Markups, Markdowns, and Bargaining in a Vertical Supply Chain*

Rémi Avignon[†], Claire Chambolle[‡], Etienne Guigue[§] and Hugo Molina[¶]

July 23, 2025

Abstract

This article bridges monopoly, monopsony, and countervailing power theories to analyze their welfare implications in a vertical supply chain. We develop a bilateral monopoly model with bargaining that accommodates upstream monopsony and downstream monopoly power. In equilibrium, the "short-side rule" applies: the quantity exchanged is determined by the firm willing to trade less. Welfare is maximized when each firm's bargaining power exactly countervails the other's market power. Otherwise, double marginalization arises in the form of *double markdownization* under excessive downstream bargaining power, or *double markupization* under excessive upstream bargaining power. We offer novel insights for price regulation and competition policy.

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

JEL classification: C78, D42, J42, L81.

---

*First version: December 20, 2024. We would like to thank Laurent Linnemer, Simon Loertscher, Massimo Motta, 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 (2024), CMA Durham Workshop, CMA-UKRI Workshop on Market Power in Supply Chains, M-POWER conference (KU Leuven, 2025), the MaCCI Annual Conference (2025), IO seminar in UC Berkeley, NHH Bergen Micro Seminar, and CRESSE conference (2025) 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.

†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; 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

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 powerful buyers charge markdowns, leading to lower output quantity and welfare.[2]

Both theories have a profound influence on 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

[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 consumer 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 European Commission's horizontal merger guidelines, which state that "buyer power would act as a countervailing factor to an increase in market power resulting from the merger" (European Commission, 2004, para. 11), and the JRC policy report on buying alliances (Daskalova et al., 2020). In the United States, the 2010 merger guidelines stated that "The Agencies consider the possibility that powerful buyers may constrain the ability [...] to raise prices" (U.S. Department of Justice and the Federal Trade Commission, 2010, page 27).

[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

grounded in different sets of assumptions, the countervailing and monopsony power theories conflict in shaping appropriate antitrust treatment of buyer power.[5]

This article develops a unified framework that incorporates both theories to provide new insights into the welfare effects of buyer and seller power in vertical supply chains. We consider a bilateral monopoly setting where an upstream firm, U, sells its product to a downstream firm, D, which then resells it to final consumers. To examine monopsony power, we depart from the canonical model of vertical contracting (e.g., Cournot, 1838; 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, 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 "short-side rule"—i.e, the minimum between what U is willing to sell and D is willing to purchase.

We demonstrate that the distribution of bargaining power between U and D affects both the magnitude and the nature of the double marginalization phenomenon highlighted by Cournot (1838) and Spengler (1950).[7] When D's bargaining power is high, the negotiated wholesale price is low, so that U is willing to supply a quantity smaller than D's demand, thereby determining the quantity exchanged in equilibrium. As a consequence, the equilibrium wholesale price and quantity move along U's marginal cost curve. U's upward-sloping marginal cost is internalized during the bargaining, and 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

(United States v. Bertelsmann SE & Co. KGaA et al., No. CV 21-2886-FYP). See also the Federal Trade Commission's lawsuit to block the merger between the supermarket giants Albertsons and Kroger (press release).

[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 heterogeneous price-taking input suppliers or workers.

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

market, resulting in a lower quantity, a higher consumer price, and a lower input price compared to what a vertically integrated firm would set in equilibrium.[8] This *double markdownization* constitutes a novel source of double marginalization distortion. In this case, Galbraith's (1952) argument does not apply: increasing D's bargaining power raises its markdown, which reduces welfare. Instead, enhancing U's bargaining power improves welfare by strengthening its ability to exercise *countervailing seller power*.[9]

When U's bargaining power is high, the logic is analogous: as the negotiated wholesale price is high, U is willing to supply a quantity exceeding D's demand, implying that the latter determines the quantity exchanged in equilibrium. As a result, the equilibrium wholesale price and quantity move along D's marginal revenue curve (i.e., its demand for U's product). In this case, alongside D's markup, U also charges a markup. This *double markup* gives rise to the classical Cournot-Spengler double marginalization problem, and Galbraith's (1952) countervailing buyer power argument applies: increasing D's bargaining power reduces U's markup, which improves welfare.

We further characterize a threshold level of U's bargaining power vis-à-vis D at which each firm fully countervails the other's market power, thereby eliminating double marginalization and replicating the vertically integrated outcome. This threshold depends on the underlying supply and demand primitives, and two limiting cases are insightful: D (resp. U) should hold all the bargaining power when U (resp. D) faces a perfectly elastic supply (resp. demand), as both firms face a constant marginal cost (resp. revenue), leaving no room for monopsony (resp. monopoly) power.

We extend our analysis in two directions. We first show that our results continue to hold qualitatively under two-part tariff contracts, provided that frictions limit the use of the fixed fee for rent extraction and prevent the full elimination of double marginalization. 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

[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). See Section 7.2 for a discussion.

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

power in the vertical supply chain. We demonstrate that there exists an optimal level of the price floor, which always increases welfare and depends on the distribution of bargaining power. The resulting welfare gains are larger under double markdownization (i.e., when D's bargaining power vis-à-vis U is high), as the price floor turns D's monopsony power into countervailing buyer power.

Contributions. We contribute to the extensive literature on vertical relationships that explores the sources and consequences of the Cournot-Spengler double marginalization and its potential remedies.[10] One strand of this literature, which studies firm-to-firm bargaining, systematically relies on the assumption of constant marginal costs of production.[11] Our main contribution is to relax this assumption. We do so by considering the presence of monopsony power in the input market, allowing us to identify a novel source of double marginalization, which we refer to as double markdownization.[12] This distortion has significant welfare implications. In particular, we uncover a novel theory of countervailing seller power and highlight the existence of a threshold level of bargaining power that eliminates double marginalization and restores bilateral efficiency. This result is complementary to Loertscher and Marx (2022), who demonstrate that equalizing bargaining power can achieve bilateral efficiency under incomplete information.[13] In contemporaneous work, Demirer and Rubens (2025) 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 notably differ in three respects. First, we leverage the voluntary exchange property to determine the equilibrium traded quantity

[10] The analysis of the double marginalization has a long tradition in the industrial organization literature (see Tirole, 1988, for a textbook and Rey and Vergé, 2008, for a review). Alongside observations of linear wholesale tariffs in vertical supply chains (e.g., Crawford and Yurukoglu, 2012; Goldberg and Hellerstein, 2013; Gowrisankaran, Nevo and Town, 2015), recent empirical evidence of the double marginalization phenomenon has been documented in Luco and Marshall (2020) and Molina (2024). Recent work in the trade and macroeconomics literature 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 (IO) literature, and Grossman, Helpman and Sabal (2024) in the trade literature.

[12] This starkly differs from the "double markup/markdown" terminology used in Kroft et al. (2023), which does not refer to the Cournot-Spengler double marginalization issue in vertical relationships.

[13] 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.

(and, in turn, the source of double marginalization), whereas they rely on participation constraints (e.g., "D (resp. U) participates in bargaining if its resulting markdown (resp. markup) is nonnegative"). Second, we consider that U exercises monopsony power in the input market, which allows us to uncover the notion of double markdownization and analyze input price floor policies. Finally, we emphasize that the standard definitions of markups and markdowns do not readily apply to bargaining settings and propose general definitions.

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 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]

Finally, we contribute to the analysis of input price floors (minimum wages), which has been extensively studied in the labor economics literature.[16] Since Robinson (1933), it is well-known that minimum wages can increase employment in the presence of labor monopsony power.[17] The incidence of minimum wage (or more broadly, input price

[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". See Section 4.1 for a discussion.

[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. Another approach is developed 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.

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

[17]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). See also

floor) policies has also been examined in oligopoly-oligopsony models, where a set of firms exert both monopoly power in the product market and monopsony power in the input market (e.g., Russo, Goodhue and Sexton, 2011; Avignon and Guigue, 2022; Hernández and Cantillo-Cleves, 2024). 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 and associated welfare gains depend critically on the balance of power within the vertical chain and the nature of the resulting double marginalization. In particular, our findings suggest that the concern for downstream buyer power is mitigated by the presence of an input price floor, shedding new light on the interdependency between competition policy and price regulation (e.g., minimum wage).

The remainder of this article is organized as follows. Section 2 provides markup and markdown definitions that accommodate both unilateral price-setting and bargaining models. Section 3 presents our bilateral monopoly model 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 examines the impact of an optimal input price floor. Section 7 discusses some of our modeling assumptions, and Section 8 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 resource 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

Loertscher and Muir (2021) for a recent theoretical contribution, highlighting the benefit of a minimum wage policy in an incomplete information framework.

the downward price distortion arising from the firm's buyer power.[18] 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, 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 inframarginal units. This latter effect leads the firm to restrict output (relative to perfect competition) and charge a markup over marginal cost.[19] 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 input price paid on inframarginal 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 a vertical supply chain setting with bargaining:

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 inframarginal units.*

[18]The marginal revenue product (MRP) is defined as the marginal product (MP)—the additional output produced by one more unit of input—multiplied by the marginal revenue (MR)—the additional revenue generated from selling that extra unit of output. Under a one-to-one production technology, $MP = 1$ so that $MRP = MR$.

[19]To illustrate the joint role of conditions (i) and (ii), consider a standard monopolist facing a downward-sloping demand curve (condition (i)) and restricted to charge a uniform unit price $p(q)$ for its output (condition (ii)). To sell an additional unit, the monopolist must lower the output price, incurring a revenue loss on inframarginal units equal to $p'(q)q$. This loss provides the monopolist with incentives to set a price above its marginal cost, thereby charging a markup. If, instead, the monopolist either faces a perfectly elastic demand (i.e., $p'(q) = 0$) or behaves as a price-taker (violating condition (i)), it neither incurs nor internalizes any revenue loss on inframarginal units, thereby eliminating the incentive to price above marginal cost. Likewise, when the monopolist can engage in perfect (first-degree) price discrimination (violating condition (ii)), 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.

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 firm's cost from inframarginal units.*

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 a monopolist, denoted by D, 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: $MR_D(q^*) = MC_D(q^*)$, where q^* denotes the equilibrium quantity and $MR_D(q^*) \equiv p(q^*) + p'(q^*)q^*$ with $p'(q^*) < 0$. Absent any (negative) effect on D's revenue from inframarginal units—that is, $p'(q^*)q^* = 0$—the minimum price at which it would be willing to sell its marginal unit of output is $\hat{p} = MC_D(q^*) < p(q^*)$. Hence, by Definition 1, D's markup is given by $\mu_D \equiv \frac{p(q^*)}{\hat{p}(q^*)} = \frac{p(q^*)}{MC_D(q^*)} \left(= \frac{p(q^*)}{MR_D(q^*)}\right)$, which coincides with the standard definition of a markup.

Symmetrically, consider a monopsonist, denoted by U, 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 $MC_U(q^*) \equiv r'(q^*)q^* + r(q^*)$ with $r'(q^*) > 0$. Absent any effect on its cost to acquire inframarginal input units—that is, $r'(q^*)q^* = 0$—the maximum price at which U would be willing to purchase its marginal unit of input is $\hat{r} = MR_U(q^*) > r(q^*)$. Thus, by 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 product market, U acts as a monopsonist in the input market, and the linear wholesale price paid by D to U is unilaterally determined by one of the two firms. In this context, as established in the early 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 markups and markdowns along the vertical supply chain (see Appendix A.8.1 for details).

In the more general case where U and D engage in bilateral negotiation, the linear 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. Consequently, it is no longer clear that the classical markup and markdown definitions remain appropriate measures of firms' market power.[20] Definitions 1 and 2 extend the notion of markups and markdowns to the context of vertical bargaining. Specifically, we obtain the following expressions for D's and U's markups and markdowns: $\mu_D = \frac{p(q^*)}{\hat{p}} = \frac{p(q^*)}{MR_D(q^*)}$, $\nu_D = \frac{\hat{w}(q^*)}{w(q^*)} = \frac{MR_D(q^*)}{w(q^*)}$, $\mu_U = \frac{w(q^*)}{\hat{w}(q^*)} = \frac{w(q^*)}{MC_U(q^*)}$, and $\nu_U = \frac{\hat{r}(q^*)}{r(q^*)} = \frac{MC_U(q^*)}{r(q^*)}$ (see Appendix A.8.2 for details).[21] Note that these definitions for markups and markdowns do not depend on the firm's position in the vertical chain. They also preserve the logic underlying the standard definitions: the upward price distortion from D's monopoly power stems from $p'(q^*)q^*$, and the downward price distortion from U's monopsony power stems from $r'(q^*)q^*$. We adopt these definitions throughout this article when referring to firms' markups and markdowns.

3 Vertical Chain and Integration Benchmark

3.1 Vertical Chain

Consider a vertical supply chain in which an upstream firm, U, purchases an input at a price r to produce a good that is sold to consumers at a price p through a downstream firm, D. We assume that U and D operate under a one-to-one production technology. Moreover, U incurs no cost other than its input price r, and D incurs no cost 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, which ensures the existence of a profit-maximizing equilibrium:

[20]For instance, as previously discussed, the upward price distortion in the product 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.

[21]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 the Cournot-Spengler canonical model of vertical contracting and its various extensions to bargaining (see Lee, Whinston and Yurukoglu, 2021, for a review).

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

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

(ii) $p'(\cdot) < 0$, $\infty > \varepsilon_p(\cdot) > 1$, *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 function, defined by $MC_U(q) \equiv r'(q)q + r(q)$, is increasing in quantity q. Note that the case where U has constant marginal costs will be treated as a limit ($r'(\cdot) \to 0$). Assumption 1.(ii) implies that D faces a decreasing inverse demand curve $p(q)$, and that its marginal revenue function, defined by $MR_D(q) \equiv p'(q)q + p(q)$, is decreasing and remains positive over the relevant range of quantities. Finally, Assumption 1.(iii) ensures that $MC_U(q)$ and $MR_D(q)$ intersect.

We define total welfare as $W(q) \equiv \int_0^q [p(q) - r(q)]dq$, and denote by q_W the welfare-maximizing quantity characterized by the condition $p(q_W) = r(q_W)$. Consumer surplus is given 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 U and D form a vertically integrated firm, denoted by I. Acting both as a monopolist in the product market and a monopsonist in 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:

$$\underbrace{p(q_I) + p'(q_I)q_I}_{MR_I(q_I)} = \underbrace{r(q_I) + r'(q_I)q_I}_{MC_I(q_I)}. \quad (1)$$

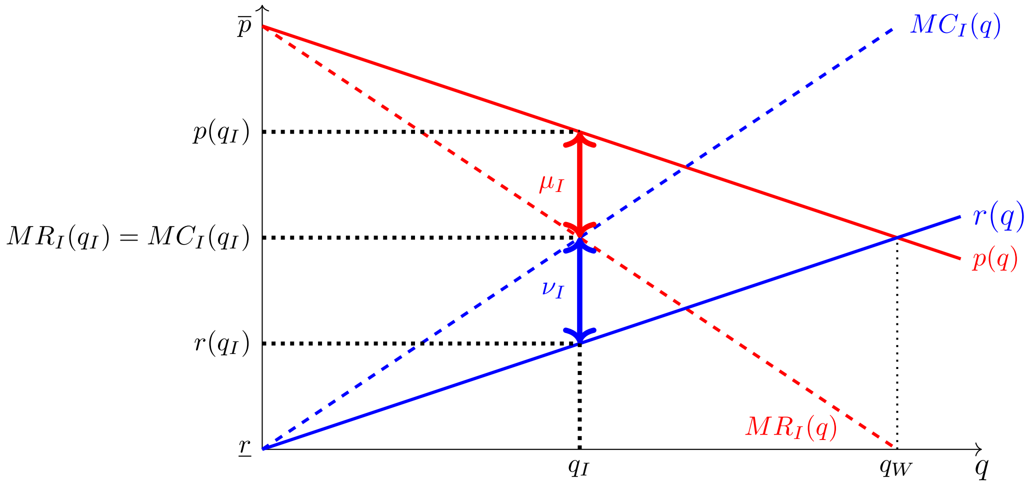

Figure 1: Monopoly and Monopsony Power under Vertical Integration

Notes: The figure is drawn under the demand function $p(q) = \overline{p} - \frac{1}{3}q$ and the supply function $r(q) = \underline{r} + \frac{1}{3}q$. The red arrow labeled μ_I and the blue arrow labeled ν_I represent, respectively, the markup and markdown wedges in differences.

where q_I denotes the equilibrium quantity at which I's marginal revenue equals its marginal cost. As discussed in Section 2, the exercise of monopoly power implies that I's marginal revenue differs from $p(q_I)$ by a wedge equal to $p'(q_I)q_I$. Similarly, the exercise of monopsony power implies that I's marginal cost differs from $r(q_I)$ by a wedge equal to $r'(q_I)q_I$. 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 resulting consumer price is $p(q_I) > p(q_W)$, and the input price is $r(q_I) < r(q_W)$. I's markup, markdown, and 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. See Appendix A.8.1 provides formal derivations of the markup and markdown expressions based on the definitions introduced in Section 2. ■

Figure 1 illustrates the insights from Proposition 1. Both monopoly and monopsony power exercised by I distort prices and reduce the equilibrium quantity q_I below the welfare-maximizing level q_W. Consequently, I charges both a markup and a markdown. The markup μ_I measures the upward distortion in the consumer price p due to I's monopoly power, which decreases with the elasticity of consumer demand ε_p. The markdown ν_I reflects the downward distortion in the input price r resulting from I's monopsony power, which decreases with the elasticity of input supply ε_r. In the limit cases, $\mu_I = 1$ when consumer demand is perfectly elastic and $\nu_I = 1$ when input supply is perfectly elastic. I's margin, denoted by M_I, summarizes the overall price distortion caused by I's market power. It is given by the ratio of the output price to the input price, and can equivalently be expressed as the product of I's markup and markdown.

4 Bargaining and Double Marginalization

We now analyze the vertical supply chain 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. We consider that U and D interact 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 each 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 Cournot-Spengler model of vertical relationships—and its extension to bargaining—as a special case when the input supply is perfectly elastic (i.e., U has a constant marginal cost). 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.[22] In Stage 2, the

[22] Since the early work on bilateral monopoly (e.g., Bowley, 1928; Tintner, 1939), the analysis of vertical relationships under linear wholesale pricing has a long-standing tradition in the vertical contracting literature (e.g., Spengler, 1950; Katz, 1987; Horn and Wolinsky, 1988; Salinger, 1988; Dobson and Waterson, 1997; O'Brien, 2014; Iozzi and Valletti, 2014; Gaudin, 2018, 2019). Furthermore, the

quantity traded is determined under voluntary exchange: neither firm can compel the other to trade more than it is willing to.[23] This assumption reflects a natural feature of most markets and is standard in both Walrasian and non-Walrasian theories (e.g., Bénassy, 1993). As discussed below, voluntary exchange is also implicit in the canonical Cournot-Spengler model of vertical relationships.

In Online Appendix OA2, 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 U and D bargain according to the random-proposer protocol of Rey and Vergé (2020).

4.1 Quantity Choice

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

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

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:

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

use of such simple contracts has been documented in the Chilean coffee market (Noton and Elberg, 2018) and fresh-egg market (Cussen and Montero, 2024), 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 7.1.

[23] We provide an alternative formulation for Stage 2 in Appendix B where, instead of announcing the quantity their are willing to trade, U and D unilaterally set the input and consumer prices, respectively.

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 equates the wholesale price w with its marginal revenue, whereas U's profit-maximizing quantity equates w with its marginal cost. As Assumption 1.(ii) implies that $MR_D(q)$ is decreasing in q, it follows that $\tilde{q}_D(w)$ is decreasing in w. Conversely, as $MC_U(q)$ is increasing in q under Assumption 1.(i), $\tilde{q}_U(w)$ is increasing in w. Given voluntary exchange, the following lemma characterizes the unique equilibrium in dominant strategies:[24]

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.[25] In particular, the equilibrium traded quantity maximizes the profit of at least one firm, thereby satisfying the market efficiency property.[26] Combined with the voluntary exchange assumption, this implies that the "short-side rule" emerges in equilibrium:

[24] It is worth noting that there exists a multiplicity of Nash equilibria in (weakly) dominated strategies. For instance, if U believes that D will announce $\hat{q} < \tilde{q}_D$, a best response for U is to also announce $\hat{q}$. The reasoning is symmetric if D believes that U will announce $\hat{q} < \tilde{q}_U$. Hence, any strategy profile $(\hat{q},\hat{q})$ with $\hat{q} < \min\{\tilde{q}_U(w),\tilde{q}_D(w)\}$ constitutes a Nash equilibrium. However, such equilibria are not trembling-hand perfect as both U and D are better off announcing $\check{q} > \hat{q}$ whenever the other firm trembles upward. Besides, 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 in Lemma 1. Finally, it is straightforward that the Pareto dominance criterion also selects the equilibria leading to the same outcome as in Lemma 1.

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

[26] 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.

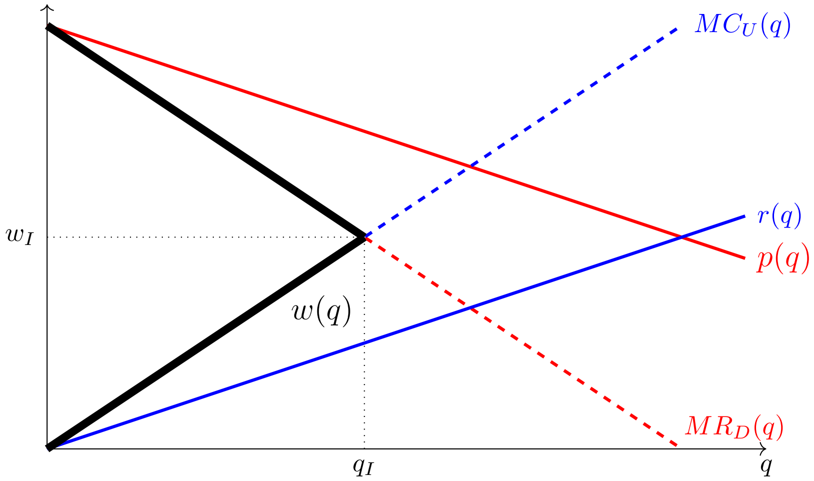

Figure 2: Short-Side Rule.

Notes: The black curve depicts the wholesale price schedule $w(q)$ under linear demand and supply functions, where $w(q) = MR_D(q)$ when $w > w_I$, and $w(q) = MC_U(q)$ when $w < w_I$.

the firm on the short side of the market realizes its profit-maximizing outcome. Second, the equilibrium quantity is (weakly) lower than the vertically integrated outcome, q_I. The reasoning is as follows. First, as Assumption 1.(iii) ensures that $MR_D(q)$ and $MC_U(q)$ intersect, there is a unique $w = w_I$ such that $\tilde{q}_D(w_I) = \tilde{q}_U(w_I) = q_I$. Second, as w affects $\tilde{q}_D(w)$ and $\tilde{q}_U(w)$ in opposite directions, either U or D is willing to trade less than q_I when $w \neq w_I$. Under voluntary exchange, it turns out that $q(w) \leq q_I$.

