The Competitive Effects of Vertical Integration Under Incomplete Hedging

Ali Hortaçsu Keisuke Ito Eric Richert *

July 13, 2026

Preliminary and Incomplete: Do Not Cite Without Permission

Abstract

Vertical integration is often thought to mitigate market power in electricity markets since integrated generators that also purchase power for retail customers have less incentive to raise wholesale prices. We show that vertical integration between generators and retailers can be anti-competitive if firms are risk averse and only have access to standard forward contracts that provide incomplete insurance. Because standard forward contracts hedge price fluctuations for a fixed quantity, risk-averse integrated firms have an incentive to hedge retail obligations by taking long positions. This hedging incentive can make integrated generators net suppliers in the spot market, softening competition. Using unit-hour level auction data, we estimate forward positions and find integrated firms indeed hold large long positions and bid less competitively than their generation-only rivals. Using our estimated model, we find that the introduction of more complete forward contracts, in the form of load-following contracts, reduces hedging by integrated firms and intensifies spot-market competition, lowering consumer costs and increasing total welfare.

JEL: L94 L22 D44 G32 L13 L42 D47

Keywords: Electricity market, vertical integration, forward contracts

*Preliminary and Incomplete. Do not cite without permission. Thank you to Gautam Gowrisankaran for a helpful discussion of the paper. We thank Koichiro Ito, Ryan Kellogg, Mar Reguant, Richard Chen and seminar participants at Cowles Models and Measurement, University of Chicago, Northwestern, Tex-Mex Conference, Berkeley Energy Camp for helpful comments. Ali Hortaçsu: University of Chicago and NBER, Email: Hortacsu@gmail.com; Keisuke Ito, University of Chicago, Email: keisuke0329ito@gmail.com ; Eric Richert, University of Chicago and NBER, Email: richert.eric@gmail.com

1 Introduction

In any market with uncertain supply and demand, both prices and quantities fluctuate, and someone must bear the risk of those fluctuations. The contracts and forward markets a regulator permits, the participants it allows to trade, the way production is divided among firms, and the vertical structure of the industry together determine firms' exposure to spot prices. That exposure plays two roles at once: it fixes who bears that risk, and it shapes firms' incentives to exercise market power. Market design choices about industry structure, participation, and forward markets arise in many markets, including natural gas, agricultural commodity, transportation, and electricity markets. In the restructuring of electricity markets, regulators have debated whether vertically integrated firms should be separated, whether and how financial forward markets should be regulated, and how far retail supply should be opened to competition. While a large literature studies how each design choice affects competition in isolation, there is much less empirical evidence on how the opportunities for hedging risk interact with market structure to shape competition.

This paper studies how vertical integration affects competition when firms use forward contracts for both risk hedging and strategic commitment. Vertical integration and forward contracting have each been the topic of long literatures in economics, but less attention has been paid to how their interaction changes firms' hedging and competitive incentives. Yet both vertical integration and forward contracting are observed in many wholesale markets, including electricity, agriculture, and airlines. When analyzed separately, forward contracts and vertical integration are each thought to increase competition by lowering wholesalers' exposure and incentives to raise spot prices. However, the forward position of an integrated firm may reflect hedging considerations driven by their retail demand in addition to wholesale pricing incentives, altering the firm's net exposure to spot prices. We analyze how forward positions, shaped by wholesale and retail considerations, determine the effect of vertical integration on competition.

We focus on deregulated electricity markets, where the interaction between vertical integration and forward contracting is especially important. In these markets, both vertical integration and forward contracts affect competition by changing the inframarginal quantity of a generator that is exposed to the spot price. When a generator sells forward contracts, it receives the contract price rather than the spot price on the contracted quantity, so forward obligations net against generation and reduce inframarginal exposure. Similarly, an inte-

grated firm earns the wholesale price on its generation but pays it on the electricity needed to serve retail demand, so retail obligations net against generation and reduce inframarginal exposure. With lower inframarginal quantity, raising the spot price produces smaller gains, and the incentive to charge markups decreases. Bushnell et al. (2008) formalize this argument for vertically integrated firms and show empirically that integrated firms bid more competitively in restructured markets. However, their analysis treats forward positions as given. In the early restructuring period they study, that was reasonable: vertical arrangements were largely predetermined and exchange-traded forward markets were thin (NGI Staff Reports (1999)). Today, many integrated firms actively choose their forward positions, and those choices reflect risk-management motives as well as wholesale pricing incentives. These risk-management motives are central in deregulated electricity markets, where crises in California and Texas bankrupted retailers exposed to volatile wholesale prices and revealed the importance of wholesale market power. We show that this interaction matters because the same forward contracts used to hedge retail risk also change an integrated firm's net exposure to spot prices, and therefore its incentive to exercise market power.

Once forward positions adjust endogenously, we show that the competitive effects of vertical integration depend on the form of forward contracts available. In most markets, fixed-quantity swaps are the dominant forward contracts. A standard swap pays the difference between the spot price and a pre-specified base price on a fixed quantity. These contracts are incomplete for an electricity retailer because the retailer's wholesale procurement cost varies with realized demand, while swap payments are tied to a fixed contract quantity. The retailer therefore remains exposed to right-tail asymmetric risk: when demand is high, prices rise sharply and unhedged losses are large, while low-demand states generate smaller offsetting gains because prices fall less dramatically. A risk-averse integrated firm responds by over-hedging its retail position, buying swaps for more than expected demand. This provides additional insurance in high-price, high-demand states, but shifts the firm's net spot-price exposure toward a longer position on average. As a result, the integrated firm has stronger incentives to raise spot prices, weakening competition.

We develop a theoretical model that characterizes the optimal swap positions of vertically integrated firms. The model shows that incomplete, fixed-quantity swaps can make vertical integration less competitive when risk-management motives induce over-hedging. It also shows that this over-hedging incentive disappears when more flexible "load-following"

contracts, which index contracted quantities to realized demand, are introduced. We provide empirical evidence consistent with these predictions using data from before and after the introduction of load-following contracts in the electricity market in Alberta, Canada.

Vertical integration between generators and retailers is common in many deregulated electricity markets, reflecting both the legacy of integrated monopolies and retailers' incentives to hedge risk by securing reliable long-term access to generation. In theory, vertical integration can also be anti-competitive due to foreclosure of rivals or increasing their operating and entry costs (Joskow (2008); Rey and Tirole (2007)). In restructured electricity markets, however, opportunities for vertical foreclosure are limited because wholesale dispatch and network access are coordinated by independent system operators, and strict regulatory oversight ensures nondiscriminatory access to the network (Mansur (2007); Bushnell et al. (2008)). The effect of vertical integration that we focus on is distinct from double marginalization, which removes the margin on an internally supplied input and lowers the retail price. Since retail markets are competitive we instead focus on how the integrated firm's net position shapes its incentive to raise the spot price.

In isolation, both vertical integration and forward positions are usually viewed as pro-competitive.[1] An integrated firm earns the wholesale price on its generation but pays that price on the electricity needed to serve retail demand, reducing its incentives to raise wholesale prices (Bushnell et al. (2008)). This logic has become influential both in subsequent empirical work and in policy analysis, where vertically integrated firms are often treated as more competitive once retail obligations are netted out of their wholesale positions.[2] In standard oligopoly models, short forward positions also intensify wholesale-market competition. Since forward positions reduce a producer's exposure to the spot price, each firm has a unilateral incentive to sell forward, committing it to greater output and forcing rivals to cede market share. In equilibrium, all firms contract forward and total output exceeds the Cournot level (Allaz and Vila (1993)). Although the magnitudes differ, the pro-competitive impact holds across a wide range of market structures (Miller and Podwol (2020), and regardless of whether standard swaps or load following futures contracts are used Brown and

[1] One exception to this is Gans and Wolak (2007) who document anti-competitive effects of vertical integration in the Australian electricity market. However, their explanation focuses on the supply of contracts and assumes no quantity risk, ruling out the channel we highlight.

[2] For example, see Brown et al. (2023); Fabra and Reguant (2014); Mansur (2007); Lundin and Tangerås (2020); Simshauser (2020); Bosco et al. (2013); NERA Economic Consulting (2025); Frontier Economics (2017).

Sappington (2023)).[3]

The natural conclusion of these benchmark models is that both vertical integration and forward contracting are pro-competitive. We show that this prediction need not hold if retailers are risk averse. In electricity markets, retailers' losses are concentrated in states where wholesale prices and retail demand are jointly high. These states threaten solvency rather than simply reducing profits. The risk of retailer insolvency reflects a standard corporate-finance rationale for hedging: when adverse states are disproportionately costly, hedging creates firm value.[4] Therefore, firms may choose forward positions not only to commit strategically in the wholesale market, but also to hedge retail risk. This additional motive can change net spot-price exposure, undoing the pro-competitive benefits of vertical integration.

We begin by separately analyzing simple models of the two main forces that drive forward contracting. Strategic commitment incentives push the integrated firm toward shorter, more negative net positions that intensify spot-market competition. In contrast, risk-hedging incentives push the integrated firm to over-hedge its downstream retail obligations by taking long forward positions, which soften spot-market competition. Over-hedging arises because the hedge available is an incomplete form of insurance.

Figure 1 provides a graphical illustration of the hedging problem. The retailer's procurement cost is the spot price $p^s(\varepsilon)$ multiplied by realized demand $q^r(\varepsilon)$. For simplicity, suppose the retailer's demand is $q^r(\varepsilon) = \mathbb{E}[q^r(\varepsilon)] + \beta\varepsilon$, where $\beta\varepsilon$ represents their share of the market-level demand shock. The retailer receives payment $p^s(\varepsilon)f$ from forward contracts with fixed quantity f. If the retailer sets $f = \mathbb{E}[q^r(\varepsilon)]$ to hedge their expected demand, they can offset the variation in $-\Delta\pi = -p^s(\varepsilon)\,\mathbb{E}[q^r]$. However, this leaves the retailer exposed to the unhedged interaction term $-p^s(\varepsilon) \cdot \beta\varepsilon$. This interaction term creates asymmetric

[3] Brown and Sappington (2023) contrast the effects of load-following and swap contracts for risk-neutral generation-only firms, focusing on the strategic role similar to Allaz and Vila (1993). However, they model load following contracts that cover the quantities produced by strategic firms, excluding the competitive fringe, whereas in practice these contracts cover retailer demand. Their assumption means contract volumes are a function of the firms' production, and as a result the load following contracts are more effective for reducing output than the standard swap.

[4] Smith and Stulz (1985) show that hedging is valuable when financial distress costs are convex. Stulz (1990) and Bessembinder (1991) model settings in which low-cash-flow states induce underinvestment due to monitoring and contracting frictions or debt overhang near default.Froot et al. (1993) show that when external finance is costly, hedging reallocates internal funds toward low-cash-flow states. Consistent with these motives, electricity retailer hedging has been extensively studied (Willems and Morbee, 2010; Oum and Oren, 2009; Zhou et al., 2017; Winzer et al., 2024; Flottmann et al., 2025).

risk. For example, with convex supply curves, When demand is high ($\varepsilon \gg 0$), spot prices rise sharply, creating large losses $-p^s(\varepsilon) \cdot \beta\varepsilon$ on the unhedged quantity. When demand is low ($\varepsilon \ll 0$), prices fall less dramatically, yielding smaller gains. The retailer can reduce this asymmetric exposure by over-hedging: setting $f^* > \mathbb{E}[q^r(\varepsilon)]$ generates additional forward payments $\Delta\pi^{\text{over}}$ precisely when demand and prices are both high, providing insurance against the costliest states.

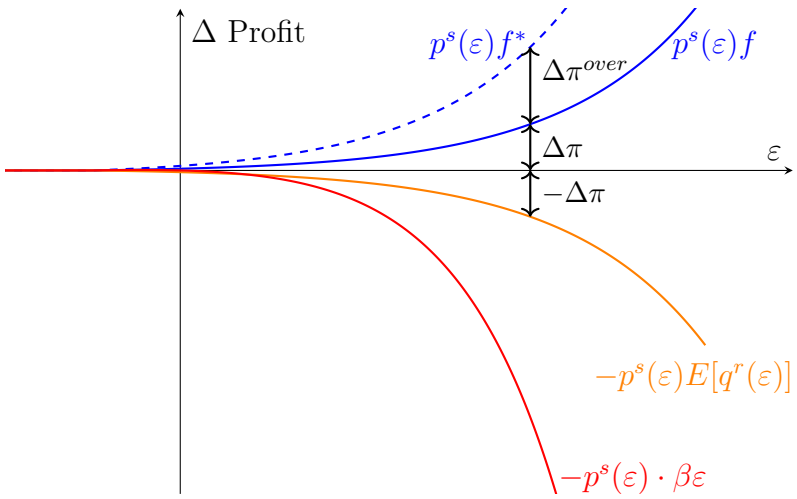


Figure 1: Asymmetric Shock Exposure and Over-hedging of Forward Positions

Notes: The figure illustrates how a demand shock ε affects a retailer's exposure to procurement costs and forward payoffs. The vertical distance of each curve from the horizontal axis represents the effect of a given realization of ε on profits. In the absence of higher-order effects (i.e., without the interaction term $-p^s(\varepsilon) \cdot \beta\varepsilon$), the retailer can fully hedge procurement risk by setting forward positions equal to expected demand, $f = \mathbb{E}[q^r(\varepsilon)]$. In this case, the linear components of profit offset exactly across states. When higher-order effects are present, exposure becomes asymmetric: losses in high-demand states are amplified relative to gains in low-demand states. As a result, the variance-minimizing forward position exceeds expected demand, $f^* > \mathbb{E}[q^r(\varepsilon)]$, generating endogenous over-hedging. The figure depicts the case of convex spot prices. The same intuition applies under symmetric price functions when the demand shock distribution is skewed.

A similar hedging incentive appears in agriculture markets (e.g., McKinnon (1967),Rolfo (1980), Hirshleifer (1988)). Farmers face market level output shocks that are correlated with their own production, for example favorable weather conditions boost everyone's yield simultaneously. These shocks induce a negative correlation between price and output through the downward sloping demand curve which creates an incentive for farmers to under-hedge their production. In electricity markets, Borenstein (2007) is, to our knowledge, the first paper to highlight this buyer-side hedging incentives. He shows that when an electricity buyer's load is positively correlated with spot prices, fixed-quantity forward contracts can

make it optimal to hedge more than expected demand. [5] Conceptually, we show that the buyer-side over-hedging motive emphasized by Borenstein (2007) can overturn the standard pro-competitive benchmark for vertical integration in Bushnell et al. (2008) once forward positions are treated as endogenous.

A natural question is why perfect insurance does not exist. In electricity markets, complete load-following contracts are largely absent from centralized trading.[6] They may be negotiated bilaterally, but they are not standardized or liquid, and an industry report notes that such products are not traded in any North American market (Alberta Utilities Commission (2018)). A key barrier is asymmetric information. Delivery must condition on the buyer's realized retail demand, which is private, limiting wide trade. Practical frictions reinforce this. Load-following contracts receive unfavorable accounting treatment, and risk sharing requires counterparties (or intermediaries) willing to absorb correlation risk, credible information-sharing to verify demand, and strong default safeguards such as collateral. These costs make centralized market design costly and keep trading thin.

We use our model to derive theoretical predictions on the equilibrium forward contract positions in four different scenarios that vary the type of contracts available and whether firms face strategic or risk-hedging incentives. We find that with strategic incentives only, both integrated and generation-only firms have similar (short) forward positions. Under risk aversion, the generation only firm takes a short position but the integrated firm over-hedges their retail demand. With load following contracts in either strategic or risk-hedging case, both integrated and generation firms take similar short forward positions.

Guided by these predictions, we study the five largest firms in the Alberta electricity market: four generation-only firms and one vertically integrated firm using unit-hour auction data. Several features make Alberta a well-suited environment for this analysis. The wholesale market clears through a single hourly uniform-price auction, generators do not receive any capacity payments, and an active exchange-traded forward market operates alongside the wholesale market. We first measure spot-market competition directly, computing markups as the wedge between bid prices and engineering marginal costs. Consistent with our theory,

[5]Related hedging motives arising from skewness in the distribution of equilibrium prices have also been documented in Bessembinder and Lemmon (2002), who analyze a period where forward contracts were much less widely traded, so the paper focuses on imperfect arbitrage rather than imperfect spot-market competition.

[6]In industry and regulatory practice, these contracts are also sometimes referred to as full-requirements or full-load contracts.

we find that prior to the introduction of load-following contracts, the vertically integrated firm exhibited the highest markups. Following the introduction, markups declined for the vertically integrated firm but not for generation-only firms. To isolate the role of forward contracts in driving this change in markups, we estimate forward positions, controlling for confounding factors such as generation portfolio composition and demand elasticity. Under the standard benchmark, netting out retail obligations would imply that the vertically integrated firm has the shortest net spot exposure. In the data, however, the vertically integrated firm bids least competitively and holds the longest forward position, consistent with the model's risk-hedging mechanism. After load-following contracts are introduced, the firm's long hedging position declines substantially while those of generation-only firms remain stable, and its bidding behavior becomes markedly more competitive, consistent with the model's predictions.

Using the model, we quantify the welfare implications of the interaction between vertical structure and risk hedging. When only standard (price-only) forward contracts are available, vertical integration increases average wholesale prices by 4.05 CAD/MWh (approximately 5 percent), raising consumer costs by approximately 332 million CAD per year and generating total welfare losses of 39 CAD million per year. However, introducing more flexible contracts reduces wholesale prices by 2.8 CAD/MWh, lowering consumer costs by 233 million CAD per year and producing total welfare gains of 35 million CAD per year. These results indicate that the adverse effects arise from the interaction between vertical integration and contract incompleteness. This implies that vertical integration is not, by itself, the sole source of the welfare losses. Instead, the losses reflect limitations in the contracts available in forward markets and can be mitigated by increasing contract completeness.

2 Institution and Data

We focus on Alberta, which has several attractive features for our analysis. The market is quite concentrated, there is a single vertically integrated firm among the dominant generators, bidding data are widely available, and we observe the market before and after the introduction of novel forward contracts that provide better protection against quantity risk. We combine several datasets on Alberta's electricity market from 2018–2021. First, we use data from the Alberta Electric System Operator (AESO) on hourly price and quantity bids

for all generation units, asset ownership, and realized market demand. We also use AESO metered-volume data to observe the vertically integrated firm's hourly retail load. We combine this with data from Brown et al. (2023) on the marginal costs of generation units. These are estimates of the hour-specific marginal costs for each plant, constructed using information on gas prices, taxes, and heat rates. Finally, we leverage forward contract pricing information from Olmstead and Yatchew (2025).

2.1 Electricity Spot Market

The wholesale electricity market in Alberta is a single real-time hourly uniform-price multi-unit auction. There is an ancillary services market to match short term fluctuations but there are no day-ahead or capacity markets. Generators submit bids in the hourly-auction as step-functions. That is, each generator provides a set of up to seven price, quantity pairs at which they are willing to supply electricity. The price they quote must be in $[0,999.99]/MWh and the generators must provide quotes for all available capacity (i.e., the largest quantity must equal the capacity). The AESO sorts the bids by price to construct an aggregate demand curve. In each minute of the hour, there is a system-marginal price, the most expensive bid that is used to clear the market. All bids with prices below this market clearing price are dispatched. Participants are paid the pool price, the time-weighted average of the system marginal price within the hour for all units dispatched. The generation market in Alberta has no location marginal prices. The AESO plans transmission investments according to a no-congestion standard, ensuring sufficient capacity to provide reasonable access to all in-merit generation (AESO (2022); Wolak (2015)). Consistent with this policy, the Market Surveillance Administrator reports that constrained-down generation accounts for only 0.2% of total load, with congestion arising primarily from discrete planned or forced transmission outages rather than insufficient baseline capacity (MSA (2014)).

The market is quite concentrated, with four large firms making up 48% of the generation capacity (see Table 1). These large firms are TransAlta, ATCO, ENMAX, and Capital Power. There is also a competitive fringe consisting of 38 smaller firms MSA (2018). Finally, there are bids from the balancing pool, a government entity which manages assets that were subject to divestiture at the time the electricity market underwent restructuring Brown et al. (2023).[7] Alberta has limited interconnection with other Canadian provinces and the United

[7]Around 2000, generation assets were divested via power purchase agreement (PPA) auctions to reduce

States. The total transmission import capacity is around 1100MW, and net imports account for only about 5% of Alberta's demand AESO (2022).

Table 1: Summary Statistics

Variable	Mean	Median	SD	Min	Max
Panel A: Hourly Market Data					
Pool price (CAD/MWh)	63.46	35.07	106.74	0.00	999.99
Total dispatched (MWh)	8340	8369	829	5855	16506
Gas dispatched	4812	4754	631	2982	9366
Coal dispatched	2609	2597	788	866	6168
Wind dispatched	583	489	448	0	3299
Hydro dispatched	176	147	102	34	636
Solar dispatched	17	0	47	0	302
Other	143	147	24	52	338
Panel B: Monthly Retail Consumption					
Total (GWh)	3964.36	3936.31	241.06	3540.89	4472.95
Vertically integrated firm	1346.47	1380.16	152.15	1046.43	1625.39
Panel C: Market Share					
3 Generation-only large firms		39% of total capacity			
Vertically integrated firm		9% of total capacity			

Notes: Sample from 2018 to 2021.

Alberta has a competitive retail market, where firms offer both fixed or variable rate contracts. The variable rate options make up a small share of consumer demand.[8] There is also a default option available to all consumers, at a price determined by the regulator. This option serves roughly 40-50% percent of households. During the period of the study, the default option had a variable (monthly) price, depending on forward market and wholesale electricity prices based on a formula determined by the regulator MSA (2015) meant to capture costs. This limits the ability of firms to charge markups on their other plans. Ap-

concentration. The auctions offered 6250MW of coal and gas capacity (roughly 64% of the market at the time). The unsold PPA capacity from these auctions (roughly 2000MW) was transferred to the balancing pool, and the government managed these assets Brown et al. (2023).

[8]If retail prices are equal to spot prices, then the risk is transferred from the retailer to households. Partial pass-through occurs in some markets in Europe: e.g., PVPC prices in Spain are based on a combination of spot and futures prices and in EDF Tempo France uses a coarse set of consumer prices based on spot market prices. This reduces risk and may affect retailers' hedging.

proximately 80% of retail electricity demand is non-residential with 61% from large industrial consumers MSA (2015).

2.2 Forward Markets

There is an active financial market for electricity futures in Alberta. The main trading platforms are the Natural Gas Exchange (NGX) and over-the-counter broker markets (MSA, 2010). On NGX, contracts are quoted and traded anonymously, and NGX acts as a clearinghouse that assumes counterparty risk. As with many financial exchanges, this is possible because NGX requires participants to post margin. Since these contracts are standardized and centrally cleared, participants can easily offset a forward position by taking on a new position in the opposite direction.

The most commonly traded forward contract is a spot-forward. These are also referred to as contracts for difference. To settle the contract, one party pays the other the difference between the agreed upon price, and the hourly pool price for the contract period. There are both flat forwards (covering every hour, each day in the contract period), as well as on-peak, off-peak, extended peak, extended off-peak and super peak contracts which provide some variation in the set of covered hours. Monthly, quarterly, and yearly flat contracts are most common. Monthly contracts account for 40% of the market, quarterly contracts for approximately 25%, and annual contracts for the remaining 35%. Participants in the forward market include not only generators and retailers, but also financial intermediaries and hedge funds. Financial participants and electricity firms each comprise roughly equal shares of total market activity MSA (2010).[9] These features, especially the exchange trading, and high participation of financial participants suggest that the forward market is competitive.

2.2.1 Testing for Arbitrage

Since our theoretical model assumes no arbitrage between the spot and forward markets, we begin by empirically verifying that this condition holds in the data. Following Longstaff and Wang (2004), we test whether the mean of the difference between spot and forward prices is zero. A histogram of the price differences is presented in Appendix Figure 7. The average

[9] Jha and Wolak (2023) provide evidence that introducing purely financial participation in California's wholesale electricity market improved agreement between day-ahead and real-time prices and reduced production costs, suggesting that financial participants can play an important role in improving forward market efficiency.

difference is -2.22, but the (heteroskedasticity and autocorrelation consistent) standard error of the estimate is 2.87, yielding a p-value of 0.44. Olmstead and Yatchew (2025) provide a different test of the unbiasedness hypothesis. If p^f_{tM} is the forward price at time t for delivery in month M, and p^s_M the spot price in month M, and unbiasedness requires $\beta_0 = 0, \beta_1 = 1$ in the regression $p^s_M = \beta_0 + \beta_1(\frac{1}{T}\sum p^f_{t,M}) + \epsilon_M$. They do not reject the null on the full sample: $\hat{\beta}_0 = -0.02\,(6.41)$ and $\hat{\beta}_1 = 1.02\,(0.09)$, yielding a p-value of 0.855, suggesting unbiasedness, which implies no arbitrage. However, in some sub-samples far before the contract is delivered they find evidence of larger slopes, and negative intercepts. This does not imply a violation of arbitrage, since it may capture price discovery, if prices far ahead are compressed towards the unconditional average, because relevant information has not yet arrived.

2.2.2 Load-Following Contracts

Part way through the sample period, the regulator authorized the introduction of a new type of forward contract into the regulated procurement process. This new contract pays the spot price minus the contracted price (like the contract for differences) but this time instead of being a contract for a fixed quantity, this price gap is paid on the realized retail demand of the firm on which the contract is written. This change was motivated by a desire to replace the administratively determined risk compensation component of the regulated retail price with a competitively determined, market-based equivalent. Specifically, any consumer not on a competitive electricity plan is served by the regulated option. Prices for this are set at a procurement cost, reflecting forward market prices, plus a small markup to the retailer for their servicing costs. In each region of the province one retailer acts as the provider of the regulated option plan. The government imposes many requirements on this plan, setting prices, and forcing the retailers to follow an agreed upon hedging strategy in their Energy Price Setting Plan (EPSP) which must be approved by the Alberta Utilities Commission (AUC) through regulatory proceedings approximately every three years.

The retailer developed the contract form and auction design with outside consultants, but the regulatory framework governing the process, the adversarial public proceeding, and the regulator's ongoing oversight of auction integrity were all essential to its adoption. Initially this move was lead by EPCOR, another large retailer (which is not vertically integrated). EPCOR transitioned to using load following futures, sold at auctions in 2018. The vertically integrated retailer followed, with the first auction held in September of 2020 for delivery in

December (see e.g., Brown et al. (2022)). As in the spot forward market, financial traders play a substantial role in the load following auctions submitting around 40% of all bids for load following contracts Brown et al. (2022).[10]

There were several frictions that made the adoption of these contracts difficult, and which may explain why despite success in centralized procurement auctions they did not arise alone in the market. First, these contracts transfer large financial risks onto the financial players (since prices and quantities are high on the same day) requiring larger margin/collateral be kept by the market maker or inducing substantial counterparty risk if contracted bilaterally NERA Economic Consulting (2019). Second, these products transfer larger risks to the firm and that may change the set of firms interested in supplying the product. Third, these products contract on the retail demand of a firm. This information is the private information of the firm and may be subject to manipulation without some trusted verification system NERA Economic Consulting (2019). Fourth, these products require good data and different modeling ability and may be complicated to price (e.g., Beblow (2017)). Fifth, because the market is still not liquid unlike a standard forward contract these contracts are difficult to unwind (one cannot simply go to the exchange and purchase an offsetting position in the other direction).

Finally, another friction that may have limited the adoption of these contracts is accounting treatment.[11] As a result of accounting treatment, the LFFC's fair value changes flow

[10]The role of financial traders in both forward markets may seem surprising given that generators may impact the settlement price by their behavior in the spot market. However, it is consistent with arguments in Allaz and Vila (1993) where traders play an active role, despite the fact that generators strategically manipulate quantities to impact the spot market price.

[11]Under accounting rules, including ASC 815 in the United States and IFRS 9 or ASPE Section 3856 in Canada, derivative contracts must be marked to market at fair value KPMG (2022). A power contract may qualify for an exemption from fair-value accounting if it is physically settled, whereas a financially settled contract generally does not. Under the "normal purchases and normal sales" election in ASC 815, the IFRS "own use" scope exemption, or the analogous ASPE exemption, a contract may be exempt from fair-value treatment if it results in actual delivery KPMG (2024); BDO Canada (2024). Since the retail sales a retailer would be hedging sit on accrual, the LFFC creates a mismatch that introduces earnings volatility. A standard swap is also a financial contract and is denied the physical exemption, but since its notional quantity is fixed, it can be designated against a specific set of forecast purchases and qualifies for hedge accounting. This allows gains and losses to be deferred until the offsetting transactions are recognized, rather than flowing immediately through net income. Because the notional volume of an LFFC varies with realized load, applying hedge accounting may be difficult. The retailer must designate sufficiently specific and highly probable forecast purchases in terms of quantity and timing, rather than simply designating a percentage of whatever load is realized. Consequently, some or all of the LFFC may not qualify for hedge accounting. PwC (2022, 2026); BDO Canada (2023).

directly through net income, while a standard swap's are deferred until the hedged purchases are recognized. This means that LFFCs induce substantial volatility in earnings. While the accounting rules apply in Alberta, the retailers that purchase LFFCs are not publicly traded and so may be less sensitive to earnings volatility than a typical retail company. Instead, their primary concern is appealing to debt-holders, who primarily evaluate their ability to repay using measures like adjusted EBITDA or funds from operations, that are less sensitive to the earnings volatility induced by LFFCs than measures like earnings per share, which are important to typical equity investors. This does not eliminate the relevance of such volatility, but it may make it less important for these retailers than it would be for firms whose equity investors closely monitor reported earnings.

Similar demand-shaped forward contracts have also been proposed for long-term resource adequacy in markets with significant intermittent renewables, for their ability to discourage the exercise of unilateral market power when realized demand exceeds forecasts (Wolak (2022); Thurber et al. (2022); Shaffer and Wolak (2024)). In addition, in several electricity markets including Pennsylvania, Ohio and New Jersey contracts variously called full-requirements, full-load, or provider of last resort are sold via regulator-run auctions to procure energy for consumers on a regulated default plan. In the auction, the bidders compete on retail prices, at which they agree to serve a share of retail demand. These contracts allow generators to lock in prices for a fixed share of realized retail demand (transferring quantity risk to the generator), similar to the sale of a LFFC. However, these contracts require physical delivery and often transfer additional risks, such as capacity-market and transmission costs, to the supplier. Moreover, because they are designed solely to procure default-service supply, rather than as symmetric financial instruments, retailers cannot use them to take flexible long or short hedging positions (see LaCasse and Wininger (2007), de Castro et al. (2008), Loxley and Salant (2004), Dormady et al. (2025)).

3 Simple Model

This section presents a simple model of optimal forward positions in swaps and load-following forward contracts (LFFCs) for producers and vertically integrated (VI) firms. The goal is to build intuition by isolating the main economic forces—strategic incentives and risk-based hedging—cleanly and separately. We combine both incentives in a richer quantitative model

in Section 5. We present four results, all proofs are in Appendix A. The first and second results solve for the optimal behavior under only strategic and risk-based hedging respectively in a market with swap contracts only. The third and fourth results repeat the same exercise, with load following contracts instead of swaps.

First, we analyze a model with only strategic behavior (i.e., risk neutral firms and imperfect competition in the spot market). Producers choose forward positions that commit themselves to larger spot-market quantities (as in Allaz and Vila (1993)). The integrated firm uses forward contracts to replicate the production-only firm's net spot-price exposure, (but at a different contracted quantity, since its retail obligations already offset part of its production exposure).

Then we analyze a model with only hedging incentives (i.e., risk averse firms, and a competitive spot market). The production-only firm takes a forward position equal to their expected production plus a small additional hedging term. On the other hand, the integrated firm hedges their retail position. Optimal hedging may differ from expected retail demand because retail profits are spot prices (which depend on the demand shock) times the demand shock, while swap forward payouts are spot prices times a constant, so swap forwards cannot perfectly insure retail profits. If demand uncertainty is right-skewed, or if prices are convex in demand, then the integrated firm can increase their payments in the high-price states where they would otherwise face large losses, by choosing a forward position that covers more than their expected retail exposure.

Finally, we analyze both models again, incorporating a flexible forward contract that adjusts to retail demand. In the model with strategic behavior, introducing load following contracts causes a small change in positions. However, in the model with risk hedging, the introduction of load following contracts reduces the short position of production-only firms, so that the integrated firm and production only firm have similar exposure to spot-market prices.

3.1 Model Setup

We consider a two period model with two strategic players, a generation-only firm (G) and a VI firm (V). In the first period, players trade forward contracts based on the prices of electricity in the second period. In the second period, players choose production quantities.

In period 1, agents face uncertainty over the residual demand state, which is resolved in

period 2. We assume forward markets are efficient: the perfect arbitrage condition holds, so expected forward profits are zero (a single risk-neutral speculator would deliver this). Retail demand is exogenously determined. The VI firm serves exogenous retail demand $q_V^r = \theta + e$, where θ is average retail demand and e is a mean-zero retail demand shock. To simplify, we assume the retail demand shock is proportional to the system demand shock: $e = \beta\varepsilon$. However, in Appendix A we prove more general forms of these results without this assumption. Following Bushnell et al. (2008), we assume that expected retail profits are zero, with retail prices p_i^r adjusted to reflect expected wholesale prices and retail demand.[12]

Throughout this section, f_i denotes firm i's forward position. Swap forward contracts pay $(p^f - p^s)f_i$, where p^s and p^f are the spot and forward prices. Therefore $f_i > 0$ corresponds to forward sales (a short position for which the firm receives a fixed price in advance and pays the spot price) and $f_i < 0$ to forward purchases (a long position for which the firm pays a fixed price in advance to receive the spot price). Subscripts "N" and "A" denote the risk-neutral and risk-averse cases, respectively.

3.2 Strategic Incentives and Swap Forward Contracts

Spot Market Equilibrium. In the second stage, firm i chooses production q_i in the spot-market to maximize profits:

$$\Pi_i = \underbrace{p_i^r q_i^r - p^s q_i^r}_{\text{Retail Profit}} + \underbrace{p^s q_i - C_i(q_i)}_{\text{Generation Profit}} + \underbrace{p^f f_i - p^s f_i}_{\text{Forward Profit}}$$

where $C_i(q_i) = b_i q_i + \frac{q_i^2}{2k_i}$ denotes total costs.[13] Following Allaz and Vila (1993) and Bushnell et al. (2008), retail quantities (q_i^r) and prices (p_i^r), and forward quantities (f_i) and prices (p^f), are sunk at the time production decisions are made and therefore drop out of the optimization over output. The maximization problem thus reduces to:

$$\max_{q_i} \; -p^s q_i^r + p^s q_i - C_i(q_i) - p^s f_i$$

[12] In Appendix B we show how retail rates relate to expected spot, swap and load following contract prices.

[13] A generation-only firm corresponds to the special case $q_i^r = 0$.

The first-order condition is:

$$\underbrace{p^s - C_i'(q_i)}_{\text{Markup}} = \underbrace{(q_i - f_i - q_i^r)}_{\text{Net position}} \cdot \left(-\frac{\partial p^s}{\partial q_i}\right)$$

This condition illustrates the standard argument for why forward contracts and vertical integration are pro-competitive. A firm's incentive to raise the spot price is determined by its net spot exposure: the smaller this exposure, the smaller the markup. Forward sales and retail obligations both reduce net spot exposure, and may even drive it negative, in which case the firm would prefer to push spot prices *below* competitive levels. The key question is therefore how forward positions are determined in equilibrium.

