Markups, Markdowns, and Bargaining in a Vertical Supply Chain

Preliminary Draft - Frequently Updated – the latest version can be found here

Rémi Avignon*, Claire Chambolle†, Etienne Guigue‡, and Hugo Molina§

This version: May 1, 2025. First version: December 20, 2024.

Abstract

This article bridges monopoly, monopsony, and countervailing power theories to analyze the welfare effects of seller and buyer power in a vertical supply chain. We develop a bilateral monopoly setting with bargaining over a linear price, where the upstream firm sources input from an increasing supply curve, exerting monopsony power mirroring the downstream firm monopoly power. In equilibrium, the short-side rule applies, meaning that the quantity traded is determined by the firm that is willing to trade the smaller amount. We show that welfare is maximized when each firm's bargaining power exactly countervails the other's market power. Otherwise, double marginalization occurs: *double markupization* arises when the upstream firm holds excessive bargaining power, and *double markdownization* in the opposite case. Our analysis yields novel insights for policy intervention and empirical research.

Keywords: Markups, Markdowns, Bargaining, Countervailing buyer power, Monopsony power, Bilateral monopoly.

JEL classification: C78, D42, J42, L81.

*INRAE-SMART, remi.avignon@inrae.fr

†Université Paris-Saclay, INRAE, AgroParisTech, Paris-Saclay Applied Economics, 91120, Palaiseau, France; claire.chambolle@inrae.fr; http://chambolle.io

‡LMU Munich, KU Leuven; etienne.guigue@econ.lmu.de

§Université Paris-Saclay, INRAE, AgroParisTech, Paris-Saclay Applied Economics, 91120, Palaiseau, France; hugo.molina@inrae.fr; https://sites.google.com/site/hugomolinaresearch

We would like to thank Laurent Linnemer, Volker Nocke, Patrick Rey, and Flavio Toxvaerd for helpful comments and discussions. We are also grateful to participants at the KU Leuven IO Seminar, CREST-Mannheim IO Days, EARIE (Amsterdam, 2024), CMA Durham Workshop, CMA-UKRI Workshop on Market Power in Supply Chains, M-POWER conference (KU Leuven, 2025), the MaCCI Annual Conference (2025), and the IO seminar in UC Berkeley for useful feedback. Etienne Guigue gratefully acknowledges financial support from Jan De Loecker's MPOWER ERC grant. We thank INSPIRE-Biosphera and the ANR-DFG PARETO for funding the project.

1 Introduction

Two prominent theories offer contrasting perspectives on the welfare effects of buyer power in vertical supply chains. The countervailing power theory, introduced by Galbraith (1952), suggests that buyer power mitigates seller market power, leading to lower markups, higher output quantity, and greater welfare.[1] In contrast, the monopsony power theory, originating with Robinson (1933), argues that buyer power increases the market power of dominant buyers, resulting in higher markdowns, lower output quantity, and a lesser welfare.[2]

Both theories have been highly influential in academic research and policymaking. For instance, a stream of research on vertical supply chains examines the factors underlying countervailing buyer power, highlighting how it reduces double marginalization and benefits consumers (see, e.g., Snyder, 2008; Smith, 2016; Lee, Whinston and Yurukoglu, 2021, for comprehensive surveys). Building on these insights, the concept of countervailing buyer power is frequently invoked in competition policy debates, either as an efficiency defense for downstream horizontal mergers or to justify the formation of buying alliances.[3] In parallel, a vast literature in labor economics documents the prevalence of monopsony power and examines the mechanisms to mitigate its adverse effects (see, e.g., Manning, 2021; Card, 2022; Azar and Marinescu, 2024, for reviews). Beyond the labor market, recent empirical work has highlighted that monopsony power is pervasive in various input markets (e.g., Morlacco, 2019; Avignon and Guigue, 2022; Treuren, 2022; Zavala, 2022; Rubens, 2023). Consequently, antitrust agencies have increasingly incorporated the concept of monopsony power into their analyses.[4] Thus, despite being grounded in different sets of assumptions, the countervailing and monop-

[1]More precisely, Galbraith's (1952) argument states that retailers (or intermediaries) with buyer power should negotiate lower prices from manufacturers and pass these benefits on to consumers through reduced output prices.

[2]Specifically, Robinson (1933) formalizes the idea that large employers have the potential to reduce employment and pay workers below their marginal revenue.

[3]See, e.g., the Horizontal Merger Guidelines of the European Commission (2004) and the JRC policy report on buying alliances (Daskalova et al., 2020).

[4]For instance, the U.S. Department of Justice sued to block a merger between two of the largest book publishers in 2021, mentioning the potential harm to American authors as the primary concern (United States v. Bertelsmann SE & Co. KGaA et al., No. CV 21-2886-FYP). See also the recent Federal Trade Commission's lawsuit to block the merger between the supermarket giants Albertsons and Kroger (press release).

sony power theories conflict in shaping appropriate antitrust treatment of buyer power.[5]

In this article, we develop a unified framework that incorporates both theories to provide new insights into the welfare effects of buyer power in vertical supply chains. Specifically, we consider a setting where an upstream monopolist, U, sells its product to a downstream monopolist, D, which then resells it to final consumers. To examine monopsony power, we depart from the canonical model of vertical contracting (e.g., Spengler, 1950), which typically assumes that U operates with constant marginal costs. Instead, we suppose that U sources its input from an upward-sloping supply curve, resulting in increasing marginal costs.[6] Mirroring D's exercise of monopoly power in the product market, U thus exercises monopsony power in the input market. We model the interactions between U and D as follows. First, U and D bargain over a linear wholesale price. Second, given this agreed wholesale price, U and D simultaneously announce the quantities they are each willing to trade. Assuming that exchange is voluntary (i.e., no firm is forced to trade more than it wants), the equilibrium quantity is determined by the minimum between what U is willing to sell (i.e., supply) and what D is willing to purchase (i.e., demand). It is worth noting that this modeling approach generalizes the canonical model of vertical relations, in which the equilibrium quantity is always determined by D. Specifically, the assumption that D dictates the quantity exchanged on the market is innocuous when U faces constant marginal costs, as it is willing to supply any quantity in equilibrium. However, we show that this is no longer the case under increasing marginal costs. In that case, for a given wholesale price, supplying an additional unit may not be profitable for U, as doing so raises the marginal costs of all other units supplied.

We highlight that the distribution of bargaining power between U and D affects both the magnitude and the nature of the double marginalization phenomenon. When D's bargaining power vis-à-vis U is low, the equilibrium wholesale price and quantity move along D's marginal revenue curve (i.e., its demand for U's product). The intuition is as follows. As the wholesale price is high, U is willing to supply a quan-

[5] As highlighted by Hemphill and Rose (2018), the Federal Trade Commission and the U.S. Department of Justice have adopted conflicting views on buyer power treatment in recent merger reviews.

[6] For instance, this increasing supply curve may result from the aggregation of individual price-taking input suppliers.

tity exceeding D's demand. As a result, the quantity exchanged in equilibrium is determined by D, which is on the "short side" of the market (i.e., the side with the willingness to trade the smaller quantity). By internalizing D's downward-sloping demand, U exercises monopoly power when selling to D by charging a markup. This markup adds up to D's markup stemming from its monopoly power in the product market, resulting in a lower quantity and a higher consumer price compared to what a vertically integrated firm would set in equilibrium. This *double markup* gives rise to the classical double marginalization problem highlighted by Spengler (1950).[7] In this case, Galbraith's (1952) countervailing buyer power argument applies: increasing D's bargaining power reduces U's markup, which improves welfare.

In contrast, when D's bargaining power vis-à-vis U is high, we find that the equilibrium wholesale price and quantity move along U's marginal cost curve. The logic is analogous: as the wholesale price is low, U is willing to supply a quantity smaller than D's demand. Thus, U is on the "short side" of the market and consequently determines the quantity exchanged in equilibrium. By internalizing U's upward-sloping marginal cost, D exercises monopsony power when purchasing from U by charging a markdown. This markdown adds up to U's markdown stemming from its monopsony power in the input market.[8] We show that this *double markdownization* mirrors the double markup scenario and constitutes a novel source of double marginalization. In this case, Galbraith's (1952) argument no longer applies: increasing D's bargaining power raises its markdown, which reduces welfare. Instead, we show that enhancing U's bargaining power vis-à-vis D improves welfare by strengthening its ability to exercise *countervailing seller power*.[9]

We further characterize the level of D's bargaining power vis-à-vis U at which each firm fully countervails the other's market power, thereby eliminating double marginalization and achieving the vertically integrated outcome. This level of bargaining power depends on the underlying supply and demand primitives, and includes two limiting

[7] As discussed by Linnemer (2022), the double marginalization phenomenon commonly attributed to Spengler (1950) is originally due to Cournot (1838) and Edgeworth (1925).

[8] More generally, D charges a markdown whenever U has increasing marginal costs, regardless of its underlying cause (e.g., monopsony power in the input market, decreasing returns to scale).

[9] This reasoning mirrors Galbraith's (1952) countervailing buyer power argument under double markup, where D's bargaining power mitigates U's markup.

cases: D should hold all the bargaining power when U faces a perfectly elastic supply curve (i.e., constant marginal cost), whereas U should hold all the bargaining power when D faces a perfectly elastic demand curve.

We extend our analysis in two directions. First, we consider the case in which U and D bargain over a two-part tariff contract. We show that our results continue to hold qualitatively in the presence of frictions that limit the use of the fixed fee as a rent-extraction device, preventing the elimination of the double marginalization problem. Second, we show that an input price floor policy aimed at protecting input suppliers eliminates markdowns. Specifically, whenever the price floor is binding, U operates under constant marginal costs, thereby precluding the exercise of monopsony power. We demonstrate that there exists an optimal level of the price floor that always increases the quantity traded and overall welfare. This optimal level depends on the underlying supply and demand primitives, as well as the distribution of bargaining power within the vertical supply chain. Moreover, its welfare-improving effect is greater under double markdownization (i.e., when D's bargaining power vis-à-vis U is high).

Contributions. We build and contribute to the extensive literature on vertical relations that, following the pioneering work of Spengler (1950), explores the sources and consequences of the double marginalization phenomenon and its potential remedies.[10] Specifically, a strand of this literature on firm-to-firm bargaining typically assumes constant marginal costs of production.[11] Our main contribution is to relax this assumption by considering the presence of monopsony power in the input market. This allows us to identify a novel source of double marginalization, which we refer to as double markdownization.[12] Importantly, we show that this new distortion has significant

[10]Double marginalization analysis has a long tradition in the industrial organization literature, see Tirole (1988) for a textbook exposition and Rey and Vergé (2008) for a review. Recent contributions to this literature include Janssen and Shelegia (2015); Crawford et al. (2018); Luco and Marshall (2020); Choné, Linnemer and Vergé (2024); Ghili and Schmitt (2024), among others. Recent work also shows how double marginalization generates aggregate distortions in input-output networks (see, e.g., Baqaee, 2018; Baqaee and Farhi, 2020; Dhyne, Kikkawa and Magerman, 2022; Arkolakis, Huneeus and Miyauchi, 2023).

[11]See, e.g., Horn and Wolinsky (1988); Dobson and Waterson (1997, 2007); Allain and Chambolle (2011); Iozzi and Valletti (2014); Gaudin (2016, 2018); Rey and Vergé (2020) in the industrial organization literature; and Grossman, Helpman and Sabal (2024) in the trade literature.

[12]Double markupization and double markdownization relate to the classical double marginalization problem in vertical relationships. Importantly, this differs from the "double markup" and "double

welfare implications. Unlike the canonical model of vertical contracting—where double marginalization arises exclusively through double markup—we find that increasing D's bargaining power under double markdownization reduces welfare. Instead, as U exercises countervailing seller power, increasing its bargaining power mitigates D's markdown and improves welfare. Our findings thus suggest that there exists a distribution of bargaining power along the vertical supply chain that eliminates double marginalization and restores efficiency. This result relates to Loertscher and Marx (2022), who demonstrate that equalizing bargaining power can enhance outcomes under incomplete information. However, a key distinction in our analysis – conducted within a complete information framework – is that the level of bargaining power that leads to efficiency is not necessarily symmetric (i.e., $1/2$), but rather depends on the underlying supply and demand primitives. In contemporaneous work, Demirer and Rubens (2025) also derive a closely related result, characterizing the existence of a level of buyer power that offsets either U's markup or D's markdown. Our articles are complementary in several respects. We develop a different, micro-founded framework with monopoly and monopsony power, uncover the notion of double markdownization, and provide general definitions and expressions for markups and markdowns in a vertical supply chain.[13] Whereas we focus on the classical sequential timing of the bilateral monopoly model, they also study the case where the wholesale and the retail prices are set simultaneously and take their model to data.

The nature of the double marginalization phenomenon (double-markup or double-markdown) depends on whether U or D ultimately sets the quantity to be traded in equilibrium, that is, which firm has the "right-to-manage". As underscored by Toxvaerd (2024), the allocation of the right-to-manage in bilateral monopolies with increasing marginal production costs and linear tariffs remains a long-standing and unresolved

markdown" in Kroft et al. (2023), where a single firm with market power in both its input and output markets marks up the input price twice to set the output price, or equivalently, marks down the output price twice to determine the input price.

[13]In particular, a distinctive feature of our approach is that the right-to-manage allocation (vertical conduct in their terminology) results from voluntary exchange, whereas Demirer and Rubens (2025) rely on participation constraints, e.g., "D [(resp. U)] participates in bargaining if its resulting markdown [(resp. markup)] is nonnegative". Moreover, we emphasize that classical markup and markdown definitions do not directly extend to bargaining settings. To address this, we propose general definitions for markups and markdowns.

issue.[14] Confronted with this modeling challenge, recent work in labor economics and international trade has exogenously assigned the right-to-manage to one or the other side of the market (e.g., Azkarate-Askasua and Zerecero, 2022; Alviarez et al., 2023; Wong, 2023). We contribute to the literature by proposing a non-cooperative allocation of the right-to-manage, grounded in the subgame perfection criterion and the natural assumption of voluntary exchange.[15] We thus restore symmetry in firms' ex ante quantity setting power, preventing a firm from compelling the other to trade more than it is willing to.[16] In doing so, we endogenize (i) the nature of the distortion arising in the vertical supply chain (markup or markdown) and, consequently, (ii) the welfare implications of each firm's bargaining power (detrimental or improving).

Our findings may also have important implications for empirical research on bargaining in vertical supply chains. As reviewed by Lee, Whinston and Yurukoglu (2021), it is common practice to assume that upstream manufacturers operate with constant marginal costs.[17] We show that the welfare consequences of the balance of power in the vertical supply chain can vary substantially depending on the slope of the marginal cost function. Our results thus call for greater flexibility in modeling cost functions in empirical work. This is particularly relevant given the prevalence of convex supply curves in many industries (e.g., Shea, 1993; Boehm and Pandalai-Nayar, 2022). In this context, inferring whether upstream or downstream firms have the right-to-manage becomes essential for estimating markups and markdowns. In addition to Demirer and Rubens (2025), a first step in this direction is developed by Atkin et al. (2024) who exploit an Argentinian import license policy that exogenously affects traded volumes

[14]In Fellner's (1947) pioneering analysis of bilateral monopolies, when the seller (resp. buyer) makes the wholesale price offer, the buyer (resp. seller) is assumed to freely determine the quantity it intends to purchase (sell) at the offered price. However, as Toxvaerd (2024) points out, no solution has been provided to the right-to-manage allocation: "it is not clear why either firm would want to cede the right to set output to the other firm, even if a wholesale price could be agreed upon". Several articles have circumvented this issue with efficient bargaining (e.g., the price and quantity are jointly negotiated as in McDonald and Solow, 1981; or the wholesale price is non-linear as in Chipty and Snyder, 1999).

[15]In contemporaneous work, Houba (2024) instead relies on a cooperative solution where firms Nash bargain over both the wholesale price and the allocation of the right-to-manage.

[16]A similar logic appears in Falch and Strøm (2007). However, their firm-union bargaining model differs markedly from our setting, as it does not account for vertical relations (and, hence, double marginalization), and both total payroll and employment directly enter the union's objective function.

[17]Among others, see Draganska, Klapper and Villas-Boas (2010); Crawford and Yurukoglu (2012); Ho and Lee (2017); Crawford et al. (2018); Noton and Elberg (2018); Sheu and Taragin (2021); Bonnet, Bouamra-Mechemache and Molina (2023).

to identify whether the importer or exporter determines the equilibrium quantity.

Finally, we contribute to the analysis of input price floors (minimum wages), which has been extensively studied in the labor economics literature.[18] Since at least Stigler (1946), it is well-known that minimum wages can increase employment in the presence of labor monopsony power.[19] The incidence of minimum wage (or more broadly, input price floor) policies has also been examined in oligopoly-oligopsony models, where a set of firms exert both monopoly power in the output market and monopsony power in the input market (e.g., Russo, Goodhue and Sexton, 2011; Avignon and Guigue, 2022; Hernández and Cantillo-Cleves, 2024).[20] To the best of our knowledge, we are the first to extend this analysis to a vertical supply chain with bargaining. In addition to showing that input price floors can improve welfare, we emphasize that the optimal design of such policies depends critically on the nature of the double marginalization phenomenon (double-markup or double-markdown).

The remainder of this article is organized as follows. Section 2 provides markup and markdown definitions that accommodate both unilateral price-setting and vertical bargaining models. Section 3 presents our vertical supply chain framework and considers a benchmark case where U and D are vertically integrated. Section 4 solves our model and characterizes the markup(s) and markdown(s) that emerge along the vertical supply chain. Section 5 analyzes the welfare implications of D's and U's bargaining power. Section 6 discusses modeling assumptions, extends the analysis to two-part tariff contracts, and examines the impact of an input price floor policy. Section 7 concludes.

2 Markups and Markdowns

Markups and markdowns measure price distortions that result from firms' exercise of market power, which leads to market failures by negatively affecting welfare and re-

[18] See Azar and Marinescu (2024) and Dube and Linder (2024) for recent surveys.

[19] Card and Krueger (1994) provide early empirical evidence of the zero or positive effect of minimum wages on employment, and Azar et al. (2024) offer the first direct evidence supporting the monopsony explanation (see, e.g., Card and Krueger, 1995; Manning, 2003, for textbook treatments).

[20] Lemos's (2008) survey of the empirical labor literature finds that firms' price responses to minimum wage increases are moderate, suggesting that cost pass-through to prices is incomplete.

source allocation (e.g., Tirole, 2015). A markup is traditionally defined as the ratio of a firm's output price to its marginal cost, measuring the upward price distortion associated with the firm's seller power. Symmetrically, a markdown is traditionally defined as the ratio of an input's marginal revenue product to its purchase price, measuring the downward price distortion arising from the firm's buyer power.[21] Markups and markdowns greater than 1 typically arise under two conditions:

(i) firms behave strategically to extract surplus from imperfectly price-elastic demand and/or supply,

(ii) surplus extraction occurs through the use of linear tariffs.

Specifically, under conditions (i) and (ii), a firm facing a downward-sloping demand curve for its output faces the following basic trade-off when deciding whether to sell an additional unit. On the one hand, selling one more unit generates extra revenue. On the other hand, doing so requires lowering the price on all units already offered for sale, thereby reducing revenue on inframarginal units. This latter effect leads the firm to restrict output (relative to perfect competition) and charge a markup over marginal cost.[22] Similarly, when purchasing its input in a market with an upward-sloping supply curve, the firm incurs a cost increase from buying an additional unit, as doing so raises the price paid on all other units purchased. This leads the firm to restrict input and charge a markdown below its marginal revenue product.

In what follows, we generalize the classical definitions of a markup and a markdown to encompass vertical supply chain settings with bargaining:

[21] The *marginal revenue product* (MRP) refers to the *marginal product* (MP)—the additional output generated by one more unit of input—multiplied by the *marginal revenue* (MR)—the additional revenue from selling the extra unit of output. In a one-to-one technology setting, we have $MP = 1$, so that $MRP = MR$.

[22] To illustrate the joint role of conditions (i) and (ii), consider a standard monopolist facing a downward-sloping (i.e., imperfectly elastic) demand curve (condition (i)) and charging a uniform unit price $x(q)$ (condition (ii)). To sell an additional unit, the monopolist must lower the output price, incurring a revenue loss on inframarginal units equal to $x'(q)q$. This loss provides the monopolist with incentives to set a price above its marginal cost, thereby charging a markup. If, instead, it faces a perfectly elastic demand (i.e., $x'(q) = 0$) or behaves as a price-taker, it does not incur, nor internalize, any revenue loss on inframarginal units, and thus no markup arises. Similarly, if the monopolist faces an imperfectly elastic demand but can engage in perfect (first-degree) price discrimination, it fully extracts consumer surplus. In that case, it has no incentive to restrict output, and the price of the last unit sold equals marginal cost—again implying no markup.

Definition 1 *The markup of firm i is defined as $\mu_i \equiv \frac{x_i}{\hat{x}_i}$, where x_i is the price at which the firm sells its marginal unit of output, and $\hat{x}_i$ is the minimum price required for this unit to be sold absent any effect on the firm's revenue from all other units offered for sale.*

Definition 2 *The markdown of firm i is defined as $\nu_i \equiv \frac{\hat{z}_i}{z_i}$, where z_i is the price at which the firm buys its marginal unit of input, and $\hat{z}_i$ is the maximum price required for this unit to be purchased absent any effect on the cost firm i incurs for all other units purchased.*

These definitions cover the standard expressions for a markup and a markdown in settings where prices are unilaterally set by a firm. For instance, consider that D is a monopolist facing a downward-sloping inverse demand curve for its output $p(q)$. D's profit-maximizing condition requires that its marginal revenue equals its marginal cost: i.e., $MR_D(q^*) = MC_D(q^*)$, where q^* denotes the equilibrium quantity and $MR_D(q^*) \equiv p(q^*) + p'(q^*)q^*$. Absent any (negative) effect on D's revenue from all other units sold, which is captured by $p'(q^*)q^*$, the minimum price at which the firm would be willing to sell its marginal unit of output is given by $\hat{p} = MC_D(q^*) < p(q^*)$. Hence, from Definition 1, D's markup is given by $\mu_D \equiv \frac{p(q^*)}{\hat{p}(q^*)} = \frac{p(q^*)}{MC_D(q^*)} (= \frac{p(q^*)}{MR_D(q^*)})$, which coincides with the standard definition of a markup. Symmetrically, consider that U is a monopsonist operating under a one-to-one production technology and facing an upward-sloping inverse supply curve for its input $r(q)$. Again, U's profit-maximizing condition is such that $MC_U(q^*) = MR_U(q^*)$, where $MC_U(q^*) \equiv r'(q^*)q^* + r(q^*)$. Absent any effect on U's cost to acquire all other input units, which is captured by $r'(q^*)q^*$, the maximum price at which the firm would be willing to purchase its marginal unit of input is given by $\hat{r} = MR_U(q^*) > r(q^*)$. Thus, from Definition 2, U's markdown is given by $\nu_U \equiv \frac{\hat{r}(q^*)}{r(q^*)} = \frac{MR_U(q^*)}{r(q^*)} (= \frac{MC_U(q^*)}{r(q^*)})$, which aligns with the standard definition of a markdown. This correspondence also holds in a bilateral monopoly setting, where D acts as a monopolist in the output market, U acts as a monopsonist in the input market, and the linear wholesale price that D pays to U is unilaterally determined by either firm. In this context, as is well-known from the earlier literature on bilateral monopoly (e.g., Bowley, 1928; Tintner, 1939), the firm that sets the wholesale price does so by equating its marginal revenue with its marginal cost. Consequently, Definitions 1 and 2 continue to yield the standard expressions for firms' markups and markdowns along the vertical supply chain (see Appendix A.7.1 for details).

In the more general case considered in this article, where U and D engage in bilateral negotiation, however, the wholesale price w is no longer pinned down by the intersection of either firm's marginal revenue and marginal cost, as it also reflects the firms' relative bargaining positions. Thus, it is no longer clear that the markup and markdown definitions developed for settings in which prices are unilaterally set by firms remain appropriate measures of firms' market power.[23] As we are not aware of any established markup and markdown definitions in the context of vertical bargaining, we rely on Definitions 1 and 2. Specifically, analyzing U and D's markups and markdowns, these definitions yield the following expressions: $\mu_D = \frac{p(q^*)}{\hat{p}} = \frac{p(q^*)}{MR_D(q^*)}$, $\nu_D = \frac{\hat{w}(q^*)}{w(q^*)} = \frac{MR_D(q^*)}{w(q^*)}$, whereas $\mu_U = \frac{w(q^*)}{\hat{w}(q^*)} = \frac{w(q^*)}{MC_U(q^*)}$ and $\nu_U = \frac{\hat{r}(q^*)}{r(q^*)} = \frac{MC_U(q^*)}{r(q^*)}$ (see Appendix A.7.2 for details).[24] As a result, these definitions allow us to keep consistent definitions of markups and markdowns that are independent of the firm's position in the vertical chain. Notably, they preserve the logic underlying the standard markup and markdown definitions that the upward price distortion from D's monopoly power in the output market stems from $p'(q^*)q^*$, and the downward price distortion from U's monopsony power in the input market stems from $r'(q^*)q^*$.

3 Vertical Chain and Integration Benchmark

3.1 Vertical Chain

Consider a vertical chain in which an upstream firm, U, purchases an input at a price r to produce a good sold to consumers at a price p through a downstream firm, D. We assume that U operates with one-to-one production technology and incurs no costs

[23] For instance, as previously discussed, the upward price distortion in the output market stems from the term $p'(q^*)q^*$ in D's marginal revenue. When $MR_D(q^*) \neq MC_D(q^*)$, it is no longer clear that $\frac{p(q^*)}{MC_D(q^*)}$ accurately reflects D's seller power.

[24] In the special case where U has no monopsony power in the input market, it is worth noting that our definition for D's markup boils down to the ratio of the output price to the (negotiated) wholesale price. Hence, we recover the markup expressions already used in Spengler's (1950) canonical model of vertical contracting and its extensions to bilateral bargains (e.g., Gaudin, 2016).

beyond the input price. Likewise, D incurs no costs beyond the wholesale price w paid to U. The inverse supply function $r(q)$ faced by U and the inverse demand function $p(q)$ faced by D satisfy the following Assumption that ensures the existence of a profit-maximizing equilibrium:

Assumption 1 *The inverse supply curve $r(q)$ and the inverse demand curve $p(q)$ are three-times differentiable and such that:*

(i) $r'(\cdot) \geq 0$ *and* $\sigma_r(\cdot) > -2$*;*

(ii) $p'(\cdot) < 0$, $\varepsilon_p(\cdot) > 1$, *and* $\sigma_p(\cdot) < 2$*;*

(iii) $p(0) > r(0) > 0$ *and* $\lim_{q \to +\infty} p(q) = 0$,

where, for any function $f(\cdot)$, $\epsilon_f(q) \equiv \frac{f(q)}{q|f'(q)|}$ *is the (inverse) elasticity of* $f(\cdot)$, *and* $\sigma_f(q) \equiv \frac{qf''(q)}{|f'(q)|}$ *is a measure of convexity of* $f(\cdot)$.