Figure 2 illustrates the logic underlying Lemma 1. When $w > w_I$, U wants to sell a quantity greater than what D is willing to purchase ($\tilde{q}_U(w) > q_I > \tilde{q}_D(w)$). Being on the short side of the market, D has the right-to-manage D—i.e., it determines the quantity exchanged in equilibrium such that w equals its marginal revenue, as described in (3). In contrast, when $w < w_I$, U prefers to sell a smaller quantity than what D wants to purchase ($\tilde{q}_D(w) < q_I < \tilde{q}_U(w)$), implying that U has the right-to-manage—i.e., it sets the equilibrium quantity such that w equals its marginal cost, as described in (5). Consequently, the firm that chooses the equilibrium quantity to be traded along the vertical supply chain directly depends on the level of w.

It is worth mentioning that Lemma 1 encompasses the canonical bilateral monopoly setting in which U has a constant marginal cost (i.e., $\varepsilon_r \to \infty$). In this case, for any $w > r$, it follows directly from (4) that U is willing to trade an infinite quantity with

D.[27] As consumer demand is not perfectly elastic, D lies on the "short" side of the market and thus always determines the equilibrium traded quantity. By contrast, when U faces increasing marginal costs, the assumption that D has the right-to-manage in all circumstances may violate voluntary exchange. This arises whenever $w < w_I$, as D would demand a quantity exceeding $\tilde{q}_U(w)$, thereby forcing U to sell such extra units at a loss (i.e., $w < MC_U(q)$ for any $q > \tilde{q}_U(w)$).[28]

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 $\pi_D(w) = [p(q(w)) - w]\, q(w)$. Although $q(w)$ is not differentiable at w_I, we demonstrate in Appendix A.2 that the above Nash product is differentiable for all $w \in [r(q(w)), p(q(w))]$. Furthermore, to ensure that (6) is well-defined, we introduce the following assumption:

Assumption 2 *U's marginal cost function $MC_U(q)$ and D's marginal revenue function $MR_D(q)$ satisfy the following conditions:*

(i) $\sigma_{MC_U} > -2$.

(ii) $\varepsilon_{MR_D} > 1$ *and* $\sigma_{MR_D} < 2$.

Analogous to Assumption 1, which ensures that I's profit function is well-defined, Assumption 2 guarantees that the Nash product in (6) admits a unique maximum. Specifically, when anticipating that U sets the quantity in Stage 2 ($w = MC_U(q)$), Assumption 1.(i) and Assumption 2.(i) ensure that the second-order condition of (6) is

[27] Formally, when $r(q) = MC_U(q) = r$ for any q, (4) implies that $\tilde{q}_U(w) = \infty$ whenever $w \geq r$. Symmetrically, when $p(q) = MR_D(q) = p$ for any q, (2) implies that $\tilde{q}_D(w) = \infty$ whenever $p \geq w$.

[28] Analogously, when D faces decreasing marginal revenue, the assumption that U always has the right-to-manage violates voluntary exchange whenever $w > w_I$, as U would supply a quantity exceeding $\tilde{q}_D(w)$, thereby forcing D to buy such extra units at a loss (i.e., $MR_D(q) < w$ for any $q > \tilde{q}_D(w)$).

satisfied (see Appendix A.4.1 for details). Similarly, when anticipating that D sets the quantity in Stage 2 ($w = MR_D(q)$), Assumption 1.(ii) and Assumption 2.(ii) ensure that the second-order condition of (6) holds (see Appendix A.5.1 for details). More precisely, $\sigma_{MR_D} < 2$ ensures that the marginal revenue of U, defined by $MR_U(q) \equiv \frac{\partial MR_D(q)q}{\partial q} = MR_D'(q)q + MR_D(q)$, is decreasing.[29] Similarly, $\sigma_{MC_U} > -2$ ensures that the marginal cost of D, defined by $MC_D(q) \equiv \frac{\partial MC_U(q)q}{\partial q} = MC_U'(q)q + MC_U(q)$, is increasing.[30] Finally, the condition $\varepsilon_{MR_D} > 1$, analogous to $\varepsilon_p > 1$, ensures that U's marginal revenue remains positive over the relevant range of quantities.[31]

The first-order condition of (6), which characterizes U's and D's joint profit and its division between them, is given by $\alpha\pi_U'(w)\pi_D(w) + (1-\alpha)\pi_D'(w)\pi_U(w) = 0$. For simplicity and to facilitate correspondence with the graphical illustrations, we re-express this first-order condition in q as follows:

$$\alpha\pi_U'(q)\pi_D(q) + (1-\alpha)\pi_D'(q)\pi_U(q) = 0 \tag{7}$$

where $\pi_U(q) = (w(q) - r(q))\,q$ and $\pi_D(q) = (p(q) - w(q))\,q$, which follows from the strict monotonicity of $q(w)$ over the range of values $w \geq w_I$ and $w \leq w_I$.[32] Interestingly, (7) embeds every factor determining U's and D's bargaining power in the vertical supply chain.[33] The first comes from the bargaining weight α, which captures any asymmetry in firms' relative bargaining ability.[34] The second factor is captured by $\pi_D(q)$ and $\pi_U(q)$, which represent D's and U's gains from trade, respectively.[35] The

[29] $MR_U(q)$ is the marginal revenue of firm U when making a take-it-or-leave-it offer to D—i.e., U faces a demand curve given by $MR_D(q)$.

[30] $MC_D(q)$ is the marginal cost of firm D when making a take-it-or-leave-it offer to U—i.e., D faces a supply curve given by $MC_U(q)$.

[31] Note that $\varepsilon_{MR_D} > 1$ can equivalently be expressed as $\varepsilon_p > 3 - \sigma_p$, which implies that consumer demand is supermodular (see e.g., Mrázová and Neary, 2017). Supermodular demand functions include, among others, the CES, translog, and AIDS demand models. Supermodularity also holds under linear demand when $\epsilon_p > 3$, and in the logit demand model for sufficiently small values of q^*.

[32] By Assumption 1, $MR_D(q)$ is strictly decreasing and $MC_U(q)$ is strictly increasing. Hence, from (3), we have $w = MR_D(q) \Leftrightarrow q = MR_D^{-1}(w)$ for $w \geq w_I$. Similarly, from (5), we have $w = MC_U(q) \Leftrightarrow q = MC_U^{-1}(w)$ for $w \leq w_I$.

[33] See also Bonnet, Bouamra-Mechemache and Molina (2025), who discuss how (7) relates to the notion of "equilibrium of fear".

[34] This bargaining weight is often deemed to reflect some imprecisely defined asymmetries in the bargaining power of firms (Roth, 1979). Using strategic models of bargaining, Binmore, Rubinstein and Wolinsky (1986) demonstrates that this parameter may capture differences in bargainers' beliefs or asymmetries in the bargaining procedure.

[35] As each firm has a single trading partner, their status quo profits reduce to zero.

last determinant of each firm's bargaining power lies in $\pi'_U(q)$ and $\pi'_D(q)$, which have opposite signs and reflect, respectively, U's and D's costs of making a concession during the negotiation (i.e., granting more favorable terms of trade to its trading partner).

Based on (3), (5), and Lemma 1, we solve Stage 1 anticipating the following three subgame equilibria: (i) U and D announce the same quantity q_I, (ii) U determines the quantity traded in Stage 2, corresponding to $w_I > w(q) = MC_U(q)$, and (iii) D determines the quantity traded in Stage 2, corresponding to $w_I < w(q) = MR_D(q)$.

4.2.1 When Bargaining Leads to Bilateral Efficiency

If both U and D announce the same quantity q_I in Stage 2, this implies that, in Stage 1, the wholesale price must satisfy $w = w_I = MR_D(q_I) = MC_U(q_I)$. Using $w = MR_D(q)$, we rewrite (7) and evaluate it at q_I, yielding:[36]

$$\alpha \underbrace{(MR_U(q_I) - MC_U(q_I))}_{\pi'_U(q_I)} \pi_D(q_I) + (1-\alpha) \underbrace{(MR_D(q_I) - MR_U(q_I))}_{\pi'_D(q_I)} \pi_U(q_I) = 0 \qquad (8)$$

As $w_I = MR_D(q_I) = MC_U(q_I)$, it follows that $\pi'_U(q_I) = -\pi'_D(q_I)$. The equality of firms' concession costs implies that profit can be transferred between U and D without reducing their joint profits. Thus, there is no bilateral inefficiency in equilibrium—that is, $\pi'_U(q_I) + \pi'_D(q_I) = 0$—and (8) 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. 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:*

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

[36] We could alternatively use $w(q) = MC_U(q)$ and evaluate (7) at q_I, as both first-order conditions are equivalent in q_I (see Appendix A.3.1 for details).

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 supply chain margin is given by $\mathcal{M} = \frac{p(q_I)}{r(q_I)} = \nu_U \times \mu_D$.

Proof. See Appendix A.3.2. ■

Proposition 2 provides two expressions for α_I, highlighting the conditions under which profit sharing and firms' incentives to trade are efficiently aligned. The first expression defines α_I as U's share of the joint profit (or gain from trade) when the equilibrium quantity is q_I: $\alpha_I = \frac{\pi_U(q_I)}{\pi_U(q_I)+\pi_D(q_I)}$.[37] The second expression, $\alpha_I = \frac{\varepsilon_p - 1}{\varepsilon_p + \varepsilon_r}$, depends solely on the supply and demand primitives. To gain intuition, consider the case where $\varepsilon_r(q_I)$ is low relative to $\varepsilon_p(q_I)$. In this case, U's incentive to exert monopsony power is stronger than D's incentive to exert monopoly power. Given U's greater incentive to reduce the quantity traded, bilateral efficiency (i.e., trading q_I) requires granting U a relatively high wholesale price (i.e., α_I is close to 1). The logic is symmetric when $\varepsilon_p(q_I)$ is low relative to $\varepsilon_r(q_I)$. A special case arises when $\varepsilon_r \to \infty$, so that U faces constant marginal costs as in the canonical bilateral monopoly model. With no scope for monopsony power, bilateral efficiency is achieved only when D holds all the bargaining power (i.e., $\alpha_I = 0$).

From Lemma 1, we know that for any $w \neq w_I$, the traded quantity falls below q_I. Consequently, any shift of bargaining power in favor of D (i.e., $\alpha < \alpha_I$) or U (i.e., $\alpha > \alpha_I$) induces a distortion in the vertical supply chain. We analyze these two cases in turn below.

4.2.2 When Bargaining Leads to Double Markdownization

If both firms anticipate that U chooses the quantity in Stage 2, this implies that, in Stage 1, the wholesale price satisfies $w < w_I$. In this case, there is a positive relationship between the negotiated wholesale price and the quantity traded, pinned down by $w =$

[37] In our baseline analysis, both U's and D's status quo profit are equal to zero. Suppose instead that U's (resp. D's) status quo profit equals $\overline{\pi}_U$ (resp. $\overline{\pi}_D$). In this case, the negotiation leads to bilateral efficiency when $\alpha = \alpha_E = \frac{\pi_U(q_I)-\overline{\pi}_U}{\pi_U(q_I)-\overline{\pi}_U+\pi_D(q_I)-\overline{\pi}_D}$, where the subscript E stands for efficiency. Hence, when $\frac{\pi_U-\overline{\pi}_U(q_I)}{\pi_D(q_I)-\overline{\pi}_D(q_I)} = \frac{\pi_U(q_I)}{\pi_D(q_I)}$, we have $\alpha_I = \alpha_E$. However, if status quo profits increase U's relative gain from trade compared to D (say, by raising $\overline{\pi}_D$), then $\alpha_E > \alpha_I$. In this case, U's bargaining position is weakened, narrowing the range of $\alpha \in]\alpha_E, 1]$ for which U is considered powerful.

$MC_U(q)$. We thus have $\pi_U(q) = MC_U(q)q - r(q)q$ and $\pi_D(q) = p(q)q - MC_U(q)q$. As $MC_U(q)$ is increasing in q, it follows that $\pi'_U(q) > 0$ and $\pi'_D(q) < 0$ over the relevant range of conflict for bargaining. More precisely, the first-order condition (7) becomes:

$$\alpha \underbrace{(MC_D(q_\nu) - MC_U(q_\nu))}_{\pi'_U(q_\nu)} \pi_D(q_\nu) + (1-\alpha) \underbrace{(MR_D(q_\nu) - MC_D(q_\nu))}_{\pi'_D(q_\nu)} \pi_U(q_\nu) = 0 \tag{9}$$

where q_ν denotes the equilibrium quantity. As $MC_U(q_\nu) < MR_D(q_\nu)$, we have $\pi'_U(q_\nu) > -\pi'_D(q_\nu)$, meaning that U's concession cost exceeds that of D in equilibrium. Bilateral efficiency would thus require D to concede a more favorable trading term to U, resulting in a larger quantity being traded. However, as D is powerful, it can drive the wholesale price below w_I, implying that $q_\nu < q_I$. To gain further insight, we rearrange (7) as follows:

$$MR_D(q_\nu) = \widetilde{MC}_D(q_\nu, \alpha) \tag{10}$$

where $\widetilde{MC}_D(q, \alpha) \equiv \beta_U(q, \alpha) MC_U(q) + (1 - \beta_U(q, \alpha)) MC_D(q)$ and $\beta_U(q, \alpha) \equiv \frac{\alpha}{1-\alpha} \frac{\pi_D(q)}{\pi_U(q)}$. In equilibrium, $\beta_U(q_\nu, \alpha) = -\frac{\pi'_D(q_\nu)}{\pi'_U(q_\nu)} \in [0, 1)$.[38] Thereby, $\beta_U(q_\nu, \alpha)$ measures U's degree of countervailing seller power, which increases in its relative bargaining weight ($\frac{\alpha}{1-\alpha}$) and D's relative gains from trade vis-à-vis U ($\frac{\pi_D}{\pi_U}$). An increase in $\beta_U(q_\nu, \alpha)$ reflects a concession from D to U, narrowing the gap between their concession costs and pushing q_ν closer to q_I.[39] Specifically, when D holds all the bargaining power, we have $\beta_U(q_{\bar{\nu}}, 0) = 0$, implying that (10) boils down to $MC_D(q_{\bar{\nu}}) = MR_D(q_{\bar{\nu}})$—i.e., D makes a take-it-or-leave-it offer to U. When $\alpha \in (0, \alpha_I)$, we have $\beta_U(q_\nu, \alpha) > 0$, which shifts $\widetilde{MC}_D(q_\nu, \alpha)$ towards $MC_U(q_\nu)$ such that $\widetilde{MC}_D(q_\nu, \alpha) \in (MC_D(q_\nu), MC_U(q_\nu))$. As $MR_D(q)$ decreases in q, the equilibrium quantity q_ν characterized by (10) increases, thereby reducing the inefficiency. Finally, when α tends to α_I, we have $\beta_U(q_\nu, \alpha_I) = 1$, and (10) reduces to $MC_U(q_\nu) = MR_D(q_\nu)$, yielding the vertically integrated outcome $q_\nu = q_I$. Based on this reasoning, we derive the following proposition:

---

[38] This is a direct rewriting of (9). See Appendix A.4.4 for a proof that $\beta_U(q_\nu, \alpha) \in [0, 1)$.

[39] As $\pi'_U(q_\nu) = MC'_U(q_\nu)q_\nu$ and $-\pi'_D(q_\nu) = MC'_U(q_\nu)q_\nu - [MR_D(q_\nu) - MC_U(q_\nu)]$, one can see that the gap between both concession costs, and therefore the inefficiency, increases in the wedge between $MR_D(q_\nu)$ and $MC_U(q_\nu)$.

Proposition 3 (Double Markdownization) *When D is powerful ($\alpha < \alpha_I$), 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, which charges a markdown given by:*

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

and adds up to U'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 ($\mu_U = \frac{w(q_\nu)}{MC_U(q_\nu)} = 1$). 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 supply chain margin is given by:

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

Proof. Appendix A.4.2 derives the expression for D's markdown, and Appendix A.4.3 characterizes the set of equilibria. Appendix A.8.2 provides formal derivations of the markup and markdown expressions based on the definitions introduced in Section 2. ■

Proposition 3 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*, which leads to the inefficient outcome $q_\nu < q_I$, is hereby identified as a novel source of double marginalization.

As in the vertically integrated outcome, U's markdown (ν_U) and D's markup (μ_D) are governed 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, D's markdown (ν_D) depends on two other factors. First, it decreases with U's countervailing seller power (β_U), ranging from $\nu_D = \frac{\varepsilon_{MC_U}+1}{\varepsilon_{MC_U}}$ when $\beta_U = 0$ to $\nu_D = 1$ when $\beta_U = 1$. As β_U increases with α and with D's relative gain from trade vis-à-vis U, ν_D also declines with both bargaining forces. Furthermore, because

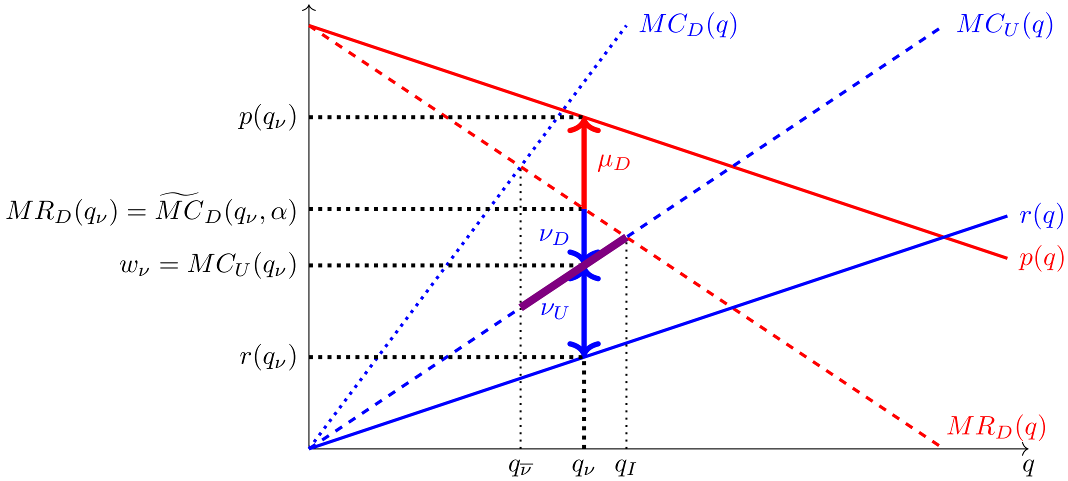

Figure 3: Equilibrium with Double Markdown ($0 < \alpha < \alpha_I$).

Notes: The figure is drawn under the demand function $p(q) = p - \frac{1}{3}q$ and supply function $r(q) = r + \frac{1}{3}q$. The purple segment represents the set of equilibrium wholesale price-quantity pairs, with $w_\nu = MC_U(q_\nu)$, as α ranges from 0 to α_I. The red arrow labeled μ_D and the blue arrows labeled ν_D and ν_U represent, respectively, the markup and markdown wedges in differences.

D's relative gain from trade vis-à-vis U increases with μ_D, D's markup and markdown are negatively related.[40] Second, in the same way that ν_U decreases with the input supply elasticity (ε_r), ν_D declines with U's supply elasticity (ε_{MC_U}).[41]

Figure 3 illustrates the equilibrium described in Proposition 3 for a given $\alpha \in (0, \alpha_I)$, using linear demand and supply functions. In equilibrium, the wholesale price is $w_\nu = MC_U(q_\nu)$, where the traded quantity q_ν satisfies (10). The purple segment represents the set of equilibrium wholesale price–quantity pairs as α ranges from 0 (left endpoint) to α_I (right endpoint). When $\alpha = 0$, $\beta_U(q_\nu, \alpha) = 0$ and (10) boils down to $MR_D(q_\nu) = MC_D(q_\nu)$, defining the equilibrium quantity $q_{\overline{\nu}}$. When $\alpha = \alpha_I$, $\beta_U(q_\nu, \alpha_I) = 1$ and (10) boils down to $MR_D(q_\nu) = MC_U(q_\nu)$, yielding q_I.

4.2.3 When Bargaining Leads to Double Markupization

If both firms anticipate that D chooses the quantity in Stage 2, this implies that, in Stage 1, the wholesale price satisfies $w > w_I$. In this case, there is a negative relationship between the negotiated wholesale price and the quantity traded, pinned

[40] See Online Appendix OA1 for further details.

[41] As $\varepsilon_{MC_U} = \frac{\varepsilon_r + 1}{\sigma_r + 2}$, it further implies that ν_D decreases with the input supply elasticity ε_r and increases with its curvature σ_r.

down by $w = MR_D(q)$. We thus have $\pi_U(q) = MR_D(q)q - r(q)q$ and $\pi_D(q) = p(q)q - MR_D(q)q$. As $MR_D(q)$ is decreasing in q, it follows that $\pi'_U(q) < 0$ and $\pi'_D(q) > 0$ over the relevant range of conflict for bargaining. More precisely, the first-order condition (7) becomes:

$$\alpha \underbrace{\left(MR_U(q_\mu) - MC_U(q_\mu)\right)}_{\pi'_U(q_\mu)} \pi_D(q_\mu) + (1-\alpha) \underbrace{\left(MR_D(q_\mu) - MR_U(q_\mu)\right)}_{\pi'_D(q_\mu)} \pi_U(q_\mu) = 0 \quad (11)$$

where q_μ denotes the equilibrium traded quantity. As $MC_U(q_\mu) > MR_D(q_\mu)$, we have $\pi'_U(q_\mu) < -\pi'_D(q_\mu)$, meaning that D's concession cost exceeds that of U in equilibrium. Bilateral efficiency would thus require U to concede a more favorable trading term to D, resulting in a larger quantity being traded. However, as U is powerful, it can drive the wholesale price above w_I, implying that $q_\mu < q_I$. As in Section 4.2.2, we can rearrange (11) as follows:

$$MC_U(q_\mu) \quad = \quad \widetilde{MR}_U(q_\mu, \alpha) \qquad (12)$$

where $\widetilde{MR}_U(q,\alpha) \equiv \beta_D(q,\alpha)MR_D(q) + (1-\beta_D(q,\alpha))\, MR_U(q)$ and $\beta_D(q,\alpha) \equiv \frac{1-\alpha}{\alpha}\frac{\pi_U(q)}{\pi_D(q)}$. In equilibrium, $\beta_D(q_\mu, \alpha) = -\frac{\pi'_U(q_\mu)}{\pi'_D(q_\mu)} \in [0,1)$.[42] Thereby, $\beta_D(q_\mu,\alpha)$ measures D's degree of countervailing buyer power. The reasoning is analogous to that described in Section 4.2.2. When $\beta_D(q_{\overline{\mu}}, 1) = 0$, U makes a take-it-or-leave-it offer to D and the equilibrium quantity $q_{\overline{\mu}}$ is defined by $MC_U(q_{\overline{\mu}}) = MR_U(q_{\overline{\mu}})$. As $MC_U(q)$ increases in q, the equilibrium quantity q_μ characterized by (12) increases, thereby reducing the inefficiency. Finally, when α tends to α_I, we have $\beta_D(q_\mu, \alpha_I) = 1$, implying that (12) reduces to $MC_U(q_\mu) = MR_D(q_\mu)$. This last case corresponds to the vertically integrated outcome, where $q_\mu = q_I$. Based on this reasoning, we derive the following proposition:

Proposition 4 (Double markupization) *When U is powerful ($\alpha > \alpha_I$), 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*

[42] This is a direct rewriting of (11). See Appendix A.5.4 for a proof that $\beta_D(q_\mu, \alpha) \in [0,1)$.

arises from U's seller power, which charges a markup given by:

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

and 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* ($\nu_D = \frac{MR_D(q_\mu)}{w_\mu} = 1$)*. 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 supply chain margin is given by:*

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

Proof. Appendix A.5.2 derives the main expression for U's markup, and Appendix A.5.3 characterizes the set of equilibria. Appendix A.8.2 provides formal derivations of the markup and markdown expressions based on the definitions introduced in Section 2. ■

Proposition 4 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 Cournot-Spengler double marginalization phenomenon, leading to an inefficient outcome ($q_\mu < q_I$).