Residual demand in the second-stage production game is:

$$p^s(Q, \varepsilon) = a + \varepsilon - Q \quad \text{and} \quad Q(p^s, \varepsilon) = a + \varepsilon - p^s$$

where Q denotes aggregate system-wide quantity, a is average system-wide demand, and ε is a mean-zero random demand shock with variance σ^2. Demand is inelastic, but the duopolists face a downward-sloping residual demand because fringe suppliers produce as long as price exceeds their marginal costs.

Forward market equilibrium. In the first stage, both firms take into account both periods. By backward induction, equilibrium quantities q_G^*, q_V^* and the spot price p^s at the production stage are functions of f_G and f_V. With swaps, the contract quantities f_G and f_V are non-random, and the first-stage problems are:

$$\text{Generator: } \max_{f_{G,N}} \mathbb{E}\Big[p^s q_G^* - C_G(q_G^*) + (p^f - p^s) f_{G,N}\Big]$$

$$\text{VI: } \max_{f_{2,N}} \mathbb{E}\Big[p_V^r q_V^r - p^s q_V^r + p^s q_V^* - C_V(q_V^*) + (p^f - p^s) f_{V,N}\Big]$$

For technical simplicity, consider the case of a symmetric cost technology (i.e., $b_G = b_V = b$, $k_G = k_V = k$), with $0 < b < a$ and $k > 0$.

Result 1 (Strategic Incentives with Swap Contracts)**.** *The equilibrium swap forward positions are:*

$$f_{G,N}^* = \frac{k^2}{5k^2 + 5k + 1}(a - b), \qquad f_{V,N}^* = f_{G,N}^* - \theta$$

Allaz and Vila (1993) (Proposition 2.3) is a special case of this result, which imposes constant marginal cost ($k \to \infty$) and generation-only firms ($q_V^r = 0$).

Since the two firms make strategic commitments to earn markups on production, and since they share the same cost function, the equilibrium net positions are identical in expectation across the generator and the VI firm. To obtain the same position, the VI firm reduces its forward position relative to the generation-only firm by exactly its expected retail quantity (i.e., $f_{G,N}^* = f_{V,N}^* + \mathbb{E}[q_V^r] > 0$). Therefore, the forward market continues to deliver the improvement in competition over the one-period Cournot game highlighted by Allaz and Vila (1993). The result also tells us that conditional on equilibrium forward market behavior, vertical integration has no effect on the level of spot market competition (net pre-committed quantities match those implied by the generator-only game). As a result, VI is neither pro nor anti-competitive in this setting.

3.3 Hedging Incentives and Swap Forward Contracts

We now consider firms that use forward contracts solely as hedging instruments. To eliminate strategic commitment, we adopt the extreme assumption that both firms are price takers and that the spot price clears the following exogenous demand:[14]

$$Q(\varepsilon) = a + \varepsilon.$$

Under the price-taking assumption ($\frac{\partial p^s}{\partial q_i} = 0$), the first-order condition at the production stage yields:

$$\frac{\partial \Pi_i}{\partial q_i} = p^s - C_i'(q_i) = 0 \iff q_i = k(p^s - b) \quad \text{for } i = G, V$$

Equating total production to total demand gives $a+\varepsilon = 2k(p^s - b)$, so the market-clearing spot price and equilibrium quantities are:

$$p^s = \frac{a+\varepsilon}{2k} + b, \qquad q_G^* = \frac{a+\varepsilon}{2}, \qquad q_V^* = \frac{a+\varepsilon}{2}$$

We assume the risk averse firm has mean-variance utility $\mathbb{E}[\Pi_i] - \lambda_i \text{Var}(\Pi_i)$. This provides

[14] While it is an important simplification of wholesale electricity markets, this assumption is standard in the finance literature on risk hedging (see Bessembinder and Lemmon (2002)) and greatly simplifies the analysis of how hedging incentives determine forward positions.

a tractable approximation to CARA utility, which we use in the empirical model. Expected profits $\mathbb{E}[\Pi_i]$ do not depend on f_i, since production-stage outcomes q_G^*, q_V^*, p^s are independent of f_G and f_V. As a result, the firms' sole concern is risk reduction, and they minimize profit variance:

$$\text{Generator: } \min_{f_{G,A}} \text{Var}\Big[p^s q_G^* - C_G(q_G^*) + (p^f - p^s) f_{G,A}\Big]$$

$$\text{VI: } \min_{f_{V,A}} \text{Var}\Big[p_V^r q_V^r - p^s q_V^r + p^s q_V^* - C_V(q_V^*) + (p^f - p^s) f_{V,A}\Big]$$

Result 2 (Hedging Incentives with Swap Contracts)**.** *The equilibrium swap forward positions are:*

$$f_{G,A}^* = \mathbb{E}[q_G^*] + \underbrace{\frac{\mathbb{E}[\varepsilon^3]}{4\,\mathbb{E}[\varepsilon^2]}}_{\textit{Incomplete generation hedge}}, \qquad f_{V,A}^* \approx f_{G,A}^* - \theta - \underbrace{\frac{\mathbb{E}[\varepsilon^2 \cdot \beta\varepsilon]}{\mathbb{E}[\varepsilon^2]}}_{\textit{Incomplete retail hedge}}$$

The generation-only firm chooses a hedge larger than its expected generation. The firm's profit depends not only on its net spot quantity but also on its generation costs. Because those costs are convex in output, profit is curved in the demand shock in a way a linear forward cannot neutralize, so the firm buys additional forwards to insure against the residual curvature it cannot hedge directly.

As in Result 1, the VI firm reduces its forward position relative to the production only firm by its expected retail quantity. However, in this setting the VI firm reduces its position *further* by $\frac{\mathbb{E}[\varepsilon^2 \cdot \beta\varepsilon]}{\mathbb{E}[\varepsilon^2]} > 0$, an ε^2-weighted average of the retail demand shock $\beta\varepsilon$.

The difference is due to the optimal hedging of retail demand using swap contracts that provide incomplete insurance. To understand the retail hedging, focus on the retail cost and forward contract terms in the integrated firms problem i.e., $-p^s(\varepsilon)q_V^r(\epsilon) + (p^f - p^s(\varepsilon))f_{V,A}$. In the retail cost term, $p^s(\varepsilon)q_V^r(\varepsilon)$, spot market prices (which depend on the residual demand state in equilibrium) are multiplied by the residual demand shock –which shifts $q_V^r(\varepsilon)$ directly. However, there is no analogous higher-order interaction with the residual demand shock in the contract payout term $(p^f - p^s(\varepsilon))f_{V,A}$, where prices are multiplied by the same constant in all demand states. This means swap futures do not perfectly insure the price times shock component.

Imposing the linear forms from the simple model, the retail exposure is $p^s(\varepsilon)(\theta+\beta\varepsilon)$ and the futures payoff is $(p^f - p^s(\varepsilon))f$. If the firm chose $f = \theta$, the firm fully offsets the $p^s\theta$ variation with their contracts but they are left with $p^s(\varepsilon) \cdot \beta\varepsilon$. This residual term, involves the spot price p^s multiplied by residual demand shocks, whereas in the futures payoff, p^s is only multiplied by a constant f. This means that if ε has a skewed distribution or $p^s(\varepsilon)$ is asymmetric in ε (say from convex costs due to the production technology), then the firm can reduce risk by distorting their forward position away from the expected retail demand.

Figure 2 illustrates the retailer hedging incentive. Each line in panel (a) represents the retail arm's payoff $\pi = p^r q^r - p^s q^r + (p^f - p^s)f$ for a given forward position f, plotted against realized market demand (Q) on the horizontal axis. Since we assume the retail price equals the expected procurement cost, all payoff profiles share the same expected profit; the curves differ only in how payoffs are distributed across states. Panel (b) shows the spot price implied by the supply curve, $p^s(Q)$ alongside the density of demand. Because demand is inelastic, each demand draw maps to a unique point on the supply curve (red line). The convexity of the supply curve means that prices fall only slightly when demand is below average but rise dramatically in high-demand states. This pattern is typical of electricity markets and produces a similar right-tailed effect on retailer profit to the skewness in residual demand states used in the simple model.

Consider, for example, a high-demand realization of $Q = 10137$ MWh, where the spot price is $p^s = 883$. With no forward positions ($f = 0$, blue dotted line), the retail arm loses 2.56 million CAD in a single hour; with forward positions equal to expected demand ($f = -E[q^r] = -2405$, purple solid line) the loss falls to 0.53 million CAD; and the variance-minimizing forward position ($f = f^* = -2983$, green dashed line) reduces it to just 0.04 million CAD. The zero line represents perfect insurance, which occurs when the forward position exactly matches retail demand ($f = -q^r$).

The swap contract must be purchased before demand is realized and does not allow the retailer to adjust their forward position to the realized demand state. Without hedging, profits fall sharply with demand: high demand coincides with high wholesale prices, generating large losses in precisely the states that matter most for a risk-averse retailer. If the retailer sets their forward position equal to negative expected retail demand ($f = -\mathbb{E}[q^r]$), their uninsured payoffs are largest in the high-demand states, and they could reduce their risk by choosing a more negative forward position (to increase payments in the high-demand

states). The optimal forward position therefore is to choose a $f < -\mathbb{E}[q^r]$. Taken together, the figure illustrates the paper's central mechanism: with right-tail asymmetric exposure to demand shocks, the retailer can reduce their payoff risk, by choosing $f < -\mathbb{E}[q^r]$ to improve their payouts in the states where demand and prices are both high.

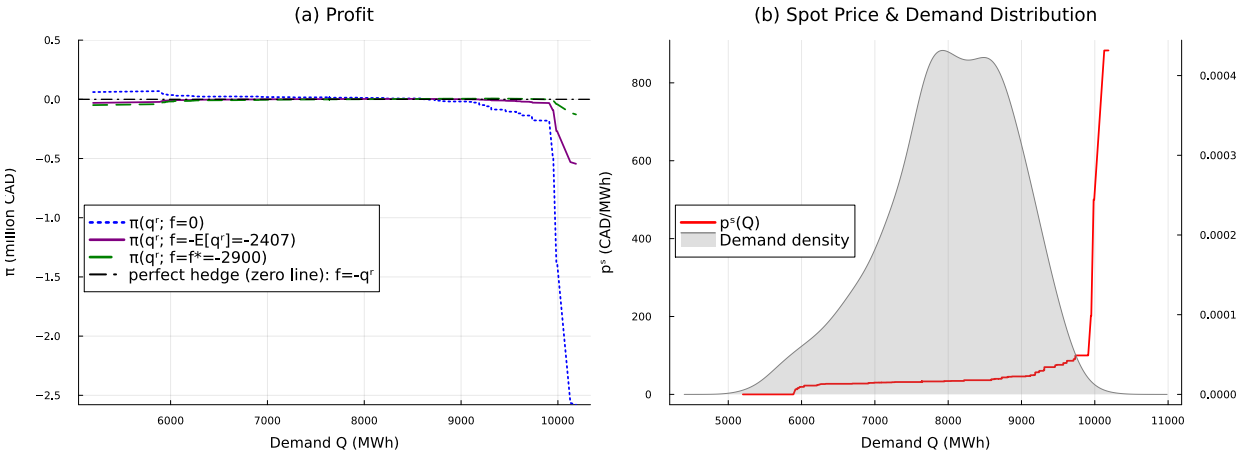


Figure 2: Retailer Payoff Profiles Under Alternative Hedging Strategies

*Notes:*This figure illustrates the retail hedging incentive in two panels. Panel (a) plots the retailer's profit across demand realizations for three forward positions: no hedging $f = 0$, hedging expected demand ($f = -E[q^r]$), and the variance-minimizing position ($f = f^* < -E[q^r]$). To generate realistic price and demand patterns, we draw 8760 hourly net demand observations (demand minus renewable production) from a single year. The zero line represents the perfect hedge ($f = -q^r$), where the forward position exactly offsets the retailer's own demand, eliminating all profit variability. Panel (b) shows the convex supply curve $p^s(Q)$ (from the bids made in the first hour of the first day in 2018) alongside the distribution of demand. The convexity of the supply curve creates asymmetric exposure across demand shocks: losses in high-demand states (where both price and unhedged quantity are large) exceed gains in low-demand states (where spot prices are low and close to the forward price). Optimal hedging reduces profit variance by loading additional forward payments into high-price, high-quantity states, providing insurance precisely when procurement costs are most severe.

It is tempting to expect that a vertically integrated firm with equal retail and generation positions would have no incentive to hedge, since its net physical spot quantity is then constant across states. But this ignores generation costs: because high-demand states raise both revenue and costs, the convexity of costs already partially offsets the curvature of generation revenue, so net generation profit is less convex than revenue alone. With equal positions, the retail-side exposure more than offsets this reduced curvature, leaving profit concave in demand shocks. The VI firm therefore still faces residual risk and over-hedges. In this example, the swap insures perfectly only at $\beta = \frac{1}{4}$, where retail curvature exactly offsets the curvature of generation profit net of convex generation costs.

3.4 Load Following Contracts

The preceding analysis showed that swaps cannot hedge all risk factors in electricity markets. In particular, retail quantity risk and the correlation between price and quantity risk remain unhedged. A natural question is: what are the effects of introducing more complete forward contracts? Specifically, how do contracts whose quantities adjust to realized demand affect equilibrium forward positions?

In a standard swap, the payment is $(p^f - p^s) \cdot f$, where p^f and f are fixed in advance. In this section, we analyze equilibrium contract positions, with load-following forward contracts (LFFC), that pay $(p^f - p^s) \cdot l q^r$, where p^f and l are fixed in advance while p^s and q^r are realized at the production stage. Setting $l = 1$ yields the perfect hedge for the retail exposure depicted in Figure 2. This contract is not merely a theoretical construct: it was recently introduced in the Alberta electricity market and other markets in North America and Australia have begun adopting similar instruments (Brown and Sappington (2023)).

Strategic Incentives and Load Following Contracts. We first consider risk-neutral firms under imperfect competition with residual demand $Q(p^s, \varepsilon) = a + \varepsilon - p^s$. The production game is equivalent to the swap model. In the first stage, the two firms solve:

$$\text{Generator: } \max_{l_{G,N}} \mathbb{E}\Big[p^s q_G^* - C_G(q_G^*) + (p^f - p^s) l_{G,N} q_V^r\Big]$$

$$\text{VI: } \max_{l_{V,N}} \mathbb{E}\Big[p_V^r q_V^r - p^s q_V^r + p^s q_V^* - C_V(q_V^*) + (p^f - p^s) l_{V,N} q_V^r\Big]$$

Result 3 (Strategic Incentives, Load Following Contracts)**.** *The expected equilibrium LFFC forward positions are:*

$$\mathbb{E}[l_{i,N} q_V^r] = f_{i,N}^* + \frac{k^2}{5k^2 + 5k + 1} \frac{-(a-b)\beta^2\sigma^2 + \theta\beta\,\sigma^2}{\theta^2 + \beta^2\sigma^2} \quad \textit{for } i = G, V$$

As in Result 1, equilibrium net positions are identical across the generator and the VI firm. The second term captures the difference between swaps and LFFCs. To build intuition for this comparison, suppose further that the retail demand shock co-moves with system demand proportionally to its mean level: $\beta = \frac{\theta}{a}$. [15] Then, we obtain $\mathbb{E}[l_{G,N} q_V^r] \approx f_{i,N}^*$ to a first-order approximation. As a result, swaps and LFFCs yield nearly identical positions in

[15] We consider a more general error structure, and provide full proofs of the results, in Appendix A.

equilibrium.

Risk Hedging and Load Following Contracts. We next consider risk-averse, price-taking firms with residual demand $Q(\varepsilon) = a + \varepsilon$. The production game is again equivalent to the swap model. Since production-stage outcomes are independent of $l_{G,A}$ and $l_{V,A}$, expected profits do not depend on the LFFC positions, and firms minimize profit variance:

$$\text{Generator: } \min_{l_{G,A}} Var\Big[p^s q_G^* - C_G(q_G^*) + (p^f - p^s) l_{G,A} q_V^r\Big]$$

$$\text{VI: } \min_{l_{V,A}} Var\Big[p_V^r q_V^r - p^s q_V^r + p^s q_V^* - C_V(q_V^*) + (p^f - p^s) l_{V,A} q_V^r\Big]$$

Result 4 (Risk Hedging with Load Following Contracts)**.** *The expected equilibrium LFFC forward positions are:*

$$\mathbb{E}[l_{G,A} q_V^r] \approx f_{G,A}^* - \mathbb{E}[q_G^*] \frac{\mathbb{E}[\varepsilon^2 \cdot \beta\varepsilon]}{\theta\, \mathbb{E}[\varepsilon^2]}, \qquad \mathbb{E}[l_{V,A} q_V^r] \approx f_{G,A}^* - \theta - \mathbb{E}[q_G^*] \frac{\mathbb{E}[\varepsilon^2 \cdot \beta\varepsilon]}{\theta\, \mathbb{E}[\varepsilon^2]} + \left(\underbrace{\frac{\mathbb{E}[\varepsilon^2 \cdot \beta\varepsilon]}{\mathbb{E}[\varepsilon^2]}}_{Correction} - \underbrace{\frac{\mathbb{E}[\varepsilon^2 \cdot \beta\varepsilon]}{\mathbb{E}[\varepsilon^2]}}_{Incomplete\text{-}hedge} \right)$$

The key insight is that LFFCs exactly correct the adjustment to retail hedging for swap incompleteness identified in Result 2. By replacing f, which was constant across states, by $l_{V,A} q_V^r$, these contracts introduce the missing interaction term into the forward contract, allowing contract payoffs to vary nonlinearly with demand shock. This has two important implications. First, the hedging adjustment driven by contract incompleteness is eliminated. Second, whereas the VI firm's forward position was below the generator's in the swap case (Result 2), under LFFCs the two firms' precommitted quantities are approximately equal in expectation ($\mathbb{E}[l_{G,A} q_V^r] \approx \mathbb{E}[l_{V,A} q_V^r] + \mathbb{E}[q_V^r]$).

3.5 Discussion of Results

Results 1 through 4 characterize equilibrium forward positions under alternative assumptions about risk preferences and market power. Figure 3 illustrates equilibrium net spot exposure (production less the retail position and equilibrium forward positions) separately for the strategic commitment case in panel (a) and the risk-hedging case in panel (b). Net spot exposure determines whether a firm has an incentive to raise or suppress the spot price. Blue circles represent the generator's net spot exposure; red squares represent the VI firm's.

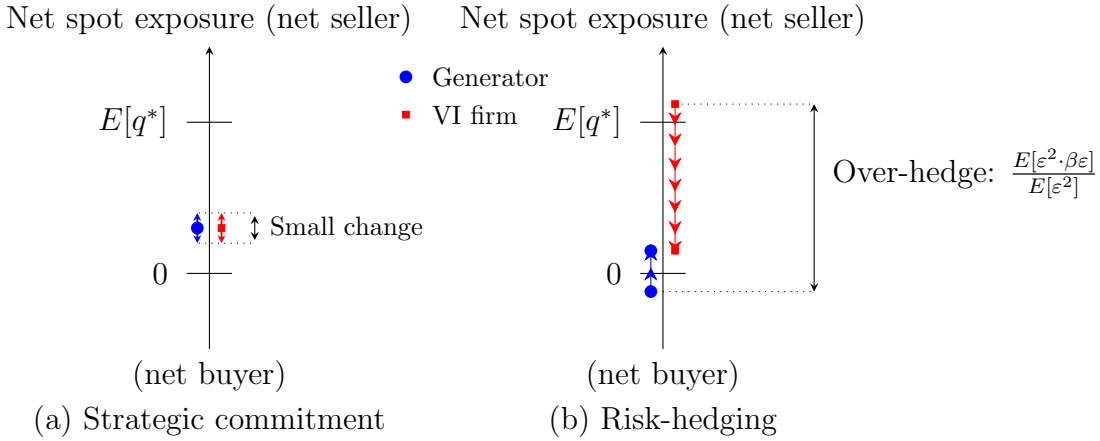
Arrows indicate the shift from optimal swap to optimal LFFC positions. Under risk neutrality with market power (panel a), both firms have identical positive net spot exposure under swaps. The impact of LFFCs is negligible and enters only at second order.

Under risk aversion (panel b), the generator's net spot exposure remains below $\mathbb{E}[q^*]$ with swaps, but the VI firm's may lie above or below $\mathbb{E}[q^*]$ depending on the magnitude of the retailer hedging. Risk-averse generators and VI firms hold *different* net positions: the incomplete-hedging term drives a wedge between them, pushing the VI firm's forward position below the generator's position. LFFCs correct this wedge, moving both the generator and the VI firm toward the same net position, as shown in panel (b).

These results suggest potential anti-competitive effects arising from the interaction between vertical integration and contract incompleteness under imperfect competition. The VI firm may substantially increase their net spot exposure due to their incentives to hedge retail risk, creating an incentive to raise spot prices. However, LFFCs can correct these adverse effects. One may wonder whether these results extend to the case where strategic and risk-based hedging are present simultaneously. In Appendix E, we empirically solve the structural Cournot model with risk-averse players and show that all these results continue to hold.

Figure 3: Graphical illustration of theoretical predictions

This figure plots the net exposure for the generator only firm (blue) and VI firm (red). The left panel shows the position under strategic commitment only. The right panel shows the position with risk-hedging. The $\mathbb{E}[q^*]$ is the expected optimal production (and differs across the two panels). It also differs within the two scenarios in the left panel, however the change in hedge positions are small and so are the changes in production. The arrows show the change from the introduction of LFFCs.



4 Empirical Evidence

In this section, we investigate contract positions and mark-ups of firms in Alberta, before and after the government initiated the load following forwards market. Our theoretical analysis in the previous sections, predicts that the integrated firm should over-hedge their retail exposure when only standard forward contracts are available. When the firm over-hedges, they would then have incentives to charge large markups in the spot market. After the introduction of load following contracts, the integrated firm should take a much smaller forward position and charge mark-ups similar to the other producing firms in the spot market.

We perform two exercises to verify these predictions. First, we analyze mark-ups over engineering estimates of generation unit-level marginal-costs from Brown et al. (2023). This allows us to test the model predictions about changes in mark-ups around the introduction of the new contracts. Second, since forward positions are private information of firms, we introduce a structural model of bidding, similar to Reguant (2014) and Hortaçsu and Puller (2008), to back-out the forward position that rationalizes the observed bidding behavior.

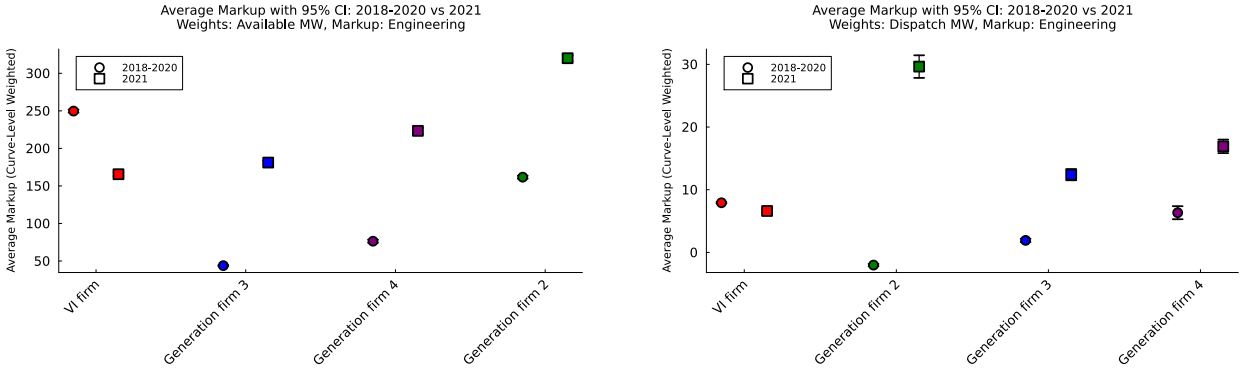
4.1 Mark-up Trends

In July 2020 the Alberta Utilities Commission approved a new hedging strategy for the regulated rate option offered by the integrated firm, which involved holding auctions for load-following forward contracts that provided quantity insurance. The first auction was held in September 2020, for contracts with delivery in December 2020. Our theoretical model predicts that the introduction of these contracts removes the incentive of the integrated firm to over-hedge their retail exposures and as a result will reduce the markups that they offer in the market.

Figure 4 plots quantity-weighted average markups (bids less engineering cost estimates) for each unit across the two periods. Panel (A) shows markups over all available quantities, while Panel (B) focuses on markups for cleared (dispatched) quantities. The results for the integrated firm are striking. They go from being the firm with the largest markups to charging the smallest markups. This is consistent with the theoretical predictions that the introduction of these contracts reduced over-hedging by the integrated firm, reducing their optimal markups. Perhaps surprisingly, the other firms increase their markups somewhat. This pattern is driven by changes in the generation capacity controlled by the generation-only

Figure 4: Engineering Markups

This figure plots the average markups (bid price minus engineering marginal cost) weighted by quantity. Panel (A) shows markups for all available capacity offered, while Panel (B) shows markups for dispatched capacity only. Marginal cost estimates are constructed following Brown et al. (2023).



Panel (A) Panel (B)

firms. Appendix Figure 8 plots the marginal cost curves of the four large firms separately for 2018–2020 and 2021. The differences reflect the expiration of the PPA agreements, which expanded the set of assets controlled by the generation-only firms but not by the integrated firm.[16] In 2021, there was also higher demand, including new all-time records in the summer and winter due to extreme weather events and some untimely outages Alberta Electric System Operator (2022).

Although these markup trends are suggestive, they alone cannot test our theoretical predictions given potential changes in market structure, generation portfolio, and market competitiveness. In the next section, we explicitly account for these potential confounders to test the implications derived in Section 3.

4.2 Estimating Forward Positions

The procedure that we use to estimate forward positions closely follows Reguant (2014). The procedure allows us to control for changes in market structure and competitiveness that cannot be addressed in the simple mark-up trend analysis. However, in this setting there are no complex bids. The market uses a uniform price auction, the price is determined by the last accepted bid needed to clear the demand.

[16]These power purchase arrangements (PPA) had previously reassigned bidding control over some generating units away from their original owners, limiting the concentration of effective market control among the large incumbent firms. See Section 2.1 for institutional details on PPA agreements and Appendix H for asset-level PPA expiration dates and ownership changes.

There are $i = \{1, \ldots, N\}$ firms, who each own a set of generator units that can produce electricity, indexed for each firm by $j = \{1, \ldots, J_i\}$. In each hourly auction (which occurs two hours before dispatch), the firms submit bids on the electricity to be produced at that hour (in MWh). Each generation unit has a minimum and maximum capacity, represented by $\underline{q}_{ij}$ and $\overline{q}_{ij}$. Whenever a unit is running, it needs to produce within the range $q_{ijh} \in [\underline{q}_{ij}, \overline{q}_{ij}]$. Each firm i chooses a bidding strategy to maximize its expected profits, conditional on a set of common public information and independent private shocks, assuming that the opposing firms submit bids according to their equilibrium strategies.

Define the collection of bids made by all firms as the array $\mathbf{b}$. Bids for each firm (i) are hour (h), unit (j), specific step functions i.e., price, quantity pairs $\{b_{ijhk}, g_{ijhk}\}$ at steps $k = 1, \ldots, K$, where K is the maximum number of steps set by the auctioneer at $K = 7$. Price bids must be between \$[0, 999.99] and the total quantity bid must be the capacity of the unit. Price bids need to be weakly increasing in quantity. Formally, bids for a given firm i are defined as:

$$\mathbf{b}_{ih} = \Bigg\{ \left(\vec{b}_{ijh}, \vec{g}_{ijh}, K_{ijh}\right) : \dim(\vec{b}_{ijh}) = \dim(\vec{g}_{ijh}) = K_{ijh} \in \{1, \ldots, 7\} \quad \forall j$$
$$b_{ijhk} \in [0, 999.99], \quad g_{ijhk} \in [\underline{q}_{ij}, \overline{q}_{ij}] \quad \forall k > 1 : b_{ijhk} > b_{ijhk-1}, \quad g_{ijhk} > g_{ijhk-1} \Bigg\}.$$

Expected profits of firm i for a given hour are:

$$\mathbb{E}_{-i}[\Pi_i(\mathbf{b})] = \mathbb{E}_{-i}\left[p_h(\mathbf{b}_h) \left(Q_{ih}(\mathbf{b}_h) - \nu_{ih}\right) - \sum_j C_{ij}(\mathbf{q}_{ij}(\mathbf{b})) \right], \tag{1}$$

where $p_h(\cdot)$ represents the equilibrium price, $Q_{ih}(\cdot)$ is the total quantity sold by firm i at hour h, ν_{ih} are the firm's pre-committed quantity (i.e., the sum of its retail demand and forward contract volume), and $C_{ij}(\cdot)$ represents the hourly costs of unit j belonging to firm i, which depends on the hourly equilibrium quantities used. We focus on the cost structure of thermal units (coal and gas) owned by large firms. We assume the cost function for unit j is:

$$C_{ij}(\mathbf{q}_{ij}) = \alpha_{ij1} q_{ijh} + \frac{\alpha_{ij2}}{2} \tilde{q}_{ijh}^2 \tag{2}$$

where $\alpha_{ij1}, \alpha_{ij2}$ represent unit-specific marginal costs of production, and $\tilde{q}_{ijh}$ represents the quantity over the minimum production level $\underline{q}_{ij}$, i.e., $\tilde{q}_{ijh} = \max\{q_{ijh} - \underline{q}_{ij}, 0\}$. Reguant

(2014) demonstrates that ramp-up costs for gas and coal units either have no effect or offset the impact of quadratic costs, such that including a ramp-cost term does not alter estimation results. We therefore exclude these costs from our specification.

The goal of the firm is to choose its bids $\mathbf{b}_i$ to maximize its expected profit.[17] A necessary condition for optimality is that a firm must be indifferent between raising or lowering the whole bidding offer at a given step of the supply curve, given by a bid b_{ijkh}, for unit j at step k and hour h:

$$\mathbb{E}_{-i}\left[\frac{\partial \Pi_i(\mathbf{b})}{\partial b_{ijkh}}\right] = 0. \tag{3}$$

which implies the bid is equal to the average marginal cost plus a markup

$$b_{ijkh} = \zeta_{ijkh} - \frac{\mathbb{E}_{-i}[Q_{ih} - \nu_{ih} \mid p_h = b_{ijkh}]}{\partial \mathbb{E}_{-i}[Q_{ih} \mid p_h = b_{ijkh}]/\partial b_{ijkh}}, \tag{4}$$

where $\zeta_{ijkh} \equiv \frac{\partial \mathbb{E}_{-i}[C_{ij}|p_h=b_{ijkh}]/\partial b_{ijkh}}{\partial \mathbb{E}_{-i}[Q_{ih}|p_h=b_{ijkh}]/\partial b_{ijkh}}$ is the weighted expected average marginal cost when b_{ijkh} sets the price. The markup depends on the expected infra-marginal quantity produced by the firm when the unit is accepted, divided by its effect on equilibrium quantities, which is equivalent to its effect on the residual demand.

We impose two additional restrictions on the functional form. First, the marginal costs function is assumed to be:

$$\zeta_{ijkht}(\boldsymbol{\alpha}_i) \equiv \begin{cases} \alpha_{j1}\, p_t^{fuel} + \alpha_{j2}\, p_t^{fuel} \tilde{q}_{ijht} + \epsilon_{ijkht}, & \text{if } j \text{ is a gas unit,} \\ \alpha_{j1} + \alpha_{j2}\, \tilde{q}_{ijht} + \epsilon_{ijkht}, & \text{if } j \text{ is a coal unit.} \end{cases} \tag{5}$$

where t represents the day of the sample, and the error term ϵ_{ijkht} can be interpreted as a shock on marginal costs known to the firm as well as optimization or specification error. p_{jht}^{fuel} denotes the daily spot price of natural gas sourced from the Alberta Natural Gas Exchange (NGX), reflecting short-run variation in fuel input costs. Coal-fired generators, by contrast, face no analogous spot market price: coal procurement in Alberta has historically operated through long-term contracts or mine-mouth arrangements that fix input costs over extended

[17] Appendix K discusses why we use a risk-neutral bidding condition while allowing forward choices to reflect risk-averse hedging motives. One interpretation is organizational: the risk-management desk chooses forward positions to manage low-frequency cash-flow risk, while the dispatch desk later solves an expected-profit continuation problem. A second interpretation allows the mean–variance objective to apply to hourly bidding, but we present evidence that the omitted mean–variance correction term is empirically small and that excluding it does not raise first-order bias concerns for the estimation.

horizons (Independent Assessment Team (1999); Alberta Energy Regulator (2025)). Consequently, for coal units, α_{j1} and α_{j2} absorb the relevant contracted fuel costs rather than prevailing market prices.

Second, we follow Reguant (2014), and assume the firm's pre-committed quantity is

$$\nu_{ht}(\boldsymbol{\gamma}_i) \equiv \gamma_i \bar{q}_{ht} + \varepsilon_{ht}, \tag{6}$$

where $\bar{q}_{ht}$ represents the expected quantity sold at the spot market. This hedging strategy allows generation firms to adjust their forward position to capture differential hedging in peak, off-peak hours and across seasons. It also captures the quantity effect of retail sales for the integrated firm in a parsimonious way. Appendix I reports estimates under two alternative specifications, including one that explicitly incorporates the integrated firm's retail load. The implied forward positions and quantity exposures are similar.

In practice the parametric restrictions mean that we have a set of linear moment equations with many more equations than parameters and the model is identified.[18] We aim to estimate the parameters of the cost function $\boldsymbol{\alpha}$, and the forward positions γ_i. The estimator has two steps. First, to construct the empirical analogue of the first-order condition in equation, we first estimate its expectation terms. To do this, we follow the multi-unit auction literature and use a resampling procedure to estimate the expectation over residual demand. Then, we use a generalized method of moments estimator based on moments constructed from these first-order conditions. Given the potential endogeneity and measurement error of markup terms, we use the publicly available demand forecast as instruments.

For resampling, hours are classified into interaction categories defined by weekday group (weekday vs. weekend), hour of day (0–23), and ownership- and portfolio-based subperiods. The ownership and portfolio subperiods are defined in Appendix 4.1 such that major ownership remains stable within each subperiod (e.g., following PPA expirations and asset exits). Within each category, we match comparable hours based on the squared distance in demand forecasts and renewable generation. This approach assumes that this conditioning is the same as the conditioning performed by bidders, i.e., there is no unobserved heterogeneity that influences bidders' expectations but not those of the econometrician.

[18] Forward positions are set at the firm level, whereas marginal costs vary more granularly across generating assets and offer steps. With additional assumptions, within-firm variation across assets and steps could be used to distinguish more flexibly the role of marginal costs from that of forward positions in determining firms' bids. In practice, however, such an approach would likely yield imprecise estimates.

Given this set of auctions, we simulate the beliefs of one firm i about the other firms' strategies using a procedure similar to Hortaçsu (2002). We resample (with replacement) each opponents' strategy independently from the set of similar hours. This gives a distribution of residual demand curves, with residual demand in bootstrap sample bs represented by $D_{ih}^{R,bs}(\cdot)$. In each simulated hour, the market clears where the quantity produced by firm i, $Q_{ih}(\cdot)$, equals the residual demand. We follow a smoothing approach that has been used in the context of electricity auctions (Reguant (2014);Wolak (2007)), and approximate demand and supply as continuous functions that depend on price. For a given bootstrap sample bs, both the residual demand faced by a firm and its supply curve are approximated as follows:

$$\hat{D}_{ih}^{R,bs}(p_h, \mathbf{b}_{-i,h} \mid s^{bs}, bw) = \sum_{l \neq i} \sum_{k} g_{ljhk} K\left(\frac{b_{ljkh} - p_h^{bs}}{bw}\right), \tag{7}$$

$$\hat{Q}_{ih}^{bs}(p_h, \mathbf{b}_{i,h} \mid s^{bs}, bw) = \sum_{j} \sum_{k} g_{ijkh} K\left(\frac{b_{ijkh} - p_h^{bs}}{bw}\right), \tag{8}$$

where K is a Kernel cumulative weight and bw is a bandwidth parameter that determines the degree of smoothing.