Assumption 1.(i) implies that U faces an increasing inverse supply curve $r(q)$ and that its marginal cost $MC_U(q)$ increases with quantity q. Note that the case of constant marginal cost is included as a special case. Assumption 1.(ii) implies that D faces a strictly decreasing inverse demand curve $p(q)$ and that its marginal revenue $MR_D(q)$ is positive and strictly decreasing in quantity. Finally, Assumption 1.(iii) implies that $MC_U(q)$ and $MR_D(q)$ intersect.

Welfare is defined by $W(q) \equiv \int_0^q [p(q) - r(q)]dq$ and q_W denotes the welfare-maximizing quantity, which is characterized by $p(q_W) = r(q_W)$. Consumer surplus is defined by $CS(q) \equiv \int_0^q p(x)\,dx - p(q)q$, and input suppliers' surplus by $SS(q) \equiv r(q)q - \int_0^q r(x)\,dx$. Both $CS(q)$ and $SS(q)$ are strictly increasing in q.

3.2 Vertical Integration Benchmark

Consider a benchmark case in which a vertically integrated fim, denoted I, purchases an input at a price r to produce a good that it sells to consumers at a price p. Acting both as a monopolist on the output market and a monopsonist on the input market, I's maximization problem is given by:

$$\max_q \pi_I = (p(q) - r(q))\,q.$$

which yields the following first-order condition:

$$\underbrace{p(q_I)\left(1-\varepsilon_p^{-1}(q_I)\right)}_{MR_I(q_I)} = \underbrace{r(q_I)\left(1+\varepsilon_r^{-1}(q_I)\right)}_{MC_I(q_I)}. \tag{1}$$

where q_I denotes the corresponding equilibrium quantity in which I's marginal revenue equals its marginal cost. The exercise of monopoly power over consumers implies that I's marginal revenue differs from the output price $p(q_I)$ by a *wedge* equal to $1-\varepsilon_p^{-1}(q_I)$. Similarly, the exercise of monopsony power over input suppliers implies that I's marginal cost differs from the input price $r(q_I)$ by a *wedge* equal to $1+\varepsilon_r^{-1}(q_I)$. From (1), we obtain the following proposition:

Proposition 1 (Vertical Integration) *The vertically integrated firm I sets the equilibrium quantity $q_I < q_W$. The consumer price is $p(q_I) > p(q_W)$, the input price is $r(q_I) < r(q_W)$, and I's markup, markdown, and total margin are given by:*

$$\begin{aligned}
\mu_I &= \frac{p(q_I)}{MC_I(q_I)} = \frac{\varepsilon_p}{\varepsilon_p - 1}, \\
\nu_I &= \frac{MR_I(q_I)}{r(q_I)} = \frac{\varepsilon_r + 1}{\varepsilon_r}, \\
M_I &\equiv \frac{p(q_I)}{r(q_I)} = \nu_I \times \mu_I.
\end{aligned}$$

Proof. Appendix A.7.1 provides a formal derivation of markup and markdown expressions based on definitions introduced in Section 2. ■

As previously defined, μ_I measures the surplus I obtains from selling the marginal output unit, or simply here I's ability to set p above its marginal cost. The markup expression indicates a negative relationship with demand elasticity: as ε_p increases, μ_I decreases. Similarly, ν_I measures the surplus I obtains from purchasing the marginal input unit, or simply here I's ability to purchase the input at a price below its marginal revenue. The markdown expression indicates a negative relationship with supply elasticity: as ε_r increases, ν_I decreases. Note that $\mu_I = 1$ in the absence of monopoly power, and $\nu_I = 1$ in the absence of monopsony power. Finally, we introduce here the definition of the margin M_I of firm I, which measures the total surplus I obtains

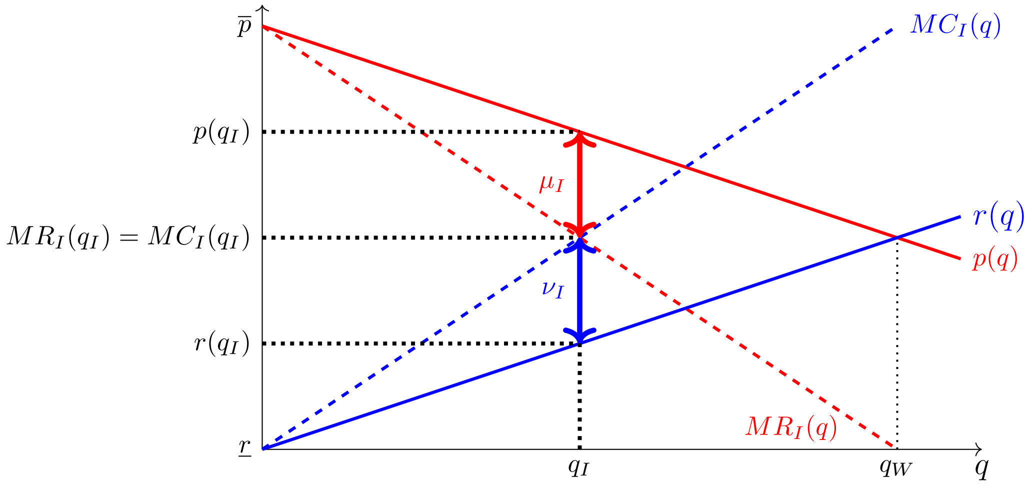


Figure 1: Monopoly and Monopsony Power in the Vertically Integrated Case
$p(q) = \overline{p} - \frac{1}{3}q$ and $r(q) = \underline{r} + \frac{1}{3}q$.

from both purchasing and selling the marginal unit. In this article's framework, a firm margin can be (i) trivially defined as the ratio of its output price (here, p) and input price (r) and (ii) written as the product of the firm markup (μ_I) and markdown (ν_I).[25] In this benchmark case, firm's I margin is thus negatively related to both demand and supply elasticities and equal to one in the absence of monopoly and monopsony power.

Figure 1 illustrates the economic forces in (1) and Proposition 1 by depicting the profit-maximizing equilibrium under linear demand and supply functions. The figure highlights the following mechanism. Firm I exercises both its monopoly and monopsony power. Given the downward-sloping inverse demand curve, I internalizes that selling one more output unit lowers the output price for all other units. Similarly, given the upward-sloping inverse supply curve, I internalizes that buying one more input unit drives up the input price for all other units. Hence, I is incentivized to reduce the quantity exchanged to q_I, generating negative welfare effects ($q_I < q_W$). This quantity reduction distorts prices, implying that consumers pay higher prices while input suppliers receive lower prices.

Figure 11 in Appendix D illustrates how variations in the elasticity of supply impact both markup and markdown. When supply becomes less elastic as compared

[25] The expression of the margin as the product of the firm markup and markdown also extends to any production function with multiple outputs and substitutable inputs. In such frameworks, the margin M obtained from selling a given output quantity q at a price p and purchasing a given variable input quantity m at a price w would be defined as $M \equiv \theta_m \frac{pq}{wm}$ with $\theta_m = \frac{\partial q}{\partial m}\frac{m}{q}$.

to Figure 1, as shown on the left side of the figure, the markdown decreases. Conversely, a more elastic supply—depicted on the right—leads to a larger markdown. In both scenarios, the markup is also marginally affected: the reduced elasticity on the left slightly increases the equilibrium quantity, while the increased elasticity on the right slightly reduces it.

Building on insights from this benchmark case, we now examine our vertical chain framework with monopoly power, monopsony power, and a general distribution of bargaining power between the upstream firm U and the downstream firm D.

4 Bargaining and Double Marginalization

We now analyze the bilateral monopoly setting introduced in Section 3.1, where U purchases an input at price $r(q)$ to produce a good sold to consumers at price $p(q)$ through D. Specifically, we consider that U and D interact in the market according to the following sequence of play:

- **Stage 1:** U and D engage in a bilateral negotiation to determine the linear wholesale price w.
- **Stage 2:** U and D simultaneously announce the quantities q_U and q_D they are willing to trade. Exchange is voluntary, implying that the quantity traded is the minimum of q_U and q_D.

This bilateral monopoly setting nests the canonical model of Spengler (1950) and its extension to bargaining (e.g., Gaudin, 2016). We now discuss each stage and introduce our equilibrium notion. In Stage 1, we use the Nash bargaining solution (Nash, 1950) to determine the linear wholesale price negotiated between U and D, where $\alpha \in [0, 1]$ denotes U's bargaining weight vis-à-vis D.[26] In Stage 2, the equilibrium quantity traded is given by the minimum of the two announced quantities, reflecting

[26]The use of simple linear wholesale tariffs has been documented in the Chilean coffee market (Noton and Elberg, 2018), the UK liquid milk market (Smith and Thanassoulis, 2015), and various other sectors (see, e.g., Mortimer, 2008; Crawford and Yurukoglu, 2012; Grennan, 2013; Gowrisankaran, Nevo and Town, 2015; Ho and Lee, 2017). We consider the case where U and D bargain over a two-part tariff contract in Section 6.2.

that neither firm can compel the other to trade more than it is willing to.[27] It is worth noting that this voluntary exchange assumption is implicit in the canonical model of bilateral monopoly, where $r'(q) = 0$ and U announces $q_U = \infty$ whenever it charges a markup.[28]

In Appendix B, we provide a microfoundation for our bilateral monopoly model. Specifically, we show that our equilibrium outcome coincides with the subgame perfect Nash equilibrium of a noncooperative game in which: (i) U and D bargain according to the random-proposer protocol of Rey and Vergé (2020), and (ii) U and D unilaterally set their input and output prices, respectively.

4.1 Quantity Choice

In Stage 2, given w, U and D simultaneously announce the quantity $q_U(w)$ and $q_D(w)$ they are willing to trade. D's optimal quantity to purchase from U and resell to consumers is given by:[29]

$$\tilde{q}_D(w) \in \underset{q_D}{\text{argmax}}\ \pi_D \equiv (p(q_D) - w)q_D, \tag{2}$$

which satisfies the following first-order condition:

$$MR_D(\tilde{q}_D(w)) = w. \tag{3}$$

Similarly, U's optimal quantity of input to purchase and sell to D is given by:[30]

$$\tilde{q}_U(w) \in \underset{q_U}{\text{argmax}}\ \pi_U \equiv (w - r(q_U))q_U, \tag{4}$$

[27]Voluntary exchange is a natural feature of most markets and a standard assumption in both the Walrasian and non-Walrasian theories (e.g., Bénassy, 1993).

[28]More precisely, as U charges a markup and its marginal cost is constant, it is willing to supply an infinite quantity. As consumer demand is not perfectly elastic, D lies on the "short" side of the market and thus always determines the equilibrium traded quantity.

[29]Note that choosing either q_D or p leads to the same result because D operates as a monopolist.

[30]Again, choosing either q_U or r leads to the same result because U operates as a monopolist.

which satisfies the following first-order condition:

$$MC_U(\tilde{q}_U(w)) = w. \tag{5}$$

As shown by (3) and (5), D's profit-maximizing quantity is determined according to its demand (MR_D), whereas U's profit-maximizing quantity is determined according to its supply (MC_U). Note that as $r(q)$ becomes flatter, $MC_U(q)$ decreases, which, according to (5), strengthens U's incentives to increase $\tilde{q}_U(w)$. In the limit case where $r(q) = MC_U(q) = r$, it follows directly from (4) that $\tilde{q}_U(w) = \infty$ whenever $w - r \geq 0$.

Given voluntary exchange, there exists a multiplicity of Nash equilibria in (weakly) dominated strategies. However, both the trembling-hand and Pareto dominance criteria select the dominant-strategy equilibrium outcome $q(w)$ characterized in the following lemma:[31]

Lemma 1 *There exists a unique subgame equilibrium in dominant strategies such that U announces $\tilde{q}_U(w)$, D announces $\tilde{q}_D(w)$, and the quantity traded is:*

$$q(w) = \min\{\tilde{q}_U(w), \tilde{q}_D(w)\} \leq q_I.$$

Proof. See Appendix A.1. ■

Two comments are in order. First, the equilibrium characterized in Lemma 1 shares many features with the typical exchange process in non-Walrasian (or rationed) equilibria.[32] In particular, the equilibrium traded quantity corresponds to the profit-maximizing quantity of at least one firm, thereby satisfying the market efficiency prop-

[31] For instance, if U believes that D will announce $\hat{q} < \tilde{q}_D$, one best response for U is to also announce $\hat{q}$, although it is a weakly dominated strategy. The reasoning is symmetric for D if it believes that U will announce $\hat{q} < \tilde{q}_U$, and hence, any strategy profile ($\hat{q}$,$\hat{q}$) with $\hat{q} < \min\{\tilde{q}_U(w),\tilde{q}_D(w)\}$ is a Nash equilibrium of the announcement game. However, such equilibria are destroyed by the trembling-hand criteria as both U and D are better off announcing $\check{q} > \hat{q}$ whenever the other firm trembles upward. In addition, when $\tilde{q}_U(w) < \tilde{q}_D(w)$, announcing any quantity in the interval $[\tilde{q}_U(w), \tilde{q}_D(w)]$ is a best response for D. Symmetrically, when $\tilde{q}_D(w) < \tilde{q}_U(w)$, announcing any quantity in the interval $[\tilde{q}_D(w), \tilde{q}_U(w)]$ is a best response for U. However, such asymmetric announcements lead to the same equilibrium outcome as described in Lemma 1. Finally, it is straightforward that the Pareto dominance criterion also selects the equilibrium leading to the same outcome as in Lemma 1.

[32] Pioneering works on non-Walrasian equilibria include Barro and Grossman (1971); Bénassy (1975); Drèze (1975); Varian (1977); Hahn (1978), among others. See, e.g., Bénassy (1986, 1990) for a textbook treatment.

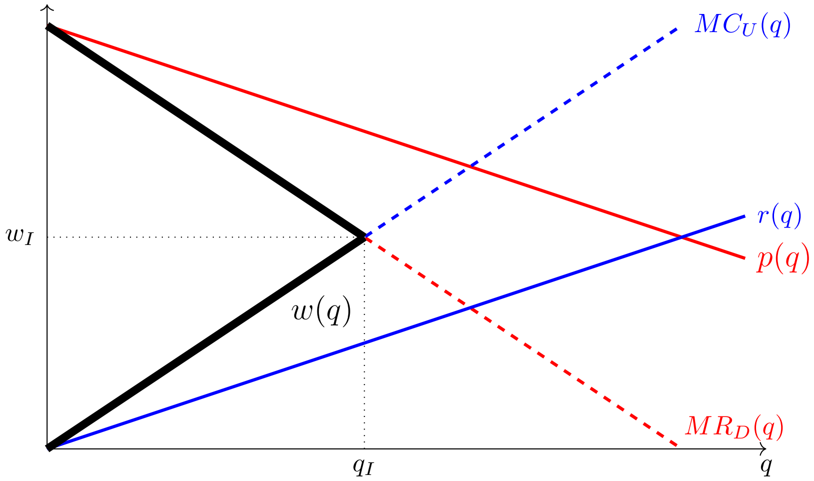


Figure 2: Short-Side Rule.

erty.[33] Combined with the voluntary exchange assumption, this implies that the "short-side" rule emerges in equilibrium: the firm on the short side of the market realizes its profit-maximizing outcome. Second, due to the double marginalization phenomenon, the equilibrium quantity is (weakly) lower than the vertically integrated outcome, q_I. To see this, consider for instance the case where $\tilde{q}_D(w) < \tilde{q}_U(w)$, so that the equilibrium quantity is determined by (3). Given that $MR_D(q)$ is decreasing and $MC_U(q)$ is increasing, it follows that $MR_D(\tilde{q}_D(w)) > MC_U(\tilde{q}_D(w))$, implying $\tilde{q}_D(w) < q_I$. We denote by w_I the wholesale price leading to the limite case where $\tilde{q}_D(w_I) = \tilde{q}_U(w_I)$, the equilibrium quantity satisfies both (3) and (5), implying that $MR_D(\tilde{q}_D(w_I)) = MC_U(\tilde{q}_D(w_I))$ and thus $\tilde{q}_D(w_I) = q_I$.

As illustrated in Figure 2, these two comments lead to a direct relationship between the wholesale price w and the allocation of the right-to-manage. When w is high ($w > w_I$), U wants to sell a quantity greater than what D is willing to purchase to maximize its profit ($\tilde{q}_U(w) > \tilde{q}_D(w)$).[34] Hence, being on the short side of the market, D determines the quantity exchanged in equilibrium (i.e., D has the right-to-manage), and $w(q) = MR_D(q)$. The reverse holds when w is low ($w < w_I$). That is, U prefers to sell a smaller quantity than what D seeks to purchase to maximize

[33] That is, there is no equilibrium situation in which both U and D are simultaneously rationed ($q(w) < \min\{\tilde{q}_U(w), \tilde{q}_D(w)\}$), as they would find profitable to continue trading until one of them reaches its profit-maximizing quantity.

[34] For any given $w < w_I$, D's marginal cost, w, exceeds its marginal revenue at U's profit-maximizing quantity (i.e., $\tilde{q}_U(w)$).

its profit ($\tilde{q}_D(w) < \tilde{q}_U(w)$), resulting in U dictating the equilibrium quantity (i.e., U has the right-to-manage), and $w(q) = MC_U(q)$.[35] Consequently, the firm that sets the equilibrium quantity is endogenously determined, depending on the level of w. This result stands in contrast to prior work in the bilateral monopoly literature, which typically assumes that, for any given w, either U or D unilaterally chooses the quantity to be traded in equilibrium (see Toxvaerd, 2024, for a review).[36]

4.2 Bargaining

We now turn to Stage 1, where U and D bargain over w, anticipating its effect on the quantity determined in Stage 2. Using the (asymmetric) Nash bargaining solution, we derive the equilibrium wholesale price from the following maximization problem:

$$\max_{w} \pi_U(w)^{\alpha} \pi_D(w)^{(1-\alpha)} \tag{6}$$

where $\pi_U(w) = (w - r(q(w)))\, q(w)$ and $\pi_D(w) = (p(q(w)) - w)\, q(w)$. For simplicity and to facilitate correspondence with the graphical illustrations, we rewrite (6) with respect to q and work with the equivalent formulation throughout the remainder of our article:

$$\max_{q} \pi_U(q)^{\alpha} \pi_D(q)^{(1-\alpha)} \tag{7}$$

where $\pi_U(q) = (w(q) - r(q))\, q$ and $\pi_D(q) = (p(q) - w(q))\, q$.

From (3), (5), and Lemma 1, we organize our analysis into three distinct cases: (i) when bargaining leads to bilateral efficiency, characterized by $w_I = MR_D(q_I) = MC_U(q_I)$; (ii) when D determines the quantity traded in Stage 2, corresponding to $w_I < w(q) = MR_D(q)$; and (iii) when U determines the quantity traded in Stage 2, corresponding to $w_I > w(q) = MC_U(q)$.

[35] Again, taking $w < w_I$ as given, U's marginal revenue, w, is lower than its marginal cost at D's profit-maximizing quantity (i.e., $\tilde{q}_D(w)$).

[36] It is worth noting that some articles have addressed this issue by considering that both q and w are jointly determined through bargaining, yielding the vertically integratedoutcome described in Section 3 (e.g., McDonald and Solow, 1981; Manning, 1987; Björnerstedt and Stennek, 2007).

4.2.1 When Bargaining Leads to Bilateral Efficiency

We look for values of α such that the quantity determined in Stage 2 is q_I, implying that $w_I = MC_U(q_I) = MR_D(q_I)$. Solving (7) under this condition yields the following proposition:

Proposition 2 (Bilateral Efficiency) *There exists a unique* $\alpha = \alpha_I \equiv \frac{\pi_U(q_I)}{\pi_U(q_I)+\pi_D(q_I)} = \frac{\varepsilon_p - 1}{\varepsilon_p + \varepsilon_r} \in [0, 1]$ *such that the wholesale price is* $w_I = MC_U(q_I) = MR_D(q_I)$*, leading to the vertically integrated outcome, as the quantity exchanged is* q_I*, the consumer price is* $p(q_I)$*, the input price is* $r(q_I)$*, and* U*'s markdown and* D*'s markup are respectively given by:*

$$
\begin{aligned}
\nu_U &= \frac{MC_U(q_I)}{r(q_I)} = \frac{\varepsilon_r + 1}{\varepsilon_r}, \\
\mu_D &= \frac{p(q_I)}{MR_D(q_I)} = \frac{\varepsilon_p}{\varepsilon_p - 1}.
\end{aligned}
$$

Consequently, U*'s margin is equal to* $M_U = \frac{w_I}{r(q_I)} = \nu_U$*,* D*'s margin is equal to* $M_D = \frac{p(q_I)}{w_I} = \mu_D$*, and the total margin of the supply chain is given by* $\mathcal{M} = \frac{p(q_I)}{r(q_I)} = \nu_U \times \mu_D$*.*

Proof. See Appendix A.2. ∎

There exists a unique $\alpha = \alpha_I$ such that the vertical chain negotiation is bilaterally efficient, replicating the vertically integrated outcome characterized in Proposition 1, i.e., a traded quantity q_I, consumer price p_I, and supplier price r_I. The resulting total margin of the supply chain $\mathcal{M}_I$ is composed of U's markdown and D's markup, respectively equal to the markup μ_I and the markdown ν_I charged by the vertically integrated firm I. Bilateral efficiency arises from the wholesale price w_I satisfying $w_I = MC_U(q_I) = MR_D(q_I)$, ensuring that in Stage 2, U and D are willing to trade the same quantity, i.e., $\tilde{q}_D(w_I) = \tilde{q}_U(w_I)$.

The uniqueness of α_I results from U and D bargaining over a linear tariff. As a consequence, the wholesale price w serves two purposes: determining the quantity traded and sharing the surplus between U and D. Lemma 1 shows that for any $w \neq w_I$, the traded quantity falls below q_I. Thus, any deviation of bargaining power in favor of U (i.e., $\alpha > \alpha_I$) or in favor of D (i.e., $\alpha < \alpha_I$) thus requires distorting the wholesale price, $w \neq w_I$, and consequently adjusting the traded quantity to accommodate profit sharing.

The level of bargaining power leading to bilateral efficiency $\alpha_I = \frac{\varepsilon_p - 1}{\varepsilon_p + \varepsilon_r} \in [0, 1]$ depends on the shape of the input supply and consumer demand. The value of α_I decreases (resp. increases) when the input supply becomes relatively more (less) elastic compared to the consumer demand. It is worth noting that, in the limit case where U has constant marginal costs (i.e., $\varepsilon_r \to \infty$), we obtain the well-established result from the canonical model of vertical contracting, i.e., $\alpha_I = 0$.

In the remainder of Section 4.2.2, we depart from $\alpha = \alpha_I$ and examine two scenarios: U being powerful when $\alpha > \alpha_I$ (Section 4.2.2), and D being powerful when $\alpha < \alpha_I$ (Section 4.2.3).

4.2.2 When Bargaining Leads to Double Markupization

Consider the case in which $w_I < w(q) = MR_D(q)$ such that D determines the equilibrium quantity in Stage 2 ($\tilde{q}_D(w) < \tilde{q}_U(w)$). In this case, the maximization of the Nash product (7) boils down to:

$$\max_q \underbrace{[MR_D(q)q - r(q)q]}_{\pi_U(q)}^{\alpha} \underbrace{[p(q)q - MR_D(q)q]}_{\pi_D(q)}^{(1-\alpha)} \tag{8}$$

for any q such that $w_I < w(q)$. Note that U and D have conflicting interests over q. Assumption 1.(i) and Assumption 1.(ii) imply that, under $w(q) = MR_D(q)$, $\pi_U(q)$ is decreasing whereas $\pi_D(q)$ is increasing in q over the range of conflict relevant for bargaining. This reflects that, at the bargaining stage and in this case, because $w(q)$ decreases in q, U prefers to restrict the traded quantity, whereas D prefers to expand it. To ensure that (8) is well-defined, we introduce the following assumption:

Assumption 2 *D's marginal revenue satisfies the following conditions:*

(i) $\varepsilon_{MR_D} > 1$.

(ii) $\sigma_{MR_D} < 2$.

Assumption 2.(i), which can equivalently be expressed as $\varepsilon_p > 3 - \sigma_p$, imposes that consumer demand is supermodular (e.g., Mrázová and Neary, 2017).[37] This condition ensures that $MR_U(q) > 0$, which in turn guarantees that (9) can be satisfied for any $\alpha \in [0, 1]$.[38] Assumption 2.(ii) ensures that the second-order condition of (8) is satisfied (see Appendix A.3.1 for further details). Given these assumptions, the first-order condition associated with (8) is:

$$\alpha[MR_U(q_\mu) - MC_U(q_\mu)]\pi_D(q_\mu) + (1-\alpha)[MR_D(q_\mu - MR_U(q_\mu)]\pi_U(q_\mu) = 0 \quad (9)$$

where q_μ denotes the equilibrium traded quantity and $MR_U(q) \equiv \frac{\partial MR_D(q)q}{\partial q} = MR'_D(q)q + MR_D(q)$ corresponds to U's marginal revenue function when it makes a take-it-or-leave-it offer to D (i.e., U faces a demand curve given by $MR_D(q)$).

To gain further insight into the equilibrium outcome, we rearrange (9) as follows:

$$MC_U(q_\mu) = \widetilde{MR}_U(q_\mu, \alpha) \quad (10)$$

where $\widetilde{MR}_U(q_\mu, \alpha) \equiv \beta_D(q_\mu, \alpha)MR_D(q_\mu) + (1 - \beta_D(q_\mu, \alpha))\, MR_U(q_\mu)$ can be interpreted as a "shadow" marginal revenue, with $\beta_D(q_\mu, \alpha) \equiv \frac{1-\alpha}{\alpha}\frac{\pi_U(q_\mu)}{\pi_D(q_\mu)}$ (see Appendix A.3.2 for details). When U has all the bargaining power, we denote $q_{\overline{\mu}}$ the corresponding equilibrium quantity. In that case, we have $\beta_D(q_{\overline{\mu}}, 1) = 0$, so that (10) boils down to $MC_U(q_{\overline{\mu}}) = MR_U(q_{\overline{\mu}})$. This corresponds to the canonical model of vertical contracting, where the inefficient outcome $q_{\overline{\mu}} < q_I$ arises because $w_{\overline{\mu}} > w_I$ (see Appendix B.2.2 for further details). When $1 > \alpha > \alpha_I$, we have $\beta_D(q_\mu, \alpha) > 0$, which shifts $\widetilde{MR}_U(q_\mu, \alpha)$ towards $MR_D(q_\mu)$, such that $MR_D(q_\mu) > \widetilde{MR}_U(q_\mu, \alpha) > MR_U(q_\mu)$. As $MC_U(q)$ increases in q, the equilibrium quantity q_μ characterized by (10) increases, thereby reducing the inefficiency. Finally, when α tends to α_I, we have $\beta_D(q_\mu, \alpha_I) = 1$, implying that (10) reduces to $MC_U(q_\mu) = MR_D(q_\mu)$. This corresponds to the vertically integrated outcome, where $q_\mu = q_I$. Based on this reasoning, we derive the following proposition:

[37] Supermodular demand functions include, among others, the CES, translog, and AIDS demand models. Supermodularity also holds for linear demand when $\epsilon_p > 3$, and in the logit demand model for sufficiently low values of q^*.

[38] Formally, $MR_U > 0 \Leftrightarrow \varepsilon_{MR_D} = \frac{\varepsilon_p - 1}{2-\sigma_p} > 1 \Leftrightarrow \varepsilon_p > 3 - \sigma_p$.

