setting Firms," *Review of Economic Studies*, 63(4), 559–580.

Stahl, D. O. (1988): "Bertrand Competition for Inputs and Walrasian Outcomes," *American Economic Review*, pp. 189–201.

Tirole, J. (2016): "From Bottom of the Barrel to Cream of the Crop: Sequential Screening with Positive Selection," *Econometrica*, 84(4), 12911343.

Yavas, A. (1996): "Matching of Buyers and Sellers by Brokers: A Comparison of Alternative Commission Structures," *Real Estate Economics*, 24, 97–112.

Online Appendix

For Online Publication

A Proofs

Proof of Proposition 1. We need to show that $\Pi_A(b)$ is maximized at $b = 1/(1+\sigma)$. The expression in (4) then follows from straight forward operations, and the second part of the proposition has already been shown in the text. That $\Pi_A(b)$ is maximized at $b = 1/(1+\sigma)$ follows for the following reasons. First, the expression $b\int_0^{1-b}[p(b,c)(1-F(p(b,c)))]dG(c)$ is maximized at $b = 1/(1+\sigma)$. Consequently, the only way a different fee could do better is that it affects the quantity traded in a profitable way. However, this is not the case. For $b > 1/(1+\sigma)$, we have $G(1-b) < 1 - F(\underline{p})$, so that $\Pi_A(b)$ becomes

$$\Pi_A(b) = \frac{b}{1-F(\underline{p})}\int_0^{1-b}(p(b,c)(1-F(p(b,c))))dG(c).$$

Because, as noted, $\underline{p}$ is independent of b, this deviation cannot be profitable. For $b < 1/(1+\sigma)$, we have

$$\Pi_A(b) = \frac{b}{G(1-b)}\int_0^{1-b}(p(b,c)(1-F(p(b,c))))dG(c).$$

But now both terms, $1/G(1-b)$ and $b\int_0^{1-b}[p(b,c)(1-F(p(b,c)))]dG(c)$ are increasing b, contradicting that the deviation is profitable. $\square$

Proof of Proposition 3. The duplication formula for the gamma-function implies that

$$\frac{\Gamma(z)}{\Gamma(2z)} = \frac{2^{1-2z}\sqrt{\pi}}{\Gamma(z+1/2)}.$$

Setting $z = 1+\sigma$ and using this fact, (5) is equivalent to

$$\frac{S_A(\sigma)}{S_W(\sigma,0)} = \frac{(1+\sigma)(1+3\sigma)}{2^\sigma(1+2\sigma)}\frac{\sqrt{\pi}}{2}\frac{\Gamma(1+\sigma)}{\Gamma(3/2+\sigma)}.$$

Because $\frac{(1+\sigma)(1+3\sigma)}{2^\sigma(1+2\sigma)}$ and $\frac{\Gamma(1+\sigma)}{\Gamma(3/2+\sigma)}$ are both decreasing in σ (with the sign of the derivative of $\Gamma(1+\sigma)/\Gamma(3/2+\sigma)$ equal to the sign of the difference between the harmonic number at σ and at $\sigma + 1/2$, which is negative), $S_A(\sigma)/S_W(\sigma, 0) \leq S_A(1)/S_W(1, 0) = 8/9$.

$\square$

Proof of Proposition 5. **Random proposer take-it-or-leave-it-offers** The marginal buyer who is indifferent to trading with the intermediary has valuation $\overline{v} = 1 - q$, the marginal seller cost $\underline{c} = q$. Buyers with a valuation below $\underline{c}$ and sellers with costs above $\overline{v}$ do not bother to enter the bilateral market. Hence, in the bilateral market, buyer valuations and seller costs are uniformly distributed on $[\underline{c}, \overline{v}]$. Denote this distribution as $\tilde{G}(x) = \tilde{F}(x) = (x - \underline{c})/(\overline{v} - \underline{c})$.

The optimal price offer $p_b(v)$ of a buyer with value $v \in [\underline{c}, \overline{v}]$ maximizes $(v - p_b)\tilde{G}(p_b)$, yielding $p_b(v) = (v + \underline{c})/2$. Likewise, the optimal offer $p_s(c)$ of a seller with cost $c \in [\underline{c}, \overline{v}]$ is the p_s that maximizes $(p_s - c)(1 - \tilde{F}(p_s))$, yielding $p_s(c) = (\overline{v} + c)/2$. Observe that the offers of the marginal traders – of the buyer with value $\overline{v}$ and the seller with cost $\underline{c}$ – are accepted with probability 1/2. Setting $\underline{c} = q$ and $\overline{v} = 1 - q$, the expected payoff for the marginal seller who is matched to a buyer therefore is

$$V_S(q) = (1-\alpha)(p_s(\underline{c})-\underline{c})(1-\tilde{F}(p_s(\underline{c})))+\alpha \int_{\underline{c}}^{\overline{v}} (p_b(v)-\underline{c})d\tilde{F}(v) = \frac{\overline{v} - \underline{c}}{2} = \frac{1 - 2q}{4}.$$

Because this is independent of α, it follows by symmetry that $V_B(q) = V_S(q)$, yielding $V(q) = 2V_S(q) = (1-2q)/2$. It is easy to check that the No Overshooting Condition and hence also the Single Crossing Condition hold.[28]

Plugging this $V(q)$ into the wholesale intermediary's profit maximization problem $(F^{-1}(1 - q) + G^{-1}(q) - \lambda V(q))q$, we get the maximizer $q^* = 1/4$ and the maximum

$$\Pi_W = \frac{2 - \lambda}{16}.$$

As derived before, under agency, the intermediary's profit is $\Pi_f = 1/12$. Hence, the intermediary prefers agency to wholesale ($\Pi_A > \Pi_W$) if $\lambda > 2/3$.

[28]The single-crossing condition is satisfied because for all $v \in [\underline{v}, \overline{v}]$, $V_B(v, 1-q, q)$ is increasing in v with a slope smaller than 1 because even conditional on being matched such a buyer trades with probability less than 1. For $v > \overline{v}$, the slope of $V_B(v, 1 - q, q)$ is also positive but depends on whether such a buyer optimally makes offers that are always accepted, which in turn depends on v and q. If the optimal offer is always accepted, the slope is 1, and less otherwise. Because the payoff $V_B(v, 1 - q, q)$ is multiplied by the matching probability λ, it follows that the single-crossing condition is always satisfied for buyers. Symmetric arguments apply to sellers.

Welfare under wholesale generated by trade via the intermediary is 3/16.[29] The surplus generated in the matching market is

$$(1-2q)\lambda\left[\alpha\frac{\int_{1/2}^{1-q}\int_{q}^{2v-(1-q)}(v-c)dcdv}{(1-2q)^2}+(1-\alpha)\frac{\int_{q}^{1/2}\int_{2c-q}^{1-q}(v-c)dvdc}{(1-2q)^2}\right]=\lambda\frac{1-2q}{8}|_{q=1/4}=\frac{\lambda}{32},$$

where the term $1-2q$ is the mass of traders in the matching market. Therefore, social welfare under wholesale is

$$S_W=\frac{3}{16}+\frac{\lambda}{32}.$$

Because welfare under agency is only 1/6, it follows that entry deterrence always decreases social welfare.