With this smooth representation, the empirical analogue to the first-order condition of unit j at step k in hour h can be constructed. The empirical analogue is:

$$m_{ijkht}(\boldsymbol{\alpha}_i, \boldsymbol{\gamma}_i; bw, B) \equiv \frac{1}{B} \sum_{bs=1}^{B} \mathbb{1}(j \text{ in}) \frac{\partial p_{ht}^{bs}}{\partial b_{ijkht}} \left[(b_{ijkht} - \zeta_{ijkht}(\boldsymbol{\alpha}_i)) \frac{\partial D_{iht}^{R,bs}}{\partial p_{ht}} + (Q_{ijkht}^{bs} - \nu_{iht}(\boldsymbol{\gamma}_i)) \right], \tag{9}$$

where B represents the number of bootstrap simulations taken for each hour. In equilibrium:

$$\frac{\partial p_{ht}}{\partial b_{ijkht}} = \frac{\partial Q_{ih}/\partial b_{ijkh}}{\partial D_{ih}^{R}/\partial p_h - \partial Q_{ih}/\partial p_h}, \tag{10}$$

which can be computed using the smooth approximation.

With all these elements, the moments are constructed. We drop all non-marginal bids made at the price floor or ceiling (these are typically used to manage start-up). Given the potential endogeneity and measurement error of markup terms, we use the publicly available demand forecast as instruments, represented with $\mathbf{Z}_{ht}$. For a given firm i, the empirical moments are:

$$\sum_{t=1}^{T} \sum_{k=1}^{K} \mathbf{Z}_{ht} m_{ijkht}(\boldsymbol{\alpha}_i, \boldsymbol{\gamma}_i; bw, B) = 0, \quad \forall j, h, \tag{11}$$

which provides a consistent estimate for $\boldsymbol{\alpha}_i$ and $\boldsymbol{\gamma}_i$. Following Reguant (2014), we use two stage least squares weights.

4.2.1 Results

We estimate the model separately for the periods before and after the introduction of the load following futures contract. Figure 5 presents the implied net spot exposure ratio based on the GMM estimates, defined as the share of expected production sold at spot prices. The four generation-only firms (average denoted by blue circles) have relatively low net spot exposure, implying they pre-sell a substantial share of their generation forward, with only a small increase after 2020. The VI firm (denoted by red squares) initially has substantially higher net spot exposure, consistent with a long forward position, but from 2021 on, its exposure falls toward that of the generation-only firms. These patterns are consistent with the predictions of our model. This analysis differs from the markup trend analysis in Figure 4 in several respects. Raw markups reflect not only forward positions but also market structure, generation portfolios, and the degree of market competitiveness. On the other hand, our GMM estimation isolates the role of forward positions and marginal costs from these confounding factors by explicitly accounting for each firm's generation portfolio and ownership structure, expected market demand, and the elasticity of residual demand.

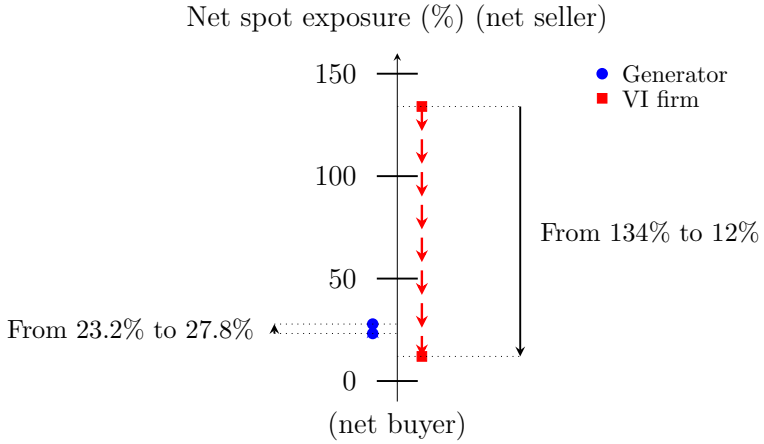


Figure 5: Net spot exposure changes with GMM estimates

Notes: The figure shows implied net spot exposure ratios based on GMM estimates, before 2021 and in 2021. Blue circles represent the average hedge ratio across generation-only firms, while red squares represent the vertically integrated firm. Each pair of points shows the net spot exposure ratio before 2021 and in 2021, with arrows indicating the direction of change. Generation-only firms decreased their net spot exposure from 23.2% to 27.8% of expected production, while the VI firm dramatically decreased from 134% to 12%.

Table 6 summarizes the GMM estimates of marginal costs and forward positions alongside

engineering estimates of marginal costs. The results are consistent with firms' annual reports. One generation-only firm directly reports its contract positions annually, with positions of 87%, 78%, and 72% for 2018–2020, respectively, which align closely with the estimates in Table 6. While the vertically integrated firm does not directly report contract positions, it discloses the earnings impact of a 10 percent change in the forward price of electricity (i.e., $0.1 \times p \times q^f$). To recover a rough estimate of the firms' forward contract positions, we use this formula, together with the average annual spot market prices, and the firm's reported losses from price changes. This yields an exposure of 22.7 million MWh in December 2017, covering more than one year of their retail load share (16 million MWh in our sample period). Given that the integrated firm's production is approximately 8 million MWh per year, this back-of-the-envelope estimate substantially exceeds net retail demand (retail load minus own production). In December 2020, the first month with load-following futures available for delivery, their position was 3 million MWh, covering roughly 8 months of net load ($13.4 - 8.9 = 4.5$ million MWh in that year).

Table 6 also reports average marginal cost estimates by fuel type. The increase in costs in 2021 are driven primarily by changes in gas prices and environmental compliance costs. The GMM estimates align well with engineering estimates, though some differences remain. Several factors may explain these differences. First, although engineering cost estimates provide useful reference points, heat rate data predate our sample period by 6–14 years, and thus do not reflect subsequent technology derating and upgrading. Second, engineering estimates do not capture unobserved cost shocks arising from factors such as temperature variation or emission market trading costs. These cost changes, however, are reflected in firms' bidding prices and are therefore captured by the GMM estimates. Overall, the GMM estimates provide evidence consistent with our theoretical predictions and yield reasonable estimates of marginal costs and forward positions when compared to engineering estimates and industry reports.

5 Quantifying the effects of Contracts and Vertical Integration

The previous section presented evidence that the VI firm over-hedges their retail demand, and that this over-hedging is corrected following the introduction of load-following contracts,

Table 2: Average marginal cost and forward position estimates for large firms

Firm	Fuel Type	Before 2021			2021		
		GMM MC ($/MWh)	Eng. MC ($/MWh)	GMM Forward (%)	GMM MC ($/MWh)	Eng. MC ($/MWh)	GMM Forward (%)
Generation firm 1	COAL	27.81	34.99	76.1%**	–	–	–
		(1.75)		(3.7%)			
	GAS	23.01	24.07		–	–	
		(0.68)					
Generation firm 2	COAL	20.91	29.13	66.0%**	27.96	37.34	70.2%**
		(2.43)		(16.2%)	(3.38)		(8.1%)
	GAS	24.98	18.90		41.62	32.06	
		(2.11)			(3.52)		
Generation firm 3	COAL	29.84	34.35	89.7%**	34.28	36.85	59.3%**
		(2.69)		(19.7%)	(7.27)		(13.8%)
	GAS	25.53	31.46		39.21	45.69	
		(1.80)			(3.39)		
Generation firm 4	COAL	29.28	31.72	75.6%**	42.63	38.08	87.3%**
		(2.27)		(11.9%)	(2.59)		(4.0%)
	GAS	40.45	32.85		55.53	49.01	
		(2.96)			(3.75)		
VI firm 1	GAS	24.57	19.57	-34.3%	40.94	32.99	88.0%**
		(3.96)		(19.9%)	(6.60)		(19.6%)

Notes: GMM marginal costs are computed at each unit's average output level and aggregated by fuel type and firm; unit-level estimates are reported in Appendix J. Because one generation firm exits and another enters during the sample period, the market is characterized by three dominant generation-only firms at any given point in time rather than four. Several assets are excluded from the firm-level aggregation: three assets are excluded due to the absence of engineering MC estimates, and three additional assets belonging to Generation-Only Firm 3 are excluded because they appear in only ∼1%, ∼3%, and ∼3% of that firm's total operating hours, respectively; their small sample sizes yield highly imprecise GMM MC estimates that would distort the firm-level MC estimates. All excluded assets are reported in Appendix J. Bootstrap standard errors in parentheses. For *GMM Forward* estimates, ** (*) denotes significance at the 1% (5%) level under H_0: forward position = 0.

at which point the VI firm's hedging position converges to levels comparable to those of generation-only firms. We now turn to quantifying the competitive and welfare implications of these findings.

When we combine risk aversion and competition we no longer obtain analytic expressions to characterize the optimal forward contract position. Therefore, we numerically solve a richer model that combines these two forces. This allows us to empirically asses the effects that arise from the interaction between vertical integration and contract incompleteness. Our model accommodates multiple strategic firms with capacity constraints and flexible marginal costs and uses estimated residual demand, marginal cost functions and demand uncertainty.

We model 3 generation only firms and 1 VI firm that are playing a Cournot Nash equilibrium. We solve the two stage model by backward induction. We observe the VI firm's hourly retail demand and thus approximate the joint distribution of system load and the VI firm's own retail load. At the second stage, demand uncertainty is realized, and conditional on the realized levels of market demand and the VI firm's retail demand, firms choose their optimal quantities to produce given their forward commitments. The firms face a residual demand curve composed of the price inelastic market demand, less the supply of fringe suppliers.

Although firms in this market bid supply schedules at the auction, we compute equilibrium under Cournot competition for tractability. This eliminates demand uncertainty at the time of bidding, which is a reasonable approximation given that bid functions allow the firm to offer conditional prices on small sets of demand states. However, by the logic of Klemperer and Meyer (1989), a firm choosing a supply schedule tends to face residual demand that is more elastic than under Cournot competition. Under Cournot, a firm takes rivals' quantities as fixed, so its residual demand shifts one-for-one with its output, but under supply-function competition, rivals' schedules slope upward and absorb part of the price movement from a quantity change, flattening the residual demand the firm faces. With more elastic residual demand, the firm will tend to charge lower markups under supply-function competition than Cournot.[19] Since the Cournot exercise produces a similar distribution of clearing prices to those observed in the data (Figure 10), we believe this approximation is reasonable.

At the first stage, firms choose forward contract positions, taking into account their effects on the second-stage equilibrium. Firms take forward positions for both strategic commitment and risk-hedging purposes. Firms are risk averse with constant absolute risk aversion (CARA) preferences, so that firm i maximizes expected utility

$$\mathbb{E}[-\exp(-\lambda_i \Pi_i)]$$

following Hughes and Kao (1997); Léautier and Rochet (2014); van Eijkel et al. (2016). We use CARA preferences because they provide an expected-utility foundation, and they capture higher-order features of the profit distribution that are not summarized by the mean and variance alone. The mean-variance specification used in the analytical model can be

[19] Klemperer and Meyer (1989) provide a sharp result that the unique supply function equilibrium is bounded between the Cournot (an upper bound on price and lower bound on quantity) and Bertrand outcomes with symmetric firms, a demand shock with sufficiently wide support and no private information (i.e., interpreting the marginal cost and forward position shocks ϵ, ε as specification errors).

interpreted as a second-order approximation to CARA expected utility and is adopted for analytical tractability. We assume that the generation only firms have a common risk-aversion parameter, and the vertically integrated firm may have a different level of risk aversion. Since the retail firm faces risk from high-prices in states when they also have to purchase a large quantity, we interpret the risk aversion parameter as reflecting concerns about default risk and large cash outflows. In line with this, in each of the California electricity crisis, 2021 ERCOT crisis, and 2020-2021 Japan wholesale price spikes, several major retailers entered bankruptcy or exited the market.

In contrast, generation-only firms face lower default risk from price fluctuations. Since wholesale prices follow a right-skewed distribution bounded between \$0 and \$999/MWh, generators' payoffs never become negative—the zero lower bound prevents large cash outflows. Price spikes, rather than threatening generators' solvency, typically generate extremely high cash inflows. So, short-term price fluctuations are unlikely to directly cause bankruptcy among generation firms. Given these asymmetric risk exposures, we view it as plausible that integrated firms (which incorporate both generation and retail operations) may have different risk aversion than generation-only firms.

We select risk aversion parameters to match the GMM estimates of forward positions reported in Table 6. Specifically, we choose λ_{GEN} and λ_{VI} to minimize the distance between the swap forward positions implied by the solution to our model, $f_i(\lambda_{GEN}, \lambda_{VI})$, and those estimated by GMM, f_i^{GMM} :

$$\min_{(\lambda_{GEN}, \lambda_{VI})} \sum_i \left(f_i(\lambda_{GEN}, \lambda_{VI}) - f_i^{GMM} \right)^2$$

We solve this problem by grid search. In Appendix E, we present results across a wide range of alternative risk aversion parameters, which show our conclusions are similar for a wide range of parameter values. We obtain the minimizing parameter pair $(\lambda^*_{GEN}, \lambda^*_{VI}) = (1.0 \times 10^{-7},\ 3.16 \times 10^{-6})$. Consistent with the cash-outflow and bankruptcy risks discussed above, these risk aversion estimates imply that the VI firm is more risk averse than the generation-only firms. [20]

Appendix L instead provides a back-of-the-envelope calculation of how increased risk

[20] In Appendix F, we further show that as the VI firm's risk aversion parameter increases, profit variance falls sharply with modest sacrifice in expected profits.

affects borrowing costs.[21] Since credit assessments place substantial weight on firms' performance in adverse states with low cash flows, a firm that does not hedge may receive a lower credit rating and face higher financing costs. We compare the annual profit distribution of a VI firm under two benchmarks: a risk-neutral benchmark and a risk-averse benchmark. The risk-neutral benchmark exposes the firm to substantially worse downside cash-flow risk, with profits $70 million CAD lower at the first percentile and $50 million CAD lower at the 10th percentile. This corresponds to a $50 million reduction in earnings before interest and taxes (EBIT). We map these EBIT reductions into interest-coverage ratios and synthetic spreads using the VI firm's 2019 financial statements, and then use the credit-rating schedules in Damodaran (2026) to calculate how this affects the firm's rating and expected borrowing costs. The exercise implies that downside realizations increase expected annual borrowing costs by $49–158 million. These magnitudes exceed the roughly $24 million reduction in expected spot-market profits incurred when moving from the risk-neutral benchmark to the risk-averse benchmark, suggesting that hedging can be value-creating by reducing the firm's financing costs.

To solve the model, we build on the algorithm of Ito and Reguant (2016). The algorithm solves a mixed-integer-linear-programming problem to solve firms' quantities from Cournot competition in the spot market given each realization of the demand shocks, then uses an outer Gauss-Seidel iteration updating each firms' choice of contracts. We provide the full details of our implementation in Appendix C. Our implementation differs from Ito and Reguant (2016) in three ways. First, rather than approximating the residual demand curve with a single linear line, we introduce a piecewise linear approximation that captures the shape of the entire residual demand curve. A 5-segment approximation provides sufficient accuracy, as shown in Appendix Figure 9. This additional flexibility is important to capture the high-price, high-demand states, which play an important role in generating risk.

Second, since we model financial forward contract choice, rather than the day-ahead choice, choices are at a much lower frequency, and we assume that the firm faces demand realizations drawn from the annual distribution of hourly demand. The demand has two parts, an system demand and the retail demand for the integrated firm. Following Ito and Reguant (2016), we assume system demand shocks are normally distributed and we approximate the

[21] Pérez-González and Yun (2013) exploit the introduction of weather derivatives in 1997 and show that risk management increased firm value for U.S. electric and gas utilities. Do et al. (2024) further show that weather-derivative hedging lowers borrowing costs for U.S. electricity and natural gas firms.

joint distribution of system demand, and demand for the vertically integrated retailer using a discrete grid of quantiles (at which, we solve the spot-market equilibrium conditional on demand).[22] Third, we introduce a risk averse utility maximization problem and assume perfect arbitrage between the forward and spot markets (due to the presence of financial arbitrageurs). In the equilibrium calculations we use the estimates of the marginal cost of all thermal generation units owned by the four firms, obtained as described in Section 4.2.

We use this model to evaluate six scenarios:

- **Swap+VI**: Three risk-averse generation-only firms with parameter λ^*_{GEN} and one risk-averse VI firm with parameter λ^*_{VI} choose swap forward positions in the first stage and produce with market power in the second stage. This scenario most closely approximates observed firm behavior before 2021.

- **Swap+VI (Neutral)**: Identical to Swap+VI, except all firms are risk neutral. This isolates the effects attributable to risk aversion alone.

- **Swap+Gen (VI risk)**: A counterfactual separation of the VI firm, into a generation firm and a retail firm where all firms operate as generation-only. Three generators have risk aversion λ^*_{GEN}, while the separated VI firm retains λ^*_{VI}. Firms choose swap forward positions in the first stage and produce with market power in the second stage.

- **Swap+Gen (Gen risk)**: Identical to Swap+Gen (VI risk), except the separated VI firm adopts λ^*_{GEN}. If risk aversion reflects bankruptcy and large cash outflows concerns, this parameterization is more appropriate, as the separated generation-only entity would face a risk profile similar to other generators.

- **LFFC+VI (VI load)**: Identical to Swap+VI, except firms choose load-following forward contract (LFFC) positions rather than swap positions in the first stage, with each contract indexed to the VI firm's own demand. This scenario most closely approximates observed firm behavior in 2021.

[22]The grid consists of evenly spaced interior deciles and two extreme quantiles, calibrated so that the lower and upper tails together correspond to realizations that occur with expected frequency once per year. To incorporate variation in the VI firm's demand beyond the system-load, we estimate the conditional expectation of its retail load given system demand at each of the grid points. Around each conditional mean we place two additional points (at plus and minus two standard deviations of retail load net of the conditional mean). In the forward contract decision and calculation of expected outcomes, the spot-market outcome at each quantile is weighted by the probability mass of the interval it represents.

- **LFFC+VI (System load)**: Identical to LFFC+VI (VI load), except the LFFC is indexed to system demand rather than to the VI firm's own demand.

- **First-best**: All firms offer production at marginal cost in the second stage.

Swap+VI serves as the baseline scenario in which the over-hedging incentive is present (similar to Alberta before 2021). Each alternative relaxes one dimension of the baseline. Swap+VI (Neutral) removes risk aversion, Swap+Gen removes vertical integration, and LFFC+VI reduces contract incompleteness.

LFFC+VI (VI load) indexed to the VI firm's own retail load is more closely tailored to that firm's actual demand and therefore reduces contract incompleteness to a greater degree. In practice, however, indexing settlement to the firm's own demand raises a moral-hazard concern: because the LFFC pays a fixed price and shifts volume risk to the wholesale counterparty, the retailer may have an incentive to adjust or game its realized demand ex post. To address this, we also consider an LFFC indexed to total system load (LFFC+VI (System load)). Because system demand is largely outside any single firm's control, this specification substantially mutes the manipulation concern, albeit at the cost of a looser hedge against the firm's own load.

As a model-validation exercise, Appendix Figure 10 compares the wholesale price distribution from the Swap+VI simulation with actual prices in 2018. The model captures the pattern of actual price distributions reasonably well, despite the fact that no wholesale price data are used to calibrate any parameter in the simulation.[23]

Figure 6, Panel A, presents the expected price difference relative to the first-best scenario under each market structure. In the Swap+VI case, the average price exceeds the first-best by \$8.12 CAD/MWh—approximately 11 percent of the first-best price. Separating the integrated firm into retail and generation-only firms (Swap+Gen (Gen risk)) reduces prices by 8.12-4.07=\$4.05 CAD/MWh, yielding the largest competitive effect among the alternatives. This effect is similar when the VI risk parameter is used for the separated firm. Switching from swaps to LFFCs also lowers prices, but by less than separation. The price premium falls to \$5.31 CAD/MWh when the LFFC is indexed to the VI firm's load and to \$5.97 CAD/MWh when it is indexed to system load, corresponding to reductions of \$2.81 and \$2.15 CAD/MWh, respectively. The Swap+VI (Neutral) scenario indicates that risk

[23]This differs from Ito and Reguant (2016), who use actual clearing prices to estimate the slope of the linear demand curve. Since we adopt a global approximation of demand, clearing prices are not required.

aversion of the VI firm accounts for roughly $2.89 CAD/MWh of the price distortion. As shown in Figure 11, these results imply that separation or more complete contracts would reduce consumer costs by approximately $332 and $233 million CAD per year, respectively.

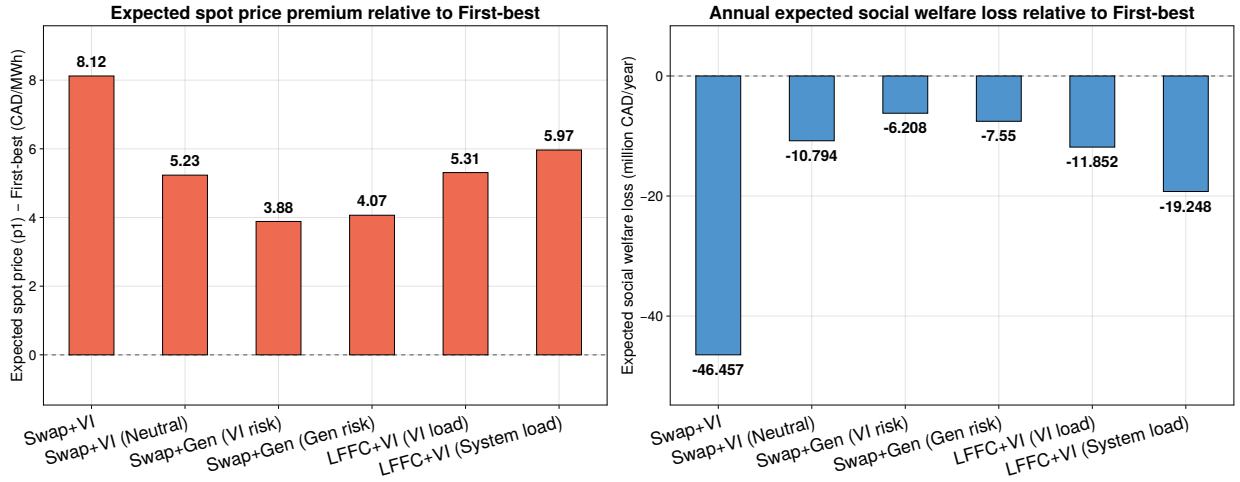


Figure 6: Price and Welfare Effects Across Simulation Scenarios

Notes: Panel A reports the expected price increase relative to the first-best benchmark under each simulation scenario. Panel B reports annual deadweight loss under each scenario. Deadweight loss compares total production costs in each scenario to those in the first-best benchmark.

Figure 6, Panel B, reports annual deadweight loss under each scenario. The efficiency implications mirror the competitive effects. Under Swap+VI, the annual welfare loss is approximately $46.46 million CAD. Separation substantially improves welfare, reducing the annual loss to $6.21 million under the VI-risk specification and $7.55 million under the generation-risk specification. These changes correspond to welfare gains of $40.25 million and $38.91 million per year, or reductions of approximately 87 percent and 84 percent in the baseline welfare loss. LFFCs also improve welfare, although the magnitude depends on the load used to index the contract. The annual welfare loss falls to $11.85 million when the LFFC is indexed to the VI firm's load and to $19.25 million when it is indexed to system load. These values imply welfare gains of $34.61 million and $27.21 million per year, or reductions of approximately 74 percent and 59 percent in the baseline welfare loss, respectively. Finally, the Swap+VI (Neutral) scenario has an annual welfare loss of $10.79 million. Relative to the Swap+VI baseline, moving to the risk-neutral specification eliminates approximately $35.66 million, or 77 percent, of the baseline welfare loss.

Table 3 summarizes the full simulation results. For each scenario, the table reports market-level outcomes—average price $\mathbb{E}[p]$ (CAD/MWh), deadweight loss ($1000 CAD/hour),

and consumer cost difference relative to the first-best ($1000 CAD/hour)—as well as firm-level outcomes (averaged over three generation-only firms versus the single VI firm): net forward position (%), thermal production (MWh), expected profit ($1000 CAD/hour), and standard deviation of profit ($1000 CAD/hour).

Several additional findings emerge beyond the welfare and competitive effects discussed above. Consistent with our theoretical predictions, the VI firm maintains a very low net forward position in the Swap+VI case. This outcome arises from the interaction of three elements: vertical integration, risk aversion, and contract incompleteness. Relaxing any one of these—through structural separation (Swap+Gen), removing risk aversion (Swap+VI Neutral), or introducing LFFCs (LFFC+VI)—yields substantially higher net forward positions.

The profit results reveal that the VI firm can substantially reduce profit variance through forward contracting with little sacrifice in expected profits. Comparing Swap+VI with Swap+VI (Neutral), risk aversion reduces the standard deviation of the VI firm's profit from 219 to 70.6 thousand CAD per hour, a reduction of approximately 68 percent, while reducing expected profit from 62.1 to 57.9 thousand CAD per hour, a reduction of approximately 6.8 percent. Although separation and LFFCs generate similar competitive and welfare improvements, LFFCs can provide better risk sharing without sacrificing expected profit. Relative to Swap+VI, the VI-load LFFC reduces the VI firm's profit standard deviation from 70.6 to 17.5 thousand CAD per hour while increasing expected profit from 57.9 to 61.7 thousand CAD per hour. The system-load LFFC also provides meaningful benefits, although it delivers less risk reduction than the VI-load LFFC. Across all scenarios, Swap+VI yields the highest expected profits for the incumbent generation-only firms.

6 Conclusion

This paper demonstrates that the conventional view of vertical integration as pro-competitive in electricity markets overlooks a critical interaction, from incomplete forward contracts and risk aversion. While prior literature has emphasized how vertical integration reduces firms' net exposure to spot prices and thereby intensifies competition, we show that this logic depends crucially on the availability of complete hedging instruments. When firms can only access standard swap contracts that hedge price but not quantity, risk-averse vertically integrated retailers optimally over-hedge their retail obligations, entering the spot market

Table 3: Summary of Outcomes Across Simulation Scenarios

Scenario	E[p] (CAD/MWh)	Net Position (%)	Production (Thermal) (MWh)	E[π] (k CAD/MWh)	SD(π) (k CAD/MWh)	DWL (k CAD/h)	Cons. cost Δ from FB (k CAD/h)
Swap+VI							
Gen-only (avg.)	83.8	65.4	591	60.3	47.2	5.30	76.4
VI		-25.5	628	57.9	70.6		
Swap+VI (Neutral)							
Gen-only (avg.)	80.9	55.0	491	57.2	77.4	1.23	49.5
VI		84.1	1046	62.1	219		
Swap+Gen (VI risk)							
Gen-only (avg.)	79.6	66.9	505	56.2	57.1	0.71	36.9
VI		74.6	1058	60.8	3.22		
Swap+Gen (Gen risk)							
Gen-only (avg.)	79.8	68.4	516	56.4	53.2	0.86	38.5
VI		65.0	1017	60.7	29.2		
LFFC+VI (VI load)							
Gen-only (avg.)	81.0	61.2	507	57.2	31.8	1.35	49.8
VI		56.7	993	61.7	17.5		
LFFC+VI (System load)							
Gen-only (avg.)	81.7	61.8	525	57.9	30.2	2.20	56.0
VI		46.0	912	60.9	67.6		
First-best							
Gen-only (avg.)	75.7	0.00	540	53.3	164	0.00	0.00
VI		0.00	1106	56.5	519		

Note: The labels "Gen-only (avg.)" and "VI" refer to the three generation-only firms and the one vertically integrated firm observed in the data, respectively. These labels are retained across all scenarios for consistency, even in counterfactual scenarios such as Swap+Gen (VI risk) and Swap+Gen (Gen risk), where the VI firm is modeled as generation-only. Production, profits, and the standard deviation of profits are reported at the firm level. For "Gen-only (avg.)," firm-level variables are reported as averages across the three generation-only firms. Expected prices, deadweight loss, and consumer costs are reported at the market level. Deadweight loss is measured as the difference in total production costs between the market power scenario and the first-best benchmark. Changes in consumer costs are also computed relative to the first-best and represent the additional payments made by consumers compared to the first-best outcome. Consumer costs are defined as the market price multiplied by total (inelastic) demand.

with long positions that soften rather than intensify competition. Our theoretical model isolates the two competing forces—strategic commitment incentives that push toward shorter positions and risk-hedging incentives that push toward longer positions—and demonstrates conditions under which the hedging motive dominates.

We provide empirical support for this mechanism from the electricity market in Alberta, Canada. We find that the vertically integrated firm maintains the longest forward position and bids least competitively among the four largest market participants, when only standard swaps are available, contradicting the predictions of standard models that treat vertical integration as pro-competitive. Critically, when more complete forward contracts become available, that allow the integrated retailer to better hedge their quantity risk, its long position shrinks substantially and its bidding behavior becomes more competitive. This

supports our theoretical prediction that contract incompleteness drives the anticompetitive effects.

The welfare effects of vertical integration under different contracts vary widely. Under the incomplete contracting regime, vertical integration increases average wholesale prices by 4.05 CAD/MWh (approximately 5 percent), with total welfare losses of 39 million CAD per year. However, introducing more complete contracts that hedge both price and quantity risk reduces wholesale prices by 2.8 CAD/MWh and generates total welfare gains of 35 million CAD per year. These findings carry important policy implications: the competitive effects of vertical integration depend critically on the design of forward markets. More broadly, our results suggest that in markets characterized by volatile prices, demand uncertainty, and convex supply curves—including agriculture, natural gas, and other commodity markets—the interaction between vertical structure and contract design may similarly influence competitive outcomes in ways not captured by existing theory.

References

AESO (2022). ATC Public Report. Technical report, Alberta Electric System Operator. Average system import capacity for 2023.

Alberta Electric System Operator (2022). AESO 2021 Annual Market Statistics. Technical report, Alberta Electric System Operator. Posted: March 8, 2022.

Alberta Energy Regulator (2025). Coal demand. Alberta Energy Outlook (ST98). Updated June 2025. Accessed February 18, 2026.

Alberta Utilities Commission (2018). Epcor energy alberta gp inc.: 2018–2021 energy price setting plan. Technical Report Decision 22357-D01-2018, Proceeding 22357, Alberta Utilities Commission, Calgary, Alberta. 16 March 2018.

Allayannis, G. and Weston, J. P. (2001). The use of foreign currency derivatives and firm market value. *The review of financial studies*, 14(1):243–276.

Allaz, B. and Vila, J.-L. (1993). Cournot competition, forward markets and efficiency. *Journal of Economic theory*, 59(1):1–16.

BDO Canada (2023). A Guide to Hedge Accounting for Private Enterprises and Not-for-Profit Organizations. Technical report, BDO Canada LLP. Accessed July 2026.

BDO Canada (2024). A Guide to Financial Instruments for Private Enterprises and Private Sector Not-for-Profit Organizations. Technical report, BDO Canada LLP. Accessed July 2026.

Beblow, J. (2017). Energy portfolio management evidence of jason beblow. Expert evidence, Office of the Utilities Consumer Advocate. In the Matter of EPCOR Energy Alberta GP Inc. 2018-2021 Energy Price Setting Plan Application, Proceeding ID 22357, Alberta Utilities Commission.

Bessembinder, H. (1991). Forward contracts and firm value: Investment incentive and contracting effects. *Journal of Financial and quantitative Analysis*, 26(4):519–532.

Bessembinder, H. and Lemmon, M. L. (2002). Equilibrium pricing and optimal hedging in electricity forward markets. *Journal of Finance*, 57(3):1347–1382.

Borenstein, S. (2007). Customer risk from real-time retail electricity pricing: Bill volatility and hedgability. *The Energy Journal*, 28(2):111–130.

Bosco, B., Parisio, L., and Pelagatti, M. (2013). Optimal pricing behavior of vertically integrated utilities: Theory and evidence from the italian electricity wholesale market. In *2013 10th International Conference on the European Energy Market (EEM)*, pages 1–7. IEEE.

Brown, D. P., Eckert, A., and Olmstead, D. E. (2022). Procurement auctions for regulated retail service contracts in restructured electricity markets. *Energy Economics*, 116:106387.

Brown, D. P., Eckert, A., and Shaffer, B. (2023). Evaluating the impact of divestitures on competition: Evidence from alberta's wholesale electricity market. *International Journal of Industrial Organization*, 89.

Brown, D. P. and Sappington, D. E. M. (2023). Load-following forward contracts. *The Energy Journal*, 44(3):187–222.

Bushnell, J. B., Mansur, E. T., and Saravia, C. (2008). Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets. *American Economic Review*, 98(1):237–266.

Carter, D. A., Rogers, D. A., and Simkins, B. J. (2006). Does hedging affect firm value? evidence from the us airline industry. *Financial management*, 35(1):53–86.

Damodaran, A. (2026). Ratings, interest coverage ratios and default spread. Data used as of January 2026.

de Castro, L., Negrete-Pincetic, M., and Gross, G. (2008). Product definition for future electricity supply auctions: The 2006 illinois experience. *The Electricity Journal*, 21(7):50–62.

Do, V., Nguyen, T. H., and Vu, T. (2024). Risk management and private debt contracts: The role of weather derivatives. *Journal of Business Finance & Accounting*, 51(9-10):2848–2883.

Dormady, N., Roa-Henriquez, A., Hoyt, M., Pesavento, M., Koenig, G., Welch, W., and Li, Z. (2025). How are retail prices formed in restructured electricity markets? *Energy Economics*, 143:108243.

Fabra, N. and Reguant, M. (2014). Pass-through of emissions costs in electricity markets. *American Economic Review*, 104(9):2872–2899.

Flottmann, J., Simshauser, P., Wild, P., and Todorova, N. (2025). The forward market dilemma in energy-only electricity markets. *Energy Economics*, 148:108676.

Frontier Economics (2017). Effects of Vertical Integration on Capacity Bidding Behaviour. Report prepared for Herbert Smith Freehills; available via the Australian Competition and Consumer Commission website.

Froot, K. A., Scharfstein, D. S., and Stein, J. C. (1993). Risk management: Coordinating corporate investment and financing policies. *the Journal of Finance*, 48(5):1629–1658.

Gans, J. S. and Wolak, F. A. (2007). A comparison of ex ante versus ex post vertical market supply: Evidence from the electricity supply industry. Working paper.

Gilje, E. P. and Taillard, J. P. (2017). Does hedging affect firm value? evidence from a natural experiment. *The Review of Financial Studies*, 30(12):4083–4132.

Hirshleifer, D. (1988). Residual risk, trading costs, and commodity futures risk premia. *Review of Financial Studies*, 1(2):173–193.

Hortaçsu, A. (2002). Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the Turkish Treasury auction market. Working paper, Department of Economics, University of Chicago.

Hortaçsu, A. and Puller, S. L. (2008). Understanding strategic bidding in multi-unit auctions: A case study of the texas electricity spot market. *The RAND Journal of Economics*, 39(1):86–114.

Hughes, J. S. and Kao, J. L. (1997). Strategic forward contracting and observability. *International Journal of Industrial Organization*, 16(1):121–133.

Independent Assessment Team (1999). Report on the determination of power purchase arrangements. Technical report, Alberta Energy and Utilities Board, Calgary, Alberta. Proceeding 990277.