Proposition 3 (Double markupization) *When U is powerful, i.e., $\alpha_I < \alpha \leq 1$, the wholesale price is $w_\mu = MR_D(q_\mu) > w_I$, the quantity exchanged is $q_\mu < q_I$, the consumer price is $p(q_\mu) > p(q_I)$, and the input price is $r(q_\mu) < r(q_I)$. Double marginalization arises from U's seller power as U charges a markup, given by:*

$$\mu_U = \frac{w_\mu}{MC_U(q_\mu)} = \frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D} - (1 - \beta_D(q_\mu, \alpha))} = \frac{\alpha\varepsilon_{MR_D}(\varepsilon_r + 1) + (1-\alpha)(\varepsilon_p - 1)\varepsilon_r}{(\varepsilon_r + 1)(\alpha(\varepsilon_{MR_D} - 1) + (1-\alpha)(\varepsilon_p - 1))},$$

which adds up to D's markup, given by $\mu_D = \frac{p(q_\mu)}{MR_D(q_\mu)} = \frac{\varepsilon_p}{\varepsilon_p - 1}$. U's markdown is equal to $\nu_U = \frac{MC_U(q_\mu)}{r(q_\mu)} = \frac{\varepsilon_r + 1}{\varepsilon_r}$, whereas D does not charge any markdown, i.e. $\nu_D = \frac{MR_D(q_\mu)}{w_\mu} = 1$. Consequently, U's margin is equal to $M_U = \frac{w_\mu}{r(q_\mu)} = \nu_U \times \mu_U$, D's margin is equal to $M_D = \frac{p(q_\mu)}{w_\mu} = \mu_D$, and the total margin of the supply chain is given by:

$$\mathcal{M} = \frac{p(q_\mu)}{r(q_\mu)} = \nu_U \times \mu_U \times \mu_D.$$

Proof. Appendix A.7.2 provides formal derivations of the markup and markdown expressions based on the definitions introduced in Section 2. Appendix A.3.3 presents the main analytical derivations, and Appendix A.3.4 characterizes the set of equilibria. ■

When U is powerful ($\alpha_I < \alpha \leq 1$), the equilibrium wholesale price is given by $w_\mu > w_I$. In this case, the short-side rule that emerges in Stage 2 implies that w_μ and q_μ co-move along D's demand, which is given by MR_D. As MR_D decreases with q, Proposition 3 establishes that U exercises monopoly power by charging a markup over its marginal cost when selling to D. This markup adds up to D's markup due to its monopoly power in the product market. The resulting *double markup* gives rise to the classical double marginalization phenomenon (Spengler, 1950), leading to an inefficient outcome ($q_\mu < q_I$).

Analogous to the vertically integrated outcome, U's markdown (ν_U) and D's markup (μ_D) are determined by the elasticities of supply and demand, respectively, reflecting U's monopsony power in the input market and D's monopoly power in the product market. Interestingly, U's markup (μ_U) depends on two factors. The first

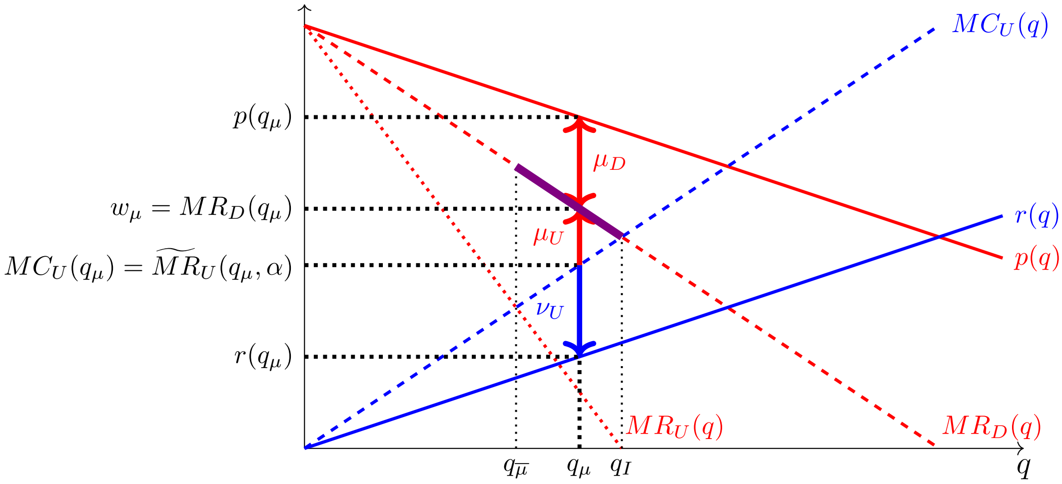


Figure 3: Equilibrium with Double Markup ($\alpha_I < \alpha \leq 1$).

is D's demand elasticity (ε_{MR_D}), reflecting U's monopoly power. The second is β_D, which captures D's ability to exert countervailing buyer power, with μ_U decreasing as β_D increases. Finally, D charges no markdown as $\nu_D = 1$.

Figure 3 illustrates the equilibrium described in Proposition 3 using linear demand and supply functions for a given $\alpha \in [\alpha_I, 1]$. In equilibrium, w_μ lies on MR_D and satisfies the condition $MC_U(q_\mu) = \widetilde{MR}_U(q_\mu, \alpha)$, which determines q_μ. The set of equilibrium wholesale prices and quantities is represented by the purple segment.

4.2.3 When Bargaining Leads to Double Markdownization

Consider now the case in which $w_I > w(q) = MC_U(q)$ such that U determines the equilibrium quantity in Stage 2 ($\tilde{q}_U(w) < \tilde{q}_D(w)$). In this case, (7) boils down to:

$$\max_q \underbrace{[MC_U(q)q - r(q)q]}_{\pi_U(q)}^{\alpha} \underbrace{[p(q)q - MC_U(q)q]}_{\pi_D(q)}^{(1-\alpha)} \tag{11}$$

for any q such that $w_I > w(q)$. Note that U and D have conflicting interests over q. Indeed, Assumption 1.(i) and Assumption 1.(ii) ensure that, under $w(q) = MC_U(q)$, $\pi_U(q)$ increases whereas $\pi_D(q)$ decreases in q over the range of conflict relevant for bargaining. This reflects that, at the bargaining stage and in this case, because $w(q)$ increases in q, U prefers to expand the traded quantity, whereas D prefers to reduce it. To ensure that the second-order condition of (11) is satisfied, we introduce the following assumption (see Appendix A.4.1 for further details):

Assumption 3 *U's marginal cost satisfies $\sigma_{MC_U} > -2$.*

Given Assumption 3, the first-order condition associated with (11) is:

$$\alpha(MC_D(q_\nu) - MC_U(q_\nu))\pi_D(q_\nu) + (1-\alpha)(MR_D(q_\nu) - MC_D(q_\nu))\pi_U(q_\nu) = 0 \quad (12)$$

where q_ν is the equilibrium traded quantity and $MC_D(q) \equiv \frac{\partial MC_U(q)q}{\partial q} = \underbrace{MC_U'(q)q}_{>0} + MC_U(q)$ corresponds to D's marginal cost function when it makes a take-it-or-leave-it offer to U (i.e., D faces a supply curve given by $MC_U(q)$).

To gain further insight into the equilibrium outcome, we rearrange (12) as follows:

$$MR_D(q_\nu) = \widetilde{MC}_D(q_\nu, \alpha) \quad (13)$$

where $\widetilde{MC}_D(q_\nu, \alpha) \equiv \beta_U(q_\nu, \alpha)MC_U(q_\nu) + (1 - \beta_U(q_\nu, \alpha))\, MC_D(q_\nu, \alpha)$ can be interpreted as a "shadow" marginal cost, with $\beta_U(q_\nu, \alpha) \equiv \frac{\alpha}{1-\alpha}\frac{\pi_D(q_\nu)}{\pi_U(q_\nu)}$ (see Appendix A.3.2 for details). When D has all the bargaining power, we have $\beta_U(q_\nu, \alpha) = 0$, implying that (13) boils down to $MC_D(q_\nu) = MR_D(q_\nu)$. The inefficient outcome $q_\nu < q_I$ arises because $w^* < w_I$ (see Appendix B.2.3 for further details). When $0 < \alpha < \alpha_I$, we have $\beta_U(q^*, \alpha) > 0$, which shifts $\widetilde{MC}_D(q_\nu, \alpha)$ towards $MC_U(q_\nu)$ such that $MC_D(q_\nu) > \widetilde{MC}_D(q_\nu, \alpha) > MC_U(q_\nu)$. As $MR_D(q)$ decreases in q, the equilibrium quantity q^* characterized by (13) increases, thereby reducing the inefficiency. Finally, when α tends to α_I, we have $\beta_U(q_\nu, \alpha_I) = 1$, implying that (13) reduces to $MC_U(q_\nu) = MR_D(q_\nu)$. This corresponds to the vertically integrated outcome, where $q_\nu = q_I$. Based on this reasoning, we derive the following proposition:

Proposition 4 (Double Markdownization) *When D is powerful, i.e., $0 \leq \alpha < \alpha_I$, the wholesale price is $w_\nu = MC_U(q_\nu) < w_I$, the quantity exchanged is $q_\nu < q_I$, the consumer price is $p(q_\nu) > p(q_I)$, and the input price is $r(q_\nu) < r(q_I)$. Double marginalization arises from D's buyer power as D charges a markdown, given by:*

$$\nu_D = \frac{MR_D(q_\nu)}{w(q_\nu)} = \frac{\varepsilon_{MC_U} + (1 - \beta_U(q_\nu, \alpha))}{\varepsilon_{MC_U}} = \frac{(\varepsilon_p - 1)(\alpha(\varepsilon_r + 1) + (1-\alpha)(\varepsilon_{MC_U} + 1))}{\alpha\varepsilon_p(\varepsilon_r + 1) + (1-\alpha)(\varepsilon_p - 1)\varepsilon_{MC_U}},$$

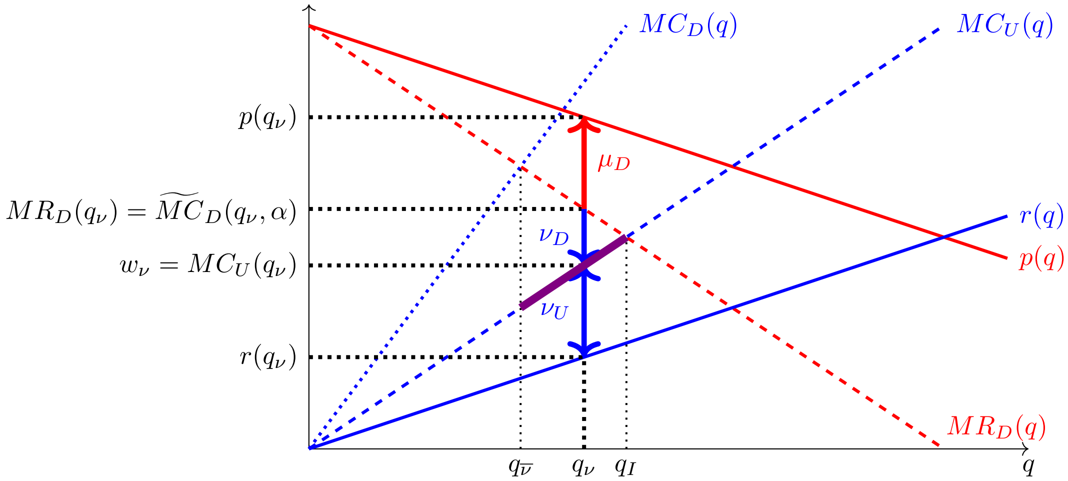


Figure 4: Equilibrium with Double Markdown ($0 < \alpha < \alpha_I$).

which adds up to D's markdown, given by $\nu_U = \frac{MC_U(q_\nu)}{r(q_\nu)} = \frac{\epsilon_r+1}{\epsilon_r}$. *D's markup is equal to* $\mu_D = \frac{p(q_\nu)}{MR_D(q_\nu)} = \frac{\varepsilon_p}{\varepsilon_p - 1}$, *whereas U does not charge any markup, i.e.* $\mu_U = \frac{w(q_\nu)}{MC_U(q_\nu)} = 1$. *Consequently, U's margin is equal to* $M_U = \frac{w_\nu}{r(q_\nu)} = \nu_U$, *D's margin is equal to* $M_D = \frac{p(q_\nu)}{w_\nu} = \nu_D \times \mu_D$, *and the total margin of the supply chain is given by:*

$$\mathcal{M} = \frac{p(q_\nu)}{r(q_\nu)} = \nu_U \times \nu_D \times \mu_D.$$

Proof. Appendix A.7.2 provides formal derivations of the markup and markdown expressions based on the definitions introduced in Section 2. Appendix A.4.3 presents the main analytical derivations, and Appendix A.4.4 characterizes the set of equilibria. ■

When D is powerful ($0 \leq \alpha \leq \alpha_I$), the equilibrium wholesale price is given by $w_\nu < w_I$. In this case, the short-side rule emerging in Stage 2 implies that w_ν and q_ν co-move along U's supply, which is given by MC_U. As MC_U increases with q, Proposition 4 establishes that D exercises monopsony power by charging a markdown below its marginal revenue when purchasing from U. This markdown adds up to U's markdown due to its monopsony power in the input market. The resulting *double markdownization* leads to an inefficient outcome ($q_\nu < q_I$) and corresponds to a novel source of double marginalization.

Again, U's markdown (ν_U) and D's markup (μ_D) are shaped by the elasticities of supply and demand, respectively, reflecting U's monopsony power in the input market and D's monopoly power in the product market. Moreover, D's markdown ν_D depends on two factors. The first is U's supply elasticity (ε_{MC_U}), reflecting D's monopsony power. The second is β_U, which captures U's ability to exert countervailing seller power, with ν_D decreasing as β_U increases. Finally, U charges no markup as $\mu_U = 1$.

Figure 4 illustrates the equilibrium described in Proposition 4 using linear demand and supply functions for a given $\alpha \in [\alpha_I, 1]$. In equilibrium, w_ν lies on MC_U and satisfies the condition $MR_D(q_\nu) = \widetilde{MC}_U(q_\nu, \alpha)$, which determines q_ν. The set of equilibrium wholesale prices and quantities is represented by the purple segment.

5 Welfare Effects of Buyer and Seller Power

We now analyze the effect of a change in the distribution of bargaining power on equilibrium outcomes. Specifically, we consider changes in the bargaining weight α (as in, e.g., Chen, 2003; Gaudin, 2018).[39] We formalize the effects of such variations, illustrated in Figure 5 where the purple (resp. green) segment represents the set of equilibria when α goes from 0 (resp. 1) to α_I, with the arrows indicating the direction of the variation, in the following corollary:

Corollary 1 *Welfare increases when α moves toward α_I. In particular:*

- *When U is powerful ($\alpha_I < \alpha \leq 1$), a decrease in α, i.e. an increase in D's bargaining power, countervails U's seller power: both U's markup μ_U and the supply chain margin $\mathcal{M} = \nu_U \times \mu_U \times \mu_D$ decrease, increasing the quantity exchanged q_μ and welfare.*

- *When D is powerful ($0 \leq \alpha < \alpha_I$), an increase in α, i.e. an increase in U's bargaining power, countervails D's buyer power: both D's markdown ν_D and*

[39] Shifts in α are exogenous changes in the distribution of bargaining power in the vertical supply chain. More broadly, changes in the distribution of bargaining power can arise from various sources, including changes in market structure (e.g., consolidation, entry, or exit) or firms' strategies (e.g., forming a buying alliance). Modeling these endogenous sources of changes in the distribution of bargaining power would require a model of vertical relations with competition at (at least) one level of the supply chain, which we leave as an avenue for future research.

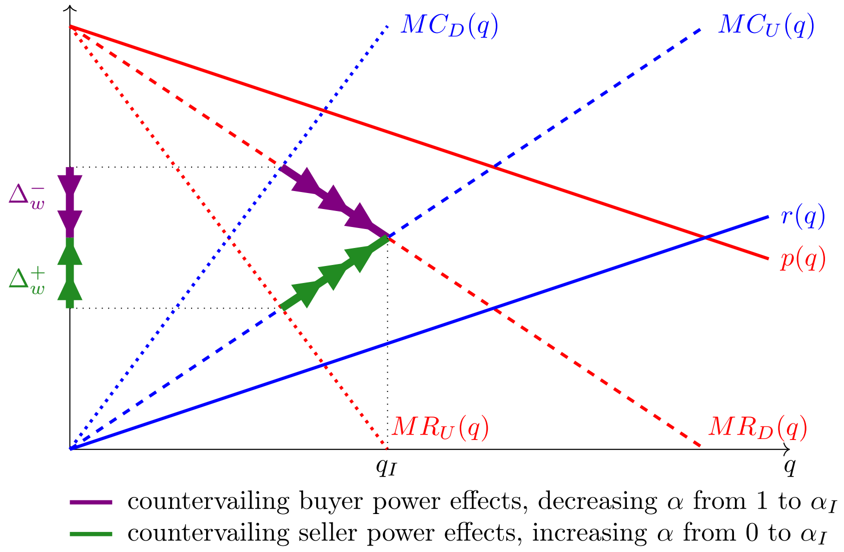


Figure 5: Effects of Countervailing Buyer and Seller Power.

the supply chain margin $\mathcal{M} = \nu_U \times \nu_D \times \mu_D$ decrease, increasing the quantity exchanged q_ν and welfare.

- *When $\alpha = \alpha_I$, welfare is maximized as U's seller power and D's buyer power fully countervails each other, i.e. $\mu_U = \nu_D = 1$, and the supply chain margin $\mathcal{M} = \nu_U \times \mu_D$ reaches its vertical integration value M_I.*

The countervailing buyer power effect emerging when U is powerful has been extensively discussed in the literature (see, e.g., Snyder, 2008). It refers to the welfare-improving effect of increasing D's bargaining power (lower α), which mitigates double marginalization by reducing U's markup and ultimately the total margin. Corollary 1 also sheds light on a novel mechanism that arises when D is powerful. In this case, although U charges no markup, D exercises monopsony power by charging a markdown (Proposition 4). Consequently, further increases in D's bargaining power exacerbate double marginalization by raising D's markdown. The countervailing seller power theory thus applies: increasing U's bargaining power offsets D's monopsony distortion, reducing D's markdown and ultimately total margin. Overall, when $\alpha = \alpha_I$, U's seller power and D's buyer power fully countervails each other. Specifically, U's countervailing seller power prevents D from exerting any markdown ($\nu_D = 1$) while D's countervailing buyer power prevents U from exerting any markup ($\mu_U = 1$). As a result, the supply chain reaches the vertical integration outcome, guaranteeing bilateral efficiency and maximizing welfare.

The results in Corollary 1 suggest that the welfare effects of buyer and seller power depend critically on the nature of the distortion that arises in equilibrium (i.e., double markup or markdown). To gain further insights on how countervailing buyer or seller power affects the different components of total welfare, we establish the following corollary:

Corollary 2 *When α moves toward α_I, distributional welfare effects are as follows:*

- *When U is powerful ($\alpha_I < \alpha \leq 1$), a decrease in α, i.e., an increase in D's bargaining power, raises consumers' and input suppliers' surpluses, and D's profit, but decreases U's profit.*

- *When D is powerful ($0 \leq \alpha < \alpha_I$), an increase in α, i.e., an increase in U's bargaining power, raises consumers' and input suppliers' surpluses, and U's profit, but decreases D's profit.*

Proof. See Appendix A.5. ■

Interestingly, Corollary 2 uncovers an additional beneficial mechanism of countervailing buyer power: by increasing the quantity exchanged when U is powerful ($\alpha_I < \alpha \leq 1$), an increase in D's bargaining power also raises input suppliers' surplus. As the countervailing buyer power theory has typically been formalized in settings with constant marginal costs, this latter effect has not been previously identified in the literature.[40] Moreover, Corollary 2 establishes the notion of countervailing seller power, emerging when D is powerful ($0 \leq \alpha < \alpha_I$). In the symmetric case illustrated in Figure 5, resulting in $\alpha_I = \frac{1}{2}$, countervailing seller and buyer power are a priori (for an unknown α) equally relevant. More generally, the emergence of one or the other effect depends on the distribution of bargaining power and the relative supply and demand elasticity. For instance, when supply is significantly less elastic than demand,

[40]The input supply function reflects the aggregation of heterogeneous individual supply functions, and when the quantity traded increases, the gain in input suppliers' surplus may as well imply larger quantities sold by input suppliers already active (intensive margin), as new suppliers being involved in the production process (extensive margin).

resulting in a high α_I, countervailing seller power may prevail even for high values of α. Finally, Corollary 2 highlights that countervailing buyer or seller power hurts the relatively powerful firm in the industry while benefiting all other parties.

Note that changes in α also indirectly affect the markup of D and the markdown of U, through the change in quantity. These indirect effects crucially depend on the characteristics of supply and demand functions. These indirect effects may reinforce or dampen the distortion, but always remain second-order: they cannot reverse the impact on welfare. We establish the following remark:

Remark 1 *The positive welfare effects of α moving toward α_I, as shown in Corollary 1, are smaller (resp. larger) when the demand and supply functions are subconvex (resp. superconvex), since D's markup μ_D increases (resp. decreases) with q when demand is subconvex (resp. superconvex), and similarly, U's markdown ν_U increases (resp. decreases) with q when supply is subconvex (resp. superconvex).*

Proof. See Appendix A.6. ■

To illustrate Remark 1, consider the case where U is powerful and α decreases, raising D's countervailing buyer power, and hereby, mitigating U's markup and the double marginalization phenomenon described in Proposition 3.[41] The extent to which the resulting decrease in the wholesale price w is passed on upstream to input suppliers and downstream to consumers ultimately depends on the shape of demand and supply. If demand is subconvex ($\frac{\partial \varepsilon_p}{\partial q} < 0$), D's markup (here also equal to its margin) increases, indicating incomplete pass-through to the consumer price. If supply is likewise subconvex ($\frac{\partial \varepsilon_r}{\partial q} < 0$), U's markdown also increases.[42] However, we find that the decrease in U's markup dominates, resulting in a decrease in both its margin and the total margin in the supply chain. Such a reduction in the distortion, and associated welfare effects,

[41] The reasoning is symmetric when D is powerful and α increases, mitigating D's markdown.

[42] The term "subconvex" demand, introduced by Mrázová and Neary (2019), refers to demand functions that are less convex than the CES demand (see also Mrázová and Neary, 2017). It is also called "Marshall's Second Law of Demand" because it captures the idea that consumers become more price-elastic at higher prices, a property most demand systems satisfy. Although supply subconvexity has received less attention in the literature, it is consistent with recent empirical evidence, such as Boehm and Pandalai-Nayar (2022) for U.S. industries and Avignon and Guigue (2022) for the French milk industry.

are even greater under superconvex demand and supply (i.e., when $\frac{\partial \varepsilon_p}{\partial q} > 0$ and $\frac{\partial \varepsilon_r}{\partial q} > 0$, respectively), as both U's markdown and D's markup decrease when α decreases.

6 Discussion

6.1 Quantity Setting: Voluntary Exchange and Right-to-Manage endogeneization

Our timing assumes that in Stage 1, firms bargain over a linear tariffs. The traded quantity is determined in Stage 2 assuming voluntary exchange: neither U nor D can be forced to trade more than it is willing to. In what follows, we discuss the implications of this assumption compared to exogenous right-to-manage frameworks.

In industrial organization and labor economics, bargaining frameworks commonly assume that either the buyer or the seller unilaterally determines the traded quantity for any given negotiated wholesale price. We refer to this as the exogenous right-to-manage (RTM) assumption. In what follows, we highlight three notable implications of adopting this alternative assumption within our framework, considering that in Stage 2, either U or D unilaterally sets the quantity.

First, welfare increases with the bargaining power of the firm that holds the RTM. Second, this result stems from a violation of the voluntary exchange condition when the powerful firm holds the RTM—specifically, when D holds the RTM and $0 \leq \alpha < \alpha_I$, or when U holds the RTM and $\alpha_I < \alpha \leq 1$. Third, the powerful firm holding the RTM faces a commitment issue: it ends up choosing a quantity that is too high, both from a unilateral and a bilateral efficiency perspective.

To see this, consider the case where the RTM is granted to D for any value of $\alpha \in [0, 1]$.[43] In Stage 2 the quantity q^* is determined by D's first-order condition $w = MR_D(q^*)$. As a result, there is a negative relation between the negotiated wholesale price w^* and the equilibrium quantity q^*. Hence, an increase in D's buyer power (i.e., a lower α) leads to a lower wholesale price and a higher quantity. When D is powerful, that is when $0 \leq \alpha < \alpha_I$ the negotiated wholesale price is $w^* < w_I$, and

[43] A symmetric reasoning applies when U holds the RTM.

the equilibrium quantity is given by $w^* = MR_D(q^*) < MC_U(q^*)$. This implies that, in q^*, U's profit is decreasing, violating voluntary exchange and leading to a traded quantity above the bilaterally efficient level: $q^* > q_I$. In the extreme case where D holds all the bargaining power ($\alpha = 0$), the equilibrium quantity is determined by $w^* = MR_D(q^*) = r(q^*)$ which maximizes welfare as $q_I < q^* < q_W$. In that case, U earns zero profit, as D's buyer power not only fully countervails U's seller power but also its monopsony power in its input market. Effectively, U has no market power: it acts as a price taker in both its output market (where D sets w^*) and its input market (where the price r adjusts so that supply matches $D's$ quantity choice q^*).[44] Moreover, when powerful, D faces a commitment issue that leads it to deteriorate the bilateral profit, hereby improving welfare. Specifically, D must lower the wholesale price w^* below w_I in Stage 1 to capture more of the bilateral surplus. However, this reduction in w^* also diminishes the bilateral surplus itself. In Stage 2, D's profit-maximizing behavior implies $w^* = MR_D(q^*)$, which drives the traded quantity above the bilaterally efficient level q_I as w^* declines.[45] Thus, holding full quantity-setting power ultimately harms D, to the extent that D would be better off if able to commit to trading a quantity $q^* \leq q_I$. Such a commitment would support an outcome where the efficient quantity $q^* = q_I$ is traded at a wholesale price $w^* = r(q_I) < MR_D(q_I)$ when $\alpha = 0$.

The commitment issue makes it unlikely that the RTM would be contractually allocated to the powerful firm in Stage 1. Indeed, for any $\alpha \in (0, 1)$, both firms would benefit from an alternative clause governing quantity determination, such as fixing the quantity at $q = q_I$ or assigning RTM to one firm under the constraint that $q \leq q_I$.

A possible rationale for the RTM being allocated to the powerful firm is that this firm possesses sufficient strength in Stage 2 to unilaterally impose the traded quantity on the other party. In line with this idea, we discuss connections with the following game inspired by Manning (1987). Firms engage in a sequential Nash bargaining process in which, in Stage 1, they agree on a price, and in Stage 2, the quantities

[44] Intuitively, by being forced to trade more than it is willing to, U loses its ability to exert monopsony power in Stage 2. U's increasing marginal cost yet continues to play a role in Stage 1 if $\alpha > 0$.

