How Does Risk Hedging Impact Operations? Insights from a Price-Setting Newsvendor Model

Liao Wang, Jin Yao

Faculty of Business and Economics, The University of Hong Kong, Pofuklam Road, Hong Kong SAR, lwang98@hku.hk, u3002692@connect.hku.hk

Xiaowei Zhang

Department of Industrial Engineering and Decision Analytics, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR, xiaoweiz@ust.hk

If a financial asset's price movement impacts a firm's product demand, the firm can respond to the impact by adjusting its operational decisions. For example, in the automotive industry, car makers decrease the selling prices of fuel-inefficient cars when the oil price rises. Meanwhile, the firm can implement a risk-hedging strategy using the financial asset jointly with its operational decisions. Motivated by this, we develop and solve a general risk-management model integrating risk hedging into a price-setting newsvendor. The optimal hedging strategy is calculated analytically, which leads to an explicit objective function for optimizing price and "virtual production quantity" (VPQ). (The latter determines the service level, i.e., the demand fulfillment probability.) We find that hedging generally reduces the optimal price when the firm sets the target mean return as its production-only maximum expected profit. With the same condition on the target mean return, hedging also reduces the optimal VPQ when the asset price trend positively impacts product demand; meanwhile, it may increase the VPQ by a small margin when the impact is negative. We construct the return-risk efficient frontier that characterizes the optimal return-risk trade-off. Our numerical study using data from a prominent automotive manufacturer shows that the markdowns in price and reduction in VPQ are small under our model and that the hedging strategy substantially reduces risk without materially reducing operational profit.

Key words: newsvendor; pricing; risk hedging; mean-variance framework

1. Introduction

Effective risk management for firms facing substantial volatility in product demand must account for various factors that impact demand. An important *exogenous* factor is the price of some tradable financial asset (e.g., stock market index, commodity, and foreign currency). The product demand may fluctuate significantly in response to the financial asset price movement. For example, the "Big

Three" in the automotive industry (General Motors, Fiat Chrysler Automobiles, and Ford Motor Company) all recognize in their annual reports that oil price is a major risk factor affecting their sales volumes (General Motors, 2020 10-K Filing; Fiat Chrysler Automobiles, 2020 10-K Filing; Ford Motor Company, 2020 10-K Filing). Both economic theory and empirical evidence support this observation. Theoretically, it is known that the demand for a product is affected by the price of a complementary good Mankiw (2014). The reason is that, in the case of automakers, a higher oil price leads to a higher gasoline price, directly increasing the expense of using a car and encouraging customers to switch from gas guzzlers to more fuel-efficient vehicles. Empirical studies that support this theory include those of Busse et al. (2009), Klier and Linn (2010), Busse et al. (2013), and Langer and Miller (2013). One can also find examples of financial asset prices impacting product demand in other industries. John Deere, the largest farming equipment manufacturer in the United States, discloses that prices of agricultural commodities (e.g., corn) significantly impact their sales, as these prices directly affect the revenues of their customers (e.g., farmers). Another example is Caterpillar Inc., the largest industrial equipment manufacturer, recognizing metal prices as demand risk factors. Our analysis in this paper is not limited to any specific industry.

A major *endogenous* factor impacting demand is the product price. Basic economic theory implies that a higher price leads to lower demand, and the slope of the demand curve describes how demand changes with the product price. Unlike the financial asset price, the product price is controllable. The firm may adjust the product price proactively in response to changes in the financial asset price to mitigate its impact on the product demand. For example, automakers often respond to the fuel price's impact on vehicle demand by setting prices strategically (Busse et al. 2009, Langer and Miller 2013, Allcott and Wozny 2014). In the scenario of rising oil prices, demand for fuel-inefficient cars decreases while demand for fuel-efficient cars increases. Thus, automakers can reduce the price of fuel-inefficient cars to offset the rising energy costs' negative impact on product demand. At the same time, they can raise the price of fuel-efficient cars in response to the increased demand.

We now demonstrate our results using data from a prominent automotive manufacturer, specifically on two of their popular car models: a sport utility vehicle (SUV) and a compact car. For the sake of brevity, we will refer to manufacturer and the car models as `AutoMFR`, `Sport`, and `Compact`, respectively. `Sport` is categorized as fuel-inefficient, having low miles per gallon (MPG), while `Compact` is fuel-efficient. The relevant financial asset is crude oil, West Texas Intermediate (WTI). For each car model, we regress the monthly sales volume and the selling price, respectively, on the monthly average WTI price. The regression results are summarized in Table 1. We find that the selling price of `Sport` has a significant negative correlation with the oil price, while `Compact` has a significant positive correlation. The same patterns hold for the sales volume. These are consistent

Car	Estimate	Regression Equation	
		Price $\sim$ WTI	Sales Volume $\sim$ WTI
Sport	slope	-40.16 (3.60)	-168.55 (23.53)
	p-value	6.6×10^{-20}	1.7×10^{-11}
Compact	slope	9.32 (1.90)	132.99 (25.20)
	p-value	4.2×10^{-6}	2.2×10^{-7}

Note. The numbers in parentheses represent the standard errors.

with our discussions above and existing empirical findings (Klier and Linn 2010, Busse et al. 2013, Allcott and Wozny 2014).

The discussion and data analysis above convey two crucial ideas. First, financial asset price movement may significantly impact product demand. Second, decisions regarding the product price may interact with the financial asset price. Namely, the firm can adjust the product price for risk management in response to the financial asset price's impact on the product demand (as in the practice of automakers discussed above). The firm may also manage risk by adopting a *risk-hedging* strategy of trading this financial asset. Since operational decisions (pricing and production) and financial decision (risk hedging) can both affect the firm's risk, risk-management strategies can be improved substantially by *jointly* optimizing these two types of decisions. Doing so accounts for endogenous and exogenous factors that impact demand simultaneously. Despite the rich literature on pricing strategies and risk management using financial assets, these two types of decisions are usually studied separately. The literature on how to jointly optimize pricing and risk hedging is scarce. An example analyzed by Guiotto and Roncoroni (2022) is a rare exception. To the best of our knowledge, our paper is the first study on how risk hedging impacts pricing in a price-setting newsvendor model.

Motivated by the above discussion, we pose the following research questions:

(i) How can a risk-management strategy be developed by jointly optimizing pricing, production, and hedging decisions using relevant financial assets?

(ii) How does risk hedging affect the pricing decision compared with the no-hedging case?

To answer these questions, we start by building a demand model that incorporates impacts from both the product pricing decision and the financial asset price. The model allows a general relationship between product demand and asset price. Based on this demand model, we formulate the risk-management problem following Markowitz's mean-variance optimization framework (Markowitz 1952). Fixing production and pricing decisions, we first analytically solve for the optimal risk-hedging strategy, which yields an explicit risk objective function for the operational decisions. We then minimize this objective function to find the optimal production and pricing decisions, which leads

4

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

to a complete characterization of the efficient frontier. We demonstrate the effectiveness of the risk-management model using real data about `Sport` and `Compact`.

1.1. Main Results and Contributions

Our main results are Theorems 1, 2 and 3. Given a pair of price and VPQ (defined in Section 2, which directly determines the service level), Theorem 1 characterizes the optimal hedging strategy and the associated variance of the total terminal return. Risk of a return is quantified by its standard deviation, so the associated variance is the squared risk. The optimal hedging strategy combines a risk-mitigation position and an investment position. The variance associated with this hedging strategy is the sum of squared investment and unhedgeable risks (i.e., the kind of risk that is irrelevant to the financial asset and thus cannot be hedged by it). This variance, as a function of the price and VPQ, provides an explicit objective to be further minimized.

Assuming the asset price trend positively impacts demand (e.g., rising oil price boosts the demand for fuel-efficient cars), Theorem 2 characterizes the optimal price and VPQ in the presence of hedging, with the target mean return set as the newsvendor's maximum expected profit. We find that both the optimal price and VPQ are *lower* in the presence of hedging than those without hedging. The reason is that hedging *offsets* the positive effect of the asset price trend on product demand, leading to a hypothetically *smaller* market than that without hedging. Adapting to the smaller market size, the optimization of the operational decision decreases the price and VPQ while reducing unhedgeable risk. Moreover, this main result holds for general operational payoff functions with certain structural properties (refer to Theorem EC.1 in the e-companion).

Theorem 3 assumes the asset price trend negatively impacts demand (e.g., rising oil price decreases the demand for fuel-inefficient cars), the opposite scenario of the one considered in Theorem 2. Interestingly, the conclusions are similar: the optimal price is lower with hedging than without hedging, and the optimal VPQ with hedging is lower than (or at most exceeds by a small margin) the optimal VPQ without hedging. We interpret this counter-intuitive result as follows. The risk-mitigation position of the hedging strategy offsets the negative effect of the asset price trend on product demand, leading to a hypothetically *larger* market size. Hence, with the target mean return set as the maximum expected profit of the base model (i.e., the profit-maximizing price-setting newsvendor without hedging), the corresponding operational decision of the base model ("NV's decision") combined with hedging induces a larger expected payoff than the target level. Thus, to attain this target mean return, with the NV's decision unchanged, the investment payoff from the hedging strategy induces a *negative* expected payoff while bearing *positive* contribution to the total risk. In this case, the joint operational and risk hedging decision can not be optimal regarding the return-risk trade-off because part of the terminal return has a negative return and positive

risk contribution, thereby highly inefficient. Therefore, to avoid a negative expected investment payoff from hedging with the enlarged market size, the optimal decision has to *leave leeway* in the target mean return for the investment payoff to fill. As a result, the pricing level needs to be adjusted downward to suppress the expected hedged production payoff. In addition, regarding the optimal VPQ, we identify a critical condition: the negative effect of the asset price is sufficiently strong. Under this condition, the optimal VPQ with hedging does not exceed that without hedging (Theorem 3(ii)). When this condition does not hold, it may exceed that without hedging but the excess amount is small (Theorem 3(iii)).

Both technical and managerial contributions are made. Technically, we develop and solve a general risk-management model that integrates pricing, production, and risk-hedging decisions. This model adopts a general diffusion process for the asset price dynamics and does not assume any specific functional relationship between demand and asset price; thus, it can incorporate various application-specific data analytics. In particular, we explicitly solve the model with the asset price following an exponential Ornstein–Uhlenbeck (EOU) process, a standard oil price model. We remark that the general asset price process (including EOU) has a *stochastic* market price of risk (MPR) process, rendering the hedging problem much more challenging than those with a constant MPR (Wang and Yao 2017) or a deterministic MPR (Caldentey and Haugh 2006).

Concerning managerial insights, the leading message of our paper is that hedging reduces the pricing level. It points to the insight that hedging not only reduces risk but also enhances a firm's competitiveness in the market. In addition, we show that the VPQ in the presence of hedging is either lower than or exceeds by a small margin the VPQ without hedging. We conduct a comprehensive numerical study using data from `AutoMFR` to demonstrate and complement our theoretical results. The analysis shows that hedging performs well: it can reduce risk by as much as 40%. In particular, markdowns for both the price and VPQ are small—not exceeding 1.03% for price reduction and 1.70% for VPQ reduction. This result is desirable, as firms usually do not want to reduce the price or service level excessively due to concerns about branding or maintaining market share. It also indicates that, while substantially reducing risk, hedging will not materially reduce a firm's operational levels.

1.2. Literature Review

This paper falls within the scope of integrated operations and financial risk management. Two main streams of this literature are related to our work. One stream is joint pricing and production/inventory management. Whitin (1955) is the first to study the fundamental connection between pricing and inventory control theory for the single-period setting. Petruzzi and Dada (1999) consider a risk-neutral firm that faces a price-dependent random demand, and they examine how the stocking

quantity decision interacts with the selling price decision. Agrawal and Seshadri (2000) study how a risk-averse retailer facing a price-dependent random demand makes order quantity and pricing decisions to maximize an expected utility. Chen et al. (2009) consider a risk-averse decision-maker similar to that analyzed by Agrawal and Seshadri (2000) but with a different risk objective. Extensions include multi-period settings (Federgruen and Heching 1999, Chen and Simchi-Levi 2004a,b) and multi-product settings (Aydin and Porteus 2008, Zhu and Thonemann 2009).

The other stream of literature concerns incorporating risk hedging by trading financial assets in operational risk management. The paper most relevant to ours is that by Wang and Yao (2017). They study a newsvendor model in which a financial asset's price partially drives product demand. Our paper differs by including product price as a decision variable and adopting a more general model for product demand and asset price. In particular, when solving a quadratic hedging problem, they assume the asset price follows a geometric Brownian motion, which has a constant MPR process. By contrast, we work with a general diffusion process, which has a stochastic MPR process. This general assumption renders the hedging problem much more technically challenging. Meanwhile, it encompasses a much broader scope of applications than that paper.

Gaur and Seshadri (2005) study how to construct an optimal trading strategy to hedge inventory risk using financial market instruments. A study closely related to ours is that of Caldentey and Haugh (2006). They formulate a general modeling framework incorporating risk hedging into operations with a quadratic utility function. However, due to the generality of the framework, they do not examine the interaction of risk hedging and any specific operational decision. Caldentey and Haugh (2009) consider a supply chain contracting problem with a supplier and a retailer engaging in a Stackelberg game. In their setting, the product market size depends on the price of some financial asset. Ding et al. (2007) consider an international firm that sells to both domestic and foreign markets and uses operational and financial hedging to manage currency exchange risk. Kouvelis and Li (2019) study a newsvendor problem with correlated operational and financial risks under value-at-risk constraints.

None of the papers reviewed above include product price as a decision variable. Two exceptions are Chen et al. (2007) and Guiotto and Roncoroni (2022). Chen et al. (2007) study joint dynamic inventory control and pricing strategies for a risk-averse decision-maker in a multi-period setting. In the extension part of their study, they assume that the model parameters are correlated with some financial asset price and then include risk hedging. Our paper differs by considering a mean-variance optimization criterion instead of an expected utility objective. In contrast to deriving the structure of the dynamic programming problem involved in pricing and inventory control, our focus is to examine how financial risk hedging affects pricing and service levels. Guiotto and Roncoroni (2022) develop a general framework of combined custom hedging for a risk-averse firm exposed to claimable

and non-claimable risks. The authors consider the pricing decision in a specific example of electricity retailers. Our paper differs from this study in two aspects. First, our setup is a price-setting newsvendor, which also involves the quantity decision, while the example in Guiotto and Roncoroni (2022) is essentially a make-to-order setup. Second, Guiotto and Roncoroni (2022) analyze the interaction of pricing and hedging via additional operational flexibility gained by optimal hedging. By contrast, we study the impact of hedging on pricing by directly analyzing the difference in the pricing levels with and without hedging.

The rest of the paper is organized as follows. In Section 2, we discuss the price-setting newsvendor model. In Section 3, we develop the demand-asset model and formulate the risk-management problem. We solve the hedging problem in Section 4 and discuss the optimal price and VPQ in the presence of hedging in Section 5. In Section 6, we apply the methodology to real data from a prominent automotive manufacturer. Concluding remarks are collected in Section 7. The e-companion includes lists of notation, supplementary analytical results (part I), and technical proofs (part II).

2. Price-Setting Newsvendor

Consider the price-setting newsvendor (NV) model, which has been studied extensively in the literature (Petruzzi and Dada 1999, Agrawal and Seshadri 2000); see also DeYong (2020) for a recent survey. Consider a newsvendor over a selling period $[0, T]$. At time 0, the newsvendor needs to decide a unit selling price P and a production quantity Q before a stochastic demand $D_T = D_T(P)$ that depends on P is realized at time T. This setup is suitable for products that do not have frequent price adjustments (e.g., the automobiles considered in this paper) but not for those that allow dynamic pricing (e.g., retail goods on e-commerce platforms). With unit production cost c and unit salvage value s, the newsvendor's payoff is $P(Q \wedge D_T) + s(Q - D_T)^+ - cQ$, where $a \wedge b = \min\{a, b\}$, $(x)^+ = \max\{x, 0\}$, and the three terms represent the revenue from sales, the salvage value, and the production cost, respectively. Because $a \wedge b = a - (a - b)^+$, the production payoff function can be expressed as: $H_T(P, Q) = (P - c)Q - (P - s)(Q - D_T)^+$.

We model the demand function as $D_T(P) = A_T - bP$, where A_T is the (random) market size independent of P, and b is a positive parameter capturing demand's sensitivity to price. We assume that A_T has a continuous distribution and let $f(a)$, $F(a)$, and $r(a) = f(a)/(1 - F(a))$ denote its density function, distribution function, and hazard rate function, respectively.

2.1. Base Model: Profit Maximization

Our *base model* assumes that the newsvendor aims to maximize the expected production payoff:

$$(P^{\mathrm{NV}}, Q^{\mathrm{NV}}) := \arg\max_{P, Q} \mathsf{E}[H_T(P, Q)] \quad \text{s.t.} \quad P \geq c, \quad Q \geq 0. \tag{1}$$

We define $R := Q + bP$ and interpret it as follows. Note that $R - A_T = Q - D_T$ represents overproduction, so $\mathsf{P}(A_T \leq R) = \mathsf{P}(D_T \leq Q)$ is the *service level* (i.e., the demand fulfillment probability). In other words, R serves A_T just like Q serves D_T. Moreover, R and Q have the same unit measuring the number of products. Therefore, R is of similar nature as Q, and we call it the *virtual production quantity (VPQ)*. In the following, we will use (P, R) instead of (P, Q) to facilitate theoretical analysis. This change of variable is standard in the literature (Petruzzi and Dada 1999, Agrawal and Seshadri 2000). In particular, problem (1) is equivalent to

$$(P^{\mathrm{NV}}, R^{\mathrm{NV}}) := \arg\max_{P,R} \mathsf{E}[H_T(P, R)] \quad \text{s.t.} \quad P \geq c, \quad R - bP \geq 0, \tag{2}$$

where

$$\mathsf{E}[H_T(P, R)] = (P - c)(R - bP) - (P - s)\mathsf{E}[(R - A_T)^+]. \tag{3}$$

It can be shown that $\mathsf{E}[H_T(P, R)]$ is concave in R for a fixed P while concave in P for a fixed R. Moreover, $\mathsf{E}[H_T(P, R)]$ is supermodular in (P, R).

We make the following assumptions throughout the paper.[1]

ASSUMPTION 1. $2r(a)^2 + r'(a) > 0$ *for all* a*. In addition,* $[1 - F(a)]^2/r(a) \to 0$ *as* $a \to \infty$ *and* $F^2(a)/r(a) \to 0$ *as* $a \to -\infty$*. Further, the three conditions also hold for distribution of* A_T *under the probability measure* P^M *defined in (12) below.*

ASSUMPTION 2. $P^{\mathrm{NV}} > c$*,* $Q^{\mathrm{NV}} = R^{\mathrm{NV}} - bP^{\mathrm{NV}} > 0$ *and* $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] > 0$*.*

ASSUMPTION 3. $\mathsf{E}[(bc - A_T)^+] \leq [2b(c - s)] \wedge \sqrt{4b\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]}$*.*

Assumption 1 is standard in the literature (Petruzzi and Dada 1999). It ensures the existence of a unique solution to problem (2). Assumption 2 excludes the trivial case that the optimal production payoff is nonpositive. Intuitively, this case only occurs when the market size A_T is very low, which is unlikely for commonly demanded products such as automobiles (the case considered in this paper). To understand Assumption 3, note that for any $P \geq c$, $A_T - bc \geq A_T - bP = D_T$ is an upper bound of the product demand, and we expect it to have a small negative part (i.e., $\mathsf{E}[(bc - A_T)^+]$ is small), which is expressed by Assumption 3. In Section 6, we numerically validate that Assumptions 1–3 are satisfied by the demand model and parameters calibrated from an automaker's data sets.

We characterize the optimal solution to problem (2) in the following proposition. Analogous results are also known in the literature (Petruzzi and Dada 1999).

---

[1] Although not imposed on the primitive parameters of the model, Assumptions 2 and 3 are easy to be verified or disproved numerically. See Section EC.3 of the e-companion (particularly Proposition EC.1) for sufficient conditions on the model primitive parameters for these two assumptions to hold.

PROPOSITION 1. *Under Assumptions 1 and 2, the profit-maximization problem in (2) has a unique solution characterized by the following optimality equations:*

$$2bP^{\mathrm{NV}} - bc = \mathsf{E}(R^{\mathrm{NV}} \wedge A_T), \qquad R^{\mathrm{NV}} = F^{-1}\Big(\frac{P^{\mathrm{NV}} - c}{P^{\mathrm{NV}} - s}\Big). \tag{4}$$

The solution to the base model specified in Proposition 1 represents a risk-neutral decision: the expected return (throughout the paper, "return" stands for profit or loss in dollar amounts rather than the rate of return) is maximized without accounting for the associated risk. In this study, our ultimate goal is to develop an effective risk-management strategy.

2.2. Risk Minimization

We follow the common practice of measuring risk by *standard deviation*. Then, minimizing the risk is equivalent to minimizing the variance (i.e., squared risk) $\mathsf{Var}[H_T(P, R)]$. Note that $\mathsf{Var}[H_T(P, R)] = (P - s)^2 \mathsf{Var}[(R - A_T)^+]$ is increasing in both P and R. A higher P leads to a higher payoff risk because it strengthens the positive (or negative) impact on the payoff from each sold (or unsold) product, increasing the payoff's volatility. Meanwhile, a higher R (i.e., VPQ) represents a higher service level and thus increases the exposure to demand volatility, amplifying the payoff risk. Managing risk without considering the associated return makes little economic sense, and the focal issue here is to consider the trade-off between the two.

Of particular interest are two optimization problems that examine the return-risk trade-off induced by the pricing (P) or VPQ (R) decisions. For the former, for $P \geq c$, we define

$$R^{\mathrm{NV}}(P) := \arg\max_R \mathsf{E}[H_T(P, R)], \ \ m(P) := \mathsf{E}[H_T(P, R^{\mathrm{NV}}(P))], \ \ v(P) := \mathsf{Var}[H_T(P, R^{\mathrm{NV}}(P))]. \tag{5}$$

In parallel, for $R \geq bc$, we define

$$P^{\mathrm{NV}}(R) := \arg\max_P \mathsf{E}[H_T(P, R)], \ \ m(R) := \mathsf{E}[H_T(P^{\mathrm{NV}}(R), R)], \ \ v(R) := \mathsf{Var}[H_T(R^{\mathrm{NV}}(R), R)]. \tag{6}$$

Further, we formulate a risk-management problem under the *conditional minimum variance (mV)* criterion[2] to explicitly express the return-risk trade-off:

$$(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}) := \arg\min_{P,R} \mathsf{Var}[H_T(P, R)] \quad \text{s.t.} \quad \mathsf{E}[H_T(P, R)] = m, \tag{7}$$

where $m \in (0, \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]]$ is the target mean return. We call (7) the *no-hedging model* to differentiate with the risk-minimization model with risk hedging in Section 3. With $P = c$ or $R = bP$, the production payoff is at most $-(c - s)\mathsf{E}[(bc - A_T)^+] \leq 0$. Thus, with $m > 0$, neither $P \geq c$ nor $R \geq bP$ can be binding and thus can be removed from the formulation in (7). The same applies to all relevant settings throughout the paper.

---

[2] There exist alternative, mathematically equivalent formulations; see Section EC.1 of the e-companion for a discussion.