Interestingly, U's markup μ_U depends on two factors. First, μ_U decreases with D's countervailing buyer power (β_D), ranging from $\mu_U = \frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D}-1}$ when $\beta_U = 0$ to $\mu_U = 1$ when $\beta_D = 1$. As β_D increases with $1-\alpha$ and with U's relative gain from trade vis-à-vis D, μ_U also decreases with these two bargaining forces. Moreover, as U's relative gain from trade vis-à-vis D increases with ν_U, there is a negative relationship between ν_U and μ_U.[43] Second, just as μ_D decreases with the elasticity of consumer demand (ε_p), μ_U reduces with D's demand elasticity (ε_{MR_D}).[44]

---

[43]Using $\nu_U = \frac{\varepsilon_r+1}{\varepsilon_r}$ and rewritting $\mu_U = \frac{\alpha\varepsilon_{MR_D}\nu_U+(1-\alpha)\varepsilon_p}{\nu_U\alpha(\varepsilon_{MR_D}-1)+(1-\alpha)(\varepsilon_p-1)}$, it follows that $\frac{\partial\mu_U}{\partial\nu_U} = -\frac{(1-\alpha)(\varepsilon_p-1)}{\nu_U^2(\varepsilon_p-1+\alpha(\varepsilon_{MR_D}-\varepsilon_p))} \leq 0$ (see Online Appendix OA1 for further details).

[44]As $\varepsilon_{MR_D} = \frac{\varepsilon_p-1}{2-\sigma_p}$, it further implies that μ_U decreases with the consumer demand elasticity ε_p and increases with its curvature σ_p.

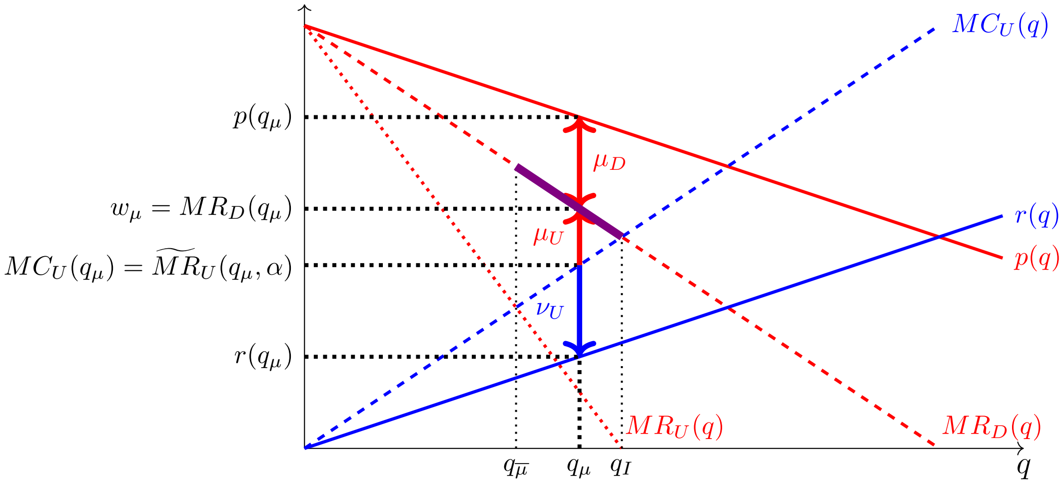

Figure 4: Equilibrium with Double Markup ($\alpha_I < \alpha < 1$).

Notes: The figure is drawn under the demand function $p(q) = p - \frac{1}{3}q$ and supply function $r(q) = r + \frac{1}{3}q$. The purple segment represents the set of equilibrium wholesale price–quantity pairs, with $w_\mu = MR_D(q_\mu)$, as α ranges from α_I to 1. The red arrows labeled μ_D and μ_U, and the blue arrow labeled ν_U represent, respectively, the markup and markdown wedges in differences.

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_μ lies on MR_D and satisfies (12), which determines q_μ. The set of equilibrium wholesale price-quantity pairs is represented by the purple segment as α ranges from α_I (right endpoint) to 1 (left endpoint).

4.2.4 Taking Stock

We summarize the key findings from Propositions 2 to 4. First, there exists a unique bargaining weight $\alpha_I \in [0, 1]$ such that U's seller power and D's buyer power exactly offset each other ($\mu_U = 1$ and $\nu_D = 1$), thereby achieving bilateral efficiency (Proposition 2). For $\alpha \neq \alpha_I$, bilateral inefficiency arises either from D's excessive buyer power ($\alpha < \alpha_I$), resulting in double markdownization (Proposition 3), or U's excessive seller power ($\alpha > \alpha_I$), resulting in double markupization (Proposition 4).

Interestingly, for given demand and supply primitives, there exist two distinct values of $\alpha \neq \alpha_I$ (i.e., one below and one above α_I) that lead to the same inefficient quantity $q < q_I$.[45] Both equilibria yield identical consumer and input supplier prices

[45] Figure 5 illustrates this point in the case of linear demand and supply functions.

($p(q)$ and $r(q)$) as well as the same degree of market power exerted at the two ends of the vertical supply chain (ν_U and μ_D). However, and most importantly, they differ in the wholesale price and in the nature of the distortion: one equilibrium involves double markdownization ($\nu_D > 1$), whereas the other features double markupization ($\mu_U > 1$). In what follows, we show that identifying whether inefficiency stems from double markup or double markdown is crucial for the design of policy interventions.

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).[46] We formalize the effects of such variations in the following corollary, which is illustrated in Figure 5, where the purple (resp. green) segment depicts the set of equilibrium pairs (w, q) as α ranges from 1 (resp. 0) to α_I and the arrows indicate the direction of the variation:

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

- *When U is powerful ($\alpha > \alpha_I$), an increase in D's bargaining power countervails U's seller power: both U's markup μ_U and the supply chain margin $\mathcal{M} = \nu_U \times \mu_U \times \mu_D$ decline, increasing the quantity traded q_μ and welfare.*

- *When D is powerful ($\alpha < \alpha_I$), an increase in U's bargaining power countervails D's buyer power: both D's markdown ν_D and the supply chain margin $\mathcal{M} = \nu_U \times \nu_D \times \mu_D$ decline, increasing the quantity traded q_ν and welfare.*

- *When $\alpha = \alpha_I$, U's seller power and D's buyer power fully countervail each other: $\mu_U = \nu_D = 1$, and both the supply chain margin $\mathcal{M} = \nu_U \times \mu_D = M_I$ and welfare reach their vertical integration value.*

[46] Shifts in α are exogenous changes in the distribution of bargaining power along the vertical supply chain. More broadly, changes in the distribution of bargaining power can arise from various sources affecting firms' relative gains from trade, 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.

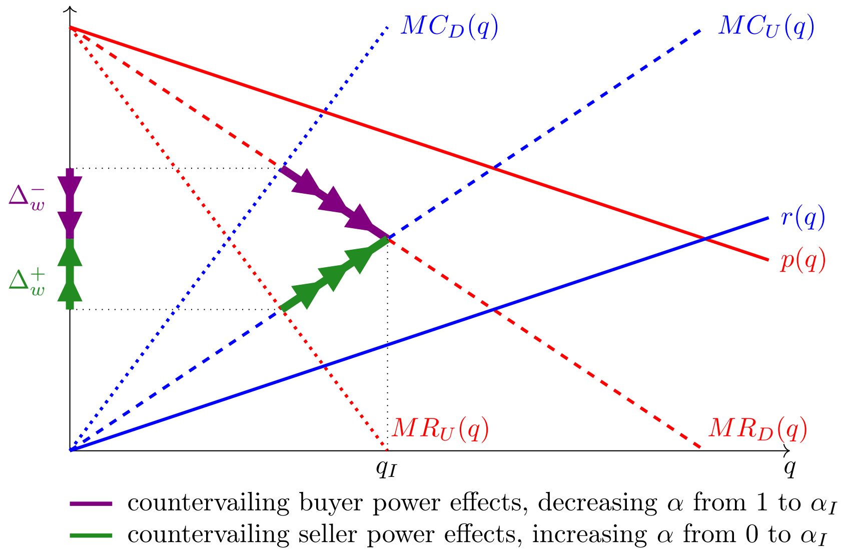

Figure 5: Effects of Countervailing Buyer and Seller Power.

Notes: The figure is drawn under the demand function $p(q) = p - \frac{1}{3}q$ and supply function $r(q) = r + \frac{1}{3}q$.

The countervailing buyer power effect that emerges 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, higher $\beta_D(q_\mu, \alpha)$), which mitigates double marginalization by reducing U's markup, and, ultimately, the supply chain margin. Corollary 1 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 3). Consequently, further increases in D's bargaining power exacerbate double marginalization by raising D's markdown. Instead, the countervailing seller power theory applies: increasing U's bargaining power (higher α, higher $\beta_U(q_\nu, \alpha)$) offsets D's monopsony distortion, thereby reducing D's markdown and ultimately the supply chain margin. In the case where $\alpha = \alpha_I$, U's seller power and D's buyer power fully countervail each other. As a result, the supply chain reaches the vertical integration outcome, achieving bilateral efficiency and maximizing welfare.[47]

Corollary 1 indicates that the welfare effects of buyer and seller power depend on the nature of the distortion arising in equilibrium (i.e., double markup or markdown). To gain further insights on how countervailing buyer or seller power affects the different

[47]It is worth noting that welfare is maximized conditional on U's monopsony power and D's monopoly power, yielding a second-best outcome where $q_I < q_W$.

welfare components, we establish the following corollary:

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

- *If U is powerful ($\alpha_I < \alpha$), increasing D's bargaining power benefits consumers, input suppliers, and D, but reduces U's profit.*
- *If D is powerful ($\alpha < \alpha_I$), increasing U's bargaining power benefits consumers, input suppliers, and U, but reduces D's profit.*

Proof. See Appendix A.6. ■

Corollary 2 uncovers an additional channel through which countervailing buyer power improves welfare: when U is powerful ($\alpha > \alpha_I$), an increase in D's bargaining power raises the quantity exchanged and, in turn, increases input suppliers' surplus. As the countervailing buyer power theory has typically been formalized in settings with constant marginal costs, to the best of our knowledge, this latter effect has not been previously identified in the literature.[48] Moreover, the corollary also highlights that both countervailing buyer and seller power hurt the powerful firm in the vertical supply chain and benefit all other agents.

By affecting the traded quantity, it is worth noting that changes in α also indirectly influence D's markup and U's markdown. These indirect effects, which crucially depend on the shape of supply and demand functions, may either amplify or attenuate the distortion stemming from U's monopsony power and D's monopoly power. However, as established in the following remark, these effects are second-order:

Remark 1 *The welfare gains from moving α toward α_I are smaller (resp. larger) when the demand and supply functions are subconvex (resp. superconvex).*

Proof. See Appendix A.7. ■

To illustrate, consider the case where U is powerful. As stated in Corollary 1, any decrease in α mitigates U's markup and the double marginalization phenomenon

[48]The input supply function reflects the aggregation of heterogeneous individual supply decisions (whether by firms or workers). Accordingly, an increase in the quantity traded can raise input suppliers' surplus both by increasing supply among those already active (intensive margin) and by drawing new participants into the production process (extensive margin).

described in Proposition 4.[49] However, the extent to which the resulting decrease in U's markup is passed on to the input and consumer prices ultimately depends on the shape of supply and demand functions. If demand is subconvex ($\frac{\partial \varepsilon_p}{\partial q} < 0$), D's markup increases with q, implying incomplete pass-through to the consumer price.[50] Likewise, if supply is subconvex ($\frac{\partial \varepsilon_r}{\partial q} < 0$), U's markdown increases with q, implying incomplete pass-through to the input price. Hence, the welfare gains from reducing U's markup by decreasing α are smaller under subconvex demand and supply. By contrast, these gains are larger under superconvex demand and supply (i.e., $\frac{\partial \varepsilon_p}{\partial q} > 0$ and $\frac{\partial \varepsilon_r}{\partial q} > 0$, respectively), as both U's markdown and D's markup decrease with q.

6 Price Floor Regulation

It is well known that minimum wages can increase employment in the presence of monopsony power (e.g., Robinson, 1933; Stigler, 1946). More broadly, price floors are often used as a policy tool to counteract the market power of firms when purchasing their inputs.[51] In this section, we use our bilateral monopoly model to examine the welfare effects of an input price floor regulation. Specifically, we determine the optimal price floor policy, defined as the price floor level that maximizes the equilibrium quantity traded. By symmetry, the analysis applies to a cap on the consumer price.[52]

We begin by characterizing the optimal price floor under vertical integration, which serves as a benchmark for analyzing this policy in our vertical supply chain framework with bargaining. In both cases, we consider a price floor $\underline{r}$ that affects the input price schedule as follows. The price floor is binding whenever U purchases a quantity $q \leq \underline{q}$ at $\underline{r}$, where $\underline{q}$ is the threshold quantity such that $\underline{r} = r(\underline{q})$. When U purchases a quantity

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

[50] 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" as 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, it is consistent with recent empirical evidence from Boehm and Pandalai-Nayar (2022) for U.S. industries and Avignon and Guigue (2022) for the French milk market.

[51] To support farmer revenues, many countries have introduced either temporary or permanent price floors in agricultural markets (e.g., the U.S. raw milk market). See Avignon and Guigue (2025) for further discussion.

[52] Price caps on food products are often proposed as a way to protect consumers, especially during periods of high inflation (e.g., Aparicio and Cavallo, 2021).

$q > \underline{q}$, the price floor is not binding, and the input price reverts to $r(q)$. Accordingly, the input supply curve becomes perfectly elastic for $q \leq \underline{q}$ (i.e., $\varepsilon_r \to \infty$), and remains unchanged otherwise.

6.1 Vertical Integration

We analyze the optimal input 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—that is, $MC_I(q) = \underline{r}$ when $q \leq \underline{q}$. We obtain the following proposition:

Proposition 5 (Optimal Input Price Floor under Vertical Integration) *Under vertical integration, the optimal input price floor is* $\underline{r_I} = r(\underline{q_I})$*, where* $\underline{q_I}$ *is defined by* $MR_I(\underline{q_I}) = r(\underline{q_I})$*. The price floor* $\underline{r_I}$ *increases the equilibrium quantity,* $\underline{q_I} \in (q_I, q_W)$*, decreases the equilibrium consumer price,* $p(\underline{q_I}) \in (p(q_I), p(q_W))$*, and increases the equilibrium input price,* $r(\underline{q_I}) \in (r(q_I), r(q_W))$*. I's markdown is eliminated,* $\underline{\nu_I} = \frac{MR_I(\underline{q_I})}{r(\underline{q_I})} = 1$ *and I's markup and margin are given by* $\underline{\mu_I} = \frac{p(\underline{q_I})}{MC_I(\underline{q_I})} = \frac{\varepsilon_p}{\varepsilon_p - 1}$*, and* $\underline{M_I} = \frac{p(\underline{q_I})}{r(\underline{q_I})} = \underline{\mu_I}$*, respectively.*

Proof. See Appendix C.1. ■

Raising the price floor $\underline{r}$ involves the following trade-off. On the one hand, it increases the cost perceived by I for any $q < \underline{q}$, which reduces the quantity I is willing to purchase. On the other hand, it also increases the quantity threshold $\underline{q}$ below which I faces a flat supply curve, preventing the exercise of monopsony power. As in Hernández and Cantillo-Cleves (2024), the optimal price floor is such that I's demand meets what input suppliers are willing to offer at that price floor, that is, $\underline{r_I} = MR_I(\underline{q_I}) = r(\underline{q_I})$. To see this, consider a (binding) price floor $\underline{r}$ set below $\underline{r_I}$. We obtain $\underline{r} = r(\underline{q}) < MR_I(\underline{q})$, implying that I is willing to purchase more than the quantity input suppliers are willing to offer at $\underline{r}$. However, as $MR_I(\underline{q}) < MC_I(\underline{q})$, I will not purchase more than $\underline{q}$. Thus, I's profit-maximizing quantity under $\underline{r}$ is $\underline{q}$, and the traded quantity can be increased by raising the price floor. In contrast, if the price floor $\underline{r}$ is set above $\underline{r_I}$, then I is willing to purchase less than $\underline{q_I}$ as $MR_I(q)$ is decreasing.[53]

[53] More precisely, I's profit-maximizing quantity is such that $MR_I(q) = \underline{r} > \underline{r_I} = MR_I(\underline{q_I})$.

In equilibrium, the optimal price floor eliminates the monopsony distortion ($\underline{\upsilon_I} = 1$) by making input supply perfectly elastic, thereby improving welfare ($\underline{q_I} > q_\mu$). However, as consumer demand remains unaffected, monopoly power persists, with $\underline{\mu_I} = \frac{p(\underline{q_I})}{r(\underline{q_I})} > 1$. This remaining distortion depresses the optimal price floor below the competitive input price level ($\underline{r} < r(q_W)$), resulting in a quantity below the competitive level, $\underline{q_I} < q_W$. This wedge narrows as demand becomes more elastic. In the limit case where $\varepsilon_p \to \infty$, the optimal price floor implements the competitive allocation, as $\underline{q_I} \to q_W$.

6.2 Vertical Supply Chain

We now examine the optimal input price floor policy within our bilateral monopoly model. When the price floor is binding, the marginal costs of U and D become constant, eliminating their ability to exert monopsony power. Specifically, as U's marginal constant marginal is flat for all $q \leq \underline{q}$, it is willing to supply $\underline{q}$ at any $w > \underline{r}$. Thus, when anticipating that the price floor is binding, U and D bargain under the expectation that D will subsequently determine the quantity traded ($w(q) = MR_D(q)$). Note that this coincides with the canonical model of vertical relationships with double markup and constant marginal costs. In this context, the optimal price floor policy is characterized in the following proposition:

Proposition 6 (Optimal Input Price Floor in a Vertical Supply Chain) *For a given α, the optimal input price floor is $\underline{r_\mu} = r(\underline{q_\mu})$, where $\underline{q_\mu}$ is defined by $\widetilde{MR}_U(\underline{q_\mu}, \alpha) = r(\underline{q_\mu})$. The wholesale price is $\underline{w_\mu} = MR_D(\underline{q_\mu}) > \underline{w_I}$, the quantity exchanged and welfare increase, $\underline{q_\mu} \in (q^\star, \underline{q_I})$, consumer surplus increases, $p(\underline{q_\mu}) \in (p(q^\star), p(\underline{q_I}))$, and input supplier surplus increases, $r(\underline{q_\mu}) \in (r(q^\star), r(\underline{q_I}))$, where $q^\star = q_\nu$ if $\alpha < \alpha_I$ and $q^\star = q_\mu$ otherwise. Neither U nor D charges a markdown, $\underline{\nu_U} = \underline{\nu_D} = 1$, but double marginalization persists due to U's seller power, resulting in 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* $\underline{M_U} = \frac{w_\mu}{r(\underline{q_\mu})} = \underline{\mu_U}$*, D's margin is equal to* $\underline{M_D} = \frac{p(\underline{q_\mu})}{\underline{w_\mu}} = \underline{\mu_D}$*, and the supply chain margin is equal to* $\underline{\mathcal{M}} = \frac{p(\underline{q_\mu})}{r(\underline{q_\mu})} = \underline{\mu_U} \times \underline{\mu_D}$.

Proof. See Appendix C.2.1. ■

Several comments are in order. We first focus on the characterization of the optimal price floor policy before discussing its welfare implications.

Under vertical integration, Proposition 5 establishes that the optimal price floor is uniquely determined by demand and supply primitives, with residual inefficiency stemming from I's monopoly power. In a vertical supply chain, Proposition 6 highlights that the optimal price floor also depends on the distribution of bargaining power between U and D, as the remaining distortion arises from double markupization, which increases with α. Specifically, the price floor is such that the quantity the vertical supply chain (U and D) is willing to purchase equals the quantity input suppliers are willing to offer at the price floor—that is, $\widetilde{MR}_U(\underline{q_\mu}) = \underline{r_\mu} = r(\underline{q_\mu})$. Following the same reasoning as under vertical integration, a higher price floor would reduce the quantity the vertical chain is willing to trade, and a lower (binding) price floor would instead reduce the quantity input suppliers are willing to offer at that price.

As in the vertical integration case, the optimal price floor eliminates the monopsony distortion along the vertical supply chain ($\underline{\nu_U} = \underline{\nu_D} = 1$). However, the monopoly distortion persists in the form of double markupization, with $\underline{\mathcal{M}} = \frac{p(\underline{q_\mu})}{r(\underline{q_\mu})} > 1$. This remaining distortion explains the dependence of the optimal price floor on α, which is summarized in the following corollary:

Corollary 3 *The optimal input price floor* $\underline{r}(\underline{q_\mu})$ *and equilibrium quantity* $\underline{q_\mu}$ *decrease in* α.