Under wholesale, consumer surplus, defined as the surplus of buyers and sellers and denoted CS_p,

$$CS_W=S_W-\Pi_W=\frac{2+3\lambda}{32}$$

while consumer surplus under agency is $CS_f=1/12$. Thus, $CS_p>CS_f$ is equivalent to $\lambda>2/9$, implying that agency is always detrimental to consumer surplus when it is preferred by the intermediary. Consumer surplus under wholesale is more responsive to increases in λ than social welfare because increases in λ not only improve the lot of the agents who are active in the matching market but also increase the consumer surplus of agents who trade with the intermediary under wholesale. In contrast, social welfare generated by intermediated trade under wholesale does not vary with λ because λ only shifts rents.

Nash bargaining Under generalized Nash bargaining, where the buyer with value v obtains α of the joint surplus $\max\{v-c,0\}$ when matched to a seller with cost c, the expected payoff of a buyer of type v conditional on being matched in the random matching market is $\alpha\int_q^v(v-c)dc/(1-2q)=\alpha(v-q)^2/(2(1-2q))$. This implies that the No Overshooting Condition holds, since an increase of v by Δv is partially bargained away and leads to an increase of the buyers utility by at most $\alpha\Delta v$ (it may be less if the probability of being matched is less than 1 and the probability that $v\geq c$ is less than 1).

The marginal buyer with value $v=1-q$ has an expected payoff from participating in the matching market, conditional on being matched, of $V_B(q)=$

[29] A simple geometric argument can be used. Social welfare under efficiency is 1/4 while Harberger's triangle due to the intermediary's monopoly power is 1/16.

$\alpha(1-2q)/2$. Analogously, $V_S(q) = (1-\alpha)(1-2q)/2$ can be established under generalized Nash bargaining where the seller gets the share $(1-\alpha)$ of the surplus. Just like with take-it-or-leave-it offers, with generalized Nash bargaining, α does not affect $V(q) = V_B(q) + V_S(q) = (1-2q)/2$, but in contrast to random take-it-or-leave-it offers, α affects the division of the sum of the marginal utilities $V(q)$ into $V_B(q)$ and $V_S(q)$. Consequently, α will affect the equilibrium prices under generalized Nash bargaining but neither the bid-ask spread nor the quantity traded. Moreover, because $V(q)$ is the same under generalized Nash bargaining and under take-it-or-leave-it offers, the intermediary's equilibrium profits will be the same and so will be the conditions under which entry deterrence occurs; that is, for $\lambda > 2/3$.

The only payoff-relevant differences regard the surplus that is generated in the random matching market.[30] With generalized Nash bargaining, this surplus is

$$(1-2q)\lambda \frac{\int_q^{1-q} \int_q^v (v-c)dcdv}{(1-2q)^2}|_{q=1/4} = \frac{\lambda}{24},$$

implying that social welfare under wholesale and generalized Nash bargaining is

$$S_W = \frac{3}{16} + \frac{\lambda}{24}.$$

A fortiori, entry-deterring agency will reduce social welfare. Consumer surplus with wholesale and Nash bargaining is

$$CS_W = S_W - \Pi_W = \frac{3+5\lambda}{48},$$

which for any $\lambda > 1/5$ is larger than the consumer surplus under agency of $1/12$.

Double Auction We focus on the linear equilibrium described in the text. The marginal traders who will be indifferent between participating and being inactive are the seller with cost $\overline{c}$ equal to $p_b(\overline{v})$; that is, $\overline{c} = \overline{v} - (\overline{v} - \underline{c})/4$, and the buyer with value $\underline{v}$ equal to $p_s(\underline{c})$; that is, $\underline{v} = \underline{c} + (\overline{v} - \underline{c})/4$. Consequently, the masses of active buyers and sellers will be

$$\overline{v} - \underline{v} = \frac{3}{4}(\overline{v} - \underline{c}) = \overline{c} - \underline{c}.$$

[30] There is another difference regarding the single-crossing condition. Because values and costs are common knowledge in a match under Nash bargaining, there is "hold-up" in the sense that an agent is never the residual claimant to the additional surplus he generates.

The expected payoff of a buyer of type $v \in [\underline{v}, \overline{v}]$ upon being matched is therefore

$$V_B(v, \overline{v}, \underline{c}) = \frac{1}{\overline{c} - \underline{c}} \int_{\underline{c}}^{p_s^{-1}(p_b(v))} [v - (p_b(v) + p_s(c))/2] dc.$$

Substituting $\underline{c} = q$ and $\overline{v} = 1 - q$, we obtain

$$V_B(q) \equiv V_B(1-q, 1-q, q) = \frac{3}{8}(1 - 2q).$$

Again, the no overshooting condition, and hence the single-crossing property, can be shown to hold for this bargaining protocol.[31]

By symmetric arguments, $V_S(q) = V_B(q)$, so that for the $k = 1/2$ double-auction,

$$V(q) = \frac{3}{4}(1 - 2q).$$

The profit maximizing quantity under wholesale is still $q^* = 1/4$. The equilibrium prices are

$$p_B^* = \frac{3}{4} - \frac{\lambda}{16} \quad \text{and} \quad p_S^* = \frac{1}{4} + \frac{3\lambda}{16}, \tag{7}$$

yielding and equilibrium spread of $(4 - 3\lambda)/8$ and thus an equilibrium profit of

$$\Pi_W = \frac{4 - 3\lambda}{32}.$$

Therefore, entry deterrence occurs in equilibrium if and only if $\lambda > 4/9$.

Equilibrium welfare under wholesale is

$$S_W = \frac{3}{16} + \lambda \frac{3}{64},$$

where

$$\frac{3\lambda}{64} = \frac{3(1-2q)}{4} \frac{\int_{q+(1-2q)/4}^{1-q} \int_q^{v-(1-2q)/4} (v - c) dc dv}{(3(1-2q)/4)^2} \Big|_{q=1/4}$$

[31]The single-crossing condition can be shown to be satisfied using analogous arguments to the case with take-it-or-leave-it offers. The payoff of a buyer with value $v \in [\underline{v}, \overline{v})$, who in equilibrium participates in the matching market, increases by less than 1 in v because such a buyer trades with probability less than 1 conditional on being matched. For $v \geq \overline{v}$, the buyer is essentially the residual claimant on the additional surplus he creates because he optimally submit the bid $p_b(\overline{v})$ and induces trade with probability 1 conditional on being matched. Therefore, the derivative of $V_B(v, 1-q, q)$ with respect v is 1 for $v \geq \overline{v}$. But, again, because the payoff $V_B(v, 1-q, q)$ is multiplied by the matching probability λ, it follows that the single-crossing condition is always satisfied for buyers, and analogously for sellers.

is the expected social surplus from the matching market, which is larger than W_A.[32]

Consumer surplus under wholesale is then easily seen to be

$$CS_W = \frac{4 + 7\lambda}{64},$$

which is larger than consumer surplus under agency for any $\lambda > 4/21$.