10

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

Before proceeding, we review the concept of the *efficient frontier* (Markowitz 1952). For a random return $W(x)$ induced by a decision x, the corresponding mV formulation is: $v(m) := \min_x \mathsf{Var}(W(x))$ subject to $\mathsf{E}[W(x)] = m$, for which m is an exogenously set target mean level of return. (In our NV's setup, $x = (P, R)$ and $W(x) = H_T(P, R)$.) If $v(m)$ is an increasing function in m, then $(m, \sqrt{v(m)})$ is said to be an efficient frontier. It means that with all decisions being optimized, every increment of return is accompanied by an increment of risk. Economically, the efficient frontier is a complete characterization of the best possible (i.e., efficient) trade-off between return and risk.

Now, we characterize these three optimization problems in the following proposition.

PROPOSITION 2. *Suppose Assumptions 1–3 hold.*

(i) *For problem (5), $R^{\mathrm{NV}}(P)$ increases in P. Both $m(P)$ and $v(P)$ increase in $P \in [\underline{P}, P^{\mathrm{NV}}]$, where $\underline{P}$ is the smallest root of $m(P) = 0$, so $(m(P), \sqrt{v(P)})$ constitutes an efficient frontier.*

(ii) *For problem (6), $P^{\mathrm{NV}}(R)$ increases in R. Both $m(R)$ and $v(R)$ increase in $R \in [\underline{R}, R^{\mathrm{NV}}]$, where $\underline{R}$ is the smallest root of $m(R) = 0$, so $(m(R), \sqrt{v(R)})$ constitutes an efficient frontier.*

(iii) *For problem (7), $\mathsf{Var}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})]$ increases in m, so $(m, \sqrt{\mathsf{Var}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})]})$ constitutes an efficient frontier for $m \in (0, \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]]$. In particular, $\frac{d}{dm}\sqrt{\mathsf{Var}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})]} = \infty$ at $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]]$.*

Proposition 2 offers insights into how operational decisions impact return-risk trade-offs. Part (i) indicates that a higher pricing level increases returns and risk simultaneously, and analogously, Part (ii) considers the impact of the VPQ.[3] Part (iii) mirrors the last two parts: higher returns induce higher risk (after P and R are optimized). It also indicates that the profit-maximizing solution to the base model is just one point on the efficient frontier, inducing both the maximum risk and the maximum *incremental* risk. As the return approaches the newsvendor's maximum profit in the base model, the additional risk affected by a slight increase in return is enormous. In other words, the solution to the base model induces an extremely risky payoff. This paper proposes a risk-hedging model that improves the efficient frontier via substantial risk reductions from the no-hedging model.

Parts (i) and (ii) of Proposition 2 are symmetrical: a higher P induces a higher R and vice versa. This property confirms the *supermodularity* of $\mathsf{E}[H_T(P, R)]$ in (P, R), which holds because the pricing level's marginal value, $\frac{\partial \mathsf{E}[H_T(P,R)]}{\partial P} = -2bP + bc + R - \mathsf{E}[(R - A_T)^+]$, is increasing in R. Economically, this means that a higher VPQ (hence also a higher service level) increases the pricing level's marginal value, and we interpret it as follows. Increasing the VPQ has opposing effects on the pricing level's marginal value. On the one hand, it increases revenue by capitalizing on the higher price from the increased VPQ (i.e., the term R in $\frac{\partial \mathsf{E}[H_T(P,R)]}{\partial P}$ above). On the other hand, the higher the service level (induced by the higher VPQ value) , the greater the expected loss due

---

[3] We expect $\underline{P}$ and $\underline{R}$ to be close to c and bc, respectively. This is confirmed in the data analysis in Section 6.2.

to overproduction (i.e., the term $\mathsf{E}[(R - A_T)^+]$ in $\frac{\partial \mathsf{E}[H_T(P,R)]}{\partial P}$ above. The positive effect outweighs the negative one, as the probability of selling a unit of product is always nonzero. In a nutshell, a higher VPQ induces a higher pricing level due to its positive impact on the pricing level's marginal value. This insight is crucial in proving our main results, Theorems 2 and 3, in Section 5.

3. Price-Setting Newsvendor with Risk Hedging

We first introduce a general continuous-time model for financial asset price and product demand (Section 3.1), and then formulate a risk-minimization problem in the presence of risk hedging (Section 3.2).

3.1. Financial Asset Price and Product Demand

We fix a filtered probability space $\{\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathsf{P}\}$ and define two independent standard Brownian motions B_t and $\tilde{B}_t$. Let $\mathcal{G}_t$ and $\mathcal{F}_t$ denote the filtrations generated by B_t and $(B_t, \tilde{B}_t)$, respectively.[4] Clearly, $\mathcal{G}_t \subseteq \mathcal{F}_t$. Suppose the asset price X_t is a continuous diffusion process driven by B_t:

$$dX_t = X_t(\mu_t dt + \sigma_t dB_t), \tag{8}$$

where μ_t and $\sigma_t > 0$ are continuous processes adapted to $\mathcal{G}_t$. Suppose also that (X_t, μ_t, σ_t) is Markovian under P. We make the following integrability assumption.

ASSUMPTION 4. $\int_0^T \mathsf{E}[(X_t\sigma_t)^2]dt < \infty$.

Let D_t be the cumulative demand up to time t. Suppose

$$dD_t = dA_t - \frac{b}{T}Pdt \quad \text{with} \quad dA_t = dC_t + \tilde{\sigma}d\tilde{B}_t, \tag{9}$$

where $\tilde{\sigma} > 0$ is a constant and C_t is a continuous process adapted to $\mathcal{G}_t$. We further assume that for all $u \leq v$, $C_v - C_u$ is adapted to the sigma-algebra generated by $\{X_t, u \leq t \leq v\}$; that is, the asset price fluctuation over any time interval impacts the demand accumulated over the same period, which is consistent with the discussion in Section 1. Note that for all $t \leq u$, $A_u = A_t + (C_u - C_t) + \tilde{\sigma}(\tilde{B}_u - \tilde{B}_t)$. Given $\mathcal{F}_t$, the distribution of A_u under P is determined by $(X_t, \mu_t, \sigma_t, A_t)$ because of the assumed Markovian property of (X_t, μ_t, σ_t) and the independence of $\tilde{B}_t$.

Three factors determine the instantaneous incremental demand dD_t in equation (9). That is, dC_t (representing the impact of the asset price fluctuation), $\tilde{\sigma}d\tilde{B}_t$ (capturing the demand's inherent volatility), and $-(b/T)Pdt$ (incorporating the sensitivity of the demand rate with respect to the selling price). Without loss of generality, we assume that $C_0 = 0$. Thus, the base demand $D_0 = 0$. In light of the demand function $D_T(P) = A_T - bP$, equation (9) models the market size A_T as

$$A_T = C_T + \tilde{\sigma}\tilde{B}_T. \tag{10}$$

---

[4] Throughout the paper, unless otherwise stated, adapted processes, martingales, and local martingales are defined with respect to $\mathcal{F}_t$.

12

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

The demand model above is *general*. We do not assume any specific functional form of C_t in $\{X_s : 0 \leq s \leq t\}$, allowing the model to be adapted to various specific application scenarios. A simple example is $C_t = g(X_t)$ for some function g; see, e.g., Gaur and Seshadri (2005). One can extend this simple model to permit dependence on the path of asset prices. For example, $C_T = \int_0^T \tilde{\mu}(X_t) dt$, where $\tilde{\mu}$ is a function of the asset price and models the demand rate (see Section 6).

Assuming the interest rate is zero, the associated MPR process is

$$\eta_t := \frac{\mu_t}{\sigma_t}. \tag{11}$$

It captures the risk-adjusted *trend* of the asset price which may positively or negatively impact product demand. For example, with rising oil prices, the demand for fuel-efficient cars (usually sedans) increases while the demand for fuel-inefficient cars decreases.

We introduce several technical definitions to formalize the impact of the asset price trend on product demand. Define the risk-neutral measure P^M for the asset price process via the Radon–Nikodym derivative Z_T. That is,

$$\frac{d\mathsf{P}^M}{d\mathsf{P}} := Z_T, \quad Z_t := e^{-\int_0^t \eta_s dB_s - \frac{1}{2} \int_0^t \eta_s^2 ds}, \ t \in [0, T]. \tag{12}$$

Suppose (X_t, σ_t) is Markovian under P^M. Then, given $\mathcal{F}_t$, the distribution of A_u under P^M is determined by (X_t, σ_t, A_t) for all $t \leq u$. As will become evident later in the paper, Z_t is crucial in determining the optimal hedging strategy. We impose a mild technical assumption on it.

ASSUMPTION 5. *Z_t is a square-integrable martingale over $[0, T]$ under P.*

Under the probability measure P^M, the asset price process X_t, which is driven by B_t, is *de-trended*, becoming a local martingale. By contrast, the change of measure does not affect the non-financial noise $\tilde{B}_t$. Moreover, let C_T^M be a random variable whose distribution under P is identical to that of C_T under P^M; that is, $\mathsf{P}(C_T^M \leq x) = \mathsf{P}^M(C_T \leq x)$ for all x.[5] We can interpret C_T^M as the version of the financial component of the demand that incorporates no asset price trend.

We also need the notion of *stochastic order* to compare random variables (Ross 1996). Recall that for two random variables X and Y, X is said to be *stochastically larger* than Y, denoted by $X \succeq Y$ or $Y \preceq X$, if $\mathsf{P}(X \leq a) \leq \mathsf{P}(Y \leq a)$ for any real a. An immediate result following this definition is that, $X \succeq Y$ if and only if $\mathsf{E}[f(X)] \geq \mathsf{E}[f(Y)]$ for any increasing function f.

To illustrate this order relation between two stochastic processes, consider two geometric Brownian motions (GBMs): $X_t^{(1)} = X_0^{(1)} \exp\left((\mu_1 - 0.5\sigma^2)t + \sigma B_t^{(1)}\right)$ and $X_t^{(2)} = X_0^{(2)} \exp\left(-0.5\sigma^2 t + \sigma B_t^{(2)}\right)$,

---

[5] For example, suppose $C_T = \int_0^T \tilde{\mu}(X_t) dt$ for some function $\tilde{\mu}$. Then, $C_T^M = \int_0^T \tilde{\mu}(X_t^M) dt$, where $\{X_t^M : t \in [0, T]\}$ is a stochastic process whose distribution under P is identical to that of $\{X_t : t \in [0, T]\}$ under P^M. In Section 6, we assume X_t is an EOU process under P, following the dynamics (29). Then, it is easy to show that the distribution of X_t under P^M equals $X_0 e^{-(1/2)\sigma^2 t + \sigma B_t^M}$, where B_t^M is a Brownian motion under P^M. Hence, the distribution of X_t^M under P equals $X_0 e^{-(1/2)\sigma^2 t + \sigma B_t}$, where B_t is a Brownian motion under P.

where $B_t^{(1)}$ and $B_t^{(2)}$ are two Brownian motions. If $X_0^{(1)} = X_0^{(2)}$ and $\mu_1 \geq 0$, then the two GBMs have the same initial value and volatility, but $X_t^{(1)}$ has a positive trend while $X_t^{(2)}$ does not. It is easy to verify that $X_t^{(1)} \succeq X_t^{(2)}$ for any $t \geq 0$. Note that in the special case that these two Brownian motions are identical in the path-wise sense, we have that $X_t^{(1)} \geq X_t^{(2)}$ path-wise, which implies the stochastic ordering between them. Furthermore, in this special case, if we assume in the market size model (10) that $A_T^{(1)} = \int_0^T \tilde{\mu}(X_t^{(1)}) dt + \tilde{\sigma} \tilde{B}_T$ and $A_T^{(2)} = \int_0^T \tilde{\mu}(X_t^{(2)}) dt + \tilde{\sigma} \tilde{B}_T$, then $A_T^{(1)} \succeq A_T^{(2)}$ for an increasing function $\tilde{\mu}$ while $A_T^{(1)} \preceq A_T^{(2)}$ for a decreasing function $\tilde{\mu}$. In Section 6, we assume that the asset price follows an EOU process, and the demand rate function $\tilde{\mu}$ is linear. In this example, it is analytically intractable to establish a stochastic order between C_T and C_T^M. Instead, we resort to a hypothesis-testing approach, applying the Mann–Whitney U-test (Mann and Whitney 1947) to simulated samples of these two quantities with parameters estimated from real world data; see Section EC.4.4 of the e-companion.

Now, we formally define the positive/negative effect of the asset price trend on product demand.

- The asset price trend is said to *positively impact* demand, if $C_T \succeq C_T^M$. In this case, $\mathsf{E}[H_T(P, R)] \geq \mathsf{E}^M[H_T(P, R)]$ for all (P, R); refer to Proposition 4 in Section 5.1.
- The asset price trend is said to *negatively impact* demand, if $C_T \preceq C_T^M$. In this case, $\mathsf{E}[H_T(P, R)] \leq \mathsf{E}^M[H_T(P, R)]$ for all (P, R); refer to Proposition 6 in Section 5.2.

In other words, the positive effect occurs if the demand incorporating the asset price trend is stochastically larger than the demand impacted by the de-trended asset price. We can interpret the negative effect likewise.

REMARK 1. The positive/negative impact of asset price trend on product demand defined above directly translates to the same stochastic order relation between the market sizes under the real-world and risk-neutral probability measures. That is, $C_T \succeq C_T^M$ implies $A_T \succeq A_T^M$, and likewise $C_T \preceq C_T^M$ implies $A_T \preceq A_T^M$. This can be easily shown using the definitions of A_T and A_T^M and the fact that stochastic ordering is preserved under convolutions of probability distributions; see Shaked and Shanthikumar (2007, page 6).

In Section 5, we study the impact of risk hedging on the operational decisions, which entails characterizing how the stochastic order of C_T between C_T^M influences the operational decisions. We show that a (stochastically) larger market induces a higher profit-maximizing price and a higher VPQ; see Propositions 4 and 6 in Section 5.

3.2. Mean-Variance Risk-Hedging Model

Having built the impact of asset price into the demand model, one can design a real-time hedging strategy using the financial asset to mitigate production payoff risk. Specifically, in problem (7), we

include a risk-hedging strategy as a third decision variable in addition to P and R:

$$\min_{P,\,R,\,\vartheta:=\{\theta_t,t\in[0,T]\}\in\mathcal{A}_X} \mathsf{Var}[H_T(P,R)+\chi_T(\vartheta)]$$

$$s.t.\quad \chi_t(\vartheta):=\int_0^t \theta_s dX_s,\quad \mathsf{E}[H_T(P,R)+\chi_T(\vartheta)]=m,$$

(13)

where θ_t is the number of shares held in the asset at time t, χ_t is the cumulative profit/loss from the hedging strategy until time t, and $\mathcal{A}_X$ is the set of all *admissible* strategies that satisfy certain conditions.[6] We call (13) the *hedging model* in contrast to the no-hedging model (7). We remark that the hedging model applies to large and financially unconstrained manufacturers; hence there is no budget limit on the execution of the hedging strategy.

To solve problem (13), we first optimize the hedging strategy with fixed P and R (see Section 4):

$$B(m,P,R):=\min_{\vartheta=\{\theta_t,t\in[0,T]\}\in\mathcal{A}_X} \mathsf{Var}[H_T(P,R)+\chi_T(\vartheta)]$$

$$s.t.\quad \chi_t(\vartheta):=\int_0^t \theta_s dX_s,\quad \mathsf{E}[H_T(P,R)+\chi_T(\vartheta)]=m.$$

(14)

Then, we minimize the optimal value of this hedging problem over P and R to further reduce risk:

$$(P_m^h, R_m^h):=\arg\min_{P,R} B(m,P,R).$$

(15)

Moreover, we examine how hedging affects optimal operational decisions compared to the no-hedging case and completely characterize the return-risk efficient frontier in Section 5.

Last, we emphasize that a primary goal of this paper (see the research question (ii) in Section 1) is to compare the price and VPQ under the no-hedging model (7) and those under the hedging model (13). Both models are formulated as a risk minimization problem subject to an exogenously set target mean return (i.e., m), following the classical Markowitz's model (Markowitz 1952). To make a fair comparison between the two models, the values of m are *identical* in both models. (Note that m is the sum of the mean operational and hedging payoffs.)

REMARK 2. The risk-minimization objective of the problem (13) may lead to a reduction in the price and VPQ (see Sections 5.1 and 5.2). However, the firm may be incentivized to expand operational activities and use risk hedging to improve the return-risk trade-off instead of minimizing risk with a *fixed* target mean return. Our model can readily capture this different attitude of the firm as follows. Suppose the firm wants to expand production and prescribes a target production quantity. Then, the only change needed in our model is to fix Q at this prescribed level in the variance function $B(m,P,R)$ in (14) and optimize the price P. Alternatively, one may adopt the *mean-variance utility maximization* (MV) formulation to encompass the objective of expanding

---

[6] For example, θ_t is an adapted process (so the strategy does not look into the future) and $\chi_T(\vartheta)$ has a finite variance; other conditions are described in Section EC.7.2.1 of the e-companion.

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

15

operational activities. It is straightforward to show that the MV formulation is mathematically equivalent to our mV formulation (13), meaning they lead to the same efficient frontier. However, because the mV formulation fixes the target mean return whereas the MV formulation does not, the impact of hedging on the operational decisions may be substantially different. See more discussion in Section EC.1 of the e-companion.

4. Solving the Hedging Problem

We apply the quadratic hedging technique (Gourieroux et al. 1998) to solve the hedging problem (14). First, define the following martingale under P^M:

$$Z_t^M := \mathsf{E}^M(Z_T \,|\, \mathcal{F}_t), \quad \text{with} \quad dZ_t^M = \zeta_t dX_t, \tag{16}$$

where the dynamics of Z_t^M (i.e., dZ_t^M) follows directly from Martingale Representation Theorem, and ζ_t is a stochastic process adapted to $\mathcal{G}_t$ (refer to Corollary EC.3 in the e-companion). The solution approach essentially relies on the *projected* production payoff process, defined as

$$V_t(P, R) := \mathsf{E}^M[H_T(P, R)\,|\, \mathcal{F}_t], \quad 0 \le t \le T. \tag{17}$$

Clearly, V_t is a martingale under P^M, and its martingale representation (summarized as the lemma below) plays a crucial role in determining the optimal hedging strategy.

LEMMA 1. $V_t(P, R)$ *defined in (17) has the following representation:*

$$V_t(P, R) = V_0(P, R) + \int_0^t \xi_s(P, R)dX_s + \int_0^t \delta_s(P, R)d\tilde{B}_s, \tag{18}$$

where $V_0(P, R) = \mathsf{E}^M[H_T(P, R)]$ and ξ_t and δ_t are adapted processes. In particular,

$$\delta_t(P, R) = \tilde{\sigma}(P - s)\mathsf{P}^M(A_T \le R\,|\, \mathcal{F}_t), \tag{19}$$

which increases in both P and R, where A_T is defined in (10) and s is the unit salvage price.

4.1. Optimal Hedging Strategy and Minimal Variance

THEOREM 1. *Suppose Assumptions 4–5 hold.*

(i) The optimal solution to the problem in (14) is

$$\theta_t^*(P, R) = \underbrace{-\xi_t(P, R)}_{\text{risk-mitigation position}} + \underbrace{\iota_t(P, R)}_{\text{investment position}}, \tag{20}$$

where, with the terms Z_t^M, ζ_t, and ξ_t being defined in (16) and (18),

$$\iota_t(P, R) = -\frac{\zeta_t}{Z_t^M}\left[\underbrace{\frac{mZ_0^M - V_0(P, R)}{Z_0^M - 1}}_{\lambda_m} - V_t(P, R) - \underbrace{\int_0^t \theta_s^*(P, R)dX_s}_{\chi_t^*}\right].$$

16

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

(ii) The optimal objective function value in (14) has the following expression:

$$B(m, P, R) = \frac{[m - V_0(P, R)]^2}{Z_0^M - 1} + \int_0^T \mathsf{E}\Big[\frac{Z_t}{Z_t^M} \delta_t^2(P, R)\Big] dt, \tag{21}$$

where $\delta_t(P, R)$ is defined in (19). In particular, $Z_0^M = \mathsf{E}(Z_T^2) > 1$ and $0 \le Z_t/Z_t^M \le 1$.

(iii) Define the following terms:

$$\chi_T^{\mathrm{rm}}(P, R) := \int_0^T (-\xi_t(P, R)) dX_t, \quad \chi_T^{\mathrm{iv}}(P, R) := \int_0^T \iota_t(P, R) dX_t, \tag{22}$$

and

$$H_T^{\mathrm{h}}(P, R) := H_T(P, R) + \chi_T^{\mathrm{rm}}(P, R). \tag{23}$$

Then, $\mathsf{Var}(\chi_T^{\mathrm{iv}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}}) \ge 0$, $\mathsf{Var}(H_T^{\mathrm{h}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}}) \ge 0$, and

$$\underbrace{\sqrt{\mathsf{Var}(\chi_T^{\mathrm{iv}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}})}}_{investment\ risk} = \Big[\underbrace{\frac{[m - V_0(P, R)]^2}{Z_0^M - 1}}_{1st\ term\ of\ B(m, P, R)}\Big]^{1/2}, \tag{24}$$

$$\underbrace{\sqrt{\mathsf{Var}(H_T^{\mathrm{h}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}})}}_{unhedgeable\ risk} = \Big[\underbrace{\int_0^T \mathsf{E}\Big[\frac{Z_t}{Z_t^M} \delta_t^2(P, R)\Big] dt}_{2nd\ term\ of\ B(m, P, R)}\Big]^{1/2}. \tag{25}$$

Part (i) of Theorem 1 is intuitively appealing. The structure in (20) reveals two aspects of the optimal hedging strategy. First, the term $\xi_t(P, R)$, which also appears in (18), captures the impact of asset price movement (i.e., dX_t) on production payoff. By taking the opposite position, $-\xi_t(P, R)$, the hedging strategy offsets such impact to mitigate risk. The second term in (20) involves the difference between λ_m (a proxy of the target mean return m) and $V_t(P, R) + \chi_t^*$ (the projected total return at time t). This term functions as an *investment* to close this gap between the target and the total return by taking a position in the asset. Wang and Yao (2017) also note these two aspects but with specific asset price (GBM) and demand models.