[45] Note that when D is powerful under voluntary exchange, such an efficiency-extraction tradeoff for D arises from U's quantity choice given w, which drives the quantity below q_I.

exchanged are negotiated.[46] Let γ denote the bargaining power of U at the quantity-setting stage, allowed to differ from α. Manning (1987) shows that if $\alpha = \gamma$, the bilateral relationship delivers the vertically integrated outcome, i.e., $q^* = q_I$. Limit cases $\gamma = 1$ and $\gamma = 0$ correspond to cases discussed above where the RTM is exogenously allocated to U or D, respectively. Our framework with voluntary exchange fundamentally differs in nature from this sequential bargaining approach. Indeed, under voluntary exchange, a firm cannot condition the trade of one unit on the trade of other units.[47] Our results would coincide with sequential bargaining framework only when $\gamma = 1$ for $\alpha \leq \alpha_I$ or $\gamma = 0$ for $\alpha > \alpha_I$. This stresses again that in our framework, the quantity-setting power γ is not exogenous but contingent on the model primitives, namely α, and demand and supply parameters.

6.2 Two-part Tariff Contract

In the absence of contractual frictions, it is well-known that a two-part tariff contract suffices to eliminate the double-marginalization problem and restore efficiency (e.g., Mathewson and Winter, 1984).[48] However, double marginalization may persist in settings where financial or contractual frictions prevail (e.g., Rey and Tirole, 1986; Bernheim and Whinston, 1998; Nocke and Thanassoulis, 2014; Calzolari, Denicolò and Zanchettin, 2020).

In this section, we demonstrate that our findings remain valid even if firms bargain over a two-part tariff contract (w, F), provided there exists frictions limiting the use of the fixed fee to transfer surplus between firms (i.e., utility is not perfectly transferable). As in Calzolari, Denicolò and Zanchettin (2020), we remain fairly agnostic about the precise source of friction and discuss potential microfoundations at the end of the

[46] Manning (1987) analyzes sequential bargaining between unions and firms, yet abstracting from monopsony power consideration. We follow Manning (1987) in referring to such a game as sequential bargaining, although he notes that such a term sometimes describes "a single bargain which has a sequence of offers and counter-offers (as in Rubinstein (1982))".

[47] To illustrate, consider for instance a case in which D orders a quantity to U and where the wholesale price is set for a period during which production and demand unfold over time, and goods are only partially storable. Thus, D cannot condition the first unit on the subsequent one. It is also the case when the production takes time and U has to set its capacity before D passes the order.

[48] In this case, the wholesale price is efficiently set at U's marginal cost, the quantity traded in equilibrium is q_I, and D's markup is given by μ_I, and U's markdown is given by ν_I.

section.

In a setting with a two-part tariff contract, the profit functions for U and D are defined, respectively, as: $\Pi_U(q) \equiv w(q)q - r(q)q + F = \pi_U(q) + F$ and $\Pi_D(q) \equiv p(q)q - w(q)q - F = \pi_D(q) - F$. We allow F to be either positive (transfer from D to U) or negative (transfer from U to D), but impose $\underline{F} \leq F \leq \overline{F}$. Specifically, when $F > 0$, we assume that the fixed fee D pays to U cannot exceed $\overline{F}$. Similarly, when $F < 0$, we assume that the fixed fee D receives from U cannot exceed $\underline{F}$.[49]

The sequence of play mirrors that described in Section 4. In Stage 1, U and D negotiate a two-part tariff contract (w, F); in Stage 2, given w, firms simultaneously announce the quantities they are willing to trade, and the short-side rule determines the quantity traded. As the fixed fee F never affects firms' quantity choice, the resolution of stage 2 is similar to that in Section 4.1. Specifically, given w, the quantity traded in equilibrium is $q(w) = \min\{\tilde{q}_U(w), \tilde{q}_D(w)\}$, as established in Lemma 1, continues to apply. In Stage 1, U and D bargain over (w, F) anticipating the effect of w on the quantity determined in Stage 2. We determine the equilibrium two-part tariff by solving the (asymmetric) Nash bargaining solution:[50]

$$\max_{q,F} \Pi_U(q, F)^{\alpha} \Pi_D(q, F)^{1-\alpha} \quad \text{subject to} \quad \underline{F} \leq F \leq \overline{F}$$

This yields the following first-order conditions on F:

$$\alpha \Pi_D(q, F) - (1 - \alpha)\Pi_U(q, F) = 0. \tag{14}$$

[49] Our modeling assumptions reflect a situation where transferring surplus between firms through F is costless as long as $\underline{F} \leq F \leq \overline{F}$, and becomes infinitely costly otherwise. Calzolari, Denicolò and Zanchettin (2020) adopt an alternative approach where U receives F when D pays $(1 + \mu)F$ (with $\mu \geq 0$), implying that the use of F creates deadweight losses. Under this alternative approach, it is worth noting that we would obtain similar results by assuming that the cost of transferring surplus, denoted by $\mu(F)$, is increasing and weakly convex in F.

[50] As in (7), we maximize the Nash product with respect to (q, F) considering in turn that $w(q) = MR_D(q)$ and $w(q) = MC_U(q)$, is equivalent to maximizing with respect to (w, F).

This enables simplifying the first order conditions on q as follows:

$$[MR_U(q) - MC_U(q)] + [MR_D(q) - MR_U(q)] = 0, \quad \text{when } w(q) = MR_D(q) \tag{15}$$

$$[MC_D(q) - MC_U(q)] + [MR_D(q) - MC_U(q)] = 0, \quad \text{when } w(q) = MC_U(q) \tag{16}$$

which characterize the equilibrium quantity $\hat{q}$, wholesale price $\hat{w} = w(\hat{q})$, and fixed fee $\hat{F}$. Three types of equilibria can arise, depending on whether: (i) the constraint on F is not binding, (ii) the lower bound is binding ($F = \underline{F}$), or (iii) the upper bound is binding ($F = \overline{F}$).

Consider first the case in which $\overline{F}$ and $\underline{F}$ are such that the constraint on F never binds. Simplifying (15) and (16), we obtain $MC_U(\hat{q}) = MR_D(\hat{q})$, implying that $\hat{q} = q_I$. As a result, for any $\alpha \in [0, 1]$, the two-part tariff contract eliminates the double-marginalization problem and restores the vertically integrated outcome described in Section 3. At $\hat{q} = q_I$, (14) yields $\hat{F} = \alpha\pi_D(q_I) - (1-\alpha)\pi_U(q_I)$. Consequently, we have $\hat{F} = 0$ if $\alpha = \alpha_I$, $\hat{F} < 0$ if $\alpha < \alpha_I$, and $\hat{F} > 0$ if $\alpha > \alpha_I$. Importantly, this shows that whenever $\underline{F} < -\pi_U(q_I)$ and $\overline{F} > \pi_D(q_I)$, the constraint on F does not play any role, and this efficient outcome constitutes the unique equilibrium under a two-part tariff contract.

Consider now that the upper and lower bounds are sufficiently tight ($\overline{F} < \pi_D(q_I)$ and $\underline{F} > -\pi_U(q_I)$) so that the constraint on F may affect the equilibrium outcome. In this case, when $\alpha > \alpha_I$ (i.e., $F > 0$), there exists a threshold $\overline{\alpha} \equiv \frac{\pi_U(q_I)+\overline{F}}{\pi_U(q_I)+\pi_D(q_I)} < 1$ such that the fixed fee D pays to U is capped at $\overline{F}$. Similarly, when $\alpha < \alpha_I$ (i.e., $F < 0$), there exists a threshold $\underline{\alpha} \equiv \frac{\pi_U(q_I)+\underline{F}}{\pi_U(q_I)+\pi_D(q_I)} > 0$ such that the fixed fee D receives from U is bounded below by $\underline{F}$.[51] Therefore, when $\overline{\alpha} > \alpha > \underline{\alpha}$, the efficient outcome described when the constraint on F never binds arises in equilibrium: $\hat{q} = q_I$ and $\hat{F} = \alpha\pi_D(q_I) - (1-\alpha)\pi_U(q_I)$. When instead $\alpha_I \leq \overline{\alpha} \leq \alpha$, we have $\hat{F} = \overline{F}$. As $\alpha_I \leq \alpha$, $w(q) = MR_D(q)$, and the first order condition boils down to:

$$MC_U(\hat{q}) \;=\; \widehat{MR}_U(\hat{q}, \hat{F}, \alpha) \tag{17}$$

[51] Note that when $\underline{F} = \overline{F} = 0$, firms are unable to use the fixed fee F to transfer surplus. Hence, this case reduces to the linear wholesale contract setting analyzed in Section 4.

where $\widehat{MR}_U(\hat{q}, \hat{F}, \alpha) \equiv \hat{\beta}_D(\hat{q}, \hat{F}, \alpha)MR_D(\hat{q}) + (1 - \hat{\beta}_D(\hat{q}, \hat{F}, \alpha))MR_U(q\hat{q})$ can be interpreted as a "shadow" marginal revenue, with $\hat{\beta}_D(\hat{q}, \hat{F}, \alpha) \equiv \frac{1-\alpha}{\alpha}\frac{\Pi_U(\hat{q},\hat{F})}{\Pi_D(\hat{q},\hat{F})}$. Interestingly, (17) mirrors (10), which yields the double markup outcome under the linear wholesale contract setting. As $\hat{\beta}_D(\hat{q}, \hat{F}, \alpha) \geq \beta_D(q_\mu, \alpha)$, we have $MR_D(\hat{q}) \geq \widehat{MR}_U(\hat{q}, \hat{F}, \alpha) \geq \widetilde{MR}_U(q_\mu, \alpha)$, implying that $q_I \geq \hat{q} \geq q_\mu$.[52] Conversely, when $\alpha_I \geq \underline{\alpha} \geq \alpha$, we have $\hat{F} = \underline{F}$. As $\alpha_I \geq \alpha$, $w(q) = MC_U(q)$, and the first order condition boils down to:

$$MR_D(\hat{q}) = \widehat{MC}_D(\hat{q}, \hat{F}, \alpha) \tag{18}$$

where $\widehat{MC}_D(\hat{q}, \hat{F}, \alpha) \equiv \hat{\beta}_U(\hat{q}, \hat{F}, \alpha)MC_U(\hat{q}) + (1 - \hat{\beta}_U(\hat{q}, \hat{F}, \alpha))MC_D(\hat{q}, \alpha)$ can be interpreted as a "shadow" marginal cost, with $\hat{\beta}_U(\hat{q}, \hat{F}, \alpha) \equiv \frac{\alpha}{1-\alpha}\frac{\Pi_D(\hat{q},\hat{F})}{\Pi_U(\hat{q},\hat{F})}$. Again, (18) reflects (13), which characterizes the double markup outcome in the linear wholesale contract setting. As $\hat{\beta}_U(\hat{q}, \hat{F}, \alpha) \geq \beta_U(q_\nu, \alpha)$, we have $MC_U(\hat{q}) \geq \widehat{MC}_D(\hat{q}, \hat{F}, \alpha) \geq \widetilde{MC}_D(q_\nu, \alpha)$, implying that $q_I \geq \hat{q} \geq q_\nu$.[53] The following proposition summarizes our results:

Proposition 5 (Two-Part Tariffs under Frictions) *When $\underline{F} > -\pi_U(q_I)$ and $\overline{F} < \pi_D(q_I)$, frictions constraining the fixed fee prevail and the set of equilibria with two-part tariff is characterized as follows:*

(i) If $\underline{\alpha} < \alpha < \overline{\alpha}$, an efficient equilibrium replicating the vertically integrated outcome arises. The quantity traded is $\hat{q} = q_I$, the wholesale price is $\hat{w} = w_I$, the fixed fee is $\hat{F} = \alpha\pi_D(q_I) - (1 - \alpha)\pi_U(q_I)$, the consumer price is $p(q_I)$, and the input price is $r(q_I)$.

(ii) If $\overline{\alpha} < \alpha < 1$, an equilibrium with double markup arises. The quantity traded $\hat{q}$ is given by (17)*, where $q_I \geq \hat{q} \geq q_\mu$, the wholesale price is $\hat{w} = MR_D(\hat{q})$, the fixed fee is $\hat{F} = \overline{F}$, the consumer price is $p(\hat{q}) \geq p(q_I)$, and the input price is $r(\hat{q}) \leq r(q_I)$.*

[52]Note first that when $\hat{F} = \overline{F} = 0$, we have $\hat{q} = q_\mu$, as the analysis reduces to the linear wholesale contract setting studied in Section 4. Moreover, whenever $\hat{F} = \overline{F} \geq 0$, it follows that $\Pi_U(\hat{q}, \hat{F}) \geq \Pi_U(\hat{q}, 0) = \pi_U(q_\mu)$ and $\Pi_D(\hat{q}, \hat{F}) \leq \Pi_D(\hat{q}, 0) = \pi_D(q_\mu)$. As a result, we obtain $\hat{\beta}_D(\hat{q}, \hat{F}, \alpha) \geq \hat{\beta}_D(\hat{q}, 0, \alpha) = \beta_D(q_\mu, \alpha)$.

[53]The reasoning parallels that in footenote 52. In particular, whenever $\hat{F} = \underline{F} \leq 0$, we have $\Pi_U(\hat{q}, \hat{F}) \leq \Pi_U(\hat{q}, 0) = \pi_U(q_\nu)$ and $\Pi_D(\hat{q}, \hat{F}) \geq \Pi_D(\hat{q}, 0) = \pi_D(q_\nu)$. Consequently, $\hat{\beta}_U(\hat{q}, \hat{F}, \alpha) \geq \hat{\beta}_U(\hat{q}, 0, \alpha) = \beta_U(q_\nu, \alpha)$.

(ii) If $0 < \alpha < \underline{\alpha}$, an equilibrium with double markdown arises. The quantity traded $\hat{q}$ is given by (18)*, where $q_I \geq \hat{q} \geq q_\nu$, the wholesale price is $\hat{w} = MC_U(\hat{q})$, the fixed fee is $\hat{F} = \underline{F}$, the consumer price is $p(\hat{q}) \geq p(q_I)$, and the input price is $r(\hat{q}) \leq r(q_I)$.*

When frictions prevent firms from setting the fixed fee to its optimal level, $\alpha\pi_D(q_I) - (1-\alpha)\pi_U(q_I)$, double-marginalization arises, leading to an inefficient outcome with $\hat{q} \leq q_I$. This distortion emerges in two distinct cases. When $\alpha \geq \overline{\alpha}$, as described in Proposition 3, U exercises monopoly power by charging a markup over its marginal cost when selling to D, resulting in the double markup outcome. In contrast, when $\alpha \leq \underline{\alpha}$, the logic follows Proposition 4, where D exercises monopsony power by imposing a markdown below its marginal revenue when purchasing from U, giving rise to the double markdown outcome. Although the distortion is of the same nature, it is less severe than under the linear wholesale contract, provided that frictions are not too extreme (i.e., $\overline{F} > 0$ and $\underline{F} < 0$). Figure 6 illustrates Proposition 5, depicting the three types of equilibria that may arise depending on the bargaining weight α.

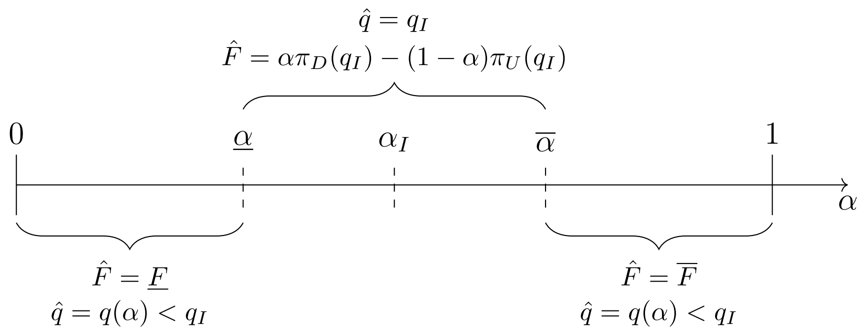


Figure 6: Equilibrium quantity and fixed fee in the presence of frictions.

Proposition 5 highlights that our main findings extend to the case in which U and D bargain over a two-part tariff contract, provided that frictions constraining the fixed fee are present. One rationale for such frictions is when the fixed fee must be paid upfront, but firms have access to imperfect financial markets, leading to liquidity constraints. An alternative microfoundation consists of introducing some uncertainty in the realization of consumer demand. For instance, consider a simple setting with two states of consumer demand: low and high demand. Suppose U and D bargain over a two-part tariff contract before demand is realized, and the fixed fee is paid only afterward. In the low-demand state, either D or U may be unable to fulfill the agreed-upon payment, especially if it is large. Anticipating this possibility, U and D may prefer to limit the fixed fee and distort the marginal price upward to avoid an ex-post breakdown of the trading relationship.[54]

6.3 Price Floors

It is well known that minimum wages can increase employment in the presence of monopsony power (see, e.g., Stigler (1946)). Generally, price floors can improve welfare when applied in markets where buyers charge a markdown when purchasing inputs, and many countries have considered or introduced either temporary or permanent price floors in agricultural markets.[55] Similarly, during periods of inflation, the implementation of price caps on specific food products is often proposed as a measure to protect consumers. In this section, we study the welfare and distributional effects of introducing a price floor in the upstream input market, allowing us to study its incidence in a vertical supply chain with bargaining. By symmetry, the entire reasoning can be applied to price caps imposed on final consumer prices.

We begin by characterizing the optimal price floor policy under vertical integration, again serving as a useful benchmark, here for the subsequent analysis of the optimal price floor under bargaining in the vertical supply chain.

For both settings, we denote $\underline{r}$ the price-floor level and $\underline{q} = r^{-1}(\underline{r})$. Such a price floor affects the input supply curve $r(q)$ in the following way. It remain unchanged when $q > \underline{q}$, whereas, for all $q \leq \underline{q}$, we have:

$$r(q) = \underline{r}.$$

When the price floor is binding, the input supply curve thus becomes perfectly elastic, i.e. $\varepsilon_r \to \infty$. As a result, the marginal cost functions of firms operating further downstream in the supply chain are similarly affected. They remain unchanged when

[54] A complete formalization of this microfoundation is available upon request.

[55] Such price floors, for instance exist in the U.S. raw milk market. See Avignon and Guigue (2025) for a public policy note extensively discussing the relevance of price floors in agricultural markets.

$q > \underline{q}$, whereas, for all $q \leq \underline{q}$, we have for all $i \in \{I, U, D\}$:

$$r(q) = MC_i(q) = \underline{r}.$$

6.3.1 Vertical Integration

We begin by analyzing the optimal price floor policy under vertical integration. When the price floor is binding, the input supply curve is flat and I's marginal cost becomes constant ($MC_I(q) = \underline{r}$ when $q \leq \underline{q}$), eliminating I's ability to exert monopsony power. We obtain the following proposition:

Proposition 6 (Vertical Integration under Optimal Price Floor) *Under vertical integration, the optimal price floor is $\underline{r_I} = r(\underline{q_I})$ where $\underline{q_I}$ is defined by $MR_D(\underline{q_I}) = r(\underline{q_I})$. When the optimal price floor is implemented, the vertically integrated firm I sets the equilibrium quantity $\underline{q_I} \in (q_I, q_W)$. The consumer price $p(\underline{q_I}) \in (p(q_I), p(q_W))$ and the input price is $r(\underline{q_I}) \in (r(q_I), r(q_W))$. I does not charge any markdown, i.e., $\underline{\nu_I} = \frac{MR_I(\underline{q_I})}{r(\underline{q_I})} = 1$ and charges a markup given by $\underline{\mu_I} = \frac{p(\underline{q_I})}{MC_I(\underline{q_I})} = \frac{\varepsilon_p}{\varepsilon_p - 1}$. Consequently, I's total margin is given by $\underline{M_I} = \frac{p(\underline{q_I})}{r(\underline{q_I})} = \underline{\mu_I}$.*

Proof. See Appendix C.1. ■

Proposition 6 stipulates that the optimal price floor fully suppresses I's ability to exercise monopsony power. Doing so, it reduces the distortion described in Proposition 1: under the optimal price floor, I only exerts a markup, and its total margin (and profit) shrinks. In contrast, the introduction of the optimal price floor benefits both consumers and input suppliers by increasing the quantity traded and raising the price received by input suppliers.

6.3.2 Bargaining

We consider now the equilibrium outcome of a price floor when U and D bargain over a linear wholesale price. Again, when the price floor is binding, U's and D's marginal costs become constant, eliminating both firms' ability to exert monopsony power. Specifically, U loses the ability to lower its input price r by reducing the quantity purchased in Stage 2. Thus, U loses incentives to restrict the quantity traded and D, in turn, faces a flat supply curve and constant marginal cost, hereby losing its ability to exert monopsony power. We now determine the welfare-maximizing price-floor policy. The optimal price floor is at a level such that, letting firms freely bargain and trade afterwards, the resulting equilibrium quantity is maximized (provided that the quantity in any case remains below q_W, the quantity that maximizes welfare).

When a binding price floor is introduced, the setting becomes similar to the canonical model of vertical relationships with double markups. Indeed, for any $w > \underline{r}$, U is willing to trade more than D, and the equilibrium quantity is determined by the condition $w = MR_D(q)$. Consequently, double markup distortion necessarily arises in equilibrium for any $\alpha > 0$. This leads to the following proposition.

Proposition 7 (Bargaining under Optimal Price Floor) *For a given α, the optimal price floor is $\underline{r_\mu} = r(\underline{q_\mu}(\alpha))$ where $\underline{q_\mu}(\alpha)$ is defined by $\widetilde{MR}_U(\underline{q_\mu}, \alpha) = r(\underline{q_\mu})$. When the optimal price floor $\underline{r_\mu}$ is implemented, the wholesale price is $\underline{w_\mu} = MR_D(\underline{q_\mu}) > \underline{w_I}$, the quantity exchanged is $\underline{q_\mu} \in (q_\mu, \underline{q_I})$, the consumer price is $p(\underline{q_\mu}) \in (p(q_\mu), p(\underline{q_I}))$, and the input price is $r(\underline{q_\mu}) = \underline{r_\mu} \in (r(q_\mu), r(\underline{q_I}))$. Neither U nor D charge a markdown, i.e., $\underline{\nu_U} = 1$ and $\underline{\nu_D} = \frac{MR_D(\underline{q_\mu})}{\underline{w_\mu}} = 1$, but double marginalization persists due to U's seller power, leading U to charge a markup, given by:*

$$\underline{\mu_U} = \frac{\underline{w_\mu}}{\underline{r_\mu}} = \frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D} - (1 - \underline{\beta_D}(\underline{q\mu}, \alpha))} = \frac{\alpha\varepsilon_{MR_D} + (1-\alpha)(\varepsilon_p - 1)}{(\alpha(\varepsilon_{MR_D} - 1) + (1-\alpha)(\varepsilon_p - 1))},$$

which adds up to D's markup, given by $\underline{\mu_D} = \frac{p(\underline{q_\mu})}{MR_D(\underline{q_\mu})} = \frac{\varepsilon_p}{\varepsilon_p - 1}$. Consequently, U's margin is equal to $\underline{M_U} = \frac{\underline{w_\mu}}{\underline{r_\mu}} = \underline{\mu_U}$, D's margin is equal to $\underline{M_D} = \frac{p(\underline{q\mu})}{\underline{w\mu}} = \underline{\mu_D}$. The total margin of the supply chain is given by:

$$\underline{\mathcal{M}} = \frac{p(\underline{q\mu})}{\underline{r_\mu}} = \underline{\mu_U} \times \underline{\mu_D}.$$

Proof. See Appendix C.2. ■

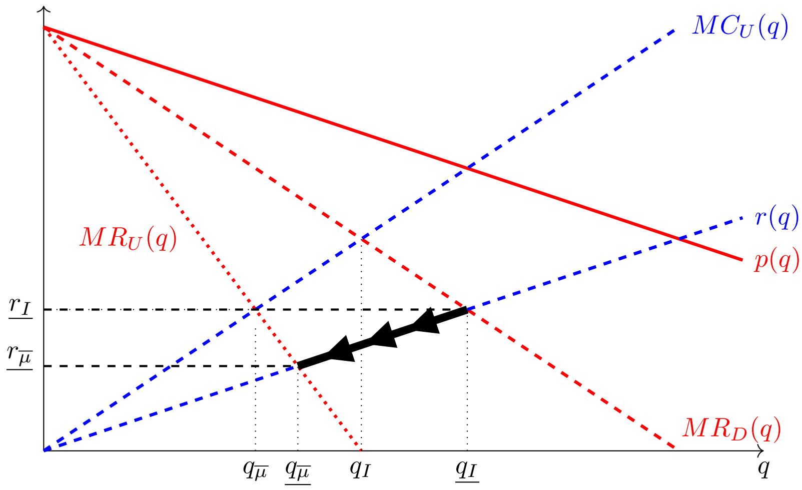


Figure 7: Optimal Price Floors.

Corollary 3 *The optimal price-floor level strictly decreases in α, i.e., increases with D's buyer power. Under the optimal price floor, welfare strictly decreases in α, i.e., increases with D's buyer power. Moreover, a decrease in α has the following distributional effects: it benefits input suppliers and consumers, and hurts U. It always benefits D when U is powerful ($\alpha_I < \alpha \leq 1$) but may hurt it when it is powerful ($0 \leq \alpha < \alpha_I$).*

Proposition 7 stipulates that the optimal price floor is a function of α, and Corollary 3 shows that the optimal price-floor level strictly decreases in α, i.e., increases with D's buyer power. Figure 7 illustrates the latter result, representing the equilibria $(\underline{r_\mu}, \underline{q_\mu})$ under optimal price-floor, with the arrows indicating the evolution when α increases. In particular, when $\alpha = 0$, the equilibrium quantity under the price floor is $\underline{q_I}$ defined by $MR_D(\underline{q_I}) = r(\underline{q_I})$ and the optimal price floor is $\underline{r_I} = r(\underline{q_I})$ as in the vertical integration case. When instead $\alpha = 1$, $\underline{q_{\overline{\mu}}}$ is defined by $MR_U(\underline{q_{\overline{\mu}}}) = r(\underline{q_{\overline{\mu}}}) = \underline{r_{\overline{\mu}}}$. More generally, the optimal price-floor level strictly decreases in α, i.e., increases with D's buyer power. The key insight is as follows. When α is high, the source of distortion in the vertical chain, even in the absence of a price floor, stems from double markupization. In this case, only U exerts monopsony power in its input market. Although a price floor eliminates U's markdown, the distortion caused by U's and D's markups remains. Therefore, if the price floor is set too high, it can exacerbate this distortion. As α decreases i.e., D's bargaining power increases, the double markup distortion decreases, and the optimal price floor can be raised. On the other hand, the price floor must be

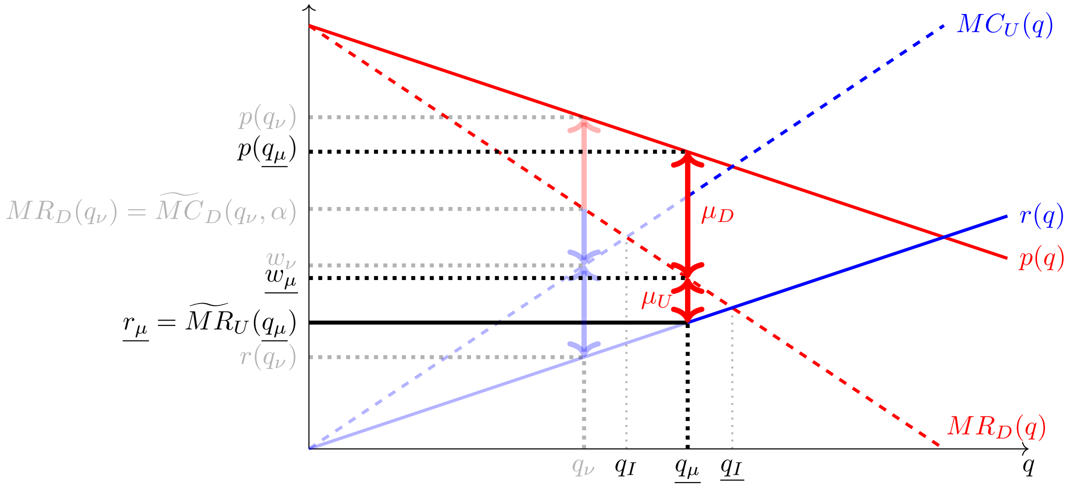


Figure 8: Optimal Price Floor under (initial) Double Markdownization ($0 < \alpha < \alpha_I$).

set at a level high enough to bind in equilibrium. Indeed, if the price floor is too low, firms are tempted to keep negotiating on the increasing portion of the marginal curve. As $\widetilde{MR}_U(q, \alpha)$ is decreasing in α, the price-floor must increase when α decreases to keep binding in equilibrium.