In passing, we note that under wholesale the double auction in the matching market generates higher welfare and higher consumer surplus than take-it-or-leave-it offers and, perhaps more surprisingly, than Nash bargaining. As mentioned, desirable properties of the double auction given the uniform distribution have been noted before (see e.g. Myerson and Satterthwaite, 1983). The present finding adds a new element to this list. As it accounts for private information held by the two parties, it cannot be as efficient as Nash bargaining because of the impossibility theorem of Myerson and Satterthwaite. But exactly because it induces bid shading by the buyer and the seller, it prevents the most inefficient buyers and sellers – the buyers with values below 3/8 and the sellers with costs above 5/8 – from entering the matching market, thereby improving sorting and surplus in the matching market relative to Nash bargaining and take-it-or-leave-it offers. □

B Optimality of Wholesale

Consider a setup with N buyers and N sellers who draw their values and costs independently from distributions F and G with support $[0, 1]$. Assume that F and G are common knowledge but that each agent's realized type is his private information. Assume that agents can only trade via the mechanism designer's intermediary and that each agent's utility of the outside option of not trading is 0. A direct mechanism is a collection of functions $\langle Q, M \rangle$ with $Q : [0, 1]^{2N} \to [0, 1]^{2N}$ specifying, for all agents i, the probability Q_i that the agent trades and $M : [0, 1]^{2N} \to \mathbb{R}^{2N}$ specifying, for all agents i, the payment M_i makes to the

[32] To see that this is true, notice that the density of buyers (and of sellers) in the matching market is 1 divided by the mass of active buyers; that is $1/(3(1 - 2q)/4)$, which is explains the term $1/(3(1 - 2q)/4)^2$. The surplus of a trade being $v - c$ and trade occurring, as noted, whenever $v \geq c + (1 - 2q)/4$ explains the upper bounds of the inner integral while the lower bound in the outer integral reflects the fact that $\underline{v} = q + (1 - 2q)/4$. Multiplying by the mass of active traders $3(1 - 2q)/4$ gives the result.

mechanism. A mechanism is feasible if $\sum_{i\in\mathcal{B}} Q_i \leq \sum_{j\in\mathcal{S}} Q_j$, where $\mathcal{B}$ ($\mathcal{S}$) is the set of buyers (sellers). By the revelation principle (see e.g. Myerson, 1981), the focus on direct mechanisms is without loss of generality. A direct mechanism $\langle Q, M\rangle$ is (Bayes-Nash) incentive compatible if, knowing his own type and the functions $\langle Q, M\rangle$ and the distributions F and G, each agent's expected payoff is maximized when reporting his type truthfully, with expectations taken with respect to F and G, assuming all other agents report truthfully. The mechanism is individually rational when this expected payoff is not less than 0.

Let $\Phi(v) := v - (1 - F(v))/f(v)$ and $\Gamma(c) := c + G(c)/g(c)$ denote the virtual value and virtual cost functions associated with F and G, respectively. We assume that the functions $\Phi(F^{-1}(1 - q))$ and $\Gamma(G^{-1}(q))$ intersect once for $q \in [0, 1]$. The regularity condition that $\Phi(v)$ and $\Gamma(c)$ are monotone functions is sufficient without being necessary for this. The assumption of a unique point of intersection guarantees that the design problem in the limit is quasiconcave. Let (p_B, p_S) be the unique solution to

$$1 - F(p_B) = G(p_S) \tag{8}$$

and

$$\Phi(p_B) = \Gamma(p_S). \tag{9}$$

With a continuum of buyers and sellers each with mass 1, the following result is true:

Proposition 6. *Consider an intermediary who faces a continuum of buyers and sellers who draw their types independently from distributions F and G. The optimal mechanism for the intermediary that respects agents' incentive compatibility and individual rationality constraints is the wholesale model with prices p_B and p_S satisfying (8) and (9).*

Proof of Proposition 6. As is well known, the revenue equivalence theorem implies that up to additive constants payments and expected revenue are pinned down by the allocation rule of a mechanism (see e.g. Myerson, 1981; Riley and Samuelson, 1981; Krishna, 2002). With a profit-maximizing mechanism, the additive constants are pinned down by the agents' individual rationality constraints. The allocation of the optimal mechanism with N buyers and N sellers who draw their types $\boldsymbol{v} = (v_1, .., v_N)$ and $\boldsymbol{c} = (c_1, ..c_N)$ independently from the distributions

F and G induces the q buyers with the highest values and the q sellers with the lowest costs to trade, where q is such that

$$\Phi(v_{(q)}) \geq \Gamma(c_{[q]}) \quad \text{and} \quad \Phi(v_{(q+1)}) < \Gamma(c_{[q+1]})$$

with $v_{(q)}$ denoting the qth highest element of $\boldsymbol{v}$ and $c_{[q]}$ denoting the qth lowest element of $\boldsymbol{c}$ (and with $c_{[0]} = 0 = v_{(N+1)}$ and $c_{[N+1]} = 1 = v_{(0)}$ to make sure q is well defined). This is a conceptually straightforward generalization of the broker-optimal mechanism derived by Myerson and Satterthwaite (1983) for the case $N = 1$.

In the dominant strategy implementation of the optimal mechanism, trading buyers pay $\max\{v_{(q+1)}, \Phi^{-1}(\Gamma(c_{[q]}))\}$ and trading sellers are paid $\min\{c_{[q+1]}, \Gamma^{-1}(\Phi(v_{(q)}))\}$.

We now show that as N goes to infinity, the buyers' payments converge to p_B^0 and the sellers' payments converge to p_S^0 and the quantity traded converges to q^*.

$$\operatorname*{plim}_{N\to\infty} v_{(q)} = \operatorname*{plim}_{N\to\infty} v_{(q+1)} = \operatorname*{plim}_{N\to\infty} \Phi^{-1}(\Gamma(c_{[q]})) =: p_B. \tag{10}$$

Similarly,

$$\operatorname*{plim}_{N\to\infty} c_{[q]} = \operatorname*{plim}_{N\to\infty} c_{[q+1]} = \operatorname*{plim}_{N\to\infty} \Gamma^{-1}(\Phi(v_{(q)})) =: p_S, \tag{11}$$

while the fraction of buyers and sellers who trade satisfy

$$\operatorname*{plim}_{N\to\infty} \frac{q}{N} = \operatorname*{plim}_{N\to\infty} \frac{\max\{i|v_{(i)} \geq p_B\}}{N} = 1 - F(p_B), \tag{12}$$

$$\operatorname*{plim}_{N\to\infty} \frac{q}{N} = \operatorname*{plim}_{N\to\infty} \frac{\max\{i|c_{[i]} \geq p_S\}}{N} = G(p_S). \tag{13}$$

(10), (11), (12), and (13) imply that the optimal mechanism converges to wholesale with p_B and p_S that satisfy $\Phi(p_B) = \Gamma(p_S)$ and $(1 - F(p_B)) = G(p_S)$.

□

C Model with Binary Types

In this appendix, we provide a model with discrete – indeed, binary – types, the main purpose being that this specification more readily admits a dynamic extension than a model with a continuum of types. It also allows us to show that our main findings are robust with respect to type distributions and dynamics.

C.1 Setup

We now assume that the type sets are $\mathcal{V} = \{v_1, v_2\}$ for buyers and $\mathcal{C} = \{c_1, c_2\}$ for sellers, and we impose symmetry in the sense that $G(c_1) = h$ and $F(v_1) = 1 - h$, so that $1 - F(v_2) = h$ and $1 - G(c_2) = 1 - h$. We also assume $v_2 > v_1 \geq c_2 > c_1$, which implies that it is efficient for all sellers to produce and for all buyers to buy one unit.

While a binary type space is more tractable for some questions, the specification does come at a cost, since one can get results which are an artefact of the discrete type space: Demand functions are not strictly downward sloping and continuous anymore, supply functions are not strictly upward sloping and continuous. This can lead to multiple prices or multiple quantities being in equilibrium in several specifications. To avoid such issues we assume that the fraction of efficient buyers v_2 and the fraction of efficient sellers c_1 is both h.

Another way of getting rid of artefacts of a discrete type space would be to add small perturbations of the type space, so that buyers' valuation would be either in an ϵ-environment of v_1 or an ϵ-environment of v_2 and the same for the seller. However, this would lead to a considerably more tedious notation.

