THE RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

Working Paper No. 555, 2001

Bilateral Oligopoly

by Jonas Björnerstedt and Johan Stennek

IUI, The Research Institute of Industrial Economics
P.O. Box 5501
SE-114 85 Stockholm
Sweden

Bilateral Oligopoly*

Jonas Björnerstedt
IUI

Johan Stennek
IUI and CEPR

May 13, 2002

Abstract

In intermediate goods markets, both buyers and sellers normally have market power, and sales are based on bilaterally negotiated contracts specifying both price and quantity. In our model, pairs of buyers and sellers meet in bilateral but interdependent Rubinstein-Ståhl negotiations. The outcome has a simple characterization (a Nash equilibrium in Nash bargaining solutions) suitable for applied work. Equilibrium quantities are efficient regardless of concentration and also with few "trading links". The law of one price does not hold. In addition to relation-specific characteristics, prices depend on both upstream and downstream concentration and on the structure of trading links. The requirements necessary for Walrasian prices are stronger than usually believed.

Key Words: bilateral oligopoly; bargaining, intermediate goods, decentralized trade, Walrasian outcome

JEL classification: C7 L1 L4 D2 D4

*Our work has benefitted greatly from discussions with Sven-Olof Fridolfsson, Paul Heidues, Lars Persson and Jörgen Weibull. We are also grateful for comments from participants in the seminars at the Stockholm School of Economics, the Universitat Autonoma de Barcelona, EEA'00 in Bolzano, EARIE'00 in Lausanne, IUI/CEPR workshop on mergers, and WZB in Berlin. We thank Nancy Adler for editorial support. Both authors would like to thank Jan Wallander's and Tom Hedelius' Foundation for financial support.
Postal address: Research Institute of Industrial Economics (IUI), Box 5501, SE-114 85 Stockholm, Sweden. E-mail: JonasB@iui.se; JohanS@iui.se.

1 Introduction

Retailing is traditionally viewed as a industry characterized by many firms and easy entry. Over the last forty years, however, concentration has greatly increased. Several retail sectors are now highly concentrated as a result of organic growth and of mergers or acquisitions. A notable example of this is the grocery sector, where the creation of buyer groups has reinforced the trend. In 1996 the five largest groups in Germany accounted for 83 percent of the market. In Belgium the corresponding figure was 82 percent, and in the Netherlands 72 percent (Dobson and Waterson, 1999). Since manufacturing is by tradition concentrated, this implies that many intermediate goods markets today are characterized by high concentration on both the seller and the buyer sides. As a result, buyers and sellers both exercise market power: the market structure is that of a bilateral oligopoly.

Contracts in intermediate goods markets are usually long-term and are negotiated bilaterally. They codify many other elements as well as price. An estimated 80 to 90 percent of all intermediate goods are traded through extended term contracts, often lasting one year or more. Spot markets, organized as exchanges or auctions, are just the tiny tip of a huge market of such one-to-one contract deals (The Economist, 2000, pp. 93-94).

Our understanding of such bilateral oligopolies is very incomplete, even when it comes to the most basic microeconomic issues. How much is traded on decentralized markets with high concentration on both sides? Does the presence of market power and externalities mean that this quantity is inefficient? What is the price, and to what extent does the distribution of the surplus depend on the number of buyers and sellers? Under what conditions will buyers concentrate their purchases to a single source, and under what conditions will they purchase from a number of different sources?

This paper presents a model of decentralized bilateral oligopoly that captures four key institutional characteristics that are common to many intermediate goods markets. First, since both sides of the market are concentrated, the firms on both sides wield market power; buyers and sellers both affect the prices at which they trade. Second, contracts are determined in decentralized negotiations between pairs of buyers and sellers; these negotiations are interdependent and the contract agreed in one negotiation constitutes part of an equilibrium prevailing in the market as a whole.[1] Third, contracts are complex, specifying the quantity and quality of the goods or services as well as the price. Fourth, the buyers and sellers in intermediate goods markets are

[1] Markets with a more centralized structure, such as auctions or exchanges, are analyzed by McAffee & Hendriks (2000) and Kranton and Minehart (2000).

both professional and well-informed parties. In the model, representatives of upstream firms meet with representatives of downstream firms to negotiate contracts specifying prices and quantities in simultaneous Rubinstein-Ståhl bargaining characterized by complete and almost-full information.[2]

We show that there exists a sequential equilibrium, implying immediate agreement on a set of contracts, one for each buyer-seller pair. Each buyer-seller pair agrees on the quantity that maximizes the sum of the two firms' profits, all other contracts being taken as given. This condition is called bilateral efficiency. They also agree on the price that splits their bilateral surplus equally. The bilateral surplus is defined as the increase in the sum of the two firms' profits that is generated by their contract, all other contracts being taken as given. Every contract can in fact be regarded as maximizing an appropriately defined bilateral Nash product, all other contracts being taken as given. The equilibrium set of contracts can thus be characterized as a Nash equilibrium in Nash bargaining solutions.[3]

The second set of results concerns market efficiency. If all upstream firms bargain with all downstream firms, the equilibrium quantities are the same as the Walrasian quantities. A bilateral oligopoly is thus socially efficient independently of market concentration, and thus despite the presence of market power. Intuitively, since every upstream firm meets with every downstream firm in a bilaterally efficient negotiation, the marginal cost for every seller is equal to the marginal value for every buyer. As costs and valuations are equalized in equilibrium, the market is socially efficient.

In a separate section we assume that downstream firms are also engaged in oligopolistic interaction in the final goods market. We show that market power in the intermediate goods market does not cause inefficiencies over and above those resulting from market power in the final goods market.

The "link structure" is a key determinant of efficiency. If some of the upstream firms cannot bargain with some of the downstream firms, the equilibrium may not necessarily be socially efficient. For example, prohibiting international trade typically implies inefficiencies. With an incomplete link structure, however, production is socially efficient conditional on the zero-quantity restrictions implied by the missing links. Moreover, full efficiency only requires a few links. On average, firms only need to trade with two

[2]Non-cooperative models of parallel bilateral bargaining have previously been studied by Horn and Wolinsky (1986b), Hart and Tirole (1990), McAfee and Schwartz (1994) and Stole and Zweibel (1996). However, this literature has not considered bilateral oligopolies with more than one firm on both sides of the market, all with bargaining power, and price/quantity contracts.

[3]We thus provide a non-cooperative foundation for the reduced form used by Davidson (1988), Horn and Wolinsky (1988a), and Dobson and Waterson (1997).

partners.

The third set of results concerns the distribution of surplus. Prices are relation-specific, which means that the law of one price does not hold. Prices depend on several factors, including the concentration of capital. Large firms get better deals than small firms. It may seem surprising that a seller with low marginal costs (large capital stock) charges higher prices than a high-cost firm. The explanation is that the incremental cost may be higher for the low-cost firm in equilibrium. Although the low-cost firm has a lower cost at each output level, it produces more in equilibrium and therefore has a higher cost at the (discrete) margin. The concentration of sales also matters. If sellers have steeper marginal cost curves than buyers, a seller with an even distribution of sales receives higher prices than a seller with more exclusive sales. Since prices are determined by the parties' incremental costs, the side with steeper marginal costs has more to gain from trading at the margin.

If all firms are small, then all prices will be close to the Walrasian price. The "small-size requirement," however, is stronger than has previously been thought. Each firm must account for a small share only of every partner's trading volume. Since the equilibrium is efficient and since trade occurs at Walrasian prices in a large but finite economy and without an auctioneer, our results represent a contribution to the strategic foundations of general equilibrium (for reviews, see Osborne and Rubinstein, 1990; Gale, 2000).

2 The Model

$U < \infty$ upstream firms produce intermediate goods. $D < \infty$ downstream firms buy intermediate goods and sell final goods. For simplicity we assume that for every unit of intermediate goods, one unit of final goods can be produced. Contracts determine the price and quantity of intermediate goods delivered from upstream to downstream firms. A *contract* at time t is a pair $c_{ud}(t) = (q_{ud}(t), p_{ud}(t))$ specifying a quantity $q_{ud}(t)$ and a price $p_{ud}(t)$. A *contract structure* $c(t)$ is a UD-tuple of contracts $c_{ud}(t)$, one for every $(u, d) \in \overline{\Omega}$, where $\overline{\Omega}$ is the set of all UD pairs of upstream and downstream firms. Likewise q and p are the vectors of all q_{ud} and p_{ud}. We write $c \backslash c_{ud}$ to indicate the contract structure given by c for all $(i, j) \neq (u, d)$ and c_{ud} for (u, d). The corresponding conventions are used for vectors q and p.

The short-run cost functions are denoted by $C^u(q)$ and $C_d(q)$ for upstream firm u and downstream firm d respectively. The marginal costs of production are positive, that is $\partial C^u / \partial q_{ud} > 0$ and $\partial C_d / \partial q_{ud} > 0$. A firm's cost is affected by its own production only, that is $\partial C^u / \partial q_{ij} = 0$ if $i \neq u$ and $\partial C_d / \partial q_{ij} = 0$ if $j \neq d$. Production is (strictly) convex if all $C^u(q)$ are (strictly) convex in $\{q_{uj}\}_{j=1}^{D}$ and all $C_d(q)$ are (strictly) convex in $\{q_{id}\}_{i=1}^{U}$. To ensure finite production we assume convexity and that $\partial^2 C^u / \partial (q_{ud})^2 > 0$ and $\partial^2 C_d / \partial (q_{ud})^2 > 0$. If the intermediate goods are homogenous, let $Q^u = \sum_{j=1}^{D} q_{uj}$ and $Q_d = \sum_{i=1}^{U} q_{id}$ be the firm-aggregate quantities.