Figure 8 illustrates Proposition 7 in the case where D is powerful, i.e., $0 \le \alpha < \alpha_I$, showing the effect of introducing an optimal price floor when double markdownization otherwise prevails. The initial equilibrium appears in semi-transparency, and the equilibrium under the optimal price floor (with double markup) appears in plain colors. In that case, the price floor neutralizes both U's and D's monopsony power, as their respective marginal costs become constant for any $q \le \underline{q}$. Yet, double marginalization persists as introducing the price floor, although improving welfare, here creates double markupization. Similarly, double markupization persists when introducing a price floor when U is powerful, as it does not address the distortion coming from both firms' seller power in their output markets.

Overall, introducing a binding price floor yields the classical bilateral monopoly framework where U supplies at constant marginal cost, and countervailing buyer power is always desirable for welfare, as stipulated in Corollary 3. Moreover, introducing a price-floor always benefits D when it is relatively weak ($\alpha_I < \alpha < 1$). When D is powerful ($0 < \alpha < \alpha_I$), D may capture a larger share of joint profits under price floor, however, the quantity traded under price floor is too high as compared to the quantity q_I

that maximizes the vertically integrated profit because D has lost its ability to impose a markdown. When $\alpha = 0$, D may thus be better off with TIOLI profit without the price floor than with the industry profit under a price floor. In contrast, the price floor hurts the industry supplier U, even when the latter has no bargaining power because U obtains no profit in that case, whereas U was able to preserve its markdown absent the price floor.

Remark 2 *The welfare benefit of introducing an optimal price floor diminishes with α for $\alpha \in [0, \alpha_I]$. Under linear demand and supply curves, the welfare benefit of introducing the optimal price floor is minimal when $\alpha = \alpha_I$.*

Absent the price-floor policy, q_ν increases in α over the interval $[0, \alpha_I]$. Instead, the optimal price floor $\underline{q_\mu}$ decreases when α increases. The welfare benefit of the price-floor policy, as captured by the gap $\underline{q_\mu} - q_\nu$, increases (resp. decreases) as α decreases (resp. increases) in the interval $[0, \alpha_I]$. Effects are more ambiguous when α lies in the interval $[\alpha_I, 1]$ as both $\underline{q_\mu}$ and q_μ are increasing as α decreases. In the linear demand and supply case, the gap $\underline{q_\mu} - q_\mu$ shrinks when α decreases, thus reducing the benefit of the optimal price-floor policy.

Overall, it is crucial to accurately assess the balance of power between U and D and, specifically, to determine whether U charges a markup or D charges a markdown, to design an effective price-floor policy. Assume, for instance, that the policymaker only observes the quantity traded $q_{\overline{\mu}}$. Such a quantity may result from two distinct distributions of bargaining power: one where U is powerful, and one where D is. In the symmetric, linear example of Figure , we have $q_{\overline{\mu}} = q_{\overline{\nu}}$. If $\alpha = 0$, introducing a price floor at the level $\underline{r_I}$, in that case optimal, creates a massive benefit for welfare as the quantity traded jumps from $q_{\overline{\nu}}$ to $\underline{q_I}$. If instead $\alpha = 1$, introducing a price floor at the level $\underline{r_I}$, in that case suboptimally too high, implies that the quantity, set by U, is pinned down by the intersection of MR_U and $\underline{r_I}$. In our specific graphical example, it coincides with the intersection of MR_U and MC_U, resulting in neutral welfare effects. This underscores the importance of assessing the nature of double marginalization for the optimal design of such a public policy. Moreover, and to further stress this point, it is worth noting that, departing from our graphical example, and specifically under asymmetric, non-linear supply and demand curves with a relatively higher elasticity of supply, the welfare effects of such a suboptimal policy could be negative.

7 Conclusion

This article provides a unified framework that allows for analysis of the interactions between monopsony, monopoly, and countervailing power theories within a vertical supply chain. Introducing an upstream firm with increasing marginal costs -exerting monopsony power in its input market- into the standard bilateral monopoly model, we demonstrate that the downstream firm no longer solely determines the quantity traded in the supply chain. Instead, we leverage subgame perfection to show how the short-side rule endogenizes which side sets quantity traded in equilibrium. These modifications to the canonical model of vertical relationships offer new perspectives on how the distribution of bargaining power shapes welfare outcomes in a vertical supply chain. Crucially, we identify the nonmonotonic welfare effects of both seller and buyer power. We show that both bilateral efficiency and welfare are maximized when each firm's bargaining power fully countervails the other's market power, which occurs for a specific distribution of bargaining power contingent on the relative degree of supply and demand elasticity. Otherwise, double marginalization occurs: *double markupization* arises when the upstream firm holds excessive bargaining power, whereas *double markdownization*, a novel type of distortion, emerges in the opposite case. Our analysis yields novel insights for policy intervention and empirical research, including calling for greater flexibility in modeling cost functions in empirical work.

Appendix

A Proofs

A.1 Proof of Lemma 1

A (weakly) dominant strategy for U and D is to announce the quantity that maximizes their profit, that is $\tilde{q}_U$ and $\tilde{q}_D$ respectively. Assume that U, say, anticipates that D will announce a quantity q_D^a, where the superscript a stands for "anticipated". If $q_D^a \leq \tilde{q}_U(w)$, announcing $\tilde{q}_U(w)$ is as good as announcing any other quantity larger or equal to q_D^a given that, anticipating the short side rule only q_D^a will then be traded. Announcing a lower quantity than q_D^a is strictly dominated as, according to the short side rule, the quantity traded would then be even further away from the optimal quantity $\tilde{q}_U(w)$. If instead U anticipates that $\tilde{q}_U(w) < q_D^a$, announcing $\tilde{q}_U(w)$ is the best strategy for U as the quantity traded in that case maximizes its profit. The reasoning is symmetric for D, and each firm announces the quantity that maximizes its profit. Because a (weakly) dominant strategy is the best response to *any* strategy the other firm might play (including the one chosen in equilibrium), the strategy profile $(\tilde{q}_U(w),\tilde{q}_D(w))$ constitutes a Nash equilibrium.

A.2 Proof of Proposition 2

Proof that $\boldsymbol{\alpha_I = \frac{\varepsilon_p(q_I)-1}{\varepsilon_p(q_I)+\varepsilon_r(q_I)}}$. When $\alpha = \alpha_I = \frac{\pi_U(q_I)}{\pi_U(q_I)+\pi_D(q_I)}$, the bargaining leads to the efficient outcome q_I such that $MC_U(q_I) = MR_D(q_I)$. We can rewrite α_I as follows:

$$\begin{aligned} \alpha_I &= \frac{(MR_D(q_I) - r(q_I))q_I}{(p(q_I) - r(q_I))q_I} \\ &= \frac{MR_D(q_I) - r(q_I)}{p(q_I) - r(q_I)}. \end{aligned}$$

Using $p(q_I) = MR_D(q_I)\frac{\varepsilon_p(q_I)}{\varepsilon_p(q_I)-1}$, $r(q_I) = MC_U(q_I)\frac{\varepsilon_r(q_I)}{\varepsilon_r(q_I)+1}$ and $MC_U(q_I) = MR_D(q_I)$, we obtain:

$$\begin{aligned} \alpha_I &= \frac{MR_D(q_I) - MR_D(q_I)\frac{\varepsilon_r(q_I)}{\varepsilon_r(q_I)+1}}{MR_D(q_I)\frac{\varepsilon_p(q_I)}{\varepsilon_p(q_I)-1} - MR_D(q_I)\frac{\varepsilon_r(q_I)}{\varepsilon_r(q_I)+1}} \\ &= \frac{\varepsilon_p(q_I) - 1}{\varepsilon_p(q_I) + \varepsilon_r(q_I)}. \end{aligned}$$

Existence. We prove below the existence of an equilibrium (w_I, q_I) in $\alpha = \alpha_I$. Formally, we check that the first order condition is weakly positive in q_I, i.e. no profitable deviation towards $w_I + \varepsilon$ or $w_I - \varepsilon$.

- We analyze a deviation from $w_I \to w_I + \varepsilon$, which means that $w(q) = MR_D(q)$ and that

$q_I \to q_I - \varepsilon$. The Nash product is:

$$\max_q [MR_D(q)q - r(q)q]^{\alpha}[p(q)q - MR_D(q)q]^{(1-\alpha)}$$

The first-order condition is given by:

$$\frac{\alpha}{\pi_U(q)}(MR_U(q) - MC_U(q)) + \frac{(1-\alpha)}{\pi_D(q)}(MR_D(q) - MR_U(q)) = 0 \tag{19}$$

Given that in q_I we have $MC_U(q_I) = MR_D(q_I)$, evaluation the first order condition (19) in q_I, there is no profitable deviation whenever:

$$\begin{aligned} \underbrace{(MR_U(q_I) - MR_D(q_I))}_{<0}(\alpha\pi_D(q_I) - (1-\alpha)\pi_U(q_I)) &\geq 0 \\ \Rightarrow (1-\alpha)\pi_U(q_I) &\geq \alpha\pi_D(q_I) \\ \Rightarrow \alpha &\leq \alpha_I \end{aligned}$$

There is no profitable deviation in $\alpha = \alpha_I$.

- We analyze a deviation towards $w_I \to w_I - \varepsilon$ which implies that $w = MC_U(q)$ and that $q_I \to q_I - \varepsilon$ The Nash product is:

$$\max_q [MC_U(q)q - r(q)q]^{\alpha}[p(q)q - MC_U(q)q]^{(1-\alpha)}$$

The first-order condition is given by:

$$\frac{\alpha}{\pi_U(q)}(MC_D(q) - MC_U(q)) + \frac{(1-\alpha)}{\pi_D(q)}(MR_D(q) - MC_D(q)) = 0 \tag{20}$$

Replacing in (20)

$$\begin{aligned} \underbrace{(MC_D(q_I) - MC_U(q_I))}_{>0}(\alpha\pi_D(q_I) - (1- \Rightarrow \alpha)\pi_U(q_I)) \geq 0 \\ \Rightarrow \alpha\pi_D(q_I) \geq (1-\alpha)\pi_U(q_I) \\ \Rightarrow \alpha \geq \alpha_I \end{aligned}$$

There is no profitable deviation in $\alpha = \alpha_I$.

A.3 Bargaining when U is Powerful ($\alpha_I < \alpha < 1$)

A.3.1 Second-Order Condition

The first-order condition given by (9) can be rearranged as follows:

$$\alpha(MR_U(q) - MC_U(q))(p(q) - MR_D(q))q + (1-\alpha)(MR_D(q) - MR_U(q))(MR_D(q) - r(q))q = 0. \quad (21)$$

We show that the first-order condition is strictly decreasing in q if $\sigma_r > -2$ and $\sigma^{MR} < 2$.

The second-order condition yields:

$$\alpha(MR'_D(2 - \sigma^{MR}) - (\sigma_r + 2)r'(q))(p(q) - MR_D(q))q + (MR_U(q) - MC_U(q))(MR_D - MR_U))$$
$$+(1-\alpha)(MR'_D(\sigma^{MR} - 1)(MR_D - r(q))q + (MR_D - MR_U)(MR_U - MC_U)) < 0. \quad (22)$$

Using $-MR'_D q = (MR_D - MR_U)$ and $a \equiv MR_U(q) - MC_U(q) \leq 0$, $b \equiv p(q) - MR_D(q) > 0$, $c \equiv MR_D(q) - MR_U(q) > 0$ and $d \equiv MR_D(q) - r(q) > 0$. Using also $a + c - d = -r'(q)q$, the first-order condition (21) simplifies as follows:

$$\alpha ab + (1-\alpha)cd = 0 \Leftrightarrow d = \frac{\alpha ab}{-(1-\alpha)c}, \quad (23)$$

and the second-order condition (22) becomes:

$$\alpha((\sigma^{MR} - 2)cb + (a + c - d)(\sigma_r + 2)b + ca) + (1-\alpha)(-(\sigma^{MR} - 1)cd + ca) < 0.$$

Using that $-c < a \Leftrightarrow MC_U < MR_D$, we find that the second-order condition (22) holds for any $\sigma_r > -2$ and $\sigma^{MR} < 2$.

A.3.2 Weight β_D

We show here that $\beta_D(q^*, \alpha)$ decreases in α. Equation (10) in the main text defines $\beta_D(q^*, \alpha)$:

$$\beta_D(q^*, \alpha) = \frac{MC_U(q^*) - MR_U(q^*)}{MR_D(q^*) - MR_U(q^*)}$$

and we note that $0 \leq \beta_D(q^*, \alpha) \leq 1$, as $MC_U(q^*) \geq MR_U(q^*)$, $MR_D(q^*) \geq MR_U(q^*)$ and $MC_U(q^*) \leq MR_D(q^*)$.

We now determine how β_D is affected by changes in α. The chain rule implies that:

$$\frac{\partial \beta_D}{\partial \alpha} = \frac{\partial \beta_D}{\partial q} \frac{\partial q}{\partial w} \frac{\partial w}{\partial \alpha}$$

with $\frac{\partial q}{\partial w} = MR'(w) < 0$ and $\frac{\partial w}{\partial \alpha} > 0$ (see proof of Proposition 8). To determine the sign of $\frac{\partial \beta_D}{\partial q}$ and hence $\dfrac{\partial \beta_D}{\partial \alpha}$, we totally differentiate (10), yielding:

$$dMC_U = (MR_D - MR_U)d\beta_D + [(1-\beta_D)MR'_U + \beta_D MR'_D]\, dq$$

Dividing both sides by dq and rearranging yields:

$$\frac{d\beta_D}{dq} = \frac{1}{MR_D - MR_U}\left[(1-\beta_D)MR'_U + \beta_D MR'_D - MC'_U\right].$$

We thus have $\dfrac{d\beta_D}{dq} > 0$ as $MR_D - MR_U < 0$ since $MR_U(q) \equiv MR'_D(q)q + MR_D(q) < MR_D(q)$, and $(1-\beta_D)MR'_U + \beta_D MR'_D - MC'_U < 0$ since $MR'_U(q) = (2-\sigma_{MR_D})MR'_D(q) < 0$ from Assumption 3, $MR'_D(q) < 0$ and $MC'_U(q) > 0$ from Assumption 1, and $\beta_D(q^*, \alpha) \geq 0$ as proved above.

Putting pieces together:

$$\frac{\partial \beta_D}{\partial \alpha} = \underbrace{\frac{\partial \beta_D}{\partial q}}_{>0} \underbrace{\frac{\partial q}{\partial w}}_{<0} \underbrace{\frac{\partial w}{\partial \alpha}}_{>0} < 0.$$

A.3.3 Markup μ_U

To determine the markup $\mu_U = \frac{MR_D(q)}{MC_U(q)}$, we divide each term of the first-order condition given by (21) by $MC_U(q)$:

$$\alpha\left(\frac{MR_U}{MC_U} - 1\right)\left(\frac{p}{MC_U} - \frac{MR_D}{MC_U}\right) + (1-\alpha)\left(\frac{MR_D}{MC_U} - \frac{MR_U}{MC_U}\right)\left(\frac{MR_D}{MC_U} - \frac{r}{MC_U}\right) = 0.$$

We then use the following simplifications:

- $\frac{MR_U}{MR_D} = \frac{MR'_D q}{MR_D} + 1 = 1 - \frac{1}{\varepsilon_{MR_D}} = \frac{\varepsilon_{MR_D}-1}{\varepsilon_{MR_D}}$,
- $MR_D = p\left(1 - \frac{1}{\varepsilon_p}\right) \Leftrightarrow \frac{p}{MR_D} = \frac{1}{1-\frac{1}{\varepsilon_p}} = \frac{\varepsilon_p}{\varepsilon_p - 1}$,
- $MC_U = r'q + r = r\left(\frac{1}{\varepsilon_r} + 1\right) \Leftrightarrow \frac{r}{MC_U} = \frac{\varepsilon_r}{\varepsilon_r + 1}$,

to rewrite:

$$\begin{aligned}
& \alpha\left(\frac{MR_D}{MC_U}\left(\frac{\varepsilon_{MR_D}-1}{\varepsilon_{MR_D}}\right) - 1\right)\left(\frac{MR_D}{MC_U}\frac{1}{\varepsilon_p - 1}\right) + (1-\alpha)\left(\frac{MR_D}{MC_U}\frac{1}{\varepsilon_{MR_D}}\right)\left(\frac{MR_D}{MC_U} - \frac{\varepsilon_r}{\varepsilon_r+1}\right) = 0 \\
\Longleftrightarrow \quad & \alpha\left(\frac{MR_D}{MC_U}\frac{\varepsilon_{MR_D}-1}{\varepsilon_{MR_D}} - 1\right)\left(\frac{1}{\varepsilon_p - 1}\right) + (1-\alpha)\left(\frac{1}{\varepsilon_{MR_D}}\right)\left(\frac{MR_D}{MC_U} - \frac{\varepsilon_r}{\varepsilon_r+1}\right) = 0 \\
\Longleftrightarrow \quad & \frac{MR_D}{MC_U}\left(\alpha\left(\frac{\varepsilon_{MR_D}-1}{\varepsilon_{MR_D}(\varepsilon_p - 1)}\right) + (1-\alpha)\left(\frac{1}{\varepsilon_{MR_D}}\right)\right) = \frac{\alpha}{\varepsilon_p - 1} + (1-\alpha)\frac{\varepsilon_r}{\varepsilon_{MR_D}(\varepsilon_r + 1)} \\
\Longleftrightarrow \quad & \frac{MR_D}{MC_U}\left(\alpha\left(\varepsilon_{MR_D}-1\right)\left(\varepsilon_r+1\right)\right) + (1-\alpha)\left(\left(\varepsilon_r+1\right)\left(\varepsilon_p-1\right)\right) = \alpha\left(\varepsilon_r+1\right)\varepsilon_{MR_D} + (1-\alpha)\varepsilon_r\left(\varepsilon_p-1\right) \\
\Longleftrightarrow \quad & \mu_U = \frac{MR_D}{MC_U} = \frac{\alpha\varepsilon_{MR_D}\left(\varepsilon_r+1\right) + (1-\alpha)\left(\varepsilon_p-1\right)\varepsilon_r}{\alpha\left(\varepsilon_{MR_D}-1\right)\left(\varepsilon_r+1\right) + (1-\alpha)\left(\varepsilon_r+1\right)\left(\varepsilon_p-1\right)}.
\end{aligned}$$

When U is powerful, its markup can also be rewritten as:

$$\mu_U = \omega_D(\alpha)\frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D}-1} + (1-\omega_D(\alpha))\frac{\varepsilon_r}{\varepsilon_r+1}, \tag{24}$$

where $\omega_D(\alpha) \equiv \frac{\alpha(\varepsilon_{MR_D}-1)}{\alpha(\varepsilon_{MR_D}-1)+(1-\alpha)(\varepsilon_p-1)}$ and with $\frac{\partial \omega_D(\alpha)}{\partial \alpha} > 0$. In the case we consider here, namely $\alpha_I < \alpha < 1$, we have $(0 <) \frac{\varepsilon_{MR_D}-1}{\varepsilon_{MR_D}+\varepsilon_r} \leq \omega_D(\alpha) \leq 1$.[56] The weight $\omega_D(\alpha)$ is a function of α, and of demand primitives through ε_p and σ_p (as $\varepsilon_{MR_D} = \frac{\varepsilon_p-1}{2-\sigma_p}$).

Equation (24) echoes expressions delivered by the exogenous right-to-manage models of Alviarez et al. (2023); Azkarate-Askasua and Zerecero (2022) and Wong (2023), whereby the bilateral markup (or markdown, see also Appendix A.4.3) is a weighted average between a monopoly (or oligopoly) markup term and a monopsony (or oligopsony) markdown term. Indeed, if $\alpha = 1$, then $\omega_D(\alpha) = 1$ and $\mu_U = \frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D}-1}$, as U can make a take-it-or-leave-it offer to D. Similarly, if $\alpha = 0$ then $\omega_D(\alpha) = 0$, and (24) would state $\mu_U = \frac{\varepsilon_r}{\varepsilon_r+1} = \nu_U^{-1}$ and thus $p = w$ and $M_U = 1$. However, by sugbgame perfection criterion and short-side rule application, such equilibrium is ruled out in our framework, as D endogenously concedes the right-to-manage if becoming too powerful, i.e, when $\alpha < \alpha_I$. Instead, at this limit, i.e., if $\alpha = \alpha_I = \frac{\varepsilon_p-1}{\varepsilon_p+\varepsilon_r}$, then $\omega_D(\alpha_I) = \frac{\varepsilon_{MR_D}-1}{\varepsilon_{MR_D}+\varepsilon_r}$ and $\mu_U = 1$, yielding the vertically integrated outcome q_I.

A.3.4 Set of Equilibria

Using the first order condition (9), we introduce the following function:

$$\Phi(q_U,\alpha) \equiv \alpha(MR_U(q_U) - MC_U(q_U))\pi_D(q_U) + (1-\alpha)(MR_D(q_U) - MR_U(q_U))\pi_U(q_U) = 0.$$

Under Assumption 2, applying the implicit function theorem, we have that:

$$Sign\left(\frac{\partial q_U}{\partial \alpha}\right) = Sign\left(\frac{\partial \Phi(q_U,\alpha)}{\partial \alpha}\right),$$

with

$$\begin{aligned}\frac{\partial \Phi(q_U,\alpha)}{\partial \alpha} &= (MR_U(q_U) - MC_U(q_U))\pi_D(q_U) - (MR_D(q_U) - MR_U(q_U))\pi_U(q_U)\\ &= (MR_U(q_U) - MR_D(q_U))\pi_U(q_U) < 0.\end{aligned}$$

Therefore, the equilibrium quantity q_U decreases in α. We know that $q_U = q_I$ when $\alpha = \alpha_I$ and therefore $q_U > q_I$ when $\alpha < \alpha_I$. As shown in Section 4.1, in that case $\tilde{q}_D(w) > \tilde{q}_U(w)$ and the initial assumption that $w = MR_D(q)$ no longer holds. The set of equilibria defined by (9) and $w_U = MR_D(q_U)$ thus only exists for $\alpha \in [\alpha_I, 1]$.

[56] $\omega_D(\alpha) > 0$ holds when the demand function is supermodular, which we assumed above.

A.4 Bargaining when D is Powerful ($0 < \alpha < \alpha_I$)

A.4.1 Second-Order Condition

The first-order condition given by (12) can be rearranged as follows:

$$\alpha(MC_D(q) - MC_U(q))\,(p(q) - MC_U(q))\,q + (1-\alpha)(MR_D(q) - MC_D(q))(MC_U(q) - r(q))q = 0. \quad (25)$$

We show that the first-order condition is strictly decreasing in q if $\sigma_p < 2$ and $\sigma^{MC} > -2$. The second-order condition yields:

$$\begin{aligned} &\alpha(MC_U'(\sigma^{MC}+1))(p(q) - MC_U(q))q + (MC_D - MC_U)(MR_D - MC_D)) \\ &+(1-\alpha)((2-\sigma_p)p'(q) - MC_U'(\sigma^{MC}+2))(MC_U(q) - r(q))q + (MR_D - MC_D)(MC_D - MC_U)). \quad (26)\end{aligned}$$

Using $MC_U'q = (MC_D - MC_U) \equiv e > 0$, $f \equiv p(q) - MC_U(q) > 0$, $g \equiv MR_D(q) - MC_D(q) < 0$ and $h \equiv MC_U(q) - r(q) > 0$. Using also $g + e - f = p'(q)q$, the first-order condition (21) simplifies as follows:

$$\alpha ef + (1-\alpha)gh = 0 \Leftrightarrow h = \frac{\alpha ef}{-(1-\alpha)g},$$

and the second-order condition (26) becomes:

$$\alpha((\sigma^{MC}+1)ef + eg) + (1-\alpha)((2-\sigma_p)(g+e-f) - e(\sigma^{MC}+2)h + eg) < 0.$$

Using that $-g < e \Leftrightarrow MC_U < MR_D$, we find that the second-order condition (26) holds for any $\sigma_p < 2$ and $\sigma^{MC} > -2$.

A.4.2 Weight β_U

We show here that $\beta_U(q^*,\alpha)$ increases in α. Equation (13) in the main text defines $\beta_U(q^*,\alpha)$:

$$\beta_U(q^*,\alpha) = \frac{MC_D(q^*) - MR_D(q^*)}{MC_D(q^*) - MC_U(q^*)}$$

with $0 \le \beta_U(q^*,\alpha) \le 1$, as $MC_D(q^*) \ge MC_U(q^*)$, $MC_D(q^*) \ge MR_D(q^*)$, and $MC_U(q^*) \le MR_D(q^*)$.