C.2 Equilibrium analysis

We now derive the equilibrium outcome for a given choice of mechanism by the market maker.

Wholesale If the intermediary wants to induce the Walrasian traded quantity (full trade) using wholesale, he optimally sets $p_B = v_1$ and $p_S = c_2$ and nets a profit of $\Pi^* = v_1 - c_2$, with w standing for Walrasian sets. If there is no random matching market, his profit when trading only with the most efficient set of traders (that is, v_2 and c_1) is the (restricted quantity) monopoly profit $\Pi_m = h(v_2 - c_1)$.

If $\Pi^* \geq \Pi_m$, a profit maximizing intermediary implements first-best. To make sure that there is a deadweight loss of monopoly, we assume $\Pi^* < \Pi_m$, which is equivalent to

$$h > \frac{v_1 - c_2}{v_2 - c_1} =: \mu, \tag{14}$$

where μ is the ratio of markups under full trade and under exclusive trade (i.e.

with efficient buyer and seller types only). From here onwards, assume that (14) holds. Note that for the continuous type model studied in the main text, there is always a deadweight loss of monopoly. Consequently, assumption (14) should be seen as avoiding an artefact of a discrete type space.

Wholesale with random matching market Let λ be the matching probability in the random matching market and assume that trade takes place at the expected price $(v_1 + c_2)/2$. This price can be due to any of the bargaining protocols mentioned in the main text, i.e. random proposal take-it-or-leave-it-offers, Nash bargaining, double auctions, or fixed-price bargaining.[33] For a buyer of type v_2 and a seller of type c_1, the expected payoffs of participating in the random matching market, denoted, respectively, as $V_B(v_2)$ and $V_S(c_1)$, are

$$V_B(v_2) = \lambda \left(v_2 - \frac{v_1 + c_2}{2} \right) \quad \text{and} \quad V_S(c_1) = \lambda \left(\frac{v_1 + c_2}{2} - c_1. \right).$$

The prices (p_B, p_S) the intermediary sets when trading with the efficient types only satisfy

$$v_2 - p_B = V_B(v_2) \quad \text{and} \quad p_S - c_1 = V_S(c_1),$$

or equivalently

$$p_B = (1 - \lambda)v_2 + \lambda \frac{v_1 + c_2}{2} \quad \text{and} \quad p_S = (1 - \lambda)c_1 + \lambda \frac{v_1 + c_2}{2}.$$

The intermediary's profit with an active random matching market (which occurs if he trades only with the efficient types) is therefore

$$\Pi_W = h(p_B - p_S) = h(1 - \lambda)(v_2 - c_1) = (1 - \lambda)\Pi_m.$$

Observe that Π_W goes to 0 as λ goes to 1. Notice also that Π^* approaches 0 as μ approaches 0. Therefore, there will be parameter constellations such that entry-deterring agency will be profitable if it leads to a strictly positive profit.

Another relevant comparison is between Π_W and the Walrasian profit Π^*. One can show that

$$h(1 - \lambda)(v_2 - c_1) = \Pi_W > \Pi^* = v_1 - c_2$$

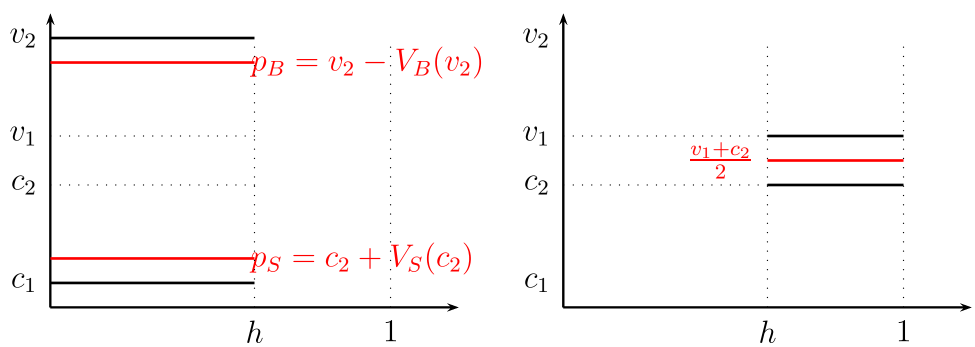
[33] It does not matter for our results, which price in the interval $[c_2, v_1]$ is chosen, but it is convenient to have a price exactly in the middle of the interval. To get this, for random proposal take-it-or-leave-it-offers, the offer probability has to be $\alpha = 1/2$. For Nash bargaining, the bargaining weight has to be $\alpha = 1/2$. For double auctions and fixed-price bargaining, any price in the interval $[c_2, v_1]$ could be the transaction price, including $(v_1 + c_2)/2$.

is equivalent to

$$\lambda < 1 - \frac{\mu}{h} = \frac{h - \mu}{h} =: \lambda^*.$$

Note that for $\mu \to 0$ (which is equivalent to $v_1 \to c_2$), this condition is always satisfied. $\mu \to 0$ is reasonable to consider, given that it implies zero profits when implementing the Walrasian allocation, which always holds for a continuous type space.

The intermediated market and the bilateral exchange with wholesale are illustrated in Fig. 7.



(a) Intermediated Market (b) Bilateral Exchange

Figure 7: The intermediated and the bilateral market with wholesale.

Agency Assume now that instead of the wholesale model, the intermediary employs a large number of brokers (with mass 1 or larger) and randomly matches buyers and sellers one-to-one across the brokers if the number of buyers and sellers joining the intermediated market is the same. (Otherwise, agents on the long side are rationed randomly). The agency model used by the brokers has the fee $\underline{\omega}$ for the low price $\underline{p}$ and the fee $\overline{\omega}$ upon the high price $\overline{p}$ (and appropriately chosen ω for any other price, which will be off equilibrium). We now derive the optimal structure of this agency model.

The intermediary wants the efficient seller to set $\underline{p}$, which will be accepted by both buyers, and the inefficient seller to set $\overline{p}$, which will only be accepted by the high type buyer. Accordingly, the incentive constraints for the sellers are

$$h(\overline{p} - \overline{\omega} - c_2) \geq \underline{p} - \underline{\omega} - c_2 \tag{15}$$

$$h(\overline{p} - \overline{\omega} - c_1) \leq \underline{p} - \underline{\omega} - c_1. \tag{16}$$

Inequalities (15) and (16) are equivalent to

$$h(\overline{p} - c_2) - \underline{p} + c_2 \geq h\overline{\omega} - \underline{\omega} \tag{17}$$

$$h(\overline{p} - c_1) - \underline{p} + c_1 \leq h\overline{\omega} - \underline{\omega}. \tag{18}$$

As $h(\overline{p}-c_2)-\underline{p}+c_2 \geq h(\overline{p}-c_1)-\underline{p}+c_1$ is equivalent to $c_2 \geq c_1$, which is assumed (with strict inequality), we know that there is a number $h\overline{\omega} - \underline{\omega}$ such that both incentive constraints are satisfied.