Throughout this paper the (short-run) cost functions are assumed to satisfy *diseconomies of size*. Firm d's incremental cost for a single product q_{ud} is given by $C_d(q) - C_d(q \backslash q_{ud} = 0)$, and d's incremental cost for a set of products $\Upsilon \subset \{1, ..., U\}$ is given by $C_d(q) - C_d\left(q \backslash \{q_{ud} = 0\}_{u \in \Upsilon}\right)$. It is assumed that the incremental cost for a product set is smaller than the sum of the incremental costs for the single products, that is

$$\sum_{u \in \Upsilon} C_d(q) - C_d(q \backslash q_{ud} = 0) \geq C_d(q) - C_d\left(q \backslash \{q_{ud} = 0\}_{u \in \Upsilon}\right), \tag{1}$$

for all quantity vectors q, subsets of upstream firms Υ, and cost functions C_d. Similarly, all cost functions C^u exhibit diseconomies of size. It is easy to show that Assumption 1 is fulfilled if goods are homogenous and the (short-run) cost function has constant or decreasing returns to scale. If goods are differentiated, the assumption is fulfilled if for example the marginal cost of producing one product is non-decreasing in the quantity of other products produced (all cross-derivatives are non-negative). The role of the assumption is to guarantee that profits are non-negative in equilibrium. Moreover, any firm's incremental profit, for any subset of products, is non-negative in equilibrium.

The per-period profit of an upstream firm u and a downstream firm d is a function of the contract structure $c(t)$ and is given by

$$\pi^u(c(t)) = \sum_{j=1}^{D} p_{uj}(t)\, q_{uj}(t) - C^u(q(t)), \tag{2a}$$

$$\pi_d(c(t)) = \sum_{i=1}^{U} \left(r_{id}(q(t)) - p_{id}(t)\right) q_{id}(t) - C_d(q(t)), \tag{2b}$$

where $r_{id}(q)$ is the price (inverse demand) for q_{id} in the final goods market. For the sake of expositional simplicity we assume that downstream firms take all retail prices $r_{id} = r > 0$ as given. (This assumption is relaxed in Section 5.) Total profit is the discounted sum of per-period profits (2a) and (2b), with the common discount factor δ.

The *bilateral surplus* of a seller-buyer pair (u, d) is the additional aggregate profit of the two firms as generated by their contract, all other contracts taken as given, i.e.

$$\left[\pi^u\left(c\backslash c_{ud}\right)+\pi_d\left(c\backslash c_{ud}\right)\right]-\left[\pi^u\left(c\backslash\left(0,0\right)\right)+\pi_d\left(c\backslash\left(0,0\right)\right)\right]. \tag{3}$$

The bilateral surplus does not depend on p_{ud}.

Second, the quantity q_{ud} is *bilaterally efficient* if it maximizes the bilateral surplus of (u,d). For a set of seller-buyer pairs $\Omega\subseteq\overline{\Omega}$, and a fixed contract structure c, let $N\left(c,\Omega\right)\subset\mathbb{R}_+^{U\times D}$ denote the set of bilaterally efficient quantity vectors, where q_{ud} is bilaterally efficient for all $(u,d)\in\Omega$, and q_{ud} is given by c for $(u,d)\in\overline{\Omega}\backslash\Omega$. A quantity q_{ud} is bilaterally efficient if, and only if, it satisfies

$$\frac{\partial\pi_d\left(c\backslash c_{ud}\right)}{\partial q_{ud}}+\frac{\partial\pi^u\left(c\backslash c_{ud}\right)}{\partial q_{ud}}\leq 0, \tag{4}$$

with equality if $q_{ud}>0$. Thus, the bilaterally efficient quantity for (u,d) depends on the contracts agreed upon by other pairs, that is to say on the contract structure c. Further, a bilaterally efficient quantity vector is a quantity, one for each pair, such that no pair can increase their aggregate profit if all other pairs agree upon their bilaterally efficient quantity. A bilaterally efficient quantity vector does not necessarily maximize aggregate profits. Following standard reasoning for the existence of a pure strategy Nash equilibrium:

Lemma 1 *For any contract structure c and any set of seller-buyer pairs $\Omega\subseteq\overline{\Omega}$, the set of bilaterally efficient quantity vectors $N\left(c,\Omega\right)$ is non-empty.*

All proofs are relegated to the Appendix.

Third, a price p_{ud} yields an *equal split* of the bilateral surplus if

$$\pi^u\left(c\backslash c_{ud}\right)-\pi^u\left(c\backslash\left(0,0\right)\right)=\pi_d\left(c\backslash c_{ud}\right)-\pi_d\left(c\backslash\left(0,0\right)\right). \tag{5}$$

The equal-split price depends on the contracts agreed upon in other negotiations. It is a function of the quantity vector q, but not of other prices.

2.1 Bargaining

A firm is represented by a separate agent in every negotiation in which it is involved. The agents, or representatives, are the players in the game, and they maximize their respective firm's profit. It is assumed that they strictly prefer agreeing at t upon a contract specifying $q_{ud}=0$ to agreeing upon $q_{ud}=0$ at $t+1$ or not agreeing at all.

Time is infinite and at every date there is a bargaining stage-game, with alternating offers. One firm suggests a contract, i.e. a quantity and a price at which the two firms will trade in all future periods. The other firm can either accept or reject the bid. Prior to acceptance there is an implicit contract specifying $q_{ud} = 0$. Once a bid is accepted, the negotiation is ended. The agreed contract is binding in all future periods, and there is no renegotiation. Production occurs at every stage, immediately after the round of negotiations and according to the (possibly implicit) contract $c_{ud}(t)$.

The *link structure* $\Omega \subseteq \overline{\Omega}$ is defined as the set of buyer-seller pairs that can negotiate. We say that the link structure is *complete* if $\Omega = \overline{\Omega}$. If the link structure is incomplete (e.g. due to trade barriers), we simply impose the restriction that $q_{ud}(t) = 0$ for all $(u, d) \in \overline{\Omega} \backslash \Omega$

At time t the bidder and the respondent both know the history h_t that describes all bids and responses in all negotiations up to $t - 1$. The respondent also knows the bid to which he must respond. On the other hand the respondent does not observe any other bids in the same stage game, not even those given to or by other representatives of his own firm. For the representative of upstream firm u in negotiation (u, d), the *strategy* $\overline{\alpha}_{ud}$ is a function that specifies for each history h_t a bid $\overline{b}_{ud}$ if u is making the bid, or a response $\overline{\rho}_{ud}$ conditional on the downstream firm's bid if d is making the bid. For the representative of the downstream firm, $\underline{\alpha}_{ud}$, $\underline{b}_{ud}$ and $\underline{\rho}_{ud}$ are similarly defined. A strategy profile α specifies a strategy for all representatives of all firms. We only consider pure strategies. As this is a game of imperfect information, we use sequential rather than subgame perfect equilibrium.[4]

Proposition 1 *For any link structure $\Omega \subseteq \overline{\Omega}$ there exists a sequential equilibrium, implying immediate agreement on quantity vector q if, and only if, q is bilaterally efficient, i.e. $q \in N(c, \Omega)$. As $\delta \to 1$, the equilibrium contracts imply an equal split of bilateral surpluses.*

There is thus a close connection between immediate agreement and bilaterally efficient quantity vectors. There exists an equilibrium with immediate agreement for *any* bilaterally efficient quantity vector, but not for any other quantity vector. The existence of such bilaterally efficient equilibria is due to complete information, price/quantity contracting, and no trade before agreement.[5]

[4] Essentially, "consistency of beliefs" implies that after receiving an out-of-equilibrium bid, a respondent should expect other bidders to have followed their equilibrium strategies (cf. Rubinstein and Wolinsky, 1990). McAfee and Schwartz (1994) label the same restriction "passive beliefs."

[5] Under two-sided asymmetric information, there may be delay and ex post inefficient agreements (Myerson and Satterthwaite, 1983). Writing contracts for both price and quantity is crucial for bilateral efficiency (McDonald and Solow, 1981). When there is trade before agreement, delay can occur in equilibrium (Haller and Holden, 1990).

To prove this proposition, Lemma 2 (in the Appendix) considers a subgame in which there is one ongoing negotiation only. It is shown that firms agree immediately on the bilaterally efficient quantity and that (as $\delta \to 1$) they split the bilateral surplus equally. This is a simple application of standard Rubinstein-Ståhl bargaining. Lemma 3 considers a subgame with many ongoing negotiations. First, it shows that there exists an equilibrium with immediate agreement on $q \in N(c, \Omega)$ and with an equal split of the bilateral surpluses. A deviation from the prescribed equilibrium in a single negotiation (u, d) will not affect the other negotiations, since it is not observed until the following period. Hence, for all other negotiations $(i, j) \neq (u, d)$, there is immediate agreement on the prescribed contracts. From the point of view of (u, d), all other negotiations can thus be treated as concluded, and Lemma 2 applies for negotiation (u, d). Second, if $q \notin N(c, \Omega)$ is prescribed, there exists at least one pair (u, d) that will improve bilateral efficiency by way of a unilateral deviation according to Lemma 2.

Immediate agreements can be ensured by imposing a Markov restriction. In the formal definition of the Markov restriction we interpret an equilibrium as a profile of beliefs. Let c_{ij}^k be player k's belief about the "outcome" of a negotiation (i, j). For convenience we restrict the "outcome" to the (possibly implicit) contract that (i, j) will implement in the current period. Likewise $c^k = \left(c_{11}^k, ..., c_{UD}^k\right)$ is k's belief about the contract structure that will be implemented in the current period.