Independent Competition and Regulatory Commission (2019). Report 1 of 2019: Technical paper — electricity model and methodology review 2018–19. Technical report, Independent Competition and Regulatory Commission. Report 1 of 2019.

Ito, K. and Reguant, M. (2016). Sequential markets, market power, and arbitrage. *American Economic Review*, 106(7):1921–1957.

Jha, A. and Wolak, F. A. (2023). Can forward commodity markets improve spot market performance? Evidence from wholesale electricity. *American Economic Journal: Economic Policy*, 15(2):292–330.

Joskow, P. L. (2008). Lessons learned from electricity market liberalization. *The Energy Journal*, 29(2):9–42.

Klemperer, P. D. and Meyer, M. A. (1989). Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57(6):1243–1277.

KPMG (2022). Hedge accounting: IFRS standards vs US GAAP. KPMG IFRS Institute. Accessed: 2026-06-29.

KPMG (2024). Own use scope exemption under IFRS accounting standards. KPMG IFRS Institute. Accessed: 2026-06-29.

LaCasse, C. and Wininger, T. (2007). Maryland versus new jersey: Is there a "best" competitive bid process? *The Electricity Journal*, 20(3):46–59.

Léautier, T.-O. and Rochet, J.-C. (2014). On the strategic value of risk management. *International Journal of Industrial Organization*, 37:153–169.

Longstaff, F. A. and Wang, A. W. (2004). Electricity forward prices: A high-frequency empirical analysis. *The Journal of Finance*, 59(4):1877–1900.

Loxley, C. and Salant, D. (2004). Default service auctions. *Journal of Regulatory Economics*, 26(2):201–229.

Lundin, E. and Tangerås, T. P. (2020). Cournot competition in wholesale electricity markets: The nordic power exchange, nord pool. *International Journal of Industrial Organization*, 68:102536.

MacKay, P. and Moeller, S. B. (2007). The value of corporate risk management. *The Journal of Finance*, 62(3):1379–1419.

Mansur, E. T. (2007). Upstream competition and vertical integration in electricity markets. *Journal of Law and Economics*, 50(1):125–156.

McKinnon, R. I. (1967). Futures markets, buffer stocks, and income stability for primary producers. *Journal of Political Economy*, 75(6):844–861.

Miller, N. H. and Podwol, J. U. (2020). Forward contracts, market structure and the welfare effects of mergers. *The Journal of Industrial Economics*, 68(2):364–407.

MSA (2010). An introduction to alberta's financial electricity market. Msa report, Market Surveillance Administrator.

MSA (2014). Q4/13 quarterly report: Oct.–Dec. 2013. Technical report, Market Surveillance Administrator, Calgary, Alberta. January 31, 2014.

MSA (2015). Retail market update 2015. Technical report, Market Surveillance Administrator. November 23, 2015.

MSA (2018). 2018 market share offer control report. Report, Alberta Market Surveillance Administrator.

NERA Economic Consulting (2019). Recommendations of NERA economic consulting with respect to ENMAX energy corporation's regulated rate option energy price setting plan. Technical report, NERA Economic Consulting. Prepared for ENMAX Energy Corporation, filed with Alberta Utilities Commission as part of Application 24721-A001.

NERA Economic Consulting (2025). Review of Level Playing Field Measures Options Paper. Report prepared for Meridian Energy Limited; available via the New Zealand Electricity Authority website.

NGI Staff Reports (1999). Nymex power futures anything but electric. *Natural Gas Intelligence*.

Olmstead, D. E. H. and Yatchew, A. (2025). Alberta's electricity futures market: An empirical analysis of price formation. *Energy Economics*, 144:108047.

Oum, Y. and Oren, S. (2009). Var constrained hedging of fixed price load-following obligations in competitive electricity markets. *Risk and Decision Analysis*, 1(1):43–56.

Pérez-González, F. and Yun, H. (2013). Risk management and firm value: Evidence from weather derivatives. *The Journal of Finance*, 68(5):2143–2176.

PwC (2022). Achieving hedge accounting in practice under IFRS 9. PwC Viewpoint. Accessed: 2026-06-29.

PwC (2026). Utilities and power companies guide, chapter 3.5: Hedge accounting. PwC Viewpoint. Accessed: 2026-06-29.

Reguant, M. (2014). Complementary bidding mechanisms and startup costs in electricity markets. *The Review of Economic Studies*, 81(4):1708–1742.

Rey, P. and Tirole, J. (2007). A primer on foreclosure. In Armstrong, M. and Porter, R., editors, *Handbook of Industrial Organization*, volume 3, pages 2145–2220. Elsevier, Amsterdam.

Rolfo, J. (1980). Optimal hedging under price and quantity uncertainty: The case of a cocoa producer. *Journal of Political Economy*, 88(1):100–116.

Shaffer, B. and Wolak, F. A. (2024). Market design for the 21st century: Recommendations for Alberta's power market. Research Paper 2, The School of Public Policy, University of Calgary.

Simshauser, P. (2020). Merchant utilities and boundaries of the firm: vertical integration in energy-only markets.

Smith, C. W. and Stulz, R. M. (1985). The determinants of firms' hedging policies. *Journal of financial and quantitative analysis*, 20(4):391–405.

Stulz, R. (1990). Managerial discretion and optimal financing policies. *Journal of financial Economics*, 26(1):3–27.

Thurber, M. C., Passow, F. H., Davis, T. L., and Wolak, F. A. (2022). Game-based investigation of standardized forward contracting for long-term resource adequacy. *The Electricity Journal*, 35(6):107154.

van Eijkel, R., Kuper, G. H., and Moraga-González, J. L. (2016). Do firms sell forward for strategic reasons? an application to the wholesale market for natural gas. *International Journal of Industrial Organization*, 49:1–35.

Willems, B. and Morbee, J. (2010). Market completeness: How options affect hedging and investments in the electricity sector. *Energy Economics*, 32(4):786–795.

Winzer, C., Ramírez-Molina, H., Hirth, L., and Schlecht, I. (2024). Profile contracts for electricity retail customers. *Energy Policy*, 195:114358.

Wolak, F. A. (2007). Quantifying the supply-side benefits from forward contracting in wholesale electricity markets. *Journal of Applied Econometrics*, 22(7):1179–1209.

Wolak, F. A. (2015). Measuring the competitiveness benefits of a transmission investment policy: The case of the Alberta electricity market. *Energy Policy*, 85:426–444.

Wolak, F. A. (2022). Long-term resource adequacy in wholesale electricity markets with significant intermittent renewables. In *Environmental and Energy Policy and the Economy*, volume 3, pages 155–220. University of Chicago Press. National Bureau of Economic Research.

Zhou, D. P., Dahleh, M. A., and Tomlin, C. J. (2017). Hedging strategies for load-serving entities in wholesale electricity markets. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 196–201. IEEE.

Appendix A Proofs of Results

Proof of Result 1: Risk-Neutral Firm with Market Power (Swap)

We consider a two-player game consisting of (i) a generation-only producer (firm 1) and (ii) a vertically integrated firm (firm 2) that operates both generation and retail activities. In addition, there is at least one risk-neutral speculator in the forward market.

Residual wholesale demand is given by

$$Q = a + \varepsilon - p^s,$$

or equivalently,

$$p^s(Q, \varepsilon) = a + \varepsilon - Q,$$

where a denotes the average market demand level and ε is an aggregate demand shock satisfying

$$\mathbb{E}[\varepsilon] = 0, \qquad \text{Var}(\varepsilon) = \sigma^2.$$

Residual demand is defined as inelastic demand a minus the quantity supplied by fringe firms, which are willing to produce at their capacity as long as the market price p exceeds their marginal cost. The shock enters through the aggregate inelastic demand level.

The VI firm's retail demand is

$$q_2^r = \theta + e,$$

where θ is the average retail demand level and e is a retail demand shock with

$$\mathbb{E}[e] = 0, \qquad \text{Var}(e) = \sigma_e^2.$$

Note that in the main text we focus on the special case $e = \beta\varepsilon$. In this appendix, we instead characterize the equilibrium forward positions for a general realization of e.

The VI firm's generation cost is quadratic:

$$C_2(q_2) = bq_2 + \frac{q_2^2}{2k_2},$$

where b is the marginal cost intercept and $k_2 > 0$ governs cost curvature.

Profit. Both firms trade forward contracts. Let f_i denote firm i's forward position, p^f the forward price, and p^s the spot price.

Producer (firm 1):

$$\Pi_1 = \underbrace{p^f f_1 - p^s f_1}_{\text{Forward profit}} + \underbrace{p^s q_1 - C_1(q_1)}_{\text{Generation profit}}.$$

Vertically integrated firm (firm 2):

$$\Pi_2 = \underbrace{p^r q_2^r - p^s q_2^r}_{\text{Retail profit}} + \underbrace{p^f f_2 - p^s f_2}_{\text{Forward profit}} + \underbrace{p^s q_2 - C_2(q_2)}_{\text{Generation profit}}.$$

Retail profit reflects the margin between the retail price p^r and the wholesale procurement cost p^s applied to realized retail demand q_2^r. Forward profit corresponds to settlement at p^f with spot adjustment at p^s. Generation profit equals wholesale revenue net of production cost.

Timing. Firms first choose forward positions (f_1, f_2). Uncertainty (ε, e) is realized. Firms then choose generation quantities (q_1, q_2) in the spot market. The spot price p^s clears the wholesale market.

Spot Market Fix forward positions $\{f_\ell\}$.

1. Producer.

The producer's spot-stage profit is

$$p^s q_1 - C_1(q_1) + (p^f - p^s) f_1 = p^s(q_1 - f_1) - b q_1 - \frac{q_1^2}{2k_1} + p^f f_1.$$

Using $\partial p^s / \partial q_1 = -1$, the first-order condition is

$$\frac{\partial \Pi_1}{\partial q_1} = p^s - (q_1 - f_1) - b - \frac{1}{k_1} q_1 = 0.$$

Rearranging yields

$$q_1^* \left(1 + \frac{1}{k_1}\right) = p^s - b + f_1,$$

so that

$$q_1^* = \frac{p^s - b + f_1}{1 + \frac{1}{k_1}}.$$

2. Vertically Integrated Firm.

The VI firm's spot-stage profit is

$$\Pi_2 = p^r q_2^r - p^s q_2^r + p^f f_2 - p^s f_2 + p^s q_2 - b q_2 - \frac{q_2^2}{2k_2}.$$

Terms independent of q_2 can be omitted. Using $\partial p^s / \partial q_2 = -1$, the first-order condition is

$$\frac{\partial \Pi_2}{\partial q_2} = p^s - (q_2 - q_2^r - f_2) - b - \frac{1}{k_2} q_2 = 0,$$

which can be written as

$$q_2^r + f_2 - q_2 + p^s - b - \frac{1}{k_2} q_2 = 0.$$

Rearranging gives

$$q_2^* \left(1 + \frac{1}{k_2}\right) = p^s - b + q_2^r + f_2,$$

and therefore

$$q_2^* = \frac{p^s - b + q_2^r + f_2}{1 + \frac{1}{k_2}}.$$

Market Clearing. For the remainder, impose symmetry in marginal-cost curvature, $k_\ell = k$ for all ℓ, and define

$$\delta := \frac{1}{1 + \frac{1}{k}}.$$

Substituting the best responses and summing outputs yields aggregate supply

$$Q^* = q_1^* + q_2^* = \delta\big(2p^s - 2b + f_1 + f_2 + q_2^r\big).$$

Market clearing requires $Q^* = a + \varepsilon - p^s$. Hence,

$$\delta\big(2p^s - 2b + f_1 + f_2 + q_2^r\big) = a + \varepsilon - p^s.$$

Collecting terms in p^s gives

$$(2\delta + 1)p^s = a + \varepsilon + \delta\big(2b - f_1 - f_2 - q_2^r\big),$$

so that the equilibrium spot price is

$$p^s = \frac{a + \varepsilon + \delta(2b - f_1 - f_2 - q_2^r)}{2\delta + 1}.$$

Individual Outputs and Forward Price. Substituting the equilibrium spot price into the firms' best responses yields output expressions.

For the VI firm,

$$q_2^* = \delta\big(p^s - b + q_2^r + f_2\big) = \frac{\big[a - b + \varepsilon - \delta f_1 + (\delta + 1)f_2 + (\delta + 1)q_2^r\big]\delta}{2\delta + 1}.$$

For the standalone producer,

$$q_1^* = \delta\big(p^s - b + f_1\big) = \frac{\big[a - b + \varepsilon - \delta f_2 + (\delta + 1)f_1 - \delta q_2^r\big]\delta}{2\delta + 1}.$$

Finally, competitive risk-neutral speculators set the forward price equal to the expected spot price,

$$p^f = \mathbb{E}[p^s] = \mathbb{E}\left[\frac{a + \varepsilon + \delta(2b - f_1 - f_2 - q_2^r)}{2\delta + 1}\right].$$

Using $\mathbb{E}[\varepsilon] = 0$ and $\mathbb{E}[q_2^r] = \theta$, this simplifies to

$$p^f = \frac{a + \delta(2b - f_1 - f_2 - \theta)}{2\delta + 1}.$$

The equilibrium spot price can be written as

$$p^s = \frac{a + \varepsilon + \delta(2b - f_2 - f_1 - \theta - e)}{2\delta + 1} \tag{12}$$

$$= \underbrace{\frac{a + \delta(2b - f_2 - f_1 - \theta)}{2\delta + 1}}_{=:\,p_0} + \underbrace{\frac{1}{2\delta + 1}}_{=:\,p_1}\varepsilon + \underbrace{\frac{-\delta}{2\delta + 1}}_{=:\,p_2} e. \tag{13}$$

The producer's equilibrium output is

$$q_1^* = \frac{\big[a - b + \varepsilon - \delta f_2 + (\delta + 1)f_1 - \delta\theta - \delta e\big]\delta}{2\delta + 1} \tag{14}$$

$$= \underbrace{\frac{\big[a - b - \delta f_2 + (\delta + 1)f_1 - \delta\theta\big]\delta}{2\delta + 1}}_{=:\,q_{1,0}} + \underbrace{\frac{\delta}{2\delta + 1}}_{=:\,q_{1,1}}\varepsilon + \underbrace{\frac{-\delta^2}{2\delta + 1}}_{=:\,q_{1,2}} e. \tag{15}$$

The VI firm's equilibrium output is

$$q_2^* = \frac{\big[a - b + \varepsilon - \delta f_1 + (\delta+1) f_2 + (\delta+1)\theta + (\delta+1)e\big]\delta}{2\delta+1} \tag{16}$$

$$= \underbrace{\frac{\big[a - b - \delta f_1 + (\delta+1) f_2 + (\delta+1)\theta\big]\delta}{2\delta+1}}_{=:\, q_{2,0}} + \underbrace{\frac{\delta}{2\delta+1}}_{=:\, q_{2,1}}\varepsilon + \underbrace{\frac{(\delta+1)\delta}{2\delta+1}}_{=:\, q_{2,2}} e. \tag{17}$$

For $\ell \in \{1, 2\}$, the price derivatives are

$$\frac{\partial p_0}{\partial f_\ell} = -\frac{\delta}{2\delta+1}, \qquad \frac{\partial p_1}{\partial f_\ell} = 0, \qquad \frac{\partial p_2}{\partial f_\ell} = 0.$$

For the producer,

$$\frac{\partial q_{1,0}}{\partial f_1} = \frac{(\delta+1)\delta}{2\delta+1}, \qquad \frac{\partial q_{1,1}}{\partial f_1} = 0, \qquad \frac{\partial q_{1,2}}{\partial f_1} = 0.$$

For the VI firm,

$$\frac{\partial q_{2,0}}{\partial f_2} = \frac{(\delta+1)\delta}{2\delta+1}, \qquad \frac{\partial q_{2,1}}{\partial f_2} = 0, \qquad \frac{\partial q_{2,2}}{\partial f_2} = 0.$$

Forward-Stage Optimization. Because $p^f = \mathbb{E}[p^s] = p_0$, the forward payoff satisfies

$$(p^f - p^s) f_\ell = -(p_1\varepsilon + p_2 e) f_\ell = \frac{-\varepsilon + \delta e}{2\delta+1} f_\ell,$$

so that

$$\mathbb{E}[(p^f - p^s) f_\ell] = 0.$$

Under risk neutrality, producers' forward positions affect expected profits only through their impact on spot market outcomes.

Producer. The producer's expected profit is

$$\mathbb{E}[\Pi_1] = \mathbb{E}\left[p^s q_1^* - \left(b q_1^* + \frac{q_1^{*2}}{2k}\right)\right].$$

Since $(p_1, q_{1,1}, p_2, q_{1,2})$ are independent of f_1, the first-order condition reduces to

$$\frac{\partial \mathbb{E}[\Pi_1]}{\partial f_1} = \frac{\partial p_0}{\partial f_1} q_{1,0} + \frac{\partial q_{1,0}}{\partial f_1} p_0 - b\frac{\partial q_{1,0}}{\partial f_1} - \frac{1}{k} q_{1,0} \frac{\partial q_{1,0}}{\partial f_1} = 0.$$

VI Firm. As discussed in the main text, and following Bushnell et al. (2008), we assume that expected retail profits are zero, with retail rates set to reflect expected wholesale prices and retail demand. Hence, $\mathbb{E}[p^r q_2^r - p^s q_2^r] = 0$. Consequently, the VI firm's forward choice affects expected profits only through its impact on spot-market outcomes.

$$\mathbb{E}[\Pi_2] = \mathbb{E}\left[p^s q_2^* - \left(b q_2^* + \frac{q_2^{*2}}{2k}\right)\right],$$

Since $(p_1, q_{2,1}, p_2, q_{2,2})$ are independent of f_2, the first-order condition reduces to

$$\frac{\partial \mathbb{E}[\Pi_2]}{\partial f_2} = \frac{\partial p_0}{\partial f_2} q_{2,0} + \frac{\partial q_{2,0}}{\partial f_2} p_0 - b\frac{\partial q_{2,0}}{\partial f_2} - \frac{1}{k} q_{2,0} \frac{\partial q_{2,0}}{\partial f_2} = 0.$$

Solving the two first-order conditions yields

$$f_1^* = \frac{(a-b)\big(k - \delta(\delta+1)\big)}{\delta\big(2\delta k + \delta + 3k + 1\big)} \qquad \text{and} \qquad f_2^* = f_1^* - \theta.$$

Thus the VI firm sells θ fewer forward units than the standalone producer, consistent with the interpretation that the VI firm's retail position partially substitutes for forward sales. Recall that

$$\delta = \frac{1}{1 + \frac{1}{k}} = \frac{k}{k+1}.$$

Substituting this expression into the general solution for f_1^* yields

$$f_1^* = \frac{k^2}{5k^2 + 5k + 1}(a - b).$$

Therefore, in the notation of Result 1,

$$f_{1,N}^* = \frac{k^2}{5k^2 + 5k + 1}(a - b), \qquad f_{2,N}^* = f_{1,N}^* - \theta.$$

This completes the proof of Result 1.

Proof of Result 2: Risk-Averse,Price-Taking Firm (Swap)

We now consider price-taking firms under swap.

Profit. Firm profits are given by

Producer (firm 1):

$$\Pi_1 = \underbrace{p^f f_1 - p^s f_1}_{\text{Forward Profit}} + \underbrace{p^s q_1 - C_1(q_1)}_{\text{Generation Profit}}.$$

Vertically integrated firm (firm 2):

$$\Pi_2 = \underbrace{p^r q_2^r - p^s q_2^r}_{\text{Retail Profit}} + \underbrace{p^f f_2 - p^s f_2}_{\text{Forward Profit}} + \underbrace{p^s q_2 - C_2(q_2)}_{\text{Generation Profit}}.$$

Demand and retail quantities. Aggregate demand is

$$Q(\varepsilon) = a + \varepsilon, \qquad \mathbb{E}[\varepsilon] = 0, \qquad \text{Var}(\varepsilon) = \sigma^2.$$

The VI firm's retail demand is

$$q_2^r = \theta + e, \qquad \mathbb{E}[e] = 0, \qquad \text{Var}(e) = \sigma_e^2.$$

Each firm has quadratic cost

$$C_i(q_i) = bq_i + \frac{q_i^2}{2k_i}.$$

Spot market Firm 1 and firm 2 are price takers in the spot market; conditional on p^s they solve

$$\max_{q_i} \ \Pi_i^s = p^s q_i - C_i(q_i) = p^s q_i - bq_i - \frac{q_i^2}{2k_i}.$$

The first-order condition yields

$$p^s - b - \frac{q_i}{k_i} = 0 \qquad \Longleftrightarrow \qquad q_i = k_i(p^s - b).$$

Assuming symmetric technologies in the spot market (so $k_1 = k_2 = k$) and using $n = 2$,

market clearing implies

$$a + \varepsilon = nk(p^s - b) \qquad \Longleftrightarrow \qquad p^s = \frac{a+\varepsilon}{nk} + b = \frac{a+\varepsilon}{2k} + b.$$

Therefore,

$$q_i = \frac{a+\varepsilon}{n} = \frac{a+\varepsilon}{2}.$$

Arbitrage With one risk-neutral speculator, the forward price satisfies

$$p^f = \mathbb{E}[p^s] = \frac{a}{2k} + b.$$

Hence,

$$p^f - p^s = -\frac{\varepsilon}{2k}.$$

Retail pricing As discussed in the main text, and following Bushnell et al. (2008), we assume that expected retail profits are zero, with retail rates set to reflect expected wholesale prices and retail demand. Hence,

$$\mathbb{E}[p^r q_2^r - p^s q_2^r] = 0$$

.

Since p^r is chosen prior to the realization of (ε, e) and $\mathbb{E}[e] = 0$, we have

$$p^r = \mathbb{E}[p^s] + \frac{\mathbb{E}[p^s e]}{\theta}$$

Thus,

$$p^r = b + \frac{a}{2k} + \frac{\mathbb{E}[\varepsilon e]}{2\theta k},$$

Forward market Using the spot-market expressions, the relevant payoffs are:

Producer (firm 1):

$$\Pi_1 = (p^f - p^s)f_1 + p^s q_1 - bq_1 - \frac{q_1^2}{2k} = -\frac{\varepsilon}{2k}f_1 + \left(\frac{a+\varepsilon}{2k} + b\right)\frac{a+\varepsilon}{2} - b\frac{a+\varepsilon}{2} - \frac{(a+\varepsilon)^2}{8k}.$$

Simplifying,

$$\Pi_1 = -\frac{\varepsilon}{2k} f_1 + \frac{(a+\varepsilon)^2}{8k}.$$

VI (firm 2):

$$\Pi_2 = p^r(\theta + e) - \left(\frac{a+\varepsilon}{nk} + b\right)(\theta + e) + (p^f - p^s) f_2 + \frac{(a+\varepsilon)^2}{8k}.$$

Using $p^f - p^s = -\varepsilon/(2k)$,

$$\Pi_2 = \left(p^r - b - \frac{a+\varepsilon}{2k}\right)(\theta + e) - \frac{\varepsilon}{2k} f_2 + \frac{(a+\varepsilon)^2}{8k}$$

Moments and optimal hedge For the producer (firm 1), expected profit is

$$\mathbb{E}[\Pi_1] = \mathbb{E}\left[\frac{(a+\varepsilon)^2}{8k} - \frac{\varepsilon}{2k} f_1\right].$$

Since $\mathbb{E}[\varepsilon] = 0$, expected profit is independent of f_1.

Keeping only the random components in Π_1:

$$\Pi_1 = -\frac{\varepsilon}{2k} f_1 + \frac{\varepsilon^2}{8k} + \frac{a\varepsilon}{4k} \; + \; \text{(non-random terms)}.$$

Equivalently,

$$\Pi_1 = \varepsilon\left[-\frac{f_1}{2k} + \frac{a}{4k}\right] + \frac{\varepsilon^2}{8k} \; + \; \text{(non-random terms)}.$$

Thus,

$$\mathrm{Var}[\Pi_1] = \left(-\frac{f_1}{2k} + \frac{a}{4k}\right)^2 \sigma^2 + \frac{1}{64k^2}\left(\mathbb{E}[\varepsilon^4] - \sigma^4\right) + 2\left(-\frac{f_1}{2k} + \frac{a}{4k}\right)\frac{1}{8k}\,\mathrm{Cov}(\varepsilon^2, \varepsilon).$$

The f_1-dependent terms are

$$\frac{\sigma^2 f_1^2}{4k^2} - \frac{2\sigma^2 f_1}{2k}\left(\frac{a}{4k}\right) - \frac{f_1}{8k^2}\,\mathbb{E}[\varepsilon^3].$$

The first-order condition is

$$2 \cdot \frac{\sigma^2 f_1}{4k^2} - \frac{2\sigma^2}{2k}\left(\frac{a}{4k}\right) - \frac{1}{8k^2}\,\mathbb{E}[\varepsilon^3] = 0,$$

so

$$f_1^* = \frac{a}{2} + \frac{\mathbb{E}[\varepsilon^3]}{4\sigma^2}.$$

For the VI firm (firm 2), profit is

$$\Pi_2 = \left(p^r - b - \frac{a+\varepsilon}{2k}\right)(\theta + e) - \frac{\varepsilon}{2k} f_2 + \frac{(a+\varepsilon)^2}{8k}.$$

Expected profit is

$$\mathbb{E}[\Pi_2] = \mathbb{E}\left[\left(p^r - b - \frac{a+\varepsilon}{2k}\right)(\theta + e) - \frac{\varepsilon}{2k} f_2 + \frac{(a+\varepsilon)^2}{8k}\right]$$

Since $\mathbb{E}[\varepsilon] = 0$, expected profit is independent of f_1.

Keeping only random components,

$$\Pi_2 = -\frac{\varepsilon e}{2k} + e\left(p^r - b - \frac{a}{2k}\right) - \frac{\theta}{2k}\varepsilon - \frac{\varepsilon}{2k} f_2 + \frac{\varepsilon^2}{8k} + \frac{a\varepsilon}{4k} \; + \; \text{(non-random terms)}.$$

Collecting the ε-linear term,

$$\Pi_2 = -\frac{\varepsilon e}{2k} + e\left(p^r - b - \frac{a}{2k}\right) + \varepsilon\left[\frac{a}{4k} - \frac{\theta + f_2}{2k}\right] + \frac{\varepsilon^2}{8k} \; + \; \text{(non-random terms)}.$$

Therefore, $\mathrm{Var}[\Pi_2]$ is a quadratic function of f_2. The f_2-dependent terms can be written as

$$\frac{\sigma^2 f_2^2}{4k^2} - \frac{\sigma^2 f_2}{k}\left(\frac{a}{4k} - \frac{\theta}{2k}\right) + \frac{f_2}{2k^2}\,\mathrm{Cov}(\varepsilon e, \varepsilon) - \left(p^r - b - \frac{a}{2k}\right)\frac{f_2}{k}\,\mathrm{Cov}(\varepsilon, e) - \frac{f_2}{8k^2}\,\mathrm{Cov}(\varepsilon^2, \varepsilon).$$

Using $\mathrm{Cov}(\varepsilon e, \varepsilon) = \mathbb{E}[\varepsilon^2 e]$, $\mathrm{Cov}(\varepsilon, e) = \mathbb{E}[\varepsilon e]$, and $\mathrm{Cov}(\varepsilon^2, \varepsilon) = \mathbb{E}[\varepsilon^3]$, this becomes

$$\frac{\sigma^2 f_2^2}{4k^2} - \frac{\sigma^2 f_2}{k}\left(\frac{a}{4k} - \frac{\theta}{2k}\right) + \frac{f_2}{2k^2}\,\mathbb{E}[\varepsilon^2 e] - \left(p^r - b - \frac{a}{2k}\right)\frac{f_2}{k}\,\mathbb{E}[\varepsilon e] - \frac{f_2}{8k^2}\,\mathbb{E}[\varepsilon^3].$$

The first-order condition is

$$2 \cdot \frac{\sigma^2 f_2}{4k^2} - \frac{\sigma^2}{k}\left(\frac{a}{4k} - \frac{\theta}{2k}\right) + \frac{1}{2k^2}\,\mathbb{E}[\varepsilon^2 e] - \left(p^r - b - \frac{a}{2k}\right)\frac{1}{k}\,\mathbb{E}[\varepsilon e] - \frac{1}{8k^2}\,\mathbb{E}[\varepsilon^3] = 0,$$

which implies

$$f_2^* = \frac{a}{2} - \theta - \frac{\mathbb{E}[\varepsilon^2 e]}{\sigma^2} + 2k\left(p^r - b - \frac{a}{2k}\right)\frac{\mathbb{E}[\varepsilon e]}{\sigma^2} + \frac{\mathbb{E}[\varepsilon^3]}{4\sigma^2}.$$

We know

$$p^r = b + \frac{a}{2k} + \frac{\mathbb{E}[\varepsilon e]}{2\theta k},$$

so that

$$2k\left(p^r - b - \frac{a}{2k}\right) = \frac{\mathbb{E}[\varepsilon e]}{\theta}.$$

Substituting yields

$$f_2^* = \frac{a}{2} - \theta - \frac{\mathbb{E}[\varepsilon^2 e]}{\sigma^2} + \frac{\mathbb{E}[\varepsilon e]^2}{\theta\sigma^2} + \frac{\mathbb{E}[\varepsilon^3]}{4\sigma^2}.$$

Approximation for the swap benchmark f_2^* (VI). Introduce normalized shocks

$$y := \frac{\varepsilon}{a}, \qquad x := \frac{e}{\theta}, \qquad v := \mathbb{E}[y^2] = \frac{\sigma^2}{a^2},$$

and the (standardized) mixed moments

$$m_{11} := \mathbb{E}[xy] = \frac{\mathbb{E}[\varepsilon e]}{a\theta}, \qquad m_{21} := \mathbb{E}[y^2 x] = \frac{\mathbb{E}[\varepsilon^2 e]}{a^2\theta}, \qquad m_{30} := \mathbb{E}[y^3] = \frac{\mathbb{E}[\varepsilon^3]}{a^3}.$$

The variance-minimizing *swap* position for the vertically integrated firm can be represented in normalized moments as

$$f_2^* = \frac{a}{2} - \theta - \theta\,\frac{m_{21}}{v} + \theta\,\frac{m_{11}^2}{v} + \frac{a}{4}\,\frac{m_{30}}{v}. \tag{18}$$

Small-noise ordering (what is first order). Define the small-noise scalings

$$\delta := \frac{\sigma}{a}, \qquad \eta := \frac{\sigma_e}{\theta},$$

and assume bounded standardized moments (so skewness, kurtosis are $O(1)$).

By Cauchy–Schwarz,

$$|m_{11}| \le \eta\delta, \qquad |m_{21}| \le O(\delta^2\eta), \qquad m_{30} = \gamma\delta^3 = O(\delta^3)$$

We also know $v = \delta^2$.

Then

$$\frac{m_{21}}{v} = O(\eta), \qquad \frac{m_{11}^2}{v} = O(\eta^2),, \qquad \frac{m_{30}}{v} = O(\delta),$$

If one wants a *first-order* approximation in (δ, η) only, drop the $O(\eta^2)$ term:

$$f_2^* = \frac{a}{2} + \frac{\mathbb{E}[\varepsilon^3]}{4\sigma^2} - \theta - \frac{\mathbb{E}[\varepsilon^2 e]}{\sigma^2} \; + \; O(\eta^2). \tag{19}$$

Using the derived result for f_1^*, we can express

$$f_2^* = f_1^* - \theta - \frac{\mathbb{E}[\varepsilon^2 e]}{\sigma^2} \; + \; O(\eta^2).$$

Therefore we have

$$f_1^* = \frac{a}{2} + \frac{\mathbb{E}[\varepsilon^3]}{4\sigma^2}, \qquad f_2^* \approx f_1^* - \theta - \frac{\mathbb{E}[\varepsilon^2 e]}{\sigma^2}$$

Finally, substituting

$$\mathbb{E}[q_i^*] = \frac{a}{2} \quad \text{and} \quad e = \beta\varepsilon$$

completes the proof of Result 2.

Proof of Result 3: Risk-Neutral Firms with Market Power (LFFC)

We now consider risk-neutral firms with market power under load-following forward contracts (LFFC). Under LFFC, firm i chooses a hedge ratio l_i that is applied to realized retail demand, $q_2^r = \theta + e$. Forward payoffs therefore scale with realized load rather than with a fixed contracted quantity.

Profit. Firm profits are given by
Producer (firm 1):

$$\Pi_1 = p^f l_1(\theta + e) - p^s l_1(\theta + e) + p^s q_1 - bq_1 - \frac{q_1^2}{2k_1}.$$

Vertically integrated firm (firm 2):

$$\Pi_2 = p^r q_2^r - p^s q_2^r + p^f l_2(\theta + e) - p^s l_2(\theta + e) + p^s q_2 - bq_2 - \frac{q_2^2}{2k_2}.$$

Relative to the swap case, the key difference is that forward positions now vary with realized retail demand $(\theta + e)$, so that spot-price exposure depends on the load shock e.

Spot Market Fix hedge ratios $\{l_i\}$.

Producer (firm 1).

Given l_1, the producer solves

$$\max_{q_1} \; p^s q_1 - bq_1 - \frac{q_1^2}{2k_1} + (p^f - p^s) l_1(\theta + e).$$

Since p^f is constant in the spot stage, this is equivalent to

$$\max_{q_1} \; p^s q_1 - bq_1 - \frac{q_1^2}{2k_1} - p^s l_1(\theta + e).$$

Using $\partial p^s / \partial q_1 = -1$, the first-order condition is

$$p^s - (q_1 - l_1(\theta + e)) - b - \frac{1}{k_1} q_1 = 0.$$

Rearranging,

$$q_1^* \left(1 + \frac{1}{k_1}\right) = p^s - b + l_1(\theta + e),$$

so that

$$q_1^* = \frac{p^s - b + l_1(\theta + e)}{1 + \frac{1}{k_1}}.$$

VI firm (firm 2).

Given l_2, the VI firm solves

$$\max_{q_2} \; (p^r q_2^r - p^s q_2^r) + p^s q_2 - bq_2 - \frac{q_2^2}{2k_2} + (p^f - p^s) l_2(\theta + e).$$

Dropping terms independent of q_2,

$$\max_{q_2} \; p^s q_2 - bq_2 - \frac{q_2^2}{2k_2} - p^s q_2^r - p^s l_2(\theta + e).$$

Using $\partial p^s / \partial q_2 = -1$, the first-order condition is

$$p^s - (q_2 - q_2^r - l_2(\theta + e)) - b - \frac{1}{k_2} q_2 = 0.$$

Rearranging,

$$q_2^* \left(1 + \frac{1}{k_2}\right) = p^s - b + q_2^r + l_2(\theta + e).$$

Since $q_2^r = \theta + e$, this becomes

$$q_2^* = \frac{p^s - b + (1 + l_2)(\theta + e)}{1 + \frac{1}{k_2}}.$$

Market Clearing. Impose symmetric curvature $k_i = k$ and define

$$\delta := \frac{1}{1 + \frac{1}{k}}.$$

Aggregating across firms gives

$$\begin{aligned} Q^* &= q_1^* + q_2^* \\ &= \delta\big(p^s - b + l_1(\theta + e) + p^s - b + (1 + l_2)(\theta + e)\big) \\ &= \delta\big(2p^s - 2b + (1 + l_1 + l_2)(\theta + e)\big). \end{aligned}$$

Market demand is

$$Q^s = a + \varepsilon - p^s.$$

Equating supply and demand yields

$$\delta\big(2p^s - 2b + (1 + l_1 + l_2)(\theta + e)\big) = a + \varepsilon - p^s$$

Collecting terms in p^s gives

$$(2\delta + 1)p^s = a + \varepsilon + \delta\big(2b - (1 + l_1 + l_2)(\theta + e)\big).$$

Hence the equilibrium spot price is

$$p^s = \frac{a + \varepsilon + \delta\big(2b - (1 + l_1 + l_2)(\theta + e)\big)}{2\delta + 1}$$

Individual Outputs. Substituting the equilibrium spot price into the firms' output choices yields expressions for equilibrium output.