Proof. See Appendix C.2.2. ■

The remaining markup distortion is greater than under vertical integration, implying that $\underline{r_\mu} \leq \underline{r_I}$ and $\underline{q_\mu} \leq \underline{q_I}$. The resulting efficiency loss decreases with D's countervailing buyer power (i.e., as α decreases), as highlighted in Proposition 4. In particular, in the limit case where $\alpha = 1$, double markup distortion is maximized, the optimal price floor is at its lowest level, and the equilibrium quantity is such that $MR_U(\underline{q}_{\overline{\mu}}) = r(\underline{q}_{\overline{\mu}})$. Conversely, when $\alpha = 0$, double markupization is eliminated, and the vertically integrated outcome under optimal price floor is achieved as the equilibrium quantity is such that $MR_D(\underline{q}_{\mu}) = r(\underline{q}_{\mu})$, with $\underline{q}_{\mu} = \underline{q}_I$.

Proposition 6 also establishes that the optimal price floor always benefits input suppliers, consumers, and welfare. Its welfare-improving effects arise from the elimination of the monopsony distortion, which increases the traded quantity and yields more favorable prices for both consumers and input suppliers relative to the situation without the price floor (i.e, a lower consumer price and a higher input price). These welfare gains are the largest when D is powerful ($\alpha < \alpha_I$), as the primary source of inefficiency in the absence of a price floor stems from monopsony power (double markdownization).[54] Importantly, the policy reverses the welfare implications of D's bargaining power: whereas it exacerbates double markdownization in the absence of regulation, it countervails U's seller power under the price floor. Hence, the welfare gains from the price floor policy decline with $\alpha \in [0, \alpha_I]$. This reasoning implies that the welfare gains from the policy are more modest when U is powerful ($\alpha > \alpha_I$). In the limiting case $\varepsilon_r \to \infty$ (i.e., $\alpha_I \to 0$), the canonical model of vertical relationships emerges, where inefficiency stems exclusively from double markup. Absent any monopsony distortion, a welfare-improving price floor policy is not feasible: any price floor set above the equilibrium input price would raise U's (constant) marginal cost, reduce the traded quantity, and lower welfare. This result underscores the importance of identifying the nature of double marginalization, as knowing the magnitude of the distortion but not its type (i.e., double markupization or markdownization) may lead to a suboptimal, or even detrimental, price floor regulation.

Finally, it is worth noting that the effects of the optimal price floor on firms' profits are ambiguous. When U is powerful ($\alpha > \alpha_I$), the optimal price floor always hurts U and benefits D (see Appendix C.2.3 for details). However, when D is powerful ($\alpha < \alpha_I$), the effects on U's and D's profits depend on the specific forms of the demand and supply functions. Under linear demand and supply, we obtain that the optimal price floor always reduces U's profit, but its effect on D's profit remains ambiguous.

[54]Figure 7 in Appendix C.2.4 provides an illustrative example.

7 Discussion

In what follows, we discuss the robustness of our main findings to two key modeling assumptions: that U and D bargain over a linear wholesale price (Section 7.1), and that U's increasing marginal cost arises from the presence of imperfectly elastic input suppliers (Section 7.2).

7.1 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).[55] However, double marginalization may persist when 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 what follows, we demonstrate that our main findings remain valid when 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).[56] 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 this section.

Under a two-part tariff contract, the profit functions for U and D are defined as $\Pi_U(q) \equiv (w(q) - r(q))\, q + F = \pi_U(q) + F$ and $\Pi_D(q) \equiv (p(q) - w(q))\, q - F = \pi_D(q) - F$, respectively. We allow F to be either positive (transfer from D to U) or negative (transfer from U to D), but assume the following restriction: $\underline{F} \leq F \leq \overline{F}$. 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 $\underline{F}$.[57]

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

[56] Importantly, the preservation of our results hinges on the fact that w is used as a surplus-sharing tool. This contrasts with settings where the upward distortion on w serves other purposes, such as incentivizing efforts to maximize U's and D's joint profit (e.g., de Cornière and Taylor, 2021).

[57] This modeling assumption reflects 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

The sequence of play mirrors that described in Section 4. Specifically, in Stage 1, U and D negotiate a two-part tariff contract (w, F). In Stage 2, firms simultaneously announce the quantities they are willing to trade and, due to voluntary exchange, the quantity traded is the minimum of the two announced quantities. As the fixed fee F never affects firms' quantity choice, the resolution of Stage 2 is similar to that described in Lemma 1, where the quantity traded in equilibrium is $q(w) = \min\{\tilde{q}_U(w), \tilde{q}_D(w)\}$. 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 following maximization problem:[58]

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

Three types of equilibria may arise, depending on whether: (i) the constraint on F is not binding, (ii) the upper bound is binding ($F = \overline{F}$), or (iii) the lower bound is binding ($F = \underline{F}$). For the sake of conciseness, we refer to Appendix D for details and summarize our results in the following proposition:

Proposition 7 (Two-Part Tariff 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 a two-part tariff is characterized as follows:*

(i) When $\underline{\alpha} < \alpha < \overline{\alpha}$, with $\underline{\alpha} \equiv \frac{\pi_U(q_I)+\underline{F}}{\pi_U(q_I)+\pi_D(q_I)}$ and $\overline{\alpha} \equiv \frac{\pi_U(q_I)+\overline{F}}{\pi_U(q_I)+\pi_D(q_I)}$, an equilibrium replicating the vertically integrated outcome arises.

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

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

---

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.

[58]It is worth noting that maximizing the Nash product with respect to (q, F), given that $w(q) = MR_D(q)$ or $w(q) = MC_U(q)$, is equivalent to maximizing with respect to (w, F).

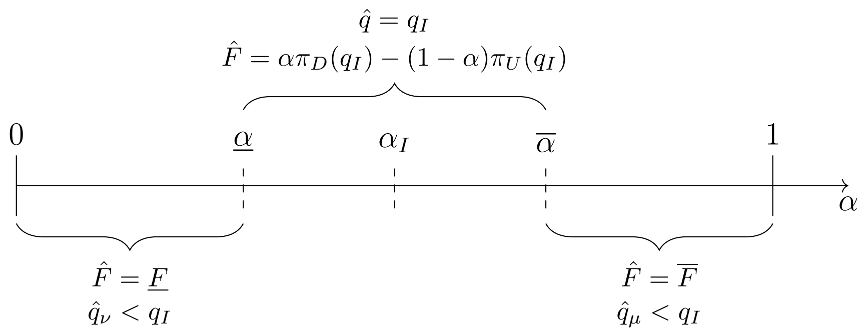

Figure 6: Equilibrium quantity and fixed fee in the presence of frictions.

When frictions prevent firms from setting the fixed fee to its optimal level—given by $\alpha\pi_D(q_I) - (1-\alpha)\pi_U(q_I)$—double marginalization arises. This distortionary outcome emerges in two distinct cases. When $\alpha \geq \overline{\alpha}$, as described in Proposition 4, U exercises monopoly power by charging a markup over its marginal cost when selling to D, resulting in the double markup outcome. Conversely, when $\alpha \leq \underline{\alpha}$, the logic follows Proposition 3, where D exercises monopsony power by charging a markdown below its marginal revenue when purchasing from U, giving rise to the double markdown outcome. Although the underlying distortion remains of the same nature, it is less severe than under a linear wholesale contract, provided that frictions are not too extreme (i.e., $\overline{F} > 0$ and $\underline{F} < 0$). Figure 6 illustrates Proposition 7 by depicting the three types of equilibria that arise depending on the bargaining weight α.

Proposition 7 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. 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 their trading relationship.[59]

7.2 Increasing Upstream Marginal Cost without Monopsony Power

The novel type of double marginalization, identified as double markdownization in Proposition 3, stems from the assumption that U faces imperfectly elastic input suppliers, implying that its marginal cost is increasing. In this section, we argue that a closely related distortion may also arise when U's marginal cost is increasing for reasons unrelated to input supply elasticity. To see this, consider a variant of our bilateral monopoly model in which U is vertically integrated with its input suppliers. Denote this integrated entity U_I.[60] In Stage 2, U_I's optimal quantity of input to produce and sell to D is given by:

$$\tilde{q}_{U_I}(w) \in \underset{q_{U_I}}{\operatorname{argmax}}\ \pi_{U_I} \equiv w q_{U_I} - \int_0^{q_{U_I}} r(x)\, dx, \tag{13}$$

which yields the first-order condition $r(\tilde{q}_{U_I}(w)) = w$. In this setting, bilateral efficiency is achieved when the quantity $\breve{q}_I$, defined by $MR_D(\breve{q}_I) = r(\breve{q}_I)$, is exchanged, as it replicates the outcome when U_I and D are vertically integrated. Proceeding analogously to Section 4.2, one can show that this efficient outcome obtains when $\alpha = \breve{\alpha}_I = \frac{\pi_{U_I}(\breve{q}_I)}{\pi_{U_I}(\breve{q}_I)+\pi_D(\breve{q}_I)}$. Otherwise, bilateral inefficiency arises in two distinct forms. When $\alpha > \breve{\alpha}_I$, U_I charges a markup, resulting in the double markupization distortion identified in Proposition 4. When $\alpha < \breve{\alpha}_I$, D charges a markdown in addition to its markup. Unlike Proposition 3, however, only a single markdown distortion arises as the integrated entity U_I does not exercise any monopsony power. Consequently, the inefficiency stemming from D's excessive level of bargaining power persists even in the absence of upstream monopsony power.

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

[60] The integrated entity U_I can be interpreted as either a union or a vertically integrated cooperative representing heterogenous workers or suppliers, abstracting from frictions that might cause its objective function to diverge from (13). See Farber (1986) and Hansmann (2000) for corresponding textbook treatments.

8 Conclusion

This article provides a unified framework to analyze monopsony, monopoly, and countervailing power theories within a vertical supply chain. We extend the canonical bilateral monopoly model with bargaining by considering that the upstream marginal cost is increasing due to an imperfectly elastic input supply curve, which allows the upstream firm to exert monopsony power in the input market. In this context, we show that the downstream firm no longer solely determines the quantity traded in the supply chain. Instead, under voluntary exchange, the "short-side rule" governs whether the upstream or downstream firm chooses the quantity exchanged in equilibrium. This insight offers new perspectives on how the distribution of bargaining power shapes market outcomes. Crucially, we identify nonmonotonic welfare effects of both seller and buyer power. Bilateral efficiency arises and welfare is maximized whenever each firm's bargaining power fully countervails the other's market power, which occurs at a specific distribution of bargaining power determined by supply and demand elasticities. Otherwise, double marginalization emerges in one of the two forms. When the downstream firm holds excessive bargaining power, *double markdownization* arises and a novel theory of countervailing seller power prevails. Conversely, when the upstream firm holds excessive bargaining power, the classical *double markupization* emerges and Galbraith's (1952) countervailing buyer power theory applies.

Our analysis yields novel insights for competition policy and price regulation. Monopsony and countervailing buyer power theories have traditionally been viewed as distinct—or even opposing (e.g., Hemphill and Rose, 2018). Notably, the 2023 U.S. Merger Guidelines place greater emphasis on monopsony power but omit any explicit reference to the notion of countervailing power, which was present in Section 8 of the 2010 version.[61] We show that these theories are not inherently at odds but instead complementary: each becomes relevant depending on how bargaining power is distributed along the vertical supply chain. Our analysis also yields new insights into the design of price floor policies. When appropriately calibrated, a price floor on input

[61] See U.S. Department of Justice and the Federal Trade Commission (2010) and U.S. Department of Justice and the Federal Trade Commission (2023).

prices can benefit both consumers and input suppliers and lead to welfare improvements. Crucially, the distribution of bargaining power along the vertical chain plays a key role, not only determining the optimal level at which the price floor should be set but also the potential welfare gains from such regulation. These gains are greater when a downstream firm holds significant bargaining power, as a price floor can turn its monopsony power into countervailing buyer power.

Our findings also have important implications for empirical research on bargaining in vertical chains. As reviewed by Lee, Whinston and Yurukoglu (2021), it is common practice to assume constant marginal costs.[62] However, we show that the welfare consequences of the distribution of bargaining power in the vertical supply chain can vary substantially depending on the slope of the upstream marginal cost function. This observation is especially relevant given the prevalence of convex supply curves in many industries (e.g., Shea, 1993; Boehm and Pandalai-Nayar, 2022). Our results thus call for greater flexibility in modeling cost functions in empirical work. In this context, inferring whether upstream or downstream firms have the right-to-manage becomes essential for characterizing the nature of the double marginalization distortion.[63] Furthermore, as the resulting markup or markdown depends on demand and supply elasticities, the joint production-demand approach proposed by De Loecker and Scott (forthcoming) offers a promising path forward for identifying market power along supply chains.

We conclude by outlining avenues for future research. A natural extension of our framework is to incorporate upstream and downstream competition. We conjecture that our core insights carry over to simple vertical structures with imperfect competition, such as competing vertical chains or upstream competition under common agency. However, introducing imperfect downstream competition and interlocking relationships raises more complex challenges (e.g., Miklós-Thal, Rey and Vergé, 2010), and existing tractable frameworks in this context typically rely on the assumption of constant marginal costs. Another promising avenue is the analysis of shock transmission along

[62] 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 (2025).

[63] 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 to identify whether the importer or exporter determines the equilibrium quantity.

the vertical supply chain. In this regard, our results suggest that the extent and type of double marginalization—shaped by supply and demand elasticities and curvatures as well as the distribution of bargaining power—play a central role in the determination of cost pass-through rates. As such analysis would imply new assumptions on the slope of demand and supply curvatures, we leave it to future research.

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.

Aparicio, Diego, and Alberto Cavallo. 2021. "Targeted Price Controls on Supermarket Products." *Review of Economics and Statistics*, 103(1): 60–71.

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.

Binmore, Kenneth George, Ariel Rubinstein, and Asher Wolinsky. 1986. "The Nash Bargaining Solution in Economic Modelling." *RAND Journal of Economics*, 17(2): 176–188.

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. 2025. "The Buyer Power Effect of Retail Mergers: An Empirical Model of Bargaining with Equilibrium of Fear." *RAND Journal of Economics*, 56(2): 194–215.

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.

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.

Cussen, Diego, and Juan-Pablo Montero. 2024. "Using List Prices to Collude or to Compete?" *Economic Journal*, 134(664): 3232–3261.

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.

de Cornière, Alexandre, and Greg Taylor. 2021. "Upstream Bundling and Leverage of Market Power." *Economic Journal*, 131(640): 3122–3144.

De Loecker, Jan, and Paul T. Scott. forthcoming. "Markup Estimation using Production and Demand Data: An Application to the US Brewing Industry." Forthcoming in the *Review of Economic Studies.*

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.

Farber, Henry S. 1986. "The Analysis of Union Behavior." *Handbook of Labor Economics*, 2: 1039–1089.

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.

Gaudin, Germain. 2019. "Vertical Relations, Opportunism, and Welfare." *RAND Journal of Economics*, 50(2): 342–358.

Goldberg, Pinelopi, and Rebecca Hellerstein. 2013. "A structural approach to identifying the sources of local currency price stability." *Review of Economic Studies*, 80(1): 175–210.

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.

Hansmann, Henry. 2000. *The Ownership of Enterprise.* Harvard University Press.

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.

Katz, Michael L. 1987. "The Welfare Effects of Third-Degree Price Discrimination in Intermediate Good Markets." *American Economic Review*, 77(1): 154–167.

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.

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

Loertscher, Simon, and Ellen V Muir. 2021. "Wage dispersion, minimum wages and involuntary unemployment: A mechanism design perspective." Mimeo.

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. 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.

Miklós-Thal, Jeanine, Patrick Rey, and Thibaud Vergé. 2010. "Vertical Relations." *International Journal of Industrial Organization*, 28: 345–349.

Molina, Hugo. 2024. "Buyer Alliances in Vertical Related Markets." Unpublished.

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.

O'Brien, Daniel P. 2014. "The welfare effects of third-degree price discrimination in intermediate good markets: the case of bargaining." *RAND Journal of Economics*, 45(1): 92–115.

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.

Roth, Alvin E. 1979. *Axiomatic Models of Bargaining.* Vol. 170 of *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag Berlin Heidelberg.

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

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.

Salinger, Michael A. 1988. "Vertical Mergers and Market Foreclosure." *The Quarterly Journal of Economics*, 103(2): 345–356.

Shea, John. 1993. "Do Supply Curves Slope Up?" *Quarterly 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.

U.S. Department of Justice and the Federal Trade Commission. 2010. "Horizontal Merger Guidelines." Washington, DC: United States Department of Justice.

U.S. Department of Justice and the Federal Trade Commission. 2023. "Merger Guidelines." Washington, DC: United States Department of Justice.

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.

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 respective profits—$\tilde{q}_U(w)$ for U and $\tilde{q}_D(w)$ for D. To see this, suppose U anticipates that D will announce q_D^a, where the superscript a stands for "anticipated". If $q_D^a \leq \tilde{q}_U(w)$, then announcing $\tilde{q}_U(w)$ is weakly optimal for U, as only q_D^a will be traded under voluntary exchange, and announcing a higher quantity yields the same outcome. Announcing less than q_D^a is strictly dominated, as it leads to a traded quantity further from $\tilde{q}_U(w)$. Conversely, if $q_D^a > \tilde{q}_U(w)$, then announcing $\tilde{q}_U(w)$ is the best strategy for U as, in that case, the quantity traded maximizes its profit. The same logic applies symmetrically to D, and each firm's (weakly) dominant strategy is to announce its profit-maximizing quantity. As a result, there exists a unique Nash equilibrium in (weakly) dominant strategies, with $q = \min\{\tilde{q}_U(w), \tilde{q}_D(w)\})$.

A.2 Differentiability of the Nash Product

The Nash product in (6) is given by $\pi_U(w)^\alpha \pi_D(w)^{(1-\alpha)}$, where $\pi_U(w) = [w - r(q(w))]q(w)$ and $\pi_D(w) = [p(q(w)) - w]q(w)$. In what follows, we show that $\pi_U(w)$ and $\pi_D(w)$ are differentiable for all $w \neq w_I$, as they are compositions of differentiable functions.

First, $\forall w \neq w_I$, we have:

$$\pi'_U(w) = q(w) + q'(w)\left[w - MC_U(q(w))\right]$$
$$\pi'_D(w) = q'(w)\left[MR_D(q(w)) - w\right] - q(w)$$

By Assumption 1, we have $q'(w) = \frac{1}{MR'_D(q)} < 0$ when D determines the traded quantity in Stage 2, and $q'(w) = \frac{1}{MC'_U(q)} > 0$ otherwise. Hence, both $\pi_U(w)$ and $\pi_D(w)$ are differentiable $\forall w \neq w_I$, and thus the Nash product in (6) is differentiable as well $\forall w \neq w_I$.

Consider now the differentiability at w_I. The right-hand derivative of $q(w)$ at w_I is given by:

$$q'_+(w_I) \equiv \lim_{\epsilon \to 0^+} \frac{MR_D^{-1}(w_I + \epsilon) - MR_D^{-1}(w_I)}{\epsilon} = MR_D^{-1'}(w_I) = \frac{1}{MR'_D(q_I)} = \frac{1}{(2 - \sigma_p(q_I))p'(q_I)} < 0,$$

which is finite and negative because $\sigma_p(q_I) < 2$ and $p'(q_I) < 0$ by Assumption 1. Similarly, the left-hand derivative of $q(w)$ at w_I is given by:

$$q'_-(w_I) \equiv \lim_{\epsilon \to 0^-} \frac{MC_U^{-1}(w_I + \epsilon) - MC_U^{-1}(w_I)}{\epsilon} = MC_U^{-1'}(w_I) = \frac{1}{MC'_U(q_I)} > 0,$$

which is finite and positive because $MC_U(q)$ is increasing (Assumption 1.(i)). Therefore, the differentiability of U's profit at w_I follows from:

$$\pi'_{U_+}(w_I) = q(w_I) + q'_+(w_I)\,[w_I - MC_U(q_I)] = q_I + \frac{w_I - MC_U(q_I)}{MR'_D(q_I)}$$

and

$$\pi'_{U_-}(w_I) = q(w_I) + q'_-(w_I)\,[w_I - MC_U(q_I)] = q_I + \frac{w_I - MC_U(q_I)}{MC'_U(q_I)}$$

both of which equal q_I as $w_I = MC_U(q_I)$. Similarly, the differentiability of D's profit at w_I follows from:

$$\pi'_{D_+}(w_I) = q'_+(w_I)\,[w_I - MR_D(q_I)] - q(w_I) = \frac{w_I - MR_D(q_I)}{MR'_D(q_I)} - q_I$$

and

$$\pi'_{D_-}(w_I) = q'_-(w_I)\,[w_I - MR_D(q_I)] - q(w_I) = \frac{w_I - MR_D(q_I)}{MC'_U(q_I)} - q_I$$

both of which equal $-q_I$ as $w_I = MR_D(q_I)$. Therefore, it follows that $\pi_U(w)$ and $\pi_D(w)$ are differentiable at w_I, and thus the Nash product in (6) is differentiable at w_I as well.

A.3 When Bargaining Leads to Bilateral Efficiency ($w = w_I$)

A.3.1 First-Order Condition

When $w(q) = MR_D(q)$, the first-order condition (7) evaluated at q_I can be written as:

$$\alpha \underbrace{(MR_U(q_I) - MC_U(q_I))}_{\pi'_U(q_I)} \pi_D(q_I) + (1-\alpha) \underbrace{(MR_D(q_I) - MR_U(q_I))}_{\pi'_D(q_I)} \pi_U(q_I) = 0 \tag{14}$$

When $w(q) = MC_U(q)$, the first-order condition (7) evaluated at q_I can be written as:

$$\alpha \underbrace{(MC_D(q_I) - MC_U(q_I))}_{\pi'_U(q_I)} \pi_D(q_I) + (1-\alpha) \underbrace{(MR_D(q_I) - MC_D(q_I))}_{\pi'_D(q_I)} \pi_U(q_I) = 0 \tag{15}$$

Given that $MC_U(q_I) = MR_D(q_I)$, (14) and (15) simplify to: $\alpha\pi_D(q_I) + (1-\alpha)\pi_U(q_I) = 0$.

A.3.2 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 $w = MC_U(q_I) = MR_D(q_I)$. Given that $\pi_U(q) = (w(q) - r(q))\, q$ and $\pi_D(q) = (p(q) - w(q))\, q$, we can rewrite α_I as follows:

$$\alpha_I \quad = \quad \frac{(MR_D(q_I) - r(q_I))q_I}{(p(q_I) - r(q_I))q_I} = \frac{MR_D(q_I) - r(q_I)}{p(q_I) - r(q_I)}.$$

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:

$$\alpha_I \quad = \quad \frac{MR_D(q_I) - MR_D(q_I)\frac{\varepsilon_r(q_I)}{\varepsilon_r(q_I)+1}}{MR_D(q_I)\frac{\varepsilon_p(q_I)}{\varepsilon_p(q_I)-1} - MR_D(q_I)\frac{\varepsilon_r(q_I)}{\varepsilon_r(q_I)+1}} = \frac{\varepsilon_p(q_I) - 1}{\varepsilon_p(q_I) + \varepsilon_r(q_I)}.$$

Equilibrium existence. We demonstrate the existence of an equilibrium (w_I, q_I) at $\alpha = \alpha_I$.