Definition 1 *Beliefs c^k are Markov if:*

1. *Belief c^k is a function of (c, Ω) only, and*

2. *Consider two states (c, Ω) and $(\widetilde{c}, \widetilde{\Omega})$ and assume that $\widetilde{\Omega} = \Omega \backslash \{(i, j)\}$ and $\widetilde{c} = c \backslash \widetilde{c}_{ij}$. If $c_{ij}^k(c, \Omega) = \widetilde{c}_{ij}$, then $c^k(c, \Omega) = c^k(\widetilde{c}, \widetilde{\Omega})$.*

A Markov strategy is a strategy which is optimal, given Markov beliefs.

The first point requires that beliefs are the same for payoff-equivalent histories, along the lines suggested by Maskin and Tirole (2000). In this model any two histories giving rise to the same (c, Ω) are payoff-equivalent. Rejected bids and the dates of previous agreements do not affect the continuation payoffs in ongoing negotiations. Markov strategies are thus functions of (c, Ω) only.[6]

[6] Two histories giving rise to (c, Ω) and (c', Ω) may in fact be payoff equivalent. As such equivalence is irrelevant to our purposes, we ignore it for the sake of notational simplicity.

Point two strengthens the Markov assumption. Given history $\widetilde{h}_t$ all players know that (i, j) have agreed on c_{ij} at some time $s < t$, and hence will implement c_{ij} in period t and all future periods. Given history h_t all players believe (with probability one) that (i, j) will agree on c_{ij} at t, and hence will implement c_{ij} in period t and all future periods. From the point of view of all players, except the representatives in negotiation (i, j), the two histories are payoff equivalent.[7] Consequently, we require that these players do not condition their behavior at t on whether they know or believe that c_{ij} will be implemented from t onwards. The only reason why one player might want to make a distinction between the two histories is if some other player makes that distinction. That is, although it is payoff-irrelevant, the distinction may be used as a signal (a sun spot) to coordinate the remaining players' behavior on different equilibria in the different subgames. Although the second point constitutes an addition to the traditional Markov restriction, the spirit is the same. The restriction reduces the information on which firms can condition their behavior (since the behavior must be the same for certain histories). Moreover, the restriction does not violate the rationality of the firms. That is, it is optimal for a firm to use a Markov strategy, if all the other firms are using them.

Proposition 2 *All equilibria prescribe immediate agreement if the players use Markov strategies, or if production is strictly convex.*

To prove that Markov is sufficient, Lemma 4 is concerned with subgames $\Gamma(h_T)$ in which equilibrium α induces delay, i.e. where play according to α implies that all negotiations do not end at t. It is shown that some subgame $\Gamma(h_S)$ of $\Gamma(h_T)$ must exist in which α induces delay with strictly fewer ongoing negotiations. In any subgame $\Gamma(h_T)$, firms have an incentive to conclude their negotiations immediately. Suppose that u proposes an agreement in (u, d) one period earlier than prescribed by α. If there is delay conditional on agreement, the lemma is proved. If not, there must be delay in a subset of negotiations in the subgame after d has rejected the deviating bid. If all ongoing negotiations are concluded at the same time, they will conform to bilateral efficiency and equal splitting, in the case of either acceptance or rejection. Both u and d will then gain by the deviation, contradicting the assumption that α was an equilibrium. The proposition is shown by repeatedly applying Lemma 4, implying each time that there exist subgames with delay

[7] All players in the ongoing negotiations have the same strategy sets, and their continuation payoffs are identical in the two subgames.

with fewer ongoing negotiations. Since the number of initial negotiations is finite, we obtain a contradiction.[8]

Without Markov and strict convexity, there may be delay in equilibrium. To see this, recall that immediate agreement on any bilaterally efficient quantity vector is still an equilibrium, without the Markov restriction. Consider for example the case of homogenous goods, which violates strict convexity. Different equilibria (bilaterally efficient quantity vectors) do then exist, corresponding to different distributions of sales (who sells to whom). Moreover, firms receive different payoffs in the different equilibria. It can be shown that if downstream firms have steeper marginal cost curves than upstream firms, then the downstream (upstream) firms' profits are higher, the more proportional (exclusive) the sales are (see Propositions 5 below). Given this multiplicity we can construct a non-Markov equilibrium entailing delay. According to this equilibrium, all firms are to conclude negotiations in round 2, making "unreasonable" bids in period 1. Deviations are punished by selecting an appropriate equilibrium in period 2. If an upstream firm makes a "reasonable" bid in period 1, the equilibrium play contingent on rejection is to agree upon proportional sales, thereby punishing the deviating upstream firm and awarding the rejecting downstream firm.

The equilibrium contract structure can also be derived using a more cooperative approach. Define the *bilateral Nash product* as

$$\left[\pi^u\left(c \backslash c_{ud}\right)-\pi^u\left(c \backslash (0,0)\right)\right]\left[\pi_d\left(c \backslash c_{ud}\right)-\pi_d\left(c \backslash (0,0)\right)\right]. \tag{6}$$

The disagreement point in the negotiation between u and d is identified with the so-called impasse point, that is to say the payoff obtained by u and d if both firms reject any offer in the (u,d) bargaining, all other contracts being taken as given. It is easy to see that $c_{ud}=(q_{ud},p_{ud})$ maximizes the bilateral Nash product if, and only if, q_{ud} is bilaterally efficient, and p_{ud} yields an equal split of the bilateral surplus. The propositions thus provide a non-cooperative foundation for the "Nash equilibrium in Nash bargaining" approach discussed in the Introduction.

[8]The proof that strict convexity is sufficient is similar and has therefore been omitted. The crucial point is that the bilaterally efficient quantity vector is unique.

3 Quantities and Efficiency

Disregarding the distribution of profits between firms, we define social welfare W as the discounted sum of profits of all upstream and downstream firms:

$$W\left(\{c(t)\}_{t=0}^{\infty}\right) = \sum_{i=1}^{U}\sum_{t=0}^{\infty}\delta^t\pi^i\left(c(t)\right) + \sum_{j=1}^{D}\sum_{t=0}^{\infty}\delta^t\pi_j\left(c(t)\right). \tag{7}$$

Note that prices affect the distribution of wealth but not social welfare and are consequently not included as an argument in the welfare function. Consumers can be disregarded since the price of final goods is constant.

Proposition 3 *Production is socially efficient, subject to the zero-quantity constraints imposed by any missing links.*

The equilibrium is thus socially efficient regardless of concentration.[9]

To understand the result, note that in all negotiations the chosen quantity equalizes the seller's marginal cost to the buyer's marginal valuation. Why this bilateral efficiency also implies social efficiency is most easily understood in the case of homogeneous goods. By equalizing its marginal valuation to the marginal costs in all negotiations, a buyer guarantees that all its partners have the same marginal cost. Similarly, every seller guarantees that all its partners have the same marginal valuations of the good. If the link structure is sufficiently complete, all sellers produce at the same marginal cost and all buyers have the same marginal valuation. The absence of coordination failure should also be noted. If the welfare function were to have several local maxima, the bilateral oligopoly might get stuck at a maximum that is not global. When all cost functions are convex, however, the welfare function is concave, and there is a unique maximum. Another reason for efficiency is the timing. If some contracts were negotiated before others (e.g. machinery before oil), there would be strategic commitment with resulting inefficiencies.[10]

[9] The efficiency result may be interpreted as including efficient transportations. Assume that goods are delivered free on board, and that the downstream firms' cost functions include all transportation costs. Then, not only the allocation of production between firms is efficient. The equilibrium trading pattern, i.e. who sells to whom, minimizes the transportation costs.

[10] Consider the case of one upstream firm u, and two downstream firms d and b. Assume that $C^u = (q_d + q_b)^2/2$, $C_d = kq_d$ and $C_b = 0$. If agreements are made simultaneously, u will only sell to the efficient firm b. If u can start negotiations with d one period earlier, they will trade a positive quantity ($q_d = r/2 - k$), thereby increasing the price that b has to pay.

Proposition 3 shows that a complete link structure is sufficient for a bilateral oligopoly to be efficient. However, real-world link structures are incomplete for various reasons,[11] and it can thus be interesting to see whether equilibria also may be efficient with incomplete link structures. To do this, let us define the *number of active links* associated with an (equilibrium) quantity vector as the number of buyer-seller pairs trading strictly positive quantities.

Proposition 4 *Consider a homogeneous goods market. There exists a socially efficient equilibrium with $U + D - 1$ active links. Generically, this is the minimum number of active links of an efficient equilibrium.*

Intuitively, for a homogeneous goods market to be efficient, all upstream firms must have the same marginal cost ($U-1$ conditions) and all downstream firms must have the same marginal cost ($D-1$ conditions). A final condition arises from the requirement that supply must equal demand. To satisfy the $U+D-1$ conditions, the same number of instruments (i.e. quantities q_{ud}) will typically be needed. Proving the Proposition is more complicated, however. To take the non-negativity constraints ($q_{ud} \geq 0$) into account, we construct an algorithm generating an efficient link structure with only $U + D - 1$ links for any market. Let us start with a link between two firms. Generically, either the buyer or the seller will not be able to satisfy its demand or its supply. In the first (second) case, add a link between the buyer (seller) and an additional seller (buyer). Generically, either the last seller or the last buyer will not be satisfied. The procedure is then repeated, stopping after $U + D - 1$ steps.

The $U + D - 1$ links cannot be chosen arbitrarily. The link structure must be connected. Assume that there is no cross-border trade between two regional markets, A and B. Production is socially efficient in each regional market, due for example to regionally complete link structures. Let p^A and p^B be the two hypothetical Walrasian prices. Production is globally efficient if, and only if, $p^A = p^B$, a condition which is not satisfied generically. The intuition for the condition is simple. If country A has relatively little upstream capacity (thus $p^A > p^B$), the upstream (downstream) firms in country A produce at a higher (lower) marginal cost than the firms in country B. Production is thus not allocated optimally. The upstream and downstream production capacities are examples of (sector) specific (production) factors. Differences in the relative endowments of the specific factors

[11] The link structure may be incomplete for political reasons, e.g. trade barriers. Incompleteness may also arise (endogenously) because it is costly to establish links. Finally, there may be strategic motives for firms not to negotiate with each other, e.g. foreclosure.

thus give rise to gains from trade, and trade liberalization enhances global welfare. (Another reason why p^A and p^B may differ is that the two countries have different production technologies.) This result is related to the theory of comparative advantage, and in particular to the Heckscher-Ohlin Theorem. However, we derive our results in a rather different environment, where all firms have market power, and where the law of one price does not hold.