VI firm (firm 2).

$$q_2^* = \frac{\delta}{2\delta+1}\Big[a - b + \varepsilon - \delta l_1(\theta+e) + (\delta+1)l_2(\theta+e) + (\delta+1)q_2^r\Big].$$

Producer (firm 1).

$$q_1^* = \frac{\delta}{2\delta+1}\Big[a - b + \varepsilon - \delta l_2(\theta+e) + (\delta+1)l_1(\theta+e) - \delta q_2^r\Big].$$

Speculators. Risk-neutral competitive speculators imply zero expected profits on load-following forward contracts:

$$\mathbb{E}\big[p^f l_i(\theta+e) - p^s l_i(\theta+e)\big] = 0.$$

Spot Price and Outputs. The equilibrium spot price becomes

$$\begin{aligned} p^s &= \frac{a + \varepsilon + \delta\big(2b - (l_1+l_2+1)(\theta+e)\big)}{2\delta+1} \\ &= \frac{a + \delta\big(2b - (l_1+l_2+1)\theta\big)}{2\delta+1} + \frac{1}{2\delta+1}\varepsilon - \frac{\delta(l_1+l_2+1)}{2\delta+1}e \\ &=: p_0 + p_1\varepsilon + p_2 e. \end{aligned}$$

The producer's equilibrium output is

$$\begin{aligned} q_1^* &= \frac{\delta}{2\delta+1}\Big[a - b + \varepsilon - \delta l_2(\theta+e) + (\delta+1)l_1(\theta+e) - \delta(\theta+e)\Big] \\ &= \underbrace{\frac{\delta}{2\delta+1}\Big[a - b - \delta l_2\theta + (\delta+1)l_1\theta - \delta\theta\Big]}_{=:q_{1,0}} + \underbrace{\frac{\delta}{2\delta+1}}_{=:q_{1,1}}\varepsilon \\ &\quad + \underbrace{\frac{\delta\big[-\delta l_2 + (\delta+1)l_1 - \delta\big]}{2\delta+1}}_{=:q_{1,2}} e. \end{aligned}$$

The VI firm's equilibrium output is

$$\begin{aligned} q_2^* &= \frac{\delta}{2\delta+1}\Big[a-b+\varepsilon-\delta l_1(\theta+e)+(\delta+1)l_2(\theta+e)+(\delta+1)(\theta+e)\Big] \\ &= \underbrace{\frac{\delta}{2\delta+1}\Big[a-b-\delta l_1\theta+(\delta+1)l_2\theta-\delta\theta\Big]+\theta\delta}_{=:q_{2,0}}+\underbrace{\frac{\delta}{2\delta+1}}_{=:q_{2,1}}\varepsilon \\ &\quad+\left(\underbrace{\frac{\delta\big[-\delta l_1+(\delta+1)l_2-\delta\big]}{2\delta+1}+\delta}_{=:q_{2,2}}\right)e \end{aligned}$$

Derivatives. For $m \in \{1,2\}$,

$$\frac{\partial p_0}{\partial l_m} = -\frac{\delta\theta}{2\delta+1}, \qquad \frac{\partial p_1}{\partial l_m} = 0, \qquad \frac{\partial p_2}{\partial l_m} = -\frac{\delta}{2\delta+1}.$$

For the producer,

$$\frac{\partial q_{1,0}}{\partial l_1} = \frac{(\delta+1)\delta\theta}{2\delta+1}, \qquad \frac{\partial q_{1,1}}{\partial l_1} = 0, \qquad \frac{\partial q_{1,2}}{\partial l_1} = \frac{(\delta+1)\delta}{2\delta+1}.$$

For the VI firm,

$$\frac{\partial q_{2,0}}{\partial l_2} = \frac{(\delta+1)\delta\theta}{2\delta+1}, \qquad \frac{\partial q_{2,1}}{\partial l_2} = 0, \qquad \frac{\partial q_{2,2}}{\partial l_2} = \frac{(\delta+1)\delta}{2\delta+1}.$$

Producer Choice of l_1. Again, by the full-arbitrage condition, expected forward profits are zero. Hence, the producer's expected profit reduces to

$$\mathbb{E}[\Pi_1] = \mathbb{E}\left[p^s q_1^* - \left(bq_1^* + \frac{q_1^{*2}}{2k_1}\right)\right].$$

Using the decompositions

$$p^s = p_0 + p_1\varepsilon + p_2 e, \qquad q_1^* = q_{1,0} + q_{1,1}\varepsilon + q_{1,2}e,$$

we obtain

$$\mathbb{E}[\Pi_1] = \mathbb{E}\Big[(p_0 + p_1\varepsilon + p_2 e)(q_{1,0} + q_{1,1}\varepsilon + q_{1,2}e) \\ - b(q_{1,0} + q_{1,1}\varepsilon + q_{1,2}e) - \frac{1}{2k_1}(q_{1,0} + q_{1,1}\varepsilon + q_{1,2}e)^2\Big].$$

Since $\mathbb{E}[\varepsilon] = \mathbb{E}[e] = 0$, only second moments remain after taking expectations.

Differentiating with respect to l_1 and using

$$\frac{\partial p_1}{\partial l_1} = \frac{\partial q_{1,1}}{\partial l_1} = 0,$$

yields

$$\frac{\partial \mathbb{E}[\Pi_1]}{\partial l_1} = q_{1,0}\left(\frac{\partial p_0}{\partial l_1} - \frac{1}{k_1}\frac{\partial q_{1,0}}{\partial l_1}\right) + \frac{\partial q_{1,0}}{\partial l_1}(p_0 - b) \\ + \left[q_{1,2}\left(\frac{\partial p_2}{\partial l_1} - \frac{1}{k_1}\frac{\partial q_{1,2}}{\partial l_1}\right) + \frac{\partial q_{1,2}}{\partial l_1}p_2\right]\sigma_e^2 \\ + \left[q_{1,1}\left(\frac{\partial p_2}{\partial l_1} - \frac{1}{k_1}\frac{\partial q_{1,2}}{\partial l_1}\right) + \frac{\partial q_{1,2}}{\partial l_1}p_1\right]\mathbb{E}[\varepsilon e].$$

The first line reflects the effect of l_1 on deterministic profits. The second and third lines capture, respectively, the variance and covariance channels arising from the load shock e and its interaction with ε.

Setting

$$\frac{\partial \mathbb{E}[\Pi_1]}{\partial l_1} = 0$$

characterizes the producer's best response l_1^*.

VI Choice of l_2. By the full-arbitrage condition and competitive retail pricing, expected forward and retail profits are zero. Therefore, the VI firm's expected profit can be written as

$$\mathbb{E}[\Pi_2] = \mathbb{E}\left[p^s q_2 - bq_2 - \frac{q_2^2}{2k_2}\right].$$

The hedge ratio l_2 affects $\mathbb{E}[\Pi_2]$ only through the equilibrium spot price and output.

Using

$$p^s = p_0 + p_1\varepsilon + p_2 e, \qquad q_2^* = q_{2,0} + q_{2,1}\varepsilon + q_{2,2}e,$$

we obtain

$$\begin{aligned}\mathbb{E}[\Pi_2] = \mathbb{E}\Big[&(p_0 + p_1\varepsilon + p_2 e)(q_{2,0} + q_{2,1}\varepsilon + q_{2,2}e) \\ &- b(q_{2,0} + q_{2,1}\varepsilon + q_{2,2}e) - \frac{1}{2k_2}(q_{2,0} + q_{2,1}\varepsilon + q_{2,2}e)^2\Big].\end{aligned}$$

Differentiating with respect to l_2 and using

$$\frac{\partial p_1}{\partial l_2} = \frac{\partial q_{2,1}}{\partial l_2} = 0,$$

yields the first-order condition

$$\begin{aligned}0 = {} & q_{2,0}\left(\frac{\partial p_0}{\partial l_2} - \frac{1}{k_2}\frac{\partial q_{2,0}}{\partial l_2}\right) + \frac{\partial q_{2,0}}{\partial l_2}(p_0 - b) \\ & + \left[q_{2,2}\left(\frac{\partial p_2}{\partial l_2} - \frac{1}{k_2}\frac{\partial q_{2,2}}{\partial l_2}\right) + \frac{\partial q_{2,2}}{\partial l_2}p_2\right]\sigma_e^2 \\ & + \left[q_{2,1}\left(\frac{\partial p_2}{\partial l_2} - \frac{1}{k_2}\frac{\partial q_{2,2}}{\partial l_2}\right) + \frac{\partial q_{2,2}}{\partial l_2}p_1\right]\mathbb{E}[\varepsilon e].\end{aligned}$$

Relative to the producer's condition, the VI first-order condition contains additional terms embedded in $q_{2,0}$ and $q_{2,2}$ that arise from the retail position. In particular, writing

$$q_{2,0} = q_{1,0} + \theta\delta, \qquad q_{2,2} = q_{1,2} + \delta,$$

Substituting into the VI first-order condition shows that it can be written as the producer first-order condition plus the incremental term

$$\delta\theta\left(\frac{\partial p_0}{\partial l_2} - \frac{1}{k_2}\frac{\partial q_{2,0}}{\partial l_2}\right) + \delta\left(\frac{\partial p_2}{\partial l_2} - \frac{1}{k_2}\frac{\partial q_{2,2}}{\partial l_2}\right)\sigma_e^2.$$

Using

$$\frac{\partial p_0}{\partial l_2} = -\frac{\delta\theta}{2\delta + 1}, \qquad \frac{\partial p_2}{\partial l_2} = -\frac{\delta}{2\delta + 1},$$

and

$$\frac{\partial q_{2,0}}{\partial l_2} = \frac{(\delta+1)\delta\theta}{2\delta+1}, \qquad \frac{\partial q_{2,2}}{\partial l_2} = \frac{(\delta+1)\delta}{2\delta+1},$$

the incremental VI term becomes

$$-\frac{\delta^2\theta^2}{2\delta+1}\left(1+\frac{\delta+1}{k_2}\right) - \frac{\delta^2}{2\delta+1}\left(1+\frac{\delta+1}{k_2}\right)\sigma_e^2.$$

This term is strictly negative. Using the producer first-order condition, the VI firms' first-order condition is

$$[a-b-\delta\theta l_2+(\delta+1)\theta l_1-\delta\theta][\delta\theta+\tfrac{1}{k_2}(\delta+1)\delta\theta]-[a+\delta(2b-(l_1+l_2+1)\theta)](\delta+1)\theta+b(2\delta+1)(\delta+1)\theta$$

$$+\{(-\delta l_2+(\delta+1)l_1-\delta)(\delta+\tfrac{1}{k_2}(\delta+1)\delta)+(\delta+1)\delta(l_1+l_2+1)\}\sigma_e^2+\left\{\tfrac{1}{k_2}(\delta+1)\delta-1\right\}\mathbb{E}[\varepsilon e]$$

$$+(2\delta+1)\delta\left[1+\frac{1}{k_2}(\delta+1)\right](\theta^2+\sigma_e^2)=0.$$

$$\Leftrightarrow$$

$$-[(\delta+1)\theta(\delta\theta+\tfrac{1}{k_2}(\delta+1)\delta\theta)+\delta(\delta+1)\theta^2+(\delta+1)(2\delta+\tfrac{1}{k_2}(\delta+1)\delta)\sigma_e^2]l_1^*$$

$$=[a-b-\delta l_2\theta-\delta\theta](\delta\theta+\tfrac{1}{k_2}(\delta+1)\delta\theta)-[a+\delta(2b-(l_2+1)\theta)](\delta+1)\theta+b(2\delta+1)(\delta+1)\theta$$

$$+\left[(-\delta l_2-\delta)(\delta+\delta\tfrac{1}{k}(\delta+1))+(\delta+1)\delta(l_2+1)\right]\sigma_e^2+\left\{\tfrac{1}{k_2}(\delta+1)\delta-1\right\}\mathbb{E}[\varepsilon e]$$

$$+(2\delta+1)\delta\left[1+\frac{1}{k_2}(\delta+1)\right](\theta^2+\sigma_e^2).$$

Two First-Order Conditions. Assume $k_1 = k_2 = k$.

From the expressions derived above, the producer's first-order condition can be written as

$$
\begin{aligned}
0 = {} & [(\delta+1)\theta(\delta\theta + \tfrac{\delta(\delta+1)}{k}\theta) + \delta(\delta+1)\theta^2 + (\delta+1)(2\delta + \tfrac{\delta(\delta+1)}{k})\sigma_e^2]\, l_1^* \\
& + [a - b - \delta l_2\theta - \delta\theta](\delta\theta + \tfrac{\delta(\delta+1)}{k}\theta) \\
& - [a + \delta(2b - (l_2+1)\theta)](\delta+1)\theta + b(2\delta+1)(\delta+1)\theta \\
& + [(-\delta l_2 - \delta)(\delta + \tfrac{\delta(\delta+1)}{k}) + (\delta+1)\delta(l_2+1)]\sigma_e^2 \\
& + \Big\{\tfrac{\delta(\delta+1)}{k} - 1\Big\}\mathbb{E}[\varepsilon e].
\end{aligned}
$$

The VI firm's first-order condition is identical except for the additional retail term:

$$
\begin{aligned}
0 = {} & [(\delta+1)\theta(\delta\theta + \tfrac{\delta(\delta+1)}{k}\theta) + \delta(\delta+1)\theta^2 + (\delta+1)(2\delta + \tfrac{\delta(\delta+1)}{k})\sigma_e^2]\, l_2^* \\
& + [a - b - \delta l_1\theta - \delta\theta](\delta\theta + \tfrac{\delta(\delta+1)}{k}\theta) \\
& - [a + \delta(2b - (l_1+1)\theta)](\delta+1)\theta + b(2\delta+1)(\delta+1)\theta \\
& + [(-\delta l_1 - \delta)(\delta + \tfrac{\delta(\delta+1)}{k}) + (\delta+1)\delta(l_1+1)]\sigma_e^2 \\
& + \Big\{\tfrac{\delta(\delta+1)}{k} - 1\Big\}\mathbb{E}[\varepsilon e] \\
& + (2\delta+1)\delta\Big(1 + \tfrac{\delta+1}{k}\Big)(\theta^2 + \sigma_e^2).
\end{aligned}
$$

Collecting terms in (l_1, l_2) and simplifying yields a linear system

$$
A\, l_1^* = B\, l_2^* + C, \qquad A\, l_2^* = B\, l_1^* + C + D,
$$

where

$$
A = \delta(\delta+1)(\theta^2 + \sigma_e^2)\Big(2 + \tfrac{\delta+1}{k}\Big),
$$

$$
B = \delta\Big(-1 + \tfrac{\delta(\delta+1)}{k}\Big)(\theta^2 + \sigma_e^2),
$$

$$
C = -[a - b - \delta\theta]\Big(\delta\theta + \tfrac{\delta(\delta+1)}{k}\theta\Big) + [a + \delta(2b - \theta)](\delta+1)\theta - b(2\delta+1)(\delta+1)\theta
$$

$$-\delta\Big(1-\tfrac{\delta(\delta+1)}{k}\Big)\sigma_e^2+\Big(1-\tfrac{\delta(\delta+1)}{k}\Big)\mathbb{E}[\varepsilon e].$$

$$D=-(2\delta+1)\delta\Big(1+\tfrac{\delta+1}{k}\Big)(\theta^2+\sigma_e^2),$$

Solving the linear system From the first equation, $l_2^*=\frac{Al_1^*-C}{B}$. Substituting into the second equation gives

$$A\left(\frac{Al_1^*-C}{B}\right)=Bl_1^*+C+D,$$

so

$$A^2l_1^*-AC=B^2l_1^*+BC+BD,$$

and therefore

$$l_1^*=\frac{BC+AC+BD}{A^2-B^2}.$$

Similarly,

$$l_2^*=\frac{BC+AC+AD}{A^2-B^2}.$$

Useful identity Note

$$\begin{aligned}A+B&=\delta(\theta^2+\sigma_e^2)\Big[(\delta+1)\Big(2+\tfrac{\delta+1}{k}\Big)+\Big(-1+\tfrac{\delta(\delta+1)}{k}\Big)\Big]\\&=\delta(\theta^2+\sigma_e^2)\Big[(2\delta+1)+\tfrac{(\delta+1)(2\delta+1)}{k}\Big]\\&=\delta(\theta^2+\sigma_e^2)(2\delta+1)\Big(1+\tfrac{\delta+1}{k}\Big).\end{aligned}$$

Hence,

$$\frac{D}{A+B}=-1.$$

Expected load-following forward positions Since $L_i=l_i(\theta+e)$ and $\mathbb{E}[e]=0$, we have $\mathbb{E}[L_i]=\theta\,l_i^*$. Using $A^2-B^2=(A-B)(A+B)$ and $\frac{D}{A+B}=1$,

$$\mathbb{E}[L_1]=\theta\left(\frac{C}{A-B}+\frac{B}{A-B}\cdot\frac{D}{A+B}\right)=\theta\left(\frac{C}{A-B}-\frac{B}{A-B}\right),$$

$$\mathbb{E}[L_2]=\theta\left(\frac{C}{A-B}+\frac{A}{A-B}\cdot\frac{D}{A+B}\right)=\theta\left(\frac{C}{A-B}-\frac{A}{A-B}\right).$$

Therefore,

$$\mathbb{E}[L_1] - \mathbb{E}[L_2] = \theta.$$

and their gap ($\mathbb{E}[L_1] - \mathbb{E}[L_2]$) is θ. This implies that, in expectation, the VI firm shades its forward position by an amount equal to its retail exposure.

Simplify LFFC

$$\mathbb{E}[L_1] = \theta \left(\frac{C}{A-B} - \frac{B}{A-B} \right)$$

with

$$A = \delta(\delta+1)\big(\theta^2 + \sigma_e^2\big) \left(2 + \frac{\delta+1}{k}\right), \qquad B = \delta \left(-1 + \frac{\delta(\delta+1)}{k}\right) \big(\theta^2 + \sigma_e^2\big),$$

$$C := -[a - b - \delta\theta]\Big(\delta\theta + \tfrac{\delta(\delta+1)}{k}\theta\Big) + [a + \delta(2b-\theta)](\delta+1)\theta - b(2\delta+1)(\delta+1)\theta$$

$$-\delta\Big(1 - \tfrac{\delta(\delta+1)}{k}\Big)\sigma_e^2 + \Big(1 - \tfrac{\delta(\delta+1)}{k}\Big)\mathbb{E}[\varepsilon e].$$

Simplification:

$$\mathbb{E}[L_1] = \frac{k^2}{5k^2+5k+1} \, \frac{\theta\big((a-b)\theta + \mathbb{E}[\varepsilon e]\big)}{\theta^2 + \sigma_e^2}$$

Let

$$S := \theta^2 + \sigma_e^2.$$

Denominator. Recall

$$A = \delta(\delta+1)S \left(2 + \frac{\delta+1}{k}\right), \qquad B = \delta \left(-1 + \frac{\delta(\delta+1)}{k}\right) S.$$

Hence,

$$A - B = \delta S \left[(\delta+1) \left(2 + \frac{\delta+1}{k}\right) - \left(-1 + \frac{\delta(\delta+1)}{k}\right) \right].$$

Distributing the minus sign yields

$$A - B = \delta S \left[(\delta+1) \left(2 + \frac{\delta+1}{k}\right) + 1 - \frac{\delta(\delta+1)}{k} \right].$$

Expanding terms,

$$(\delta+1)\left(2+\frac{\delta+1}{k}\right)=2(\delta+1)+\frac{(\delta+1)^2}{k}.$$

Therefore,

$$A-B=\delta S\left[2(\delta+1)+1+\frac{(\delta+1)^2-\delta(\delta+1)}{k}\right].$$

Since

$$(\delta+1)^2-\delta(\delta+1)=(\delta+1),$$

we obtain

$$A-B=\delta S\left[2\delta+3+\frac{\delta+1}{k}\right].$$

Writing over a common denominator k gives

$$A-B=\frac{\delta S\big(2\delta k+\delta+3k+1\big)}{k}.$$

Numerator We simplify

$$C=-[a-b-\delta\theta]\Big(\delta\theta+\tfrac{1}{k}(\delta+1)\delta\theta\Big)+[a+\delta(2b-\theta)](\delta+1)\theta-b(2\delta+1)(\delta+1)\theta$$

$$-\Big\{-\delta\Big(\delta+\tfrac{1}{k}(\delta+1)\delta\Big)+(\delta+1)\delta\Big\}\sigma_e^2-\Big\{\tfrac{1}{k}(\delta+1)\delta-1\Big\}\mathbb{E}[\varepsilon e].$$

The first line of C (terms without σ_e^2 and $\mathbb{E}[\varepsilon e]$) is

$$-[a-b-\delta\theta]\Big(\delta\theta+\tfrac{\delta(\delta+1)}{k}\theta\Big)+[a+\delta(2b-\theta)](\delta+1)\theta-b(2\delta+1)(\delta+1)\theta$$

$$=-\theta\Big\{[a-b-\delta\theta]\Big(\delta+\tfrac{\delta(\delta+1)}{k}\Big)-(\delta+1)[a+\delta(2b-\theta)]+b(2\delta+1)(\delta+1)\Big\}$$

$$=-\theta\Big[(a-b)\Big(\delta+\tfrac{\delta(\delta+1)}{k}\Big)-(\delta+1)(a-b)-\delta\theta\Big(\delta+\tfrac{\delta(\delta+1)}{k}\Big)+(\delta+1)\delta\theta\Big]$$

$$=-\theta\Big[(a-b)\Big(\tfrac{\delta(\delta+1)}{k}-1\Big)+\delta\theta\Big(1-\tfrac{\delta(\delta+1)}{k}\Big)\Big]=\theta\Big(1-\tfrac{\delta(\delta+1)}{k}\Big)\Big[(a-b)-\delta\theta\Big].$$

For the σ_e^2 term, the coefficient is

$$\delta\left(\delta+\frac{(\delta+1)\delta}{k}\right)-(\delta+1)\delta=-\delta\left[1-\frac{\delta(\delta+1)}{k}\right],$$

so the σ_e^2 contribution in C is

$$-\delta\left(1-\frac{\delta(\delta+1)}{k}\right)\sigma_e^2.$$

Finally,

$$-\Big\{\tfrac{1}{k}(\delta+1)\delta-1\Big\}\mathbb{E}[\varepsilon e]=\left(1-\frac{\delta(\delta+1)}{k}\right)\mathbb{E}[\varepsilon e].$$

Thus,

$$C=\theta\Big(1-\tfrac{\delta(\delta+1)}{k}\Big)\Big[(a-b)-\delta\theta\Big]-\delta\left(1-\frac{\delta(\delta+1)}{k}\right)\sigma_e^2+\left(1-\frac{\delta(\delta+1)}{k}\right)\mathbb{E}[\varepsilon e]$$

$$=\theta\Big(1-\tfrac{\delta(\delta+1)}{k}\Big)(a-b)-\delta\left(1-\frac{\delta(\delta+1)}{k}\right)(\theta^2+\sigma_e^2)+\left(1-\frac{\delta(\delta+1)}{k}\right)\mathbb{E}[\varepsilon e]$$

$$=\Big(\tfrac{k-\delta(\delta+1)}{k}\Big)\Big(\theta(a-b)+\mathbb{E}[\varepsilon e]\Big)+\delta\left(\frac{\delta(\delta+1)-k}{k}\right)(\theta^2+\sigma_e^2)$$

$$=\Big(\tfrac{k-\delta(\delta+1)}{k}\Big)\Big(\theta(a-b)+\mathbb{E}[\varepsilon e]\Big)+B.$$

So

$$C-B=\Big(\tfrac{k-\delta(\delta+1)}{k}\Big)\Big(\theta(a-b)+\mathbb{E}[\varepsilon e]\Big)$$

Therefore,

$$\begin{aligned}\mathbb{E}[L_1]&=\theta\left(\frac{C}{A-B}-\frac{B}{A-B}\right)\\&=\theta\Big(\frac{\Big(\frac{k-\delta(\delta+1)}{k}\Big)\Big(\theta(a-b)+\mathbb{E}[\varepsilon e]\Big)}{\frac{\delta S\left(2\delta k+\delta+3k+1\right)}{k}}\Big)\\&=\theta\Big(\frac{(k-\delta(\delta+1))\Big(\theta(a-b)+\mathbb{E}[\varepsilon e]\Big)}{\delta S\left(2\delta k+\delta+3k+1\right)}\Big)\\&=\frac{k^2}{5k^2+5k+1}\cdot\frac{\theta\big((a-b)\theta+\mathbb{E}[\varepsilon e]\big)}{\theta^2+\sigma_e^2}.\end{aligned}$$

Now subtract the swap forward position from the load-following:

$$\mathbb{E}[L_1]-f_1^*=\frac{k^2}{5k^2+5k+1}\left[\frac{\theta\big((a-b)\theta+\mathbb{E}[\varepsilon e]\big)}{\theta^2+\sigma_e^2}-(a-b)\right].$$

We can rewrite the bracketed term as

$$\frac{\theta\big((a-b)\theta+\mathbb{E}[\varepsilon e]\big)}{\theta^2+\sigma_e^2}-(a-b)=\frac{\theta\,\mathbb{E}[\varepsilon e]-(a-b)\sigma_e^2}{\theta^2+\sigma_e^2}.$$

So the wedge is exactly

$$\mathbb{E}[L_1]=f_1^*+\frac{k^2}{5k^2+5k+1}\,\frac{\theta\,\mathbb{E}[\varepsilon e]-(a-b)\sigma_e^2}{\theta^2+\sigma_e^2}$$

Equivalently,

$$\mathbb{E}[L_1]-f_1^*=\frac{k^2}{5k^2+5k+1}\,\frac{\theta\,\mathbb{E}[\varepsilon e]-(a-b)\sigma_e^2}{\theta^2+\sigma_e^2}$$

Result 3 is obtained as a special case under the assumption $e=\beta\varepsilon$.

Impose the proportionality assumption: Assume the retail demand shock is proportional to the market demand shock

$$e=\frac{\theta}{a}\varepsilon+\eta\ \text{ with }\ E[\eta]=0, E[\varepsilon\eta]=0$$

which says that the part of e that comoves with ε has slope θ/a. Plug into the condition:

$$\mathbb{E}[\varepsilon e]=\frac{a}{\theta}\sigma_e^2.$$

Then the key numerator becomes

$$\theta\,\mathbb{E}[\varepsilon e]-(a-b)\sigma_e^2=b\sigma_e^2$$

Let $\eta=\frac{\sigma_e}{\theta}$. Assume small η for approximation. We have

$$\mathbb{E}[L_1]-f_1^*=\frac{k^2}{5k^2+5k+1}\,\frac{b\sigma_e^2}{\theta^2+\sigma_e^2}=\frac{k^2}{5k^2+5k+1}\,\frac{b\eta^2}{1+\eta^2}=\frac{k^2}{5k^2+5k+1}\,\frac{O(\eta^2)}{1+O(\eta^2)}$$

Since

$$\frac{1}{1+O(\eta^2)}=1+O(\eta^2)$$

We can get

$$\mathbb{E}[L_1]-f_1^*=O(\eta^2)$$

Therefore,

$$\mathbb{E}[L_1] \approx f_1^* \qquad \text{to first order}$$

Just in case, the exact (non-approximated) results are reported below. Using $\frac{k^2}{5k^2+5k+1} > 0$

- if $b > 0$: $\mathbb{E}[L_1] - f_1^* > 0$ so LFFC > Swap;
- if $b = 0$: $\mathbb{E}[L_1] - f_1^* = 0$ so LFFC = Swap;
- if $b < 0$: $\mathbb{E}[L_1] - f_1^* < 0$ so LFFC < Swap.

This completes the proof of Result 3.

Proof of Result 4: Risk-Averse, Price-Taking Firms (LFFC)

We now consider price-taking firms under load-following forward contracts (LFFC). Under LFFC, firm i chooses a hedge ratio l_i applied to realized retail demand $(\theta + e)$. Forward payoffs therefore scale with realized load rather than being fixed quantities.

Profit. Firm profits are given by
Producer (firm 1):

$$\Pi_1 = \underbrace{p^f\, l_1(\theta + e) - p^s\, l_1(\theta + e)}_{\text{forward profit}} + \underbrace{p^s q_1 - C(q_1)}_{\text{generation profit}}.$$

Vertically integrated firm (firm 2):

$$\Pi_2 = \underbrace{p^r q_2^r - p^s q_2^r}_{\text{retail profit}} + \underbrace{p^f\, l_2(\theta + e) - p^s\, l_2(\theta + e)}_{\text{forward profit}} + \underbrace{p^s q_2 - C(q_2)}_{\text{generation profit}}.$$

Market clearing (perfect competition). Similarly to the proof of Result 2, and given that $(C(\cdot))$ is identical across firms, the spot price and output levels satisfy

$$p^s = \frac{a + \varepsilon}{2k} + b, \qquad q_1 = q_2 = \frac{a + \varepsilon}{2}.$$

Arbitrage and retail pricing Risk-neutral competitive speculators imply zero expected profits on load-following forward contracts:

$$\mathbb{E}\big[p^f l_i(\theta + e) - p^s l_i(\theta + e)\big] = 0.$$

Since l_i is chosen prior to the realization of (ε, e) and $\mathbb{E}[e] = 0$, dividing by l_i yields

$$p^f \theta = \mathbb{E}[p^s(\theta + e)] = \theta\, \mathbb{E}[p^s] + \mathbb{E}[p^s e].$$

Hence,

$$p^f = \mathbb{E}[p^s] + \frac{\mathbb{E}[p^s e]}{\theta}.$$

Substituting the equilibrium spot price under LFFC, we obtain

$$p^f = \mathbb{E}[p^s] + \frac{\mathbb{E}[p^s e]}{\theta} = \frac{a}{2k} + b + \frac{\mathbb{E}[\varepsilon e]}{2k\,\theta},$$

As discussed in the main text, and following Bushnell et al. (2008), we assume that expected retail profits are zero, with retail rates set to reflect expected wholesale prices and retail demand. Hence,

$$\mathbb{E}[p^r q_2^r - p^s q_2^r] = 0$$

.

Since p^r is chosen prior to the realization of (ε, e) and $\mathbb{E}[e] = 0$, we have

$$p^r = \mathbb{E}[p^s] + \frac{\mathbb{E}[p^s e]}{\theta}$$

Thus,

$$p^f = p^r$$

Risk aversion and objective. By the full-arbitrage condition and competitive retail pricing, expected forward and retail profits are zero. Moreover, under the price-taking assumption, forward positions do not affect spot market outcomes and hence are independent of l_1 and l_2. As a result, expected profits do not depend on l_1 or l_2 in this setup (as can be

verified analytically, following the same steps as in the proof of Result 2). Consequently, for both the producer and the vertically integrated firm, the mean–variance objective reduces to choosing l_1 and l_2 to minimize the variance of profits.

Common decomposition The forward profit is obtained using

$$p^f - p^s = \frac{1}{2k}\left(\frac{\mathbb{E}[\varepsilon e]}{\theta} - \varepsilon\right).$$

Hence,

$$\begin{aligned}(p^f - p^s)(\theta + e) &= \frac{1}{2k}\left(\frac{\mathbb{E}[\varepsilon e]}{\theta} - \varepsilon\right)(\theta + e) \\ &= \frac{1}{2k}\Big(\mathbb{E}[\varepsilon e] + R\Big),\end{aligned}$$

where the stochastic component is defined as

$$R := -\varepsilon e + \frac{\mathbb{E}[\varepsilon e]}{\theta}\, e - \theta\varepsilon.$$

The generation profit is

$$\begin{aligned}p^s q_1 - C(q_1) &= \left(\frac{a+\varepsilon}{2k} + b\right)\frac{a+\varepsilon}{2} - b\frac{a+\varepsilon}{2} - \frac{1}{2k}\left(\frac{a+\varepsilon}{2}\right)^2 \\ &= \frac{(a+\varepsilon)^2}{8k}.\end{aligned}$$

Since constants do not affect variances, we drop deterministic terms and define the stochastic generation component

$$G := \frac{\varepsilon^2 + 2a\varepsilon}{8k}.$$

Firm 1 (Producer): Optimal l_1. Firm 1's profit can be written as

$$\Pi_1 = \frac{l_1}{2k}\big(\mathbb{E}[\varepsilon e] + R\big) + \frac{(a+\varepsilon)^2}{8k}.$$

Because $\mathbb{E}[\varepsilon e]$ affects only the mean of profits, it does not enter the variance. The

variance-relevant component is therefore

$$\frac{l_1}{2k}R + G,$$

where G denotes the stochastic component of generation profits defined above.