Part (ii) of Theorem 1 presents the minimal variance of the hedging problem (14). Its structure is closely related to that of the optimal hedging strategy and can be interpreted from the perspective of risk decomposition, clarified in Part (iii). By definition, the total return can be rewritten as

$$H_T + \chi_T^* = H_T^{\mathrm{h}} + \chi_T^{\mathrm{iv}}, \tag{26}$$

where χ_T^{iv} is the payoff from the investment position ι_t of the optimal hedging strategy θ_t^*, and H_T^{h} represents the *hedged production payoff* for the following reason. Note that

$$H_T^{\mathrm{h}} = H_T + \chi_T^{\mathrm{rm}} = V_0 + \int_0^T \xi_t dX_t + \int_0^T \delta_t d\tilde{B}_t + \int_0^T (-\xi_t) dX_t = V_0 + \int_0^T \delta_t d\tilde{B}_t, \tag{27}$$

where the first equality is the definition (23), the second equality holds because of the fact that $H_T = V_T$, Lemma 1, and the definition of χ_T^{rm} in (22). The calculations in (27) reveals that χ_T^{rm} (which is the payoff from the risk-mitigation payoff of the optimal hedging strategy) offsets $\int_0^T \xi_t dX_t$ (which is the part of production payoff H_T that is driven by the asset price movement). Furthermore, taking expectations in (27) yields $V_0 = \mathsf{E}(H_T^{\mathrm{h}})$, so V_0 carries the economic meaning of *expected hedged production payoff.*

We now provide a risk-decomposition interpretation of (21). According to the definition of $B(m, P, R)$ in (14) and the total return's representation in (26), we have

$$B(m, P, R) = \mathsf{Var}(H_T^{\mathrm{h}} + \chi_T^{\mathrm{iv}}) = \left[\mathsf{Var}(\chi_T^{\mathrm{iv}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}}) \right] + \left[\mathsf{Var}(H_T^{\mathrm{h}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}}) \right].$$

Namely, we may decompose the variance of the total return into two parts,[7] one associated with the investment payoff χ_T^{iv} and the other with the hedged production payoff H_T^{h}. In particular, we call $\sqrt{\mathsf{Var}(\chi_T^{\mathrm{iv}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}})}$ *investment risk* and call $\sqrt{\mathsf{Var}(H_T^{\mathrm{h}}) + \mathsf{Cov}(H_T^{\mathrm{h}}, \chi^{\mathrm{iv}})}$ *unhedgeable risk.*

Furthermore, (24) and (25) show that these two types of risk directly translate to the two terms of the decomposition of the total variance $B(m, P, R)$ in (21). Their expressions are intuitive and revealing. First, note that the expected investment payoff needs to fill in the gap between the target mean return m and the expected hedged production payoff V_0, that is,

$$\mathsf{E}(\chi_T^{\mathrm{iv}}) = m - V_0. \tag{28}$$

Hence, the investment risk is proportional to $m - V_0$. (Proposition 7 in Section 5 implies that $m - V_0 > 0$ for the optimal (P, R) in the presence of hedging.) This implies that a larger gap (i.e., $m - V_0$) needs a higher investment payoff, inducing higher investment risk. Second, regarding the unhedgeable risk, because it does not involve m and increases in the demand noise $\tilde{\sigma}$ via δ_t in (19), it can not be reduced by trading financial asset. Moreover, the unhedgeable risk increases in both P and R, meaning that a higher operational level increases the exposure to the demand uncertainty that is unhedgeable. Assuming the product price is exogenously given, Wang and Yao (2017) and Wang (2021) note results similar to (21) under particular asset price models. However, these two papers do not establish the kind of relationship like (24) and (25) between the risk decomposition and the two terms of the minimum variance (21); they do not recognize V_0 as the expected hedged production payoff from an economic perspective.

In the presence of hedging, the optimal operational decisions minimizing $B(m, P, R)$ essentially assumes a market of size A_T^M, the version of A_T under the risk-neutral probability measure P^M

---

[7] This kind of variance decomposition is also used in Risk Parity, a widely applied portfolio management strategy in the financial industry; see Asness et al. (2012) and the white paper by Wealthfront, LLC. (`https://research.wealthfront.com/whitepapers/risk-parity`).

(formally, A_T^M is defined via $\mathsf{P}(A_T^M \leq a) = \mathsf{P}^M(A_T \leq a)$ for all a), because the operations-related terms in (21) (i.e., $V_0(P, R)$ and $\delta_t(P, R)$) only depend on the distribution of A_T under P^M. Hedging—particularly the risk mitigating term $-\xi_t$ in (20)—offsets any impact of the asset price trend on production payoff (via the impact on the market size A_T). Consider the following examples.

- The asset is a stock index, reflecting the general economic condition. Suppose its price X_t follows a GBM, and a higher asset price boosts the product demand. Then, the asset price trend's positive effect on the distribution of the market size A_T under P can be captured by an upward trend of X_t (i.e., $\mu > 0$). In the presence of hedging, A_T is evaluated under P^M, which corresponds to a GBM with zero drift (i.e., the asset price under P^M has no trend).

- The asset is a commodity (e.g., crude oil) with mean-reverting prices. The price tends to return to its long-term mean when deviating from the latter, impacting product demand. For instance, when the oil price rises towards its long-term mean, the demand for fuel-inefficient cars will fall. Suppose the price X_t follows an EOU process. Then, the asset price negatively impacts the demand if the initial price X_0 is below the long-term mean (see Sections 6). Similar to the last example, the distribution of A_T under P^M also corresponds to a GBM with zero drift.

In both examples, hedging removes the asset price trend's positive/negative effect on production payoff, and the firm optimizes operational decisions assuming that the market size will be as if the asset price has no trend, remaining at the current level in expectation. In other words, when the optimization of the operational decision (P, R) is concerned, hedging effectively replaces the actual market size A_T in the real world (i.e., under P) with a market size A_T^M in the risk-neutral world (i.e., under P^M) where the asset price has no trend. This insight will be crucial in analyzing optimal price and VPQ in Section 5.

REMARK 3. We solve the hedging problem (14) for a general continuous diffusion process having a *stochastic* MPR process. Our solution procedure is significantly more challenging than that of Wang and Yao (2017)—where the asset price is assumed to follow a GBM, inducing a *constant* MPR process—mainly concerning the characterization of ζ_t defined in (16). Specifically, ζ_t must be such that the optimal hedging strategy, θ_t^* in (20), is admissible (see the discussion below equation (13)). While it is straightforward to verify this property for GBMs, doing so for general diffusion processes is considerably more difficult (see Lemma EC.2 in the e-companion). Due to the generality, our model encompasses a much broader scope of applications than that of Wang and Yao (2017). For example, in Section 4.2, we apply Theorem 1 to an EOU process capturing asset prices' mean-reverting property, whereas the corresponding result of Wang and Yao (2017) does not cover mean-reverting processes. In particular, verifying Assumption 5 for EOU processes is nontrivial because of the difficulty of calculating critical terms including ζ_t/Z_t^M in the hedging strategy (20) and Z_t/Z_t^M in the variance function (21) (see Section EC.7.4 of the e-companion). By

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

19

contrast, Assumption 5 trivially holds for GBMs. Again, the difference stems from the nature of the asset price's MPR process: it is stochastic for EOU processes, whereas constant for GBMs.

4.2. Example: EOU Asset Price Model

To align with the automakers' example in Section 1, we assume the financial asset is crude oil, and its price X_t follows an EOU process, commonly used to model the mean-reverting feature of commodity prices (Schwartz 1997). Suppose $X_t = e^{Y_t}$ with $dY_t = \kappa(\alpha - Y_t)dt + \sigma dB_t$. Then,

$$dX_t = \kappa\left(\alpha + \frac{\sigma^2}{2\kappa} - \log X_t\right)X_t dt + \sigma X_t dB_t, \tag{29}$$

where κ, α, and σ are all positive constants and the dynamics of the asset price X_t directly follows from Itô's Lemma. In particular, α represents the long-term mean of Y_t and κ is the mean-reversion coefficient. The MPR η_t in (11) corresponding to the EOU process is

$$\eta_t = \frac{\kappa}{\sigma}\left(\alpha + \frac{\sigma^2}{2\kappa} - \log X_t\right). \tag{30}$$

When the mean-reversion coefficient κ approaches zero, Y_t in (29) reduces to σB_t, X_t in (29) reduces to a GBM, and η_t in (30) becomes $\sigma/2$. All the results in this section hold for this limiting case.

To apply Theorem 1 to the EOU model, we need to verify Assumptions 4 and 5. The former holds trivially, and for the latter, we impose the following sufficient condition (see Lemma EC.5 in the e-companion):

$$\kappa T < \frac{\pi}{4}. \tag{31}$$

PROPOSITION 3. *Suppose that X_t follows the EOU process (29) and that the condition (31) holds.*

(i) The optimal solution to the problem in (14) is

$$\theta_t^*(P,R) = -\xi_t(P,R) - \frac{a(T-t) + \frac{\kappa}{\sigma^2}(\alpha - Y_t)b(T-t)}{X_t}[\lambda_m - V_t(P,R) - \chi_t^*],$$

where $\chi_t^* = \int_0^t \theta_s^*(P,R)dX_s$, $\xi_t(P,R)$ is defined in (18), λ_m is defined in (20), and a and b are functions defined as

$$a(\tau) = \frac{1}{2} + \frac{1}{\cos\kappa\tau - \sin\kappa\tau}, \quad b(\tau) = \frac{\cos\kappa\tau + \sin\kappa\tau}{\cos\kappa\tau - \sin\kappa\tau}, \quad \tau \in [0,T]. \tag{32}$$

(ii) The optimal objective function value of the problem in (14) is

$$B(m,P,R) = \frac{[m - V_0(P,R)]^2}{Z_0^M - 1} + \int_0^T \mathsf{E}[e^{-f_0(T-t) - f_1(T-t)Y_t - f_2(T-t)Y_t^2}\delta_t^2(P,R)]dt, \tag{33}$$

where $\delta_t(P,R)$ is defined in (18), and f_0, f_1, and f_2 are functions defined as

$$
\begin{aligned}
f_0(\tau) &= -\alpha - \left(\frac{1}{2}\kappa + \frac{1}{4}\sigma^2\right)\tau - \frac{1}{2}\log[\cos\kappa\tau - \sin\kappa\tau] + \frac{\alpha + \left(\frac{\alpha^2\kappa}{\sigma^2} + \frac{\sigma^2}{2\kappa}\right)\sin\kappa\tau}{\cos\kappa\tau - \sin\kappa\tau}, \\
f_1(\tau) &= \frac{-1 + \cos\kappa\tau - \left(\frac{2\kappa\alpha}{\sigma^2} + 1\right)\sin\kappa\tau}{\cos\kappa\tau - \sin\kappa\tau}, \quad f_2(\tau) = \frac{\kappa}{\sigma^2}\frac{\sin\kappa\tau}{\cos\kappa\tau - \sin\kappa\tau}, \quad \tau \in [0,T].
\end{aligned}
\tag{34}
$$

Functions in (32) and (34) are well defined under the condition (31). By further specifying the market size A_T in (10), we can derive $\xi_t(P,R)$ and $\delta_t(P,R)$ explicitly (see Section EC.4.1 of the e-companion).

20

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

5. Optimal Price and VPQ in the Presence of Hedging

In this section, we focus on the operational task and solve problem (15) to find the jointly optimal price and VPQ in the presence of hedging.[8] We present the main results of this paper regarding optimal price and VPQ in the presence of hedging, discussing their properties depending on whether the asset price trend positively (Section 5.1) or negatively (Section 5.2) impacts the demand. We also fully characterize the efficient frontier for the hedging model (Section 5.3).

Sections 5.1 and 5.2 focus on the case that the target mean return is set as $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$. This return level is an important benchmark for decision-making as it is the maximum expected profit that the firm can attain (without hedging) in the base model (1). It is economically sensible for a non-financial firm (e.g., a manufacturer) to target its earnings around this level. Setting a substantially lower target is certainly inconsistent with the profit-generating goal of the firm. On the other hand, targeting at a significantly higher level than m will alter the non-financial nature of the firm since the extra income can only come from trading the financial asset, which does not align with the firm's primary operations. With $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$, it is straightforward to see that the risk-minimizing price and VPQ in the no-hedging model (7) are identical to the profit-maximizing decisions in the base model (1), both equal to $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$. Hence, Theorems 2 and 3 compare $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ with the risk-minimizing decisions (P^h_m, R^h_m), and their conclusions are based on the condition $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$.

REMARK 4. There exist situations in which other choices of m may better reflect the firm's attitude. For example, if it anticipates the product demand is low (hence a low value of $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$), the firm may want to set m significantly higher than $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$ and use hedging to fill in the gap between the operational payoff and the target. In other words, the firm views hedging as a tool for profit improvement instead of risk minimization. For another example, if its risk tolerance is low, the firm may prefer a target m that is substantially below $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$ to reduce risk. When $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$ does not reasonably reflect the firm's return or risk preferences, or when it is not straightforward to select a proper m, the mean-variance utility maximization (MV) formulation may be a better choice, provided that the firm's risk-aversion coefficient is known or easy to estimate. In this case, the efficient frontier generated by our mV formulation can be used to solve the MV formulation; see Section EC.1 of the e-companion.

5.1. Asset Price Trend Positively Impacting Demand

We first examine the case of the asset price trend positively impacting demand (i.e., $C_T \succeq C_T^M$). A leading example is automakers: when oil prices exhibit a downward trend, demand for fuel-inefficient vehicles such as SUVs or pickup trucks will increase. We compare the risk-minimizing

---

[8] For a given $m \geq 0$, numerically minimizing $B(m, P, R)$ over (P, R) can be reduced to a line search over a finite interval—instead of a two-dimensional grid search—because $B(m, P, R)$ is convex with respect to P within a specific range; see Proposition EC.2 in Section EC.4.3 of the e-companion.

operational decisions in the hedging and no-hedging models, that is, (P_m^h, R_m^h) versus $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ for $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$. To this end, we introduce an intermediate set of operational decisions. Define the profit-maximizing price and VPQ under the risk-neutral measure P^M as

$$(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}}) := \arg\max_{P,R} V_0(P, R).$$

Recall that, by definition, $V_0(P, R) = \mathsf{E}^M[H_T(P, R)]$ is the expected production payoff under the risk-neutral measure, so $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$ maximizes the expected production payoff without considering the asset price trend. In the following, we present two sets of comparisons: one comparing $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ with $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$, the other comparing $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$ with (P_m^h, R_m^h).

PROPOSITION 4. *If $C_T \succeq C_T^M$, then (i) $P^{\mathrm{NV}} \geq P^{\mathrm{NV(M)}}$ and $R^{\mathrm{NV}} \geq R^{\mathrm{NV(M)}}$; (ii) $\mathsf{E}[H_T(P, R)] \geq V_0(P, R)$ for all (P, R).*

Part (i) of Proposition 4 follows from an analysis on the marginal effect of the price and VPQ on the expected production payoff $\mathsf{E}[H_T(P, R)]$. In particular, $\frac{\partial \mathsf{E}(H_T)}{\partial P} = R - 2bP + bc - \mathsf{E}[(R - A_T)^+]$ and $\frac{\partial \mathsf{E}(H_T)}{\partial R} = (P - c) - (P - s)\mathsf{P}(A_T \leq R)$ are both increasing in A_T, indicating that a larger market size induces more significant marginal benefits from increases in P and R. As a result, the optimal price and VPQ are higher in the presence of a stochastically larger market size. The economic intuition is as follows. A larger market size provides more buffer to cope with the negative impact of a higher price; hence, the firm has more room to increase the price and thus generate a higher maximum profit. In addition to increasing the price, the firm can simultaneously increase the production quantity Q to capture the greater demand. The VPQ (recall, $R = Q + bP$) integrates the combined effect of pricing and production decisions and thus increases with a larger market size. We know from Remark 1 that if $C_T \succeq C_T^M$, then $A_T \succeq A_T^M$, so the profit-maximizing price and VPQ are lower with a market size A_T^M than those with a market size A_T. By definition, the former decisions are $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$ and the latter are $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$.

Regarding Part (ii) of Proposition 4, we know from its expression in (3) that $\mathsf{E}[H_T(P, R)]$ is increasing in A_T for any price and VPQ. Hence, with $A_T \succeq A_T^M$, the expected production payoff is greater in the real-world than in the risk-neutral world. By definition, the former is $\mathsf{E}[H_T(P, R)]$ and the latter is $\mathsf{E}^M[H_T(P, R)] = V_0(P, R)$.

PROPOSITION 5. *$P_m^h \leq P^{\mathrm{NV(M)}}$ and $R_m^h \leq R^{\mathrm{NV(M)}}$ for all $m \geq 0$.*

Proposition 5 holds regardless of the relationship between C_T and C_T^M. It asserts that the risk-minimizing price and VPQ in the presence of hedging are bounded above by the profit-maximizing price and VPQ under the risk-neutral measure, respectively. From an economic perspective, among the set of operational decisions (P, R) that attain the same expected hedged production payoff

22

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

$V_0(P, R)$, those with smaller values are preferred as they induce lower exposure to unhedgeable uncertainties. Hence, values higher than $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$, which maximizes $V_0(P, R)$, correspond to over-pricing and over-stocking, thereby not return-risk efficient.

To further clarify the intuition behind Proposition 5, consider the particular case in which $C_T \succeq C_T^M$ and $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$ (i.e., the setting of Theorem 2). In this case, we can show[9] that $V_0(P, R) \leq m$ for all (P, R). Because $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$ maximizes $V_0(P, R)$, increasing (P, R) beyond $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$ would make $V_0(P, R)$ begin to decrease, which increases the value of $|m - V_0(P, R)| = m - V_0(P, R)$ and thus increases the investment risk defined in (24). Meanwhile, the unhedgeable risk defined in (25) is always increasing in both P and R. Thus, increasing (P, R) beyond $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$ would make the total risk measured by $B(m, P, R)$ in (21) begin to increase, so (P_m^h, R_m^h), which minimizes $B(m, P, R)$, must not exceed $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$.

THEOREM 2. *Suppose Assumptions 1–5 hold and the asset price trend positively impacts demand (i.e., $C_T \succeq C_T^M$). Let $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$. Then, $P_m^h \leq P^{\mathrm{NV}}$ and $R_m^h \leq R^{\mathrm{NV}}$.*

Theorem 2 is an immediate result of Propositions 4 and 5. It indicates that the risk-minimizing operational activities in the presence of hedging are lower than the profit-maximizing operational activities without hedging, when the asset price positively impacts demand. On the one hand, as the discussion following Theorem 1 reveals, with risk hedging, the firm operates as if in the risk-neutral world with a market size A_T^M, which is (stochastically) smaller than the market size A_T in the real world when $C_T \succeq C_T^M$. This implies that the profit-maximizing price and VPQ are lower in the presence of hedging than without hedging; that is, $P^{\mathrm{NV}} \geq P^{\mathrm{NV(M)}}$ and $R^{\mathrm{NV}} \geq R^{\mathrm{NV(M)}}$ (Proposition 4). On the other hand, through the risk decomposition in Theorem 1, we know that increasing the price and VPQ beyond $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$ would increase both the investment risk and the unhedgeable risk, so the risk-minimizing price and VPQ in the presence of hedging must not exceed $(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$, meaning $P_m^h \leq P^{\mathrm{NV(M)}}$ and $R_m^h \leq R^{\mathrm{NV(M)}}$ (Proposition 5). Therefore, the risk-minimizing operational activities must be at most the profit-maximizing levels of the base model to avoid unnecessary risk that does not bring any additional return.

REMARK 5. The insight of Theorem 2—that is, within our risk-minimization framework, hedging reduces operational activities when the asset price positively impacts demand—can be generalized beyond the price-setting newsvendor problem. Indeed, it holds for general operational payoff functions with certain structural properties; see Section EC.5 of the e-companion.

---

[9] First, because $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ maximizes $\mathsf{E}[H_T(P, R)]$, we have $m \geq \mathsf{E}[H_T(P, R)]$ for all (P, R). Next, by Proposition 4, if $C_T \succeq C_T^M$, $\mathsf{E}[H_T(P, R)] \geq V_0(P, R)$ for all (P, R).

5.2. Asset Price Trend Negatively Impacting Demand

Now we consider the case of the asset price trend negatively impacting demand (i.e., $C_T \preceq C_T^M$). In this case, hedging can offset the negative impact on the production payoff. For example, when oil prices exhibit an upward trend, the demand for fuel-inefficient cars such as SUVs decreases. The firm can enter a long position[10] for oil shares to hedge against the oil price's negative impact on the demand for SUVs. By doing so, the firm receives a positive return from the hedging, making the expected hedged production payoff higher than the expected production payoff (without hedging).

PROPOSITION 6. *If $C_T \preceq C_T^M$, then (i) $P^{\mathrm{NV}} \leq P^{\mathrm{NV(M)}}$ and $R^{\mathrm{NV}} \leq R^{\mathrm{NV(M)}}$; $\mathsf{E}[H_T(P,R)] \leq V_0(P,R)$ for all (P,R).*

Proposition 6 is analogous to Proposition 4 and follows from the same argument based on analyzing the marginal effect of the price and VPQ on the expected production payoff.

PROPOSITION 7. *$V_0(P_m^h, R_m^h) \leq m$ for all $m \geq 0$.*

Proposition 7 is crucial for establishing Theorem 3, which analyzes the optimal operational decisions in the presence of hedging when the asset price trend negatively impacts the demand. Recall from the discussion following Theorem 1 that $V_0(P_m^h, R_m^h)$ represents the *expected hedged production payoff*. Thus, if $V_0(P_m^h, R_m^h) > m$, we can *increase* the target mean return from m to $m' = V_0(P_m^h, R_m^h)$ while retaining the same operational decisions (P_m^h, R_m^h). By doing so, the investment risk defined in (24) is reduced to zero, and the unhedgeable risk, which is defined in (25) and only depends on the operational decisions, remains the same, implying $B(m', P_m^h, R_m^h) < B(m, P_m^h, R_m^h)$. In other words, when the target mean return is $m' = V_0(P_m^h, R_m^h)$, the operational decisions (P_m^h, R_m^h) along with the hedging strategy $\vartheta^*(m', P_m^h, R_m^h)$ lead to an expected payoff $m' > m$ and and a variance $B(m', P_m^h, R_m^h) < B(m, P_m^h, R_m^h)$. Therefore, if $V_0(P_m^h, R_m^h) > m$, the set of integrated operational and hedging decisions $(P_m^h, R_m^h, \vartheta^*(m, P_m^h, R_m^h))$ would not be efficient, because it is dominated by another set of decisions $(P_m^h, R_m^h, \vartheta^*(m', P_m^h, R_m^h))$ which induces a higher return and lower risk.

To better understand why such inefficiency would arise, note that $\mathsf{E}[\chi_T^{\mathrm{iv}}(P_m^h, R_m^h)] = m - V_0(P_m^h, R_m^h)$ represents the *expected investment payoff* from hedging (see (22) and (28)), and it would become *negative* if $V_0(P_m^h, R_m^h) > m$. While the investment payoff's contribution to the total risk is always nonnegative, a negative return from the investment is clearly detrimental to the return-risk trade-off of the total return, which can be expressed as $H_T^{\mathrm{h}} + \chi_T^{\mathrm{iv}}$, the sum of the hedged production payoff and the investment payoff; see (26). Consequently, the joint decisions $(P_m^h, R_m^h, \vartheta^*(P_m^h, R_m^h))$ cannot be return-risk efficient because part of the payoff (i.e., the investment payoff χ_T^{iv}) is highly inefficient as it bears a negative return and a positive contribution to the total risk. Wang and Yao (2017) and

---

[10] The risk mitigation component of the hedging strategy, i.e., the term $-\xi_t$ in (20), essentially reverses the effect of asset price on the production payoff.

24

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

Wang (2021) also note that the condition $V_0(P_m^h, R_m^h) \leq m$ is necessary for the optimality based on the observation that increasing the target mean return from m to V_0 would induce a better return-risk trade-off. However, these two papers do not provide an economic interpretation of this result as we do here, due to their lack of the risk decomposition presented in Parts (ii) and (iii) of Theorem 1.

THEOREM 3. *Suppose Assumptions 1–5 hold and the asset price trend negatively impacts demand (i.e., $C_T \preceq C_T^M$). Let $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$ and let $\bar{P} = P - s$ for any P. Define*

$$P^\circ := P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) \quad and \quad r^\circ := \frac{\bar{P}^\circ}{\bar{P}^{\mathrm{NV}}}. \tag{35}$$

(Note: $r^\circ \geq 1$.) Then, the following properties hold.

(i) $P_m^h \leq P^{\mathrm{NV}}$.