4 Prices and Distribution

To study the distribution of surplus, we solve for the equilibrium price:

$$p_{ud} = r/2 - \frac{C_d(q) - C_d(q \backslash q_{ud} = 0)}{q_{ud}}/2 + \frac{C^u(q) - C^u(q \backslash q_{ud} = 0)}{q_{ud}}/2, \quad (8)$$

where the two parentheses contain the average incremental costs of producing q_{ud}, and where all quantities are at their equilibrium levels. As a benchmark we use the Walrasian price which is defined as the market clearing price, assuming that both upstream and downstream firms are price-takers. The Walrasian price is denoted by p^*_{ud} and characterized by[12]

$$p^*_{ud}(q) = r/2 - \frac{\partial C_d(q)}{\partial q_{ud}}/2 + \frac{\partial C^u(q)}{\partial q_{ud}}/2, \quad (9)$$

where q is the socially efficient quantity vector. In the case of homogenous goods, the unique Walrasian price is denoted p^*.

Consider first the case in which goods are homogenous and all upstream (downstream) firms have access to the same technology. The short-run cost function for firm u is given by $\overline{C}(Q^u, k^u)$, where k^u is u's capital level, and the equivalent applies for downstream firms. The incremental costs are given by $\underline{C}(Q_d, k_d) - \underline{C}(Q_d - q_{ud}, k_d)$ and $\overline{C}(Q^u, k^u) - \overline{C}(Q^u - q_{ud}, k^u)$ respectively, suggesting that the prices are determined by the distribution of capital and the distribution of sales. To show this formally, we conduct three experiments. As it turns out, the result hinges on the relative shape of downstream and upstream technologies. We say that u has a *steeper marginal cost curve* than d, if

$$\overline{C}_{qq}(Q^u - x, K^u) > \underline{C}_{qq}(Q_d - x, K_d) \quad \text{for all } x \leq \min\{Q^u, Q_d\}. \quad (10)$$

[12] The first order condition for a price taking upstream firm is $p_{ud} - \partial C^u(q)/\partial q_{ud} = 0$ and the first order condition for a price taking downstream firm is $r - p_{ud} - \partial C_d(q)/\partial q_{ud} = 0$. To derive equation (9), subtract one from the other.

Proposition 5 *In an efficient homogeneous goods market, trade may occur at different prices. In particular:*

1. *A seller with more capital charges a higher (lower) price to the same buyer for the same quantity than a seller with less capital, if* $\overline{C}_{qk}(Q^u, k) / \overline{C}_{qq}(Q^u, k)$ *is falling (increasing) in* Q^u.

2. *A seller distributing* q *units evenly between two buyers of the same size earns a higher revenue than when selling different amounts, if it has a steeper marginal cost curve than the downstream firms.*

3. *Assume that there are constant returns to scale both upstream and downstream, with average variable costs* $\overline{a}$ *and* $\underline{a}$ *respectively. Individual prices, as well as the average price, can be anywhere in the interval* $\left(\frac{p^*+\underline{a}}{2}, \frac{r+p^*-\overline{a}}{2}\right)$.

The first two points show that the law of one price does not hold, and that prices are not completely determined by aggregate supply and demand.

The first point shows that the *distribution of capital* between firms does matter. To demonstrate this we compare the prices charged by two upstream firms which have different amounts of capital and are selling the same quantity to a particular downstream firm. Under reasonable conditions (e.g. if there are constant returns to scale, or if marginal costs are linear), big buyers (sellers) get better deals than small buyers (sellers) at any given quantity. Since the comparison is made at given quantities, this prediction is different from quantity discounts. This result is surprising at first sight. It shows that firms with low marginal costs (i.e. high capital) may charge higher prices than high-cost firms for selling the same amount of goods to the same customer. The explanation is that the incremental cost may be higher for the low-cost firm in equilibrium. Although the low-cost firm has a lower cost at each output level, it produces more in equilibrium and therefore has a higher cost at the (discrete) margin.

The second point shows that the *distribution of sales* (who sells to whom) does matter. In the case of homogeneous goods, there are multiple equilibria with different sales distributions. Even though all equilibria are efficient, prices vary with the equilibrium chosen. To make this point, we compare the prices charged by two upstream firms with different trading patterns, but the same amount of capital. Consider the case in which upstream firms have increasing marginal costs but downstream firms have constant marginal costs. An upstream firm distributing its sales equally among many downstream firms will have a high incremental cost in all negotiations, and will therefore charge a high price.

In the third experiment we manipulate both the distribution of capital and the distribution of sales to show that not only individual prices but also the average price, may differ from the Walrasian price.

Social efficiency in combination with relation-specific prices can be understood in the light of perfect price discrimination. In contrast to the standard model, however, we allow for competing buyers and sellers and show that efficiency is independent of the number of firms. Moreover, not only the sellers but also the buyers have bargaining power, and both capture a share of the social surplus. The result, namely that price discrimination leads to efficiency, suggests that policies condemning price discrimination should be implemented with caution in markets with bilateral market power.

There is a close connection between the negotiated prices, p_{ud}, and the Walrasian prices, p^*_{ud}. A comparison of equations (8) and (9) shows that the only difference is that p_{ud} is determined by the average incremental costs of producing q_{ud}, while p^*_{ud} is determined by the marginal costs of production. For small q_{ud} the average incremental costs and the marginal costs are approximately the same. Given an efficient equilibrium in which q_{ud} is close to zero, p_{ud} is close to the Walrasian price.

One way to model this more formally is to consider a set of cost functions $\{C^u\}_{u=1}^{\infty}$ and $\{C_d\}_{d=1}^{\infty}$. Let a sequence of economies for $s = 1, 2, ...$ consist of all firms $\{C^u\}_{u=1}^{U_s}$ and $\{C_d\}_{d=1}^{D_s}$, where $U_{s+1} > U_s$ and $D_{s+1} > D_s$. Let $\{p^*_{ud}(s)\}$ be the associated sequence of Walrasian prices, and $\{(q(s), p(s))\}$ an associated sequence of equilibrium contract structures.

Proposition 6 *Consider a sequence of economies, and an associated sequence of socially efficient contract structures. If $q_{ud}(s) \to 0$ for all u and d,* $\max_{u,d} |p_{ud}(s) - p^*_{ud}(s)| \to 0$.

Note that $q_{ud}(s) \to 0$ is crucial for Walrasian prices. In particular, consider an economy with $U_s = D_s = s$, where all upstream firms are identical and all downstream firms are identical, and assume that upstream firm $u \leq s$ sells exclusively to downstream firm $d = u$ in each economy (which is a socially efficient equilibrium). For every s in the sequence the economy then consists of s bilateral monopolies all trading at the bilateral monopoly price, which typically is different from the Walrasian price. Although each firm buys or sells only a small share of the total market quantity, namely $1/s$, the prices differ from the Walrasian price.

Propositions 3 (efficient quantities) and 6 (Walrasian prices) provide a foundation for Walrasian equilibrium on two grounds. First, in our model the agents set prices themselves without the help of an auctioneer. Second, Proposition 6 is a limit theorem and not a theorem in the limit. That is, the Walrasian price is an approximation of a large but finite economy. Efficiency is attained even when there are only a few firms. This result is thus much stronger than the limit results reported in the previous literature (Gale, 2000).[13]

The "small-size requirement" for Walrasian prices is more demanding than has previously been understood, however. It is often said that if every firm in the industry makes a trifling fraction of the industry's sales or purchases, a single price will rule in a market (see e.g. Stigler, 1968). In our model it is not sufficient that each firm's sales or purchases are small in relation to the aggregate market quantity. Every firm must only account for a small share of every partner's trading volume.

5 Imperfect Competition in the Final Goods Market

Assume that downstream firms compete à la Cournot in the downstream market. The first order condition for bilateral efficiency is then given by

$$\frac{\partial\left[\pi_d+\pi^u\right]}{\partial q_{ud}}=r_{ud}\left(q\right)+\sum_{i=1}^{U}\frac{\partial r_{id}\left(q\right)}{\partial q_{ud}}q_{id}-\frac{\partial C_d\left(q\right)}{\partial q_{ud}}-\frac{\partial C^u\left(q\right)}{\partial q_{ud}}=0. \tag{11}$$

The second term is present because the firms affect prices in the final goods market. We do not establish conditions for the existence of a bilaterally efficient quantity vector, but simply presume its existence. It is possible, however, to verify the existence of a unique bilaterally efficient quantity vector in simple examples. It is also possible to show that a sequential equilibrium exists implying immediate agreement on q if, and only if, q is bilaterally efficient. The proof is similar to the proof of Proposition 1, and has therefore been omitted here.

Our main result shows that a bilateral oligopoly is efficient also when there is market power in the final goods market. As a benchmark we use the Walrasian equilibrium, which presumes that all upstream and downstream firms take the prices of intermediate goods (but not final goods) as given.[14]

[13] Thus, our results contribute to the strategic foundations of general equilibrium (or the theory of decentralized trade). The previous literature, reviewed by Osborne and Rubinstein (1990) and Gale (2000), consider random matching of buyers and sellers (endowed with one unit of the indivisable good), each match being an ultimatum game. See Westermark (2000) for an application to labor markets.