Hence,

$$\mathrm{Var}(\Pi_1) = \mathrm{Var}\left(\frac{l_1}{2k}R + G\right).$$

Minimizing variance with respect to l_1 yields the first-order condition

$$\frac{l_1}{(2k)^2}\,\mathrm{Var}(R) + \frac{1}{2k}\,\mathrm{Cov}(R, G) = 0.$$

Multiplying through by $(2k)^2$ gives

$$l_1\ \mathrm{Var}(R) + 2k\ \mathrm{Cov}(R, G) = 0,$$

and therefore

$$l_1^* = -2k\,\frac{\mathrm{Cov}(R, G)}{\mathrm{Var}(R)}.$$

Normalization and Moments. Define the normalized shocks

$$x := \frac{e}{\theta}, \qquad y := \frac{\varepsilon}{a}, \qquad v := \mathbb{E}[y^2], \qquad v_e := \mathbb{E}[x^2].$$

Let the mixed moments be

$$m_{11} := \mathbb{E}[xy], \quad m_{21} := \mathbb{E}[y^2x], \quad m_{30} := \mathbb{E}[y^3], \quad m_{31} := \mathbb{E}[y^3x],$$

$$m_{12} := \mathbb{E}[x^2y], \quad m_{22} := \mathbb{E}[x^2y^2].$$

These normalizations imply

$$\sigma^2 = a^2v, \qquad \sigma_e^2 = \theta^2v_e, \qquad \mathbb{E}[\varepsilon e] = a\theta m_{11}, \qquad \mathbb{E}[\varepsilon^2 e] = a^2\theta m_{21},$$

$$\mathbb{E}[\varepsilon^3] = a^3m_{30}, \qquad \mathbb{E}[\varepsilon^3 e] = a^3\theta m_{31}.$$

Using these definitions, the stochastic components can be written as

$$R = a\theta S, \qquad S := -yx + m_{11}x - y,$$

$$G = \frac{a^2}{8k}T, \qquad T := y^2 + 2y.$$

Therefore,

$$\begin{aligned}\frac{\mathrm{Cov}(R,G)}{\mathrm{Var}(R)} &= \frac{\mathrm{Cov}(a\theta S,\ \frac{a^2}{8k}T)}{\mathrm{Var}(a\theta S)} \\ &= \frac{a}{8k\,\theta}\,\frac{\mathrm{Cov}(S,T)}{\mathrm{Var}(S)}.\end{aligned}$$

Substituting into the producer's first-order condition yields

$$l_1\theta = -\frac{a}{4}\,\frac{\mathrm{Cov}(S,T)}{\mathrm{Var}(S)}. \tag{20}$$

Bounding and approximation Assume small-noise scalings

$$\delta := \sqrt{v} = \frac{\sigma}{a}, \qquad \eta := \sqrt{v_e} = \frac{\sigma_e}{\theta},$$

and bounded standardized moments (bounded skewness, kurtosis, and the sixth moment)

$$\gamma := \frac{\mathbb{E}[\varepsilon^3]}{\sigma^3} = \frac{\mathbb{E}[y^3]}{v^{3/2}}, \qquad \kappa := \frac{\mathbb{E}[\varepsilon^4]}{\sigma^4} = \frac{\mathbb{E}[y^4]}{v^2}, \qquad \tau := \frac{\mathbb{E}[\varepsilon^6]}{\sigma^6} = \frac{\mathbb{E}[y^6]}{v^3}, \qquad \kappa_e := \frac{\mathbb{E}[e^4]}{\sigma_e^4} = \frac{\mathbb{E}[x^4]}{v_e^2},$$

with $\gamma, \kappa, \tau, \kappa_e = O(1)$. By Cauchy–Schwarz,

$$|m_{11}| \le \eta\delta, \qquad |m_{21}| \le O(\delta^2\eta), \qquad m_{30} = \gamma\delta^3 = O(\delta^3), \qquad |m_{31}| \le O(\delta^3\eta).$$

We compute

$$\begin{aligned}\mathrm{Cov}(S,T) &= \mathrm{Cov}(-yx + m_{11}x - y,\ y^2 + 2y) \\ &= \underbrace{-2v}_{\text{leading}} + A,\end{aligned}$$

where

$$A := -m_{30} - 2m_{21} + 2m_{11}^2 + \underbrace{\left(-m_{31} + m_{11}m_{21} + m_{11}v\right)}_{\text{higher order}}.$$

Similarly,

$$\begin{aligned}\mathrm{Var}(S) &= \mathrm{Var}(-yx + m_{11}x - y) \\ &= \underbrace{v}_{\text{leading}} + B,\end{aligned}$$

where

$$B := 2m_{21} - 3m_{11}^2 + m_{22} + \underbrace{\left(m_{11}^2 v_e - 2m_{11}m_{12}\right)}_{\text{higher order}}.$$

The bounds above imply

$$A = O(\delta^3) + O(\delta^2\eta) + O(\delta^2\eta^2) + O(\delta^3\eta), \qquad B = O(\delta^2\eta) + O(\delta^2\eta^2) + O(\delta^2\eta^3) + O(\delta^2\eta^4).$$

Approximation step. Assume δ is small and $|B| < v/2$ so that $\mathrm{Var}(S)$ is bounded away from zero. Then

$$\begin{aligned}\frac{\mathrm{Cov}(S,T)}{\mathrm{Var}(S)} = \frac{-2v + A}{v + B} &= \left(-2 + \frac{A}{v}\right)\left(1 + \frac{B}{v}\right)^{-1} \\ &= \left(-2 + \frac{A}{v}\right)\left(1 - \frac{B}{v} + O\left(\frac{B^2}{v^2}\right)\right) \\ &= -2 + \frac{A + 2B}{v} + O\left(\frac{|A||B|}{v^2} + \frac{B^2}{v^2}\right).\end{aligned}$$

Using the definitions of A and B, the leading part of $A + 2B$ is

$$A + 2B = -m_{30} + 2m_{21} - 4m_{11}^2 + 2m_{22} + (\text{higher order}).$$

Therefore,

$$\frac{\mathrm{Cov}(S,T)}{\mathrm{Var}(S)} = -2 - \frac{m_{30}}{v} + 2\frac{m_{21}}{v} - 4\frac{m_{11}^2}{v} + 2\frac{m_{22}}{v} + \mathcal{R}, \tag{21}$$

where

$$\mathcal{R} = O(\delta\eta) + O(\delta\eta^2) + O(\eta^2) + O(\eta^3) \quad \text{and also contains } O(B^2/v^2) \text{ terms.}$$

Firm 1 conclusion and comparison to the swap benchmark Combining (20) and (21) yields

$$\begin{aligned} l_1\,\theta &= -\frac{a}{4}\left(-2-\frac{m_{30}}{v}+2\frac{m_{21}}{v}-4\frac{m_{11}^2}{v}+2\frac{m_{22}}{v}+\mathcal{R}\right) \\ &= \frac{a}{2}+\frac{a}{4}\frac{m_{30}}{v}-\frac{a}{2}\frac{m_{21}}{v}+a\frac{m_{11}^2}{v}-\frac{a}{2}\frac{m_{22}}{v}-\frac{a}{4}\mathcal{R}. \end{aligned}$$

The variance-minimizing *swap* position for firm 1 is (exactly)

$$f_1^* = \frac{a}{2}+\frac{a}{4}\,\frac{m_{30}}{v}. \tag{22}$$

Taking the difference and retaining the first nontrivial term ,

$$\begin{aligned} l_1\theta - f_1^* &= -\frac{a}{2}\frac{m_{21}}{v}+O(\eta^2)+O(\delta\eta) \\ &= -\frac{a}{2}\,\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]}+O(\eta^2)+O(\delta\eta). \end{aligned}$$

This implies

$$l_1\theta \approx f_1^* - \frac{a}{2}\,\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]}$$

to first order, with the remaining approximation error of second order.

Firm 2 (VI): optimal l_2 The only difference relative to the standalone producer is the retail profit term $p^r q_2^r - p^s q_2^r$. Under competitive forward and retail markets, we have already established that $p^r = p^f$. Consequently, Firm 2's profit can be written as

$$\Pi_2 = (p^f - p^s)(\theta+e) + (p^f - p^s)\,l_2(\theta+e) + \frac{(a+\varepsilon)^2}{8k}.$$

Using the decomposition above, the variance-relevant component is

$$\frac{1+l_2}{2k}R + G.$$

Hence

$$\mathrm{Var}(\Pi_2) = \mathrm{Var}\left(\frac{1+l_2}{2k}R+G\right),$$

and the first-order condition is

$$(1 + l_2)\ \mathrm{Var}(R) + 2k\ \mathrm{Cov}(R, G) = 0, \qquad \Longleftrightarrow \qquad 1 + l_2 = -2k\,\frac{\mathrm{Cov}(R, G)}{\mathrm{Var}(R)}.$$

Using the producer characterization, this implies

$$1 + l_2^* = l_1^*.$$

Multiplying by θ,

$$l_2\,\theta = l_1\,\theta - \theta. \tag{23}$$

This shows that, in expectation, the VI firm shades its load-following forward position by the level of expected retail demand.

Firm 2 comparison to the swap benchmark The variance-minimizing *swap* position for firm 2 is (exactly)

$$f_2^* = -\theta + \frac{a}{2} + \theta\,\frac{m_{11}^2}{v} - \theta\,\frac{m_{21}}{v} + \frac{a}{4}\,\frac{m_{30}}{v}. \tag{24}$$

Combining (23) with the expansion for $l_1\theta$ and subtracting (24) yields,

$$\begin{aligned} l_2\theta - f_2^* &= \left(\theta - \frac{a}{2}\right)\frac{m_{21}}{v} + O(\eta^2) + O(\delta\eta) \\ &= \left(\theta - \frac{a}{2}\right)\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]} + O(\eta^2) + O(\delta\eta) \\ &= -\frac{a}{2}\,\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]} + \frac{\mathbb{E}[\varepsilon^2 e]}{\mathbb{E}[\varepsilon^2]} + O(\eta^2) + O(\delta\eta). \end{aligned}$$

This implies

$$l_2\theta \approx f_2^* - \frac{a}{2}\,\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]} + \frac{\mathbb{E}[\varepsilon^2 e]}{\mathbb{E}[\varepsilon^2]}$$

to first order, with the remaining approximation error of second order.

Finally, we have the following first-order approximations:

$$l_1\theta \approx f_1^* - \frac{a}{2}\,\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]},$$

$$l_2\theta \approx f_2^* - \frac{a}{2}\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]} + \frac{\mathbb{E}[\varepsilon^2 e]}{\mathbb{E}[\varepsilon^2]}.$$

Using the identity

$$l_2\theta = l_1\theta - \theta,$$

we can equivalently write

$$l_2\theta \approx f_1^* - \theta - \frac{a}{2}\frac{\mathbb{E}[\varepsilon^2 e]}{\theta\,\mathbb{E}[\varepsilon^2]}.$$

Finally, substituting

$$\mathbb{E}[l_{i,A}q_2^r] = l_i\theta, \quad e = \beta\varepsilon, \quad \text{and} \quad \mathbb{E}[q_i^*] = \frac{a}{2},$$

completes the proof of Result 4.

Appendix B Additional Results

Proposition 1. *Under a load-following forward, the retail rate giving zero expected profit is the LFFC per-unit price,*

$$\frac{\mathbb{E}[p^s q^r]}{\mathbb{E}[q^r]}.$$

This rate strictly exceeds $\mathbb{E}[p^s]$*. However, charging* $p^r = \mathbb{E}[p^s]$ *yields an expected loss for any amount of retailer hedging.*

Proof. **Part 1: The LFFC per-unit price is** $\frac{\mathbb{E}[p^s q^r]}{\mathbb{E}[q^r]}$**.** The LFFC pays $\left(p^s - p^f\right) l\, q^r$, where l denotes a strip. Full arbitrage requires that this payoff have zero expected value:

$$\mathbb{E}\left[\left(p^s - p^f\right) l\, q^r\right] = 0 \quad \Longrightarrow \quad p^f = \frac{\mathbb{E}[p^s q^r]}{\mathbb{E}[q^r]}.$$

Part 2: The zero-profit rate is the LFFC price. The retailer sells at p^r, buys the realized load at the spot price, and holds a LFFC that insures its retail demand. The LFFC settles at $\left(p^s - p^f\right) l\, q^r$, and with risk-neutral arbitrageurs it is priced so that this settlement is zero in expectation. The retailer's profit is therefore

$$\Pi^{\text{retail}} = \underbrace{p^r q^r}_{\text{revenue}} - \underbrace{p^s q^r}_{\text{spot cost}} + \underbrace{\left(p^s - p^f\right) l\, q^r}_{\text{LFFC settlement}}.$$

Taking expectations, the settlement term vanishes by full arbitrage, so

$$\mathbb{E}[\Pi^{\text{retail}}] = p^r\,\mathbb{E}[q^r] - \mathbb{E}[p^s q^r].$$

and setting retail profits to zero gives

$$p^r_* = \frac{\mathbb{E}[p^s q^r]}{\mathbb{E}[q^r]}.$$

Splitting the numerator with $\mathbb{E}[p^s q^r] = \mathbb{E}[p^s]\,\mathbb{E}[q^r] + \mathrm{Cov}(p^s, q^r)$,

$$p^r_* = \frac{\mathbb{E}[p^s]\,\mathbb{E}[q^r] + \mathrm{Cov}(p^s, q^r)}{E[q^r]} = \mathbb{E}[p^s] + \frac{\mathrm{Cov}(p^s, q^r)}{\mathbb{E}[q^r]}.$$

Since p^s is increasing in q^r, we have $\mathrm{Cov}(p^s, q^r) > 0$, and therefore

$$p^r_* = E[p^s] + \frac{\mathrm{Cov}(p^s, q^r)}{E[q^r]} > E[p^s].$$

Part 3: $\mathbb{E}[p^s]$ is less than this price and results in expected losses for the retailer. Now suppose the retailer charges $p^r = E[p^s]$ and, unable to hedge the random load exactly, holds a swap forward for an *arbitrary* fixed volume f priced fairly at $p^f = \mathbb{E}[p^s]$. Its payoff is $(p^s - p^f)f$, so

$$\Pi^{\text{retail}} = p^r q^r - p^s q^r + \big(p^s - \mathbb{E}[p^s]\big) f.$$

Taking expectations, the settlement term vanishes by full arbitrage, so

$$\mathbb{E}[\Pi^{\text{retail}}] = \mathbb{E}[p^s]\ \mathbb{E}[q^r] - \mathbb{E}[p^s q^r] = -\,\mathrm{Cov}(p^s, q^r).$$

Because p^s is increasing in q^r, $\mathrm{Cov}(p^s, q^r) > 0$, hence $E[\Pi^{\text{retail}}] < 0$ and the retailer makes an expected loss. □

Appendix C Simulation Algorithm

We solve a two-stage game among $N = 4$ strategic firms. In Stage 1, firms choose their forward positions. In Stage 2, system demand and the vertically integrated firm's retail load are realized, and firms compete in quantities in the spot market. For each uncertainty state,

we formulate the spot-market equilibrium as a mixed-integer optimization problem. We then nest these state-specific problems within a Gauss–Seidel algorithm that computes sequential best responses in the forward market, following Ito and Reguant (2016).

Appendix C.1 Uncertainty States

We approximate the joint distribution of system demand and the vertically integrated firm's retail load using a discrete grid. This discretization is needed because firms choose their forward positions before uncertainty is realized and therefore evaluate expected utility by integrating profits over the joint distribution of system demand and retail load. Because each possible realization requires solving a separate spot-market equilibrium, evaluating this integral directly over a continuous distribution would be computationally infeasible. We therefore approximate the expectation using a finite weighted sum over representative states.

Let S denote system demand and let R denote the vertically integrated firm's retail load. We model retail load conditional on system demand rather than imposing a constant retail share. Let $\mu_S = \mathbb{E}[S]$ and $\sigma_S = \mathrm{sd}(S)$, and define standardized system demand as $z = \frac{S-\mu_S}{\sigma_S}$.

For each year, we estimate the conditional retail-load relationship

$$R = \alpha + \beta \exp(\kappa z) + u, \qquad \mathbb{E}[u \mid S] = 0, \qquad \mathrm{Var}(u \mid S) = \sigma_R^2.$$

Conditional on κ, the model is linear in α and β. We therefore estimate α and β by OLS for each candidate value of κ and select κ using a one-dimensional grid search. We impose $\beta\kappa \geq 0$, so that predicted retail load is weakly increasing in system demand.

We approximate the distribution of system demand by a normal distribution and construct eleven quantile points at probabilities $\{p_{\text{tail}}, 0.1, 0.2, \ldots, 0.9, 1 - p_{\text{tail}}\}$ where $p_{\text{tail}} = \frac{1}{2\times 8760}$. The two tail points together represent approximately one extreme hour per year, divided symmetrically between the lower and upper tails. For quantile probability p_j, define

$$z_j = \Phi^{-1}(p_j), \qquad S_j = \mu_S + \sigma_S z_j.$$

Each system-demand point is assigned probability weight w_j^S based on the probability masses between the midpoints of adjacent quantile levels. Let S_{50} denote system demand at the median. In the spot-market simulation, we define the demand shock relative to the median as $\epsilon_j = S_j - S_{50}$.

At each system-demand point, the fitted conditional retail load is

$$R_{j,\text{mid}} = \alpha + \beta \exp(\kappa z_j).$$

To capture residual retail-load uncertainty, we construct lower and upper realizations according to

$$R_{j,\text{low}} = R_{j,\text{mid}} - 2\sigma_R, \qquad R_{j,\text{high}} = R_{j,\text{mid}} + 2\sigma_R.$$

Each of the three conditional retail-load realizations receives one-third of the probability assigned to the corresponding system-demand point. This produces a 33-state grid consisting of eleven system-demand levels and three conditional retail-load realizations. Indexing the resulting states by s, their probabilities satisfy $\sum_s w_s = 1$. We construct this joint grid prior to the equilibrium simulation and use it as an input to the state-specific spot-market problems.

Appendix C.2 Spot-Market Equilibrium

Appendix C.2.1 Residual demand and production costs

Let q_{ijs} denote the production of firm i from cost block j in state s, and define

$$q_{is} = \sum_{j=1}^{K} q_{ijs}, \qquad Q_s = \sum_{i=1}^{N} q_{is}.$$

Here, Q_s is total production by the strategic firms. Nonstrategic production is incorporated into the residual-demand curve rather than added to Q_s.

The inverse residual-demand curve contains a price-cap segment, three downward-sloping segments, and a zero-price segment:

$$P_s(Q) = \begin{cases} \bar{P}, & 0 \le Q \le \kappa_{1s}, \\ \dfrac{A_{ms} - Q}{b_m}, & \kappa_{m-1,s} \le Q \le \kappa_{ms}, \quad m = 2, 3, 4, \\ 0, & \kappa_{4s} \le Q \le \overline{Q}. \end{cases}$$

The parameters $b_m > 0$ determine the slopes of the three interior segments. The demand shock shifts the residual-demand curve horizontally.

Each firm has K ordered production blocks. Marginal cost on block (i, j) is linear:

$$MC_{ij}(q_{ijs}) = a_{ij} + c_{ij}q_{ijs},$$

where $\overline{q}_{ij}$ is the block's capacity. The associated production cost is

$$C_{ij}(q_{ijs}) = a_{ij}q_{ijs} + \frac{1}{2}c_{ij}q_{ijs}^2.$$

Let f_{is} denote firm i's forward position in state s, q_i^{infra} its nonstrategic inframarginal production, and R_{is} its retail load. Retail load is zero for generation-only firms. Define the firm's precommitted position as

$$h_{is} = f_{is} - q_i^{\text{infra}} + R_{is}.$$

The firm's net exposure to the spot price is therefore

$$q_{is} - h_{is} = q_{is} + q_i^{\text{infra}} - f_{is} - R_{is}.$$

A larger forward sale or retail obligation reduces the firm's net exposure to a higher spot price and therefore reduces its incentive to withhold production.

Appendix C.2.2 Mixed-integer equilibrium conditions

Let firms be indexed by $i = 1, \ldots, N$, production blocks by $j = 1, \ldots, K$, uncertainty states by $s = 1, \ldots, 33$, and residual-demand segments by $m = 1, \ldots, 5$. Let q_{ijs} denote production from block (i, j) in state s, $q_{is} = \sum_j q_{ijs}$ firm i's total production, and $Q_s = \sum_i q_{is}$ total strategic production. The spot price is P_s. Each block has capacity $\overline{q}_{ij}$ and marginal cost $a_{ij} + c_{ij}q_{ijs}$

For residual-demand segment m, let $\underline{Q}_{ms}$ and $\overline{Q}_{ms}$ denote its lower and upper quantity bounds, A_{ms} its state-specific intercept, and $b_m > 0$ its inverse-demand slope parameter. Segments $m = 1$ and $m = 5$ are the price-cap and zero-price segments, respectively, while $m = 2, 3, 4$ are downward sloping. The constant M is a sufficiently large number used to deactivate constraints associated with segments or production blocks that are not selected.

Recall that firm i's precommitted position is $f_{is} - q_i^{\text{infra}} + R_{is}$, where f_{is} is its forward position, q_i^{infra} is its nonstrategic inframarginal production, and R_{is} is its retail load. Thus,

$q_{is} - h_{is}$ is the firm's net exposure to the spot price.

Let $z_{ms} \in 0, 1$ indicate that residual-demand segment m is active in state s. Let $u_{ijs} \in \{0, 1\}$ indicate that production block (i, j) is used, and let $\overline{u}_{ijs} \in \{0, 1\}$ indicate that the block is used at full capacity. Let $\eta_{ijs} \geq 0$ denote the shadow value of the block's upper-capacity constraint.

For each downward-sloping residual-demand segment $m = 2, 3, 4$, define

$$G_{ijms} = P_s - \frac{q_{is} - h_{is}}{b_m} - a_{ij} - c_{ij} q_{ijs} - \eta_{ijs}. \tag{25}$$

The spot-market equilibrium conditions are

$$\text{[Segment selection]} \quad \sum_{m=1}^{5} z_{ms} = 1 \tag{26}$$

$$\text{[Segment range 1]} \quad Q_s \geq \underline{Q}_{ms} - M(1 - z_{ms}), \qquad m = 1, \ldots, 5 \tag{27}$$

$$\text{[Segment range 2]} \quad Q_s \leq \overline{Q}_{ms} + M(1 - z_{ms}), \qquad m = 1, \ldots, 5 \tag{28}$$

$$\text{[Inverse demand]} \quad -M(1 - z_{ms}) \leq Q_s + b_m P_s - A_{ms} \leq M(1 - z_{ms}), \qquad m = 2, 3, 4 \tag{29}$$

$$\text{[FOC 1]} \quad G_{ijms} \leq M(1 - z_{ms}), \qquad \forall i, j, \quad m = 2, 3, 4 \tag{30}$$

$$\text{[FOC 2]} \quad G_{ijms} \geq -M(2 - z_{ms} - u_{ijs}), \qquad \forall i, j, \quad m = 2, 3, 4 \tag{31}$$

$$\text{[Complementarity]} \quad 0 \leq \eta_{ijs} \leq M\overline{u}_{ijs}, \qquad \forall i, j \tag{32}$$

$$\text{[Range 1]} \quad q_{ijs} \leq \overline{q}_{ij} u_{ijs}, \qquad \forall i, j \tag{33}$$

$$\text{[Range 2]} \quad q_{ijs} \geq \overline{q}_{ij} \overline{u}_{ijs}, \qquad \forall i, j \tag{34}$$

$$\text{[Block consistency]} \quad \overline{u}_{ijs} \leq u_{ijs}, \qquad \forall i, j \tag{35}$$

$$\text{[Merit order]} \quad u_{ijs} \leq \overline{u}_{i,j-1,s}, \qquad \forall i, \quad j = 2, \ldots, K \tag{36}$$

For the price-cap and zero-price segments, the corresponding inverse-demand conditions are

$$\text{[Price cap]} \quad -M(1 - z_{1s}) \leq P_s - \overline{P} \leq M(1 - z_{1s}) \tag{37}$$

$$\text{[Price floor]} \quad -M(1 - z_{5s}) \leq P_s \leq M(1 - z_{5s}) \tag{38}$$

When both residual-demand segment m and block (i, j) are active, equations (30) and (31) imply

$$P_s - \frac{q_{is} + q_i^{\text{infra}} - f_{is} - R_{is}}{b_m} = MC_{ij}(q_{ijs}) + \eta_{ijs}. \tag{39}$$

The range, block-consistency, and merit-order conditions imply

$$\overline{u}_{ijs} \leq u_{ijs} \leq \overline{u}_{i,j-1,s} \leq u_{i,j-1,s}.$$

Thus, a higher-cost production block can be used only after the preceding block has been filled. Blocks classified as must-run are fixed at their available capacity, with the corresponding lower-bound optimality conditions included in the formulation.

At the price-cap and price-floor kinks, the profit function is not differentiable. We introduce additional kink indicators that force aggregate quantity to equal the relevant knot. At the price-cap kink, optimality requires that the profit derivative evaluated from the flat price-cap segment be weakly positive and the derivative evaluated from the adjacent downward-sloping segment be weakly negative. The analogous one-sided derivative conditions are imposed at the zero-price kink. These conditions ensure that zero belongs to the subdifferential of each firm's profit at the selected kink.

Supply-side shocks are beyond the scope of this paper. Nevertheless, when the price cap binds, nameplate capacity may overstate the quantity a firm can realistically produce. Some capacity may be unavailable because of outages, deratings, or other operating constraints, and even available capacity may not be fully dispatched because the market-clearing quantity can be reached before every firm produces at its nameplate capacity. To account for these possibilities, we cap each firm's production in price-cap states at a fraction δ of its modeled capacity. In the main specification, δ is set to the 25th percentile of the observed ratio of dispatched quantity to maximum capacity in price-cap hours, approximately 65 percent.

Multiple segment-consistent equilibria can arise at the knots of the residual-demand curve. We select the equilibrium on the rightmost feasible segment, which corresponds to the lower-price branch, and then select the solution with the lowest production cost.

Appendix C.3 Profits and Contract Specifications

Firm i's state-contingent total profit consists of generation and inframarginal profit, forward-contract profit, and retail profit:

$$\pi_{is} = \underbrace{P_s q_{is} - \sum_{j=1}^{K} C_{ij}(q_{ijs}) + P_s q_i^{\text{infra}}}_{\text{generation and inframarginal profit}} + \underbrace{(F_i - P_s) f_{is}}_{\text{forward-contract profit}} + \underbrace{(F^R - P_s) R_{is}}_{\text{retail profit}}. \tag{40}$$

Here, F_i is the settlement price for firm i's forward contract, while F^R is the retail price applied to the vertically integrated firm's retail position. The forward quantity f_{is} may be fixed or state dependent, depending on the contract specification. For generation-only firms, $R_{is} = 0$, so the retail-profit term vanishes.

Firm i evaluates its Stage 1 position using CARA expected utility over the distribution of state-contingent profits. For $\lambda_i > 0$, expected utility is

$$EU_i = -\sum_s w_s \exp\left(-\lambda_i \pi_{is}\right). \tag{41}$$

Because the logarithm is strictly increasing, maximizing expected utility is equivalent to maximizing the certainty equivalent

$$CE_i = -\frac{1}{\lambda_i} \log\left(\sum_s w_s \exp\left(-\lambda_i \pi_{is}\right)\right). \tag{42}$$

The certainty equivalent is expressed in profit units and is therefore more convenient for computation and interpretation. In the risk-neutral limit, $\lambda_i = 0$, the certainty equivalent becomes expected profit:

$$CE_i = \sum_s w_s \pi_{is}.$$

We consider three contract specifications. Under a fixed-volume swap, firm i chooses a

quantity x_i and

$$f_{is} = x_i, \qquad F^{\text{swap}} = \sum_s w_s P_s. \tag{43}$$

Under a load-following forward contract indexed to the vertically integrated firm's retail load,

$$f_{is} = x_i R_s, \qquad F^R = \frac{\sum_s w_s P_s R_s}{\sum_s w_s R_s}. \tag{44}$$

Here, x_i is the share of the reference retail load covered by firm i's contract. Under a contract indexed instead to system demand,

$$f_{is} = x_i S_s, \qquad F^S = \frac{\sum_s w_s P_s S_s}{\sum_s w_s S_s}. \tag{45}$$

The retail position itself is settled using the retail-load-weighted price F^R. The settlement prices are recomputed for every candidate profile of Stage 1 positions because the equilibrium spot prices depend on those positions.

Firms have constant absolute risk aversion. For a risk-aversion parameter $\lambda_i > 0$, firm i maximizes the certainty equivalent

$$CE_i = -\frac{1}{\lambda_i} \log \left(\sum_s w_s \exp(-\lambda_i \pi_{is}) \right). \tag{46}$$

The risk-neutral limit is

$$CE_i = \sum_s w_s \pi_{is} \qquad \text{when } \lambda_i = 0.$$

We allow the vertically integrated firm and generation-only firms to have different risk-aversion parameters, denoted by λ_{VI} and λ_{gen}, respectively.

Appendix C.4 Stage 1 Forward-Market Equilibrium

We solve for a Nash equilibrium in Stage 1 using sequential best responses. Let x_i denote firm i's Stage 1 choice. For a swap, x_i is a fixed forward quantity. For a load-following contract, x_i is the share of the reference load covered by the contract.

Algorithm 1 Stage 1 Equilibrium via Gauss–Seidel

Initialize $x_i^{(0)} = 0$ for all i
Solve the spot-only equilibrium and set $\tau = 10^{-2} \|\mathbf{q}^{\text{spot}}\|$
Set iter $= 0$ and crit $= \infty$
while crit $> \tau$ and iter < 20 **do**
    Set $\widetilde{\mathbf{x}} = \mathbf{x}^{(\text{iter})}$
    for $i = 1, \ldots, N$ **do**
        Hold the other firms' current positions $\widetilde{\mathbf{x}}_{-i}$ fixed
        For each candidate $x_i \in [L_i, U_i]$:
            Solve the Stage 2 equilibrium in each of the 33 uncertainty states
            Recompute the relevant settlement prices and state-contingent profits
            Compute $CE_i(x_i, \widetilde{\mathbf{x}}_{-i})$
        Solve $x_i^* = \arg\max_{x_i \in [L_i, U_i]} CE_i(x_i, \widetilde{\mathbf{x}}_{-i})$
        Update $\widetilde{x}_i = x_i^*$
    end for
    Set crit $= \left\|\widetilde{\mathbf{x}} - \mathbf{x}^{(\text{iter})}\right\|$
    Set $\mathbf{x}^{(\text{iter}+1)} = \widetilde{\mathbf{x}}$
    Set iter $=$ iter $+ 1$
end while

Here, $\mathbf{q}^{\text{spot}}$ is the vector of expected firm quantities in the spot-only equilibrium. The algorithm stops when the change in Stage 1 positions is less than one percent of the norm of the spot-only equilibrium quantities or after 20 iterations.

Define $K_i = \sum_j \overline{q}_{ij}$ as firm i's total generation capacity. In the swap market, we restrict the forward position of a generation-only firm to $x_i \in [0, K_i]$ and that of the vertically integrated firm to $x_i \in [-3K_i, K_i]$. In the load-following markets, x_i is a coverage share rather than a quantity. We restrict the vertically integrated firm's coverage share to $x_i \in [-1, 1]$ and the coverage shares of generation-only firms to $x_i \in [0, 1]$.

Appendix D Additional Figures

Appendix D.1 Forward Price Diagnostics

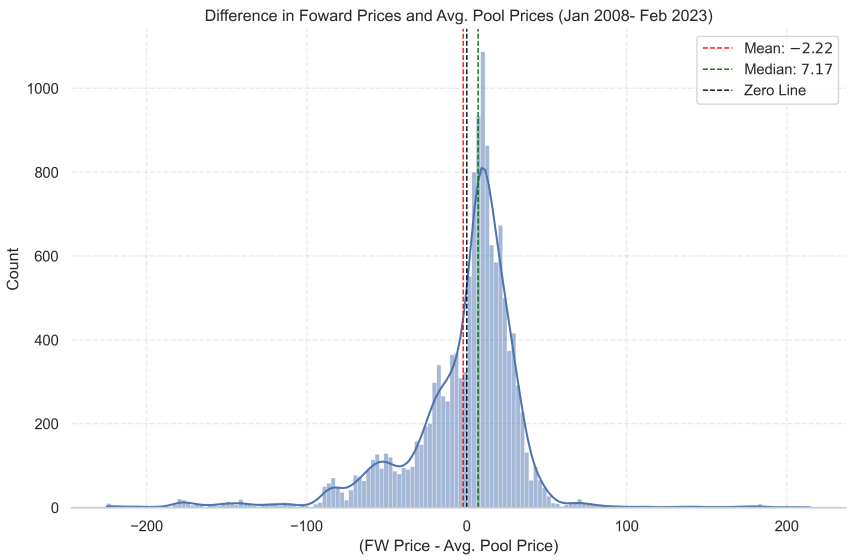


Figure 7: Forward vs Spot Prices

Appendix D.2 Changes in Marginal Costs from the End of PPA

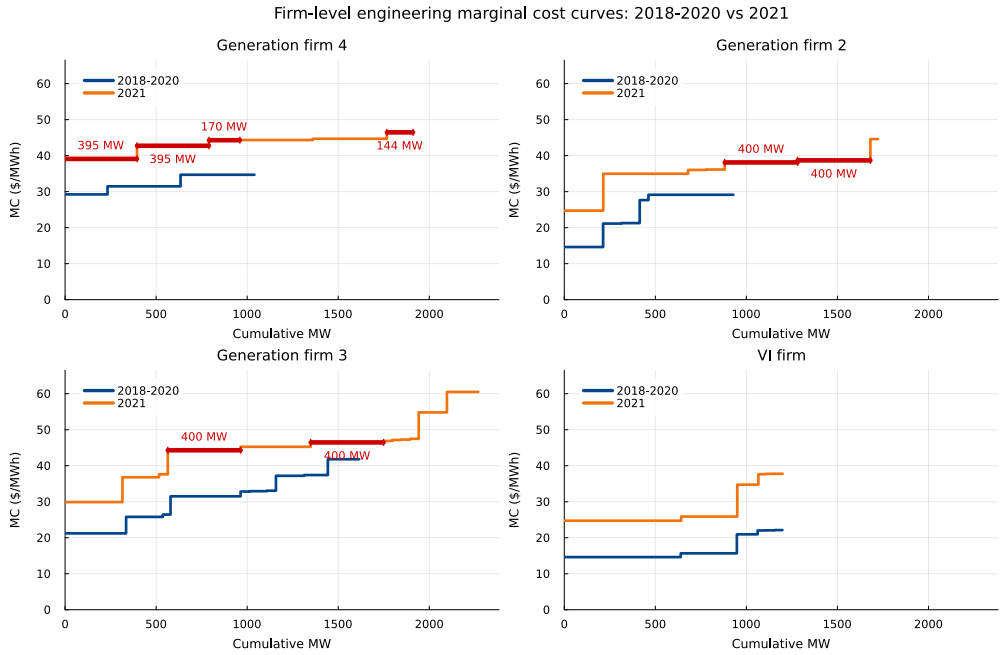


Figure 8: Changes in Marginal Costs from the End of PPA

Notes: The figure illustrates the firm level engineering marginal cost curves before and after the end of the PPA agreements at the end of 2020.

Appendix D.3 Simulation Approximation and Validation

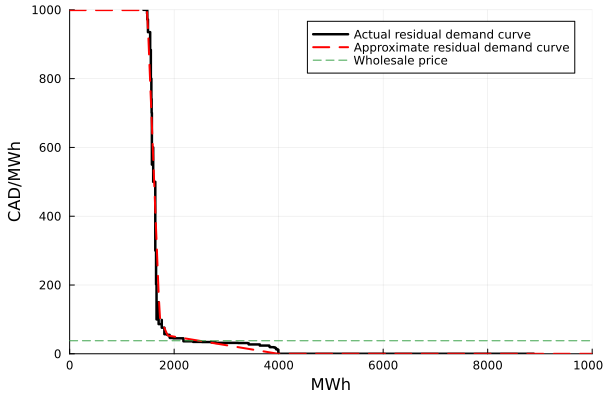


Figure 9: Residual Demand Approximation

Notes: The figure illustrates the piecewise linear approximation to the residual demand curve used in the simulation algorithm. The approximation captures the nonlinear shape of residual demand, including the high-price, high-demand region that is important for firms' hedging incentives.



Figure 10: Simulated and Actual Price Distributions

Notes: The figure compares the wholesale price distribution generated by the Swap+VI simulation with actual Alberta wholesale prices in 2018. The simulation does not use actual wholesale prices to calibrate the model parameters.

Appendix D.4 Additional Simulation Results

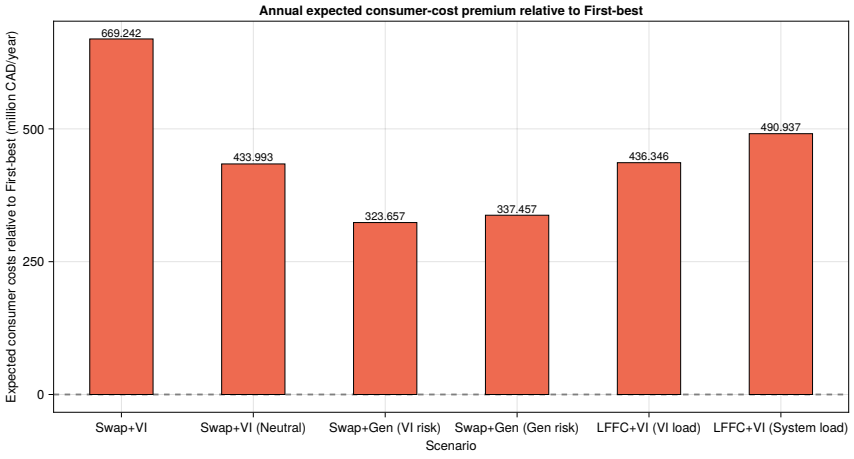


Figure 11: Average Annual Consumer Cost Increase Relative to the First-Best

Appendix E Empirical Application: Combining Risk Aversion and Imperfect Competition

We compute equilibrium forward positions over a wide grid of risk-aversion parameters, allowing separate values for generation-only firms and the vertically integrated firm. The grid ranges from $\lambda_i = 0$ to $\lambda_i = 0.001$, and we report results for each pair of risk-aversion parameters below.