(ii) $R_m^h \leq R^{\mathrm{NV}}$, *if*

$$\left[\frac{\bar{P}^{\mathrm{NV}}}{r^\circ + \sqrt{(r^\circ)^2 - 1}}\right] \cdot \mathsf{P}^M(A_T \geq R^{\mathrm{NV}}) \leq c - s. \tag{36}$$

(iii) $R_m^h \leq R^\circ$, *if (36) does not hold, where R° (which exists uniquely) is determined by*

$$\frac{\mathsf{E}^M\left[A_T \mathbf{1}\{A_T \leq R^\circ\}\right]}{\mathsf{P}^M(A_T \geq R^\circ)} = \frac{\bar{P}^{\mathrm{NV}}(2b\bar{P}^\circ + 2bs - bc)}{(c-s)(r^\circ + \sqrt{(r^\circ)^2 - 1})} - R^{\mathrm{NV}}.$$

In particular, $R^{\mathrm{NV}} \leq R^\circ \leq R^{\mathrm{NV(M)}}$.

Part (i) of Theorem 3 indicates that hedging reduces the pricing level when the asset price negatively impacts demand. This is counter-intuitive because the assumption here ($C_T \preceq C_T^M$) is opposite to the assumption of Theorem 2 ($C_T \succeq C_T^M$), and yet the conclusions regarding the risk-minimizing price are the same. Nevertheless, the economic intuition here is different from that of Theorem 2.

When the asset price negatively impacts demand, with risk hedging, the firm operates as if in the risk-neutral world where the market size A_T^M is (stochastically) larger than the actual market size A_T in the real world. In this case, the expected production payoff is greater in the risk-neutral world than in the real world regardless of the operational decisions; that is, $V_0(P, R) \geq \mathsf{E}[H_T(P, R)]$ for all (P, R) (Proposition 6). In particular, we have $V_0(P^{\mathrm{NV}}, R^{\mathrm{NV}}) \geq \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] = m$, so the expected investment payoff (defined in (22)) associated with $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ would be $\mathsf{E}[\chi_T^{\mathrm{iv}}(P^{\mathrm{NV}}, R^{\mathrm{NV}})] = m - V_0(P^{\mathrm{NV}}, R^{\mathrm{NV}}) \leq 0$. However, by Proposition 7, the expected investment payoff associated with the risk-minimizing operational decisions must satisfy $\mathsf{E}[\chi_T^{\mathrm{iv}}(P_m^h, R_m^h)] = m - V_0(P_m^h, R_m^h) \geq 0$. In other words, the profit-maximizing operational decisions $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ are overly aggressive for risk minimization, because they generate more production payoff (in the risk-neutral world) than necessary to meet the target m, resulting in a negative expected investment payoff and introducing

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

25

unnecessary risk. Instead, the firm should decrease the price from P^{NV} to reduce the expected hedged production payoff V_0. Doing so would *leave leeway* in the target mean return for the hedging to provide a positive expected investment payoff while simultaneously reducing the investment and unhedgeable risks.

Both Parts (ii) and (iii) of Theorem 3 characterize the optimal VPQ in the presence of hedging. However, they differ in the assumption regarding the magnitude of the asset price trend's negative effect, particularly whether the condition (36) holds. This condition is interpreted as follows. Suppose the asset price trend's negative effect on product demand strengthens. (For example, if we fix the initial oil price below the long-term average level, any increment in this level will negatively affect the demand for fuel-inefficient cars more due to the upward oil price trend.) Then, the market size under the risk-neutral measure (i.e., A_T^M) is not affected since it does not involve the oil price trend while the actual market size (i.e., A_T) decreases, which induces a decreased VPQ R^{NV}. Then, because It then follows from the fact that $P^{\text{NV}} = [\mathsf{E}(R^{\text{NV}} \wedge A_T) + bc]/(2b)$ and $P^\circ = [\mathsf{E}(R^{\text{NV}} \wedge A_T^M) + bc]/(2b)$ that both P^{NV} and P° decrease. Because A_T decreases while A_T^M is not affected, we expect P^{NV} to decrease more than P° does, so r°, defined in (35), will increase as the negative effect strengthens. Thus, the term in the bracket on the left-hand side of the inequality (36), which is the main factor controlling the magnitude of the left-hand side (since the probability term never exceeds 1), decreases.

In a nutshell, (36) means that the asset price trend's negative effect on product demand is strong. Part (ii) states that under this circumstance, the optimal VPQ is lower in the presence of hedging than without hedging. This is consistent with the insight from Part (i). When the negative effect of asset price on product demand is strong, the expected hedged production payoff $V_0(P, R)$ is substantially larger than the mean of actual production payoff ($\mathsf{E}[H_T(P, R)]$) because hedging contributes significant return in this scenario. Consequently, to prevent the hedged production payoff from exceeding m (i.e., the benchmark production payoff without hedging), the exposure to asset price needs to be contained to reduce the exposure to the product market. This is achieved by—in addition to the price markdown—further suppressing the VPQ.

On the other hand, if the asset price trends' negative effect is not strong enough, Part (iii) states that we can upper bound the optimal VPQ R_m^h by a value lower than $R^{\text{NV(M)}}$, the optimal VPQ under the risk-neutral measure. In this case, the market size under the risk-neutral measure is not substantially larger than the actual one due to the mild negative effect on product demand, and we expect R^{NV} to be not far below $R^{\text{NV(M)}}$. Therefore, even if R_m^h exceeds R^{NV}, we expect the exceeding amount to be moderate.

In summary, Parts (ii) and (iii) of Theorem 3 indicate that, in the presence of hedging, either the optimal VPQ is lower than R^{NV} when the negative effect is strong, or otherwise it exceeds R^{NV} by a moderate amount (refer to Section EC.4.5 of the e-companion for a numerical illustration).

26

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

5.3. Efficient Frontier

PROPOSITION 8. $B(m, P_m^h, R_m^h)$ *increases in* m; *thus,* $(m, B(m, P_m^h, R_m^h))$ *is an efficient frontier.*

Proposition 8 indicates that, at optimality, higher risk accompanies a higher return. This efficient frontier lies lower than the one specified in Part (iii) of Proposition 2. That is, $B(m, P_m^h, R_m^h)$ is smaller than $\mathsf{Var}(H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}))$ for any target mean return $m \geq 0$ and the gap between the two represents the risk reduction achieved by the hedging model from the no-hedging model. In addition to characterizing the return-risk trade-off, we can use this efficient frontier to optimize other risk objectives, such as mean-variance or risk-averse utility functions; see Section EC.1 of the e-companion.

6. Numerical Study

We now implement the hedging model developed in this paper using real-world financial and automotive sales/price data sets. With the asset and demand models calibrated from data, we conduct a comprehensive numerical study[11] to illustrate various aspects of the analytical results derived in Sections 4 and 5. We briefly introduce the data and model calibration in Sections 6.1 and 6.2, respectively, and relegate the details to the e-companion. We examine the risk-minimizing operational decisions in the hedging model when the target return is the newsvendor's maximum profit in the base model, and compare them with the counterparts in the no-hedging model (Section 6.3). Then, we examine the hedging performance and compare the efficient frontiers for the hedging and no-hedging models (Section 6.4).

6.1. Data Description

Two sets of data are used: one is financial data, and the other is operational data from `AutoMFR`, a prominent automotive manufacturer. The financial asset in our context is the WTI crude oil, a primary global oil benchmark. The data source is the Federal Reserve Bank of St. Louis database. The financial data includes *daily* spot prices from 2010 to 2019. `AutoMFR`'s operational data, including *monthly* sales volumes and manufacturer-suggested retail prices (MSRPs), was purchased from a commercial vendor specializing in automotive business data. The operational data includes brands, models, versions, MSRPs, and combined MPG. We focus on two popular models manufactured by `AutoMFR`—`Sport` (fuel-inefficient, low MPG) and `Compact` (fuel-efficient, high MPG). The data for `Sport` ranges from January 2011 to December 2019, and the data for `Compact` ranges from January 2010 to May 2018. We deseasonalize the sales data using the standard STL method (Cleveland et al. 1990). In addition, we adjust the MSRP for inflation using monthly Consumer Price Index

---

[11] There are two crucial differences between our numerical study and that by Wang and Yao (2017). First, we use real-world data sets for model calibration, whereas they only use simulated data. Second, we use the EOU process to model oil prices, whereas they use GBM, which is more suitable for modeling stock prices.

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

27

(CPI) data ranging from January 2010 to December 2019 for all urban consumers from the Federal Reserve Bank of St. Louis database. Specifically, the adjusted price of a given month is the original price divided by the CPI index of that month, then multiplied by the CPI index of December 2019. In other words, each price is adjusted to December 2019 by accounting for the inflation rate prevailing from that month to December 2019. The sales volume for each model is aggregated from sales across various versions and the price is the weighted (by sales volume) average price of different versions. The summary statistics of these two car models can be found in Section EC.4.2 of the e-companion .

6.2. Model Calibration

In the automaker's case under consideration, one selling period is one month (i.e., $T = 1/12$). We assume that the price of WTI—the relevant financial set—follows an EOU process. We also adopt the specification of Wang and Yao (2017) to model the market size A_T defined in (10) as

$$A_T = \int_0^T \tilde{\mu}(X_t)dt + \tilde{\sigma}\tilde{B}_T = \int_0^T (\mu_0 + \mu_1 X_t)dt + \tilde{\sigma}\tilde{B}_T, \tag{37}$$

where $\tilde{\mu}(x)$ is the *demand rate* function and we assume it to be linear in the asset price X_t.

Although, in reality, the unsold units of each month can be carried over to subsequent months, we do not model this inventory carryover in this numerical study. Instead, we assume for simplicity that the firm makes a myopic decision each month without considering the sales revenue generated by inventory carried over to and sold in later months. With this assumption, we set the salvage value to zero (i.e., $s = 0$) to reflect that units sold in later months are not accounted for in the current month's payoff (i.e., H_T). Furthermore, the monthly production quantity is unavailable in the data set, so we can not match its model-implied values to the real ones. Instead, the calibrating procedure matches, besides the model-implied and real prices, the model-implied average monthly sales volumes and the real ones. Despite the compromises made in the calibrating procedure, the calibrated model does capture qualitative characteristics of the real-world scenario. For example, the price sensitivity b is positive, and μ_1 in (37) is positive (resp., negative) for fuel-efficient (resp., fuel-inefficient) car model, i.e., `Compact` (resp., `Sport`). In particular, the estimated average gross profit margins match the real-world value. The calibrating procedure for the asset price and demand models and the estimated parameters are detailed in Section EC.4.2 of the e-companion.[12]

---

[12] We have also verified that Assumptions 1–3 and Assumption 5 (to be exact, its sufficient condition (31)) hold for the estimated parameters and the initial oil prices used in Sections 6.3 and 6.4. In addition, for $\underline{P}$ and $\underline{R}$ defined in Proposition 2, we have $(\underline{P} - c)/c = 0.01\%$ and $(\underline{R} - bc)/(bc) = 0.1\%$; thus, $\underline{P}$ and $\underline{R}$ are close to c and bc, respectively.

28

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

6.3. Assessing the Hedging Model: Fixed Target Return

Suppose the target return is the newsvendor's maximum profit in the base model (i.e., $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$). We consider three initial oil prices: $X_0 = 40, 70$, and 100. The estimated long-run average oil price is around \$70 (see Section EC.4.2.1 of the e-companion). Thus, the case of $X_0 = 40$ represents an upward asset price trend; in this case, the trend positively impacts the demand for `Compact` but negatively affects the demand for `Sport`. By contrast, the case of $X_0 = 100$ represents a downward asset price trend; in this case, the effects of the trend on the demands for these two models are reversed. The positive/negative effect (formally $C_T \succeq C_T^M$ or $C_T \preceq C_T^M$) is confirmed by the Mann–Whitney U-test applied to the simulated samples of C_T and C_T^M with the calibrated parameters (see Section EC.4.4 of the e-companion). The case of $X_0 = 70$ represents a negligible trend scenario as it is very close to the long-run average. With these three initial oil prices, we compute the optimal pricing/VPQ/production decisions for both `Sport` and `Compact`. The results are summarized in Table 2 and we discuss them below.

Table 2 Production (Q), Price (P), VPQ (R), Risks and Return with $\mathsf{m} = \mathsf{E}[\mathbf{H_T}(\mathbf{P}^{\mathrm{NV}}, \mathbf{R}^{\mathrm{NV}})]$.

Car	X_0	Model	Q	P	R	Risk	Unhedgeable Risk	Return (m)
Sport	40	NV	15,548	42,079	100,549	40.86	40.52	103.41
		Hedging	(-5.28%)	(-1.04%)	(-1.69%)	(-36.73%)	(-43.73%)	(96.38%)
	70	NV	12,552	40,627	94,618	34.09	33.72	65.58
		Hedging	(-1.99%)	(-0.34%)	(-0.56%)	(-10.18%)	(-13.88%)	(99.49%)
	100	NV	9,516	39,155	88,608	28.72	28.05	35.97
		Hedging	(-6.19%)	(-0.78%)	(-1.36%)	(-23.24%)	(-32.20%)	(96.55%)
Compact	40	NV	9,826	21,946	154,511	7.99	7.97	12.90
		Hedging	(-6.90%)	(-0.48%)	(-0.89%)	(-39.86%)	(-51.09%)	(94.19%)
	70	NV	12,263	22,313	159,364	9.30	9.19	20.49
		Hedging	(-1.42%)	(-0.13%)	(-0.23%)	(-8.47%)	(-13.94%)	(99.63%)
	100	NV	14,646	22,670	164,105	11.09	10.67	29.55
		Hedging	(-3.23%)	(-0.33%)	(-0.59%)	(-32.44%)	(-36.27%)	(98.08%)

Note. The percentages in the columns labeled "Q", "P", "R", 'Risk", and "Unhedgeable Risk" represent the reduction relative to the no-hedging model; see Section EC.2 for the formal definitions of risk and unhedgeable risk for the no-hedging model. The percentages in the column labeled "Return" represent the contribution from the production payoff. The units of "Risk", "Unhedgeable Risk", and "Return" are all one million dollars.

First, when the initial oil price for a particular month significantly deviates from the long-run average (i.e., $X_0 = 40$ or 100), the asset price exhibits a prominent trend towards its long-run level due to the mean-reverting effect. For such cases, the risk reduction achieved by the hedging model is significantly more than the risk reductions achieved for the case of $X_0 = 70$ (which is close to the long-run average). The key reason is that a more salient asset price trend leads to a higher return-to-risk ratio of the investment, meaning subject to the same level of investment risk, one

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

29

earns a higher investment payoff $\mathsf{E}(\chi_T^{\mathrm{iv}})$ and thus requires a lower production payoff V_0 to achieve the target for the total expected payoff m. (Recall from (28) that $\mathsf{E}(\chi_T^{\mathrm{iv}}) + V_0 = m$.) Therefore, the higher return-to-risk ratio of the investment payoff induces more reductions in the operational levels in the presence of hedging, thereby reducing the unhedgeable risk (and thus the total risk) by a more significant amount.

We now explain why the case of $X_0 = 40$ or 100 leads to a higher return-to-risk ratio of the investment than the case of $X_0 = 70$. By the expressions for the expected investment payoff and the investment risk (see (28) and (24), respectively), we know that the return-to-risk ratio of the investment is $\sqrt{Z_0^M - 1}$. In the example that X_t follows an EOU process, we can show—using the fact that $Z_0^M = \mathsf{E}(Z_T^2)$ and Jensen's inequality—that $Z_0^M \geq \exp(\frac{1}{2} \int_0^T \mathsf{E}^M(\eta_t^2) dt)$, where $\eta_t = \frac{\kappa}{\sigma}(\alpha + \frac{\sigma^2}{2\kappa} - \log X_t)$ is the market price of risk (see (30)). Hence, when $X_0 = 40$ or 100, $\log X_t$ deviates significantly from the long-run average α,[13] leading to a large value of η_t^2. Then, the above lower bound suggests a large value of Z_0^M, meaning a high return-to-risk ratio of the investment.

A second observation from Table 2 is that, for both car models, the optimal operational decisions (price P, VPQ R and production Q) in the presence of hedging are adjusted downward relative to those without hedging. However, the decrements are small, around 7% at most, so the production payoff still contributes most of the total return. This is reassuring, as hedging does not excessively decrease the operational level. A particularly desirable trait is that the price markdown is small: while price reduction enhances market competitiveness, automakers are usually reluctant to reduce prices too much due to the concern of protecting their brands' values. Moreover, the "Unhedgeable Risk" column of Table 2 shows that even small markdowns in the operational levels can substantially reduce the unhedgeable risk. The reason is as follows. The slope of the efficient frontier for the no-hedging model is exploding as m approaches the maximum profit level $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$ (see Proposition 2(iii)). Thus, decreasing the operational levels even slightly, which reduces the expected production payoff, can substantially reduce the exposure to the noise in the product demand, which is unhedgeable.

Last, when the asset price trend's impact on product demand is strong ($X_0 = 40$ or 100), hedging makes more significant adjustments to the operational decision than the weaker impact case ($X_0 = 70$). This is consistent with the finding that hedging is more effective in the former case.

6.4. Assessing the Hedging Model: Varying Target Returns

6.4.1. Efficient Frontiers
Figure 1 plots the efficient frontiers for both car models with an initial oil price $X_0 = 40$. The target return m ranges from 90% to 100% of $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$, the

---

[13] It is easy to verify that $\log X_t = -\frac{1}{2}\sigma^2 t + \sigma B_t^M$. Then, $\eta_t = \frac{\kappa}{\sigma}(\alpha + \frac{\sigma^2}{2\kappa} + \frac{1}{2}\sigma^2 t - \sigma B_t^M)$. When $X_0 = 40$, $\log X_t$ is significantly below α, and η_t tends to become more positive as t increases. When $X_0 = 100$, $\log X_t$ is above α; thus, so long as the gap is significant and t is small ($t \leq \frac{1}{12}$ in our case), η_t still tends to take a significantly negative value.

30

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

newsvendor's maximum profit in the base model. The risk is measured by the standard deviation of the terminal return. (The right ends of the curves correspond to the results in Table 2). All curves are upward sloping, which is the hallmark of efficient frontiers: when all decisions are optimized, an increase in return is always accompanied by increased risk.

Figure 1   Efficient Frontiers.
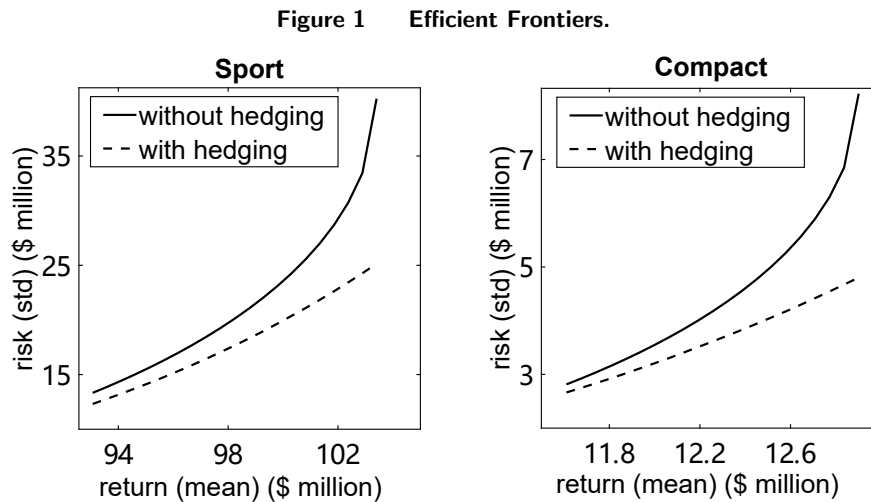
Note. Initial WTI price is $X_0 = 40$.

Figure 1 clearly shows, as expected, that the hedging model outperforms the no-hedging model: the frontiers for the former lie below the frontiers for the latter. More significantly, it sheds light on the difference in *incremental risk* (i.e., the slope of a frontier) between the two models. Specifically, for Sport or Compact, the frontier for the hedging model increases more slowly than that for the no-hedging model, especially towards the right ends of the frontiers. In other words, the hedging model's incremental risk is smaller than the no-hedging model's, and this phenomenon is more prominent for high-return cases. The reason is that, in the presence of hedging, the operational levels P and R need not grow as fast as those in the no-hedging model (see Figure 3 and the related discussion) to reach higher values of the target mean return m. Such smaller increments in P and R induce the slower growth of the unhedgeable risk, thereby, the slower growth of the total risk.

The smaller incremental risks of the hedging model drive the risk reduction to increase in the target mean return m: the gap between the frontiers for the two models widens as m increases, especially towards the frontier's right end. (Recall from Proposition 2(iii) that the no-hedging model's incremental risk explodes as m approaches the newsvendor's maximum profit.) The same pattern also holds in relative terms: the percentage risk reduction increases in m. Hence, the hedging model's advantage over the no-hedging model is more significant when the target mean return is higher. This also demonstrates the hedging model's effectiveness: on top of its first-order advantage (i.e., risk reduction), it also exhibits a second-order advantage (i.e., reduction in incremental risk).

Figure 2 illustrates the contribution of production payoff in (3) to total return for varying target returns. In particular, production payoff accounts for at least 94% of total return in all instances, indicating that optimized operations are the primary source of profit for the automakers. We can also observe that as the target return increases, the contribution from hedging increases (i.e., the contribution from production decreases). This is consistent with how hedging reduces risk. In the no-hedging model, as the return increases, risk increases due to higher price and VPQ; the increment in risk is most prominent when the return approaches the newsvendor's maximum profit (see Proposition 2 (iii)). By contributing more to the return as it increases, hedging suppresses the growth of production payoff, which therefore suppresses the growth of unhedgeable risk to control the total risk. This is reflected by the increased percentage contribution from hedging when the target return is higher.

Figure 2   Production Payoff Contribution with Risk Hedging.
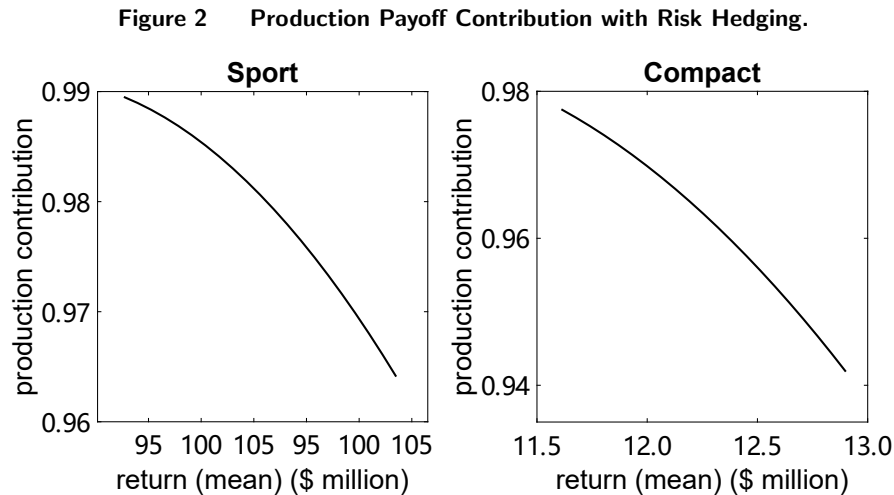
Note. Initial WTI price is $X_0 = 40$.

6.4.2.   Optimal Operational Decisions Figure 3 plots the optimal prices (P), VPQ (R), and production quantities (Q) for both the hedging and no-hedging models (with an initial oil price $X_0 = 40$) over a range of target returns. Theorems 2 and 3 show that if the target return equals the newsvendor's maximum profit in the base model, the optimal price with hedging is lower than without hedging. Despite the absence of analytical results for other values of the target return, Figure 3 shows that over the given range of target returns, the optimal price with hedging never exceeds the optimal price without hedging, and the price markdown is small.