Studying how $\beta_U(q^*,\alpha)$ is affected by changes in α similarly to what we did for $\beta_D(q^*,\alpha)$ in Appendix A.3.2, we can show that $\dfrac{\partial \beta_U}{\partial \alpha} = \underbrace{\dfrac{\partial \beta_U}{\partial q}}_{>0}\underbrace{\dfrac{\partial q}{\partial w}}_{>0}\underbrace{\dfrac{\partial w}{\partial \alpha}}_{>0} > 0.$

A.4.3 Markdown ν_D

To determine the markdown $\nu_D = \frac{MR_D(q)}{MC_U(q)}$, we divide each term of the first-order condition given by (25) by $MC_U(q)$:

$$\alpha\left(\frac{MC_D(q)}{MC_U(q)}-1\right)\left(\frac{p(q)}{MC_U(q)}-1\right)+(1-\alpha)\left(\frac{MR_D(q)}{MC_U(q)}-\frac{MC_D(q)}{MC_U(q)}\right)\left(1-\frac{r(q)}{MC_U(q)}\right)=0. \quad (27)$$

We then use the following simplifications:

- $\frac{MC_D(q)}{MC_U(q)} = \frac{MC_U'(q)q+MC_U}{MC_U} = \left(\frac{1}{\varepsilon_{MC_U}}+1\right)$,
- $MR_D(q) = p(q)\left(1-\frac{1}{\varepsilon_p}\right) \Leftrightarrow \frac{p(q)}{MC_U(q)} = \frac{MR_D}{MC_U}\left(\frac{\varepsilon_p}{\varepsilon_p-1}\right)$,
- $MC_U = r'(q)q + r(q) = r(q)\left(\frac{1}{\varepsilon_r}+1\right) \Leftrightarrow r(q) = MC_U(q)\frac{\varepsilon_r}{\varepsilon_r+1}$,

to rewrite:

$$\begin{aligned}
&\alpha\left(\frac{1}{\varepsilon_{MC_U}}\right)\left(\frac{MR_D}{MC_U}\left(\frac{\varepsilon_p}{\varepsilon_p-1}\right)-1\right)+(1-\alpha)\left(\frac{MR_D}{MC_U}-\left(\frac{\varepsilon_{MC_U}+1}{\varepsilon_{MC_U}}\right)\left(\frac{1}{\varepsilon_r+1}\right)\right)=0\\
\Longleftrightarrow\quad &\frac{MR_D}{MC_U}\left(\alpha\left(\frac{1}{\varepsilon_{MC_U}}\right)\left(\frac{\varepsilon_p}{\varepsilon_p-1}\right)+(1-\alpha)\left(\frac{1}{\varepsilon_r+1}\right)\right)=\alpha\left(\frac{1}{\varepsilon_{MC_U}}\right)+(1-\alpha)\left(\frac{\varepsilon_{MC_U}+1}{\varepsilon_{MC_U}}\right)\left(\frac{1}{\varepsilon_r+1}\right)\\
\Longleftrightarrow\quad &\frac{MR_D}{MC_U}\left(\alpha\varepsilon_p(\varepsilon_r+1)+(1-\alpha)(\varepsilon_{MC_U})(\varepsilon_p-1)\right)=\alpha(\varepsilon_p-1)(\varepsilon_r+1)+(1-\alpha)(\varepsilon_{MC_U}+1)(\varepsilon_p-1)\\
\Longleftrightarrow\quad &\nu_D=\frac{MR_D}{MC_U}=\frac{\alpha(\varepsilon_p-1)(\varepsilon_r+1)+(1-\alpha)(\varepsilon_{MC_U}+1)(\varepsilon_p-1)}{\alpha\varepsilon_p(\varepsilon_r+1)+(1-\alpha)(\varepsilon_{MC_U})(\varepsilon_p-1)}.
\end{aligned}$$

When D is powerful, its markdown can be rewritten as:

$$\nu_D = \omega_U(\alpha)\frac{\varepsilon_{MC_U}+1}{\varepsilon_{MC_U}}+(1-\omega_U(\alpha))\frac{\varepsilon_p-1}{\varepsilon_p}, \quad (28)$$

where $\omega_U(\alpha) \equiv \frac{\alpha(\varepsilon_r+1)\varepsilon_p}{\alpha(\varepsilon_r+1)\varepsilon_p+(1-\alpha)(\varepsilon_p-1)\varepsilon_{MC_U}}$, with $\frac{\partial\omega_U(\alpha)}{\partial\alpha} < 0$. In the case we consider here, namely $0 < \alpha < \alpha_I$, we have $(0 <) \frac{\varepsilon_p}{\varepsilon_{MR_U}+\varepsilon_p} \leq \omega_U(\alpha) \leq 1$.[57] This weight $\omega_U(\alpha)$ is a function of α, and of demand primitives through ε_r and σ_r (as $\varepsilon_{MC_U} = \frac{\varepsilon_r+1}{\sigma_r+2}$).

As (24), (28) echoes expressions delivered by the exogenous right-to-manage models of Alviarez et al. (2023); Azkarate-Askasua and Zerecero (2022) and Wong (2023), whereby the bilateral markdown (or markup, see also Appendix A.3.3) is a weighted average between a monopoly (or oligopoly) markup term and a monopsony (or oligopsony) markdown term. Indeed, if $\alpha = 1$, then $\omega_U(\alpha) = 1$ and $\nu_D = \frac{\varepsilon_{MC_U}+1}{\varepsilon_{MC_U}}$, as D can make a take-it-or-leave-it offer to U. Similarly, if $\alpha = 0$ then $\omega_U(\alpha) = 0$ and (28) would state $\nu_D = \mu_D^{-1}$ and thus $r = w$ and $M_D = 1$. However, by sugbgame perfection criterion and short-side rule application, such equilibrium is ruled out in our framework, as U endogenously concedes the right-to-manage if becoming too powerful, i.e. when $\alpha > \alpha_I$. Instead, at this limit, i.e.,

[57] $\omega_U(\alpha) > 0$ holds when the demand function is supermodular, which we assumed above.

if $\alpha = \alpha_I = \frac{\varepsilon_p - 1}{\varepsilon_p + \varepsilon_r}$, then $\omega_U(\alpha_I) = \frac{\varepsilon_p}{\varepsilon_{MR_U} + \varepsilon_p}$, and $\nu_D = 1$, yielding the vertically integrated outcome q_I.

A.4.4 Set of Equilibria

Using the first order condition (12), we introduce the following function:

$$\Psi(q_D, \alpha) \equiv \alpha(MC_D(q_D) - MC_U(q_D))\pi_D(q_D) + (1-\alpha)(MR_D(q_D) - MC_D(q_D))\pi_U(q_D) = 0.$$

Under Assumption 3, applying the implicit function theorem, we have that

$$Sign(\frac{\partial q_D}{\partial \alpha}) = Sign(\frac{\partial \Psi(q_D, \alpha)}{\partial \alpha}),$$

and

$$\begin{aligned} \frac{\partial \Phi(q_D, \alpha)}{\partial \alpha} &= (MC_D(q_D) - MC_U(q_D))\pi_D(q_D) - (MR_D(q_D) - MC_D(q_D))\pi_U(q_D) \\ &= (MC_D(q_D) - MR_D(q_U))\pi_U(q_U) > 0. \end{aligned}$$

Therefore, the equilibrium quantity q_D increases in α. We know that $q_D(\alpha_I) = q_I$ when $\alpha = \alpha_I$ and therefore $q_D > q_I$ when $\alpha > \alpha_I$. As shown in Section 4.1, in that case $\tilde{q}_D(w) < \tilde{q}_U(w)$ and the initial assumption that $w = MC_U(q)$ no longer holds. The set of equilibria defined by (9) and $w_D = MC_U(q_D)$ thus only exists for $\alpha \in [0, \alpha_I]$.

A.5 Proof of Corollary 2

The proof succesively considers the case when U is powerful and when D is powerful.

When U is powerful

Note that $\pi_U(q) = [MC_U(q) - r(q)]q$. However, we have $\frac{\partial \pi_U}{\partial q} = MC_D(q) - MCU(q) > 0$ which is positive under Assumption 1.(i) and Assumption 1.(ii). Regarding the profit of D, we have $\pi_D(q) = [p(q) - MC_U(q)]q$ which is maximized when $MR_D(q) = MC_D(q)$ in $q_{\overline{\nu}}$. Here $\frac{\partial \pi_D}{\partial q} = MRD(q) - MRU(q) < 0$ which is negative for all $q > q_{\overline{\nu}}$ under Assumption 1.(i) and Assumption 1.(ii).

When D is powerful

Note that $\pi_U(q) = [MR_D(q) - r(q)]q$ is maximized in $q_{\overline{\mu}}$. However, we have $\frac{\partial \pi_U}{\partial q} = MRU(q) - MCU(q)$ which is negative under Assumption 1.(i) and Assumption 1.(ii) for all $q \geq q_{\overline{\mu}}$. Regarding the profit of D, we have $\pi_D(q) = [p(q) - MR_D(q)]q$. Here $\frac{\partial \pi_D}{\partial q} = MRD(q) - MRU(q) > 0$ which is positive for all $q > 0$.

A.6 Proof of Remark 1

The proof succesively considers the case when U is powerful and when D is powerful. Except if mentioned otherwise, all derivative signs used in the proof follow from Assumption 1 and Corollary 1.

When U is powerful

(i) As $\mathcal{M} \equiv \frac{p}{r}$, with $\frac{\partial p}{\partial \alpha} = \underbrace{\frac{\partial p}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} < 0$ and $\frac{\partial r}{\partial \alpha} = \underbrace{\frac{\partial r}{\partial q}}_{+} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} > 0$, we necessarily have $\frac{d\mathcal{M}}{d\alpha} < 0$.

(ii) As $\nu_D \equiv \frac{MC_U}{w} = \frac{MC_U}{MR_D}$, with $\frac{\partial MC_U}{\partial \alpha} = \underbrace{\frac{\partial MC_U}{\partial q}}_{+} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} > 0$ and $\frac{\partial MR_D}{\partial \alpha} = \underbrace{\frac{\partial MR_D}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} < 0$, we necessarily have $\frac{d\nu_D}{d\alpha} < 0$. Again similarly, as $M_D \equiv \frac{p}{w} = \frac{p}{MC_U}$, with $\frac{\partial p}{\partial \alpha} = \underbrace{\frac{\partial p}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} < 0$ and $\frac{\partial MC_U}{\partial \alpha} = \underbrace{\frac{\partial MC_U}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} > 0$, we necessarily have $\frac{dM_D}{d\alpha} < 0$.

(iii) In the subconvex demand and supply case, where $\frac{\partial \varepsilon_f}{\partial q} < 0$ for every $f \in \{p, r\}$, we have $\frac{dM_U}{d\alpha} = \frac{d\nu_U}{d\alpha} = \underbrace{\frac{\partial \nu_U}{\partial \varepsilon_r}}_{-} \underbrace{\frac{\partial \varepsilon_r}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} > 0$, and similarly, $\frac{d\mu_D}{d\alpha} = \underbrace{\frac{\partial \mu_D}{\partial \varepsilon_p}}_{-} \underbrace{\frac{\partial \varepsilon_p}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{+} > 0$. Derivations follow a similar logic for the superconvex case, where $\frac{\partial \varepsilon_f}{\partial q} > 0$, and CES case where $\frac{\partial \varepsilon_f}{\partial q} = 0$, for every $f \in \{p, r\}$.

When D is powerful

(i) As $\mathcal{M} \equiv \frac{p}{r}$, with $\frac{\partial p}{\partial \alpha} = \underbrace{\frac{\partial p}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} > 0$ and $\frac{\partial r}{\partial \alpha} = \underbrace{\frac{\partial r}{\partial q}}_{+} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} < 0$, we necessarily have $\frac{d\mathcal{M}}{d\alpha} > 0$.

(ii) As $\mu_U \equiv \frac{w}{MC_U} = \frac{MR_D}{MC_U}$, with $\frac{\partial MR_D}{\partial \alpha} = \underbrace{\frac{\partial MR_D}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} > 0$ and $\frac{\partial MC_U}{\partial \alpha} = \underbrace{\frac{\partial MC_U}{\partial q}}_{+} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} < 0$, we necessarily have $\frac{d\mu_U}{d\alpha} > 0$. Again similarly, as $M_U \equiv \frac{w}{r} = \frac{MR_D}{r}$, with $\frac{\partial MR_D}{\partial \alpha} = \underbrace{\frac{\partial MR_D}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} > 0$ and $\frac{\partial r}{\partial \alpha} = \underbrace{\frac{\partial r}{\partial q}}_{+} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} < 0$, we necessarily have $\frac{dM_U}{d\alpha} > 0$.

(iii) In the subconvex demand and supply case, where $\frac{\partial \varepsilon_f}{\partial q} < 0$ for every $f \in \{p, r\}$, we have $\frac{dM_D}{d\alpha} = \frac{d\mu_D}{d\alpha} = \underbrace{\frac{\partial \mu_D}{\partial \varepsilon_p}}_{-} \underbrace{\frac{\partial \varepsilon_p}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} < 0$, and similarly, $\frac{d\nu_U}{d\alpha} = \underbrace{\frac{\partial \nu_U}{\partial \varepsilon_r}}_{-} \underbrace{\frac{\partial \varepsilon_r}{\partial q}}_{-} \underbrace{\frac{\partial q}{\partial \alpha}}_{-} < 0$. Derivations follow a similar logic for the superconvex case, where $\frac{\partial \varepsilon_f}{\partial q} > 0$, and CES case where $\frac{\partial \varepsilon_f}{\partial q} = 0$, for every $f \in \{p, r\}$.

A.7 Markup and Markdown Expressions

A.7.1 Vertical Integration and Take-it-or-leave-it Offers

Although we consider here the vertically integrated firm I, the proofs contained in this Appendix A.7.1 directly extend to any firm unilaterally choosing its prices or quantity facing demand and supply curves.

Here, I faces an increasing inverse supply curve $r(q)$ and a decreasing inverse demand curve $p(q)$.

I's profit maximization gives the equilibrium:

$$\max_q \pi_I(q) = p(q)q - r(q)q.$$

The equilibrium quantity q_I is such that:

$$MR_I(q_I) = MC_I(q_I).$$

Markup. To implement our markup definition, we consider the following hypothetical case: I can sell an infinitesimal quantity ϵ at a given price $\overline{p}$, different from the price p at which I sells other output units. Thus, I does not incur any revenue loss on other units when selling ϵ. In what follows, we determine the minimum price $\hat{p}$ such that I supplies the marginal equilibrium output unit leading to the equilibrium quantity $q^* + \varepsilon$ with $\epsilon \to 0$. Formally, we look for $\hat{p}$, the minimal value of $\overline{p}$ such that:

$$\overline{\pi}_I(q_I, \epsilon, \overline{p}) \geq \pi_I(q_I) \tag{29}$$

where $\overline{\pi}_I(q_I, \epsilon, \overline{p})$ is I's profit for a quantity $q^* + \varepsilon$ in the hypothetical case, defined by:

$$\overline{\pi}_I(q_I, \epsilon, \overline{p}) \equiv p(q_I)(q_I) + \overline{p}\epsilon - r(q_I + \epsilon)(q_I + \epsilon). \tag{30}$$

Using (30), the inequality (29) can thus be rewritten as:

$$\overline{p} \geq \frac{r(q_I + \epsilon) - r(q_I)}{\epsilon} q_I + r(q_I + \epsilon). \tag{31}$$

Examining the limit of the right-hand side of (31) when $\epsilon \to 0$, we obtain:

$$lim_{\epsilon \to 0} \frac{r(q_I + \epsilon) - r(q_I)}{\epsilon} q_I + r(q_I + \epsilon) = r'(q_I)q_I + r(q_I) = MC_I(q_I).$$

The minimum price at which I would supply the marginal output unit is thus $\hat{p} = MC_I(q_I)$. It follows by definition that $\mu_I(q_I) \equiv \frac{p}{\hat{p}} = \frac{p(q_I)}{MC_I(q_I)}$.

Markdown. To implement our markdown definition, we consider the following hypothetical case: I can buy an infinitesimal quantity ϵ at a given price $\overline{r}$, different from the price r at which I buys other input units. Thus, I does not incur any cost increase on other units when buying ϵ. In what follows, we determine the maximum price $\hat{r}$ such that I purchases the marginal input unit leading to the equilibrium quantity $q^* + \varepsilon$ with $\epsilon \to 0$. Formally, we look for $\hat{r}$, the maximal value of $\overline{r}$ such that:

$$\overline{\pi}_I(q_I, \epsilon, \overline{r}) \geq \pi_I(q_I), \tag{32}$$

where $\overline{\pi}_I(q_I, \epsilon, \overline{r})$ is I's profit for a quantity $q^* + \varepsilon$ in the hypothetical case, defined by:

$$\overline{\pi}_I(q_I, \epsilon, \overline{r}) \equiv p(q_I + \epsilon)(q_I + \epsilon) - r(q_I)(q_I) - \overline{r}\epsilon. \tag{33}$$

Using (33), the inequality (32) can thus be rewritten as:

$$\frac{p(q_I + \epsilon) - p(q_I)}{\epsilon} q_I + p(q_I + \epsilon) \geq \overline{r}. \tag{34}$$

Examining the limit of the left-hand side of (34) when $\epsilon \to 0$, we obtain:

$$lim_{\epsilon \to 0} \frac{p(q_I + \epsilon) - p(q_I)}{\epsilon} q_I + p(q_I + \epsilon) = p'(q_I)q_I + p(q_I) = MR_I(q_I).$$

The maximum price at which I would purchase the marginal input unit is thus $\hat{r} = MR_I(q_I)$. It follows by definition that $\nu_I(q_I) \equiv \frac{\hat{r}}{r} = \frac{MR_I(q_I)}{r(q_I)}$.

A.7.2 Bargaining

When U and D bargain over a linear tariff w, the equilibrium value of the Nash product $N(q^*)$ is the following:

$$N(q^*) \equiv \max_q \pi_U(q)^\alpha \pi_D(q)^{(1-\alpha)} \quad \text{s.t} \quad w(q) = \begin{cases} MC_U(q) \text{ if } \tilde{q}_u(w) < \tilde{q}_D(w) \\ MR_D(q) \text{ if } \tilde{q}_u(w) > \tilde{q}_D(w) \end{cases} \tag{35}$$

where $\pi_U(q) = (w(q) - r(q))q$ and $\pi_D(q) = (p(q) - w(q))q$.

Markup of D. To implement our markup definition, we consider the following hypothetical case: D can sell an infinitesimal quantity ϵ at a given price $\overline{p}$, different from the price p at which D sells other output units. Thus, D does not incur any revenue loss on other units when selling ϵ. In what follows, we determine the minimum price $\hat{p}$ such that D supplies the marginal output unit leading to the equilibrium quantity $q^* + \varepsilon$ with $\epsilon \to 0$. Formally, we look for $\hat{p}$, the minimal value of $\overline{p}$ such that:

$$\overline{N}(q^*, \epsilon, \overline{p}) \geq N(q^*), \tag{36}$$

where $\overline{N}(q^*, \epsilon, \overline{p})$ corresponds to the value of the Nash product for a quantity $q^* + \varepsilon$ in the hypothetical case, defined by:

$$\overline{N}(q^*, \epsilon, \overline{p}) \equiv [\underbrace{(w(q^* + \epsilon)(q^* + \epsilon) - r(q^* + \epsilon)(q^* + \epsilon)}_{\pi_U(q^* + \epsilon)}]^\alpha [\underbrace{(p(q^*)(q^*) + \overline{p}\epsilon - w(q^* + \epsilon)(q^* + \epsilon))}_{\overline{\pi}_{D(q^*, \epsilon, \overline{p})}}]^{1-\alpha}. \tag{37}$$

Applying the Taylor formula to the equilibrium Nash product in (35) yields:

$$N(q^*+\epsilon) = N(q^*) + \frac{\partial N(q^*)}{\partial q}\epsilon + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^*)}{n!} = N(q^*) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^*)}{n!}. \tag{38}$$

Using (38), we can rewrite (36) as:

$$\overline{N}(q^*, \epsilon, \overline{p}) \geq N(q^*+\epsilon) - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^*)}{n!},$$

and then use (37) to obtain:

$$\pi_U(q^*+\epsilon)^\alpha \overline{\pi}_D(q^*, \epsilon, \overline{p})^{1-\alpha} \geq \pi_U(q^*+\epsilon)^\alpha \pi_D(q^*+\epsilon)^{1-\alpha} - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^*)}{n!}. \tag{39}$$

Defining $R_D(q^*+\epsilon) \equiv p(q^*+\epsilon)(q^*+\epsilon)$ and applying the Taylor formula yields:

$$R_D(q^*+\epsilon) = p(q^*)q^* + \underbrace{[p'(q^*)q^* + p(q^*)]}_{MR_D(q^*)}\epsilon + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{R_D^{(n)}(q^*)}{n!}. \tag{40}$$

Using (40), and $\sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^*)}{n!} \to 0$ and $\sum_{n\geq 2}^{\infty} \epsilon^n \frac{R^{(n)}(q^*)}{n!} \to 0$ as $\epsilon \to 0$, (39) boils down to:

$$\begin{aligned}
& [p(q^*)q^* + \overline{p}\epsilon - w(q^*+\epsilon)(q^*+\epsilon)]^{1-\alpha} \geq [p(q^*)q^* + MR_D(q^*)\epsilon - w(q^*+\epsilon)(q^*+\epsilon)]^{1-\alpha} \\
\Leftrightarrow \quad & p(q^*)q^* + \overline{p}\epsilon - w(q^*+\epsilon)(q^*+\epsilon) \geq p(q^*)q^* + MR_D(q^*)\epsilon - w(q^*+\epsilon)(q^*+\epsilon) \\
\Leftrightarrow \quad & \overline{p}\epsilon \geq MR_D(q^*)\epsilon \\
\Leftrightarrow \quad & \overline{p} \geq MR_D(q^*).
\end{aligned}$$

The minimum price at which D supplies the marginal output unit is thus $\hat{p} = MR_D(q^*)$. It follows by definition that $\mu_D(q^*) \equiv \frac{p}{\hat{p}} = \frac{p(q^*)}{MR_D(q^*)}$.

Markup of U. To implement our markup definition, we consider the following hypothetical case: U can sell an infinitesimal quantity ϵ to D at a given price $\overline{w}$, different from the price w at which U sells other output units to D. Thus, when selling ϵ, U does not incur any revenue loss on other units.[58] In what follows, we determine the minimum price $\hat{w}$ such that U supplies the *marginal* output unit leading to the equilibrium quantity $q^* + \varepsilon$ with $\epsilon \to 0$. Formally, we look for $\hat{w}$, the minimal value of

[58] In the bargaining, U and D share a joint profit, and trading a marginal unit affect the profit of both firms. The revenue loss on inframarginal units supported by U when selling an additional unit is therefore twofold, via its direct and indirect (through internalization of D's profit) effects on U's profit. Consistent with our markup definition, the hypothetical case we consider fully eliminates such a revenue loss.

$\overline{w}$ such that:

$$\overline{N}(q^*, \epsilon, \overline{w}) \geq N(q^*), \tag{41}$$

Where $\overline{N}(q^*, \epsilon, \overline{p})$ corresponds to the value of the Nash product for a quantity $q^*+\varepsilon$ in the hypothetical case, defined by:

$$\overline{N}(q^*, \epsilon, \overline{w}) = (w(q^*)q^* + \overline{w}\epsilon - r(q^* + \epsilon)(q^* + \epsilon))^{\alpha} (p(q^* + \epsilon)(q^* + \epsilon) - w(q^*)q^* - \overline{w}\epsilon)^{1-\alpha}. \tag{42}$$

Substituting in (41), we obtain:

$$(w(q^*)q^* + \overline{w}\epsilon - r(q^* + \epsilon)(q^* + \epsilon))^{\alpha} (p(q^* + \epsilon)(q^* + \epsilon) - w(q^*)q^* - \overline{w}\epsilon)^{1-\alpha} \geq N(q^*). \tag{43}$$

Applying the Taylor formula to $C_U(q^* + \epsilon) \equiv r(q^* + \epsilon)(q^* + \epsilon)$ yields:

$$r(q^* + \epsilon)(q^* + \epsilon) = r(q^*)q^* + \underbrace{[r'(q^*)q^* + r(q^*)]}_{MC_U(q^*)} \epsilon + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{C_U^{(n)}(q^*)}{n!}. \tag{44}$$

Using Taylor formulas applied to R_D (40) and C_U (44), we can rewrite (43) as:

$$\left[(\overline{w} - MC_U(q^*))\epsilon + \pi_U(q^*) - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{C_U^{(n)}(q^*)}{n!}\right]^{\alpha} \left[(MR_D(q^*) - \overline{w})\epsilon + \pi_D(q^*) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{R_D^{(n)}(q^*)}{n!}\right]^{1-\alpha} \geq N(q^*).$$

Further simplifying (43) requires distinguinshing two cases depending on the value of $\overline{w}$, specifically whether $\tilde{q}_D(\overline{w}) \leq \tilde{q}_U(\overline{w})$ or $\tilde{q}_D(\overline{w}) \geq \tilde{q}_U(\overline{w})$.

Assume first that $\overline{w}$ such that $q^* + \epsilon = \tilde{q}_D(\overline{w}) \leq \tilde{q}_U(\overline{w})$. Stage 2 of the game implies that $\overline{w} = MR_D(q^* + \epsilon)$. Applying the Taylor formula to $MR_D(q^* + \epsilon)$ yields:

$$\begin{aligned} MR_D(q^* + \epsilon) &= MR_D(q^*) + \epsilon MR_D'(q^*) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{MR_D^{(n)}(q^*)}{n!} \\ \Leftrightarrow \quad MR_D(q^*) - \overline{w} &= -\epsilon MR_D'(q^*) - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{R_D^{(n)}(q^*)}{n!}. \end{aligned} \tag{45}$$

Substituting (45) in (44), and using $N(q^*) = \pi_U(q^*)^{\alpha} \pi_D(q^*)^{1-\alpha}$ yields:

$$(\epsilon(\overline{w} - MC_U(q^*)) + \pi_U(q^*))^{\alpha} \left(-\epsilon^2 MR_D'(q^*) - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{R_D^{(n)}(q^*)}{n!} + \pi_D(q^*)\right)^{1-\alpha} \geq \pi_U(q^*)^{\alpha} \pi_D(q^*)^{1-\alpha}.$$

Considering $\epsilon \to 0$, we obtain:

$$\begin{aligned} &(\epsilon(\overline{w} - MC_U(q^*)) + \pi_U(q^*))^{\alpha}\pi_D(q^*)^{1-\alpha} \geq \pi_U(q^*)^{\alpha}\pi_D(q^*)^{1-\alpha} \\ \Leftrightarrow \quad &\overline{w} \geq MC_U(q^*). \end{aligned} \tag{46}$$

If the minimum price $\hat{w}$ at which the marginal unit is supplied by U is such that $q^* + \epsilon = \tilde{q}_U(\hat{w}) \leq \tilde{q}_D(\hat{w})$, then $\hat{w} = MC_U(q^*)$.

Assume now that $\overline{w}$ such that $q^* + \epsilon = \tilde{q}_U(\overline{w}) \leq \tilde{q}_D(\overline{w})$. Stage 2 of the game implies that $\overline{w} = MC_U(q^* + \epsilon)$. Applying the Taylor formula to $MC_U(q^* + \epsilon)$ yields:

$$\begin{aligned} &MC_U(q^* + \epsilon) = MC_U(q^*) + \epsilon MC_U'(q^*) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{MC_U^{(n)}(q^*)}{n!} \\ \Leftrightarrow \quad &\overline{w} - MC_U(q^*) = \epsilon MC_U'(q^*) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{MC_U^{(n)}(q^*)}{n!} \end{aligned} \tag{47}$$

Substituting (47) in (44) yields:

$$\left(\epsilon^2 MC_U'(q^*) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{MC_U^{(n)}(q^*)}{n!} + \pi_U(q^*)\right)^{\alpha} (\epsilon(MR_D(q^*) - \overline{w}) + \pi_D(q^*))^{1-\alpha} \geq \pi_U(q^*)^{\alpha}\pi_D(q^*)^{1-\alpha}. \tag{48}$$

Considering $\epsilon \to 0$, we obtain:

$$\begin{aligned} &(\epsilon(MR_D(q^*) - \overline{w}) + \pi_D(q^*))^{1-\alpha}\,\pi_U(q^*)^{\alpha} \geq \pi_D(q^*)^{1-\alpha}\pi_U(q^*)^{\alpha} \\ \Leftrightarrow \quad &MR_D(q) \geq \overline{w}. \end{aligned} \tag{49}$$

If the minimum price $\hat{w}$ at which U would supply the marginal unit is such that $q^* + \epsilon = \tilde{q}_U(\hat{w}) \geq \tilde{q}_D(\hat{w})$, then (49) yields an upper bound to $\overline{w}$. $\hat{w}$ is thus given by $\hat{w} = w(q^*) = MC_U(q^*)$.