[14] We use the Walrasian equilibirum as a benchmark rather than a social planner, since

Proposition 7 *If a bilaterally efficient equilibrium exists, it is as efficient as a Walrasian equilibrium.*

That is, market power in the intermediate goods market does not necessarily give rise to inefficiencies over and above those resulting from market power in the final goods market. The efficiency of the intermediate goods market is due in part to the absence of so-called double marginalization, which in turn may be explained by the fact that firms contract for both prices and quantities.[15]

The price yielding an equal split of the bilateral surplus is given by

$$p_{ud} = \frac{r_{ud}(q)}{2} + \frac{1}{2}\frac{C^u(q\backslash q_{ud}) - C^u(q\backslash 0)}{q_{ud}} - \frac{1}{2}\frac{C_d(q\backslash q_{ud}) - C_d(q\backslash 0)}{q_{ud}}. \\ + \frac{1}{2}\sum_{i\neq u}\frac{r_{id}(q\backslash q_{ud}) - r_{id}(q\backslash 0)}{q_{ud}}q_{id}. \qquad (12)$$

The last term indicates that the relative bargaining power of u and d is determined by the substitutability between the different goods sold by d, i.e. the differentiation between upstream firms.[16] In particular, the more substitutable q_{ud} and q_{vb}, the lower is p_{ud} (given q and r). When consumers see the products of the upstream firms as close substitutes, a downstream firm will not suffer much by not coming to an agreement with a particular upstream firm. Not selling u's product increases the demand and the price for the other products sold by d. In this case, downstream firms have a strong bargaining position. The last term of equation (12) has implications for the preferences of firms among different trading patterns. Assume that all firms have constant marginal costs, then downstream firms will prefer to buy intermediate goods from several upstream firms, while upstream firms prefer exclusive dealing.[17]

---

it is not possible to let a social planner choose quantities in the intermediate goods market, without at the same time correcting for the Cournot inefficiency.

[15] One may ask why an upstream monopoly cannot induce a monopoly in the downstream sector by selling the monopoly quantity to one downstream firm, for example d, and nothing to any other firm, for example b. The reason is that u has an incentive to cheat on d by agreeing upon the bilaterally efficient quantity q_{ub} with b. However, if we would make the link structure endogenous, it may be profitable for u to form one link only, thereby committing itself not to cheat.

[16] Note that for q_{ud} small, $[r_{id}(q\backslash q_{ud}) - r_{id}(q\backslash 0)]/q_{ud}$ is approximately equal to the cross derivative of demand.

[17] In the present model it is only the degree of substitutability between the upstream firms' products that is important for bargaining power in the intermediate goods market.

6 Concluding Remarks

The aim of this paper has been to construct a model that captures four key institutional characteristics of many intermediate goods markets, and to address some basic microeconomic issues concerning the efficiency and distribution of surplus in such markets. In doing so, we have derived some more precise predictions amenable to empirical testing. The finding that bilateral oligopolies are socially efficient can be tested, even in the absence of marginal cost data. Aggregate production should vary with the total amount of capital, but not as a result of mergers and spin-offs. The findings that large firms get better deals than small firms, and that a seller with an even distribution of sales obtains higher prices than a seller with a more exclusive sales (under certain conditions) are testable, but do require data on contract prices. Again, a more indirect test is feasible. It is easy to show that horizontal mergers increase the profits of the merging firms and reduce the profits of their trading partners. This may in fact explain an earlier empirical result that seller profitability is negatively related to buyer concentration (see Lustgarten, 1975; Schumacher, 1991; and for a review see Scherer and Ross, 1990). Previously, this regularity has been understood in light of the theory of collusion described in Stigler (1964), (sellers are more likely to collude if there are many customers). Our results show that the empirical regularity is consistent with non-collusion between firms. Furthermore, under additional assumptions (linear marginal costs), the average price is determined by the difference in concentration (Hirfindahl index) between upstream and downstream firms. This prediction may be used as a basis for new and more structurally oriented empirical studies of bilateral oligopolies.

The most important step for further theoretical research is to endogenize the link structure. The link structure has been shown to be a key determinant of social efficiency in bilateral oligopolies, as well as an important determinant of the distribution of surplus. Endogenizing the links is likely to be non-trivial, due to the strategic externalities that are likely to exist between different links. The studying of links is also policy-relevant. Trade policies such as tariffs and quotas affect the link structure. Antitrust action against vertical foreclosure may contribute to a more complete link structure. Investments in infrastructure reduce the cost for distant firms forming links with each other. The importance of such policies can only be studied in a model with endogenous links. Other important applications involve the

---

In reality, however, high substitutability between different downstream firms (e.g. retailers that are located close to each other) may give upstream firms high bargaining power (cf. Porter, 1976).

effects of the Internet, which has presumably reduced the cost of links.

References

Binmore, K.: Perfect equilibrium in bargaining models, *in* K. Binmore P. Dasgupta (eds), *The Economics of Bargaining*, Basil Blackwell, chapter 5, 1987, 77–105.

Davidson, C.: Multiunit Bargaining in Oligopolistic Industries, *Journal of Labor Economics* 6, 1988, 397-422.

Dobson, Paul; Waterson, Michael: Countervailing power and consumer prices, *Economic Journal* 107(441), 1997, 418–30.

Dobson, Paul; Waterson, Michael: Retail Power: Recent Developments and Policy Implications, *Economic Policy* 28, April 1999, 135-164.

Economist: B2B Exchanges: The Container Case, *The Economist*, October 21st 2000, 93-94.

Gale, Douglas: *Strategic Foundations of General Equilibrium: Dynamic Matching and Bargaining Games*, Cambridge University Press, Cambridge, 2000.

Haller, Hans; Holden, Steinar: A Letter to the Editor on Wage Bargaining, *Journal of Economic Theory* 52, 1990, 232-236.

Hart, Oliver; Tirole, Jean: Vertical integration and market foreclosure, *Brookings Papers on Economic Activity* Special Issue, 1990, 205–76.

Horn, Henrik; Wolinsky, Asher: Bilateral monopolies and incentives for merger, *Rand Journal of Economics* 19(3), 1988a, 408–19.

Horn, Henrik; Wolinsky, Asher: Worker substitutability and patterns of unionisation, *Economic Journal* 98(391), 1988b, 484–97.

Kranton, Rachel E.; Minehart, Deborah F.: A Theory of Buyer-Seller Networks, forthcoming in *American Economic Review.*

Lustgarten, Steven H.: The Inpact of Buyer Concentration in Manufacturing Industries, *Review of Economics and Statistics* 57, May 1975, 125-132.

Maskin, Eric; Tirole, Jean: Markov Perfect Equilibrium, I: Observable Actions, forthcoming in *Journal of Economic Theory.*

McAfee, Preston; Schwartz, Marius: Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity, *American Economic Review* 84(1), March 1994, 210-230.

McAfee, Preston; Hendriks, Ken: A Theory of Bilateral Oligopoly, with Applications to Vertical Mergers, mimeo., June 2000.

McDonald, Ian M.; Solow, Robert, M.: Wage Bargaining and Employment, *American Economic Review* 71(5), December 1981, 896-908.

Myerson, Roger B.; Satterthwaite, Mark A.: Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory*; 29(2), April 1983, 265-81.

Osborne, M. J. Rubinstein, A: *Bargaining and Markets*, Economic Theory, Econometrics, and Mathematical Economics series, Harcourt Brace Jovanovich, Academic Press, 1990.

Porter, Michael E.: *Interbrand Choice, Strategy, and Bilateral Market Power*, Harward University Press, Cambridge, Massachusetts, 1976.

Rubinstein, Ariel; Wolinsky, Asher: Decentralized trading, strategic behaviour and the walrasian outcome, *Review of Economic Studies* 57(1), 1990, 63–78.

Scherer, F.M.; Ross, David: *Industrial Market Structure and Economic Performance*, Third Edition, Boston: Houghton Mifflin Company, 1990.

Schumacher, Ute: Buyer Structure and Seller Performance in U.S. Manufacturing Industries, The Review of Economics and Statistics, 1991, 277-84.

Stigler, George: A Theory of Oligopoly, *Journal of Political Economy* 72, February 1964, 44-61.

Stigler, George: *The Organization of Industry*, Chicago: The University of Chicago Press, 1968.

Stole, Lars; Sweibel, J.: Intra-Firm Bargaining Under Non-Binding Contracts, *Review of Economic Studies* 63, 1996: 375-410.

Westermark, Andreas: Bargaining, Binding Contracts and Competitive Wages, Mimeo. Uppsala University, 2000.

Proofs

Preliminaries

At date t one firm makes a bid $b_{ud}(t)$ which is a pair $(q_{ud}(t), p_{ud}(t)) \in \mathbb{R}^2_+$ where $q_{ud}(t)$ is a quantity and $p_{ud}(t)$ is the price. The other firm is allowed either to accept or reject the bid, $\rho_{ud}(t) \in \{y, n\}$. The *action* at time t, is the UD-tuple $(a_{11}(t), ..., a_{ud}(t))$, where $a_{ud}(t) = (b_{ud}(t), \rho_{ud}(t))$. A *history* at time t, denoted h_t, is a t-tuple of actions $(a_0, ..., a_{t-1})$, with h_0 denoting the "empty" history at $t = 0$. Let H_t be the set of possible h_t. Let $c_{ud}(h_t) = b_{ud}(T)$ if $r_{ud}(T) = y$ for some $T < t$. Let $\Gamma(h_T)$ denote the *subgame* induced by the history h_T at time T. The representative of firm u in negotiation (u, d) has the following strategy:

$$\begin{aligned} \overline{b}_{ud}(h_t) &: H_t \to \mathbb{R}^2_+, \text{ and} \\ \overline{\rho}_{ud}(h_t, b) &: H_t \times \mathbb{R}^2_+ \to \{y, n\}, \end{aligned}$$

and similarly for upstream firms' representatives.