Because firms are heterogeneous in size, we report net equilibrium spot exposure as a share of expected production—that is, the fraction of expected output sold at the spot price. Figure 12 shows equilibrium spot exposure ratios for generation-only firms and the vertically integrated firm across a wide grid of risk-aversion parameters, unifying strategic commitment and risk-hedging incentives. These figures correspond to the normalized counterparts of Figure 3.

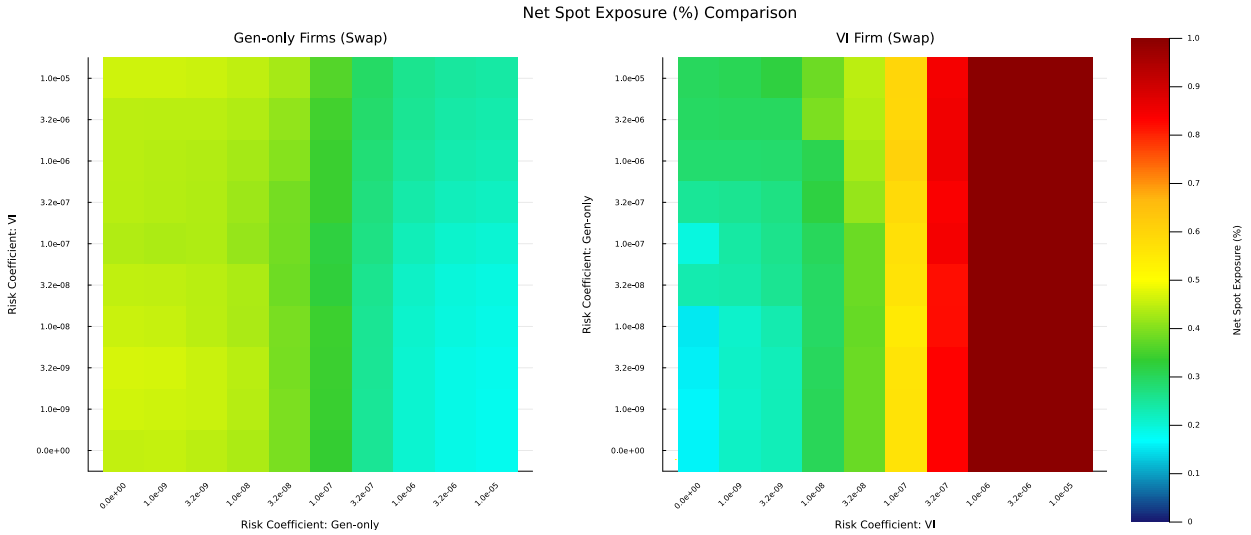


Figure 12: Equilibrium Spot Exposure Ratios under Swap Contracts

Notes: Using the simulations described in Section 5, we compute equilibrium spot exposure ratios—defined as equilibrium spot exposure relative to expected production— under swap contracts for each pair of risk-aversion parameters. The left panel reports the average ratio across the three generation-only firms, while the right panel reports the ratio for the vertically integrated firm.

The left panel reports the average equilibrium swap-based spot exposure ratio across generation-only firms for each pair of risk-aversion parameters. The model predicts spot exposure ratios of roughly 20–50 percent, which fall within a reasonable range compared with existing estimates in the literature (e.g., Reguant, 2014; Wolak, 2007) and industry

reports. The right panel reports the equilibrium spot exposure ratio for the VI firm. When risk aversion is low, the VI firm maintains a spot position comparable to that of generation-only firms. As risk aversion increases, however, the VI firm's spot exposure rises substantially, reflecting the over-hedging effect. These patterns are consistent with the predictions of our theoretical model: generation-only firms reduce their net spot exposure when using swaps for both strategic commitment and risk hedging, whereas the VI firm, despite reducing spot exposure for strategic reasons, may increase it under higher risk aversion due to over-hedging induced by contract incompleteness.

Figure 13 presents the corresponding results under load-following forward contracts. The left panel shows spot exposure ratios for generation-only firms, while the right panel shows those for the vertically integrated firm. Relative to the swap case, the VI firm's equilibrium spot exposure ratios are substantially closer to those of generation-only firms. Moreover, the comparative statics with respect to the VI firm's risk aversion are reversed: the equilibrium spot exposure ratio now decreases with risk aversion. Overall, these results are consistent with the theoretical prediction that load-following forward contracts correct the VI firm's over-hedging incentives and align its spot exposure more closely with that of generation-only firms.

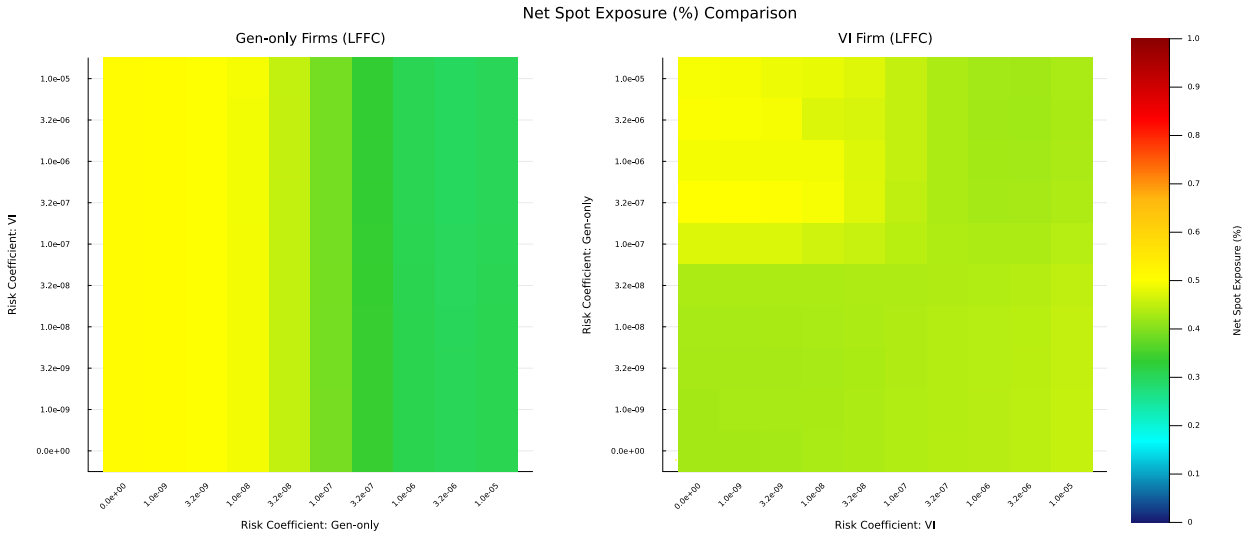


Figure 13: Equilibrium Spot Exposure Ratios under LFFC

Notes: Using the simulations described in Section 5, we compute equilibrium spot exposure ratios—defined as equilibrium spot exposure relative to expected production—under load-following forward contracts for each pair of risk-aversion parameters. The left panel reports the average ratio across the three generation-only firms, while the right panel reports the ratio for the vertically integrated firm.

Appendix F Robustness of Risk Aversion Parameters

We use the simulation model described in Section 5 to assess the robustness of our results to alternative risk-aversion parameters. Specifically, we solve

$$\min_{(\lambda_{GEN},\lambda_{VI})} \sum_i \left(f_i(\lambda_{GEN},\lambda_{VI}) - f_i^{GMM}\right)^2$$

by grid search. The minimizing parameter pair is $(\lambda^*_{GEN},\lambda^*_{VI}) = (1.0\times10^{-7},\ 3.16\times10^{-6})$, under which firms maximize CARA utility. Although the actual hedging strategies of retailers and vertically integrated firms are not publicly observed, industry and regulatory reports indicate that minimum-variance hedging is a common benchmark and provides a reasonable approximation to retailers' heuristic practices (Independent Competition and Regulatory Commission, 2019). Since $\lambda^*_{VI} > \lambda^*_{GEN}$, our estimates imply that vertically integrated firms are more risk averse than generation-only firms, which tend toward risk neutrality. This is consistent with the view that retailers are highly risk averse, though not necessarily risk minimizing.

To evaluate robustness, we conduct a series of sensitivity exercises across alternative risk-aversion values. Figure 14 reports the norm distance between the GMM-implied and model-predicted forward positions across a wide grid of parameter pairs, with the red star denoting the minimizing pair. Figures 15 and 16 show that both average profits and profit standard deviations decline as risk aversion rises, highlighting a trade-off between expected profit and risk reduction.

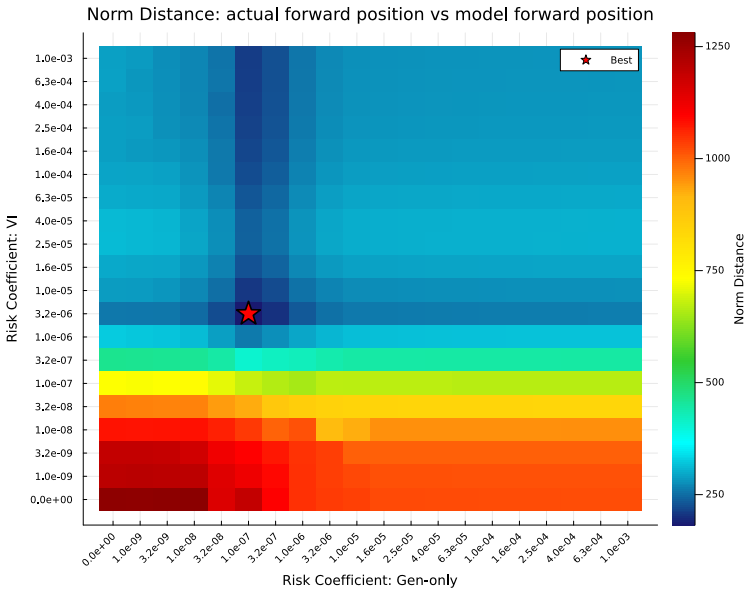


Figure 14: Norm Distance Between GMM and Model-Implied Forward Positions

*Notes:*This figure illustrates the norm distance between the GMM-implied and model-predicted forward positions, defined as $\sqrt{\left(f_i(\lambda_{GEN}, \lambda_{VI}) - f_i^{GMM}\right)^2}$. The simulation model is described in Section 5. The red star denotes the parameter pair $(\lambda_{GEN}, \lambda_{VI})$ that minimizes the norm distance.

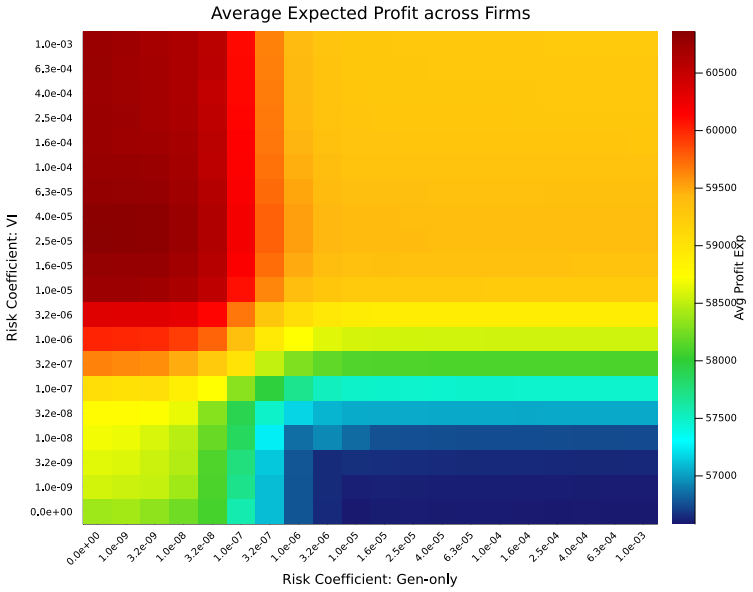


Figure 15: Average Profit Across Firms Under Alternative Risk Aversion Parameters

Notes: We use the simulation model described in Section 5. The figure reports the average expected profit across firms for alternative pairs of risk aversion parameters, $(\lambda_{GEN}, \lambda_{VI})$.

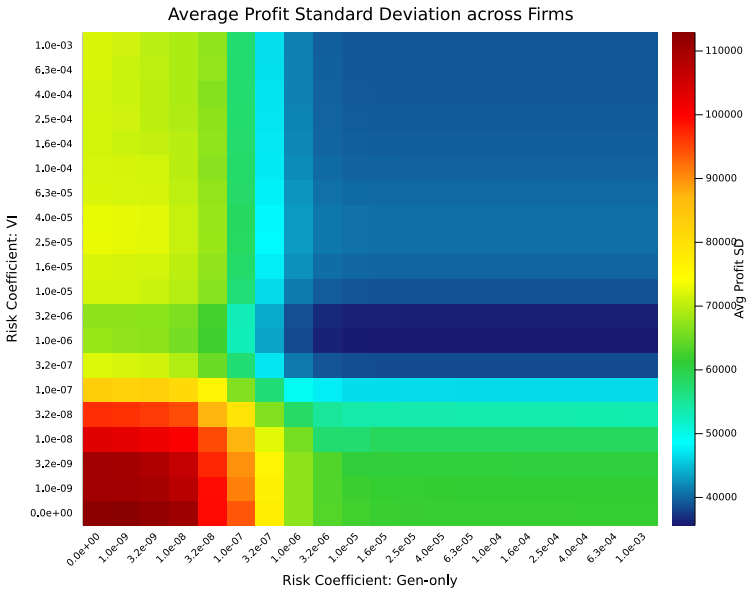


Figure 16: Average Profit Standard Deviation Across Firms Under Alternative Risk Aversion Parameters

*Notes:*We use the simulation model described in Section 5. The figure reports the average profit standard deviaion across firms for alternative pairs of risk aversion parameters, $(\lambda_{GEN}, \lambda_{VI})$.

As shown in Figures 17 and 18, increases in the VI firm's risk aversion parameter have only a modest effect on expected profits, while substantially reducing the variance of profits. This pattern indicates that the VI firm can achieve significant risk reduction with little sacrifice in expected profitability.

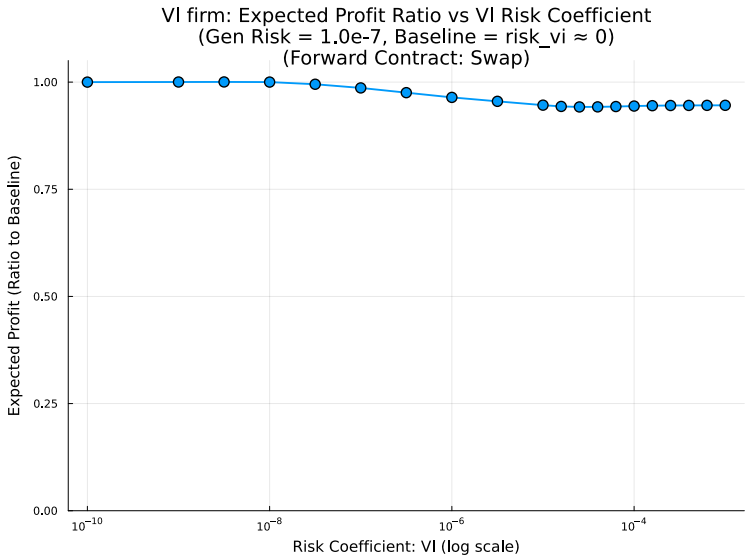


Figure 17: Expected Profit of the VI Firm Across VI Risk Aversion Parameters

*Notes:*This figure reports the VI firm's expected profit relative to the benchmark in which the VI firm is risk neutral. The producer's risk aversion is fixed at $\lambda_{GEN} = 1.0 \times 10^{-7}$, the value that best matches the GMM estimates.

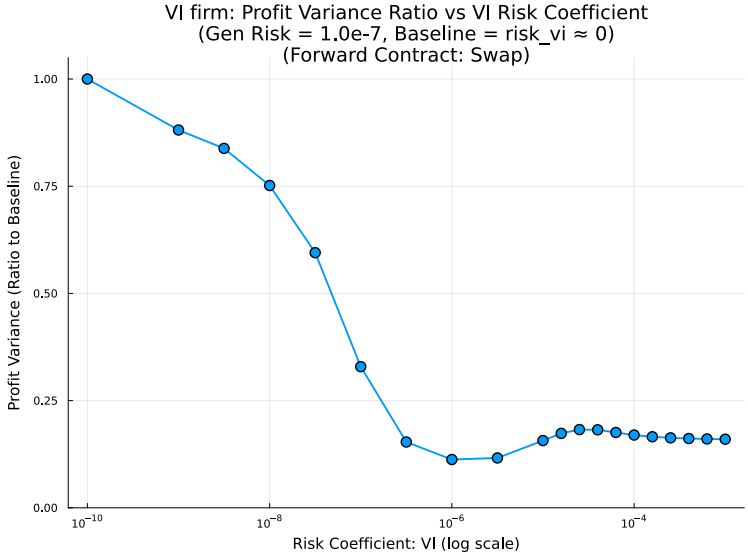


Figure 18: Profit Variance of the VI Firm Across VI Risk Aversion Parameters

*Notes:*This figure reports the VI firm's profit variance relative to the benchmark in which the VI firm is risk neutral. The producer's risk aversion is fixed at $\lambda_{GEN} = 1.0 \times 10^{-7}$, the value that best matches the GMM estimates.

Appendix G Additional Simulation Scenarios

In addition to Table 3, we consider two further scenarios:

- **Single market**: We shut down all forward markets, so firms compete only in the production market.

- **Swap+VI (LFFC 40%)**: Identical to Swap+VI, except the VI firm is required to hedge 40% of its own demand with LFFCs.

The single-market scenario captures the competitive and welfare impacts of abolishing the entire forward market. In the data, we observe when LFFCs become available, but we do not know how much the VI firm holds in LFFCs versus swaps, since a large part of their contract positions is not disclosed—which is why we estimate the forward positions. For the regulated option alone, however, the VI firm reports to the regulator the percentage of its regulated load hedged through LFFCs, which it states is 40%. Accordingly, in Swap+VI (LFFC 40%), we consider a scenario in which 40% of the VI firm's demand is procured through LFFCs and the remainder is hedged with swaps.

Table 4 reports the results. Relative to Swap+VI in Table 3, abolishing the forward market yields higher prices but lower deadweight loss. Swap+VI (LFFC 40%) suggests that even a 40% LFFC share achieves competitive and welfare impacts as large as those under LFFC+VI (System load).

Table 4: Summary of Outcomes Across Additional Simulation Scenarios

Scenario	E[p] (CAD/MWh)	Net Position (%)	Production (Thermal) (MWh)	E[π] (k CAD/MWh)	SD(π) (k CAD/MWh)	DWL (k CAD/h)	Cons. cost Δ from FB (k CAD/h)
Single market							
Gen-only (avg.)	84.5	0.00	407	58.8	163	3.77	84.1
VI		0.00	1156	66.9	514		
Swap+VI (LFFC 40%)							
Gen-only (avg.)	81.7	67.3	557	58.2	48.9	2.64	56.1
VI		29.7	816	59.8	41.7		

Note: The labels "Gen-only (avg.)" and "VI" refer to the three generation-only firms and the one vertically integrated firm observed in the data, respectively. These labels are retained across all scenarios for consistency, even in counterfactual scenarios such as Swap+Gen (VI risk) and Swap+Gen (Gen risk), where the VI firm is modeled as generation-only. Production, profits, and the standard deviation of profits are reported at the firm level. For "Gen-only (avg.)," firm-level variables are reported as averages across the three generation-only firms. Expected prices, deadweight loss, and consumer costs are reported at the market level. Deadweight loss is measured as the difference in total production costs between the market power scenario and the first-best benchmark. Changes in consumer costs are also computed relative to the first-best and represent the additional payments made by consumers compared to the first-best outcome. Consumer costs are defined as the market price multiplied by total demand.

Appendix H Track Ownership Changes of Assets

This appendix documents ownership changes for major thermal generation assets during the sample period. The table focuses on thermal assets owned by ATCO, Heartland Generation Ltd., ENMAX, Capital Power, TransAlta, and the Balancing Pool, which constitute the primary large firms in the Alberta market examined in the paper.

Timeline of Ownership and Asset Status Changes

- **2018-01-01**: Beginning of the analysis period.
- **2018-04-01**: Power Purchase Agreements (PPAs) for SD4, SD5, and SD6 expired; SD3 and SD5 exited the market.
- **2018-10-01**: PPA for BR5 expired.
- **2019-10-01**: Joint ownership of GN3 and KH3 was dissolved. GN3 became wholly owned by Capital Power, and KH3 became wholly owned by TransAlta.
- **2019-11-07**: ATCO's thermal assets were transferred to Heartland Generation Ltd.
- **2019-12-15**: BR3 exited the market.
- **2021-01-01**: PPAs for SH1, SH2, GN1, GN2, KH1, and KH2 expired.
- **2021-03-11**: HSM1 exited the market.
- **2021-12-31**: End of the analysis period. ‘‘‘

Based on these ownership transitions and asset exits, we redefine the re-sampling pool accordingly to ensure that firm ownership and asset availability are correctly aligned with each subperiod of the analysis.

Appendix I Alternative Forward Contract Specifications

In this section, we present estimates of forward contract positions and quantity exposures under alternative parametric restrictions on how forward positions vary over time. Let $\tau(t) \in \{\text{pre}, \text{post}\}$ indicate the sample period, $m(t)$ denote the month containing observation

t, and $s(t) \in \{\text{peak}, \text{off}\}$ indicate whether observation t occurs during the peak or off-peak period. Let $\bar{q}_m$ denote monthly expected production and $\bar{q}^R_m$ denote monthly retail load. For the peak/off-peak specifications, let $\bar{q}_{m,s}$ and $\bar{q}^R_{m,s}$ denote the corresponding monthly expected production and retail load. As in the main text, $\bar{q}_{ht}$ denotes expected production at observation t. Our baseline specification assumes that firm i's pre-committed quantity takes the form

$$\nu_t(\boldsymbol{\gamma}_i) = \gamma_{i,\tau(t)}\bar{q}_{ht} + \varepsilon_t. \tag{47}$$

Thus, all firms are assigned the same functional relationship between pre-committed quantities and expected production, while the coefficient $\gamma_{i,\tau(t)}$ is allowed to differ across firms and sample periods.

We use this common hedge-ratio specification in the main analysis for three reasons. First, it follows Reguant (2014), whose method provides the basis for our estimation of forward positions from bidding behavior. Using the same specification facilitates comparison with that paper and with the related literature.

Second, applying different functional forms to vertically integrated and generation-only firms would confound two sources of variation: differences in their actual forward positions and differences mechanically induced by the functional form imposed on each firm type. A common specification therefore provides an apples-to-apples comparison of their estimated forward positions. Importantly, the common specification does not require the two firm types to have the same estimated hedge ratio, since each firm has its own coefficient $\gamma_{i,\tau(t)}$.

Third, the common hedge-ratio specification corresponds directly to the prediction of the Allaz–Vila benchmark extended to include vertical integration. In this framework, the vertically integrated firm's retail obligation acts as a built-in forward position that is a perfect substitute for an explicit forward contract in the spot-market first-order condition. The producer's pre-committed quantity is

$$\nu_1 = f^*_{1,N}, \tag{48}$$

whereas the vertically integrated firm's pre-committed quantity is

$$\nu_2 = f^*_{2,N} + \theta. \tag{49}$$

Result 1 establishes that

$$f^*_{1,N} = f^*_{2,N} + \theta, \tag{50}$$

and hence that $\nu_1 = \nu_2$. Under the symmetric benchmark, expected production is also the same across the two firms, implying

$$\frac{\nu_1}{\mathbb{E}[q^*_1]} = \frac{\nu_2}{\mathbb{E}[q^*_2]}. \tag{51}$$

Because the Allaz–Vila model is the standard benchmark for strategic forward contracting, imposing a common functional form provides a direct test of its predicted common hedge ratio rather than building differences between the two firm types into the estimation specification.

We nevertheless examine two alternative specifications to assess whether the estimated forward positions and quantity exposures are sensitive to the baseline restriction.

We consider two alternative specifications. The first, restricts firms to choose a fixed share of monthly expected peak and off-peak demand to hedge:

$$\nu_t(\boldsymbol{\gamma}_i) = \gamma_{i,\tau(t)} \bar{q}_{m(t),s(t)} + \varepsilon_t.$$

This implies two different levels of forward positions (one in peak and one in off-peak hours). The specification is designed to capture the fact that the most commonly traded forward product in the data are monthly contracts. This could make it difficult for a firm to hedge a fixed share of the expected retail demand which moves at the hourly level.

The second, incorporates data on the retail demand of the integrated firm. In particular we obtained data from every hour on the metered volume attributable to the integrated firm. Then, we parameterize the short forward position f_t as

$$f_t(\boldsymbol{\gamma}_i) = \gamma_{i,\tau(t)} \left(\bar{q}^R_{m(t),s(t)} - \bar{q}_{m(t),s(t)} \right).$$

which we assume maps into the pre-committed quantity as:

$$\nu_t(\boldsymbol{\gamma}_i) = \bar{q}^R_t - f_t(\boldsymbol{\gamma}_i) + \varepsilon_t. \tag{52}$$

This specification allows the firm to hedge its net physical exposure, the gap between retail demand and expected production, rather than production alone. The specification

parametrizes the hedge off a predetermined month-by-segment block rather than an hour-varying quantity (reflecting the use of monthly contracts) so the forward position is constant within a segment and does not inherit hour-to-hour variation, but the pre-committed quantity moves with changes in expected retail demand for the integrated firm.

The results are presented in Table 5. Although the parameters have different interpretations across specifications, the economically relevant implied quantities, including forward positions, pre-committed quantities, physical exposure, and inframarginal quantities, are similar. In particular, explicitly incorporating the vertically integrated firm's retail load or restricting forward positions to monthly peak and off-peak blocks does not materially change its estimated exposure. These results indicate that the substantive conclusions are not driven by the common hedge-ratio specification used in the baseline analysis.

Table 5: Implied Short Forward Positions and Inframarginal Quantities

Spec.	Period	Estimate	$\widehat{q}^{\text{pre}}$ pre-committed quantity	$\widehat{f}$ short forward position	$\bar{q} - \bar{q}^{ret}$ physical exposure	$\bar{q} - \widehat{q}^{\text{pre}}$ inframarginal quantity
Baseline	Pre	−0.343	−291.1	−2496.9	−1357.1	1139.8
Baseline	Post	0.880	787.6	−1026.4	−918.9	107.4
1	Pre peak	−0.342	−309.7	−2644.7	−1429.6	1215.1
1	Pre off-peak	−0.342	−251.5	−2198.7	−1212.0	986.7
1	Post peak	0.939	884.0	−1040.9	−983.5	57.4
1	Post off-peak	0.939	753.4	−839.7	−790.7	48.9
2	Pre peak	1.897	−376.9	−2712.0	−1429.6	1282.4
2	Pre off-peak	1.897	−351.9	−2299.1	−1212.0	1087.1
2	Post peak	0.905	1034.8	−890.1	−983.5	−93.4
2	Post off-peak	0.905	877.5	−715.6	−790.7	−75.1

Notes: The table reports estimates for the vertically integrated firm under the baseline specification and the two alternative specifications described in the text. "Estimate" denotes the estimated coefficient $\widehat{\gamma}_{i,\tau}$. Its interpretation varies across specifications: in the baseline it scales hourly expected production, in Specification 1 it scales monthly expected production separately for peak and off-peak hours, and in Specification 2 it scales the firm's monthly net retail position. $\widehat{q}^{\text{pre}}$ denotes the implied pre-committed quantity, and $\widehat{f}$ denotes the implied explicit short forward position. Positive values of $\widehat{f}$ indicate net forward sales, while negative values indicate net forward purchases. $\bar{q}$ and $\bar{q}^{\text{ret}}$ denote average expected production and average retail load, respectively. Physical exposure, $\bar{q} - \bar{q}^{\text{ret}}$, is the firm's net physical position before accounting for explicit forward contracts. The pre-committed quantity, $\widehat{q}^{\text{pre}}$, is the portion of expected production offset by retail obligations and explicit forward contracts before spot-market competition. Physical exposure, $\bar{q} - \bar{q}^{\text{ret}}$, measures the firm's net wholesale position based only on expected production and retail load, before accounting for financial forward contracts. Inframarginal quantity, $\bar{q} - \widehat{q}^{\text{pre}}$, measures the firm's remaining net spot-price exposure after accounting for both retail obligations and explicit forward positions. Positive values indicate net selling exposure, while negative values indicate net buying exposure. All quantity measures are reported in MWh.

Appendix J Unit-Level Marginal Cost Estimates

Table 6: Unit-Level Marginal Cost Estimates for Large Firms

			Before 2021			In 2021		
Firm	Unit	Fuel	GMM MC ($/MWh)	(SE)	Eng. MC ($/MWh)	GMM MC ($/MWh)	(SE)	Eng. MC ($/MWh)
Generation firm 1	Unit 1	GAS	15.37	(0.73)	19.40	–	–	–
	Unit 2	GAS	17.61	(1.45)	25.23	–	–	–
	Unit 3	COAL	19.33	(4.32)	37.48	–	–	–
	Unit 4	GAS	19.98	(0.76)	32.81	–	–	–
	Unit 5	GAS	20.18	(1.86)	25.23	–	–	–
	Unit 6	GAS	24.09	(0.86)	15.50	–	–	–
	Unit 7	GAS	26.25	(1.61)	19.98	–	–	–
	Unit 8	GAS	27.25	(2.39)	25.23	–	–	–
	Unit 9	GAS	27.27	(1.58)	25.23	–	–	–
	Unit 10	COAL	27.47	(2.06)	31.46	–	–	–
	Unit 11	GAS	29.09	(1.08)	28.02	–	–	–
	Unit 12	COAL	36.61	(0.47)	36.02	–	–	–
Generation firm 2	Unit 1	COAL	23.30	(2.68)	–	–	–	–
	Unit 2	COAL	20.91	(2.43)	29.13	27.88	(3.72)	34.93
	Unit 3	COAL	–	–	–	26.77	(4.04)	38.55
	Unit 4	GAS	20.66	(2.17)	14.55	34.41	(3.62)	24.10
	Unit 5	COAL	–	–	–	29.21	(4.01)	38.55
	Unit 6	GAS	27.04	(2.31)	21.08	45.04	(3.84)	36.04
	Unit 7	GAS	27.26	(2.68)	21.08	45.41	(4.47)	36.04
Generation firm 3	Unit 1	GAS	16.77	(1.10)	25.86	24.86	(1.83)	36.80
	Unit 2	GAS	21.98	(1.62)	41.85	32.57	(2.70)	60.49
	Unit 3	COAL	27.47	(2.06)	31.39	–	–	–
	Unit 4	GAS	24.74	(1.15)	21.21	36.67	(1.92)	29.91
	Unit 5	GAS	24.90	(4.82)	36.61	36.91	(8.04)	52.74
	Unit 6	COAL	32.20	(4.74)	37.30	–	–	–
	Unit 7	GAS	26.96	(4.11)	32.85	39.67	(6.85)	46.45
	Unit 8	COAL	–	–	–	34.28	(7.27)	36.85
	Unit 9	GAS	29.83	(2.55)	31.28	44.60	(4.24)	45.15
	Unit 10	GAS	33.49	(2.10)	26.55	49.64	(3.49)	37.83
	Unit 11	GAS	–	–	–	48.74	(7.55)	51.89
	Unit 12	GAS	43.21	(13.67)	32.81	64.04	(22.78)	47.10
	Unit 13	GAS	54.99	(133.01)	32.81	81.50	(221.59)	47.10
	Unit 14	GAS	139.39	(311.75)	32.81	206.59	(519.37)	47.10
Generation firm 4	Unit 1	COAL	27.03	(0.80)	29.13	–	–	–
	Unit 2	COAL	27.77	(0.57)	–	27.77	(1.17)	–
	Unit 3	COAL	29.38	(3.77)	31.35	32.05	(6.45)	35.63
	Unit 4	COAL	31.42	(3.18)	34.69	41.89	(3.60)	42.11
	Unit 5	COAL	–	–	–	39.54	(2.90)	38.95
	Unit 6	GAS	–	–	–	41.25	(2.61)	45.87
	Unit 7	COAL	–	–	–	42.43	(1.74)	36.85
	Unit 8	GAS	–	–	–	45.42	(4.87)	51.83
	Unit 9	GAS	40.45	(2.96)	32.85	59.53	(4.94)	46.45
	Unit 10	COAL	–	–	–	57.24	(4.97)	36.85
	Unit 11	GAS	–	–	–	75.91	(8.39)	51.89
VI firm	Unit 1	GAS	15.28	(1.70)	20.93	25.46	(2.83)	34.73
	Unit 2	GAS	15.33	(1.88)	15.56	25.53	(3.13)	25.78
	Unit 3	GAS	17.62	(1.74)	14.55	29.36	(2.90)	24.10
	Unit 4	GAS	26.29	(6.72)	22.12	43.80	(11.19)	37.77
	Unit 5	GAS	29.01	(7.81)	22.12	48.33	(13.02)	37.77
	Unit 6	GAS	43.91	(7.59)	22.12	73.15	(12.64)	37.77

Notes: The sample covers 2018–2021. Marginal costs are estimated at the unit level using the GMM estimator. The GMM marginal costs are computed at each unit's average output level. Generation Firm 1 exits during the sample period. Several assets either retire or convert to different fuel types. Engineering cost estimates are unavailable for two assets.

Appendix K Risk Aversion in Forward Choices and Risk-Neutral Bidding

In the Cournot model, we assume that firms choose forward positions using a risk-averse objective in the first stage, but choose production according to a profit-maximization objective without uncertainty in the second stage. This distinction provides a tractable model of forward contracting and bidding. In practice, however, firms may still face residual uncertainty at the bidding stage, even after most demand uncertainty has been resolved. This residual uncertainty may arise from forecast errors, renewable output deviations, or private information held by competitors. This raises the question of whether a risk-neutral bidding condition is a good approximation when estimating forward positions.

One interpretation of this modeling choice is organizational. The firm as a whole may be risk neutral, but the forward-stage risk-averse objective provides a reduced-form approximation to the risk-management desk's hedge choice. The parameter λ then captures the shadow cost of low-cash-flow states that matters for forward contracting, rather than risk aversion in every firm decision. The risk-management desk chooses the forward position taking these low-frequency cash-flow concerns into account, while anticipating that the dispatch desk will later choose production or bids according to an expected-profit continuation problem. This separation is also consistent with industry practice. Hortaçsu and Puller (2008) report, based on interviews with ERCOT market participants, that traders who negotiate forward contracts and traders who bid in the real-time balancing market are distinct and often belong to different trading desks.

Under this interpretation, risk aversion governs the hedge decision, whereas the local hourly bidding first-order condition is governed by expected profits. The risk-management desk has already purchased the relevant insurance through the forward position, so the dispatch desk does not need to obtain additional insurance through local hourly bid distortions. Instead, the dispatch desk takes the forward position as given and solves the expected-profit continuation problem. The model is dynamically consistent because the bidding behavior anticipated by the risk-management desk coincides with the continuation problem solved by the dispatch desk. Below, we consider a second interpretation in which the risk-averse objective also applies to hourly bidding, and show that the risk-neutral bidding condition remains a good approximation.

Intuitively, forward hedging decisions are made at a longer horizon, when firms have limited information about realized demand and system conditions. By contrast, real-time bidding takes the form of a supply schedule. A local bid perturbation affects market outcomes only in states in which that bid sets the market-clearing price with positive probability. Hence, a necessary condition for optimality is that no local deviation is profitable: at each step of the supply curve, the firm must not gain by marginally raising or lowering the offer. The relevant risk is therefore conditional on the step being marginal, rather than the unconditional risk faced when choosing forward positions. This conditioning reveals information about the realized residual demand curve and partially conditions the firm's bidding incentive on the aggregate state. In the continuous-schedule benchmark of Hortaçsu and Puller (2008), the supply curve provides full state contingency by mapping each market-clearing state into a quantity, so the risk-averse and risk-neutral bidding conditions coincide.