In all cases (whether $q^* + \epsilon = \tilde{q}_U(\hat{w}) \leq \tilde{q}_D(\hat{w})$ or $q^* + \epsilon = \tilde{q}_U(\hat{w}) \geq \tilde{q}_D(\hat{w})$), the minimum price at which U would supply the marginal output unit to D is thus $\hat{w} = MC_U(q^*)$. It follows by definition that $\mu_U(q^*) \equiv \frac{w}{\hat{w}} = \frac{w(q^*)}{MC_U(q^*)}$.

Markdown of *D*. To implement our markdown definition, we consider the following hypothetical case: D can buy an infinitesimal quantity ϵ at a given price $\overline{w}$, different from the price w at which D buys other input units. Thus, when buying ϵ, D does not incur any cost increase on other units.[59]

[59] In the bargaining, U and D share a joint profit, and trading a marginal unit affect the profit of both firms. The cost increase on inframarginal units supported by D when purchasing an additional unit is therefore twofold, via its direct and indirect (through internalization of U's profit) effects on D's profit. Consistent with our markdown definition, the hypothetical case we consider fully eliminates such a cost increase.

In what follows, we determine the maximum price $\hat{w}$ such that D purchases the *marginal* input unit leading to the equilibrium quantity $q^* + \varepsilon$ with $\epsilon \to 0$. Formally, we look for $\hat{w}$, the maximal value of $\overline{w}$ such that (41) is satisfied. Indeed, it is worth noting, as w is both the output price of U and the input price of D, the Nash product $\overline{N}(q^*, \epsilon, \overline{w})$ to be considered for characterizing D's markdown is given by (42) appearing in U's markup characterization, with the key difference that, consistent with our markdown definition, we here look for the maximal (as opposed to minimal) value of $\overline{w}$ which satisfies (41).

Again, two cases must be treated depending on the value of $\overline{w}$, specifically whether $\overline{w}$ is such that $\tilde{q}_D(\overline{w}) \leq \tilde{q}_U(\overline{w})$ or $\tilde{q}_D(\overline{w}) \geq \tilde{q}_U(\overline{w})$.

Assume first that $q^* + \epsilon = \tilde{q}_D(\overline{w}) \leq \tilde{q}_U(\overline{w})$. Stage 2 of the game implies that $\overline{w} = MR_D(q^* + \epsilon)$, and (41) is satisfied when:

$$\overline{w} \geq MC_U(q^*). \tag{50}$$

If the maximum price $\hat{w}$ at which the marginal input unit is purchased by D is such that $q^* + \epsilon = \tilde{q}_D(\overline{w}) \leq \tilde{q}_U(\overline{w})$, then (50) yields a lower bound to $\overline{w}$. $\hat{w}$ is thus given by $\hat{w} = w(q^*) = MR_D(q^*)$.

Assume now that $q^* + \epsilon = \tilde{q}_U(\overline{w}) \leq \tilde{q}_D(\overline{w})$. Stage 2 of the game implies that $\overline{w} = MR_D(q^* + \epsilon)$, and (41) is satisfied when:

$$\overline{w} \leq MR_D(q^*).$$

If the maximum price $\hat{w}$ at which the marginal unit is purchased by D is such that $q^* + \epsilon = \tilde{q}_U(\overline{w}) \leq \tilde{q}_D(\overline{w})$, then $\hat{w} = MR_D(q^*)$.

In all cases (whether $q^* + \epsilon = \tilde{q}_U(\hat{w}) \leq \tilde{q}_D(\hat{w})$ or $q^* + \epsilon = \tilde{q}_U(\hat{w}) \geq \tilde{q}_D(\hat{w})$), the maximum price at which D would purchase the marginal unit is thus $\hat{w} = MR_D(q^*)$. It follows by definition that $\nu_D(q^*) \equiv \frac{\hat{w}}{w} = \frac{MR_D(q^*)}{w(q^*)}$.

Markdown of *U*. To implement our markdown definition, we consider the following hypothetical case: U can buy an infinitesimal quantity ϵ at a given price $\overline{r}$, different from the price r at which U buys other input units. Thus, when buying ϵ, U does not incur any cost increase on other units. In what follows, we determine the maximum price $\hat{r}$ such that U purchases the *marginal* input unit leading to the equilibrium quantity $q^* + \varepsilon$ with $\epsilon \to 0$. Formally, we look for $\hat{r}$, the maximal value of $\overline{r}$ such that:

$$\overline{N}(q^*, \epsilon, \overline{r}) \geq N(q^*). \tag{51}$$

By definition, we have:

$$\overline{N}(q^*, \epsilon, \overline{r}) = [\underbrace{(w(q^* + \epsilon)(q^* + \epsilon) - r(q^*)(q^*) - \overline{r}\epsilon}_{\overline{\pi}_U(q^*, \epsilon, \overline{r})}]^{\alpha} [\underbrace{(p(q^* + \epsilon)(q^* + \epsilon) - w(q^* + \epsilon)(q^* + \epsilon))}_{\pi_D(q^* + \epsilon)}]^{1-\alpha}. \tag{52}$$

The derivation of $\hat{r}$ is similar to that of $\hat{p}$ in D's markup characterization of Appendix A.7.2. Using again Taylor formulas and considering $\epsilon \to 0$, one can show that (51) boils down to:

$$\overline{r} \leq MC_U(q^*).$$

The maximum price at which U would purchase the marginal input unit is thus $\hat{r} = MC_U(q^*)$. It follows by definition that $\nu_U(q^*) \equiv \frac{\hat{r}}{r} = \frac{MC_U(q^*)}{r(q^*)}$.

B Microfoundation

In this Appendix, we demonstrate that the equilibrium outcome of our model introduced in Section 4 coincides with the subgame perfect Nash equilibrium of a variant of the non-cooperative game developed in Rey and Vergé (2020). The next section introduces this game, and subsequent sections solve it, proceeding backward.

B.1 Timing

We consider that U and D play the following game:

- **Stage 1: Bargaining** The wholesale price w is determined through a bilateral negotiation between U and D according to the following protocol.

  1.1 U makes a take-it-or-leave-it (TIOLI) offer to D, which either accepts or rejects.

  1.2 If D rejects the offer, Nature selects one side to make an ultimate TIOLI offer. U is selected with probability ϕ and D with probability $1 - \phi$

  1.3 The selected firm makes the ultimate TIOLI offer to its counterpart, which either accepts or rejects.

- **Stage 2: Input and retail price setting**

  D sets the retail price p, and U sets the input price r, simultaneously. Given the input quantity that U can procure, D purchases from U to meet consumer demand.

We look for the subgame perfect Nash equilibrium of this two-stage game. Given $\phi \in [0, 1]$ and $\alpha \in [0, 1]$, we show that there exists $\phi \in [0, 1]$ such that the equilibrium outcome of the model developed in Section 4 coincides with the subgame perfect Nash equilibrium of the above noncooperative game.

B.2 Resolution

We proceed backward to look for the subgame perfect Nash equilibrium of our two-stage game. Stage 2 is solved in Appendix B.2.1. Stage 1.3 when U makes the (ultimate) take-it-or-leave-it offer and when D makes the (ultimate) take-it-or-leave-it offer are considered in Appendix B.2.2 and B.2.3, respectively. Stage 1.2 involves no strategic player choices, so we then focus on Stage 1.1 in Appendix B.2.4.

B.2.1 Stage 2

As long as input supply and consumer demand are not perfectly elastic, it is worth noting that the input price r determines the maximum quantity that U can procure from its input suppliers, while the retail price p determines the maximum quantity that D will purchase from U. Accordingly, D internalizes that the maximum quantity it can purchase from U is constrained by r.[60] Similarly, U recognizes that the maximum quantity it can sell to D is limited by p.[61] Hence, given w, each firm sets its profit-maximizing price while anticipating the pricing decision of the other. Formally, anticipating that $r = r^a$, D's maximization problem is as follows:

$$\max_{p} \ (q_D(p) - w)\, q_D(p) \quad \text{subject to} \quad q_D(p) \leq q_U(r^a) \tag{53}$$

where the constraint reflects that D cannot sell more quantity than what U is able to procure at r^a. An interior solution to (53) arises when $q_D(\tilde{p}) \leq q_U(r^a)$, where $\tilde{p}$ satisfies the following first-order condition:

$$MR_D(q_D(\tilde{p})) = w$$

Otherwise we have a corner solution where $q_D(\check{p}) = q_U(r^a)$, implying that $MR_D(q_D(\check{p})) > w$ as MR_D is decreasing in quantity (Assumption 1). Similarly, anticipating that $p = p^a$, U's maximization problem is as follows:

$$\max_{r} \ (w - q_U(r))\, q_U(r) \quad \text{subject to} \quad q_U(r) \leq q_D(p^a) \tag{54}$$

where the constraint reflects that U cannot sell more quantity than what D is willing to purchase to meet consumer demand at p^a. Again, an interior solution to (54) arises when $q_U(\tilde{r}) \leq q_D(p^a)$, where $\tilde{r}$ satisfies the following first-order condition:

$$MC_U(q_U(\tilde{r})) = w$$

[60] For instance, if both p and r are low, U may be unable to meet D's demand.

[61] For instance, if both p and r are high, U may be able to procure more quantity than what D is willing to purchase.

Otherwise we have a corner solution where $q_U(\check{r}) = q_D(p^a)$, implying that $MC_U(q_U(\check{r})) < w$ as MR_C is increasing in quantity (Assumption 1).

As in Section 4.1, there exists a multiplicity of Nash equilibria. For instance, if D believes that U will set an input price r^a such that $q_U(r^a) \leq q_D(\tilde{p})$, its best response is to set $\check{p} \leq \tilde{p}$ so that $q_D(\check{p}) = q_U(r^a)$. Similarly, if U believes that D will set a retail price p^a such that $q_U(\tilde{r}) \geq q_D(p^a)$, its best response is to set $\check{r} \geq \tilde{r}$ so that $q_U(\check{r}) = q_D(p^a)$. Hence, any strategy profile $(\check{p}, \check{r})$ satisfying $q_D(\check{p}) = q_U(\check{r})$, with $\check{p} \leq \tilde{p}$ and $\check{r} \geq \tilde{r}$, constitutes a Nash equilibrium. However, it is straightforward to verify that any such equilibrium with $\check{p} < \tilde{p}$ and $\check{r} > \tilde{r}$ is Pareto dominated by equilibria in which at least one firm sets its unconstrained optimal price (i.e., $\tilde{p}$ or $\tilde{r}$). In any such equilibrium, the quantity traded is given by $q = \min\{q_D(\tilde{p}), q_U(\tilde{r})\}$, which coincides with the equilibrium outcome in Lemma 1 of Section 4.1.

Remark. Instead of considering that U and D set their prices simultaneously, one can suppose a sequence of play where one firm chooses its profit-maximizing price before the other. For instance, consider the following timing:

2.1 Given w, U chooses the input price r.

2.2 Given w and r, D sets the output price p.

Proceeding backward, it can be shown that there exists a unique equilibrium where the traded quantity coincides with that described in Lemma 1 of Section 4.1.

B.2.2 Stage 1.3: Take-it-or-Leave-it Offer from U

Proceeding backward, we now solve Stage 1.3 where U makes a take-it-or-leave-it offer to D. Note also that this section solves Stage 1 of our primary framework introduced in Section 4 when U has all the bargaining power (i.e. $\alpha = 1$), as both problems coincide.

Lemma 2 *When U makes a take-it-or-leave-it to D, the equilibrium wholesale price is* $\overline{w} = MC_U(\overline{q}) \left(\frac{\varepsilon_{MR_D(\overline{q})}}{1 - \varepsilon_{MR_D(\overline{q})}} \right)$ *with $\overline{q} \equiv q(\overline{w})$ the equilibrium quantity and $\overline{w} = MR_D(\overline{q}) > w_I$.*

Proof. The maximization problem faced by U is the following:

$$\overline{w} \equiv \underset{w}{\text{argmax}}\ \pi_U(w) \quad \text{subject to } w = \begin{cases} MC_U(q(w)), \text{ for } w \leq w_I \\ MR_D(q(w)), \text{ for } w \geq w_I \end{cases}$$

By definition,

$$\pi_U(w) = wq(w) - r(q(w))q(w),$$

and

$$\frac{\partial \pi_U}{\partial w} = q(w) + q'(w)[w - MC_U(q)].$$

Assuming first that $w \leq w_I$, we have from Stage 2 that $w = MC_U(q)$, and thus that $\dfrac{\partial \pi_U}{\partial w} = q(w) > 0$, implying that $w \leq w_I$ is not profit-maximizing for U. The optimal w chosen by U is thus such that $w > w_I$, which implies that $w(q) = MR_D(q)$ from Stage 2.[62]

U maximization problem thus simplifies to:

$$\overline{w} \equiv \underset{w}{\text{argmax}}\ wq(w) - r(q(w))q(w) \quad \text{subject to } w = MR_D(q)$$

The first-order condition yields:

$$\overline{w}\left(1 + \frac{q(\overline{w})}{q'(\overline{w})\overline{w}}\right) = MC_U(q(\overline{w})). \tag{55}$$

Using the constraint and $\overline{q} \equiv q(\overline{w})$, the first-order condition can be rewritten:

$$\overline{w} = MC_U(\overline{q})\left(\frac{\varepsilon_{MR_D}(\overline{q})}{\varepsilon_{MR_D}(\overline{q}) - 1}\right),$$

with $\overline{w} = MR_D(\overline{q}) > w_I$.

The second-order condition yields:

$$\frac{\partial^2 \pi_U(w)}{\partial w^2} = q''(w)[w(q) - MC_U(q)] + q'(w)[2 - q'(w)MC'_U(q)] < 0$$

$$\iff \sigma_{MR_D} < \frac{2 - \frac{MC'_U}{MR'_D}}{\left(\frac{MC_U}{MR_D} + 1\right)\epsilon_{MR_D}} \tag{56}$$

where we used that $w(q) = MR_D(q)$, and thus $q'(w) = \frac{1}{MR'_D(q)} < 0$ and $q''(w) = \frac{\sigma_{MR_D}}{q[MR'_D(q)]^2}$. We also have $\varepsilon_{MR_D} \equiv \frac{MR_D(q)}{q|MR'_D(q)|}$ and $\sigma_{MR_D} \equiv \frac{qMR''_D(q)}{|MR'_D(q)|}$ as defined in the main text. Using (55), which implies $\left(\frac{MC_U}{MR_D} + 1\right)\epsilon_{MR_D} = 1$, (56) simplifies to:

$$\sigma_{MR_D} < 2 - \frac{MC'_U}{MR'_D}.$$

As $MC'_U(q) > 0$ and $MR'_D(q) < 0$, it is straightforward that Assumption 3, which stipulates that $\sigma_{MR_D} < 2$, is sufficient to guarantee the second-order condition validity.[63] ■

In this extreme case, the expression of the equilibrium markup of U simplifies, so that:[64]

$$\mu_U \equiv \frac{w(q^*)}{MC_U(q^*)} = \frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D} - 1} = \frac{\varepsilon_p - 1}{\varepsilon_p + \sigma_p - 3}.$$

[62] Stage 2 results are summarized in Lemma 1.

[63] Assumption 3 imposes that U's marginal revenue is decreasing when making a take-it-or-leave-it offer to D. It guarantees U's profit concavity even if upstream supply is perfectly elastic.

[64] Apart from the endogenous quantity adjustment, other markup and markdown expressions in terms of primitives remain similar in this polar case to the ones derived in Proposition 3.

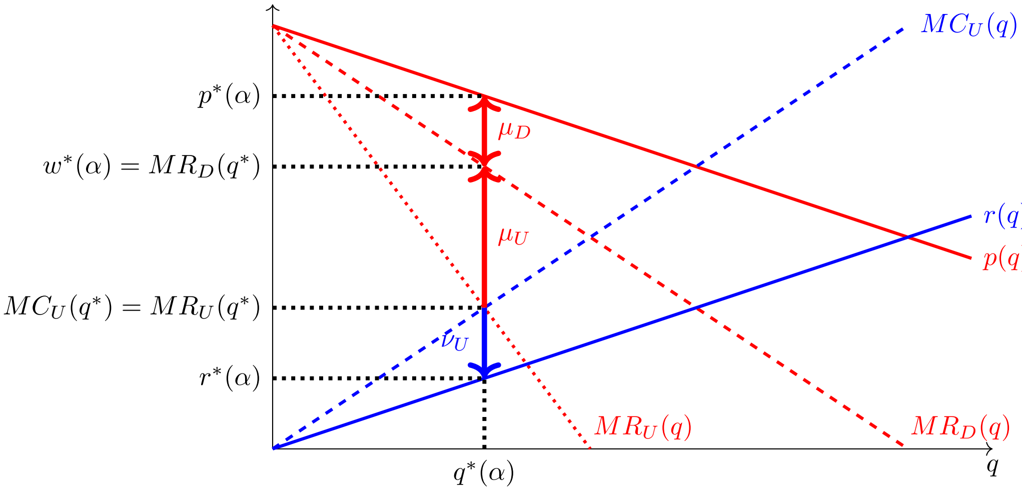


Figure 9: Take-it-or-Leave-it Offer from U ($\alpha = 1$).

Here, U makes a take-it-or-leave-it offer to D, and can thus equalize its marginal revenue and marginal cost, as visible in Figure 9. As a result, the markup of U only depends on demand primitives through $(\varepsilon_p, \sigma_p)$. It contrasts with the intermediate case where U is $\alpha_I < \alpha < 1$, and where the supply elasticity (ε_r) affects the markup of U, as D exercises countervailing power via the Nash-bargaining.

B.2.3 Stage 1.3: Take-it-or-Leave-it Offer from D

We now solve Stage 1.3 where D makes a take-it-or-leave-it offer to U. Note also that this section solves Stage 1 of our primary framework introduced in Section 4 when D has all the bargaining power (i.e. $\alpha = 0$), as both problems coincide.

Lemma 3 *When D makes a take-it-or-leave-it to U, the equilibrium wholesale price is $\underline{w} = MR_D(\underline{q})\left(\frac{\varepsilon_{MC_U(\underline{q})}}{1+\varepsilon_{MC_U(\underline{q})}}\right)$ with $\underline{q} \equiv q(\underline{w})$ the equilibrium quantity and $\underline{w} = MC_U(\underline{q}) < w_I$.*

Proof. The maximization problem faced by D is the following:

$$\underline{w} \equiv \underset{w}{\text{argmax}}\ \pi_D(w) \quad \text{subject to } w = \begin{cases} MC_U(q(w)), \text{ for } w \leq w_I \\ MR_D(q(w)), \text{ for } w \geq w_I \end{cases}$$

By definition,

$$\pi_D = p(q(w))q(w) - wq(w),$$

and

$$\frac{\partial \pi_D}{\partial w} = q'(w)[MR_D(q) - w] - q(w).$$

Assuming first that $w \geq w_I$, we have from Stage 2 that $w = MR_D(q)$ and thus that $\frac{\partial \pi_D}{\partial w} =$

$-q(w) < 0$, implying that $w \geq w_I$ is not profit-maximizing for D. The optimal w chosen by D is thus such that $w < w_I$ which implies that $w(q) = MC_U(q)$ from Stage 2.[65]

D maximization problem thus simplifies to:

$$\underline{w} \equiv \underset{w}{\text{argmax}}\ p(q(w))q(w) - wq(w) \quad \text{subject to } w = MC_U(q).$$

The first-order condition yields:

$$MR_D(\underline{q}) = \underline{w}\left(1 + \frac{q(\underline{w})}{q'(\underline{w})\underline{w}}\right). \tag{57}$$

Using the constraint and $\underline{q} = q(\underline{w})$, the first-order condition can be rewritten as:

$$\underline{w} = MR_D(\underline{q})\left(\frac{\varepsilon_{MC_U}(\underline{q})}{\varepsilon_{MC_U}(\underline{q}) + 1}\right), \tag{58}$$

with $\underline{w} = MC_U(\underline{q}) < w_I$.

The second-order condition yields:

$$\frac{\partial^2 \pi_D(w)}{\partial w^2} = q''(w)[MR_D(q) - w(q)] + q'(w)[q'(w)MR'_D(q) - 2] < 0$$
$$\iff \sigma_{MC_U} > \frac{\frac{MR'_D}{MC'_U} - 2}{\left(\frac{MR_D}{MC_U} - 1\right)\epsilon_{MC_U}} \tag{59}$$

where we used that $w(q) = MC_U(q)$, and thus $q'(w) = \frac{1}{MC'_U(q)} > 0$ and $q''(w) = -\frac{\sigma_{MC_U}}{q[MC'_U(q)]^2}$. We also have $\varepsilon_{MC_U} \equiv \frac{MC_U(q)}{q|MC'_U(q)|}$ and $\sigma_{MC_U} \equiv \frac{qMC''_U(q)}{|MC'_U(q)|}$ as defined in the main text. Using (58), which implies $\left(\frac{MR_D}{MC_U} - 1\right)\epsilon_{MC_U} = 1$, (59) simplifies to:

$$\sigma_{MC_U} > \frac{MR'_D}{MC'_U} - 2.$$

As $MR'_D(q) < 0$ and $MC'_U(q) > 0$, it is straightforward that Assumption 2, which stipulates that $\sigma_{MC_U} > -2$, is sufficient to guarantee the second-order condition validity.[66] ■

In this extreme case, the expression of the equilibrium markdown of D simplifies, so that:[67]

$$\nu_D \equiv \frac{MR(q^*)}{w(q^*)} = \frac{\varepsilon_{MC_U} + 1}{\varepsilon_{MC_U}} = \frac{\sigma_r + \varepsilon_r + 3}{\varepsilon_r + 1}.$$

Here, D makes a take-it-or-leave-it offer to U, and can thus equalize its marginal revenue and

[65] Stage 2 results are summarized in Lemma 1.

[66] Assumption 2 imposes that D's marginal cost is increasing when making a take-it-or-leave-it offer to U. It guarantees D's profit concavity even if the final demand is perfectly elastic.

[67] Apart from the endogenous quantity adjustment, other markup and markdown expressions in terms of primitives remain similar in this polar case to the ones derived in Proposition 4.

marginal cost, as visible in Figure 10. As a result, the markdown of D only depends on the supply primitives through $(\varepsilon_r, \sigma_r)$. It contrasts with the intermediate case where $0 < \alpha < \alpha_I$, and where the demand elasticity (ε_p) affects the markdown of D, as U exercises countervailing power via the Nash-bargaining.

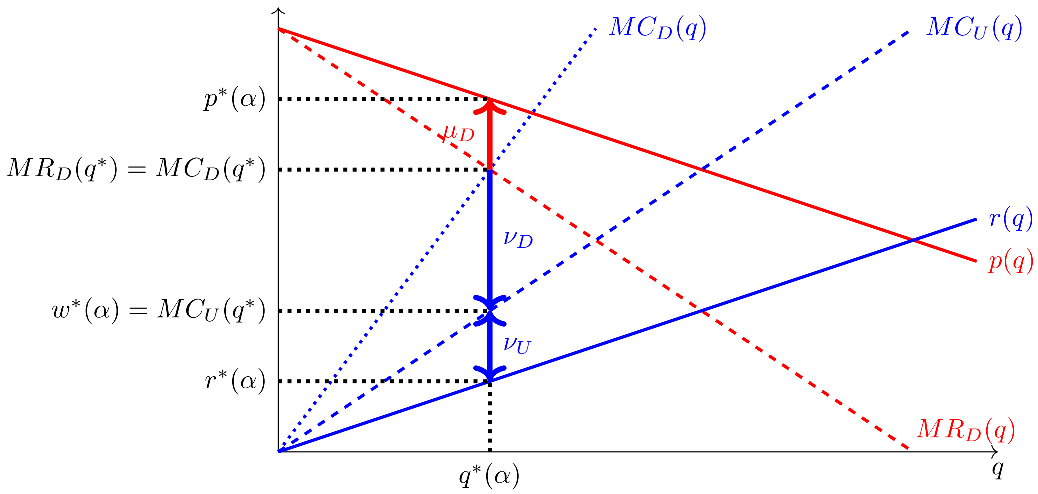


Figure 10: Take-it-or-Leave-it Offer from D $(\alpha = 0)$.

B.2.4 Stage 1.1

Proceeding backward, in Stage 1.1, U solves the following maximization problem:

$$\max_{w} \pi_U(w) \quad \text{subject to} \begin{cases} \pi_D(w) \geq \phi \pi_D(\overline{w}) + (1-\phi)\pi_D(\underline{w}) \\ w = \begin{cases} MC_U(q(w)) \text{ if } w \leq w_I \\ MR_D(q(w)) \text{ if } w \geq w_I \end{cases} \end{cases}$$

where $\pi_D(\overline{w})$ and $\pi_D(\underline{w})$ are the profits of D when U makes the TIOLI offer and when D makes the TIOLI offer, respectively.

Proposition 8 *For any Nash-bargaining solution $w^* \in [\underline{w}, \overline{w}]$ there exists a unique $\phi \in [0, 1]$ such that the non-cooperative game solution $w^{**} = w^*$.*

Proof. It can first be shown that there exists a unique $w = w^{**} \in [\underline{w}, \overline{w}]$ which (i) satisfies D's participation constraint and (ii) solves U's maximization problem, and (iii) satisfies $w^{**} = MC_U(q^{**})$ if $w^{**} \leq w_I$, or $MR_D(q^{**})$ if $w^{**} \geq w_I$, where $q^{**} \equiv q(w^{**})$. Indeed, we know from

Appendix B.2.2 and Appendix B.2.3 that:

$$\begin{cases} \frac{\partial \pi_U(w)}{\partial w} < 0 \text{ and } \frac{\partial \pi_D(w)}{\partial w} < 0 \text{ for } w > \overline{w} \\ \frac{\partial \pi_U(w)}{\partial w} > 0 \text{ and } \frac{\partial \pi_D(w)}{\partial w} < 0 \text{ for } \underline{w} < w < \overline{w} \\ \frac{\partial \pi_U(w)}{\partial w} > 0 \text{ and } \frac{\partial \pi_D(w)}{\partial w} > 0 \text{ for } w < \underline{w} \end{cases}$$

First, note that it implies $\underline{w} \leq w^{**} \leq \overline{w}$ because for $w \leq \underline{w}$ the two firms prefer a higher w and for $w \geq w$ the two firms prefer a lower w. Second, note that, in $[\underline{w}, \overline{w}]$, profit derivative signs are invariant to which case of (iii) prevails in equilibrium and that (iii) has thus no bite in the rest of the proof. Third, suppose that $w^{**} \in [\underline{w}, \overline{w}]$ and satisfying (i), (ii), (iii). Considering deviations, any $w > w^{**}$ violates D's participation constraint whereas, any $w < w^{**}$ brings a lower profit to U.

We show now that there exists a unique $\phi \in [0, 1]$ allowing to reach any $w^{**} \in [\underline{w}, \overline{w}]$. Defining $C(\phi) \equiv \phi \pi_D(\overline{w}) + (1 - \phi)\pi_D(\underline{w})$, D's participation constraint can be rewritten as $\pi_D(w) \geq C(\phi)$. In equilibrium, D's participation constraint is binding, and $\pi_D(w^{**}) = C(\phi)$. The value of $\phi \in [0, 1]$ thus governs the value of w^{**}. As $C'(\phi) = \pi_D(\overline{w}) - \pi_D(\underline{w}) < 0$ and $\pi'_D(w) < 0$ (for $w > \underline{w}$), w^{**} is increasing in ϕ. For $\phi = 0$, D's participation constraint implies that $C(0) = \pi_D(\overline{w})$, and thus $w^{**} = \underline{w}$. Similarly for $\phi = 1$, $C(1) = \pi_D(\overline{w})$ and $w^{**} = \overline{w}$.