In addition, as return increases, the optimal price and VPQ grow more slowly with hedging than without hedging. In particular, they increase in the target return m at most linearly in the hedging model (dashed curves), in contrast to their faster growths without hedging (solid curves), especially when m is reaching the right ends of the curves. This can be explained as follows. By the (partial)

concavity of the expected production payoff $\mathsf{E}[H_T(P, R)]$ in P and R, as the target mean return m approaches NV's maximum profit $\mathsf{E}[H_T(P^{\text{NV}}, R^{\text{NV}})]$, P and R in the no-hedging model need to increase sharply for the expected production payoff to reach this maximum point. This is consistent with the sharp increases in P and R around the right ends of the solid curves observed in Figure 3. On the other hand, in the hedging model, part of the target mean return m stems from the hedging payoff. Thus, the operational levels P and R can be adjusted downward relative to the no-hedging model because the production payoff can be lower than the target mean return m. In other words, as m increases, the optimal operational levels of the hedging model need not grow as fast as those in the no-hedging model.

The observation above is also consistent with the concavity of the curves in Figure 2. Since the operational levels with hedging increase slowly in m, the percentage contribution from the production decreases at an accelerating pace. Thus, when the firm demands a higher return, the price increment with hedging need not be as high as that without hedging, which helps the manufacturer to stay competitive in the market. Similar patterns also hold for the optimal R and Q.

Figure 3   Optimal Price (P), VPQ (R), and Production Quantity (Q).
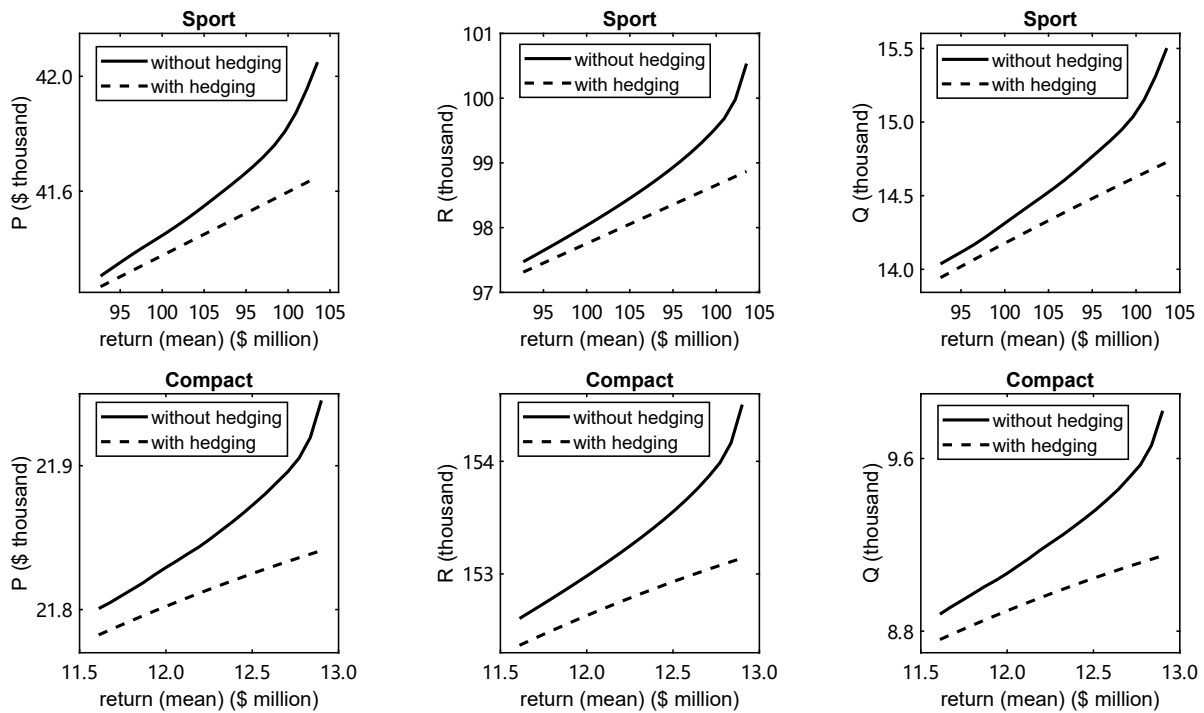
Note. Initial WTI price is $X_0 = 40$.

7.   Concluding Remarks

In this study, we propose and solve a general model that integrates pricing, production, and risk hedging using financial assets. We completely characterize the return-risk efficient frontier. We find that the pricing level is lower with hedging than without hedging when the asset price trend either positively or negatively impacts product demand. The price reduction is desirable for firms that operate in a competitive market. In addition, the service level is lower (which is induced by a lower VPQ) with hedging than without hedging when the asset price trend has a positive effect on product demand. However, when the asset price trend negatively impacts demand, the service level with hedging may exceed the service level without hedging, but only by a small margin. Our numerical study using data sets of `AutoMFR` shows that the hedging model performs substantially better than a price-setting newsvendor without hedging. The markdowns in pricing and service levels are small, which are appealing because hedging does not materially decrease operational profit while substantially reducing risk.

We conclude with a discussion on potential research directions. One direction is to consider the setup where the time horizon $[0, T]$ is a *production period*, and the pricing decision is made at time $t = T$ instead of at $t = 0$. Two problems need to be solved in this setup. One is to maximize payoff by determining the price at time $t = T$ after all the uncertainties are resolved; the other is an mV risk-management problem at time $t = 0$ to determine the production quantity and hedging strategy. This new setup focuses on the impact of hedging on the production decision because it is through this decision that hedging impacts the pricing.

Another potential research direction is to incorporate dynamic hedging into a dynamic pricing setup. As pointed out earlier, the model developed in this paper is suitable for products that do not allow frequent price adjustments. Integrating dynamic hedging with dynamic pricing for risk management is significant for products or services with such price adjustments. Under this setup, one usually models the product demand as a Poisson process with an intensity decreasing in the selling price (Gallego and van Ryzin 1994). One may further allow the intensity to depend on the asset price to incorporate the impact of the financial market. For example, if the intensity is an increasing function in the asset price, then the asset price trend positively impacts product demand. Incorporating asset price into the intensity turns the demand model into a doubly stochastic Poisson process. While the hedging methodology specified in this paper can potentially be applied to this dynamic-pricing setup, the key challenge lies in minimizing the variance function to find the optimal pricing policy, which is now a stochastic process, making the problem analytically intractable. In this regard, a feasible approach is to introduce heuristics, such as approximating the optimal pricing decision by linear functions in the asset price.

Acknowledgments

References

Agrawal V, Seshadri S (2000) Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. *Manuf. Serv. Oper. Manag.* 2(4):410–423.

Allcott H, Wozny N (2014) Gasoline prices, fuel economy, and the energy paradox. *Rev. Econ. Stat.* 96(5):779–795.

Asness C, Frazzini A, Pedersen L (2012) Leverage aversion and risk parity. *Financial Analysts Journal* 68(1):47–59.

Aydin G, Porteus EL (2008) Joint inventory and pricing decisions for an assortment. *Oper. Res.* 56(5):1247–1255.

Busse MR, Knittel CR, Zettelmeyer F (2009) Pain at the pump: How gasoline prices affect automobile purchasing in new and used markets. NBER Working Paper, available at SSRN: `https://ssrn.com/abstract=1525768`.

Busse MR, Knittel CR, Zettelmeyer F (2013) Are consumers myopic? Evidence from new and used car purchases. *Am. Econ. Rev.* 103(1):220–256.

Caldentey R, Haugh M (2006) Optimal control and hedging of operations in the presence of financial markets. *Math. Oper. Res.* 31(2):285–304.

Caldentey R, Haugh MB (2009) Supply contracts with financial hedging. *Oper. Res.* 57(1):47–65.

Chen X, Sim M, Simchi-Levi D, Sun P (2007) Risk aversion in inventory management. *Oper. Res.* 55(5):828–842.

Chen X, Simchi-Levi D (2004a) Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Oper. Res.* 52(6):887–896.

Chen X, Simchi-Levi D (2004b) Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Math. Oper. Res.* 29(3):698–723.

Chen YF, Xu M, Zhang ZG (2009) Technical note—A risk-averse newsvendor model under the CVaR criterion. *Oper. Res.* 57(4):1040–1044.

Cleveland R, Cleveland W, McRae J, Terpenning I (1990) A seasonal-trend decomposition procedure based on loess. *J. Offic. Stat.* 6(1):3–73.

DeYong GD (2020) The price-setting newsvendor: review and extensions. *Int. J. Prod. Res.* 58(6):1776–1804.

Ding Q, Dong L, Kouvelis P (2007) On the integration of production and financial hedging decisions in global markets. *Oper. Res.* 55(3):470–489.

Wang, Yao, and Zhang: *How Does Risk Hedging Impact Operations?*
Article submitted to *Management Science*; manuscript no. MS-FIN-22-00115.R2

35

Federgruen A, Heching A (1999) Combined pricing and inventory control under uncertainty. *Oper. Res.*
47(3):454–475.

Gallego G, van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite
horizons. *Manag. Sci.* 40(8):999–1020.

Gaur V, Seshadri S (2005) Hedging inventory risk through market instruments. *Manuf. Serv. Oper. Manag.*
7(2):103–120.

Gourieroux C, Laurent JP, Pham H (1998) Mean-variance hedging and numéraire. *Math. Finance* 8(3):179–200.

Guiotto P, Roncoroni A (2022) Combined custom hedging: Optimal design, noninsurable exposure, and
operational risk management. *Oper. Res.* 70(1):38–54.

Klier T, Linn J (2010) The price of gasoline and new vehicle fuel economy: Evidence from monthly sales data.
American Economic Journal: Economic Policy 2(3):134–153.

Kouvelis P, Li R (2019) Integrated risk management for newsvendors with value-at-risk constraints. *Manuf.
Serv. Oper. Manag.* 21(4):816–832.

Langer A, Miller NH (2013) Automakers' short-run responses to changing gasoline prices. *Rev. Econ. Stat.*
95(4):1198–1211.

Mankiw NG (2014) *Principles of Economics* (Cengage Learning).

Mann H, Whitney D (1947) On a test of whether one of two random variables is stochastically larger than
the other. *Ann. Math. Stat.* 18(1):50–60.

Markowitz H (1952) Portfolio selection. *J. Finance* 7(1):77–91.

Petruzzi NC, Dada M (1999) Pricing and the newsvendor problem: A review with extensions. *Oper. Res.*
47(2):183–194.

Ross SM (1996) *Stochastic Processes* (John Wiley & Sons).

Schwartz ES (1997) The stochastic behavior of commodity prices: Implications for valuation and hedging. *J.
Finance* 52(3):923–973.

Shaked M, Shanthikumar JG (2007) *Stochastic Orders* (Springer, New York).

Wang L (2021) Mean–variance hedging for production planning with multiple products. *Prod. Oper. Manag.*
30(10):3497–3522.

Wang L, Yao DD (2017) Production with risk hedging—Optimal policy and efficient frontier. *Oper. Res.*
65(4):1095–1113.

Whitin TM (1955) Inventory control and price theory. *Manag. Sci.* 2(1):61–68.

Zhu K, Thonemann UW (2009) Coordination of pricing and inventory control across products. *Naval Res.
Logist.* 56(2):175–190.

E-Companion

Lists of Notation and Assumptions

Table EC.1 Basic Notation.

c	unit production cost
s	unit salvage value
D_t	demand accumulated over time period $[0, t]$ for $0 \le t \le T$
b	positive parameter capturing demand's sensitivity to price
m	target mean return
P	real world probability measure
P^M	risk-neutral probability measure
μ_t	stochastic return rate of the asset price X_t
σ_t	stochastic volatility of the asset price X_t
η_t	market price of risk, $\eta_t = \mu_t / \sigma_t$
A_t	market size accumulated over time period $[0, t]$ for $0 \le t \le T$
A_t^M	version of A_t under P^M, i.e., $\mathsf{P}(A_t^M \le a) = \mathsf{P}^M(A_t \le a)$, $\forall a$.
C_t	component of A_t that is determined by the financial asset only
C_t^M	version of C_t under P^M, i.e., $\mathsf{P}(C_t^M \le x) = \mathsf{P}^M(C_t \le x)$, $\forall x$.
$\tilde{\sigma}$	non-financial noise volatility coefficient of the cumulative demand process
$\tilde{\mu}(\cdot)$	demand rate function

Table EC.2 Notation for Operational Decisions.

P	pricing level
Q	production quantity
R	virtual production quantity (VPQ) , $R := Q + bP$
P^{NV}	optimal P to maximize expected NV profit under P
$P^{\mathrm{NV(M)}}$	optimal P to maximize expected NV profit under P^M
Q^{NV}	optimal Q to maximize expected NV profit under P
$Q^{\mathrm{NV(M)}}$	optimal Q to maximize expected NV profit under P^M
R^{NV}	optimal R to maximize expected NV profit under P
$R^{\mathrm{NV(M)}}$	optimal R to maximize expected NV profit under P^M
P_m^{NV}	optimal P to minimize payoff variance subject to target mean return m in the no-hedging model
Q_m^{NV}	optimal Q to minimize payoff variance subject to target mean return m in the no-hedging model
R_m^{NV}	optimal R to minimize payoff variance subject to target mean return m in the no-hedging model
P_m^h	optimal P to minimize payoff variance subject to target mean return m in the hedging model
Q_m^h	optimal Q to minimize payoff variance subject to target mean return m in the hedging model
R_m^h	optimal R to minimize payoff variance subject to target mean return m in the hedging model

Note. NV stands for newsvendor.

Table EC.3   Assumptions for Main Results

Theoretical Results	Assumptions Involved
Proposition 1	Assumptions 1–2
Proposition 2	Assumptions 1–3
Lemma 1, Theorem 1 and Proposition 3	Assumptions 4–5
Propositions 4–7 and Theorems 2–3	Assumptions 1–5

Table EC.4   List of Definitions.

service level	$\mathsf{P}(D_T \le Q)$, which equals $\mathsf{P}(A_T \le R)$
mV	minimization of variance subject to an exogenously set target mean return
	i.e., $\min \mathsf{Var}$ subject to $\mathsf{E} = m$, where m is exogenously set
MV	mean-variance utility maximization (i.e., $\max \mathsf{E} - \gamma \mathsf{Var}$)
efficient frontier	In mV, let $v(m)$ be the minimum variance. If $v(m)$ increases in m, then $(m, \sqrt{v(m)})$ constitutes an efficient frontier
risk-mitigation payoff (χ_T^{rm})	defined as $\chi_T^{\mathrm{rm}} = \int_0^T (-\xi_t) dt$ in (22)
investment payoff (χ_T^{iv})	defined as $\chi_T^{\mathrm{iv}} = \int_0^T \iota_t dX_t$ in (22)
hedged production payoff (H_T^{h})	defined as $H_T^{\mathrm{h}} = H_T + \chi_T^{\mathrm{rm}}$ in (23)
expected hedged production payoff (V_0)	$V_0 := \mathsf{E}^M(H_T) = \mathsf{E}(H_T^{\mathrm{h}})$
investment risk	defined as $\sqrt{\mathsf{Var}(\chi_T^{\mathrm{iv}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}})}$ in (24)
unhedgeable risk	defined as $\sqrt{\mathsf{Var}(H_T^{\mathrm{h}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}})}$ in (25)

Part I: Supplementary Materials

EC.1.   Equivalents to the Mean-Variance Formulation

The mV optimization problem (7) adopted throughout the paper has the form of

$$\min_x \mathsf{Var}[W(x)] \quad \text{s.t.} \quad \mathsf{E}[W(x)] = m, \tag{EC.1.1}$$

where $W(x)$ represents a random return that depends on a decision x. With the above formulation, a decision-maker aims to minimize the variance of his return while meeting a pre-specified target on the mean return m. The following two alternative formulations are mathematically equivalent to (EC.1.1).

The decision-maker may choose to maximize the mean return with a constraint on the variance:

$$\max_x \mathsf{E}[W(x)] \quad \text{s.t.} \quad \mathsf{Var}[W(x)] \le v. \tag{EC.1.2}$$

It is easy to show that for any m, there exists v such that the optimal solution to problem (EC.1.1) is optimal for problem (EC.1.2), and vice versa. Thus, the efficient frontier constructed from the

mV problem by varying the target mean return m is identical to the efficient frontier constructed from the latter problem by varying the risk tolerance coefficient v.

Another formulation is the mean-variance utility maximization (MV) :

$$\max_{x} \mathsf{E}[W(x)] - \gamma \mathsf{Var}[W(x)], \tag{EC.1.3}$$

where $\gamma \geq 0$ is a given risk-aversion coefficient. It is straightforward to show that any pair of $(\mathsf{E}[W(x^*(\gamma))], \mathsf{Var}[W(x^*(\gamma))])$ induced by the optimal solution $x^*(\gamma)$ of the problem in (EC.1.3), for a given γ, must also be on the efficient frontier generated by the mV model, with $\mathsf{E}[W(x^*(\gamma))]$ being the corresponding target mean return m in the mV formulation. In particular, $x^*(\gamma)$ and the optimal solution of mV model with $m = \mathsf{E}[W(x^*(\gamma))]$ coincide. Therefore, the mV and MV models are mathematically equivalent. In particular, in our setup, with the efficient frontier built from the mV model, the MV model (with a given γ) can be translated to the following problem (recall, $B(m, P, R)$ in (14) is the minimum variance of mV with target mean return set as m):

$$\max_{m} \quad m - \gamma B(m, P_m^h, R_m^h).$$

Accounting for that (P_m^h, R_m^h) minimizes $B(m, P, R)$, the problem above is clearly equivalent to

$$\max_{m, P, R} \quad m - \gamma B(m, P, R).$$

This problem is in general not concave. Moreover, the evaluation of $B(m, P, R)$ can be computationally expensive as it involves nested simulation. So, numerical global optimization procedure is needed. There is an established body of literature on the numerical global optimization with computationally expensive objective function (e.g., Jones et al. (1998)) and such methods are implemented in many software packages. Once the optimal m of the problem above, denoted m_γ, is found, the MV problem reduces to an mV with $m = m_\gamma$.

In practice, however, unless the decision-maker has a reasonable estimate of γ a priori, the mV formulation has a modeling advantage as the parameter m carries a more straightforward economic meaning. In addition, a decision-maker usually maximizes a risk-averse *utility function*, which can be approximated by its first two moments. The problem directly translates to selecting points from the efficient frontier generated by the mV formulation in such a situation (Markowitz 1952, 2014).

EC.2. Risk Decomposition for the No-hedging Model

In parallel with the risk decomposition of the total return in the hedging model as described in Part (ii) of Theorem 1, in particular the decomposition in (24) and (25), here we define analogous decomposition of risk for the no-hedging model. Below, all arguments P and R are dropped as they are given.

By Lemma 1 and definition of V_t in (17), H_T has the following representation:

$$H_T = V_T = V_0 + \int_0^T \xi_t dX_t + \int_0^T \delta_t d\tilde{B}_t.$$

As discussed in Section 4.1, ξ_t captures the impact from the asset price movement on product demand and δ_t is associated with the intrinsic demand noise that can not be reduced by hedging. Therefore, in parallel with the decomposition of total return of the hedging model, $H_T + \chi_T^*$, into hedged production payoff and the investment payoff, here we decompose H_T into a *financial* part, H_T^{fin}, and an *unhedgeable* part, H_T^{u}:

$$H_T^{\mathrm{fin}} := V_0 + \int_0^T \xi_t dX_t \quad \text{and} \quad H_T^{\mathrm{u}} := \int_0^T \delta_t d\tilde{B}_t. \tag{EC.2.1}$$

Then, we decompose the risk of H_T as:

$$\Big[\underbrace{\sqrt{\mathsf{Var}(H_T)}}_{\text{total risk}}\Big]^2 = \mathsf{Var}(H_T^{\mathrm{fin}} + H_T^{\mathrm{u}}) = \Big[\underbrace{\sqrt{\mathsf{Var}(H_T^{\mathrm{fin}}) + \mathsf{Cov}(H_T^{\mathrm{fin}}, H_T^{\mathrm{u}})}}_{\text{financial risk}}\Big]^2 + \Big[\underbrace{\sqrt{\mathsf{Var}(H_T^{\mathrm{u}}) + \mathsf{Cov}(H_T^{\mathrm{fin}}, H_T^{\mathrm{u}})}}_{\text{unhedgeable risk}}\Big]^2.$$

In the numerical implementation of the model in Section 6, ξ_t and δ_t are simulated according to their expressions in (EC.4.2) and (EC.4.1), respectively. Then, the H_T^{fin} and H_T^{u} are simulated based on the generated sample paths of ξ_t and δ_t, with which we compute the variances and the covariance to evaluate the financial and unhedgeable risks of H_T defined above.

EC.3. Sufficient Conditions for Assumptions 2 and 3

Strictly speaking, Assumptions 2 and 3 are not explicitly on primitive model parameters as they involve P^{NV}, R^{NV} and $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$, which are functions of the model primitives. In this section, we derive sufficient conditions explicitly on the model primitives that guarantee Assumptions 2 and 3 to hold. Note that it is sufficient to derive conditions under which $\mathsf{E}[H_T(P, R)] > 0$ for *some* P and R. Clearly, under such condition $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] > 0$ while $P = c$ or $R = bc$ leads to $\mathsf{E}[H_T(P, R)] < 0$, hence $P^{\mathrm{NV}} > c$ and $R^{\mathrm{NV}} > bc$ hold also.

To this end, define $h(P, R) = \mathsf{E}[H_T(P, R)]$ and derive (with $\mu_A := \mathsf{E}(A_T)$ and $\sigma_A^2 = \mathsf{Var}(A_T)$):

$$h(P, R) = (P - c)(R - bP) - (P - s)\mathsf{E}[(R - A_T)^+]$$
$$\geq -bP^2 + bcP + (P - c)R - (P - s)(R - \mu_A)^+ - (P - s)\mathsf{E}[(\mu_A - A_T)^+]$$
$$\geq -bP^2 + bcP + (P - c)R - (P - s)(R - \mu_A)^+ - \frac{1}{2}(P - s)\sigma_A.$$

The first inequality is due to the fact that $(R - A_T)^+ = (R - \mu_A + \mu_A - A_T)^+ \leq (R - \mu_A)^+ + (\mu_A - A_T)^+$; the second inequality follows from the fact that for any random variable X, $\mathsf{E}[(E(X) - X)^+] = (1/2)\mathsf{E}(|X - \mathsf{E}(X)|) \leq (1/2)\sqrt{\mathsf{Var}(X)}$.