Let $h_t(h_T, \alpha)$ with $t \geq 0$ be the history such that (i) for $t \leq T$ it is on the path to h_T, and (ii) for $t > T$ it is induced by α contingent on h_T having been reached.

Let $\Omega(h_T) \subseteq \Omega$ denote the set of ongoing negotiations (u, d) at the beginning of subgame $\Gamma(h_T)$.

Lemma 1

Let $B_{ud}(q)$ denote the set of bilaterally efficient quantities for (u, d), and let $B(q)$ be the cartesian product of all B_{ud}. As the bilateral surplus for (u, d) is a continuous and concave function of q, $B(q)$ by the theorem of the maximum is a convex-valued, upper-hemi continuous correspondence. By the Kakutani fixed point theorem, there exists a fixed point.

Proposition 1

Consider a subset of negotiations $\Omega \subseteq \overline{\Omega}$ and a fixed contract structure c, with the associated vector of quantities q. Let $\widehat{q}(c, \Omega)$ be a selection in N, which is non-empty by Lemma 1, i.e. $\widehat{q}(c, \Omega) \in N(c, \Omega)$. Define the price $\hat{p}_{ud}(\widehat{q}, t)$ by

$$\hat{p}_{ud}(\widehat{q}, t) = \begin{cases} \widetilde{p}_{ud}(\widehat{q}) + \varepsilon_{ud}(\widehat{q}, \delta) & \text{if } u \text{ is bidding at } t \\ \widetilde{p}_{ud}(\widehat{q}) - \varepsilon_{ud}(\widehat{q}, \delta) & \text{if } d \text{ is bidding at } t \end{cases} \tag{13}$$

where

$$\widetilde{p}_{ud}(\widehat{q}) = \tfrac{1}{2}\frac{C^u(\widehat{q}) - C^u(\widehat{q}\backslash q_{ud}=0)}{\widehat{q}_{ud}} - \tfrac{1}{2}\left(\frac{C_d(\widehat{q}) - C_d(\widehat{q}\backslash q_{ud}=0)}{\widehat{q}_{ud}} - r\right), \tag{14}$$

and

$$\varepsilon_{ud}(\widehat{q},\delta) = \frac{1}{2}\frac{\delta-1}{\delta+1}\left[\frac{C_d(\widehat{q}) - C_d(\widehat{q}\backslash q_{ud}=0)}{\widehat{q}_{ud}} - r + \frac{C^u(\widehat{q}) - C^u(\widehat{q}\backslash q_{ud}=0)}{\widehat{q}_{ud}}\right]. \tag{15}$$

Let $\Omega(h_T) \subseteq \Omega$ denote the set of ongoing negotiations (u,d) at the beginning of subgame $\Gamma(h_T)$. The next lemma considers the case when only one negotiation remains.

Lemma 2 *Assume $\Omega(h_T) = \{(u,d)\}$ and $c(h_T) = c$ with $c_{ud} = (0,0)$. If $\widehat{q}_{ud}(c,\{(u,d)\}) \neq 0$ there exists a unique subgame-perfect equilibrium in the subgame $\Gamma(h_T)$, implying immediate agreement. According to this equilibrium, $\overline{b}_{ud}(h_t) = (\overline{p}_{ud}(\widehat{q}(c,\{(u,d)\})), \widehat{q}(c,\{(u,d)\}))$ and $\underline{b}_{ud}(h_t) = \left(\underline{p}_{ud}(\widehat{q}(c,\{(u,d)\})), \widehat{q}(c,\{(u,d$ Moreover, firms accept (reject) any offer giving them a higher (lower) profit than implied by the equilibrium bids.*

Proof: The existence of a unique subgame-perfect equilibrium follows from Binmore (1987). The equilibrium prescribes immediate agreement on a Pareto-efficient outcome in every subgame. In the present context, this is equivalent to quantities being bilaterally efficient. According to standard Rubinstein-Ståhl reasoning, prices are determined by

$$\frac{1}{1-\delta}\pi^u(c\backslash\underline{b}_{ud}) = \pi^u(c\backslash(0,0)) + \frac{\delta}{1-\delta}\pi^u\left(c\backslash\overline{b}_{ud}\right) \tag{16}$$

$$\frac{1}{1-\delta}\pi_d\left(c\backslash\overline{b}_{ud}\right) = \pi_d(c\backslash(0,0)) + \frac{\delta}{1-\delta}\pi_d(c\backslash\underline{b}_{ud}). \tag{17}$$

Using the fact that both firms propose the bilaterally efficient quantity, that is $\overline{q}_{ud} = \underline{q}_{ud} = \widehat{q}_{ud}$, the system may be solved to yield equation (13). Prices $\left(\underline{p}_{ud}, \overline{p}_{ud}\right)$ do not depend on time or history.

Lemma 3 *There exists a sequential equilibrium with all $(u,d) \in \Omega(h_T)$ agreeing immediately on $c_{ud} = (q_{ud}, p_{ud})$ at T if, and only if, $q_{ud} = \hat{q}_{ud}$ with $\widehat{q} \in N(c(h_T), \Omega(h_T))$. In addition, $p_{ud} = \hat{p}_{ud}(\widehat{q}, T)$ for all $(u,d) \in \Omega(h_T)$ where $\hat{q}_{ud} \neq 0$.*

Proof: Consider $(u,d) \in \Omega(h_T)$. Deviations from prescribed equilibrium at T by u or d will not affect h_T by the informational assumptions. Hence, for all other negotiations $(i,j) \neq (u,d)$ there is immediate agreement on contract $c_{ij}(h_T)$. Lemma 2 thus applies for negotiation (u,d).

Assume immediate agreement on $q \notin N(c,\Omega)$. There then exists at least one pair $(u,d) \in \Omega$ that will improve bilateral efficiency by a unilateral deviation according to Lemma 2. Likewise, any $p_{ud} \neq \widehat{p}_{ud}(q,T)$ is either not maximizing the bidders profits or will be rejected.

Proposition 2

We say that strategy profile α *induces delay* in $\Gamma(h_T)$ if $\Omega(h_{T+1}(h_T,\alpha)) \neq \emptyset$.

Lemma 4 *Consider an equilibrium α and a subgame $\Gamma(h_T)$. If α induces delay at T, then there exists a subgame $\Gamma(h_S)$ of $\Gamma(h_T)$ (with $S > T$) such that α induces delay in $\Gamma(h_S)$ and $\emptyset \neq \Omega(h_S) \subset \Omega(h_T)$.*

Proof: Assume that there exists some date $t > T$ such that some, but not all, negotiations in $\Omega(h_T)$ are concluded. The Lemma then follows immediately, since there is delay in subgame $\Gamma(h_t)$. Two cases remain to be considered.

Case 1: Assume that α prescribes that everybody agrees at $t > T$. A deviation specifying the same $\hat{q}_{ud}$ at $t-1$ will increase payoff for (u,d), as by the Markov assumption conditional on disagreement, the actions of everybody else will be the same. Thus α cannot be an equilibrium.

Case 2: Assume that nobody ever agrees in subgame $\Gamma(h_T)$. Consider negotiation $(u,d) \in \Omega(h_{S-1}(h_T,\alpha))$ in some period $S-1 \geq T$. Let h_S be the history in which u suggests $q_{ud} = 0$ in period $S-1$, and d accepts (all others play according to α). Let h_S^- be the history in which d rejects. In both subgames, three outcomes are logically possible: immediate agreement, no agreement, or agreement at different times. In the case of the last possibility, the Lemma is immediately proved.

Assume that α prescribes no agreement in $\Gamma\left(h_S^-\right)$. Then it cannot be the case that α prescribes immediate agreement or no agreement in $\Gamma(h_S)$. Thus α induces delay in $\Gamma(h_S)$ with fewer active negotiations. Immediate agreement or no agreement in $\Gamma(h_S)$ would imply that d's acceptance of the bid $q_{ud} = 0$ is equilibrium play. If d accepts and all negotiations end at S, Lemma 3 implies that the agreement is on $\widehat{q} \in N(c(h_S), \Omega(h_S))$ and the corresponding prices $\widehat{p}$. An agreement on bilaterally efficient quantity vector $\widehat{q}$ cannot reduce the profits of u and d relative to the profits under no agreement at all, according to Assumption 1. Finally, d accepting the bid implies that it is better for u to suggest $q_{ud} = 0$ in period $S-1$ than equilibrium play (no agreements).

Now assume that α prescribes immediate agreement in $\Gamma\left(h_S^-\right)$. Similar arguments then show that α cannot prescribe immediate agreement or no agreement in $\Gamma\left(h_S\right)$.

Proof of the Proposition Assume on the contrary that α is an equilibrium strategy profile that induces delay in $\Gamma\left(h_T\right)$. According to Lemma 4, there exists a subgame with delay with fewer negotiations. Repeated application of the lemmas give an infinite sequence of subgames $\Gamma\left(h_{T_1}\right)$, $\Gamma\left(h_{T_2}\right)$,... where $\Omega\left(h_{T_{k+1}}\right) \subset \Omega\left(h_{T_k}\right)$ and $\Omega\left(h_{T_k}\right) \neq \emptyset$ for all T_k. As $\Omega\left(h_T\right)$ is a finite set, we obtain a contradiction.

Proposition 3

The welfare function is maximized when the sum (across firms) of per-period profits is maximized. Moreover, $\partial W/\partial q_{ud} = \partial \pi_d/\partial q_{ud} + \partial \pi^u/\partial q_{ud}$ as $\partial \pi^u/\partial q_{ij} = 0$ for $i \neq u$ and $\partial \pi_d/\partial q_{ij} = 0$ for $j \neq d$. If production is convex (strictly convex), it can be shown that the welfare function W is concave (strictly concave).