Consider first a deviation $\tilde{w} = w_I + \epsilon$ with $\epsilon \to 0$, implying that $w(q) = MR_D(q)$ (i.e., D sets the traded quantity in Stage 2) and $\tilde{q} = q_I - \varepsilon$. In this case, consider that the first-order condition (7) at $\tilde{q}$ is satisfied for a given $\tilde{\alpha}$, that is:

$$\tilde{\alpha}(MR_U(\tilde{q}) - MC_U(\tilde{q}))\pi_D(\tilde{q}) + (1-\tilde{\alpha})(MR_D(\tilde{q}) - MR_U(\tilde{q}))\pi_U(\tilde{q}) = 0 \tag{16}$$

As $MC_U(q_I) = MR_D(q_I)$, the left-hand side of (16) evaluated at q_I boils down to:

$$\tilde{\alpha}(MR_U(q_I) - MR_D(q_I))\,(\tilde{\alpha}\pi_D(q_I) - (1-\tilde{\alpha})\pi_U(q_I))$$

As $MR_U(q_I) = MR_D(q_I) + MR'_D(q_I)q_I$ and $MR'_D(q_I) < 0$ (Assumption 2.(ii)), we have $MR_U(q_I) - MR_D(q_I) < 0$ implying that (16) is satisfied at q_I only if $\tilde{\alpha}\pi_D(q_I) - (1-\tilde{\alpha})\pi_U(q_I) = 0 \Leftrightarrow \tilde{\alpha} = \frac{\pi_U(q_I)}{\pi_D(q_I)+\pi_U(q_I)} = \alpha_I$. As a result, there is no $\tilde{w} > w_I$ (implying $\tilde{q} < q_I$) at α_I that satisfies the first-order condition for the maximization of the Nash product.

We now consider a deviation $\tilde{w} = w_I - \epsilon$ with $\epsilon \to 0$, which implies that $w = MC_U(q)$ (i.e., U sets the traded quantity in Stage 2) and $\tilde{q} = q_I - \epsilon$. Again, consider in this case that the first-order condition (7) at $\tilde{q}$ is satisfied for a given $\tilde{\alpha}$, that is:

$$\tilde{\alpha}(MC_D(\tilde{q}) - MC_U(\tilde{q}))\pi_D(\tilde{q}) + (1-\tilde{\alpha})(MR_D(\tilde{q}) - MC_D(\tilde{q}))\pi_U(\tilde{q}) = 0 \tag{17}$$

As $MR_D(q_I) = MC_U(q_I)$, the left-hand side of (17) evaluated at q_I boils down to:

$$\tilde{\alpha}(MC_D(q_I) - MC_U(q_I))\,(\tilde{\alpha}\pi_D(q_I) - (1-\tilde{\alpha})\pi_U(q_I))$$

As $MC_D(q_I) = MC_U(q_I) + MC'_U(q_I)q_I$ and $MC'_U(q_I) > 0$ (Assumption 1.(i)), we have $MC_D(q_I) -$

$MC_U(q_I) > 0$ implying that (17) is satisfied at q_I only if $\tilde{\alpha}\pi_D(q_I) - (1-\tilde{\alpha})\pi_U(q_I) = 0 \Leftrightarrow \tilde{\alpha} = \frac{\pi_U(q_I)}{\pi_D(q_I)+\pi_U(q_I)} = \alpha_I$. As a result, there is no $\tilde{w} < w_I$ (implying $\tilde{q} < q_I$) at α_I that satisfies the first-order condition for the maximization of the Nash product.

A.4 Bargaining when D is Powerful ($w = MC_U(q) < w_I$)

A.4.1 Second-Order Condition

In the case where $w = MC_U(q)$, the first-order condition in (10) can be rearranged as follows:

$$MR_D(q_\nu) - \widetilde{MC}_D(q_\nu) = 0 \tag{18}$$

We show that Assumptions 1 and 2.(i) are sufficient to ensure that the second-order condition for the maximization of the Nash product holds (i.e., $MR_D(q) - \widetilde{MC}_D(q)$ is strictly decreasing in q).

Recall that $\widetilde{MC}_D(q) = \beta_U(q,\alpha)MC_U(q) + (1-\beta_U(q,\alpha))MC_D(q)$, where $\beta_U(q,\alpha)) = \frac{\alpha\pi_D(q)}{(1-\alpha)\pi_U(q)}$. Differentiating $\widetilde{MC}_D(q)$ with respect to q, we obtain:

$$\frac{\partial \widetilde{MC}_D(q)}{\partial q} = \frac{\partial \beta_U(q,\alpha)}{\partial q}(MC_U(q) - MC_D(q)) + \beta_U(q,\alpha)(MC_U'(q) + MC_D'(q)). \tag{19}$$

As shown in Section 4.2.2, under $w(q) = MC_U(q)$, we have $\pi_U'(q) > 0$ and $\pi_D'(q) < 0$, implying that $\frac{\partial \beta_U(q,\alpha)}{\partial q} = \frac{\alpha\pi_D'(q)\pi_U(q) - (1-\alpha)\pi_D(q)\pi_U'(q)}{(\alpha\pi_U(q))^2} < 0$. As $MC_D(q) = MC_U'(q)q + MC_U(q)$ and $MC_U(q)$ is increasing (Assumption 1.(i)), we have $MC_U(q) - MC_D(q) < 0$, implying that the first term in (19) is positive. Turning to the second term, we have $MC_U'(q) = r''(q)q + 2r'(q) = r'(q)(\sigma_r + 2) > 0$ whenever $\sigma_r > -2$ (Assumption 1). Similarly, we have $MC_D'(q) = MC_U''(q)q + 2MC_U'(q) = MC_U'(q)(\sigma_{MC_U} + 2) > 0$ whenever $\sigma_{MC_U} > -2$ (Assumption 2.(i)). Under these conditions, we thus obtain $\frac{\partial \widetilde{MC}_D(q)}{\partial q} > 0$. Moreover, we have $MR_D'(q) = p''(q)q + 2p'(q) = -p'(q)(\sigma_p - 2) < 0$ whenever $\sigma_p < 2$ (Assumption 1). Together, these conditions ensure that the second-order condition for the maximization of the Nash product is always satisfied.

A.4.2 D's Markdown ν_D

To show that D's markdown is given by $\nu_D = \frac{MR_D(q_\nu)}{w(q_\nu)} = \frac{MR_D(q_\nu)}{MC_U(q_\nu)}$, we divide each term of the first-order condition in (9) by $MC_U(q_\nu)$:

$$\alpha\left(\frac{MC_D(q_\nu)}{MC_U(q_\nu)} - 1\right)\left(\frac{p(q_\nu)}{MC_U(q_\nu)} - 1\right) + (1-\alpha)\left(\frac{MR_D(q_\nu)}{MC_U(q_\nu)} - \frac{MC_D(q_\nu)}{MC_U(q_\nu)}\right)\left(1 - \frac{r(q_\nu)}{MC_U(q_\nu)}\right) = 0 \tag{20}$$

We then use the following simplifications (omitting the argument q_ν in ε_p, ε_r, and ε_{MC_U} for notational simplicity):

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

and rearrange (20) as follows:

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

A.4.3 Set of Equilibria

Suppose that U and D anticipate that U sets the traded quantity in Stage 2. This implies that, in Stage 1, the negotiated wholesale price satisfies $w = MC_U(q) < w_I$. Based on the first-order condition in (9), we define the function:

$$\Psi(q, \alpha) \equiv \alpha(MC_D(q) - MC_U(q))\pi_D(q) + (1-\alpha)(MR_D(q) - MC_D(q))\pi_U(q)$$

The equilibrium quantity q_ν is determined by $\Psi(q_\nu, \alpha) = 0$. By Assumption 2, and applying the implicit function theorem, we obtain:

$$Sign(\frac{\partial q_\nu}{\partial \alpha}) = Sign(\frac{\partial \Psi(q_v, \alpha)}{\partial \alpha}),$$

and $\frac{\partial \Psi(q_\nu,\alpha)}{\partial \alpha} = (MC_D(q_\nu) - MC_U(q_\nu))\pi_D(q_\nu) - (MR_D(q_\nu) - MC_D(q_\nu))\pi_U(q_\nu)$. Using the equilibrium condition $\Psi(q_\nu, \alpha) = 0$, we have $(MR_D(q_\nu) - MC_D(q_\nu))\pi_U(q_\nu) = -\frac{\alpha}{1-\alpha}(MC_D(q_\nu) - MC_U(q_\nu))\pi_D(q_\nu)$, which implies $\frac{\partial \Psi(q_\nu,\alpha)}{\partial \alpha} = \frac{1}{1-\alpha}(MC_D(q_\nu) - MC_U(q_\nu))\pi_D(q_\nu) > 0$ because $MC_D(q_\nu) = MC_U(q_\nu) + MC'_U(q_\nu)q_\nu$ and $MC_U(q_\nu)$ is increasing (Assumption 1.(i)). It follows that the equilibrium quantity q_ν increases with α. Moreover, because $q_\nu = q_I$ when $\alpha = \alpha_I$, we have $q_\nu < q_I$ for $\alpha < \alpha_I$ and $q_\nu > q_I$ for $\alpha > \alpha_I$. However, when $\alpha > \alpha_I$, voluntary exchange implies that D sets the traded quantity in Stage 2 according to $w = MR_D(q)$ (see Section 4.1). Therefore, the set of equilibria characterized by (9) and $w_\nu = MC_U(q_\nu)$ exists only for $\alpha \in [0, \alpha_I]$.

A.4.4 U's degree of Countervailing Seller Power β_U

The first-order condition in (10) defines $\beta_U(q_\nu)$ as:

$$\beta_U(q_\nu, \alpha) \equiv \frac{\alpha}{1-\alpha}\frac{\pi_D(q_\nu)}{\pi_U(q_\nu)} = -\frac{\pi'_D(q_\nu)}{\pi'_U(q_\nu)} = \frac{MC_D(q_\nu) - MR_D(q_\nu)}{MC_D(q_\nu) - MC_U(q_\nu)} = \frac{-g(q_\nu)}{e(q_\nu)}$$

where $e(q) \equiv MC_D(q) - MC_U(q)$ and $g(q) \equiv MR_D(q) - MC_D(q)$. We have $e(q) > 0$ (Assumption 1.(i)), and $g(q) \leq 0$ due to the concavity of $\pi_D(q)$ under Assumptions 1.(ii) and 2.(ii). As $MC_U(q_\nu) \leq MR_D(q_\nu)$ because $q_\nu \leq q_I$, it follows that $0 \leq \beta_U(q_\nu, \alpha) \leq 1$. By the chain rule, we have $\frac{d\beta_U}{d\alpha}\Big|_{q=q_\nu} = \frac{\partial\beta_U}{\partial q}\Big|_{q=q_\nu}\frac{dq}{d\alpha}\Big|_{q=q_\nu}$. As established in Appendix A.4.3, we have $\frac{dq}{d\alpha}\Big|_{q=q_\nu} > 0$. Moreover, we have:

$$\begin{aligned}
\frac{\partial\beta_U}{\partial q}\Big|_{q=q_\nu} &= \frac{e(q_\nu)(MC'_D(q_\nu) - MR'_D(q_\nu)) + g(q_\nu)(MC'_D(q_\nu) - MC'_U(q_\nu))}{e(q_\nu)^2}\\
\Leftrightarrow \quad \frac{\partial\beta_U}{\partial q}\Big|_{q=q_\nu} &= \frac{MC'_D(q_\nu)(MR_D(q_\nu) - MC_U(q_\nu)) - MR'_D(q_\nu)e(q_\nu) - g(q_\nu)MC'_U(q_\nu)}{e(q_\nu)^2}.
\end{aligned}$$

Given that $e(q_\nu) > 0$, $g(q_\nu) \leq 0$, and $MR_D(q_\nu) > MC_U(q_\nu)$, we obtain $\frac{\partial\beta_U}{\partial q}\Big|_{q=q_\nu} > 0$. As a result:

$$\frac{d\beta_U}{d\alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial\beta_U}{\partial q}\Big|_{q=q_\nu}}_{>0}\underbrace{\frac{dq}{d\alpha}\Big|_{q=q_\nu}}_{>0} > 0$$

A.5 Bargaining when U is Powerful ($w = MR_D(q) > w_I$)

A.5.1 Second-Order Condition

In the case where $w = MR_D(q)$, the first-order condition in (12) can be rearranged as follows:

$$\widetilde{MR}_U(q_\mu) - MC_U(q_\mu) = 0 \tag{21}$$

We show below that Assumptions 1 and 2.(ii) are sufficient to ensure that the second-order condition for the maximization of the Nash product holds (i.e., that $\widetilde{MR}_U(q) - MC_U(q)$ is strictly decreasing in q).

Recall that $\widetilde{MR}_U(q) = \beta_D(q,\alpha)MR_D(q) + (1-\beta_D(q,\alpha))MR_U(q)$, where $\beta_D(q,\alpha)) = \frac{(1-\alpha)\pi_U(q)}{\alpha\pi_D(q)}$. Differentiating $\widetilde{MR}_U(q)$ with respect to q, we obtain:

$$\frac{\partial\widetilde{MR}_U(q)}{\partial q} = \frac{\partial\beta_D(q,\alpha)}{\partial q}(MR_D(q) - MR_U(q)) + \beta_D(q,\alpha)(MR'_D(q) + MR'_U(q)). \tag{22}$$

As shown in Section 4.2.3, under $w(q) = MR_D(q)$, we have $\pi'_U(q) < 0$ and $\pi'_D(q) > 0$, implying that $\frac{\partial\beta_D(q,\alpha)}{\partial q} = \frac{(1-\alpha)\pi'_U(q)\pi_D(q) - \alpha\pi_U(q)\pi'_D(q)}{(\alpha\pi_D(q))^2} < 0$. Given that $MR_U = MR'_D(q)q + MR_D(q)$ and $MR'_D(q)$

is decreasing (Assumption 1.(ii)), we have $MR_D(q) - MR_U(q) > 0$, implying that the first term in (22) is negative. Turning to the second term, we have $MR_D'(q) = p''(q)q + 2p'(q) = -p'(q)(\sigma_p - 2) < 0$ whenever $\sigma_p < 2$ (Assumption 1). Similarly, we have $MR_U'(q) = MR_D''(q)q + 2MR_D'(q) = -MR_D'(q)(\sigma_{MR_D} - 2) < 0$ whenever $\sigma_{MR_D} < 2$ (Assumptions 2.(ii)). Therefore $\widetilde{MR}_U(q)$ is decreasing. Finally we have $MC_U'(q) = r''(q)q + 2r'(q)) = r'(q)(\sigma_r + 2) > 0$ whenever $\sigma_r > -2$ (Assumption 1). Together, these conditions ensure that the second-order condition for the maximization of the Nash product is always satisfied.

A.5.2 U's Markup μ_U

To show that U's markup is given by $\mu_U = \frac{w(q_\mu)}{MC_U(q_\mu)} = \frac{MR_D(q_\mu)}{MC_U(q_\mu)}$, we divide each term of the first-order condition in (21) by $MC_U(q_\mu)$:

$$\alpha\left(\frac{MR_U(q_\mu)}{MC_U(q_\mu)} - 1\right)\left(\frac{p(q_\mu)}{MC_U(q_\mu)} - \frac{MR_D(q_\mu)}{MC_U(q_\mu)}\right) + (1-\alpha)\left(\frac{MR_D(q_\mu)}{MC_U(q_\mu)} - \frac{MR_U(q_\mu)}{MC_U(q_\mu)}\right)\left(\frac{MR_D(q_\mu)}{MC_U(q_\mu)} - \frac{r(q_\mu)}{MC_U(q_\mu)}\right) = 0. \quad (23)$$

We then use the following simplifications (omitting the argument q_ν in ε_p, ε_r, and ε_{MR_D} for notational simplicity):

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

and rearrange (23) as follows:

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

A.5.3 Set of Equilibria

Suppose that U and D anticipate that D sets the traded quantity in Stage 2. This implies that, in Stage 1, the negotiated wholesale price satisfies $w = MR_D(q) < w_I$. Based on the first-order condition in (11), we define the following function:

$$\Phi(q, \alpha) \equiv \alpha(MR_U(q) - MC_U(q))\pi_D(q) + (1-\alpha)(MR_D(q) - MR_U(q))\pi_U(q)$$

The equilibrium quantity q_μ is determined by $\Phi(q_\mu, \alpha) = 0$. By Assumption 2, and applying the implicit function theorem, we obtain:

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

and $\frac{\partial \Phi(q_\mu,\alpha)}{\partial \alpha} = (MR_U(q_\mu) - MC_U(q_\mu))\pi_D(q_\mu) - (MR_D(q_\mu) - MR_U(q_\mu))\pi_U(q_\mu)$. Using the equilibrium condition $\Phi(q_\mu, \alpha) = 0$, we have $(MR_U(q_\mu) - MC_U(q_\mu))\pi_D(q_\mu) = -\frac{1-\alpha}{\alpha}(MR_D(q_\mu) - MR_U(q_\mu))\pi_U(q_\mu)$, which implies $\frac{\partial \Psi(q_\nu,\alpha)}{\partial \alpha} = \frac{1}{\alpha}(MR_U(q_\mu) - MR_D(q_\mu))\pi_U(q_\mu) < 0$ because $MR_U(q_\mu) = MR_D(q_\nu) + MR'_D(q_\mu)q_\mu$ and $MR_D(q_\mu)$ is decreasing (Assumption 1.(ii)). It follows that the equilibrium quantity q_μ decreases with α. Moreover, because $q_\mu = q_I$ when $\alpha = \alpha_I$, we have $q_\mu > q_I$ for $\alpha < \alpha_I$ and $q_\mu < q_I$ for $\alpha > \alpha_I$. However, when $\alpha < \alpha_I$, voluntary exchange implies that U sets the traded quantity in Stage 2 according to $w = MC_U(q)$ (see Section 4.1). Therefore, the set of equilibria characterized by (11) and $w_\mu = MR_D(q_\mu)$ exists only for $\alpha \in [\alpha_I, 1]$.

A.5.4 D's degree of Countervailing Buyer Power β_D

The first-order condition in (12) defines $\beta_D(q_\mu, \alpha)$ as:

$$\beta_D(q_\mu, \alpha) = \frac{MC_U(q_\mu) - MR_U(q_\mu)}{MR_D(q_\mu) - MR_U(q_\mu)} = \frac{-a(q_\mu)}{c(q_\mu)}$$

where $a(q) \equiv MR_U(q) - MC_U(q)$ and $c(q) \equiv MR_D(q) - MR_U(q) > 0$. Note that $a(q) \leq 0$ due to the concavity of $\pi_U(q)$ under Assumptions 1.(i) and 2.(i), and $c(q) > 0$ (Assumption 1.(ii)). As $MC_U(q_\mu) \leq MR_D(q_\mu)$ because $q_\mu \leq q_I$, we have $0 \leq \beta_D(q_\mu, \alpha) \leq 1$. By the chain rule, we have $\frac{d\beta_D}{d\alpha}\Big|_{q=q_\mu} = \frac{\partial \beta_D}{\partial q}\Big|_{q=q_\mu} \frac{dq}{d\alpha}\Big|_{q=q_\mu}$. As established in Appendix A.5.3, we have $\frac{dq}{d\alpha}\Big|_{q=q_\mu} < 0$. Moreover, we have:

$$\begin{aligned}
\frac{\partial \beta_D}{\partial q}\Big|_{q=q_\mu} &= \frac{(MC'_U(q_\mu) - MR'_U(q_\mu))c + a(MR'_D(q_\mu) - MR'_U(q_\mu))}{c(q_\mu)^2} \\
\Leftrightarrow \quad \frac{\partial \beta_D}{\partial q}\Big|_{q=q_\mu} &= \frac{MC'_U(q_\mu)c(q_\mu) + MR'_D(q_\mu)a(q_\mu) - MR'_U(q_\mu)(MR_D(q_\mu) - MC_U(q_\mu))}{c(q_\mu)^2}.
\end{aligned}$$

Given that $a(q_\mu) \leq 0$, $c(q_\mu) > 0$, and $MC_U(q_\mu) < MR_D(q_\mu)$, we have $\frac{\partial \beta_D}{\partial q}\Big|_{q=q_\mu} > 0$. As a result:

$$\frac{d\beta_D}{d\alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial \beta_D}{\partial q}\Big|_{q=q_\mu}}_{>0} \underbrace{\frac{dq}{d\alpha}\Big|_{q=q_\mu}}_{<0} < 0.$$

A.6 Proof of Corollary 2

When D is powerful ($w = MC_U(q) < w_I$). As U sets the traded quantity in Stage 2, we have $w = MC_U(q)$. This implies that $\pi_U(q) = (w - r(q))\, q = (MC_U(q) - r(q))\, q$ and $\pi_D(q) = (p(q) - w)\, q = (p(q) - MC_U(q))\, q$. First, we have $\frac{d\pi_U}{d\alpha}\Big|_{q=q_\nu} = \frac{\partial \pi_U}{\partial q}\Big|_{q=q_\nu} \frac{dq}{d\alpha}\Big|_{q=q_\nu}$. As $\frac{dq}{d\alpha}\Big|_{q=q_\nu} > 0$ (Appendix A.4.3) and $\frac{\partial \pi_U}{\partial q}\Big|_{q=q_\nu} = MC_D(q_\nu) - MC_U(q_\nu) > 0$, because $MC_D(q_\nu) = MC_U(q_\nu) + MC_U'(q_\nu)q_\nu$ and $MC_U(q_\nu)$ is increasing (Assumption 1.(i)), it follows that $\frac{d\pi_U}{d\alpha}\Big|_{q=q_\nu} > 0$. Second, we have $\frac{d\pi_D}{d\alpha}\Big|_{q=q_\nu} = \frac{\partial \pi_D}{\partial q}\Big|_{q=q_\nu} \frac{dq}{d\alpha}\Big|_{q=q_\nu}$. As $\frac{dq}{d\alpha}\Big|_{q=q_\nu} > 0$ and $\frac{\partial \pi_D}{\partial q}\Big|_{q=q_\nu} = MR_D(q_\nu) - MC_D(q_\nu) \leq 0$ due to the concavity of $\pi_D(q)$ under Assumptions 1.(ii) and 2.(ii), it follows that $\frac{d\pi_D}{d\alpha}\Big|_{q=q_\nu} \leq 0$. Finally, as $CS(q) \equiv \int_0^q p(x)\, dx - p(q)q$ and $SS(q) \equiv r(q)q - \int_0^q r(x)\, dx$ are strictly increasing in q, and $\frac{dq}{d\alpha}\Big|_{q=q_\nu} > 0$, consumers and input suppliers benefit when α increases toward α_I.