The individual rationality constraints for the sellers are

$$\overline{p} - \overline{w} - c_2 \geq 0 \tag{19}$$

$$\underline{p} - \underline{w} - c_1 \geq 0. \tag{20}$$

Making the individual rationality constraint for the inefficient seller type binding, we get $\overline{w} = \overline{p} - c_2$. Using this and making the incentive constraint for the efficient seller type bind, we get $\underline{\omega} = \underline{p} - c_1 - h(c_2 - c_1)$. It is easy (and routine) to verify that the individual rationality constraint for the efficient seller type and the incentive constraint for the inefficient seller type will be satisfied with slack. The intermediary's profit under entry-deterring agency is therefore

$$\Pi_A \quad = \quad h\underline{\omega} + h(1-h)\overline{\omega} = h(\underline{p} + \overline{p} - c_1 - c_2) - h^2(\overline{p} - c_1),$$

since with probability h, the seller is efficient and trades for sure (fee $\underline{\omega}$), and with probability $1-h$, the seller is inefficient and trades with probability h (fee $\overline{\omega}$). Of course, $\overline{p} = v_2$ and $\underline{p} = v_1$, so that

$$\underline{\omega} = v_1 - c_1 + h(c_2 - c_1) \quad \text{and} \quad \overline{\omega} = v_2 - c_2. \tag{21}$$

Plugging this into Π_A yields

$$\Pi_A = h(1-h)(v_2 - c_1) + h(v_1 - c_2) = h(v_2 - c_1)(1 - h + \mu) = (1 - h + \mu)\Pi_m,$$

which is less than Π_m under condition (14).

Equilibrium mechanisms If we assume that $\mu = 0$, we get

$$\Pi_A = h(1-h)(v_2 - c_1). \tag{22}$$

Therefore, under the assumption that $\mu = 0$, $\Pi_A > \Pi_W$ is equivalent to $\lambda > h$. Under this assumption, the fees simplify to $\overline{w} = v_2 - c_2$ and $\underline{w} = (1-h)(v_1 - c_1)$. In general, $\Pi_A > \Pi_W$ can be rearranged to

$$\lambda > \lambda_\Pi := h - \mu$$

Note that $\lambda_\Pi > 0$ by condition (14), which states that $h - \mu > 0$. Summarizing, we have established the following result.

Proposition 7. *Entry-deterring agency is profitable for the intermediary if and only if the competing exchange were otherwise sufficiently efficient, that is if and only if $\lambda > h - \mu$.*

A further comparison worthwhile making concerns agency and wholesale with the Walrasian quantity. We know that for λ sufficiently large, the intermediary prefers Walrasian wholesale (v_1, c_2) to posting a large spread (v_2, c_1). One may wonder whether Walrasian wholesale may be preferred to agency. This is, however, not the case for the following reason. Fee setting generates higher profits than Walrasian wholesale if

$$h(v_2 - c_1)(1 - h + \mu) = \Pi_W > \Pi^* = v_1 - c_2,$$

which can be rearranged to $h > \mu$, which is satisfied by assumption. This means that agency is always preferred to Walrasian wholesale.

Note further that whenever wholesale with a large spread is preferred to agency ($\lambda < \lambda_\Pi$), it is also preferred to Walrasian wholesale, since $\lambda^* > \lambda_\Pi$. (This actually also follows from the fact that agency is always preferred to Walrasian wholesale.)

The agency model is illustrated in Fig. 8.

C.3 Dynamic random matching

Dynamic effects arise naturally in the context of intermediated trade because agents who are not matched or do not trade today have the option of trading tomorrow.[34] We now extend the static model to account for these effects.

[34]See, for example, Spulber (1996), Rust and Hall (2003) and Duffie, Garlenau, and Pedersen (2005).

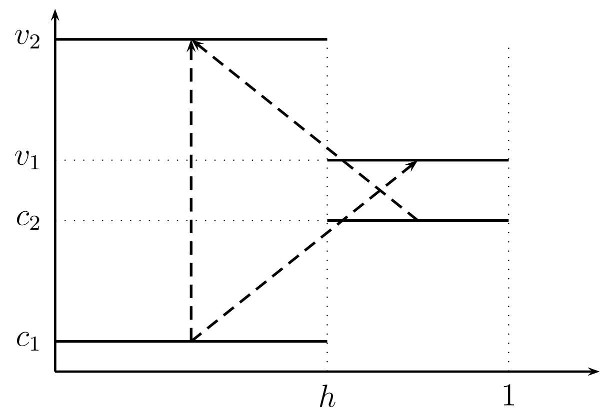


Figure 8: Trade in the intermediated market with agency.

Setup We consider the following dynamic random matching extension of our baseline model. There is an infinite horizon model with periods $t = 0, 1, 2, ...$. Initially both the intermediated and the direct market are empty. In each period mass 1 of buyers and mass 1 of sellers enter the market. If a trader trades, he leaves the market. If he does not trade, he stays in the market with the survival probability δ. With probability $1-\delta$ a buyer or seller who does not trade drops out for exogenous reasons and gets utility 0. These assumptions are similar to those in Satterthwaite and Shneyerov (2008)) and Shneyerov and Wong (2010). The assumptions on the distribution of types are the same as in the static model, that is every buyer is of type v_2 with probability h and of type v_1 with probability $1-h$ while every seller is of type c_1 with probability h and of type c_2 with probability $1-h$, where $v_2 > v_1 \geq c_2 > c_1$.

Before we solve the dynamic random matching model, we briefly discuss the rationale for these assumptions. The literature on dynamic random matching typically uses one of the following three modeling assumptions to model impatience of participants in a market: (i) fixed search costs per period of participation, (ii) discounting, (iii) random drop-out of traders in every period. The third modeling assumption has several advantages in our setup. First, it is a parsimonious way of having both something similar to discounting (as in (ii)) and making sure that no traders stay in the market forever (as in (i)). The assumption also has the effect that the intermediary is also impatient: if a trader does not trade in a certain period, he may drop out in the next, which means a potentially foregone opportunity to extract rents for the intermediary. Another advantage is that a model with drop-out nests the static model we discussed before: if the probability of dropping out after a period without trade is 1, we are essentially in the static setup (repeated indefinitely). A further advantage is that the intermediary can simply focus on the stationary equilibrium of the market.[35] With discounting, one would have to consider both profits in the stationary (limiting) equilibrium and profits on the transition path to this equilibrium. An intermediary whose impatience stems only from drop-outs can ignore the transition path.

One can also think of a larger model, in which all three sources of impatience exists, but search costs converge to zero and the discount factor converges to 1.

Wholesale In the intermediated market, buyers with v_2 and sellers with c_1 enter and trade immediately.[36] In the non-intermediated market, buyers with v_1 and sellers with c_2 enter. Their per period probability of trade is λ. If they do not trade, they stay in the market with probability δ and may trade in any of the subsequent periods, provided they do not drop out. Hence, the ultimate probability of trade is

$$\hat{\lambda} = \sum_{t=0}^{\infty} [(1-\lambda)\delta]^t \lambda = \frac{\lambda}{1-\delta(1-\lambda)}$$

It is easy to see that for $\delta \to 0$, $\hat{\lambda} \to \lambda$, and for $\delta \to 1$, $\hat{\lambda} \to 1$.

Since only λ is replaced by $\hat{\lambda}$ and everything else remains the same, the intermediary's per period profits are

$$\Pi_W = h(1-\hat{\lambda})(v_2 - c_1).$$

Note that the usual subtleties due to per period vs per cohort profits do not occur here, since traders joining the intermediary's platform trade immediately, so that per period and per cohort profits are the same.

[35] One has to be careful, when analyzing the stationary equilibrium though, since one has to look at per entering cohort profits rather than per period profits. We will discuss this later on.

[36] For $\lambda = 0$, It can be shown that this is the optimal mechanism, irrespective of δ, see e.g. Niedermayer and Shneyerov (2014), even if one allows for non-stationary, non-anonymous mechanisms.