Now, choose $R = \mu_A$, we have

$$
\begin{aligned}
\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] &\geq h(\mu_A, P) \geq -bP^2 + bcP + (P-c)\mu_A - \frac{1}{2}(P-s)\sigma_A \\
&= -bP^2 + \left(bc + \mu_A - \frac{1}{2}\sigma_A\right)P - \left(c\mu_A - \frac{1}{2}s\sigma_A\right) =: f(P) \qquad \text{(EC.3.1)}
\end{aligned}
$$

which is a concave quadratic function in P. Denote the discriminant of $f(P)$ by $\Delta(\mu_A)$ (the argument μ_A is to stress its dependence on μ_A):

$$
\Delta(\mu_A) = \left(bc + \mu_A - \frac{1}{2}\sigma_A\right)^2 - 4b\left(c\mu_A - \frac{1}{2}s\sigma_A\right) = \mu_A^2 - (\sigma_A + 2bc)\mu_A + \frac{1}{4}\sigma_A^2 - bc\sigma_A + b^2c^2 + 2bs\sigma_A.
$$

For existence of P such that $f(P) > 0$, it is necessary that $\Delta(\mu_A) > 0$. Note that $\Delta(\mu_A)$ is a convex quadratic function in μ_A with discriminant

$$
\Delta' = (\sigma_A + 2bc)^2 - \sigma_A^2 + 4bc\sigma_A - 4b^2c^2 - 8bs\sigma_A = 8b(c-s)\sigma_A > 0.
$$

Then,

$$
\mu_A > \frac{\sigma_A + 2bc + \sqrt{8b(c-s)\sigma_A}}{2} \Rightarrow \Delta(\mu_A) > 0. \qquad \text{(EC.3.2)}
$$

Then, for $f(P) > 0$ (which implies $g(P) > 0$ hence $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$), it is sufficient that

$$
P \in \left(c \vee \frac{bc + \mu_A - \frac{1}{2}\sigma_A - \sqrt{\Delta(\mu_A)}}{2b}, \frac{bc + \mu_A - \frac{1}{2}\sigma_A + \sqrt{\Delta(\mu_A)}}{2b}\right).
$$

Note that the interval above exists due to:

$$
\begin{aligned}
\frac{bc + \mu_A - \frac{1}{2}\sigma_A + \sqrt{\Delta(\mu_A)}}{2b} - c &= \frac{1}{2b}\left(bc + \mu_A - \frac{1}{2}\sigma_A + \sqrt{\Delta(\mu_A)} - 2bc\right) \\
&= \frac{1}{2b}\left(\mu_A - \frac{1}{2}\sigma_A - bc + \sqrt{\Delta(\mu_A)}\right) \\
&> \frac{1}{2b}\left(\sqrt{2b(c-s)\sigma_A} + \sqrt{\Delta(\mu_A)}\right) > 0,
\end{aligned}
$$

where the first inequality is due to the condition in the left hand side of (EC.3.2). In summary, under the condition on the left hand side of (EC.3.2), there exists P and R such that $\mathsf{E}[H_T(P, R)] > 0$ and thus Assumption 2 holds.

Now we proceed to deriving sufficient conditions for Assumption 3. To start with, note that

$$
\mathsf{E}[(bc - A_T)^+] = \mathsf{E}[(bc - A_T^+)^+] + \mathsf{E}[(-A_T)^+] \leq bc\mathsf{P}(A_T \leq bc) + \mathsf{E}[(-A_T)^+].
$$

Let $p_0 := \mathsf{P}(A_T \leq bc)$ and $\epsilon_0 := \mathsf{E}[(-A_T)^+]$, then the following conditions are clearly sufficient for Assumption 3:

$$
bcp_0 + \epsilon_0 \leq 2b(c-s) \quad \text{and} \quad bcp_0 + \epsilon_0 \leq 2\sqrt{b \max_P f(P)},
$$

where $f(P)$ is defined in (EC.3.1). Since $f(P)$ is concave quadratic function, it is straightforward to derive its maximum value:

$$
\begin{aligned}
f^* := \max_P f(P) = f(P^*) &= \frac{1}{4b}\left[\left(\mu_A - bc - \frac{1}{2}\sigma_A\right)^2 - 2b(c-s)\sigma_A\right] > 0, \\
P^* = \frac{\mu_A - \frac{1}{2}\sigma_A + bc}{2b} &> c.
\end{aligned}
\tag{EC.3.3}
$$

In (EC.3.3) above, both $f^* > 0$ and $P^* > c$ follow immediately from the condition on the left hand side of (EC.3.2). Clearly, $\mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] > f^*$.

Summarizing the analysis above, we have the following result.

PROPOSITION EC.1. *Recall, $\mu_A := \mathsf{E}(A_T)$ and $\sigma_A^2 := \mathsf{Var}(A_T)$; in addition, $p_0 := \mathsf{P}(A_T \le bc)$ and $\epsilon_0 := \mathsf{E}[(-A_T)^+]$. Then, Assumptions 2 and 3 are met under the following conditions:*

$$
\mu_A > \frac{\sigma_A + 2bc + \sqrt{8b(c-s)\sigma_A}}{2} \quad \text{and} \quad bcp_0 + \epsilon_0 \le 2b(c-s) \wedge 2\sqrt{bf^*},
\tag{EC.3.4}
$$

where the expression of f^* is in (EC.3.3).

EC.4. Supplementary Materials for Section 6
EC.4.1. Expressions of $\delta_t(P,R)$ and $\xi_t(P,R)$

Here we give explicit expressions of $\delta_t(P,R)$ and $\xi_t(P,R)$ in Section 4.2; recall, these two quantities are involved in the martingale representation of V_t in (18). With the dynamics of X_t specified in (29) and the demand model specified in (37), we can obtain more explicit expressions for δ_t and ξ_t as follows. Define $f(t,x,a) := \mathsf{E}^M[(R - A_T)^+ \mid X_t = x, A_t = a]$. Note that since in (29), $\sigma_t = \sigma$ is a constant, we can remove σ_t from the arguments of the function $f(\cdot)$. Then,

$$
\begin{aligned}
f(t,x,a) &= \mathsf{E}^M\left[\left(R - a - \int_t^T \tilde{\mu}(xe^{\sigma(B_u^M - B_t^M) - \frac{1}{2}\sigma^2(u-t)}) - \tilde{\sigma}(\tilde{B}_T - \tilde{B}_t)\right)^+ \,\Big|\, X_t = x, A_t = a\right] \\
&= \mathsf{E}\left[\left(R - a - \int_t^T \tilde{\mu}(xe^{\sigma(B_u - B_t) - \frac{1}{2}\sigma^2(u-t)}) - \tilde{\sigma}(\tilde{B}_T - \tilde{B}_t)\right)^+\right],
\end{aligned}
$$

where the second equality uses the facts that B_t^M is a Brownian motion under P^M, the independence between $\{B_t\}$ and $\{\tilde{B}_t\}$, and that both $\{B_u - B_t, t \le u \le T\}$ and $\tilde{B}_T - \tilde{B}_t$ are independent of $\mathcal{F}_t$. Using the fact that $\tilde{B}_T - \tilde{B}_t$ follows Normal distribution with zero mean and $\sqrt{T-t}$ as standard deviation, taking derivative with respect to a leads to

$$
f_a(t,x,a) = -\mathsf{E}\left[\Phi\left(\frac{R - a - \int_t^T \tilde{\mu}(xe^{\sigma(B_u - B_t) - \frac{1}{2}\sigma^2(u-t)})du}{\tilde{\sigma}\sqrt{T-t}}\right)\right],
$$

where $\Phi(x)$ is the distribution function of standard normal random variable. Furthermore, note that the dX_t-term in (18) is produced by f_x. Taking derivative with respect to x leads to

$$
f_x(t,x,a) = -\mathsf{E}\left[\Phi\left(\frac{R - a - \int_t^T \tilde{\mu}(xe^{\sigma(B_u - B_t) - \frac{1}{2}\sigma^2(u-t)})du}{\tilde{\sigma}\sqrt{T-t}}\right)\right.
$$

$$\cdot \int_t^T \tilde{\mu}'(xe^{\sigma(B_u - B_t) - \frac{1}{2}\sigma^2(u-t)})e^{\sigma(B_u - B_t) - \frac{1}{2}\sigma^2(u-t)}du\Big].$$

Therefore, by Lemma 1,

$$\delta_t = -\tilde{\sigma}(P-s)f_a(t, X_t, A_t) = \tilde{\sigma}(P-s)\mathsf{E}\Big[\Phi\Big(\frac{R - A_t - \int_t^T \tilde{\mu}(X_t X_{u-t}^M)du}{\tilde{\sigma}\sqrt{T-t}}\Big)\Big|X_t, A_t\Big], \qquad \text{(EC.4.1)}$$

and

$$\xi_t = -(P-s)f_x(t, X_t, A_t)$$
$$= (P-s)\mathsf{E}\Big[\Phi\Big(\frac{R - a - \int_t^T \tilde{\mu}(X_t X_{u-t}^M)du}{\tilde{\sigma}\sqrt{T-t}}\Big)\int_t^T \tilde{\mu}'(X_t X_{u-t}^M)X_{u-t}^M du\Big|X_t, A_t\Big]. \qquad \text{(EC.4.2)}$$

Here, $X_{u-t}^M := e^{\sigma(B_u - B_t) - \frac{1}{2}\sigma^2(u-t)}$.

EC.4.2. Parameter Estimation

The summary statistics of the data set is presented in Table EC.5.

Table EC.5   Summary Statistics.

Car	Sample Size	Price ($)	Sales Volume	MPG
Sport	108	40,615 (1,239)	15,933 (3,516)	16.89 (1.96)
Compact	101	22,300 (444)	16,190 (3,378)	24.92 (3.34)

Note. The numbers in parentheses are standard deviations.

EC.4.2.1.   WTI Price To estimate the parameters of the EOU process in (29), we calibrate the model to the log WTI price data by applying Ordinary Least Squares (OLS) method to the discretized equation produced by (29). Recall, $Y_t = \log X_t$ is modeled in (29). To calibrate this model to WTI prices, we start with discretizing Y_t:

$$Y_{t+\nu} = (1 - \kappa\nu)Y_t + \kappa\alpha\nu + \sigma\sqrt{\nu}\cdot\epsilon_t, \qquad \text{(EC.4.3)}$$

where $\nu = 1/252$ is the time step size (i.e., one trading day), and ϵ_t is a standard normal random variable independent of Y_t. Then, (EC.4.3) leads to the following linear regression model (n stands for the n-th day in the data set):

$$Y_{n+1} = aY_n + b + \epsilon_n,$$

where $a = 1 - \kappa\nu$, $b = \kappa\alpha\nu$, and ϵ_n's are i.i.d zero-mean random noises with variance $\sigma^2\nu$. Applying OLS produces estimators of a, b, and standard error of residuals, which are denoted by $\hat{a}$, $\hat{b}$, and

$\hat{\sigma}^2$, respectively. Then, the estimators (denoted by corresponding symbols with hats) for κ, α, and σ^2 are

$$\hat{\kappa} = \frac{1-\hat{a}}{\nu}, \quad \hat{\alpha} = \frac{\hat{b}}{1-\hat{a}}, \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{N\nu} \sum_{n=1}^{N} (Y_{n+1} - (\hat{a}Y_n + \hat{b}))^2,$$

where $N = 2600$ is the number of observations.

Here we remark that the regression model above is essentially an autoregressive model (with lag 1). Although for such models, OLS estimators in general are *not* unbiased due to the past errors influencing future values of the process, they are *consistent* estimators; see Hamilton (1994). Considering the large sample size ($N = 2600$), we deem that the OLS estimators are appropriate under the current context.

The estimated parameters are as follows:

$$\hat{\kappa} = 0.5356, \quad \hat{\alpha} = 4.1847, \quad \text{and} \quad \hat{\sigma} = 0.3327.$$

The estimated parameters suggests that the long-run average of WTI is

$$\mathsf{E}[X_\infty] = e^{\alpha + \frac{\sigma^2}{2\kappa}} = 72.82.$$

EC.4.2.2.   Demand Model Calibration Recall, market size is assumed to take the form shown in (37). To evaluate the integral involved in (37) in each period (month), we apply the discretized numerical integration:

$$C_T = \int_0^T \tilde{\mu}(X_t) dt \approx T(\mu_0 + \mu_1 \bar{X}) \equiv A + B\bar{X}, \tag{EC.4.4}$$

where $\bar{X}$ is the average asset price within one month: $\bar{X} = \left(\sum_{j=0}^{J} X_j \right)/(J+1)$, and $J = 20$—that is, there are 21 trading days in a month. Thus, $\mu_0 = A/T$, and $\mu_1 = B/T$, with $T = 1/12$.

Substituting (EC.4.4) to (37), we have the following linear regression model:

$$D = A + B\bar{X} - bP + \tilde{\sigma}\sqrt{T}\epsilon, \tag{EC.4.5}$$

where ϵ is a standard normal error that is independent of $\bar{X}$. The pricing level P in (EC.4.5) and the production quantity Q are determined by the manufacturer, which we assume to be accord with the newsvendor's profit-maximization problem (recall, the salvage value s is zero):

$$(P^*, Q^*) = \arg\max_{P,Q} \mathsf{E}[(P-c)Q - P(Q-D)^+]. \tag{EC.4.6}$$

The set of model parameters to be calibrated is $\{A, B, b, c, \tilde{\sigma}\}$. Given the oil price X_0 on the first trading day of month i, the model-implied price and production quantity for the same month are

$$(P_i^*, Q_i^*) = \arg\max_{P,Q} \mathsf{E}[(P-c)Q - P(Q-D_i)^+],$$

where D_i is simulated given $\{A, B, b, c, \tilde{\sigma}\}$ and X_0. In principle, we need to minimize the distance between the model-implied P and Q and those observed in data. While P is observable, Q is not available in our data set. Thus, instead of fitting Q, we turn to minimizing the error of the demand model in (EC.4.5). Here, we use sales volume to proxy demand. This is appropriate because the inventory-to-sales ratio in the automotive industry in the U.S. market has been around 2.5, which is substantially above 1, for decades (Dunn and Vine (2016)), suggesting that the inventory held has been able to cover the demand. Therefore, we minimize the distance between the model-implied expected demand $A + B\bar{X}_i - bP_i^*$ and observed sales volume S_i. Thus, our calibration model is

$$\min_{A, B, b, c, \tilde{\sigma}} \sum_{i=1}^{n} [(P_i^* - P_i)^2 + (A + B\bar{X}_i - bP_i^* - S_i)^2].$$

Then, we apply off-the-shelf global optimization algorithm to solve the problem above to obtain the solutions.

The calibrated parameters are summarized in Table EC.6. The cost c is the average direct production cost that includes raw material and labor/assembly costs.

Table EC.6   Calibrated Model Parameters.

Car	$A \,(= \mu_0 T)$	$B \,(= \mu_1 T)$	b	c (cost)	$\tilde{\sigma}$
Sport	111,155.66	-185.42	2.02	34,543.91	11,577.37
Compact	151,887.67	157.41	6.59	20,467.10	8,619.46

Note. $D_T = \int_0^T (\mu_0 + \mu_1 X_t) dt + \tilde{\sigma} \tilde{B}_T - bP$, and $T = 1/12$.

Note that the sign of the parameter B determines how asset price impacts demand. Specifically, the plus (resp., minus) sign of B represents the positive (resp., negative) impact of oil price on product demand. That is, the demand of Sport (resp., Compact), the fuel-inefficient (resp., fuel-efficient) model, is negatively (resp., positively) impacted by oil price. This is consistent with the economic intuition and empirical evidence discussed in Section 1. Another point worth noting is that the estimated production costs (c) for both cars are around 85% of the average selling price, which matches the profit margin observed in the automotive industry (Margins by Sector, Auto & Truck, `https://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/margin.html`). With the parameters in Table EC.6, we simulate samples of C_T and C_T^M and apply Mann–Whitney U-test (Mann and Whitney (1947)) to test the stochastic dominance relationships between these two quantities, respectively for the positive ($C_T \succeq C_T^M$) and negative ($C_T \preceq C_T^M$) cases. The test results support such relationships for all model instances; see Section EC.4.4 for details.

EC.4.3. Numerical Procedure

We apply Monte Carlo simulation to evaluate the expectations involved in the variance function $B(m, P, R)$. To simulate sample paths, we set the discretized time step size at $\nu = 1/252$ and assume that each month has 21 trading days. To evaluate the first term of $B(m, P, R)$, in particular V_0, Monte Carlo simulation is readily applied with simulated values of D_T, which is part of the simulated paths (X_t, D_t) (the simulation procedure is described below). Evaluating the second term of $B(m, P, R)$ in (33) involves the following steps:

(i) Use Monte Carlo simulation to generate N_1 sample paths of (Y_t, X_t, A_t) according to (29), (10), and (EC.4.4).

(ii) At each time point $t \leq T$, given each realized value of $(Y_t = y, X_t = x, A_t = a)$, generate N_2 paths of X_u^M for $u \in [t, T]$ with initial value $X_0^M = 1$ and evaluate the integral $\int_t^T [\mu_0 + \mu_1 \cdot (x X_{u-t}^M)] du$ along each path of $X_u^M := e^{\sigma(B_u - B_t) - \frac{1}{2}\sigma^2(u-t)}$; the integral is evaluated by trapezoidal rule.

(iii) Evaluate $\delta_t(P, R)$ in (EC.4.1) via simulation using the paths generated in step (ii) for each path $(X_t = x, A_t = a, Y_t = y)$ generated in step (i).

(iv) Evaluate the function values in (34) and use $\delta_t(P, R)$ to evaluate, for $t \in [0, T]$,

$$\mathsf{E}[e^{-f_0(T-t)-f_1(T-t)Y_t - f_2(T-t)Y_t^2}\delta_t^2(P, R)].$$

(v) Evaluate the second term in $B(m, P, R)$ via the trapezoidal rule.

To minimize $B(m, P, R)$ over (P, R), we follow two steps. First, given R, we perform a line search for the corresponding optimal P over $[c, P^{\mathrm{NV(M)}}]$. This is fairly efficient, as $B(m, P, R)$ is convex in P given m and R (see Proposition EC.2). Then, we perform a line search for the optimal R over $[bc, R^{\mathrm{NV(M)}}]$.

PROPOSITION EC.2. *For a given R, $\bar{P}(R)$ is the smallest root of $V_0(P, R) = m$ if $m \leq \max_P V_0(P, R)$; otherwise, let $\bar{P}(R) = \arg\max_P V_0(P, R)$. Then, $\arg\min_P B(m, P, R) \leq \bar{P}(R)$ and $B(m, P, R)$ is convex in P over $[c, \bar{P}(R)]$.*

EC.4.4. Stochastic Dominance Tests for C_T and C_T^M

In this section, we test the stochastic dominance relationship between C_T and C_T^M by applying the classical Mann–Whitney U-test (Mann and Whitney (1947)) on samples of C_T and C_T^M simulated with calibrated parameters (see Section EC.4.2.1, Section EC.4.2.2 and Table EC.6). As the testing procedures are similar, in the following, we detail the test results for calibrated parameters and briefly report the conclusion for the modified parameters.

For the positive (resp., negative) case, we want to test

$$H_0 : C_T^M \stackrel{d}{=} C_T \quad (\text{resp.,} \ C_T \stackrel{d}{=} C_T^M) \quad \text{versus} \quad H_1 : C_T^M \preceq C_T \quad (\text{resp.,} \ C_T \preceq C_T^M).$$

With the initial oil price at the levels of $X_0 = 40$ and 100, and calibrated parameters of WTI price in Section EC.4.2.1, and those of demand model in Table EC.6, for each problem instance, we simulate 100000 realizations for C_T and C_T^M independently. The Mann–Whitney U test (of the null hypothesis of equal distributions against appropriate alternatives) is implemented in many software packages and we utilize the Python package (i.e., scipy.stats.mannwhitneyu) to conduct the test. The p-values are summarized in Table EC.7.

Table EC.7   P-values for the Mann–Whitney U test.

Car	$X_0 = 40$	$X_0 = 100$
Sport	9.70×10^{-4}	1.72×10^{-153}
Compact	0	9.56×10^{-173}

In this table, two cases of asset price trends positively impacting the demands are reported: Compact ($X_0 = 40$) and Sport ($X_0 = 100$). For these two cases, all the p-values are extremely small, which strongly reject the null hypothesises in favor of the alternative ones (i.e., $C_T \succeq C_T^M$). The other two are the cases of asset price's negative effects on the demands: Sport ($X_0 = 40$) and Compact ($X_0 = 100$). The p-values of both tests also provide strong evidences to reject the null hypothesises and support the alternative ones, $C_T^M \succeq C_T$. In the Section EC.4.5, hypothetical model parameters are used. We conduct the same tests for all problem instances. We find that all the p-values are extremely small, so conclude that the tests reject the null hypothesis and support the alternative one, $C_T^M \succeq C_T$ (note that all problem instances in this part are for the case of asset price trend negatively impacting demand).

EC.4.5.   Numerical Illustration of Theorem 3

Theorem 3 characterizes how risk hedging affects the optimal price and VPQ when the asset price trend negatively impacts demand. Part (i) of the theorem shows that hedging induces a price markdown. However, as shown in Parts (ii) and (iii), the effect of hedging on the optimal VPQ is more involved, depending on whether the asset price trend's negative effect on product demand is sufficiently strong. We now illustrate Parts (ii) and (iii) using the same demand and asset models as in Section 6, but with different parameter values.[14]

---

[14] To make the asset price trend more prominent to strengthen the negative effect, we amplify the value of the mean-reverting coefficient κ in (29) by a factor of 10. To control for the magnitude of MPR process (30), we also increase the volatility σ by a factor of 10. In addition, we change the parameters in equation (37) to make the demand more sensitive to the asset price while keeping $\mathsf{E}(A_T)$ at the same level as in the calibrated model in Section 6.

We first define the following metric to measure the negative effect of the asset price trend on the demand, which we call "hurting factor", because "negatively impacting demand" can be expressed informally as "hurting demand":

$$\Delta := c - s - \left[\frac{\bar{P}^{\mathrm{NV}}}{r^\circ + \sqrt{(r^\circ)^2 - 1}} \right] \cdot \mathsf{P}^M(A_T \geq R^{\mathrm{NV}}).$$

Note that Δ is the difference between the right- and left-hand sides of the inequality (36). Recall from the discussion following Theorem 3 that $\Delta \geq 0$ indicates that the negative effect is sufficiently strong. In this case (Part (ii)), the optimal VPQ with hedging is not higher than that without hedging, that is, $R_m^h \leq R^{\mathrm{NV}}$. Otherwise (Part (iii)), R_m^h may exceed R^{NV}, but by only a moderate amount at most. In the following numerical illustration, we first show Δ indeed captures the negative effect and then examine how R°, R^{NV}, and $R^{\mathrm{NV(M)}}$, quantities involved in Part (iii), respond as Δ varies.

In the EOU model (29), a simple way to vary the magnitude of the asset price's negative effect is to vary the parameter α. For `Sport`, fixing the initial oil price $X_0 = 30$, we increase α to increase the long-run average of oil prices, strengthening the negative effect of the oil price trend on product demand. On the other hand, for `Compact`, we fix $X_0 = 250$ to introduce a prominent downward oil price trend and then decrease α to strengthen the negative effect.

Figure EC.1 illustrates Part (ii) of Theorem 3. We first interpret the first column. The upward-sloping WTI price trend, reflected by $E(X_\infty - X_0)$, becomes more prominent as α increases, boosting the negative effect of the WTI price trend on the demand for `Sport`. In the meantime, the hurting factor Δ increases as shown in the first graph of the first column. The second graph of the first column shows the same pattern, although for `Compact`, a stronger negative effect is reflected by a larger value of $E(X_0 - X_\infty)$. Therefore, we can indeed use Δ to measure the magnitude of the asset price trend's negative effect on product demand. The second column of the graphs shows that the difference in expectations between A_T^M (market size under risk-neutral measure) and A_T (the actual market size) increases as Δ increases. This is because the strengthening asset price trend reduces A_T but does not affect A_T^M. The third column shows that r°, defined in (35), increases with the negative effect for both car models, which aligns with the discussion following Theorem 3.