When U is powerful ($w = MR_D(q) > w_I$). As D sets the traded quantity in Stage 2, we have $w = MR_D(q)$. This implies that $\pi_U(q) = (w - r(q))\, q = (MR_D(q) - r(q))\, q$ and $\pi_D(q) = (p(q) - w)\, q = (p(q) - MR_D(q))\, q$. First, we have $\frac{d\pi_D}{d\alpha}\Big|_{q=q_\mu} = \frac{\partial \pi_D}{\partial q}\Big|_{q=q_\mu} \frac{dq}{d\alpha}\Big|_{q=q_\mu}$. As $\frac{dq}{d\alpha}\Big|_{q=q_\mu} < 0$ (Appendix A.5.3) and $\frac{\partial \pi_D}{\partial q}\Big|_{q=q_\mu} = MR_D(q_\mu) - MR_U(q_\mu) > 0$, because $MR_U(q_\mu) = MR_D(q_\mu) + MR_D'(q_\mu)q_\mu$ and $MR_D(q_\mu)$ is decreasing (Assumption 1.(ii)), it follows that $\frac{d\pi_D}{d\alpha}\Big|_{q=q_\mu} < 0$. Second, we have $\frac{d\pi_U}{d\alpha}\Big|_{q=q_\mu} = \frac{\partial \pi_U}{\partial q}\Big|_{q=q_\mu} \frac{dq}{d\alpha}\Big|_{q=q_\mu}$. As $\frac{dq}{d\alpha}\Big|_{q=q_\mu} < 0$ (Appendix A.5.3) and $\frac{\partial \pi_U}{\partial q}\Big|_{q=q_\mu} = MR_U(q_\mu) - MC_U(q_\mu) \leq 0$ due to the concavity of $\pi_U(q)$ under Assumptions 1.(i) and 2.(i), it follows that $\frac{d\pi_U}{d\alpha}\Big|_{q=q_\mu} \geq 0$. Finally, as $CS(q) \equiv \int_0^q p(x)\, dx - p(q)q$ and $SS(q) \equiv r(q)q - \int_0^q r(x)\, dx$ are strictly increasing in q, and $\frac{dq}{d\alpha}\Big|_{q=q_\mu} < 0$, consumers and input suppliers benefit when α decreases toward α_I.

A.7 Proof of Remark 1

Unless otherwise stated, all derivative signs in the proof follow from Assumption 1, Corollary 1, and Appendix A.4.3 and A.5.3.

When D is powerful ($w = MC_U(q) < w_I$). In what follows, we analyze how the supply chain margin, D's margin, and U's margin vary with α.

- **Supply chain margin.** By definition, we have $\mathcal{M} \equiv \frac{p(q_\nu)}{r(q_\nu)}$. As $\frac{\partial p}{\partial \alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial p}{\partial q}\Big|_{q=q_\nu}}_{<0} \underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0} < 0$ and $\frac{\partial r}{\partial \alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial r}{\partial q}\Big|_{q=q_\nu}}_{>0} \underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0} > 0$, we obtain $\frac{d\mathcal{M}}{d\alpha}\Big|_{q=q_\nu} < 0$.
- **D's margin.** By definition, we have $M_D \equiv \mu_D \times \nu_D = \frac{p(q_\nu)}{w(q_\nu)} = \frac{p(q_\nu)}{MC_U(q_\nu)}$. As $\frac{\partial p}{\partial \alpha}\Big|_{q=q_\nu} =$

$\underbrace{\frac{\partial p}{\partial q}\Big|_{q=q_\nu}}_{<0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0} < 0$ and $\frac{\partial MC_U}{\partial \alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial MC_U}{\partial q}\Big|_{q=q_\nu}}_{<0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0} > 0$, it follows that $\frac{dM_D}{d\alpha}\Big|_{q=q_\nu} < 0$.

By definition, we have $\nu_D \equiv \frac{MR_D(q_\nu)}{w(q_\nu)} = \frac{MR_D(q_\nu)}{MC_U(q_\nu)}$. As $\frac{\partial MC_U}{\partial \alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial MC_U}{\partial q}\Big|_{q=q_\nu}}_{>0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0} > 0$ and $\frac{\partial MR_D}{\partial \alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial MR_D}{\partial q}\Big|_{q=q_\nu}}_{<0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0} < 0$, we obtain $\frac{d\nu_D}{d\alpha}\Big|_{q=q_\nu} < 0$.

By definition, we have $\mu_D = \frac{p(q_\nu)}{MR_D(q_\nu)} = \frac{\varepsilon_p}{\varepsilon_p - 1}$ and $\frac{d\mu_D}{d\alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial \mu_D}{\partial \varepsilon_p}\Big|_{q=q_\nu}}_{<0}\frac{\partial \varepsilon_p}{\partial q}\Big|_{q=q_\nu}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0}$. Hence, the sign of $\frac{d\mu_D}{d\alpha}\Big|_{q=q_\nu}$ depends on $\frac{\partial \varepsilon_p}{\partial q}\Big|_{q=q_\nu}$, which is negative under subconvex demand, zero under CES demand, and positive under superconvex demand.

- ***U*'s margin.** By definition, we have $M_U = \underbrace{\mu_U}_{=1} \times \nu_U = \frac{MC_U(q_\nu)}{r(q_\nu)} = \frac{\varepsilon_r + 1}{\varepsilon_r}$ and $\frac{dM_U}{d\alpha}\Big|_{q=q_\nu} = \frac{d\nu_U}{d\alpha}\Big|_{q=q_\nu} = \underbrace{\frac{\partial \nu_U}{\partial \varepsilon_r}\Big|_{q=q_\nu}}_{<0}\frac{\partial \varepsilon_r}{\partial q}\Big|_{q=q_\nu}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\nu}}_{>0}$. Hence, the sign of $\frac{dM_U}{d\alpha}\Big|_{q=q_\nu}$ depends on $\frac{\partial \varepsilon_r}{\partial q}\Big|_{q=q_\nu}$, which is negative under subconvex supply, zero under CES supply, and positive under superconvex supply.

As a result, while the supply chain margin and D's margin always decrease with α, U's margin (or markdown) and D's markup may decrease or increase, depending on the curvature of the demand and supply functions.

When U is powerful ($w = MR_D(q) > w_I$). Again, in what follows, we analyze how the supply chain margin, U's margin, and D's margin vary with α.

- **Supply chain margin.** By definition, we have $\mathcal{M} \equiv \frac{p(q_\mu)}{r(q_\mu)}$. As $\frac{\partial p}{\partial \alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial p}{\partial q}\Big|_{q=q_\mu}}_{<0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0} > 0$ and $\frac{\partial r}{\partial \alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial r}{\partial q}\Big|_{q=q_\mu}}_{>0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0} < 0$, we obtain $\frac{d\mathcal{M}}{d\alpha}\Big|_{q=q_\mu} > 0$.

- ***U*'s margin.** By definition, we have $M_U \equiv \mu_U \times \nu_U = \frac{w_\mu}{r(q_\mu)} = \frac{MR_D(q_\mu)}{r(q_\mu)}$. As $\frac{\partial MR_D}{\partial \alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial MR_D}{\partial q}\Big|_{q=q_\mu}}_{<0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0} > 0$ and $\frac{\partial r}{\partial \alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial r}{\partial q}\Big|_{q=q_\mu}}_{>0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0} < 0$, it follows that $\frac{dM_U}{d\alpha} > 0$.

  By definition, we have $\mu_U \equiv \frac{w_\mu}{MC_U(q_\mu)} = \frac{MR_D(q_\mu)}{MC_U(q_\mu)}$. As $\frac{\partial MR_D}{\partial \alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial MR_D}{\partial q}\Big|_{q=q_\mu}}_{<0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0} > 0$ and $\frac{\partial MC_U}{\partial \alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial MC_U}{\partial q}\Big|_{q=q_\mu}}_{>0}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0} < 0$, we obtain $\frac{d\mu_U}{d\alpha}\Big|_{q=q_\mu} > 0$.

  By definition, we have $\nu_U = \frac{MC_U(q_\nu)}{r(q_\nu)} = \frac{\varepsilon_r + 1}{\varepsilon_r}$ and $\frac{d\nu_U}{d\alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial \nu_U}{\partial \varepsilon_r}\Big|_{q=q_\mu}}_{<0}\frac{\partial \varepsilon_r}{\partial q}\Big|_{q=q_\mu}\underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0}$.

Hence, the sign of $\frac{d\nu_U}{d\alpha}\Big|_{q=q_\mu}$ depends on $\frac{\partial \varepsilon_r}{\partial q}\Big|_{q=q_\mu}$, which is negative under subconvex supply, zero under CES supply, and positive under superconvex supply.

- ***D*'s margin.** By definition, we have $M_D = \mu_D \times \underbrace{\nu_D}_{=1} = \frac{p(q_\mu)}{MR_D(q_\mu)} = \frac{\varepsilon_p}{\varepsilon_p - 1}$ and $\frac{dM_D}{d\alpha}\Big|_{q=q_\mu} = \frac{d\mu_D}{d\alpha}\Big|_{q=q_\mu} = \underbrace{\frac{\partial \mu_D}{\partial \varepsilon_p}\Big|_{q=q_\mu}}_{<0} \frac{\partial \varepsilon_p}{\partial q}\Big|_{q=q_\mu} \underbrace{\frac{\partial q}{\partial \alpha}\Big|_{q=q_\mu}}_{<0}$. Hence, the sign of $\frac{d\mu_D}{d\alpha}\Big|_{q=q_\mu}$ depends on $\frac{\partial \varepsilon_p}{\partial q}\Big|_{q=q_\mu}$, which is negative under subconvex demand, zero under CES demand, and positive under superconvex demand.

As a result, whereas the supply chain margin and U's margin always increase with α, D's margin (or markup) and U's markdown may decrease or increase, depending on the curvature of the demand and supply functions.

A.8 Markup and Markdown Definitions

A.8.1 Vertical Integration and Take-it-or-leave-it Offers

Consider a vertically integrated firm, I, as studied in Section 3.2. I faces an increasing inverse supply curve $r(q)$ and a decreasing inverse demand curve $p(q)$. Its profit maximization problem can be written as $\max_q \pi_I(q) = p(q)q - r(q)q$, which yields the following first-order condition: $MR_I(q_I) = MC_I(q_I)$, where q_I denotes the equilibrium quantity.

Markup. Consider a hypothetical scenario in which, in addition to selling q_I units at price p, I decides whether to sell an infinitesimal quantity ϵ to consumers at a distinct price $\overline{p}$. Importantly, under this scenario, I incurs no revenue loss on inframarginal units when selling the additional quantity ϵ, as the price charged for the q_I units remains unchanged. Based on this hypothetical scenario, we determine the minimum price $\hat{p}$ at which I is just willing to sell this additional quantity ϵ, resulting in a total quantity sold of $q_I + \varepsilon$ (with $\epsilon \to 0$). Formally, $\hat{p}$ is the smallest value of $\overline{p}$ such that:

$$\overline{\pi}_I(q_I, \epsilon, \overline{p}) \geq \pi_I(q_I) \tag{24}$$

where $\pi_I(q_I) = p(q_I)q_I - r(q_I)q_I$ and $\overline{\pi}_I(q_I, \epsilon, \overline{p})$ corresponds to I's profit from selling $q_I + \varepsilon$ in the hypothetical scenario, which is given by:

$$\overline{\pi}_I(q_I, \epsilon, \overline{p}) \equiv p(q_I)q_I + \overline{p}\epsilon - r(q_I + \epsilon)q_I - r(q_I + \epsilon)\epsilon. \tag{25}$$

Using (25), the inequality in (24) becomes:

$$\overline{p} \geq \frac{r(q_I + \epsilon) - r(q_I)}{\epsilon} q_I + r(q_I + \epsilon). \tag{26}$$

Taking the limit of the right-hand side of (26) as $\epsilon \to 0$, we obtain:

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

As a result, the minimum price at which I would be willing to sell the marginal unit ϵ is $\hat{p} = MC_I(q_I)$. According to our definition, it follows that I's markup is given by: $\mu_I(q_I) \equiv \frac{p}{\hat{p}} = \frac{p(q_I)}{MC_I(q_I)}$.

Markdown. Consider a hypothetical scenario in which, in addition to purchasing q_I units at price r from its input suppliers, I decides whether to purchase an infinitesimal quantity ϵ at a distinct input price $\overline{r}$. Importantly, under this scenario, I incurs no cost increase on inframarginal units when buying the additional quantity ϵ, as the purchasing price for the q_I units remains unchanged. Based on this hypothetical scenario, we determine the maximum price $\hat{r}$ at which I is just willing to buy this additional quantity ϵ, resulting in a total quantity purchased of $q_I + \varepsilon$. Formally, $\hat{r}$ is the highest value of $\overline{r}$ such that:

$$\overline{\pi}_I(q_I, \epsilon, \overline{r}) \geq \pi_I(q_I) \tag{27}$$

where $\pi_I(q_I) = p(q_I)q_I - r(q_I)q_I$ and $\overline{\pi}_I(q_I, \epsilon, \overline{r})$ corresponds to I's profit from purchasing $q_I + \varepsilon$ in the hypothetical scenario, which is given by:

$$\overline{\pi}_I(q_I, \epsilon, \overline{r}) \equiv p(q_I + \epsilon)q_I + p(q_I + \epsilon)\epsilon - r(q_I)q_I - \overline{r}\epsilon. \tag{28}$$

Using (28), the inequality in (27) becomes:

$$\frac{p(q_I + \epsilon) - p(q_I)}{\epsilon} q_I + p(q_I + \epsilon) \geq \overline{r}. \tag{29}$$

Taking the limit of the left-hand side of (29) as $\epsilon \to 0$, we obtain:

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

As a result, the maximum price at which I would be willing to purchase the marginal unit ϵ is $\hat{r} = MR_I(q_I)$. According to our definition, it follows that I's markdown is given by: $\nu_I(q_I) \equiv \frac{\hat{r}}{r} = \frac{MR_I(q_I)}{r(q_I)}$.

Although we focus on the vertically integrated case, the logic extends to any firm that unilaterally sets its price or quantity, such as in a vertical supply chain with take-it-or-leave-it offers.

A.8.2 Vertical Supply Chain with Bargaining

Consider a vertical supply chain as studied in Section 4, where U and D negotiate over a wholesale price according to the Nash bargaining solution. The Nash product is given by $N(q) = \pi_U(q)^{\alpha}\pi_D(q)^{(1-\alpha)}$, where $\pi_U(q) = (w(q) - r(q))q$ and $\pi_D(q) = (p(q) - w(q))q$. We denote by $q^\star = \{q_\mu, q_\nu\}$ the quantity that maximizes the Nash product, where $q^\star = q_\mu$ when D sets the quantity traded in Stage 2 (i.e., $w = MR_D(q)$), and $q^\star = q_\nu$ when U sets the quantity traded in Stage 2 (i.e., $w = MC_U(q)$).

***D*'s markup $\boldsymbol{\mu_D}$.** Consider a hypothetical scenario in which, in addition to selling $q^\star$ units at price p, D decides whether to order an infinitesimal quantity ϵ to U and sell it at price $\overline{p}$. Hence, D incurs no revenue loss on inframarginal units when selling the additional quantity ϵ, as the selling price for the $q^\star$ units remains unchanged. Based on this hypothetical scenario, we determine the minimum price $\hat{p}$ such that D sells this additional quantity ϵ, resulting in a total quantity sold of $q^\star + \epsilon$. Formally, $\hat{p}$ is the smallest value of $\overline{p}$ such that:

$$\overline{N}(q^\star, \epsilon, \overline{p}) \geq N(q^\star), \tag{30}$$

where $N(q^\star) = \pi_U(q^\star)^\alpha \pi_D(q^\star)^{(1-\alpha)}$ and $\overline{N}(q^\star, \epsilon, \overline{p})$ corresponds to the value of the Nash product for a quantity $q^\star + \epsilon$ in the hypothetical scenario, which is given by:

$$\overline{N}(q^\star, \epsilon, \overline{p}) \equiv [\underbrace{(w(q^\star + \epsilon)(q^\star + \epsilon) - r(q^\star + \epsilon)(q^\star + \epsilon)}_{\pi_U(q^\star+\epsilon)}]^\alpha [\underbrace{(p(q^\star)q^\star + \overline{p}\epsilon - w(q^\star + \epsilon)(q^\star + \epsilon))}_{\overline{\pi}_D(q^\star,\epsilon,\overline{p})}]^{1-\alpha}. \tag{31}$$

Using a Taylor expansion of the Nash product in (31) around $q^\star$ yields:

$$N(q^\star + \epsilon) = N(q^\star) + \frac{\partial N(q^\star)}{\partial q}\epsilon + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^\star)}{n!} = N(q^\star) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^\star)}{n!}. \tag{32}$$

Using (32), we can rewrite (30) as:

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

and then use (31) to obtain:

$$\pi_U(q^\star + \epsilon)^\alpha \overline{\pi}_D(q^\star, \epsilon, \overline{p})^{1-\alpha} \geq \pi_U(q^\star + \epsilon)^\alpha \pi_D(q^\star + \epsilon)^{1-\alpha} - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^\star)}{n!}. \tag{33}$$

Defining $R_D(q^\star + \epsilon) \equiv p(q^\star + \epsilon)(q^\star + \epsilon)$ and applying a Taylor expansion around $q^\star$ yields:

$$R_D(q^\star + \epsilon) = p(q^\star)q^\star + \underbrace{[p'(q^\star)q^\star + p(q^\star)]}_{MR_D(q^\star)}\epsilon + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{R_D^{(n)}(q^\star)}{n!}. \tag{34}$$

Using (34), and given that $\sum_{n\geq 2}^{\infty} \epsilon^n \frac{N^{(n)}(q^\star)}{n!} \to 0$ and $\sum_{n\geq 2}^{\infty} \epsilon^n \frac{R^{(n)}(q^\star)}{n!} \to 0$ as $\epsilon \to 0$, we can simplify (33) as follows:

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

As a result, the minimum price required for D to supply the marginal unit ϵ is $\hat{p} = MR_D(q^\star)$. According to our definition, it follows that D's markup is given by: $\mu_D(q^\star) \equiv \frac{p(q^\star)}{\hat{p}(q^\star)} = \frac{p(q^\star)}{MR_D(q^\star)}$.

***U*'s markup $\boldsymbol{\mu_U}$.** Consider a hypothetical scenario in which, in addition to selling $q^\star$ units to D at price w, U decides whether to sell an infinitesimal quantity ϵ at price $\overline{w}$. Hence, U incurs no revenue loss on inframarginal units when selling ϵ to D.[64] Based on this hypothetical scenario, we determine the minimum price $\hat{w}$ such that U offers this additional quantity ϵ, resulting in a total quantity sold of $q^\star + \varepsilon$ to D. Formally, $\hat{w}$ is the smallest value of $\overline{w}$ such that:

$$
\overline{N}(q^\star, \epsilon, \overline{w}) \geq N(q^\star) \tag{35}
$$

where $N(q^\star) = \pi_U(q^\star)^\alpha \pi_D(q^\star)^{(1-\alpha)}$ and $\overline{N}(q^\star, \epsilon, \overline{w})$ corresponds to the value of the Nash product for a quantity $q^\star + \varepsilon$ in the hypothetical scenario, which is given by:

$$
\overline{N}(q^\star, \epsilon, \overline{w}) = (w(q^\star)q^\star + \overline{w}\epsilon - r(q^\star+\epsilon)(q^\star+\epsilon))^\alpha \left(p(q^\star+\epsilon)(q^\star+\epsilon) - w(q^\star)q^\star - \overline{w}\epsilon\right)^{1-\alpha}. \tag{36}
$$

Substituting (36) in (35), we obtain:

$$
(w(q^\star)q^\star + \overline{w}\epsilon - r(q^\star+\epsilon)(q^\star+\epsilon))^\alpha \left(p(q^\star+\epsilon)(q^\star+\epsilon) - w(q^\star)q^\star - \overline{w}\epsilon\right)^{1-\alpha} \geq N(q^\star). \tag{37}
$$

Defining $C_U(q^\star+\epsilon) \equiv r(q^\star+\epsilon)(q^\star+\epsilon)$ and applying a Taylor expansion around $q^\star$ yields:

$$
r(q^\star+\epsilon)(q^\star+\epsilon) = r(q^\star)q^\star + \underbrace{[r'(q^\star)q^\star + r(q^\star)]}_{MC_U(q^\star)}\epsilon + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{C_U^{(n)}(q^\star)}{n!}. \tag{38}
$$

[64] When bargaining, U and D share their joint profit, so trading an additional unit affects the profits of both firms. In particular, U's revenue loss on inframarginal units when selling an additional unit is now twofold: a direct effect and an indirect effect through the internalization of D's profit, as D also incurs a revenue loss on inframarginal units when selling the additional quantity to consumers. Consistent with our markup definition, the hypothetical scenario we consider eliminates both sources of such a revenue loss.

Using (34) and (38), we can rewrite (37) as:

$$\left[(\overline{w} - MC_U(q^\star))\epsilon + \pi_U(q^\star) - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{C_U^{(n)}(q^\star)}{n!}\right]^\alpha \left[(MR_D(q^\star) - \overline{w})\epsilon + \pi_D(q^\star) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{R_D^{(n)}(q^\star)}{n!}\right]^{1-\alpha} \geq N(q^\star). \tag{39}$$

If $q^\star = q_\mu$, we can further simplify (39) using $\overline{w} = MR_D(q_\mu + \epsilon)$. In particular, applying a Taylor expansion to $MR_D(q_\mu + \epsilon)$ around q_μ yields:

$$\begin{aligned} & MR_D(q_\mu + \epsilon) = MR_D(q_\mu) + \epsilon MR'_D(q_\mu) + \sum_{n\geq 2}^{\infty} \epsilon^n \frac{MR_D^{(n)}(q_\mu)}{n!} \\ \Leftrightarrow \quad & MR_D(q_\mu) - \overline{w} = -\epsilon MR'_D(q_\mu) - \sum_{n\geq 2}^{\infty} \epsilon^n \frac{R_D^{(n)}(q_\mu)}{n!}. \end{aligned} \tag{40}$$

Substituting (40) in (39), and using $N(q_\mu) = \pi_U(q_\mu)^\alpha \pi_D(q_\mu)^{1-\alpha}$ yields:

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

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

$$\begin{aligned} & (\epsilon(\overline{w} - MC_U(q_\mu)) + \pi_U(q_\mu))^\alpha \pi_D(q_\mu)^{1-\alpha} \geq \pi_U(q_\mu)^\alpha \pi_D(q_\mu)^{1-\alpha} \\ \Leftrightarrow \quad & \overline{w} \geq MC_U(q_\mu). \end{aligned} \tag{41}$$

As a result, the minimum price at which U and D agree to exchange the marginal unit ϵ is $\hat{w} = MC_U(q_\mu)$. According to our definition, it follows that U's markup is given by: $\mu_U(q_\mu) \equiv \frac{w}{\hat{w}} = \frac{w(q_\mu)}{MC_U(q_\mu)}$. The reasoning is the same when instead $q^\star = q_\nu$, $w = MC_U(q_\nu)$, and $\mu_U(q_\nu) \equiv \frac{w}{\hat{w}} = \frac{w(q_\nu)}{MC_U(q_\nu)} = 1$ (proof available upon request).