Agency Next, consider agency. We will consider the same type of equilibrium as in the static setup: all traders join the intermediary's platform, sellers with low costs c_1 set a low price $\underline{p}$, sellers with high costs c_2 set a high price $\overline{p}$. In equilibrium, high valuation buyers accept both the high and the low price, low valuation buyer only accept the low price.

The dynamic random matching model differs from the static setup in two ways. First, the distribution of types in the market differ from the entrant population as less efficient traders spend more time in the market. Second, traders have an option value of delaying trade and trading with a potentially more attractive future trading partner. Note that by assuming that all impatience stems from the drop-out probability, we can focus on per cohort profits in the steady state equilibrium rather than on the complicated transition path to the steady state.

Because of symmetry of the buyer and seller probabilities of being efficient, it is sufficient to analyze efficient vs. inefficient traders. The analysis then applies to both the buyer and the seller side. Note that efficient traders (v_2 and c_1) trade immediately, so the mass of efficient traders in the market is equal to the mass of efficient entering traders, h. Denote the mass of inefficient traders in the market as m and the fraction of efficient traders in the market as $\tilde{h}$. The following has to hold:

$$\frac{m}{m+h} = 1 - \tilde{h},$$

which is equivalent to

$$m = h\frac{1-\tilde{h}}{\tilde{h}}.$$

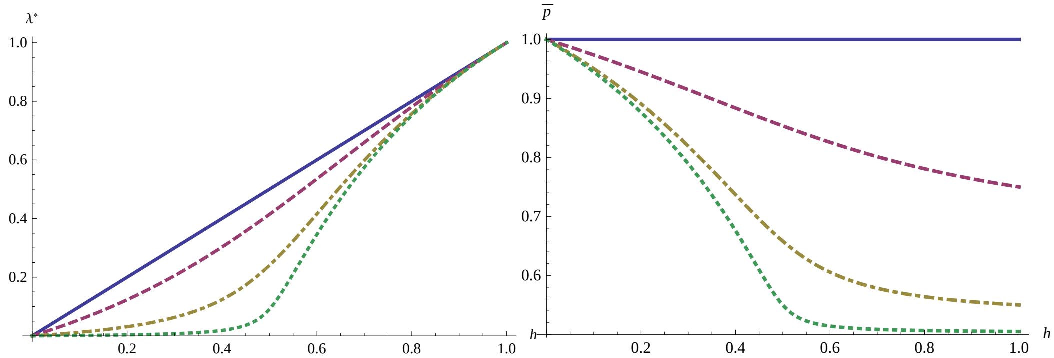
In a stationary equilibrium, the inflow and the outflow of a certain type of agents has to be equal. For the efficient types, this clearly holds, since they trade with probability 1. For the inefficient agents, the inflow has mass $1 - h$. The outflow is given by the mass m of agents in the market and the probability of not staying in the market. The probability of staying in the market is given by the probability of not trading $1 - \tilde{h}$ and the probability of not dropping out δ. The inflow-outflow equilibrium equation is hence

$$1 - h = m(1 - (1 - \tilde{h})\delta)$$

Plugging in the expression for m yields a quadratic equation in $\tilde{h}$. Rearranging and solving for $\tilde{h}$ gives the solution[37]

$$\tilde{h} = \frac{2\delta h - 1 + \sqrt{(2\delta h - 1)^2 + 4\delta(1-\delta)h^2}}{2\delta h}$$

It is easy to check th at for $\delta \to 0$, $\tilde{h} \to h$ and for $\delta \to 1$, $\tilde{h} \to 2 - 1/h$. The



(a) λ_Π as a function of h for $\delta = 0$ (solid), $\delta = 0.5$ (dashed), $\delta = 0.9$ (dash-dotted), and $\delta = 0.99$ (dotted).

(b) $\overline{p}$ as a function of h for $\delta = 0$ (solid), $\delta = 0.5$ (dashed), $\delta = 0.9$ (dash-dotted), and $\delta = 0.99$ (dotted).

Figure 9: λ_Π and $\overline{p}$.

latter implies that the system is stable for $\delta \to 1$ if $h \in (1/2, 1)$. It is also easy to check that for $h \to 0$, $\tilde{h} \to 0$ and for $h \to 1$, $\tilde{h} \to 1$. From here onwards, we therefore assume

$$\max\{1/2, \mu\} < h < 1$$

Though $\tilde{h}$ is the probability that a trader meets an efficient potential trading partner in a given period, what matters for the decisions of traders is the ultimate probability of meeting an efficient type. This probability is given as

$$\hat{h} = \sum_{t=0}^{\infty} \left[\delta(1-\tilde{h})\right]^t \tilde{h} = \frac{\tilde{h}}{1-\delta(1-\tilde{h})}$$

Plugging in $\tilde{h}$ yields

$$\hat{h} = \frac{2h}{1 + \sqrt{1 - 4(1-h)h\delta}}$$

[37]The two candidates to the quadratic equation are $\tilde{h}_{1,2} = \frac{2\delta h - 1 \pm \sqrt{(2\delta h-1)^2 + 4\delta(1-\delta)h^2}}{2\delta h}$. Since the square root is greater than $2\delta h - 1$, the solution with a minus sign would yield a negative value, contradicting that it is a probability. Therefore, the solution is given by the expression with the plus sign.

Observe that for $h \geq 1/2$, $4(1-h)h\delta \leq 1$. Therefore, the square root is always real.

Next, we can derive the incentive compatibility and individual rationality constraints of traders. Note that since the setup is stationary, we only need to check whether a high cost seller has the incentive to set a high price in every period rather than a low price in every period, since the trade-off is the same in every period. Similarly, we only need to check that a low cost seller has the incentive to set a low price in every period rather than a high price in every period. Further, one can simply replace h with $\hat{h}$ for most of the analysis. This yields the incentive compatibility constraints for the sellers

$$\hat{h}(\overline{p} - \overline{w} - c_2) \geq \underline{p} - \underline{w} - c_2 \tag{23}$$

$$\hat{h}(\overline{p} - \overline{w} - c_1) \leq \underline{p} - \underline{w} - c_1 \tag{24}$$

The individual rationality constraints can be written analogously. By the same logic as for the static setup, we get the fees

$$\overline{w} = \overline{p} - c_2, \tag{25}$$

$$\underline{w} = \underline{p} - c_1 - \hat{h}(c_2 - c_1). \tag{26}$$

For the inefficient buyer, the option value of future trade is 0 as he will never get an offer below $\underline{p}$ no matter how long he waits. Hence, by the same logic as before, incentive compatibility and individual rationality constraints imply that $\underline{p} = v_1$.