Figure EC.2 illustrates Part (iii) of Theorem 3. When the negative effect is not strong enough (i.e., $\Delta < 0$), R° is close to R^{NV} relative to its distance to $R^{\mathrm{NV(M)}}$, confirming that in this case R_m^h can possibly exceed R^{NV}, but only by a small margin (recall that $R_m^h \leq R^\circ$).

EC.4.6.   Risk Decomposition

Recall, by Part (iii) of Theorem 1, the squared total risk (i.e., variance) is the sum of squared investment and unhedgeable risks. Figure EC.3 shows that the squared unhedgeable risk (see (25))

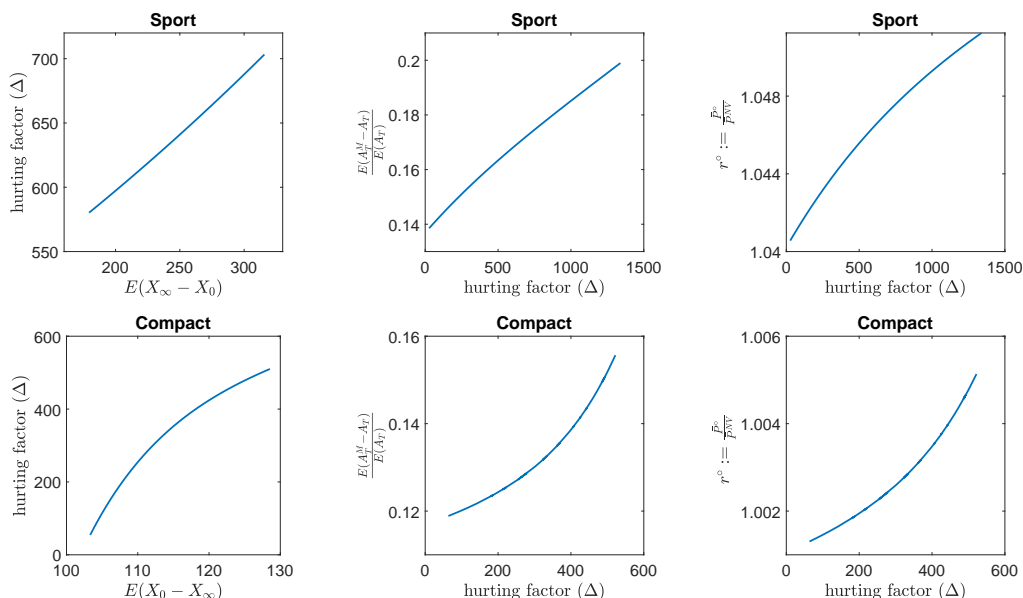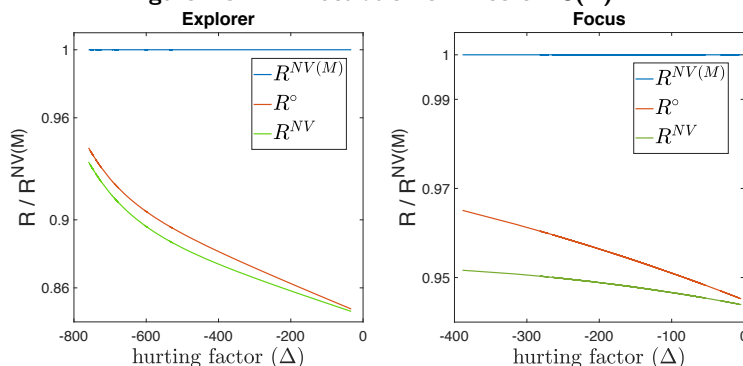**Figure EC.1   Illustration of Theorem 3(ii).**

Figure EC.2   Illustration of Theorem 3(iii).

contributes to most of the squared total risk $B(m, P, R)$ for all instances—at least 77% for `Sport` and 64% for `Compact`—indicating that the main risk factor (after being hedged) is intrinsic demand volatility. In other words, the investment risk (see (21) and (24)) constitutes only a moderate part of the total risk. This trait is desirable: a manufacturer, which is non-financial in nature, does not want to bear too much risk from financial investment. The risk decomposition for the instances of the other two starting oil prices ($X_0 = 70, 100$) exhibits analogous pattern as that in Figure EC.3.

EC.5.   General Operational Payoff Functions

In this section, we extend the methodology developed so far to incorporate general operational payoff functions beyond the price-setting newsvendor. We solve the corresponding hedging problem and identify crucial structural properties of the payoff function under which risk hedging reduces optimal operational levels when the asset price trend positively impacts the stochastic operational

Figure EC.3   Percentage Contribution of Squared Unhedgeable Risk to Total Variance.
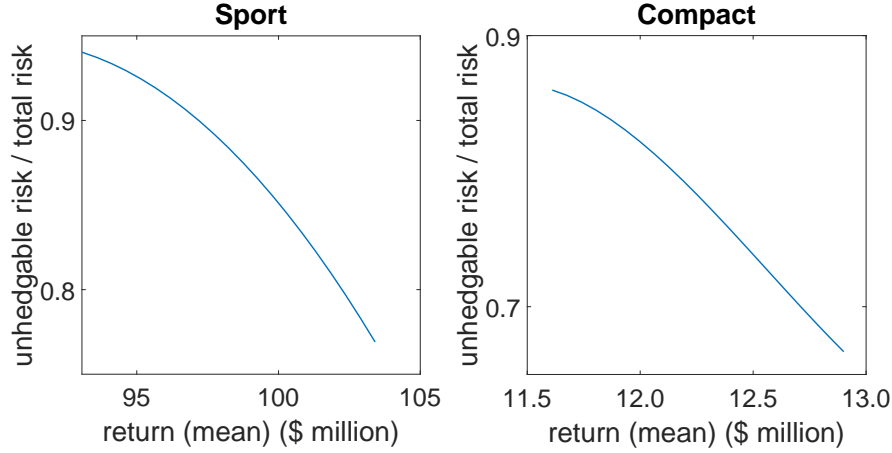
Note. Initial WTI price is $X_0 = 40$.

factors. We apply the general result to two examples—(i) price-setting newsvendor with price-based substitutable products and (ii) price-advertising optimization.

Consider a non-financial firm operating over a time horizon $[0, T]$. At time 0, the firm needs to make an operational decision $\mathbf{q}$, an n-dimensional vector. Without loss of generality, we assume $\mathbf{q} \geq \mathbf{0}$; that is, $q_i \geq 0$ for each $i = 1, \ldots, n$. At time T, the firm receives an operational payoff $H_T(\mathbf{q}, \mathbf{A}_T)$, where $\mathbf{A}_T = (A_{jT})_{j=1}^{T}$ is a J-dimensional random vector, representing stochastic operational factors that may affect the payoff, such as demands, market sizes, costs, and exchange rates. Suppose these factors partially depend on some financial asset price and are modeled as

$$\mathbf{A}_T = \mathbf{C}_T + \tilde{\boldsymbol{\Sigma}}\tilde{\mathbf{B}}_T \tag{EC.5.1}$$

with the associated dynamics

$$d\mathbf{A}_t = d\mathbf{C}_t + \tilde{\boldsymbol{\Sigma}}d\tilde{\mathbf{B}}_t, \tag{EC.5.2}$$

where $\tilde{\mathbf{B}}_t$ is a K-dimensional standard Brownian motion, independent of B_t that drives the financial asset price X_t in (8). In equation (EC.5.1), $\tilde{\boldsymbol{\Sigma}}\tilde{\mathbf{B}}_T$ represents operational noise independent of the financial market (particularly $\{X_t, 0 \leq t \leq T\}$), where $\tilde{\boldsymbol{\Sigma}} = [\tilde{\sigma}_{jk}]_{j,k=1,1}^{J,K}$ is a J-by-K matrix with $\tilde{\boldsymbol{\sigma}}_j = (\tilde{\sigma}_{jk})_{k=1}^{K}$ denoting the j-th row. In equation (EC.5.2), $\mathbf{C}_t = (C_{jt})_{j=1}^{J}$ is a K-dimensional stochastic process, representing the operational components determined by X_t. That is, each component C_{jt} is adapted to the filtration generated by B_t and we assume that $\mathbf{C}_v - \mathbf{C}_u$ is adapted to the sigma-algebra generated by $\{X_t, u \leq t \leq v\}$.

A standard operations management problem is $\max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_T)]$. However, maximizing the expected payoff without risk management may lead to highly risky decisions. To develop a risk-management framework for the general setup, we extend the mV optimization problem (13) to

$$\min_{\mathbf{q} \geq \mathbf{0}, \vartheta \in \mathcal{A}_X} \mathsf{Var}[H_T(\mathbf{q}, \mathbf{A}_T) + \chi_T(\vartheta)] \quad s.t. \quad \chi_t(\vartheta) := \int_0^t \theta_s dX_s, \quad \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_T) + \chi_T(\vartheta)] = m. \tag{EC.5.3}$$

We first fix $\mathbf{q}$ and consider the hedging problem (with $\mathbf{q}$ dropped from the arguments of H_T):

$$B(m, \mathbf{q}) := \min_{\vartheta \in \mathcal{A}_X} \mathsf{Var}[H_T + \chi_T(\vartheta)] \quad s.t. \quad \chi_t(\vartheta) := \int_0^t \theta_s dX_s, \quad \mathsf{E}[H_T + \chi_T(\vartheta)] = m. \tag{EC.5.4}$$

The solution is characterized by Theorem EC.2 in Section EC.10 of the e-companion, which extends Theorem 1. In Section EC.5.1, we minimize the variance function $B(m, \mathbf{q})$ over $\mathbf{q}$ to find the optimal operational decision in the presence of hedging.

EC.5.1. Optimal Operational Decision in the Presence of Hedging

Theorem 2 shows that for a price-setting newsvendor, hedging reduces both the optimal price and VPQ when asset price trend *positively* impacts demand. We extend this result to general operational payoff functions under two technical assumptions, one concerning the payoff function's structure while the other the relationship between the unhedgeable risk and the operational decision.

ASSUMPTION EC.1. *The operational payoff function $H_T(\mathbf{q}, \mathbf{a})$ satisfies the following conditions.*
(i) *For any $\mathbf{a}$, $H_T(\mathbf{q}, \mathbf{a})$ is concave in q_i for all $i = 1, \ldots, n$.*
(ii) *For any $\mathbf{q}$, $H_T(\mathbf{q}, \mathbf{a})$ is increasing in a_j for all $j = 1, \ldots, J$.*
(iii) *For any given $\mathbf{a}$, $H_T(\mathbf{q}, \mathbf{a})$ is supermodular in $\mathbf{q}$.*
(iv) *$H_T(\mathbf{q}, \mathbf{a})$ has increasing differences in $(\mathbf{q}, \mathbf{a})$; that is, $H_T(\mathbf{q}_1, \mathbf{a}_2) - H_T(\mathbf{q}_1, \mathbf{a}_1) \leq H_T(\mathbf{q}_2, \mathbf{a}_2) - H_T(\mathbf{q}_2, \mathbf{a}_1)$ for all $\mathbf{q}_1 \leq \mathbf{q}_2$ and $\mathbf{a}_1 \leq \mathbf{a}_2$.*

In Part (i) of Assumption EC.1, the concavity of the payoff function in each decision variable reflects the "too-little-or-too-much" trade-off. For instance, the newsvendor payoff function is concave in the production quantity, reflecting the underage or overage. In Part (ii), the monotonicity of the payoff function in the stochastic operational factors is common for many practical applications. In addition to the price-setting newsvendor example for which the payoff increases with the market size, a manufacturer's payoff may increase with the exchange rate or decrease with the (random) cost of raw materials. (In the latter case, we can replace a_i with $-a_i$ to recover the increasing property.) In Part (iii), the supermodularity property means that the operational variables are complementary to each other. Examples include the production quantities of complementary goods. In Part (iv), the increasing-difference property indicates that the marginal gain from increasing the operational levels is higher for higher levels of the stochastic operational factors. For example, marginal profit from increasing production is higher with a larger market size.

Under Assumption EC.1, we can define the following standard profit-maximizing decision:

$$\mathbf{q}^* := \min \left\{ \arg\max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_T)] \right\}, \qquad (\text{EC.5.5})$$

where the minimization takes effect entry-wise if the inner maximization problem has multiple optimal solutions. The supermodularity property of $H_T(\mathbf{q}, \mathbf{a})$ ensures that the optimal solutions of the inner maximization problem form a sublattice, so there exists a smallest optimal solution and $\mathbf{q}^*$ is well defined (Topkis 1978). We select the smallest optimal solution based on the economic intuition that a firm prefers a lower operational level to reach the same profit level. For example, a manufacturer prefers to hold less inventory to reach the same level of production payoff.

The following assumption generalizes the relationship between the unhedgeable risk and the price/VPQ in the price-setting newsvendor model. In this model, the unhedgeable risk is defined in equation (25) and increases in both P and R. This definition naturally extends to the general setting, in which the unhedgeable risk is a function of the operational decision $\mathbf{q}$ (see Theorem EC.2 in Section EC.10 of the e-companion).

ASSUMPTION EC.2. *The unhedgeable risk increases in each component of* $\mathbf{q}$.

This assumption is based on the intuition that a higher operational level induces more exposure to the uncertainties associated with the operational noises, leading to higher unhedgeable risk.

THEOREM EC.1. *Suppose Assumptions EC.1–EC.2 hold and the asset price trend positively impacts the stochastic operational factors (i.e., $\mathbf{C}_T^M \preceq \mathbf{C}_T$, or equivalently, $\mathsf{P}(\mathbf{C}_T^M \in U) \leq \mathsf{P}(\mathbf{C}_T \in U)$ for any upper set $U \subseteq \mathbb{R}^J$). Let $m = \mathsf{E}[H_T(\mathbf{q}^*, \mathbf{A}_T)]$ and $\mathcal{S}_m^h = \arg\min_{\mathbf{q} \geq \mathbf{0}} B(m, \mathbf{q})$, where $\mathbf{q}^*$ is defined in (EC.5.5). Then, $\mathbf{q} \leq \mathbf{q}^*$ for all $\mathbf{q} \in \mathcal{S}_m^h$.*

REMARK EC.1. While Theorem 2 extends to general operational payoff functions, it is not the case for Theorem 3 for the following reason. When the asset price trend has a positive effect (the setting of Theorems 2 and EC.1), hedging reduces the levels of the operational factors, thereby decreasing the optimal operational decision due to the increasing-difference property in Assumption EC.1. Technically, a crucial step in proving Theorem EC.1 is to show that if $\mathbf{C}_T^M \preceq \mathbf{C}_T$, then the mV-optimal operational decision with hedging is lower than $\mathbf{q}^{*M}$ and $\mathbf{q}^{*M} \leq \mathbf{q}^*$, where $\mathbf{q}^{*M}$ is the counterpart of $\mathbf{q}^*$ under the risk-neutral measure. This is analogous to Proposition 4 (see the discussion therein). However, when the asset price trend has a negative effect (the setting of Theorem 3), although the mV-optimal operational decision with hedging is still lower than $\mathbf{q}^{*M}$, the relationship between $\mathbf{q}^{*M}$ and $\mathbf{q}^*$ is reversed (i.e., $\mathbf{q}^{*M} \geq \mathbf{q}^*$ if $\mathbf{C}_T^M \succeq \mathbf{C}_T$). Thus, the relationship between $\mathbf{q}^*$ and the mV-optimal operational decision with hedging is generally unclear. This relationship becomes clear in Theorem 3 (i.e., the price-markdown result for a price-setting newsvendor when the asset price trend negatively impacts demand) because we can exploit specific structures of the newsvendor payoff function. The analysis, however, does not extend to general payoff functions.

EC.5.2. Example: Price-Setting Newsvendor with Price-Based Substitutable Products

A firm manufactures J products and sells them to the market over a time horizon $[0, T]$. At time $t = 0$, the firm must decide the production quantities $\mathbf{Q} := (Q_j)_{j=1}^{J}$ and prices (net costs) $\mathbf{P} := (P_j)_{j=1}^{J}$. All the product demands are realized at time $t = T$. The demand for product j is $D_{jT}(\mathbf{P}) = A_{jT} - \sum_{i=1}^{J} \gamma_{ji} P_i$, where A_{jT} is the market size of this product and $\mathbf{\Gamma} = [\gamma_{ij}]_{i,j=1}^{J,J}$ is a J-by-J matrix. For each j, $\gamma_{jj} > 0$ represents the sensitivity of the demand for product j relative to its price P_j, and γ_{ji} (with $j \neq i$) represents the effect of the demand for product i on the demand for product j. We also define $R_j = Q_j + \sum_{i=1}^{J} \gamma_{ji} P_i$, representing the VPQ for product j, and let $\mathbf{R} = (R_j)_{j=1}^{J}$. The payoff function is

$$H_T((\mathbf{P}, \mathbf{R}), \mathbf{A}_T) = \sum_{j=1}^{J} \left[P_j \left(R_j - \sum_{i=1}^{J} \gamma_{ji} P_i \right) - (P_j + c_j)(R_j - A_{jT})^+ \right], \tag{EC.5.6}$$

where c_j is the unit production cost net unit salvage value. Then, the mV risk-management problem for this multi-product price-setting newsvendor model is in the form of (EC.5.3) with $\mathbf{q} = (\mathbf{P}, \mathbf{R})$.

COROLLARY EC.1. *Suppose $\gamma_{ij} < 0$ and $\tilde{\boldsymbol{\sigma}}_i \cdot \tilde{\boldsymbol{\sigma}}_j \geq 0$ for all $i \neq j$, where $\tilde{\boldsymbol{\sigma}}_j$ is the j-th row of $\tilde{\mathbf{\Sigma}}$ in (EC.5.1). Let $m = \mathsf{E}[H_T((\mathbf{P}^*, \mathbf{R}^*), \mathbf{A}_T)]$ and $\mathcal{S}_m^h = \arg\min_{\mathbf{P} \geq \mathbf{0}, \mathbf{R} \geq \mathbf{0}} B(m, (\mathbf{P}, \mathbf{R}))$. Let $\mathbf{q}^* = (\mathbf{P}^*, \mathbf{R}^*)$ be as defined in (EC.5.5). If the asset price trend positively impacts the market sizes (i.e., $\mathbf{C}_T^M \preceq \mathbf{C}_T$), then $\mathbf{P} \leq \mathbf{P}^*$ and $\mathbf{R} \leq \mathbf{R}^*$ for all $(\mathbf{P}, \mathbf{R}) \in \mathcal{S}_m^h$.*

Corollary EC.1 indicates that if the asset price trend positively impacts each product's market size, risk hedging reduces all products' optimal price and VPQ. The condition $\gamma_{ij} < 0$ is a standard assumption in the literature on multi-product problems (Aydin and Porteus 2008, Dong et al. 2009). It means that different products are *price-based substitutes*: a price increase in one product boosts the demand for another. The condition $\tilde{\boldsymbol{\sigma}}_i \cdot \tilde{\boldsymbol{\sigma}}_j \geq 0$ means that the unhedgeable factors of different product demands are nonnegatively correlated.

EC.5.3. Example: Price-Advertising Optimization

The example below is adapted from the classical price-advertising optimization problem in Dorfman and Steiner (1954); see also Bagwell (2007) for a survey of the optimal advertising literature. Consider a make-to-order manufacturer which must decide, at time $t = 0$, the unit selling price P (net the unit production cost) and an *advertising level* γ. The demand at time $t = T$ is D_T with

$$D_T = d(A_T, P, \gamma), \tag{EC.5.7}$$

where A_T is the market size following the model (EC.5.1) and $d(\cdot)$ is a twice-differentiable function, increasing in A_T and γ and decreasing in P. The manufacturer pays advertising cost $k(\gamma)$ that is

increasing in γ and does not depend on D_T. The profit function is $H_T((P,\gamma),A_T) = Pd(A_T,P,\gamma) - k(\gamma)$. Then, the mV risk-management problem for this pricing-advertising model is in the form of (EC.5.3) with $\mathbf{q} = (P,\gamma)$.

COROLLARY EC.2. *Suppose that $d(A_T,P,\gamma)$ defined in (EC.5.7) is increasing in A_T, concave and decreasing in P, concave and increasing in γ, and supermodular in (A_T,P,γ). Suppose also that $k(\gamma)$ is convex and increasing in γ. Let $m = \mathsf{E}[H_T((P^*,\gamma^*),A_T)]$ and $\mathcal{S}_m^h = \arg\min_{P\geq 0,\gamma\geq 0} B(m,(P,\gamma))$. Let $\mathbf{q}^* = (P^*,\gamma^*)$ be defined as in (EC.5.5). If the asset price trend positively benefits the market size (i.e., $C_T \succeq C_T^M$), then $P \leq P^*$ and $\gamma \leq \gamma^*$ for all $(P,\gamma) \in \mathcal{S}_m^h$.*

Corollary EC.2 states that if the asset price trend positively impact the market size, hedging decreases both the optimal pricing and advertising levels. The concavity of $d(A_T,P,\gamma)$ in P means that the demand is more sensitive to price increment at higher pricing levels. The concavity of $d(A_T,P,\gamma)$ in γ suggests a diminishing marginal benefit of advertising on product demand. The convexity of $k(\gamma)$ means that the marginal advertising cost increases as advertising level increases.

Now we explain the supermodularity of $d(A_T,P,\gamma)$ in all three arguments. This property is equivalent to the supermodularity of the function in each pair of the arguments, and we discuss them one by one. First, the supermodularity in the pricing and advertising levels (P,γ) indicates that the demand is more sensitive to the advertising level with a higher selling price. The same condition is also assumed in Topkis (1998); see its Example 3.3.10, where the "quality" there can stand for the advertising level. Second, the supermodularity in the market size and advertising level (A_T,γ) means that the marginal benefit on product demand from a higher advertising level increases with the market size. This is intuitive: a larger population of potential customers implies that advertising achieves a broader audience coverage.

Last, the supermodularity in the market size and pricing level (A_T,P) means that the detrimental effect on product demand from a price markup decreases with the market size. It may be attributed to factors—either asset-related or unrelated to the financial market—that boost the market size and decrease customers' sensitivity to price increments. For example, as the economy strengthens, the broad market stock index rises, and the financial component of the market size increases. Meanwhile, income increases amid better economic conditions, and people are less sensitive to price markups. For another example, when customers' preferences and tastes shift in favor of the product in concern, market size increases due to the increment of its non-financial component. Meanwhile, favoring this product, customers tend to be less sensitive to its price increment.