D's markdown ν_D. Consider a hypothetical scenario in which, in addition to buying $q^\star$ units to U at price w, D decides whether to purchase an infinitesimal quantity ϵ at price $\overline{w}$. Hence, D incurs no cost increase on inframarginal units when buying the additional quantity ϵ.[65] Based on this hypothetical scenario, we determine the maximum price $\hat{w}$ such that D purchases this additional quantity ϵ, resulting in a total quantity purchased of $q^\star + \varepsilon$ to U.

Formally, $\hat{w}$ is the highest value of $\overline{w}$ such that (35) is satisfied. Assume first that $q^\star = q_\mu$, implying that $\overline{w} = MR_D(q_\mu + \epsilon)$ and (41) is satisfied. If the maximum price $\hat{w}$ at which the additional

[65] Again, when bargaining, U and D share their joint profit, so trading an additional unit affects both firms' profits. In particular, D's cost increase on inframarginal units when purchasing an additional unit is now twofold: a direct effect and an indirect effect through the internationalization of U's profit, as U also incurs a cost increase on inframarginal units when purchasing this additional quantity to input suppliers. Consistent with our markdown definition, the hypothetical scenario we consider eliminates both sources of such a cost increase.

unit ϵ is purchased by D is such that $q_\mu + \epsilon = \tilde{q}_D(\overline{w}) \leq \tilde{q}_U(\overline{w})$, then (41) yields a lower bound on $\overline{w}$ and $\hat{w} = w(q_\mu) = MR_D(q_\mu)$. According to our definition, it follows that D's markdown is equal to $\nu_D \equiv \frac{\hat{w}(q_\mu)}{w(q_\mu)} = \frac{MR_D(q_\mu)}{MR_D(q_\mu)} = 1$.

Assume instead that $q^\star = q_\nu$, implying that $\overline{w} = MR_D(q_\nu + \epsilon)$ and (35) is satisfied when:

$$\overline{w} \leq MR_D(q_\nu).$$

If the maximum price $\hat{w}$ at which the additional unit ϵ is purchased by D is such that $q_\nu + \epsilon = \tilde{q}_U(\overline{w}) \leq \tilde{q}_D(\overline{w})$, then $\hat{w} = MR_D(q_\nu)$. According to our definition, it follows that D's markdown is equal to $\nu_D(q_\nu) \equiv \frac{\hat{w}}{w} = \frac{MR_D(q_\nu)}{w(q_\nu)}$.

***U*'s markdown ν_U.** Consider a hypothetical scenario in which, in addition to buying $q^\star$ units to input suppliers at price r, U decides whether to buy an infinitesimal quantity ϵ at price $\overline{r}$. Hence, U incurs no cost increase on inframarginal units when buying the additional quantity ϵ, as the purchasing price for the $q^\star$ units remains unchanged. Based on this hypothetical scenario, we determine the maximum price $\hat{r}$ such that U purchases the additional quantity ϵ, resulting in a total quantity purchased of $q^\star + \varepsilon$. Formally, $\hat{r}$ is the highest value of $\overline{r}$ such that:

$$\overline{N}(q^\star, \epsilon, \overline{r}) \geq N(q^\star). \tag{42}$$

where, analogously to (31), we have:

$$\overline{N}(q^\star, \epsilon, \overline{r}) = [\underbrace{(w(q^\star + \epsilon)(q^\star + \epsilon) - r(q^\star)q^\star - \overline{r}\epsilon}_{\overline{\pi}_U(q^\star, \epsilon, \overline{r})}]^\alpha [\underbrace{(p(q^\star + \epsilon)(q^\star + \epsilon) - w(q^\star + \epsilon)(q^\star + \epsilon))}_{\pi_D(q^\star + \epsilon)}]^{1-\alpha}. \tag{43}$$

Using (42) and (43), the derivation of $\hat{r}$ is similar to that of $\hat{p}$ in the characterization of D's markup. Specifically, based on Taylor expansions and considering $\epsilon \to 0$, one can show that (42) boils down to:

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

As a result, the maximum price at which U would purchase the marginal unit ϵ is $\hat{r} = MC_U(q_\mu)$. According to our definition, it follows that U's markdown is given by: $\nu_U(q^\star) \equiv \frac{\hat{r}}{r} = \frac{MC_U(q^\star)}{r(q^\star)}$.

B Alternative Formulation for Stage 2

In Section 4, we consider a bilateral monopoly model where, in Stage 2, U and D each announce the quantity they are willing to trade. Holding Stage 1 unchanged, we propose an equivalent formulation, where the input and consumer prices are set directly, and show that it yields the same equilibrium outcome as described in Lemma 1.

Stage 2 (alternative): Input and consumer price setting. D sets the consumer 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.

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 consumer 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.[66] Similarly, U recognizes that the maximum quantity it can sell to D is limited by p.[67] 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:

$$\max_{p} \; (p - w)\, q_D(p) \quad \text{subject to} \quad q_D(p) \leq q_U(r^a) \tag{44}$$

where the constraint reflects that D cannot sell more quantity than what U is able to procure at r^a. An interior solution to (44) 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 (Assumption 1.(ii)). Similarly, anticipating that $p = p^a$, U's maximization problem is as follows:

$$\max_{r} \; (w - q_U(r))\, q_U(r) \quad \text{subject to} \quad q_U(r) \leq q_D(p^a) \tag{45}$$

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 (45) arises when $q_U(\tilde{r}) \leq q_D(p^a)$, where $\tilde{r}$ satisfies the following first-order condition:

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

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_U is increasing (Assumption 1.(i)).

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 $q_U(\tilde{r}) \geq q_D(p^a)$,

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

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

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 profit-maximizing 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 described in Lemma 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 consumer 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.

C Optimal Input Price Floor

C.1 Vertical Integration

Absent any price floor, recall that I's profit maximization leads to $MR_I(q_I) = MC_I(q_I)$. Hence, we say that the price floor $\underline{r}$ is binding when $r(q_I) < r(\underline{q})$ implying that $\underline{q} > q_I$. As $MR_I(q)$ is decreasing and $MC_I(q)$ is increasing we have $MR_I(\underline{q}) \leq MC_I(\underline{q})$. We now demonstrate that the optimal price floor is such that $\underline{r_I} = r(\underline{q_I})$, with $MR_I(\underline{q_I}) = r(\underline{q_I})$.

- Consider a deviation towards a higher price floor $\underline{r_I^+} > \underline{r_I}$. I's profit-maximizing quantity is such that $MR_I(q) = \underline{r_I^+} \Rightarrow \underline{q_I^+} < \underline{q_I}$, as $MR_I(q)$ is decreasing in q. Hence, a deviation towards $\underline{r_I^+}$ reduces welfare.

- Consider a deviation towards a lower price floor $\underline{r_I^-} < \underline{r_I}$, such that $\underline{r_I^-}$ is binding (i.e., $r(q_I) < r(\underline{r_I^-})$). Note that $q_I < \underline{q_I^-} < \underline{q_I}$ as $r(q)$ is increasing in q (Assumption 1).

  For any $q > \underline{q_I^-}$, I's marginal cost is given by $MC_I(q) > MR_I(q)$, implying that I never chooses $q > \underline{q_I^-}$. For any $q < \underline{q_I^-}$, I's marginal cost is given by $\underline{r_I^-} < MR_I(q)$, implying that I never chooses $q < \underline{q_I^-}$. So I always chooses $q = \underline{q_I^-} < \underline{q_I}$. Hence, $\underline{r_I^-}$ reduces welfare compared to $\underline{r_I}$.

C.2 Vertical Supply Chain

Consider a price floor $\underline{r}$ that is binding. For any $q \leq \underline{q}$, as $MC_U(q) = \underline{r}$, we know that D determines the traded quantity in Stage 2 according to $w = MR_D(q)$. In that case, the first-order condition for the maximization of the Nash product under $\underline{r}$ is given by:

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

where $\underline{\beta_D}(q, \alpha, \underline{r}) = \frac{(1-\alpha)\underline{\pi_U}(q,\underline{r})}{\alpha \pi_D(q)}$ and $\underline{\pi_U}(q, \underline{r}) = (MR_D(q) - \underline{r})q$.

Note that the second-order condition for the maximization of the Nash product under a price floor $\underline{r}$ is satisfied whenever $\sigma_{MR_D} < 2$ (Assumption 2.(ii)). To see this, we have $\frac{\partial \underline{\widetilde{MR}_U}(q,\alpha,\underline{r})}{\partial q} = \frac{\partial \underline{\beta_D}(q,\alpha,\underline{r})}{\partial q}(MR_D(q) - MR_U(q)) + \underline{\beta_D}(q, \alpha, \underline{r}) MR'_D(q) + (1 - \underline{\beta_D}(q, \alpha, \underline{r})) MR'_U(q)$ and $\frac{\partial \underline{\beta_D}(q,\alpha,\underline{r})}{\partial q} = \frac{(1-\alpha)\underline{\pi'_U}(q,\underline{r})\pi_D(q) - \alpha \pi'_D(q) \underline{\pi_U}(q,\underline{r})}{\alpha^2 \pi_D(q)^2}$. As $\underline{\pi'_U}(q, \underline{r}) < 0$ and $\pi'_D(q) > 0$, we have $\frac{\partial \underline{\beta_D}(q,\alpha,\underline{r})}{\partial q} < 0$. Moreover, $\sigma_{MR_D} < 2$ implies $MR'_U(q) < 0$ and $\sigma_p(q) < 2$ guarantees that $MR'_D(q) < 0$. As $MR_D(q) - MR_U(q) > 0$, it follows that $\frac{\partial \underline{\widetilde{MR}_U}(q,\alpha,\underline{r})}{\partial q} < 0$.

Note also that, as $\underline{\pi_U}(q, \underline{r})$ decreases in $\underline{r}$, we have $\frac{\partial \underline{\widetilde{MR}_U}(q,\alpha,\underline{r})}{\partial \underline{r}} = \frac{\partial \underline{\beta_D}(q,\alpha,\underline{r})}{\partial \underline{r}}(MR_D(q) - MR_U(q)) < 0$. Furthermore, as $\underline{\pi_U}(\underline{q}, \underline{r}) = \pi_U(\underline{q})$, we have $\underline{\widetilde{MR}_U}(\underline{q}, \alpha, \underline{r}) = \widetilde{MR}_U(\underline{q}, \alpha)$. Based on these properties, we define the optimal input price floor $\underline{r_\mu} = r(\underline{q_\mu})$, where $\underline{q_\mu}$ solves:

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

C.2.1 Proof of Proposition 6

In what follows, we demonstrate that (i) $\underline{r_\mu}$ increases welfare, and (ii) no alternative price floor yields higher welfare.

Proof that $\underline{q_\mu} > q^\star = \{q_\mu, q_\nu\}$. When $\alpha \in [\alpha_I, 1]$, the equilibrium absent price floor is given by $\widetilde{MR}_U(q_\mu, \alpha) = MC_U(q_\mu) > r(q_\mu)$. As $\widetilde{MR}_U(q, \alpha)$ is decreasing in q (Appendix A.5.1), we obtain $q_\mu < \underline{q_\mu}$ for $\alpha \in [\alpha_I, 1]$. When $\alpha = \alpha_I$, as $q_\mu = q_I$, we also have $\underline{q_\mu} > q_I$. Following arguments similar to those developed in Appendix A.5.3, we also know that $\underline{q_\mu}$ is decreasing in α. As for $\alpha \in [0, \alpha_I]$, the equilibrium quantity absent price floor is $q_\nu \leq q_I$, we obtain $\underline{q_\mu} > q_I \geq q_\nu$.

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

- Consider a deviation towards a higher price floor $\underline{r_\mu^+} > \underline{r_\mu}$. The maximization of the Nash product is such that $\underline{\widetilde{MR}_U}(\underline{q_\mu^+}, \alpha, \underline{r_\mu^+}) = \underline{r_\mu^+} > \underline{r_\mu} = \underline{\widetilde{MR}_U}(\underline{q_\mu}, \alpha, \underline{r_\mu}) > \underline{\widetilde{MR}_U}(\underline{q_\mu}, \alpha, \underline{r_\mu^+})$, as $\underline{\widetilde{MR}_U}(q, \alpha, \underline{r})$ is decreasing in $\underline{r}$. Moreover, as $\underline{\widetilde{MR}_U}(q, \alpha, \underline{r})$ is decreasing in q, we have $q_\lambda < \underline{q_\mu}$, implying that the deviation is not welfare-improving.

- Consider a deviation towards a lower price floor $\underline{r_\mu^-} < \underline{r_\mu}$. With $\underline{q_\mu^-}$ the quantity such that $\underline{r_\mu^-} = r(\underline{q_\mu^-})$, we have $\underline{q_\mu^-} < \underline{q_\mu}$ as $r(q)$ is increasing. It implies that any binding price floor $\underline{r_\mu^-} < \underline{r_\mu}$ would generate an equilibrium quantity $\underline{q_\mu^-} < \underline{q_\mu}$.

  Suppose now that the deviation is such that the price floor is no longer binding. Depending on α, the equilibrium quantity absent a price floor q_ν or q_μ may arise. However, we proved that $q_\mu < \underline{q_\mu}$ for all $\alpha \in [\alpha_I, 1]$. We also proved that $q_\nu \leq q_I < \underline{q_\mu}$ for any $\alpha \in [0, \alpha_I]$. The deviation is not welfare-improving.

C.2.2 Variation of $\underline{q_\mu}$ and $\underline{r}(\underline{q_\mu})$ with respect to α

The first-order partial derivative of the Nash product under $\underline{r_\mu}$ is given by: $\Gamma(\underline{q_\mu}, \alpha) = \widetilde{MR}_U(\underline{q_\mu}, \alpha) - r(\underline{q_\mu})$. By the implicit function theorem, we have: $\frac{d\underline{q_\mu}}{d\alpha} = -\frac{\frac{\partial \Gamma(\underline{q_\mu}, \alpha)}{\partial \alpha}}{\frac{\partial \Gamma(\underline{q_\mu}, \alpha)}{\partial q}}$. The second-order condition for the maximization of the Nash product ensures that $\frac{\partial \Gamma(\alpha, \underline{q_\mu})}{\partial q} < 0$. Hence, $Sign\left(\frac{d\underline{q_\mu}(\alpha)}{d\alpha}\right) = Sign\left(\frac{\partial \Gamma(\alpha, \underline{q_\mu})}{\partial \alpha}\right) = Sign\left(\frac{\partial \beta_D}{\partial \alpha}(MR_D - MR_U)\right) = Sign\left(\frac{-\pi_U(\underline{q_\mu})\pi_D(\underline{q_\mu})}{\alpha^2 \pi_D(\underline{q_\mu})^2}(MR_D(\underline{q_\mu}) - MR_U(\underline{q_\mu}))\right) < 0$, as $MR_D(\underline{q_\mu}) > MR_U(\underline{q_\mu})$.

C.2.3 Effect of the Optimal Price Floor on Firms' Profits and Welfare

Consider first that $\alpha_I < \alpha < 1$. In this case, absent price floor, U's profit is maximized at $q_{\overline{\mu}}$, $\frac{\partial \pi_U}{\partial q} = MR_U(q) - MC_U(q) \leq 0$ for $q \geq q_{\overline{\mu}}$ (Assumption 1.(i) and Assumption 1.(ii)), and its equilibrium profit is $\pi_U(q_\mu) = (MR_D(q_\mu) - r(q_\mu))q_\mu$. Under optimal price floor, U's equilibrium profit is $\underline{\pi_U}(\underline{q_\mu}, \underline{r_\mu}) = (MR_D(\underline{q_\mu}) - r(\underline{q_\mu}))\underline{q_\mu} = \pi_U(\underline{q_\mu})$. Given that $q_{\overline{\mu}} < q_\mu < \underline{q_\mu}$, we have $\pi_U(\underline{q_\mu}) < \pi_U(q_\mu)$, which implies that U's profit is negatively affected by the optimal price floor. Regarding D's profit, absent price floor, we have $\pi_D(q) = (p(q) - MR_D(q))\, q$ and, under Assumption 1.(i) and Assumption 1.(ii), $\frac{\partial \pi_D}{\partial q} = MR_D(q) - MR_U(q) > 0$ for $q > 0$. Under optimal price floor, D's equilibrium profit is $\underline{\pi_D}(\underline{q_\mu}, \underline{r_\mu}) = (p(\underline{q_\mu}) - MR_D(\underline{q_\mu}))\underline{q_\mu} = \pi_D(\underline{q_\mu})$. Given that $q_\mu < \underline{q_\mu}$, we have $\pi_D(\underline{q_\mu}) > \pi_D(q_\mu)$, which implies that D's profit is positively affected by the optimal price floor. Finally, as $q_\mu < \underline{q_\mu}$, consumers and input suppliers also benefit from the optimal price floor.

Consider now that $0 < \alpha < \alpha_I$. The optimal price floor affects U's profit through two channels: the equilibrium quantity and U's margin. The quantity effect is positive, as $q_\nu < \underline{q_\mu}$. However, the margin effect can be negative and outweigh the quantity effect. For instance, when $\alpha = 0$, it is clear that the price floor reduces U's margin and we obtain $\pi(\underline{q_\mu}) = 0 < \pi_U(q_\nu)$. The price floor also negatively affects U's profit when $\alpha = \alpha_I$ as $\underline{\pi_U}(\underline{q_\mu}, \underline{r_\mu}) < \pi_U(q_\mu) = \pi_U(q_\nu) = \pi_U(q_I)$. More generally, for $0 < \alpha < \alpha_I$, the effect of the optimal price floor on U's profit remains undetermined, depending on demand and supply primitives. Focusing on linear demand and supply functions, we find that U's profit is always negatively affected by the optimal price floor. The optimal price floor on D's profit follows a similar logic. The quantity effect is positive, but the margin effect remains ambiguous and depends on demand and supply primitives. Intuitively, when demand is highly elastic (e.g., $\epsilon_p \to \infty$), D's profit mainly comes from its markdown ν_D. In this case, by eliminating D's ability to exert monopsony power, the optimal price floor always reduces D's profit. In the linear case, the price floor negatively affects D's profit when both α and the demand slope are relatively low, and has a positive effect otherwise. Finally, as $q_\nu < \underline{q_\mu}$, consumers and input suppliers always benefit from the optimal price floor.

C.2.4 Graphical Representation

Figure 7 illustrates the effect of introducing an optimal input price floor when D is powerful ($\alpha < \alpha_I$), i.e., when double markdownization otherwise prevails. The equilibrium without price floor is displayed in semi-transparency, whereas the equilibrium with an optimal input price floor is displayed in plain colors.

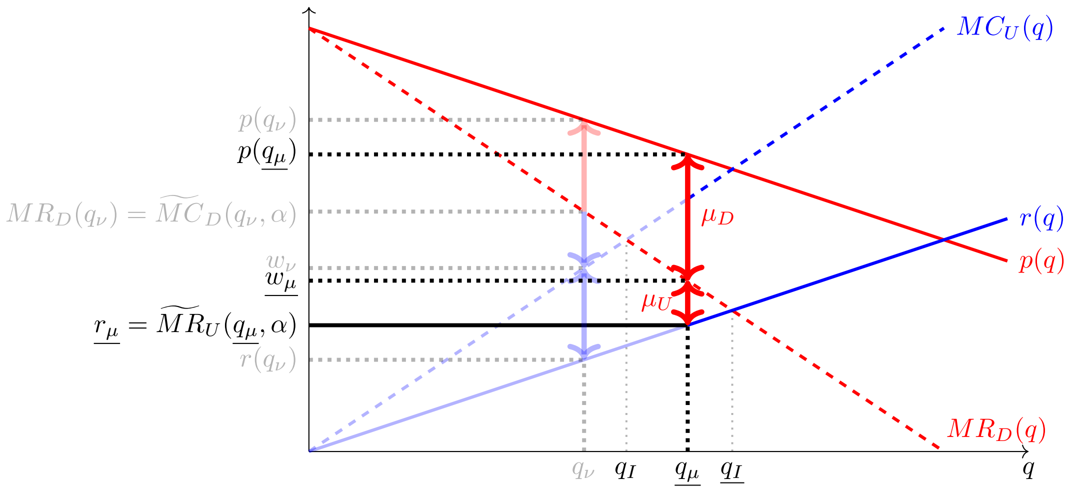

Figure 7: Optimal Price Floor Compared to Double Markdownization ($0 < \alpha < \alpha_I$).

Notes: The figure is drawn under the demand function $p(q) = p - \frac{1}{3}q$ and the supply function $r(q) = r + \frac{1}{3}q$. The equilibrium without price floor is displayed in semi-transparency, with the red arrow and blue arrows representing the markup and markdown wedges in differences, respectively. The equilibrium with an optimal input price floor is displayed in plain colors, with the red arrows labeled μ_D and μ_U representing the resulting markup wedges in differences.

D Two-part Tariff Contract

Consider the bilateral monopoly setting introduced in Section 3.1 and Section 4. Instead of bargaining over a linear wholesale price w, suppose that U and D bargain over a two-part tariff (w, F). Assume further that the use of the fixed fee is subject to frictions, such that $\underline{F} \leq F \leq \overline{F}$. Proceeding backwards, we first analyze the quantity choice stage (Stage 2) before turning to the bargaining stage (Stage 1).

Stage 2: Quantity choice. As the fixed fee F never affects firms' quantity choice, Lemma 1 applies and the quantity traded in equilibrium is $q(w) = \min\{\tilde{q}_U(w), \tilde{q}_D(w)\}$.