The choice set is convex since it is defined by non-negativity ($q_{ud} \geq 0$) and equality ($q_{ud} = 0$ for $(u, d) \in \overline{\Omega} \backslash \Omega$) constraints. The choice set is bounded in the case of strictly convex production. By the maximum theorem we see that as the welfare function W is concave, there exists a non-empty compact convex set of quantities in $\mathbb{R}_+^{U \times D}$ maximizing W.

According to Proposition 1 all equilibria are bilaterally efficient. The set of first order (Kuhn-Tucker) conditions for bilateral efficiency is the same as the set of first order conditions for social welfare.

Proposition 4

Genericity is taken to mean that if $\alpha \neq U$ or $\beta \neq D$, then $\sum_{u=1}^{\alpha} Q^u \neq \sum_{d=1}^{\beta} Q_d$. Maximizing social welfare, we get: $r - C_d'\left(Q_d\right) - C_u'\left(Q^u\right) = 0$. Similarly $r - C_d'\left(Q_d\right) - C_v'\left(Q^v\right) = 0$. This implies

$$\begin{aligned} C_u'\left(Q^u\right) &= p^* \\ C_d'\left(Q_d\right) &= r - p^* \\ \sum_u Q^u &= \sum_d Q_d. \end{aligned}$$

As both C_d and C_u are assumed to be monotonic, we have a unique set of Q^u and Q_d satisfying the equations.

As we have $U + D - 1$ independent equations, we cannot generically add more than $UD - U - D + 1$ independent equations such as $q_{ij} = 0$ without getting an over-specified system of equations. Thus the number of active links cannot generically be less than $U + D - 1$. The set of quantity vectors maximizing social welfare are thus given by:

$$N(c, \Omega) = \left\{ q : \sum_d q_{ud} = Q^u \text{ and } \sum_u q_{ud} = Q_d \right\},$$

where $Q^u = S^u(p^*)$ and $Q_d = D_d(r - p^*)$.

To prove the proposition, we construct an efficient quantity vector $\widehat{q}$ with $U + D - 1$ active links using the following algorithm.

1. Initial step: Set all $\widehat{q}_{ud} = 0$. Set all $\widehat{Q}^u = 0$ and $\widehat{Q}_d = 0$. Set $\alpha = 1$ and $\beta = 1$.

2. Determine $\widehat{q}_{\alpha\beta}$, by requiring firm α to deliver the maximum possible quantity to firm β. Let $\widehat{q}_{\alpha\beta} = \min\left\{Q^\alpha - \widehat{Q}^\alpha, Q_\beta - \widehat{Q}_\beta\right\}$.[18]

3. Increase α or β by one. (a) Set $\widehat{Q}^\alpha = \widehat{Q}^\alpha + \widehat{q}_{\alpha\beta}$ and $\widehat{Q}_\beta = \widehat{Q}_\beta + \widehat{q}_{\alpha\beta}$. (b) If $\widehat{Q}^\alpha = Q^\alpha$, set $\alpha = \alpha + 1$; (c) If $\widehat{Q}_\beta = Q_\beta$, set $\beta = \beta + 1$.

4. If $\beta > D$ or $\alpha > U$ stop. Otherwise, go to step 2.

Note that the algorithm ends after a finite number of iterations. Steps 2a and 3a imply that $\widehat{Q}_\beta = Q_\beta$ or $\widehat{Q}^\alpha = Q^\alpha$. Hence, α or β is increased by one in each iteration (3b-c). The algorithm stops if $\beta > D$ or $\alpha > U$ (step 4).

Next, we show that the generated $\widehat{q}$ satisfies $\sum_u \widehat{q}_{ud} = Q_d$ and $\sum_d \widehat{q}_{ud} = Q^u$ for all u and d. The algorithm stops if and only if $\alpha = U + 1$ or $\beta = D + 1$ (step 4). If both equalities hold, then from 3b-c we see that $\sum_u \widehat{q}_{ud} = Q_d$ and $\sum_d \widehat{q}_{ud} = Q^u$ for all $d \leq D$ and $u \leq U$. Now we will show that when the algorithm stops, $\alpha = U + 1$ and $\beta = D + 1$. Assume on the contrary that the algorithm stops with $\alpha \leq U$ and $\beta = D + 1$. Then $Q_d = \widehat{Q}_d$ for all $d \leq D$. It follows that

$$\sum_{d=1}^{D} Q_d = \sum_{d=1}^{D} \widehat{Q}_d = \sum_{u=1}^{U} \widehat{Q}^u. \tag{18}$$

[18] Note that: $\widehat{q}_{\alpha\beta} > 0$, since $Q_\beta - \widehat{Q}_\beta > 0$ and $Q^\alpha - \widehat{Q}^\alpha > 0$.

The second equality follows from the fact that the sum of delivered goods is equal to the sum of received goods. Moreover

$$\sum_{u=1}^{U} Q^u \geq \sum_{u=1}^{U} \widehat{Q}^u + \left(Q^\alpha - \widehat{Q}^\alpha\right).$$

Note that $Q^\alpha - \widehat{Q}^\alpha > 0$ due to $\alpha \leq U$. (The final time step 3b is executed, then either $Q^\alpha > \widehat{Q}^\alpha$ or $Q^\alpha = \widehat{Q}^\alpha$. In the latter case α is increased by one, and then $Q^\alpha > \widehat{Q}^\alpha = 0$). Using (18), we get

$$\sum_{u=1}^{U} Q^u \geq \sum_{u=1}^{\alpha} Q^u = \sum_{d=1}^{D} Q_d + \left(Q^\alpha - \widehat{Q}^\alpha\right),$$

a contradiction (since $\sum_{u=1}^{U} Q^u = \sum_{d=1}^{D} Q_d$). Similarly, we can not have $\beta \leq D$.

Finally, we show that the number of active links will be $U + D - 1$. Note that only one of α and β is increased by one (in step 3 of every iteration) when $\alpha < U$ and $\beta < D$. Assume the contrary. Then $\widehat{Q}^\alpha = Q^\alpha$ and $\widehat{q}^\beta = q^\beta$. Moreover

$$\sum_{u=1}^{\alpha} Q^u = \sum_{u=1}^{\alpha} \widehat{Q}^u = \sum_{d=1}^{\beta} \widehat{q}^d = \sum_{d=1}^{\beta} Q_d,$$

a contradiction since, generically, $\sum_{u=1}^{\alpha} Q^u \neq \sum_{d=1}^{\beta} Q_d$. Next, note that the algorithm ends after $U + D - 1$ iterations. The first time step 2 is executed, $\alpha + \beta = 2$. The last time step 2 is executed, $\alpha + \beta = U + D$. As exactly one quantity $\widehat{q}_{ud}$ is made strictly positive each time step 2 is performed, the number of active links will be $U + D - 1$.

Proposition 5

Point 1: We want to compare the prices that two different upstream firms h and l charge to the same buyer d, assuming that $q_{hd} = q_{ld} \equiv q_d$.

$$\Delta \equiv 2\left(p_{hd} - p_{ld}\right) q_d = \left[\overline{C}\left(Q^h, k^h\right) - \overline{C}\left(Q^h - q_d, k^h\right)\right] - \left[\overline{C}\left(Q^l, k^l\right) - \overline{C}\left(Q^l - q_d, k^l\right)\right]. \tag{19}$$

That is, the difference in price is determined by the difference in incremental cost that the two firms have for supplying q_d.

Assume that $k_l < k_h$. Then, $Q^l = Q\left(k^l\right) < Q\left(k^h\right) = Q^h$. To see this note that the first order conditions for efficiency $\overline{C}_q(Q(k), k) = p^*$ implies

$$Q'(z) = -\frac{\overline{C}_{qk}(Q(k), k)}{\overline{C}_{qq}(Q(k), k)} > 0. \tag{20}$$

since $\overline{C}_{qk} < 0$ and $\overline{C}_{qq} > 0$. The price difference is

$$\Delta = \int_{Q^h - q_d}^{Q^h} \overline{C}_q\left(x, k^h\right) dx - \int_{Q^l - q_d}^{Q^l} \overline{C}_q\left(x, k^l\right) dx \tag{21}$$

$$= \int_{Q^l - q_d}^{Q^l} \overline{C}_q\left(\left(Q^h - Q^l\right) + x, k^h\right) dx - \int_{Q^l - q_d}^{Q^l} \overline{C}_q\left(x, k^l\right) dx \tag{22}$$

Using $Q^u = Q(k^u)$

$$\Delta = \int_{Q^l - q_d}^{Q^l} \left[\overline{C}_q\left(\left(Q\left(k^h\right) - Q\left(k^l\right)\right) + x, k^h\right) - \overline{C}_q\left(\left(Q\left(k^l\right) - Q\left(k^l\right)\right) + x, k^l\right)\right] dx. \tag{23}$$

The expression within brackets can by defining $f(z) = \overline{C}_q\left(\left(Q(z) - Q\left(k^l\right)\right) + x, z\right)$ and using $f\left(k^h\right) - f\left(k^l\right) = \int_{k^l}^{k^h} f'(z)\, dz$ where

$$f'(z) = \overline{C}_{qq}\left(\left(Q(z) - Q\left(k^l\right)\right) + x, z\right) Q'(z) + \overline{C}_{qk}\left(\left(Q(z) - Q\left(k^l\right)\right) + x, z\right). \tag{24}$$

Using equation (20) to substitute for Q' yields

$$f'(z) = -\overline{C}_{qq}\left(\left(Q(z) - Q\left(k^l\right)\right) + x, z\right) \frac{\overline{C}_{qk}(Q(z), z)}{\overline{C}_{qq}(Q(z), z)} + \overline{C}_{qk}\left(\left(Q(z) - Q\left(k^l\right)\right) + x, z\right) \tag{25}$$

Hence:

$$\Delta = \int_{Q^l - q_d}^{Q^l} \int_{k^l}^{k^h} \left[-\overline{C}_{qq}\left(\left(Q(z) - Q^l\right) + x, z\right) \tfrac{\overline{C}_{qk}(Q(z), z)}{\overline{C}_{qq}(Q(z), z)} + \overline{C}_{qk}\left(\left(Q(z) - Q^l\right) + x, z\right)\right] dz dx \tag{26}$$

$$= \int_{Q^l - q_d}^{Q^l} \int_{k^l}^{k^h} \overline{C}_{qq}\left(\left(Q(z) - Q^l\right) + x, z\right) \left[\tfrac{\overline{C}_{qk}\left(\left(Q(z) - Q^l\right) + x, z\right)}{\overline{C}_{qq}\left(\left(Q(z) - Q^l\right) + x, z\right)} - \tfrac{\overline{C}_{qk}(Q(z), z)}{\overline{C}_{qq}(Q(z), z)}\right] dz dx. \tag{27}$$

Since $\overline{C}_{qq}$ is positive, the sign of the right-hand side is determined by the expression within brackets. Note that $\left(Q(z) - Q\left(k^l\right)\right) + x < Q(z)$ since $x \in \left[Q^l - q_d, Q^l\right]$. Hence, $\Delta > 0$ if $\overline{C}_{qk}(Q, k) / \overline{C}_{qq}(Q, k)$ is falling in Q.