For the efficient buyer, the situation is somewhat more complicated than in the static setup: an efficient buyer getting a high price offer $\overline{p}$ has the option of delaying trade and potentially getting a low offer $\underline{p}$ in the future. Hence, the efficient buyer's incentive compatibility constraint is

$$v_2 - \overline{p} \geq \delta\hat{h}(v_2 - \underline{p}),$$

where the left-hand side is the utility from accepting a high offer immediately and the right-hand side is the value of waiting a period and then staying in the market until getting a low offer. Again, stationarity makes sure that this condition is sufficient since rejecting a high offer in the current period and accepting a high offer some time in the future cannot be optimal. Rearranging yields

$$\overline{p} = (1 - \delta\hat{h})v_2 + \delta\hat{h}v_1.$$

This allows us to write per cohort profits. Note that considering per cohort profits is the right measure, since the mass of entering traders per period is exogenously given. The mass of efficient sellers entering per period is h, each trading immediately and generating profits $\underline{w}$. The mass of inefficient sellers entering per period is $1-h$, each generating profit $\overline{w}$ with the ultimate probability $\hat{h}$. Hence, profits are

$$\Pi_A = h\underline{w} + (1-h)\hat{h}\overline{w} = h(\underline{p} - c_1) + \hat{h}((1-h)\overline{p} - (c_2 - hc_1)). \tag{27}$$

Entry-deterrence We now compare profits for wholesale and agency. Note that Π_W decreases in $\hat{\lambda}$ and $\hat{\lambda}$ increases with λ, whereas Π_A is independent of $\hat{\lambda}$ (and λ). Therefore, if $\Pi_W = \Pi_A$ for some λ_Π, then $\Pi_W < \Pi_A$ for all $\lambda > \lambda_\Pi$. One can show that λ_Π exists and is unique, since for $\lambda = 1$, $\Pi_W = 0 < \Pi_A$, for $\lambda = 0$ we know that the optimal mechanism is wholesale, and Π_W is continuous and strictly decreasing in λ.

We can get λ_Π by solving $\Pi_W = \Pi_A$ in closed form:

$$\lambda_\Pi = \frac{2(h-\mu)(1-\delta)}{1 - 2\delta(h-\mu) + \sqrt{1 - 4(1-h)h\delta}}.$$

or

$$\lambda_\Pi = \frac{2(h-\mu)(1-\delta)}{2\delta\mu + \sqrt{(2\delta h - 1)^2 + 4\delta(1-\delta)h^2} - (2\delta h - 1)}.$$

With some algebra, it can be shown that $\lambda_\Pi \in (0,1)$ if $h > \mu$. This also implies that $\Pi_A > 0$, since for $\lambda = 1$, $\Pi_W = 0$.

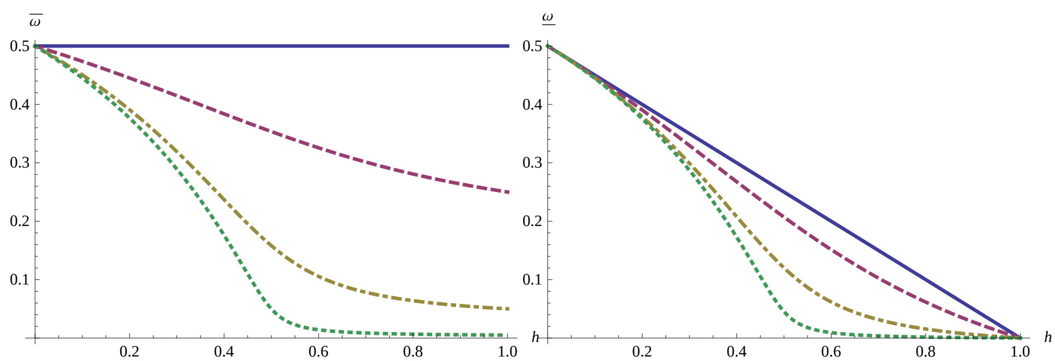
One can also show that agency is always preferred to Walrasian wholesale, just as for the static setup. Note that for Walrasian wholesale, profits are the same as in the static setup since traders trade immediately. That is, profits are $v_1 - c_2$. Again, it can be shown that $h > \mu$ implies $\Pi_A > v_1 - c_2$.

For the special case $v_1 = c_2$ (or, equivalently, $\mu = 0$), this further simplifies to

$$\lambda_\Pi = \frac{2h(1-\delta)}{1 - 2h\delta + \sqrt{1 - 4(1-h)h\delta}}.$$

We can plot λ_Π as a function of h for different values of δ as depicted in Figure 9. For $\delta = 0$ we are back in the static setup and $\lambda_\Pi = h$ as before. The figure illustrates that as δ increases, λ_Π becomes lower, that is as the market becomes more dynamic (or as frictions become smaller), it is more likely that agency is preferred by the intermediary. This can be shown to hold in general, that is, $\partial\lambda_\Pi/\partial\delta < 0$ holds under the assumption that $h > \mu$.

Additionally, one may wonder how prices and fees change as δ changes. To provide some numerical examples, set $v_2 = 1$, $c_1 = 0$ and $v_1 = c_2 = 1/2$. The high price $\overline{p}$ as a function of h is plotted in panel (b) of Figure 9 for the same values of δ as in panel (a). Note that $\underline{p} = v_1$ for any δ and h. $\overline{w}$ and $\underline{w}$ are plotted in panels (a) and (b) of Figure



(a) $\overline{w}$ as a function of h for $\delta = 0$ (solid), $\delta = 0.5$ (dashed), $\delta = 0.9$ (dash-dotted), and $\delta = 0.99$ (dotted).

(b) $\underline{w}$ as a function of h for $\delta = 0$ (solid), $\delta = 0.5$ (dashed), $\delta = 0.9$ (dash-dotted), and $\delta = 0.99$ (dotted).

Figure 10: $\overline{\omega}$ and $\underline{\omega}$.

C.4 Welfare

We now turn to the analysis of the effects of entry-deterring agency on social welfare and on consumer surplus, and we provide a brief assessment of the quantitative effects of entry-deterring agency in our model. We conclude the section with a discussion of policy implications of our simple model. Throughout this section, we confine attention to the basic (that is, static) model.

C.4.1 Effects of Agency on Welfare and Consumer Surplus

When the intermediary posts the prices $p_B = (1-\lambda)v_2 + \lambda\frac{v_1+c_2}{2}$ and $p_S = (1-\lambda)c_1 + \lambda\frac{v_1+c_2}{2}$, welfare S_W under wholesale is

$$S_W = h(v_2 - c_1) + \lambda(1-h)(v_1 - c_2) = (v_2 - c_1)[h + \lambda(1-h)\mu]. \qquad (28)$$

Under wholesale, consumer surplus, defined as the sum of surplus grasped by buyers and sellers and denoted CS_W, is

$$CS_W = \lambda(v_2 - c_1)h + \lambda(1-h)(v_1 - c_2) = \lambda(v_2 - c_1)(h + (1-h)\mu). \qquad (29)$$

Welfare S_A under agency, on the other hand, is

$$\begin{aligned} S_A &= h[hv_2 + (1-h)v_1 - c_1] + (1-h)h[v_2 - c_2] \\ &= h(v_2 - c_1) + h(1-h)(v_1 - c_2) = (v_2 - c_1)[h + h(1-h)\mu]. \qquad (30) \end{aligned}$$

The first equality in (30) follows because the fraction h of the sellers have low costs and trade with all buyers, thereby generating a surplus of $hv_2 + (1-h)v_1 - c_1$ while the remaining sellers have high costs and only trade if matched to a buyer with a high valuation. The second equality follows after rearranging terms. Consumer surplus under agency, denoted CS_A and defined like under wholesale as the sum of buyer and seller surplus, is

$$\begin{aligned} CS_A &= h[hv_2 + (1-h)v_1 - c_1 - \underline{\omega}] + (1-h)h[v_2 - c_1 - \overline{\omega}] \\ &= h^2(v_2 - v_1 + c_2 - c_1) = h^2(v_2 - c_1)(1-\mu), \qquad (31) \end{aligned}$$

where the second line follows after plugging in the value for $\underline{\omega}$ and $\overline{\omega}$ given in (21) and simplifying.