Stage 1: Bargaining. 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 following maximization problem:[68]

$$\max_{q,F} \Pi_U(q,F)^{\alpha}\Pi_D(q,F)^{1-\alpha} \quad \text{subject to} \quad \underline{F} \leq F \leq \overline{F} \tag{46}$$

The first-order condition on F is given by:

$$\alpha\Pi_D(q,F) - (1-\alpha)\Pi_U(q,F) = 0 \tag{47}$$

Using (47), the first-order condition on q when $w(q) = MR_D(q)$ (i.e., $\tilde{q}_D(w) < \tilde{q}_U(w)$) reduces to:

$$(MR_U(\hat{q}_\mu) - MC_U(\hat{q}_\mu)) + (MR_D(\hat{q}_\mu) - MR_U(\hat{q}_\mu)) = 0 \tag{48}$$

where $\hat{q}_\mu$ denotes the equilibrium traded quantity. Similarly, when $w(q) = MC_U(q)$ (i.e., $\tilde{q}_D(w) > \tilde{q}_U(w)$), the first-order condition on q is given by:

$$(MC_D(\hat{q}_\nu) - MC_U(\hat{q}_\nu)) + (MR_D(\hat{q}_\nu) - MC_D(\hat{q}_\nu)) = 0 \tag{49}$$

where $\hat{q}_\nu$ denotes the equilibrium traded quantity. Three types of equilibria may 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 (that is, (47) always holds). From (48) and (49), we obtain $MC_U(\hat{q}_\mu) = MR_D(\hat{q}_\mu)$ and $MC_U(\hat{q}_\nu) = MR_D(\hat{q}_\nu)$, implying that $\hat{q}_\mu = \hat{q}_\nu = q_I$. 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 3.2. At $q = q_I$, (47) 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 $\underline{\alpha} \equiv \frac{\pi_U(q_I)+\underline{F}}{\pi_U(q_I)+\pi_D(q_I)} > 0$

[68] Note that maximizing 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).

such that the fixed fee D receives from U is bounded below by $\underline{F}$.[69] Therefore, when $\overline{\alpha} > \alpha > \underline{\alpha}$, (47) holds and the vertically integrated outcome arises: i.e., the equilibrium quantity is q_I and the equilibrium fixed fee is $\hat{F} = \alpha\pi_D(q_I) - (1-\alpha)\pi_U(q_I)$. When instead $\alpha_I \leq \overline{\alpha} \leq \alpha$, we have $\hat{F} = \overline{F}$. As $\alpha_I \leq \alpha$, we have $w(q) = MR_D(q)$, implying that the first-order condition of (46) with respect to w is given by:

$$MC_U(\hat{q}_\mu) \quad = \quad \widehat{MR}_U(\hat{q}_\mu, \overline{F}, \alpha) \tag{50}$$

where $\widehat{MR}_U(\hat{q}_\mu, \overline{F}, \alpha) \equiv \hat{\beta}_D(\hat{q}_\mu, \overline{F}, \alpha) MR_D(\hat{q}_\mu) + (1 - \hat{\beta}_D(\hat{q}_\mu, \overline{F}, \alpha)) MR_U(\hat{q}_\mu)$, with $\hat{\beta}_D(\hat{q}_\mu, \overline{F}, \alpha) \equiv \frac{1-\alpha}{\alpha} \frac{\Pi_U(\hat{q}_\mu, \overline{F})}{\Pi_D(\hat{q}_\mu, \overline{F})}$. Interestingly, (50) mirrors (12), which yields the double markup outcome under the linear wholesale contract setting. As $\hat{\beta}_D(\hat{q}_\mu, \overline{F}, \alpha) \geq \beta_D(q_\mu, \alpha)$, we have $MR_D(\hat{q}_\mu) \geq \widehat{MR}_U(\hat{q}_\mu, \overline{F}, \alpha) \geq \widetilde{MR}_U(q_\mu, \alpha)$, implying that the equilibrium quantity $\hat{q}_\mu$ is such that $q_I \geq \hat{q}_\mu \geq q_\mu$.[70] Conversely, when $\alpha_I \geq \underline{\alpha} \geq \alpha$, we have $\hat{F} = \underline{F}$. As $\alpha_I \geq \alpha$, we have $w(q) = MC_U(q)$, implying that the first-order condition of (46) with respect to w is given by:

$$MR_D(\hat{q}_\nu) = \widehat{MC}_D(\hat{q}_\nu, \underline{F}, \alpha) \tag{51}$$

where $\widehat{MC}_D(\hat{q}_\nu, \underline{F}, \alpha) \equiv \hat{\beta}_U(\hat{q}_\nu, \underline{F}, \alpha) MC_U(\hat{q}_\nu) + (1 - \hat{\beta}_U(\hat{q}_\nu, \underline{F}, \alpha)) MC_D(\hat{q}_\nu, \alpha)$, with $\hat{\beta}_U(\hat{q}_\nu, \underline{F}, \alpha) \equiv \frac{\alpha}{1-\alpha} \frac{\Pi_D(\hat{q}_\nu, \underline{F})}{\Pi_U(\hat{q}_\nu, \underline{F})}$. Again, (51) reflects (10), which characterizes the double markdown outcome in the linear wholesale contract setting. As $\hat{\beta}_U(\hat{q}_\nu, \underline{F}, \alpha) \geq \beta_U(q_\nu, \alpha)$, we have $MC_U(\hat{q}_\nu) \geq \widehat{MC}_D(\hat{q}_\nu, \underline{F}, \alpha) \geq \widetilde{MC}_D(q_\nu, \alpha)$, implying that the equilibrium quantity $\hat{q}_\nu$ is such that $q_I \geq \hat{q}_\nu \geq q_\nu$.[71]

[69] 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.

[70] Note first that when $\hat{F} = \overline{F} = 0$, we have $\hat{q}_\mu = q_\mu$, 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 $\Pi_U(\hat{q}_\mu, \overline{F}) \geq \Pi_U(\hat{q}_\mu, 0) = \pi_U(q_\mu)$ and $\Pi_D(\hat{q}_\mu, \overline{F}) \leq \Pi_D(\hat{q}_\mu, 0) = \pi_D(q_\mu)$. As a result, we obtain $\hat{\beta}_D(\hat{q}_\mu, \overline{F}, \alpha) \geq \hat{\beta}_D(\hat{q}_\mu, 0, \alpha) = \beta_D(q_\mu, \alpha)$.

[71] The reasoning parallels that in footnote 70. In particular, whenever $\hat{F} = \underline{F} \leq 0$, we have $\Pi_U(\hat{q}_\nu, \underline{F}) \leq \Pi_U(\hat{q}_\nu, 0) = \pi_U(q_\nu)$ and $\Pi_D(\hat{q}_\nu, \underline{F}) \geq \Pi_D(\hat{q}_\nu, 0) = \pi_D(q_\nu)$. Consequently, $\hat{\beta}_U(\hat{q}_\nu, \underline{F}, \alpha) \geq \hat{\beta}_U(\hat{q}_\nu, 0, \alpha) = \beta_U(q_\nu, \alpha)$.

Online Appendix for "Markups, Markdowns and Bargaining in a Vertical Chain"

Rémi Avignon, Claire Chambolle, Etienne Guigue, and Hugo Molina

July 23, 2025

OA1 Alternative Expressions of Markups and Markdowns

When D is powerful ($\alpha < \alpha_I$), its markdown can be rewritten as:

$$\nu_D = \omega_D(q_\nu, \alpha)\frac{\varepsilon_{MC_U} + 1}{\varepsilon_{MC_U}} + (1 - \omega_D(q_\nu, \alpha))\frac{\varepsilon_p - 1}{\varepsilon_p}, \tag{52}$$

where $\omega_D(q_\nu, \alpha) \equiv \frac{\alpha(\varepsilon_r+1)\varepsilon_p}{\alpha(\varepsilon_r+1)\varepsilon_p+(1-\alpha)(\varepsilon_p-1)\varepsilon_{MC_U}} \in (0, 1)$, with $\frac{\partial \omega_D(q_\nu,\alpha)}{\partial \alpha} < 0$.[72] Equation (52) echoes expressions delivered by the exogenous right-to-manage models of Alviarez et al. (2023); Azkarate-Askasua and Zerecero (2022) and Wong (2023). Specifically, here, the bilateral distortion is a weighted average between a markup and a markdown term. If $\alpha = 1$, we obtain $\omega_D(q_\nu, \alpha) = 1$ and $\nu_D = \frac{\varepsilon_{MC_U}+1}{\varepsilon_{MC_U}}$, as D makes a take-it-or-leave-it offer to U. If $\alpha = 0$, we obtain $\omega_D(q_\nu, \alpha) = 0$ and (52) states $\nu_D = \mu_D^{-1}$ and thus $r = w$ and $M_D = 1$. However, under voluntary exchange, such an equilibrium is ruled out, as U endogenously concedes the right-to-manage if becoming too powerful, i.e., when $\alpha > \alpha_I$. Instead, at the relevant limit, i.e., if $\alpha \to \alpha_I = \frac{\varepsilon_p-1}{\varepsilon_p+\varepsilon_r}$, we have $\omega_D(\alpha_I) = \frac{\varepsilon_p}{\varepsilon_{MR_U}+\varepsilon_p}$, and thus $\nu_D = 1$, yielding the vertically integrated outcome q_I. Overall, when $\alpha \in (0, \alpha_I)$, D's markdown ν_D negatively depends on its markup μ_D, reflecting the influence of bargaining on the joint profit sharing.

Similarly, when U is powerful ($\alpha > \alpha_I$), its markup can be rewritten as:

$$\mu_U = \omega_U(q_\mu, \alpha)\frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D} - 1} + (1 - \omega_U(q_\mu, \alpha))\frac{\varepsilon_r}{\varepsilon_r + 1}, \tag{53}$$

where $\omega_U(q_\mu, \alpha) \equiv \frac{\alpha(\varepsilon_{MR_D}-1)}{\alpha(\varepsilon_{MR_D}-1)+(1-\alpha)(\varepsilon_p-1)} \in (0, 1)$ and $\frac{\partial \omega_U(q_\mu,\alpha)}{\partial \alpha} > 0$.[73] Similar to (52), (53) shows that U's markup is a weighted average between a markup and a markdown term. If $\alpha = 1$, then $\omega_U(q_\mu, \alpha) = 1$ and $\mu_U = \frac{\varepsilon_{MR_D}}{\varepsilon_{MR_D}-1}$, as U makes a take-it-or-leave-it offer to D. If $\alpha = 0$ then $\omega_U(q_\mu, \alpha) = 0$, and (53) would state $\mu_U = \frac{\varepsilon_r}{\varepsilon_r+1} = \nu_U^{-1}$ and thus $p = w$ and $M_U = 1$. Again, under voluntary

[72] $\omega_U(q_\mu, \alpha) > 0$ holds under Assumption 1.

[73] $\omega_U(q_\mu, \alpha) > 0$ holds when $\varepsilon_{MR_D} > 1$, i.e., the demand function is supermodular (Assumption 2).

exchange, such an equilibrium is ruled out, as D endogenously concedes the right-to-manage when too powerful, i.e, $\alpha < \alpha_I$. Instead, at this limit, i.e., if $\alpha \to \alpha_I = \frac{\varepsilon_p - 1}{\varepsilon_p + \varepsilon_r}$, we have $\omega_U(\alpha_I) = \frac{\varepsilon_{MR_D} - 1}{\varepsilon_{MR_D} + \varepsilon_r}$ and thus $\mu_U = 1$, yielding the vertically integrated outcome q_I. Overall, when $\alpha \in (\alpha_I, 1)$, U's markup μ_U negatively depends on its markdown ν_U, as a consequence of bargaining.

OA2 Microfoundation

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). Specifically, 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: Quantity choice.** U and D simultaneously announce the quantities q_U and q_D they are each willing to trade. Exchange is voluntary, implying that the quantity traded is the minimum of q_U and q_D.

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 non-cooperative game described above.

As Stage 2 remains unchanged compared to Section 4, we focus on Stage 1, where U and D bargain over w, anticipating its impact on the equilibrium quantity characterized in Lemma 1 (note that the alternative formulation of Stage 2 in Appendix B could also be considered).

OA2.1 Stage 1.3: Take-it-or-Leave-it Offers

OA2.1.1 Take-it-or-Leave-it Offer from U

Proceeding backward, we now solve Stage 1.3 by considering the case where Nature selects U to make the ultimate TIOLI to D.

Lemma 2 *When U makes a take-it-or-leave-it offer to D, the equilibrium wholesale price is:*

$$w_{\overline{\mu}} = MC_U(q_{\overline{\mu}}) \left(\frac{\varepsilon_{MR_D(q_{\overline{\mu}})}}{1 - \varepsilon_{MR_D(q_{\overline{\mu}})}} \right)$$

where $q_{\overline{\mu}} \equiv q(w_{\overline{\mu}})$ is the equilibrium quantity and $w_{\overline{\mu}} = MR_D(q_{\overline{\mu}}) > w_I$.

Proof. The maximization problem faced by U is given by:

$$w_{\overline{\mu}} \equiv \underset{w}{\text{argmax}}\ \pi_U(w)$$

where $\pi_U(w) = wq(w) - r(q(w))q(w)$ and $\frac{\partial \pi_U}{\partial w} = q(w) + q'(w)[w - MC_U(q)]$. Assume first that $w \leq w_I$. From Stage 2, this implies $w = MC_U(q)$, so that $\frac{\partial \pi_U}{\partial w} = q(w) > 0$. Hence, setting $w \leq w_I$ does not maximize U's profit. Instead, the optimal choice for U must satisfy $w > w_I$, which, as discussed in Section 4.1 of the main text, implies that $w(q) = MR_D(q)$. In this case, the first-order condition $\frac{\partial \pi_U}{\partial w} = 0$ yields:

$$w_{\overline{\mu}} = MC_U(q_{\overline{\mu}}) \left(\frac{\varepsilon_{MR_D}(q_{\overline{\mu}})}{\varepsilon_{MR_D}(q_{\overline{\mu}}) - 1} \right),$$

where $w_{\overline{\mu}} = MR_D(q_{\overline{\mu}}) > w_I$ and $\varepsilon_{MR_D}(q_{\overline{\mu}}) \equiv \frac{MR_D(q_{\overline{\mu}})}{q_{\overline{\mu}} |MR_D'(q_{\overline{\mu}})|}$. Note that the second-order condition for U's profit-maximization problem is $\frac{\partial^2 \pi_U}{\partial w^2} = q''(w)[w(q) - MC_U(q)] + q'(w)[2 - q'(w) MC_U'(q)] < 0$. As $w(q) = MR_D(q)$, $q'(w) = \frac{1}{MR_D'(q)} < 0$, and $q''(w) = \frac{\sigma_{MR_D}}{qMR_D'(q)^2}$, this second-order condition simplifies to:

$$\sigma_{MR_D}(q) < \frac{2 - \frac{MC_U'(q)}{MR_D'(q)}}{\left(1 - \frac{MC_U(q)}{MR_D(q)}\right) \epsilon_{MR_D}(q)} \tag{54}$$

where $\sigma_{MR_D}(q) \equiv \frac{qMR_D''(q)}{|MR_D'(q)|}$. From the first-order condition, we obtain $\left(1 - \frac{MC_U(q_{\overline{\mu}})}{MR_D(q_{\overline{\mu}})}\right) \epsilon_{MR_D}(q_{\overline{\mu}}) = 1$, implying that (54) simplifies to:

$$\sigma_{MR_D}(q_{\overline{\mu}}) < 2 - \frac{MC_U'(q_{\overline{\mu}})}{MR_D'(q_{\overline{\mu}})}.$$

As $MC_U'(q_{\overline{\mu}}) > 0$ and $MR_D'(q_{\overline{\mu}}) < 0$, it is straightforward that Assumption 2 in the main text, which stipulates that $\sigma_{MR_D} < 2$, is sufficient to ensure that the second-order condition for U's profit-maximization holds. ■

OA2.1.2 Take-it-or-Leave-it Offer from *D*

We now consider the case where Nature selects D to make the ultimate take-it-or-leave-it offer to U.

Lemma 3 *When D makes a take-it-or-leave-it offer to U, the equilibrium wholesale price is:*

$$w_{\overline{\nu}} = MR_D(q_{\overline{\nu}})\left(\frac{\varepsilon_{MC_U(q_{\overline{\nu}})}}{1+\varepsilon_{MC_U(q_{\overline{\nu}})}}\right)$$

where $q_{\overline{\nu}} \equiv q(w_{\overline{\nu}})$ is the equilibrium quantity and $w_{\overline{\nu}} = MC_U(q_{\overline{\nu}}) < w_I$.

Proof. The maximization problem faced by D is given by:

$$w_{\overline{\nu}} \equiv \underset{w}{\text{argmax}}\ \pi_D(w)$$

where $\pi_D = p(q(w))q(w) - wq(w)$ and $\frac{\partial \pi_D}{\partial w} = q'(w)[MR_D(q) - w] - q(w)$. Assume first that $w \geq w_I$. From Stage 2, this implies that $w = MR_D(q)$, so that $\frac{\partial \pi_D}{\partial w} = -q(w) < 0$. Hence, setting $w \geq w_I$ does not maximize D's profit. Instead, the optimal choice for D must satisfy $w < w_I$, which, as discussed in Section 4.1 of the main text, implies that $w(q) = MC_U(q)$. In this case, the first-order condition $\frac{\partial \pi_D}{\partial w} = 0$ yields:

$$w_{\overline{\nu}} = MR_D(q_{\overline{\nu}})\left(\frac{\varepsilon_{MC_U}(q_{\overline{\nu}})}{\varepsilon_{MC_U}(q_{\overline{\nu}})+1}\right) \tag{55}$$

where $\underline{w} = MC_U(\underline{q}) < w_I$ and $\varepsilon_{MC_U} \equiv \frac{MC_U(q)}{q|MC'_U(q)|}$. Note that the second-order condition for D's profit-maximization is $\frac{\partial^2 \pi_D(w)}{\partial w^2} = q''(w)[MR_D(q) - w(q)] + q'(w)[q'(w)MR'_D(q) - 2] < 0$. As $w(q) = MC_U(q)$, $q'(w) = \frac{1}{MC'_U(q)} > 0$, and $q''(w) = -\frac{\sigma_{MC_U}(q)}{qMC'_U(q)^2}$, this second-order condition simplifies to:

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

where $\sigma_{MC_U}(q) \equiv \frac{qMC''_U(q)}{|MC'_U(q)|}$. From the first-order condition, we obtain $\left(\frac{MR_D(q_{\overline{\nu}})}{MC_U(q_{\overline{\nu}})} - 1\right)\epsilon_{MC_U}(q_{\overline{\nu}}) = 1$, implying that (56) simplifies to:

$$\sigma_{MC_U}(q_{\overline{\nu}}) > \frac{MR'_D(q_{\overline{\nu}})}{MC'_U(q_{\overline{\nu}})} - 2.$$

As $MR'_D(q_{\overline{\nu}}) < 0$ and $MC'_U(q_{\overline{\nu}}) > 0$, it is straightforward that Assumption 2, which stipulates that $\sigma_{MC_U} > -2$, is sufficient to ensure that the second-order condition for D's profit-maximization holds.
■

OA2.2 Stage 1.1

Given that D can reject the offer and make a (ultimate) counter-offer with probability $1-\phi$, U's profit-maximization problem can be written as:

$$\max_w \pi_U(w) \quad \text{subject to} \quad \pi_D(w) \geq \phi\pi_D(w_{\overline{\mu}}) + (1-\phi)\pi_D(w_{\overline{\nu}}) \tag{57}$$

where $\pi_D(w_{\overline{\mu}})$ and $\pi_D(w_{\overline{\nu}})$ denote D's profit in the subgames where U and D, respectively, make the (ultimate) TIOLI offer. The constraint in (57) represents D's participation constraint stemming from its ability to reject U's offer and make a counter-offer with probability $1-\phi$. From (57), we obtain the main result of this section:

Proposition 8 *There exists a unique $\phi \in [0,1]$ such that the bargaining outcome of the non-cooperative game $w^{\star\star}$ replicates the Nash bargaining solution $w^{\star} \in [w_{\overline{\nu}}, w_{\overline{\mu}}]$,*

Proof. We first show that there exists a unique $w^{\star\star} \in [w_{\overline{\nu}}, w_{\overline{\mu}}]$ which (i) satisfies D's participation constraint, (ii) solves U's profit-maximization problem in (57). From Online Appendix OA2.1.1 and OA2.1.2, we know that:

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

As both firms prefer a higher w when $w < w_{\overline{\nu}}$ and a lower w when $w > w_{\overline{\mu}}$, it is straightforward that $w_{\overline{\nu}} \leq w^{\star\star} \leq w_{\overline{\mu}}$. As $\frac{\partial \pi_U(w)}{\partial w} > 0$ and $\frac{\partial \pi_D(w)}{\partial w} < 0$ for $w \in [w_{\overline{\nu}}, w_{\overline{\mu}}]$, it implies that any solution $w^{\star\star}$ to (57) must bind D's participation constraint, that is, $\pi_D(w^{\star\star}) = \phi\pi_D(w_{\overline{\mu}}) + (1-\phi)\pi_D(w_{\overline{\nu}})$. Note that any $w > w^{\star\star}$ would violate D's participation constraint and any $w < w^{\star\star}$ would lower U's profit.

We show now that, for a given $w^{\star\star} \in [w_{\overline{\nu}}, w_{\overline{\mu}}]$, there exists a unique $\phi \in [0,1]$. Defining $C(\phi) \equiv \phi\pi_D(w_{\overline{\mu}}) + (1-\phi)\pi_D(w_{\overline{\nu}})$, D's participation constraint can be rewritten as $\pi_D(w) \geq C(\phi)$. In equilibrium, D's participation constraint is binding, that is, $\pi_D(w^{\star\star}) = C(\phi)$. Hence, the value of ϕ determines $w^{\star\star}$. Specifically, as $C'(\phi) = \pi_D(w_{\overline{\mu}}) - \pi_D(w_{\overline{\nu}}) < 0$ and $\frac{\partial \pi_D(w)}{\partial w} < 0$ for $w > w_{\overline{\nu}}$, it follows that $w^{\star\star}$ is increasing in ϕ. When $\phi = 0$, D's participation constraint reduces to $C(0) = \pi_D(w_{\overline{\mu}})$, implying that $w^{\star\star} = w_{\overline{\mu}}$. Similarly, when $\phi = 1$, we have $C(1) = \pi_D(w_{\overline{\mu}})$, implying that $w^{\star\star} = w_{\overline{\mu}}$. Consequently, for any given Nash bargaining solution $w^{\star} \in [w_{\overline{\nu}}, w_{\overline{\mu}}]$, there exists a unique $\phi \in [0,1]$ such that $w^{\star\star} = w^{\star}$. ■

As Stage 2 remains unchanged, and because for any given ϕ the bargaining outcome $w^{\star\star}$ in the non-cooperative game replicates the Nash bargaining solution used in Section 4.2 of the main text, it turns out that the equilibrium outcome of the bilateral monopoly setting introduced in Section 4 of the main text coincides with the subgame perfect Nash equilibrium of the non-cooperative game developed in this Online Appendix.