For any Nash-bargaining solution $w^* \in [\underline{w}, \overline{w}]$ there thus exists a unique $\phi \in [0, 1]$ such that $w^{**} = w^*$. ■

Corollary 4 *The non-cooperative game solution is:*

$$w^{**} = \begin{cases} MC_U(q^{**}), \ \textit{for } \phi \leq \phi_I, \\ w_I = MC_U(q_I) = MR_D(q_I), \ \textit{for } \phi = \phi_I, \\ MR_D(q^{**}), \ \textit{for } \phi \geq \phi_I, \end{cases}$$

where

$$\phi_I = \frac{\pi_D(\underline{w}) - \pi_D(w_I)}{\pi_D(\underline{w}) - \pi_D(\overline{w})}.$$

Proof. Let us assume first that $w^{**} \geq w_I$, so that we must have $w^{**} = MR_D(q^{**})$. As D's participation constraint is binding in equilibrium from Proposition 8 (see proof), we necessarily have:

$$w^{**} = MR_D(q^{**}) = p(q^{**}) - \frac{C(\phi)}{q^{**}}$$

The condition $w^{**} \geq w_I$ implies that $MR_D(q^{**}) = p(q^{**}) - \frac{C(\phi)}{q^{**}} \geq w_I$, and at the limit:

$$w^{**} = MR_D(q^{**}) = p(q^{**}) - \frac{C(\phi)}{q^{**}} = w_I,$$

which holds true if and only if $q^{**} = q_I$. Such an equilibrium case thus prevails for a threshold value ϕ_I defined by:

$$\begin{aligned} & MR_D(q_I) && = p(q_I) - \frac{C(\phi_I)}{q_I} \\ \iff & C(\phi_I) && = \underbrace{p(q_I)q_I - MR_D(q_I)q_I}_{\pi_D(w_I)} \\ \iff & \phi_I && = \frac{\pi_D(\underline{w}) - \pi_D(w_I)}{\pi_D(\underline{w}) - \pi_D(\overline{w})}. \end{aligned}$$

Note that $\phi_I \in (0,1)$ as $\pi_D(\underline{w}) > \pi_D(w_I) > \pi_D(\overline{w}) > 0$. Indeed, for $w^{**} = w_I$, D's binding participation constraint implies that $\pi(w_I)$ is a convex combination of $\pi_D(\overline{w})$ and $\pi_D(\underline{w})$, with $\pi_D(\overline{w}) < \pi_D(\underline{w})$.

This case where $w^{**} = MR_D(q^{**})$ and $w^{**} \geq w_I$ prevails in equilibrium for $\phi \in [\phi_I, 1]$. A similar reasoning straightforwardly applies to the case where $w^{**} = MC_U(q^{**})$ and $w^{**} \leq w_I$, which prevails in equilibrium for $\phi \in [0, \phi_I]$ (with ϕ_I uniquely defined as $MR_D(q_I) = MC_U(q_I)$ from Proposition 2). ■

C Price Floors

For a given price floor $\underline{r}$, we denote the quantity $\underline{q} = r^{-1}(\underline{r})$. The price floor is binding whenever the quantity traded is lower than $\underline{q}$.

C.1 Vertical integration case

Assume that the price floor is binding, the maximization of firm I's profit leads to $MR_D(q) = \underline{r}$. The welfare-maximizing price floor is such that $\underline{r_I} = r(\underline{q_I})$ with $MR_D(\underline{q_I}) = r(\underline{q_I})$.

- Consider first a deviation towards a higher price floor $\tau > \underline{r_I}$; It is immediate that $MR_D(q) = \tau \Rightarrow q < \underline{q_I}$ since $MR_D(q)$ is decreasing in q. The deviation is not profitable for welfare.
- Consider now a deviation towards a lower price floor $\tau < \underline{q_I}$; In that case we define $q^\tau = r^{-1}(\tau) < \underline{q_I}$ and the vertically integrated firm chooses q^τ such that $MR_D(q^\tau) = \tau$. The deviation is not profitable for welfare.

C.2 Bargaining case

When the price floor is binding, it destroys the equilibrium defined in Propositions 3 and 4.

Assume now that the price floor is binding. We know that U will supply up to $\underline{q}$ as long as $w \geq \underline{r}$. Up to $\underline{q}$, this is D that selects the equilibrium quantity according to $w = MR_D(q)$. In that case, the bargaining between U and D is modified as follows:

$$\underline{\widetilde{MR}}_U(q,\alpha) = \underline{\beta_D}(q,\alpha)MR_U(q) + (1-\underline{\beta_D})MR_D(q) = \underline{r}$$

where $\underline{\beta_D}(q,\alpha) = \frac{(1-\alpha)\underline{\pi_U}(q)}{\alpha\pi_D(q)}$ and $\underline{\pi_U}(q) = (MR_D(q) - \underline{r})q$. Under the assumption that the price floor is binding, we have $\underline{r} \geq r(q)$ and therefore $\underline{\pi_U}(q) \leq \pi_U(q)$ and $\underline{\beta_D}(q,\alpha) \leq \beta_D(q,\alpha)$. Because $\underline{r} = r(\underline{q})$, however, we have $\underline{\pi_U}(\underline{q}) = \pi_U(\underline{q})$ and therefore $\widetilde{MR}_U(\underline{q},\alpha) = \underline{\widetilde{MR}}_U(\underline{q},\alpha)$.

Proof that $\underline{q_\mu} > q_\mu$. We defined $\underline{q_\mu}(\alpha)$ as the solution of:

$$\underline{\widetilde{MR}}_U(\underline{q_\mu},\alpha) = \widetilde{MR}_U(\underline{q_\mu}(\alpha),\alpha) = r(\underline{q_\mu}).$$

The unconstrained equilibrium being defined as $\widetilde{MR}_U(q_\mu,\alpha) = MC_U(q_\mu) > r(q_\mu)$, and $\widetilde{MR}_U(q_\mu,\alpha)$ decreasing in q, we therefore have that $q_\mu(\alpha) < \underline{q_\mu}(\alpha)$ for all $\alpha \in [0,1]$.

Proof that $\underline{r_\mu} = r(\underline{q_\mu})$ is the welfare-maximizing price floor.

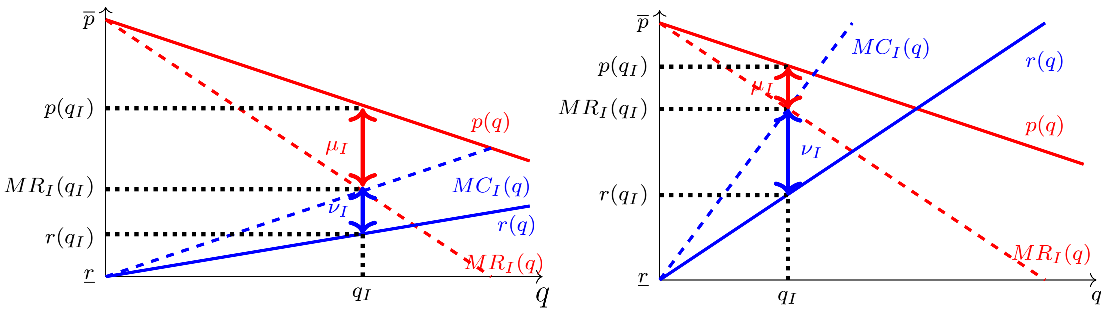
- Consider first a deviation towards a higher price floor $\tau > \underline{r_\mu}$. It is immediate that $\underline{\widetilde{MR}}_U(q_\tau,\alpha) = \tau \Rightarrow q_\tau < \underline{q_\mu}(\alpha)$ since $\underline{\widetilde{MR}}_U(q,\alpha)$ is decreasing in q and $\underline{\widetilde{MR}}_U(\underline{q_\mu},\alpha) = \underline{r_\mu} < \tau$. The deviation is not profitable for welfare.
- Consider now a deviation towards a lower price floor $\tau < \underline{r_\mu}(\alpha)$. In that case we define $q_\tau = r^{-1}(\tau) < \underline{q_\mu}(\alpha)$. As by definition $\underline{\widetilde{MR}}_U(\underline{q_\mu},\alpha) = \widetilde{MR}_U(\underline{q_\mu},\alpha) = r(\underline{q_\mu}) > r(q_\tau)$, it implies that the price-floor equilibrium $\underline{q_\mu}$ no longer exists and that only price floor equilibria leading to $q \leq q_\tau$ exist. The deviation is not profitable for welfare.

For all $\underline{r_\mu}(\alpha)$, the equilibrium quantity $\underline{q_\mu}(\alpha)$ is traded, and the price floor is always binding in equilibrium.

Effect of the optimal price floor on profits, surpluses, and welfare. Note that $\pi_U(q) = [MR_D(q) - r(q)]q$ is maximized in $q_{\underline{\mu}}$. However, we have $\frac{\partial \pi_U}{\partial q} = MRU(q) - MCU(q) < 0$ under Assumption 1.(i) and Assumption 1.(ii) for all $q \geq q_{\overline{\mu}}$. Therefore, whenever $\alpha_I < \alpha < 1$, given that the price floor increases the quantity traded, it hurts U. When $0 < \alpha < \alpha_I$, $\pi_U(q) = [MC_U(q) - r(q)]q$ and the derivative is $MC_D(q) - MC_U(q) > 0$ which means that the larger the quantity the better. Given that the equilibrium quantity under price floor keeps increasing when α decreases, firm U's profit also keeps decreasing. However, U's profit is strictly lower under the optimal price floor in q_I, and we know it is strictly lower under the optimal price floor in $\alpha = 0$ since U is able to preserve a positive markdown absent the price floor, whereas it gets no profit under the price floor. For regular profit functions, it is likely that U is unambiguously hurt.

Regarding the profit of D, we have $\pi_D(q) = [p(q) - MR_D(q)]q$. Here $\frac{\partial \pi_D}{\partial q} = MRD(q) - MRU(q) > 0$ which is positive for all $q > 0$. Therefore, when $\alpha_I < \alpha < 1$, the quantity traded increases under the price floor, which benefits D. However, when $0 < \alpha < \alpha_I$, we have $\pi_D(q) = [p(q) - MC_U(q)]q$ which is maximized in $q_{\overline{\nu}}$. The comparison of the profit of D in $\alpha = 0$ is not obvious. Indeed, in $\alpha = 0$, the quantity $\underline{q_I}$ being larger than q_I, D obtains the full vertical profit, but it is less than the vertically integrated profit since it loses the ability to impose a markdown. As the optimal price floor always increases the quantity traded up to $\underline{q_I} < q_W$, it always benefits consumers, input suppliers, and total welfare.

D Supplementary Figures



(i) $r(q) = \underline{r} + \frac{1}{6}q$, lower elasticity of supply.

(ii) $r(q) = \underline{r} + \frac{2}{3}q$, larger elasticity of supply.

Figure 11: Markup and Markdown when the elasticity of supply changes

References

Allain, Marie-Laure, and Claire Chambolle. 2011. "Anti-Competitive Effects of Resale-Below-Cost Laws." *International Journal of Industrial Organization*, 29(4): 373–385.

Alviarez, Vanessa I., Michele Fioretti, Ken Kikkawa, and Monica Morlacco. 2023. "Two-Sided Market Power in Firm-to-Firm Trade." Unpublished. Available at `http://www.nber.org/papers/w31253`.

Arkolakis, Costas, Federico Huneeus, and Yuhei Miyauchi. 2023. "Production Network Formation, Trade, and Welfare." Unpublished. Available at `https://www.nber.org/papers/w30954`.

Atkin, David, Joaquin Blaum, Pablo D. Fajgelbaum, and Augusto Ospital. 2024. "Trade Barriers and Market Power: Evidence from Argentina's Discretionary Import Restrictions." Unpublished. Available at `https://www.nber.org/papers/w32037`.

Avignon, Rémi, and Etienne Guigue. 2022. "Markups and Markdowns in the French Dairy Market." Unpublished. Available at `https://drive.google.com/file/d/1kFjynavf_M7LR8CZs_R0vZ_qU1rr3QdK/view`.

Avignon, Rémi, and Etienne Guigue. 2025. "Prix planchers dans les filières agro-alimentaires: une mesure d'efficacité?" HAL.

Azar, José, and Ioana Marinescu. 2024. "Monopsony Power in the Labor Market: From Theory to Policy." *Annual Review of Economics*, 16(1): 491–518.

Azar, José, Emiliano Huet-Vaughn, Ioana Marinescu, Bledi Taska, and Till Von Wachter. 2024. "Minimum Wage Employment Effects and Labour Market Concentration." *Review of Economic Studies*, 91(4): 1843–1883.

Azkarate-Askasua, Miren, and Miguel Zerecero. 2022. "Union and Firm Labor Market Power." Unpublished. Available at `https://ssrn.com/abstract=4323492`.

Baqaee, David Rezza. 2018. "Cascading Failures in Production Networks." *Econometrica*, 86(5): 1819–1838.

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *Quarterly Journal of Economics*, 135(1): 105–163.

Barro, Robert J, and Herschel I Grossman. 1971. "A General Disequilibrium Model of Income and Employment." *American Economic Review*, 61(1): 82–93.

Bénassy, Jean-Pascal. 1975. "Neo-Keynesian Disequilibrium Theory in a Monetary Economy." *Review of Economic Studies*, 42(4): 503–523.

Bénassy, Jean-Pascal. 1986. *Macroeconomics: An Introduction to the Non-Walrasian Approach.* . 1 ed., Elsevier.

Bénassy, Jean-Pascal. 1990. "Non-Walrasian Equilibria, Money, and Macroeconomics." *Handbook of Monetary Economics*, 1: 103–169.

Bénassy, Jean-Pascal. 1993. "Nonclearing Markets: Microeconomic Concepts and Macroeconomic Applications." *Journal of Economic Literature*, 31(2): 732–761.

Bernheim, B.Douglas, and Michael D. Whinston. 1998. "Exclusive Dealing." *Journal of Political Economy*, 106(1): 64–103.

Björnerstedt, Jonas, and Johan Stennek. 2007. "Bilateral Oligopoly – The Efficiency of Intermediate Goods Markets." *International Journal of Industrial Organization*, 25(5): 884–907.

Boehm, Christoph E, and Nitya Pandalai-Nayar. 2022. "Convex Supply Curves." *American Economic Review*, 112(12): 3941–3969.

Bonnet, Céline, Zohra Bouamra-Mechemache, and Hugo Molina. 2023. "The Buyer Power Effect of Retail Mergers: An Empirical Model of Bargaining with Equilibrium of Fear." Forthcoming in the *RAND Journal of Economics.*

Bowley, Arthur L. 1928. "Bilateral Monopoly." *Economic Journal*, 38(152): 651–658.

Calzolari, Giacomo, Vincenzo Denicolò, and Piercarlo Zanchettin. 2020. "The Demand-Boost Theory of Exclusive Dealing." *RAND Journal of Economics*, 51(3): 713–738.

Card, David. 2022. "Who Set Your Wage?" *American Economic Review*, 112(4): 1075–1090.

Card, David, and Alan B. Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania."

Card, David, and Alan B Krueger. 1995. *Myth and Measurement: The New Economics of the Minimum Wage.* Princeton University Press.

Chen, Zhiqi. 2003. "Dominant Retailers and the Countervailing-Power Hypothesis." *RAND Journal of Economics*, 34(4): 612–625.

Chipty, Tasneem, and Christopher M. Snyder. 1999. "The Role of Firm Size in Bilateral Bargaining: A Study of the Cable Television Industry." *Review of Economics and Statistics*, 81(2): 326–340.

Choné, Philippe, Laurent Linnemer, and Thibaud Vergé. 2024. "Double Marginalization, Market Foreclosure, and Vertical Integration." *Journal of the European Economic Association*, 22(4): 1884–1935.

Cournot, Augustin. 1838. "Recherches sur les Principes Mathématiques de la Théorie des Richesses."

Crawford, Gregory S, and Ali Yurukoglu. 2012. "The Welfare Effects of Bundling in Multichannel Television Markets." *American Economic Review*, 102(2): 643–685.

Crawford, Gregory S., Robin S. Lee, Michael D. Whinston, and Ali Yurukoglu. 2018. "The Welfare Effects of Vertical Integration in Multichannel Television Markets." *Econometrica*, 86(3): 891–954.

Daskalova, Victoria I, Liesbeth Colen, Zohra Bouamra-Mechemache, and Kjersti Nes. 2020. *Retail Alliances in the Agricultural and Food Supply Chain.* Publications Office of the European Union.

Demirer, Mert, and Michael Rubens. 2025. "Welfare Effects of Buyer and Seller Power." Unpublished. Available at `https://www.nber.org/papers/w33371`.

Dhyne, Emmanuel, Ayumu Ken Kikkawa, and Glenn Magerman. 2022. "Imperfect competition in firm-to-firm trade." *Journal of the European Economic Association*, 20(5): 1933–1970.

Dobson, Paul W., and Michael Waterson. 1997. "Countervalling Power and Consumer Prices." *Economic Journal*, 107(441): 418–430.

Dobson, Paul W., and Michael Waterson. 2007. "The Competition Effects of Industry-Wide Vertical Price Fixing in Bilateral Oligopoly." *International Journal of Industrial Organization*, 25(5): 935–962.

Draganska, Michaela, Daniel Klapper, and Sofia Berto Villas-Boas. 2010. "A Larger Slice or a Larger Pie? An Empirical Investigation of Bargaining Power in the Distribution Channel." *Marketing Science*, 29(1): 57–74.

Drèze, Jacques H. 1975. "Existence of an Exchange Equilibrium under Price Rigidities." *International Economic Review*, 16(2): 301–320.

Dube, Arindrajit, and Attila Linder. 2024. "Minimum Wages in the 21st Century." In *Handbook of Labor Economics*. Vol. 5, , ed. Christian Dustmann and Thomas Lemieux, Chapter 4, 261–383. Elsevier.

Edgeworth, Francis Y. 1925. "The Pure Theory of Monopoly." In *Papers relating to political economy.* Vol. 1, 111–142. Macmillan for the Royal Economic Society.

European Commission. 2004. "Guidelines on the Assessment of Horizontal Mergers Under the Council Regulation on the Control of Concentrations Between Undertakings." Official Journal of the European Union 2004/C 31/03.

Falch, Torberg, and Bjarne Strøm. 2007. "Wage Bargaining and Monopsony." *Economics Letters*, 94(2): 202–207.

Fellner, William. 1947. "Prices and Wages under Bilateral Monopoly." *Quarterly Journal of Economics*, 61(4): 503–532.

Galbraith, John. 1952. *American Capitalism: The Concept of Countervailing Power.*

Gaudin, Germain. 2016. "Pass-through, Vertical Contracts, and Bargains." *Economics Letters*, 139(2): 1–4.

Gaudin, Germain. 2018. "Vertical Bargaining and Retail Competition: What Drives Countervailing Power?" *Economic Journal*, 128(614): 2380–2413.

Ghili, Soheil, and Matt Schmitt. 2024. "Risk Aversion and Double Marginalization." *Journal of Industrial Economics*, 72(2): 762–806.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. "Mergers When Prices are Negotiated: Evidence from the Hospital Industry." *American Economic Review*, 105(1): 172–203.

Grennan, Matthew. 2013. "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices." *American Economic Review*, 103(1): 145–177.

Grossman, Gene M, Elhanan Helpman, and Alejandro Sabal. 2024. "Optimal resilience in multitier supply chains." *Quarterly Journal of Economics*, 139(4): 2377–2425.

Hahn, Frank. 1978. "On Non-Walrasian Equilibria." *Review of Economic Studies*, 45(1): 1–17.

Hemphill, C. Scott, and Nancy L. Rose. 2018. "Mergers that Harm Sellers." *Yale Law Journal*, 127(7): 2078–2109.

Hernández, Carlos Eduardo, and Santiago Cantillo-Cleves. 2024. "A Toolkit for Setting and Evaluating Price Floors." *Journal of Public Economics*, 232: 105084.

Ho, Kate, and Robin S Lee. 2017. "Insurer Competition in Health Care Markets." *Econometrica*, 85(2): 379–417.

Horn, Henrick, and Asher Wolinsky. 1988. "Bilateral Monopolies and Incentives for Merger." *RAND Journal of Economics*, 19(3): 408–419.

Houba, Harold. 2024. "Bargaining and Market Power in Bilateral Monopoly." Unpublished.

Iozzi, Alberto, and Tommaso Valletti. 2014. "Vertical Bargaining and Countervailing Power." *American Economic Journal: Microeconomics*, 6(3): 106–135.

Janssen, Maarten, and Sandro Shelegia. 2015. "Consumer Search and Double Marginalization." *American Economic Review*, 105(6): 1683–1710.

Kroft, Kory, Yao Luo, Magne Mogstad, and Bradley Setzler. 2023. "Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry." Forthcoming in the *American Economic Review.*

Lee, Robin S., Michael D. Whinston, and Ali Yurukoglu. 2021. "Structural Empirical Analysis of Contracting in Vertical Markets." In *Handbook of Industrial Organization.* Vol. 4, , ed. Kate Ho, Ali Hortaçsu and Alessandro Lizzeri, Chapter 9, 673–742. Elsevier.

Lemos, Sara. 2008. "A Survey of the Effects of the Minimum Wage on Prices." *Journal of Economic Surveys*, 22(1): 187–212.

Linnemer, Laurent. 2022. "Doubling Back on Double Marginalization." *Review of Industrial Organization*, 61: 1–19.

Loertscher, Simon, and Leslie Marx. 2022. "Incomplete Information Bargaining with Applications to Mergers, Investment, and Vertical Integration." *American Economic Review*, 112: 616– 649.

Luco, Fernando, and Guillermo Marshall. 2020. "The Competitive Impact of Vertical Integration by Multiproduct Firms." *American Economic Review*, 110(7): 2041–2064.

Manning, Alan. 1987. "An Integration of Trade Union Models in a Sequential Bargaining Framework." *Economic Journal*, 97(385): 121–139.

Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton University Press.

Manning, Alan. 2021. "Monopsony in Labor Markets: A Review." *ILR Review*, 74(1): 3–26.

Mathewson, G. Frank, and Ralph A. Winter. 1984. "An Economic Theory of Vertical Restraints." *RAND Journal of Economics*, 15(1): 27–38.

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

Morlacco, Monica. 2019. "Market Power in Input Markets: Theory and Evidence from French Manufacturing." Unpublished. Available at https://www.dropbox.com/scl/fi/n2m83dj07eb5xgaxdgd1v/BP_draft_01102020.pdf?rlkey=waw8c3f8wlwtk6nilvm8t2avu&e=1&dl=0.

Mortimer, Julie Holland. 2008. "Vertical Contracts in the Video Rental Industry." *Review of Economic Studies*, 75(1): 165–199.

Mrázová, Monika, and J. Peter Neary. 2019. "Selection Effects with Heterogeneous Firms." *Journal of the European Economic Association*, 17(4): 1294–1334.

Mrázová, Monika, and J. Peter Neary. 2017. "Not So Demanding: Demand Structure and Firm Behavior." *American Economic Review*, 107(12): 3835–74.

Nash, John F. 1950. "The Bargaining Problem." *Econometrica*, 18(2): 155–162.

Nocke, Volker, and John Thanassoulis. 2014. "Vertical Relations under Credit Constraints." *Journal of the European Economic Association*, 12(2): 337–367.

Noton, Carlos, and Andrés Elberg. 2018. "Are Supermarket Squeezing Small Suppliers? Evidence from Negotiated Wholesale Prices." *Economic Journal*, 128(610): 1304–1330.

Rey, Patrick, and Jean Tirole. 1986. "The Logic of Vertical Restraints." *American Economic Review*, 76(5): 921–939.

Rey, Patrick, and Thibaud Vergé. 2008. "Economics of Vertical Restraints." In *Handbook of Antitrust Economics.* , ed. Paolo Buccirossi, Chapter 9, 353–390. The MIT Press.

Rey, Patrick, and Thibaud Vergé. 2020. "Secret Contracting in Multilateral Relations." Unpublished. Available at: https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/by/rey/744_version2020.pdf.

Robinson, Joan. 1933. *The Economics of Imperfect Competition.* Springer.

Rubens, Michael. 2023. "Market Structure, Oligopsony Power, and Productivity." *American Economic Review*, 113(9): 2382–2410.

Rubinstein, Ariel. 1982. "Perfect equilibrium in a bargaining model." *Econometrica: Journal of the Econometric Society*, 97–109.

Russo, Carlo, Rachael E. Goodhue, and Richard J. Sexton. 2011. "Agricultural Support Policies in Imperfectly Competitive Markets: Why Market Power Matters in Policy Design." *American Journal of Agricultural Economics*, 93(5): 1328–1340.

Shea, John. 1993. "Do Supply Curves Slope Up?" *Quaterly Journal of Economics*, 108(1): 1–32.

Sheu, Gloria, and Charles Taragin. 2021. "Simulating Mergers in a Vertical Supply Chain with Bargaining." *RAND Journal of Economics*, 52(3): 596–632.

Smith, Howard. 2016. "The Economic of Retailer-Supplier Pricing Relationships: Theory and Evidence." In *Handbook on the Economics of Retailing and Distribution.* Edward Elgar Publishing Limited.

Smith, Howard, and John Thanassoulis. 2015. "Prices, Profits, and Pass-Through of Costs Along A Supermarket Supply Chain: Bargaining and Competition." *Oxford Review of Economic Policy*, 31(1): 64–89.

Snyder, Christopher M. 2008. "Countervailing Power." In *The New Palgrave Dictionary of Economics.* 1–3. Palgrave Macmillan UK.

Spengler, Joseph J. 1950. "Vertical Integration and Antitrust Policy." *Journal of Political Economy*, 58(4): 347–352.

Stigler, George J. 1946. "The Economics of Minimum Wage Legislation." *American Economic Review*, 36(3): 358–365.

Tintner, Gerhard. 1939. "The Problem of Bilateral Monopoly: A Note." *Journal of Political Economy*, 47(2): 263–270.

Tirole, Jean. 1988. *The Theory of Industrial Organization.* MIT press.

Tirole, Jean. 2015. "Market Failures and Public Policy." *American Economic Review*, 105(6): 1665–1682.

Toxvaerd, Flavio. 2024. "Bilateral Monopoly Revisited: Price Formation, Efficiency and Countervailing Powers." Unpublished. Available at https://cepr.org/publications/dp19338.

Treuren, Leonard. 2022. "Wage Markups and Buyer Power in Intermediate Input Markets." Unpublished. Available at https://feb.kuleuven.be/research/economics/ces/documents/DPS/DPS2022/dps%202206.pdf.

Varian, Hal R. 1977. "Non-Walrasian Equilibria." *Econometrica*, 45(3): 573–590.

Wong, Horng Chern. 2023. "Understanding High-Wage Firms." Unpublished. Available at https://horngchernwong.github.io/horngchernwong.com/HWFs.pdf.

Zavala, Lucas. 2022. "Unfair Trade? Monopsony Power in Agricultural Value Chains." Unpublished. Available at https://dropbox.com/s/o8ez4kjmpo5a6mg/JMP.pdf?raw=1.