Point 2: In the rest of the proof we omit the capital level as an argument in the cost function. Consider an upstream firm u in its dealings with two downstream firms 1 and 2. We want to find the most profitable distribution of sales (q_{u1}, q_{u2}) given that $q_{u1} + q_{u2} = x$. Reallocations of sales between different buyers do not affect cost, and we can focus on the revenues:

$$\sum_{d=1}^{D} p_{ud} q_{ud} = \sum_{d=1}^{D} \frac{1}{2} r q_{ud} - \frac{1}{2}\left[\underline{C}\left(\frac{Q}{D}\right) - \underline{C}\left(\frac{Q}{D} - q_{ud}\right)\right] + \frac{1}{2}\left[\overline{C}\left(\frac{Q}{U}\right) - \overline{C}\left(\frac{Q}{U} - q_{ud}\right)\right].$$

After an affine transformation, and assuming $q_{u1} + q_{u2} = x$, the revenues may be written

$$R(q_{u1}) = \underline{C}\left(\frac{Q}{D} - q_{u1}\right) + \underline{C}\left(\frac{Q}{D} - x + q_{u1}\right) - \overline{C}\left(\frac{Q}{U} - q_{u1}\right) - \overline{C}\left(\frac{Q}{U} - x + q_{u1}\right),$$

for all $q_{u1} \in [0, x]$.

Note that

$$R''(q_{u1}) = \left[\underline{C}_{qq}\left(\frac{Q}{D} - q_{u1}\right) - \overline{C}_{qq}\left(\frac{Q}{U} - q_{u1}\right)\right] + \left[\underline{C}_{qq}\left(\frac{Q}{D} - x + q_{u1}\right) - \overline{C}_{qq}\left(\frac{Q}{U} - x + q_{u1}\right)\right].$$

Hence, if the downstream (upstream) firms' marginal costs are steeper in the range of x units, then the revenue function is strictly convex (concave). In these cases there exists a unique interior extremum. It is simple to verify that $R'(x/2) = 0$ and hence that the interior extremum is at the symmetric allocation.

Note that

$$R'(0) = -R'(x) = \int_0^x \left[\underline{C}_{qq}\left(\frac{Q}{D} - z\right) - \overline{C}_{qq}\left(\frac{Q}{U} - z\right)\right] dz.$$

Hence, if the downstream (upstream) firms' marginal costs are steeper in the range of x units, then the revenue function is strictly increasing (decreasing) at $q_{u1} = 0$ and strictly decreasing (increasing) at $q_{u1} = 0$.

Thus, if the downstream (upstream) firms' marginal costs are steeper in the range of x units, then the revenue function has two minima (maxima) at the endpoints and a maximum (minimum) at the symmetric allocation.

Point 3: First, note that if the technology has (long-run) constant returns to scale, firm u's short run cost function can be written $k^u \overline{C}(Q^u/k^u)$. (This follows from the fact that the short-run cost function is homogeneous of degree one in Q^u and k^u.) Second, in equilibrium, firm u produces $Q^u = \left(k^u/\overline{K}\right) Q$. Since Q is efficient, it does not depend on the distribution of capital or the distribution of sales. Third, in equilibrium all upstream firms produce at the same average variable cost

$$\frac{k^u\overline{C}\left(Q^u/k^u\right)-k^u\overline{C}\left(0\right)}{Q^u}=\frac{\overline{K}C\left(Q/\overline{K}\right)-\overline{K}C\left(0\right)}{Q}\equiv\overline{a}$$

since $Q^u=\left(k^u/\overline{K}\right)Q$, and the equivalent applies for downstream firms.

When there are constant returns to scale, the price is given by

$$p_{ud}=\frac{r}{2}-\frac{1}{2}\left(\frac{\underline{C}\left(\frac{Q}{\underline{K}}\right)-\underline{C}\left(\frac{Q}{\underline{K}}-\frac{q_{ud}}{Q_d}\frac{Q}{\underline{K}}\right)}{\frac{q_{ud}}{Q_d}\frac{Q}{\underline{K}}}\right)+\frac{1}{2}\left(\frac{\overline{C}\left(\frac{Q}{\overline{K}}\right)-\overline{C}\left(\frac{Q}{\overline{K}}-\frac{q_{ud}}{Q^u}\frac{Q}{\overline{K}}\right)}{\frac{q_{ud}}{Q^u}\frac{Q}{\overline{K}}}\right).\tag{28}$$

Note that p_{ud} is monotonically decreasing in q_{ud}/Q^u and monotonically increasing in q_{ud}/Q_d. This follows from the fact that the marginal costs are increasing, implying that average incremental costs are higher at higher output levels.

Thus, to find an upper bound on p_{ud} assume that $q_{ud}/Q_d=1$ and that $q_{ud}/Q^u\approx 0$. Then

$$p_{ud}=\frac{r}{2}-\frac{1}{2}\frac{\underline{C}\left(Q/\underline{K}\right)-\underline{C}\left(0\right)}{\left(Q/\underline{K}\right)}+\frac{1}{2}\overline{C}'\left(Q/\overline{K}\right).\tag{29}$$

Since $\overline{C}'\left(Q/\overline{K}\right)=p^*$, $p_{ud}=\left(p^*+r-\underline{a}\right)/2$. The lower bound can be derived in a similar way.

To show that average prices can also attain the upper bound, assume that all upstream capital is concentrated in one firm, and that the downstream capital is split equally among D downstream firms. Then, $q_{ud}/Q_d=1$ and $q_{ud}/Q^u=Q_d/Q^u=\left(k_d/\overline{K}\right)/\left(k^u/\overline{K}\right)=\left(k_d/\overline{K}\right)=1/D$ for all (u,d). As $D\rightarrow\infty$, $q_{ud}/Q^u\rightarrow 0$ for all (u,d).

Proposition 6

For each economy s, select an equilibrium contract structure $\left(q\left(s\right),p\left(s\right)\right)$ that is socially efficient, and assume that $q_{ud}\left(s\right)\rightarrow 0$. The sequence of Walrasian prices $p^*\left(s\right)$ is characterized by $p^*_{ud}\left(s\right)=r/2-\left(\partial C_d\left(q\left(s\right)\right)/\partial q_{ud}\right)/2+\left(\partial C^u\left(q\left(s\right)\right)/\partial q_{ud}\right)/2$. The equilibrium prices are given by

$$p_{ud}\left(s\right)=\frac{r}{2}-\frac{1}{2}\left(\frac{C_d\left(q\left(s\right)\right)-C_d\left(q\left(s\right)\backslash q_{ud}=0\right)}{q_{ud}\left(s\right)}\right)+\frac{1}{2}\left(\frac{C^u\left(q\left(s\right)\right)-C^u\left(q\left(s\right)\backslash q_{ud}=0\right)}{q_{ud}\left(s\right)}\right).\tag{30}$$

According to the Mean Value Theorem, $C_d(q(s)) - C_d(q(s)\backslash q_{ud} = 0) = \frac{\partial C_d(q(s)\backslash \widetilde{q}_{ud}(s))}{\partial q_{ud}} q_{ud}(s)$ for some $\widetilde{q}_{ud}(s) \in (0, q_{ud}(s))$ and the equivalent applies for the upstream firm. Hence

$$p_{ud}(s) = \frac{r}{2} - \frac{1}{2}\frac{\partial C_d(q(s)\backslash \widetilde{q}_{ud}(s))}{\partial q_{ud}} + \frac{1}{2}\frac{\partial C^u(q(s)\backslash \widetilde{q}_{ud}(s))}{\partial q_{ud}}. \tag{31}$$

Since $q_{ud}(s) \to 0$, it follows that $\widetilde{q}_{ud}(s) \to q_{ud}(s)$, which proves the proposition.

Proposition 7

Assume now that both upstream and downstream firms are price takers on the intermediate goods market. Upstream firms will chose q_{ud} to satisfy $p_{ud} - \partial C^u(q^u)/\partial q_{ud} = 0$. Downstream firms will chose q_{ud} to satisfy $r_{ud}(q) + \sum_{i=1}^{U} q_{id}\partial r_{id}(q)/\partial q_{ud} - \partial C_d(q)/\partial q_{ud} - p_{ud} = 0$. Substituting $p_{ud} = \partial C^u(q^u)/\partial q_{ud}$ into the second equation yields the first order condition for bilateral efficiency.