The remainder of this section provides an empirical diagnostic for this second interpretation. We show that, conditional on a bid step setting the market-clearing price, the residual variation relevant for the local bid perturbation is small. Thus, the risk relevant for forward-contracting decisions is much larger than the residual risk relevant for marginal real-time bidding incentives, and the risk-averse correction to the bidding first-order condition is quantitatively negligible.

Consider the CARA expected-utility objective

$$\mathbb{E}[-\exp(-\lambda\,\pi_i(\mathbf{b}))]\,, \qquad \lambda > 0, \tag{53}$$

where λ is the coefficient of absolute risk aversion. In the step-function bidding problem, a marginal change in bid b_{ijkh} affects market outcomes only in states in which that bid step sets the market-clearing price. Therefore, after dropping the strictly positive multiplicative constant λ, the CARA first-order condition can be written as

$$\mathbb{E}[\exp(-\lambda\,\pi_i)\,\Delta_{ijkh} \mid p_h = b_{ijkh}] = 0, \tag{54}$$

where

$$\Delta_{ijkh} = (Q_{ih} - \nu_{ih})\frac{\partial p_h}{\partial b_{ijkh}} + \big(p_h - C_i'\big)\frac{\partial Q_{ih}}{\partial b_{ijkh}},$$

and

$$\pi_i(\mathbf{b}) = p_h(\mathbf{b}_h)\, Q_{ih}(\mathbf{b}_h) - \sum_j C_{ij}\big(q_{ij}(\mathbf{b})\big) - \big(p_h(\mathbf{b}_h) - p^f\big)\nu_{ih}.$$

Define conditional expected profit and the corresponding conditional profit deviation as

$$\mu_{ijkh} = \mathbb{E}[\pi_i \mid p_h = b_{ijkh}]$$

and

$$Z_{ijkh} = \pi_i - \mu_{ijkh}.$$

By construction,

$$\mathbb{E}[Z_{ijkh} \mid p_h = b_{ijkh}] = 0.$$

Since

$$\pi_i = \mu_{ijkh} + Z_{ijkh},$$

we have

$$\exp(-\lambda\, \pi_i) = \exp(-\lambda\, \mu_{ijkh}) \exp(-\lambda\, Z_{ijkh}).$$

Conditional on $p_h = b_{ijkh}$, the term $\exp(-\lambda\, \mu_{ijkh})$ is a strictly positive constant. Thus, equation (54) is equivalent to

$$\mathbb{E}[\exp(-\lambda\, Z_{ijkh})\Delta_{ijkh} \mid p_h = b_{ijkh}] = 0. \tag{55}$$

For small λ, a Taylor expansion around $\lambda = 0$ gives[24]

$$\exp(-\lambda\, Z_{ijkh}) = 1 - \lambda\, Z_{ijkh} + r_{ijkh}(\lambda),$$

[24]The accuracy of the approximation depends on the magnitude of the dimensionless product λZ_{ijkh}, where $Z_{ijkh} \equiv \pi_i - \mathbb{E}[\pi_i \mid p_h = b_{ijkh}]$. Suppose that λ is the same risk-aversion parameter governing the firm's forward-position choice. Expected profit is measured in monetary units, whereas the variance of profit is measured in squared monetary units. The parameter λ therefore rescales the unconditional variance penalty so that it is comparable to expected profit. Because the unconditional variance of profit is large relative to the level of expected profit, a firm that trades off expected profit against variance, rather than simply minimizing variance, must have a numerically small λ. Consistent with this scaling, our estimates are $(\lambda^*_{\text{GEN}},\ \lambda^*_{\text{VI}}) = \big(1.0 \times 10^{-7},\ 3.16 \times 10^{-6}\big)$.

Moreover, Z_{ijkh} is a profit *deviation* measured in levels (monetary units), not a squared quantity, and it is defined conditionally: it captures profit variation only across states in which the same bid step sets the market-clearing price. Conditioning removes the large variation that arises across different price-setting segments, so Z_{ijkh} is small relative to the overall dispersion of profit. Thus, the product λZ_{ijkh} is plausibly small, supporting the first-order Taylor approximation around $\lambda = 0$.

where

$$r_{ijkh}(\lambda) = \frac{\lambda^2 Z_{ijkh}^2}{2} - \frac{\lambda^3 Z_{ijkh}^3}{6} + \cdots .$$

Substituting this expansion into equation (55) gives

$$\begin{aligned} 0 = \mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}] \\ - \lambda\, \mathbb{E}[Z_{ijkh}\Delta_{ijkh} \mid p_h = b_{ijkh}] \\ + \mathbb{E}[r_{ijkh}(\lambda)\Delta_{ijkh} \mid p_h = b_{ijkh}]\,. \end{aligned} \tag{56}$$

The first-order approximation of equation (56) becomes

$$\mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}] - \lambda\, \mathbb{E}[Z_{ijkh}\Delta_{ijkh} \mid p_h = b_{ijkh}] \approx 0. \tag{57}$$

Because Z_{ijkh} has conditional mean zero,

$$\mathbb{E}[Z_{ijkh}\Delta_{ijkh} \mid p_h = b_{ijkh}] = \mathrm{Cov}(\pi_i, \Delta_{ijkh} \mid p_h = b_{ijkh})\,.$$

Therefore, the approximated CARA first-order condition is

$$\mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}] - \lambda\ \mathrm{Cov}(\pi_i, \Delta_{ijkh} \mid p_h = b_{ijkh}) \approx 0. \tag{58}$$

Let

$$\widetilde{\pi}_i = p_h(\mathbf{b}_h)\, Q_{ih}(\mathbf{b}_h) - \sum_j C_{ij}\big(q_{ij}(\mathbf{b})\big)$$

denote variable profit excluding the forward payoff. Since

$$\pi_i = \widetilde{\pi}_i - (p_h - p^f)\nu_{ih},$$

and the forward payoff is constant conditional on $p_h = b_{ijkh}$, we have

$$\mathrm{Cov}(\pi_i, \Delta_{ijkh} \mid p_h = b_{ijkh}) = \mathrm{Cov}(\widetilde{\pi}_i, \Delta_{ijkh} \mid p_h = b_{ijkh})\,.$$

Thus,

$$\mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}] - \lambda\ \mathrm{Cov}(\widetilde{\pi}_i, \Delta_{ijkh} \mid p_h = b_{ijkh}) \approx 0. \tag{59}$$

The first term,

$$\mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}],$$

is the risk-neutral marginal-profit term used in our estimation. The second term is the first-order risk-aversion correction implied by CARA utility.

By Cauchy–Schwarz,

$$|\mathrm{Cov}(\widetilde{\pi}_i, \Delta_{ijkh} \mid p_h = b_{ijkh})| \leq \sqrt{\mathrm{Var}[\widetilde{\pi}_i \mid p_h = b_{ijkh}]}\sqrt{\mathrm{Var}[\Delta_{ijkh} \mid p_h = b_{ijkh}]}.$$

Thus, the absolute contribution of the CARA covariance correction is bounded by

$$\begin{aligned} &|\lambda \ \mathrm{Cov}(\widetilde{\pi}_i, \Delta_{ijkh} \mid p_h = b_{ijkh})| \\ &\leq \lambda \sqrt{\mathrm{Var}[\widetilde{\pi}_i \mid p_h = b_{ijkh}]}\sqrt{\mathrm{Var}[\Delta_{ijkh} \mid p_h = b_{ijkh}]}. \end{aligned} \tag{60}$$

We use this bound to assess whether the first-order CARA correction is quantitatively important relative to the risk-neutral marginal-profit term. For each bid step, define

$$R^{\mathrm{CARA}}_{ijkh} \equiv \frac{\lambda \sqrt{\mathrm{Var}[\widetilde{\pi}_i \mid p_h = b_{ijkh}]}\sqrt{\mathrm{Var}[\Delta_{ijkh} \mid p_h = b_{ijkh}]}}{|\mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}]|}. \tag{61}$$

The ratio R^{CARA}_{ijkh} is the upper bound on the absolute CARA covariance correction divided by the absolute risk-neutral marginal-profit term.

The intuition is that conditioning on $p_h = b_{ijkh}$ removes the main source of payoff variation: variation in the market-clearing price. Once the marginal bid is fixed, the remaining variation in $\widetilde{\pi}_i$ and Δ_{ijkh} comes primarily from quantity movements within the set of states in which the same bid step sets the price. In particular, the residual variation is driven by fluctuations in the dispatched quantity of the marginal segment, holding the price fixed at b_{ijkh}. This substantially reduces the conditional variance of profits and therefore limits the possible magnitude of the covariance correction. If R^{CARA}_{ijkh} is small, then the second term in (59) is small relative to the first term, and the risk-neutral first-order condition provides a good approximation to the risk-averse bidding condition.

The objects in $\widetilde{\pi}_i$, Δ_{ijkh}, and κ_{ijkh} depend on the structural cost parameters α_i and on the forward position ν_{ih}. To evaluate the diagnostic under plausible parameter values, we

use engineering cost values for α_i. For forward positions, we consider hedge ratios

$$\gamma_i \in \{-1, -0.5, 0, 0.5, 1\},$$

as well as the estimated hedge ratios $\gamma_i = -0.34$ and $\gamma_i = 0.77$. We then recompute R^{CARA}_{ijkh} under each specification. This exercise tests whether the covariance correction remains small across a wide range of cost and forward position assumptions.

In the empirical exercise, we report R^{CARA}_{ijkh} using the estimated value of λ. Although λ is estimated within the model, the implied forward positions are consistent with firms' annual reports and industry descriptions. This external consistency suggests that the estimated value of λ captures economically plausible hedging incentives rather than being driven solely by within-model variation. It is therefore best to interpret λ as a plausible order-of-magnitude benchmark rather than a precise point estimate. As we show below, the diagnostic is not sensitive to this choice: scaling λ up by a factor of 10 or 100 leaves the conclusion unchanged.

Table 7 reports summary statistics for R^{CARA}_{ijkh} across alternative values of γ_i. The ratios are small across the relevant bid steps and parameter specifications. Thus, the risk-aversion correction is quantitatively negligible relative to the marginal effect of bids on expected profits.

A useful benchmark for interpreting Table 7 is the analogous relative-magnitude diagnostic in Reguant (2014). The omitted term in that setting is different: it captures the effect of marginal changes in simple hourly bids on expected profits through changes in the probability that a complex bid is binding. In our setting, the omitted term is the risk-averse covariance correction. Nevertheless, the logic of the diagnostic is similar: compare the component omitted from the estimating FOC with the marginal effect of bids on expected profits, which is the component used in estimation. Reguant reports that the omitted component is no greater than 10.55 percent of the marginal-profit component on average, with median and interquartile range equal to zero. Our exercise applies the same type of comparison in a more conservative way: we report a Cauchy–Schwarz upper bound on the covariance correction rather than the covariance correction itself; we fix marginal costs at engineering estimates rather than choosing the cost parameters that minimize the GMM criterion; and we evaluate the diagnostic not only at the estimated forward position parameters but also across a range of candidate forward position values, including values far from the estimated hedge ratios.

Thus, Table 7 should be read as a conservative robustness check showing that the omitted risk-averse correction is small relative to the risk-neutral marginal-profit component of the bidding FOC.

To make this comparison transparent, Tables 8 and 9 report the two components of the risk-averse FOC separately. Table 8 reports the Cauchy–Schwarz upper bound on the absolute covariance correction, while Table 9 reports the absolute value of the risk-neutral marginal-profit term, $|\mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}]|$. These tables can be read as residual tables for the local bidding FOC. For a candidate value of γ_i, holding the engineering cost parameters fixed, Table 9 reports the sample analogue of the risk-neutral FOC residual. In theory, the correct cost and forward position parameters set this object equal to zero. In the data, the sample analogue remains nonzero because of marginal cost shocks, forward position shocks, measurement error, and other approximation errors. The GMM estimator chooses parameters that make these FOC residuals as small as possible, but some residual variation is irreducible.

This residual-table interpretation is visible in the estimates. For example, for Generation firm 4, $\gamma = 0.77$ is close to the GMM-estimated forward position parameter. At this value, Table 9 reports a mean risk-neutral FOC residual of 488, compared with 1780 at $\gamma = -1$, a clearly implausible forward position value. Relative to typical hourly profits of about 60,000 CAD for generation-only firms, reported in Table 3, these values correspond to about 0.8 percent and 3.0 percent of profits, respectively. Thus, moving toward the empirically relevant forward position value substantially reduces the risk-neutral FOC residual. By contrast, Table 8 reports covariance-correction bounds of about 0.5 at both values of γ, which is negligible on the same profit scale. The same logic applies to the vertically integrated firm. The covariance correction remains small relative to the corresponding risk-neutral marginal-profit residual.

The scale of these objects should be interpreted as a profit scale. The local FOC component can be written as

$$\Delta_{ijkh} = (Q_{ih} - \nu_{ih}) + \big(p_h - C_i'\big)\frac{\partial Q_{ih}}{\partial b_{ijkh}},$$

Thus, Δ_{ijkh} is the local change in profit from perturbing a bid. The price–MC wedge enters multiplied by the local quantity response, $\partial Q_{ih}/\partial b_{ijkh}$, and the FOC also contains the infra-marginal exposure term $Q_{ih} - \nu_{ih}$. A natural benchmark for these magnitudes is therefore

the firm's profit from the relevant hourly market outcome. On this scale, the risk-neutral residuals in Table 9 are the economically relevant residuals that the GMM criterion is designed to minimize, while the covariance-correction bounds in Table 8 are much smaller. This comparison reinforces the conclusion that the empirically relevant component of the bidding first-order condition is the risk-neutral marginal-profit term, and that the omitted risk-averse correction is quantitatively second order.

Table 7: Relative Magnitudes in the CARA FOC

Firm	γ	Mean	Median	p25	p75
VI firm 1	-1.00	0.0016	0.0000	0.0000	0.0005
	-0.50	0.0008	0.0000	0.0000	0.0005
	-0.34	0.0003	0.0000	0.0000	0.0005
	0.00	0.0006	0.0000	0.0000	0.0004
	0.50	0.0004	0.0000	0.0000	0.0003
	0.77	0.0007	0.0000	0.0000	0.0003
	1.00	0.0002	0.0000	0.0000	0.0002
Generation firm 1	-1.00	0.0000	0.0000	0.0000	0.0000
	-0.50	0.0000	0.0000	0.0000	0.0000
	-0.34	0.0000	0.0000	0.0000	0.0000
	0.00	0.0000	0.0000	0.0000	0.0000
	0.50	0.0001	0.0000	0.0000	0.0000
	0.77	0.0000	0.0000	0.0000	0.0000
	1.00	0.0000	0.0000	0.0000	0.0000
Generation firm 2	-1.00	0.0000	0.0000	0.0000	0.0000
	-0.50	0.0275	0.0000	0.0000	0.0000
	-0.34	0.0001	0.0000	0.0000	0.0000
	0.00	0.0001	0.0000	0.0000	0.0000
	0.50	0.0002	0.0000	0.0000	0.0000
	0.77	0.0001	0.0000	0.0000	0.0000
	1.00	0.0002	0.0000	0.0000	0.0001
Generation firm 3	-1.00	0.0000	0.0000	0.0000	0.0000
	-0.50	0.0000	0.0000	0.0000	0.0000
	-0.34	0.0000	0.0000	0.0000	0.0000
	0.00	0.0007	0.0000	0.0000	0.0000
	0.50	0.0000	0.0000	0.0000	0.0000
	0.77	0.0001	0.0000	0.0000	0.0000
	1.00	0.0001	0.0000	0.0000	0.0000
Generation firm 4	-1.00	0.0007	0.0000	0.0000	0.0000
	-0.50	0.0010	0.0000	0.0000	0.0001
	-0.34	0.0035	0.0000	0.0000	0.0001
	0.00	0.0032	0.0000	0.0000	0.0001
	0.50	0.0130	0.0001	0.0000	0.0003
	0.77	0.0021	0.0001	0.0000	0.0007
	1.00	0.0045	0.0001	0.0000	0.0008

Notes: This table reports summary statistics of R_{ijkh}^{CARA}. The sample is restricted to bid-step cells with at least 20 observations.

Table 8: Upper Bound on the Absolute Covariance Correction

Firm	γ	Mean	Median	p25	p75
VI firm 1	-1.00	0.2296	0.0000	0.0000	0.4353
	-0.50	0.2296	0.0000	0.0000	0.4353
	-0.34	0.2296	0.0000	0.0000	0.4353
	0.00	0.2296	0.0000	0.0000	0.4353
	0.50	0.2296	0.0000	0.0000	0.4353
	0.77	0.2296	0.0000	0.0000	0.4353
	1.00	0.2296	0.0000	0.0000	0.4353
Generation firm 1	-1.00	0.0085	0.0031	0.0000	0.0133
	-0.50	0.0085	0.0031	0.0000	0.0133
	-0.34	0.0085	0.0031	0.0000	0.0133
	0.00	0.0085	0.0031	0.0000	0.0133
	0.50	0.0085	0.0031	0.0000	0.0133
	0.77	0.0085	0.0031	0.0000	0.0133
	1.00	0.0085	0.0031	0.0000	0.0133
Generation firm 2	-1.00	0.0485	0.0074	0.0034	0.0232
	-0.50	0.0485	0.0074	0.0034	0.0232
	-0.34	0.0485	0.0074	0.0034	0.0232
	0.00	0.0485	0.0074	0.0034	0.0232
	0.50	0.0485	0.0074	0.0034	0.0232
	0.77	0.0485	0.0074	0.0034	0.0232
	1.00	0.0485	0.0074	0.0034	0.0232
Generation firm 3	-1.00	0.0239	0.0000	0.0000	0.0000
	-0.50	0.0239	0.0000	0.0000	0.0000
	-0.34	0.0239	0.0000	0.0000	0.0000
	0.00	0.0239	0.0000	0.0000	0.0000
	0.50	0.0239	0.0000	0.0000	0.0000
	0.77	0.0239	0.0000	0.0000	0.0000
	1.00	0.0239	0.0000	0.0000	0.0000
Generation firm 4	-1.00	0.5448	0.0158	0.0000	0.0579
	-0.50	0.5448	0.0158	0.0000	0.0579
	-0.34	0.5448	0.0158	0.0000	0.0579
	0.00	0.5448	0.0158	0.0000	0.0579
	0.50	0.5448	0.0158	0.0000	0.0579
	0.77	0.5448	0.0158	0.0000	0.0579
	1.00	0.5448	0.0158	0.0000	0.0579

Notes: This table reports summary statistics for the Cauchy–Schwarz upper bound on the absolute covariance correction in the CARA bidding first-order condition. The reported object is $\lambda\sqrt{\mathrm{Var}\left[\widetilde{\pi}_i \mid p_h = b_{ijkh}\right]}\sqrt{\mathrm{Var}\left[\Delta_{ijkh} \mid p_h = b_{ijkh}\right]}$. These objects bound the absolute values of $\lambda \,\mathrm{Cov}\left(\widetilde{\pi}_i, \Delta_{ijkh} \mid p_h = b_{ijkh}\right)$. The sample is restricted to bid-step cells with at least 20 observations.

Table 9: Absolute Marginal-Profit Term in the Bidding FOC

Firm	γ	Mean	Median	p25	p75
VI firm 1	-1.00	949.7771	1035.3556	567.0984	1287.3775
	-0.50	801.5358	765.7693	408.1279	1077.2406
	-0.34	773.7729	710.1492	330.4651	1120.8542
	0.00	771.2757	544.7862	240.6045	1158.6670
	0.50	978.5596	680.8463	337.3647	1599.7927
	0.77	1148.2656	966.4960	451.7778	1855.7508
	1.00	1328.2118	1162.9306	673.7742	2057.3718
Generation firm 1	-1.00	1049.4842	1314.3588	468.1678	1533.7785
	-0.50	973.8631	999.5500	614.3259	1370.1686
	-0.34	956.3041	888.9027	651.4045	1234.5686
	0.00	918.9911	731.7405	612.4558	1069.6428
	0.50	872.8369	709.2934	312.8987	1344.1938
	0.77	905.7502	660.6273	140.7558	1490.0948
	1.00	982.5938	704.0433	202.9241	1626.9382
Generation firm 2	-1.00	2125.1873	2294.1475	1620.6331	2658.0779
	-0.50	1778.1810	1729.1986	1312.4363	1989.5072
	-0.34	1669.7152	1554.6790	1279.1467	1791.2325
	0.00	1440.2451	1198.9050	959.5611	1426.9144
	0.50	1132.3133	588.4333	446.7715	1082.4674
	0.77	1016.1731	290.0734	165.2487	1321.4909
	1.00	1058.3277	289.6719	113.9285	1567.7835
Generation firm 3	-1.00	1927.0744	2070.0230	1605.7316	2327.6530
	-0.50	1392.4910	1479.0603	1131.4383	1706.3321
	-0.34	1226.2490	1284.7289	988.1705	1505.6635
	0.00	883.2657	899.2454	695.6529	1084.9647
	0.50	434.8735	384.1771	291.2076	525.1349
	0.77	251.2420	145.6769	83.6963	246.1758
	1.00	315.9812	182.5672	103.9003	379.7253
Generation firm 4	-1.00	1780.9407	1699.0492	1316.9135	2251.3144
	-0.50	1339.4702	1266.3942	943.1526	1667.7451
	-0.34	1201.5267	1108.3288	828.2380	1469.5453
	0.00	920.3738	813.1379	553.0042	1068.0167
	0.50	576.9920	377.0444	212.3986	498.3509
	0.77	488.8893	170.7788	69.8428	394.5811
	1.00	580.3695	211.7623	110.1940	583.8508

Notes: This table reports summary statistics for $|\mathbb{E}[\Delta_{ijkh} \mid p_h = b_{ijkh}]|$, the absolute value of the risk-neutral marginal-profit term in the bidding first-order condition. The sample is restricted to bid-step cells with at least 20 observations. dropped.

Appendix L Back-of-the-Envelope Calculation of the Creditworthiness Value of Hedging

Fully endogenizing firms' risk aversion is beyond the scope of the paper. This appendix instead provides a back-of-the-envelope calculation that quantifies the magnitude of the creditworthiness channel. The exercise asks whether the variation in retail profits generated by unhedged exposure is large enough to affect the VI firm's borrowing costs.

The finance literature provides several reasons why hedging can create firm value. Smith and Stulz (1985) show that hedging is valuable when expected financial distress costs are convex. Stulz (1990) and Bessembinder (1991) study environments in which low-cash-flow states generate underinvestment because of monitoring frictions, contracting frictions, or debt overhang. Froot et al. (1993) show that when external finance is costly, hedging reallocates internal funds toward states in which financing constraints are most severe.

Empirical evidence is consistent with this channel. Corporate risk management has been shown to increase firm value in settings with large hedgeable risks, including weather-driven risks among electric and gas utilities (Pérez-González and Yun, 2013; Do et al., 2024), commodity-price risk among oil and gas firms (MacKay and Moeller, 2007; Gilje and Taillard, 2017), fuel-price risk among airlines (Carter et al., 2006), and foreign-exchange exposure among nonfinancial firms (Allayannis and Weston, 2001). Most closely related to our setting, Pérez-González and Yun (2013) exploit the introduction of weather derivatives in 1997 and show that risk management increased firm value for U.S. electric and gas utilities. Do et al. (2024) further show that weather-derivative hedging lowers borrowing costs for U.S. electricity and natural gas firms. These findings are consistent with a creditworthiness channel: by reducing cash-flow volatility and downside risk, hedging can improve access to external finance and lower borrowing costs.

We use this channel to interpret the VI firm's hedge position. The calculation proceeds in three steps. First, we use simulated annual retail profits to quantify the magnitude of downside cash-flow states. Second, we map these cash-flow reductions into interest-coverage ratios and synthetic credit spreads. Third, we compare the implied borrowing-cost savings with the expected-profit cost of hedging in the counterfactual model.

Appendix L.1 Why Downside Scenarios Matter for Creditworthiness

Credit ratings are particularly sensitive to downside cash-flow realizations. This is important because the value of hedging does not come primarily from raising average cash flows. It comes from reducing the probability that the firm enters states in which credit metrics deteriorate enough to increase financing costs.

Rating agencies' through-the-cycle methodologies formalize this focus. Under a through-the-cycle rating philosophy, ratings are intended to measure default risk over a full economic cycle rather than at a single point in time. In practice, this means that agencies place substantial weight on how the borrower performs in adverse states. A Federal Reserve study of credit-risk ratings at large U.S. banks describes this approach as estimating the borrower's condition at the worst point in the economic or industry cycle and assigning the rating according to the risk posed in that state.

Moody's corporate methodology is also asymmetric in this sense. Its fundamental credit analysis evaluates credit performance across a range of plausible economic scenarios, and scorecard outcomes may be stronger than the assigned rating at the top of the cycle and weaker at the bottom of the cycle. This asymmetry is natural for credit analysis: unusually strong cash-flow states rarely reduce borrowing costs below the relevant rating bucket, whereas weak cash-flow states can trigger downgrades, covenant pressure, or refinancing difficulties.

The same logic appears in DBRS's treatment of the VI firm. DBRS applies downside stress assumptions to cash-flow profiles and reports sensitivity analyses under declines in net cash flow. In its 2018 rating report, DBRS notes that a negative rating action could occur if key credit metrics deteriorated, while a positive rating action was considered unlikely over the medium term. The report's discussion of challenges emphasizes wholesale price volatility, Alberta market uncertainty, and regulatory risk. These features make downside scenarios the relevant states for evaluating the creditworthiness value of hedging.

Appendix L.2 Simulated Annual Profits

We next quantify the size of downside cash-flow states using the estimates of the Cournot model from Section 5. We fix the risk-aversion parameter for generation-only firms at

the estimated value, $\lambda^*_{GEN} = 1.0 \times 10^{-7}$, and compare two benchmarks for the vertically integrated firm: the risk-neutral benchmark, $\lambda = 0$, and the risk-averse benchmark, $\lambda = \lambda^*_{VI} = 3.16 \times 10^{-6}$. In each case, firms choose swap forward positions in the first stage and make production decisions in the second stage.

Let π^0_h denote hourly profit under $\lambda = 0$, and let π^*_h denote hourly profit under $\lambda = 3.16 \times 10^{-6}$. The Cournot simulation gives the mean and variance of hourly profits,

$$\mathbb{E}[\pi^0_h], \quad \mathbb{E}[\pi^*_h], \qquad \mathrm{Var}(\pi^0_h), \quad \mathrm{Var}(\pi^*_h).$$

Assuming hourly profits are independent and identically distributed, annual profits are given by

$$\Pi^j = \sum_{h=1}^{8760} \pi^j_h, \qquad j \in \{0, *\}.$$

Thus,

$$\mathbb{E}[\Pi^j] = 8760\, \mathbb{E}[\pi^j_h], \qquad \mathrm{Var}(\Pi^j) = 8760\ \mathrm{Var}(\pi^j_h).$$

We use these implied annual moments to simulate annual profit realizations under the two benchmarks.

Figure 19 plots simulated annual profits relative to their means, sorted by magnitude. The figure shows that the risk-neutral benchmark generates substantially larger downside risk than the risk-averse benchmark. In the lower tail of the annual profit difference, profits differ by approximately \$70 million CAD at the 1st percentile and approximately \$50 million CAD at the 10th percentile. These magnitudes motivate the borrowing-cost calculation below: if exposure to downside cash-flow states can reduce annual EBIT by \$50 million CAD, then even modest effects on credit risk or borrowing rates may be economically meaningful.

Appendix L.3 Mapping Cash-Flow Losses into Borrowing Costs

We map cash-flow losses into borrowing costs using the interest-coverage ratio, defined as

$$\text{Interest Coverage Ratio} = \frac{\text{EBIT}}{\text{Interest Expense}}.$$

This ratio measures the extent to which operating earnings cover interest obligations. For example, a coverage ratio of 2.4 means that earnings are 2.4 times interest expense. A ratio

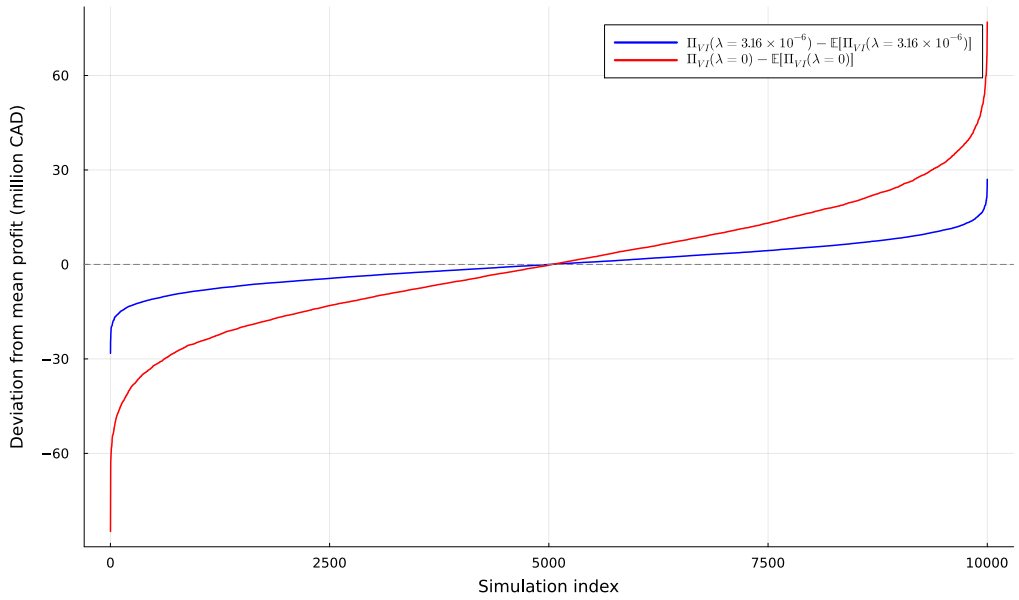


Figure 19: Simulated annual profits relative to the benchmark-specific mean. The figure compares the risk-neutral benchmark, $\lambda = 0$, with the risk-averse benchmark, $\lambda = 3.16 \times 10^{-6}$. The lower tail shows that annual profits can differ by roughly \$50–70 million CAD across the two benchmarks.

of 1.0 means that earnings just cover interest expense.

Table 10 reports the synthetic rating schedule from Damodaran (2026). The table maps interest-coverage ratios into rating categories and default spreads. The final column adds a Government of Canada 10-year yield of 3.47 percent to the default spread, producing an implied borrowing rate for each coverage bucket.

The calculation is mechanical. Let Δ denote the EBIT reduction induced by a downside retail-profit realization. The counterfactual coverage ratio is

$$\text{Coverage}(\Delta) = \frac{\text{EBIT} - \Delta}{\text{Interest Expense}}.$$

We then assign the firm to the corresponding spread bucket in Table 10. The implied increase in annual borrowing costs is

$$\Delta\text{Borrowing Cost} = [s(\text{Coverage}(\Delta)) - s(\text{Coverage}(0))] \times \text{Long-Term Debt},$$

where $s(\cdot)$ denotes the default spread associated with a coverage-ratio bucket.

Table 10: Synthetic Ratings and Borrowing Rates by Interest Coverage Ratio

Coverage Ratio	Rating	Default Spread	Borrowing Rate
< 0.20	D	19.00%	22.47%
$0.20 - 0.65$	C	16.00%	19.47%
$0.65 - 0.80$	CC	12.61%	16.08%
$0.80 - 1.25$	CCC	8.85%	12.32%
$1.25 - 1.50$	B-	5.09%	8.56%
$1.50 - 1.75$	B	3.21%	6.68%
$1.75 - 2.00$	B+	2.75%	6.22%
$2.00 - 2.25$	BB	1.84%	5.31%
$2.25 - 2.50$	BB+	1.38%	4.85%
$2.50 - 3.00$	BBB	1.11%	4.58%
$3.00 - 4.25$	A-	0.89%	4.36%
$4.25 - 5.50$	A	0.78%	4.25%
$5.50 - 6.50$	A+	0.70%	4.17%
$6.50 - 8.50$	AA	0.55%	4.02%
> 8.50	AAA	0.40%	3.87%

Appendix L.4 Borrowing-Cost Effects for the VI Firm

We apply this mapping to the VI firm's 2021 and 2019 financial statements. The exercise takes interest expense, EBIT, and long-term debt from the financial statements, computes the baseline coverage ratio, and then reduces EBIT by \$50 million. These reductions match the approximate size of the downside retail-profit exposure in Figure 19.

2021 Financial Report

The 2021 inputs are

$$\text{Interest Expense} = \$118\text{M}+\$11\text{M} = \$129\text{M}, \qquad \text{EBIT} = \$300\text{M}, \qquad \text{Long-Term Debt} = \$3{,}550\text{M}.$$

The baseline coverage ratio is

$$\frac{300}{129} = 2.33,$$

which falls in the BB+ bucket, with a default spread of 1.38 percent.

A \$50 million reduction in EBIT gives

$$\frac{250}{129} = 1.94,$$

which falls in the B+ bucket, with a default spread of 2.75 percent. The implied increase in annual borrowing costs is therefore

$$(2.75\% - 1.38\%) \times \$3{,}550\text{M} = 1.37\% \times \$3{,}550\text{M} = \$48.6\text{M}.$$

2019 Financial Report

The 2019 inputs are

$$\text{Interest Expense} = \$117\text{M}+\$19\text{M} = \$136\text{M}, \qquad \text{EBIT} = \$208\text{M}, \qquad \text{Long-Term Debt} = \$2{,}796\text{M}.$$

The baseline coverage ratio is

$$\frac{208}{136} = 1.53,$$

which falls in the B bucket, with a default spread of 3.21 percent.

A \$50 million reduction in EBIT gives

$$\frac{158}{136} = 1.16,$$

which falls in the CCC bucket, with a default spread of 8.85 percent. The implied increase in annual borrowing costs is

$$(8.85\% - 3.21\%) \times \$2{,}796\text{M} = 5.64\% \times \$2{,}796\text{M} = \$157.7\text{M}.$$

Thus, a \$50 million downside EBIT realization can increase implied annual borrowing costs by \$48.6 million using the 2021 financial statements and by \$157.7 million using the 2019 financial statements. The effect is discrete because borrowing costs change only when the cash-flow loss moves the firm across a coverage-ratio threshold.

Appendix L.5 Does Hedging Create Firm Value?

We finally compare these borrowing-cost effects with the expected-profit cost of hedging in the counterfactual model. In the Cournot counterfactual, increasing the VI firm's risk-aversion parameter from $\lambda = 0$ to $\lambda = \lambda^*_{VI} = 3.16 \times 10^{-6}$ reduces expected profit by roughly \$24 million per year while substantially reducing profit variance. Figures 17 and 18 illustrate

this tradeoff. Expected profit declines modestly as risk aversion increases, while the variance of profit falls sharply.

The back-of-the-envelope calculation suggests that this tradeoff is plausibly value-creating. A $50 million downside cash-flow realization can move the firm across rating thresholds and increase annual borrowing costs by $49–158 million across the firm's debt stack. By contrast, the expected-profit cost of hedging in the counterfactual model is approximately $24 million per year. The cost of hedging is therefore plausibly smaller than the financing-cost consequence of remaining exposed to downside retail-profit states. In this sense, hedging can create firm value not by increasing expected operating cash flows, but by reducing the probability of costly deterioration in credit metrics.