Observe that for $\mu = 0$, we have $S_W = S_A$. For $\mu > 0$, S_W increases in λ while S_A is independent of λ. At $\lambda = h$, we have $S_W = S_A$ and $S_W > S_A$ for any $\lambda > h$, assuming $v_1 > c_2$.[38] Since the condition for agency to be profitable is $\lambda > h - \mu$, this means that (i) agency is profitable whenever it reduces welfare but also, and somewhat surprisingly, (ii) that profitable and welfare enhancing agency is possible, the latter occurring when $h - \mu < \lambda < h$. The intuition seems to be that if h is large relative to λ, the intermediary under agency is a better match maker than is the random matching market.[39]

Notice also that CS_A is independent of λ while CS_W increases in λ. Since at $\lambda = h$, $CS_W > CS_A$, it follows that agency decreases consumer surplus whenever

[38]At $v_1 = c_2$ welfare is the same with agency and wholesale because both mechanisms always induce trade by the efficient agents.

[39]Conditional on being an inefficient type, an agent's probability of trade in the matching market under wholesale is λ while his probability of trading at the market maker with agency is h. Since the efficient types trade with probability 1 regardless of the mechanism used by the intermediary, the result follows.

it decreases total welfare. Because CS_W is continuous in λ, it follows also that the parameter space for which agency is detrimental to consumer surplus is larger than the parameter space for which it is detrimental to welfare. This reflects a theme from Loertscher and Niedermayer (2016), where agency emerges as a tool to extract rents from buyers and sellers.

Somewhat tedious algebra reveals that $CS_A = CS_W$ at $\lambda = \lambda_{CS}$ with

$$\lambda_{CS} := h^2 \frac{1-\mu}{h+(1-h)\mu}. \tag{32}$$

Since $\lambda_\Pi = h - \mu < \lambda_{CS}$, entry deterrence is profitable and increases consumer surplus for $\lambda \in \left(h-\mu, h^2\frac{1-\mu}{h+(1-h)\mu}\right)$. Summarizing, we have established the following two propositions, where the comparisons are made with welfare and, respectively, consumer surplus under (selective) wholesale.

Proposition 8. 1. *("Bad wholesale equilibrium outcome") If $\lambda < h - \mu$, agency achieves higher welfare than wholesale, but a profit maximizing intermediary will choose wholesale.*

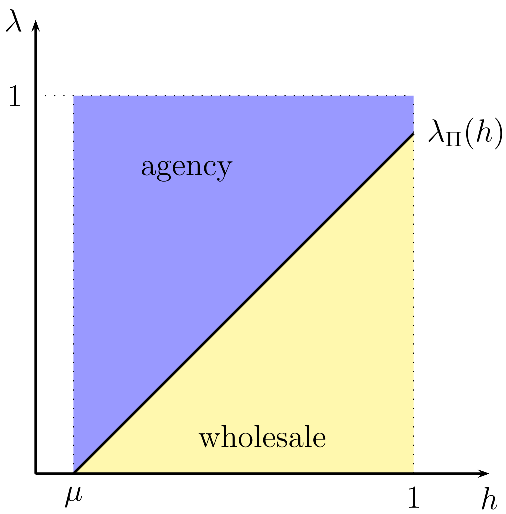
2. *("Good agency equilibrium outcome") If $\lambda \in [h-\mu, h]$, agency achieves higher welfare and a profit maximizing intermediary will choose agency.*

3. *("Entry-deterring agency equilibrium outcome") If $\lambda > h$, wholesale achieves higher welfare, but a profit maximizing intermediary will choose agency.*

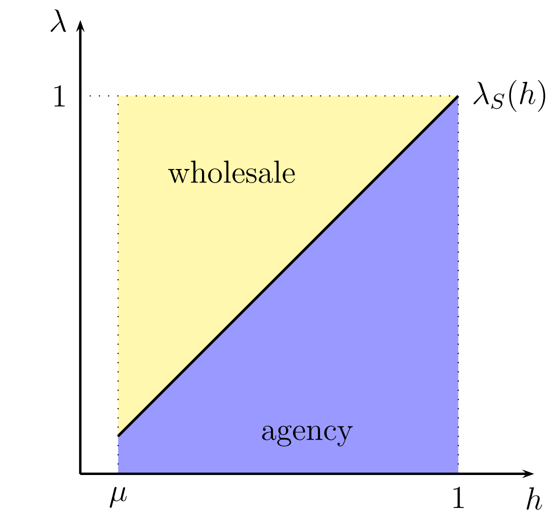
The possibility of positive equilibrium effects on welfare and consumer surplus of entry deterrence are reminiscent of Cabral and Riordan (1994, 1997), who show that in the presence of a learning curve these effects can go either way. Note, however, that in continuous type space model, the profit generated by choosing the Walrasian quantity is zero, which corresponds to a μ close to 0 in our setup. For μ close to zero, the middle case ("good agency equilibrium outcome") vanishes.

Proposition 9. *For $\lambda \in \left(h-\mu, h^2\frac{1-\mu}{h+(1-h)\mu}\right)$, entry-deterring agency increases consumer surplus while for $\lambda > h^2\frac{1-\mu}{h+(1-h)\mu}$ it decreases consumer surplus.*

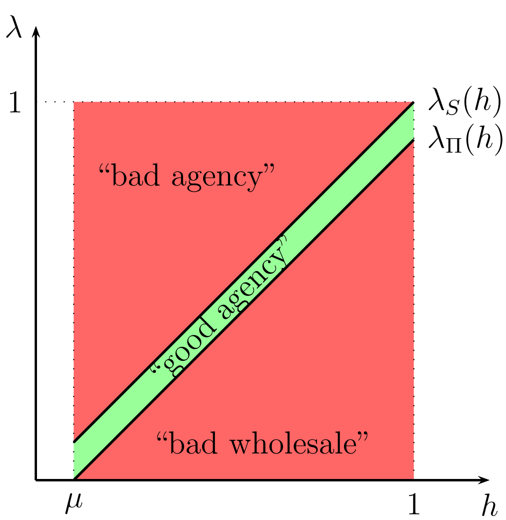
Note again that for $\mu \to 0$, both λ_Π and λ_{CS} go to h, which means that consumer surplus always decreases when the intermediary prefers agency.



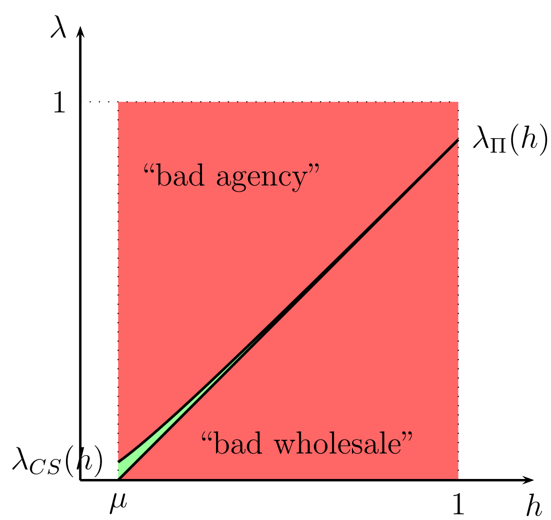
(a) Profits



(b) Welfare



(c) Profits vs Welfare



(d) Profits vs Consumer Surplus

Figure 11: The optimal mechanism in term of profits and welfare (Subfigures 11a and 11b); comparison of profits vs welfare and profits vs consumer surplus (Subfigures 11c and 11d) with $\lambda_\Pi(h) = h - \mu$, $\lambda_S(h) = h$, and $\lambda_{CS}(h) = h - \frac{h\mu}{h(1-\mu)+\mu}$