Part II: Proofs

EC.6. Proofs for Section 2
EC.6.1. Proof of Proposition 1

The expected profit, as a function of (P, R), is

$$\mathsf{E}[H_T(P, R)] = (P - c)(R - bP) - (P - s)\mathsf{E}[(R - A_T)^+].$$

Given P, differentiate $\mathsf{E}[H_T(P, R)]$ with respect to R:

$$\frac{\partial \mathsf{E}[H_T(P, R)]}{\partial R} = (P - c) - (P - s)F(R). \tag{EC.6.1}$$

Then, take second-order derivative to R:

$$\frac{\partial^2 \mathsf{E}[H_T(P, R)]}{\partial R^2} = -(P - s)f(R) < 0 \tag{EC.6.2}$$

Thus, for a given P, the expected profit is a concave function in R, and the optimal solution, denoted $R^{\mathrm{NV}}(P)$, is solved from setting $\partial \mathsf{E}[H_T(P, R)]/\partial R$ to zero. This leads to

$$R^{\mathrm{NV}}(P) = F^{-1}(\frac{P - c}{P - s}).$$

Clearly, $R^{\mathrm{NV}}(P)$ increases in P. Define:

$$m(P) := \mathsf{E}[H_T(P, R^{\mathrm{NV}}(P))] = (P - c)(R^{\mathrm{NV}}(P) - bP) - (P - s)\mathsf{E}[(R^{\mathrm{NV}}(P) - A_T)^+].$$

As $P \to c$, $R^{\mathrm{NV}}(c) \to -\infty$, $(P - s)\mathsf{E}[(R^{\mathrm{NV}}(P) - A_T)^+] \to 0$, then,

$$\lim_{P \to c} m(P) = \lim_{P \to c}(P - c)(R^{\mathrm{NV}}(P) - bP) = \lim_{P \to c} \frac{R^{\mathrm{NV}}(P) - bP}{\frac{1}{P - c}} = \lim_{P \to c} \frac{\frac{(c - s)f(R^{\mathrm{NV}}(P))}{(1 - F(R^{\mathrm{NV}}(P))^2)} - b}{-\frac{1}{(P - c)^2}}$$

$$= -\lim_{P \to c} \frac{1}{f(R)}(c - s)(\frac{P - s}{P - c})^2 + \lim_{P \to c} b(P - c)^2 = -(c - s)\lim_{R \to -\infty} \frac{F^2(R)}{f(R)}$$

$$= -(c - s)\lim_{R \to -\infty} \frac{F^2(R)}{r(R)}(1 - F(R)).$$

Under Assumption 1, $F^2(a)/r(a) \to 0$ as $a \to -\infty$, so the limit above is 0.

Next, differentiate $m(P)$ with respect to P:

$$m'(P) = R^{\mathrm{NV}}(P) - 2bP + bc - \mathsf{E}[(R^{\mathrm{NV}}(P) - A_T)^+] \tag{EC.6.3}$$

As $P \to c$, $R^{\mathrm{NV}}(c) \to -\infty$ and $m'(c) \to -\infty$.

Differentiate $m'(P)$ with respect to P:

$$m''(P) = -2b + (1 - F(R^{\mathrm{NV}}(P)))\frac{1}{\frac{dP}{dR^{\mathrm{NV}}(P)}} = -2b + \frac{(1 - F(R^{\mathrm{NV}}(P)))^3}{(c - s)f(R^{\mathrm{NV}}(P))} = -2b + \frac{(1 - F(R^{\mathrm{NV}}(P)))^2}{(c - s)r(R^{\mathrm{NV}}(P))} \tag{EC.6.4}$$

where we use the definition of the hazard rate for A_T, i.e., $r(a) = \frac{f(a)}{1-F(a)}$. Further differentiating $m''(P)$ with respect to P:

$$\frac{dm''(P)}{dP} = -\frac{(2r(R^{\mathrm{NV}}(P))^2 + r'(R^{\mathrm{NV}}(P)))(1 - F(R^{\mathrm{NV}}(P)))^2}{(c-s)r^2(R^{\mathrm{NV}}(P))}\frac{dR^{\mathrm{NV}}(P)}{dP}.$$

By Assumption 1, we have $\frac{dm''(P)}{dP} < 0$, so $m''(P)$ is decreasing in P. Thus, as P increases from c to ∞, $m''(P)$ decreases from ∞ to $-2b$. (As $P \to c$, $R^{\mathrm{NV}}(P) \to -\infty$, $r(R^{\mathrm{NV}}(P)) \to 0$, leading to $m''(P) \to \infty$; As $P \to \infty$, $\frac{(1-F(R^{\mathrm{NV}}(P)))^3}{f(R^{\mathrm{NV}}(P))} \to 0$ by Assumption 1). Combining the above, $m'(P)$ increases first and then decreases with $m'(P) \to -\infty$ as $P \to c$ and $m'(P) \to -\infty$ as $P \to \infty$.

Combining the analysis above and Assumption 2, there must exist some $P^0 > c$ such that $m'(P^0) > 0$, otherwise, the optimal expected profit is negative. Therefore, $m'(P) = 0$ has two zeros for $P > c$. The smaller one is the minimizer, the larger one is the maximizer. As P increases over c, $m(P)$ first decreases from 0 to a negative value, which is the minimum level of $m(P)$. Then, it increases to a positive value, which is the maximum level of $m(P)$ and the associated optimal price is denoted as P^{NV}, which solves the following optimality equation:

$$m'(P^{\mathrm{NV}}) = R^{\mathrm{NV}}(P) - 2bP^{\mathrm{NV}} + bc - \mathsf{E}[(R^{\mathrm{NV}}(P) - A_T)^+] = 0. \tag{EC.6.5}$$

The equation above is equivalent to:

$$2bP^{\mathrm{NV}} - bc = \mathsf{E}(R^{\mathrm{NV}}(P^{\mathrm{NV}}) \wedge A_T), \tag{EC.6.6}$$

and $m(P)$ decreases when $P > P^{\mathrm{NV}}$. Let $\underline{P}$ be the smaller zero of $m(P) = 0$ and then $P^{\mathrm{NV}} > \underline{P} > c$. Let $R^{\mathrm{NV}} = R^{\mathrm{NV}}(P^{\mathrm{NV}})$, which satisfies the other optimality equation:

$$P^{\mathrm{NV}} - c - (P^{\mathrm{NV}} - s)F(R^{\mathrm{NV}}) = 0.$$

Combining the optimality equations of $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ leads to (4), and this completes the proof. $\square$

EC.6.2.   Proof of Proposition 2

We prove the three results stated in this proposition one by one.

For Part (i), we have already proved that $m(P)$ is increasing in $P \in [\underline{P}, P^{\mathrm{NV}}]$ in Section EC.6.1. The variance of production payoff is:

$$\mathsf{Var}[H_T(P, R)] = (P - s)^2 \{\mathsf{E}[[(R - A_T)^+]^2] - E^2[(R - A_T)^+]\}.$$

Differentiate $\mathsf{Var}[H_T(P, R)]$ with respect to P:

$$\frac{\mathsf{Var}[H_T(P, R)]}{\partial P} = 2(P - s)[\mathsf{E}[(R - A_T)^+]^2 - E^2[(R - A_T)^+]] > 0.$$

Differentiate $\mathsf{Var}[H_T(P,R)]$ with respect to R:

$$\frac{\partial \mathsf{Var}[H_T(P,R)]}{\partial R} = 2(P-s)^2\{\mathsf{E}[(R-A_T)^+](1-F(R)) > 0.$$

Next, differentiating $v(P) = \mathsf{Var}[H_T(P, R^{\mathrm{NV}}(P))]$ with respect to P, we have

$$\frac{dv(P)}{dP} = \frac{\partial \mathsf{Var}[H_T(P, R^{\mathrm{NV}}(P))]}{\partial P} + \frac{\partial \mathsf{Var}[H_T(P, R^{\mathrm{NV}}(P))]]}{\partial R}\frac{dR^{\mathrm{NV}}(P)}{dP} > 0 \qquad (\text{EC.6.7})$$

Then $\mathsf{Var}[H_T(P, R^{\mathrm{NV}}(P))]$ is increasing in P for $P \geq c$.

For Part (ii), fixing R, define

$$m(R) = \mathsf{E}[H_T(P^{\mathrm{NV}}(R), R)] = (P^{\mathrm{NV}}(R) - c)(R - bP^{\mathrm{NV}}(R)) - (P^{\mathrm{NV}}(R) - s)\mathsf{E}[(R-A_T)^+].$$

Recall, $P^{\mathrm{NV}}(R) = (\mathsf{E}[R \wedge A_T] + bc)/(2b)$. It is straightforward to verify that $P^{\mathrm{NV}}(bc) < c$. Further, from Assumption 3, we have:

$$\mathsf{E}[(bc - A_T)^+] \leq 2b(c-s).$$

Then,

$$P^{\mathrm{NV}}(bc) = \frac{1}{2b}(\mathsf{E}[bc \wedge A_T] + bc) = \frac{1}{2b}(bc - \mathsf{E}[(bc - A_T)^+] + bc) > s,$$

so

$$m(bc) = b(P^{\mathrm{NV}}(bc) - c)(c - P^{\mathrm{NV}}(bc)) - (P^{\mathrm{NV}}(bc) - s)\mathsf{E}[(bc - A_T)^+] < 0.$$

Next, differentiate $m(R)$ with respect to R and taking into account $\partial \mathsf{E}[H_T(P^{\mathrm{NV}}(R), R)]/\partial P = 0$:

$$\frac{dm(R)}{dR} = \frac{d\mathsf{E}[H_T(P^{\mathrm{NV}}(R), R)]}{dR} = \frac{\partial \mathsf{E}[H_T(P^{\mathrm{NV}}(R), R)]}{\partial R} = (P^{\mathrm{NV}}(R) - c) - (P^{\mathrm{NV}}(R) - s)F(R).$$

Let

$$J(R) =: \frac{dm(R)}{dR} = (P^{\mathrm{NV}}(R) - c) - (P^{\mathrm{NV}}(R) - s)F(R),$$

and differentiate $J(R)$ with respect to R:

$$\frac{dJ(R)}{dR} = f(R)\Big[\frac{1-F(R)}{2br(R)} - (P^{\mathrm{NV}}(R) - s)\Big].$$

Then, differentiate $\frac{dJ(R)}{dR}$ with respect to R:

$$\frac{d^2J(R)}{dR^2} = \frac{df(R)}{dR}\frac{\frac{dJ(R)}{dR}}{f(R)} - \frac{f(R)(1-F(R))}{2br^2(R)}(2r^2(R) + r'(R)).$$

We claim that $J(R)$ is either monotone or unimodal in R. First note for any R, $\frac{f(R)(1-F(R))}{2br^2(R)}(2r^2(R) + r'(R)) > 0$ by Assumption 1. Then, there are two cases. The first one is that $\frac{dJ(R)}{dR} > 0$ or $\frac{dJ(R)}{dR} < 0$ for all R, then $J(R)$ is monotone in R. The other case is that there exists R such that $\frac{dJ(R)}{dR} = 0$. Then, at this R, $\frac{d^2J(R)}{dR^2} < 0$, which indicates that for any stationary point R, it must be a local

maximizer. Thus, $J(R)$ is unimodal for this case. Note, $J(bc) < 0$ and $J(R) \to -(c-s) < 0$ as $R \to \infty$. Then, by Assumption 2, there must exist some $R > bc$ such that $J(R) > 0$. Otherwise $J(R) < 0$ for all $R \geq bc$ which implies that $m(R)$ is decreasing in R and $m(R) \leq m(bc) < 0$, and the optimal VPQ is $R = bc$, which contradicts this assumption. By the analysis above, we conclude that $J(R)$ is unimodal in $R \geq bc$. Furthermore, there are two roots of $J(R) = 0$ and let R_1^0 (resp., R_2^0) be the smaller (resp., larger) one. Clearly, $m(R)$ is decreasing in $R < R_1^0$, increasing in $R_1^0 < R < R_2^0$ and decreasing in $R > R_2^0$. Since $m(bc) < 0$, there are two roots for $m(R) = 0$ for $R \geq bc$ and let $\underline{R}$ denote the smaller one. By Proposition 1, $R^{NV} = R_2^0$. Thus, we conclude that $m(R)$ is increasing in $\underline{R} < R < R^{NV}$ (and decreasing in $R > R^{NV}$).

Analogous to Proposition 2 (i), it is straightforward to prove that $\mathsf{Var}[H_T(P^{\mathrm{NV}}(R), R)]$ increase in R.

For Part (iii), let $\bar{m} = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$. For any $m \in (0, \bar{m})$, we claim that both $\partial \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}))]/\partial P$ and $\partial \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}]/\partial R$ are positive. Otherwise, suppose $\partial \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}))]/\partial P \leq 0$. As $\mathsf{E}[H_T(P, R_m^{\mathrm{NV}})]$ is concave in P and $\mathsf{E}[H_T(c, R_m^{\mathrm{NV}})] \leq 0 < \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})] = m$, we know that $\mathsf{E}[H_T(P, R_m^{\mathrm{NV}})]$ is increasing in $P \in (c, P^{\mathrm{NV}}(R_m^{\mathrm{NV}}))$ and decreasing in $P > P^{\mathrm{NV}}(R_m^{\mathrm{NV}})$. Then $P_m^{\mathrm{NV}} \geq P^{\mathrm{NV}}(R_m^{\mathrm{NV}})$ as we have assumed $\partial \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}))]/\partial P \leq 0$. So there exists P_m' such that $P_m' \leq P^{\mathrm{NV}}(R_m^{\mathrm{NV}}) \leq P_m^{\mathrm{NV}}$ and we have:

$$\mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})] = \mathsf{E}[H_T(P_m', R_m^{\mathrm{NV}})] = m, \ \mathsf{Var}[H_T((P_m', R_m^{\mathrm{NV}}))] \leq \mathsf{Var}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})].$$

In other words, keeping the target return m, $(P_m', R_m^{\mathrm{NV}})$ has a smaller variance than $(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})$ does, and thus $(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})$ cannot be the solution to the problem in (7). By contradiction, $\partial \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}))]/\partial P$ must be positive. Similarly, $\partial \mathsf{E}[H_T((P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}))]/\partial R > 0$ always holds.

Next, introducing the Lagrange multiplier λ, the Lagrangian function of the problem in (7) is:

$$\mathcal{L} = \mathsf{Var}[H_T(P, R)] - \lambda(\mathsf{E}[H_T(P, R)] - m). \tag{EC.6.8}$$

Then, $(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}}, \lambda_m^*)$ satisfies the Karush–Kuhn–Tucker (KKT) equations:

$$\frac{\partial \mathsf{Var}[H_T(P, R)]}{\partial P} - \lambda \frac{\partial \mathsf{E}[H_T(P, R)]}{\partial P} = 0; \tag{EC.6.9}$$

$$\frac{\partial \mathsf{Var}[H_T(P, R)]}{\partial R} - \lambda \frac{\partial \mathsf{E}[H_T(P, R)]}{\partial R} = 0; \tag{EC.6.10}$$

$$\mathsf{E}[H_T(P, R)] - m = 0. \tag{EC.6.11}$$

we have $\lambda_m^* > 0$ as we have proved that $\partial \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})]/\partial P > 0$ and $\partial \mathsf{E}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})]/\partial R > 0$ hold. By Envelop Theorem, we have:

$$\frac{d}{dm} \mathsf{Var}[H_T(P_m^{\mathrm{NV}}, R_m^{\mathrm{NV}})] = \lambda^* > 0$$

this proves that $\mathsf{Var}[H_T(P^{\mathrm{NV}}_m, R^{\mathrm{NV}}_m)]$ is increasing in m for $m \in (0, \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})])$.

When $(P, R) \rightarrow (P^{\mathrm{NV}}, R^{\mathrm{NV}})$, $m \rightarrow \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]]$, $\partial \mathsf{E}[H_T(P, R)]/\partial P \rightarrow 0$ and $\partial \mathsf{E}[H_T(P, R)]/\partial R \rightarrow 0$. Therefore, the KKT condition (EC.6.9) or condition (EC.6.10) implies that $\lambda^* \rightarrow \infty$. Next, note that by Assumption 2, $\mathsf{P}(A_T \leq R^{\mathrm{NV}}) = (P^{\mathrm{NV}} - c)/(P^{\mathrm{NV}} - s) \in (0, 1)$, hence $\mathsf{Var}(H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})) = (P^{\mathrm{NV}} - s)^2 \mathsf{Var}[(R^{\mathrm{NV}} - A_T)^+] > 0$. Also, $\mathsf{Var}[(H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}}))] = (P^{\mathrm{NV}} - s)^2 \mathsf{Var}[(R^{\mathrm{NV}} - A_T)]^+ \leq (P^{\mathrm{NV}} - s)^2 \mathsf{Var}(A_T) < \infty$. So, as $m \rightarrow \infty$, $\mathsf{Var}[H_T(P, R)] \rightarrow \mathsf{Var}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] \in (0, \infty)$. Hence, $d\sqrt{\mathsf{Var}[H_T(P, R)]}/dm = \lambda^*/(2\sqrt{\mathsf{Var}[H_T(P, R)]}) \rightarrow \infty$. This proves that the incremental risk approaches infinity when the expected payoff approaches the newsvendor's maximum profit in the base model. $\qquad\square$

EC.7. Proofs for Section 4
EC.7.1. Martingale Representation

Here we show that any martingale under P^M admits a martingale representation based on dX_t and $d\tilde{B}_t$. The result is summarized in the following proposition.

PROPOSITION EC.3. *Let Z_t and P^M be defined in (12) and suppose Z_t satisfies Assumption 5.*
(i) For any martingale $\tilde{M}_t$ under P^M, it has the following representation:

$$\tilde{M}_t = \tilde{M}_0 + \int_0^t a_s dX_s + \int_0^t b_s d\tilde{B}_s \qquad (\text{EC.7.1})$$

for some adapted processes $\{a_t\}$ and $\{b_t\}$.

(ii) For any martingale $\{\tilde{M}_t\}$ with respect to $\mathcal{G}_t$ under P^M (recall, $\mathcal{G}_t$ is generated by $\{B_t\}$ and $\mathcal{G}_t \subseteq \mathcal{F}_t$),

$$\tilde{M}_t = \tilde{M}_0 + \int_0^t \gamma_s dX_s \qquad (\text{EC.7.2})$$

for some process $\{\gamma_t\}$ that is adapted to $\mathcal{G}_t$.

Proof. To start with, note that $\{Z_t\}$ is also a martingale with respect to $\mathcal{G}_t$ under P. To see this, first note that $\{Z_t\}$ is clearly adapted to $\mathcal{G}_t$, and

$$\mathsf{E}[Z_T | \mathcal{G}_t] = \mathsf{E}[\mathsf{E}[Z_T | \mathcal{F}_t] | \mathcal{G}_t] = \mathsf{E}[Z_t | \mathcal{G}_t] = Z_t,$$

where the first equality is the iterated conditioning and the second one follows from Assumption 5; the last one is due to $Z_t \in \mathcal{G}_t$. Applying Itô's Lemma to $1/Z_t = \exp\{\int_0^t \eta_s dB_s + (1/2) \int_0^t \eta_s^2 ds\}$:

$$d\frac{1}{Z_t} = \frac{1}{Z_t}[\eta_t dB_t + \frac{1}{2}\eta_t^2 dt] + \frac{1}{2}\frac{1}{Z_t}\eta_t^2 dt = \frac{1}{Z_t}[\eta_t^2 dt + \eta_t dB_t].$$

Next, by Girsanov Theorem (Pham 2009), $B_t^M := B_t + \int_0^t \eta_s ds$ is a Brownian motion under P^M. Note that since the MPR process associated with $\tilde{B}_t$ is 0, $\{\tilde{B}_t\}$ is also a Brownian motion under P^M, i.e., the change of measure does not affect $\{\tilde{B}_t\}$. Therefore,

$$dX_t = \sigma_t X_t(\eta_t dt + dB_t) = \sigma_t X_t dB_t^M. \qquad (\text{EC.7.3})$$

Clearly, $\{X_t\}$ is adapted to $\mathcal{G}_t$.

Now we prove Part (i). Define $M_t = Z_t \tilde{M}_t$, which is clearly an adapted process. In particular, $\mathsf{E}[|M_t|] = \mathsf{E}[Z_t|\tilde{M}_t|] = \mathsf{E}^M[|\tilde{M}_t|] < \infty$ hence $\{M_t\}$ is integrable under P. Then, for any $0 \le t \le u \le T$,

$$\mathsf{E}[M_u \,|\, \mathcal{F}_t] = \mathsf{E}[Z_u \tilde{M}_u \,|\, \mathcal{F}_t] = \mathsf{E}\Big[\frac{Z_u}{Z_t} Z_t \tilde{M}_u \,\Big|\, \mathcal{F}_t\Big] = Z_t \mathsf{E}^M[\tilde{M}_u \,|\, \mathcal{F}_t] = Z_t \tilde{M}_t = M_t.$$

The third equality follows from $Z_t \in \mathcal{F}_t$ and Bayes formula for change of measure (Pham 2009). Thus, M_t is a martingale under P and admits the following martingale representation (Pham 2009):

$$dM_t = \Gamma_t dB_t + \Lambda_t d\tilde{B}_t.$$

$\{\Gamma_t\}$ and $\{\Lambda_t\}$ are adapted processes. Then, apply Itô's Lemma on $\tilde{M}_t = M_t \frac{1}{Z_t}$:

$$\begin{aligned}
d\tilde{M}_t = d\Big(M_t \frac{1}{Z_t}\Big) &= M_t d\frac{1}{Z_t} + \frac{1}{Z_t} dM_t + dM_t d\frac{1}{Z_t} \\
&= M_t \frac{1}{Z_t}(\eta_t^2 dt + \eta_t dB_t) + \frac{1}{Z_t}(\Gamma_t dB_t + \Lambda_t d\tilde{B}_t) + \Gamma_t \frac{\eta_t}{Z_t} dt \\
&= \tilde{M}_t \eta_t(\eta_t dt + dB_t) + \frac{\Gamma_t}{Z_t}(dB_t + \eta_t dt) + \frac{\Lambda_t}{Z_t} d\tilde{B}_t \\
&= \frac{\big[\tilde{M}_t \eta_t + \frac{\Gamma_t}{Z_t}\big]}{\sigma_t X_t} dX_t + \frac{\Lambda_t}{Z_t} d\tilde{B}_t.
\end{aligned}$$

This proves the first part of the proposition.

Now we proceed to Part (ii), under which $\{\tilde{M}_t\}$ is an P^M-martingale with respect to $\mathcal{G}_t$. We have:

$$\tilde{M}_t = \mathsf{E}^M[\tilde{M}_T | \mathcal{G}_t] = \mathsf{E}\Big[\frac{Z_T}{Z_t}\tilde{M}_T \Big| \mathcal{G}_t\Big] = \frac{1}{Z_t}\mathsf{E}(Z_T \tilde{M}_T | \mathcal{G}_t).$$

To show the second equality, we verify that the definitions of conditional expectation are satisfied. First, $\mathsf{E}\Big[\frac{Z_T}{Z_t}\tilde{M}_T \Big| \mathcal{G}_t\Big]$ is clearly adapted to $\mathcal{G}_t$. And, for any set $B \in \mathcal{G}_t$,

$$\mathsf{E}^M\Big[\mathsf{E}\Big[\frac{Z_T}{Z_t}\tilde{M}_T \Big| \mathcal{G}_t\Big]\mathbf{1}_B\Big] = \mathsf{E}\Big[Z_t \mathsf{E}\Big[\frac{Z_T}{Z_t}\mathbf{1}_B \tilde{M}_T \Big| \mathcal{G}_t\Big]\Big] = \mathsf{E}\Big[\frac{Z_t}{Z_t}\mathsf{E}\Big[Z_T \mathbf{1}_B \tilde{M}_T \Big| \mathcal{G}_t\Big]\Big] = \mathsf{E}\Big[Z_T \mathbf{1}_B \tilde{M}_T\Big] = \mathsf{E}^M(\mathbf{1}_B \tilde{M}_T).$$

The first equality is change of measure and the second one follows from $Z_t \in \mathcal{G}_t$; the third equality is the iterated conditioning and the last one is again the change of measure.

Define $M_t := \mathsf{E}(Z_T \tilde{M}_T | \mathcal{G}_t)$, which is clearly a martingale with respect to $\mathcal{G}_t$ under P hence admits the martingale representation:

$$dM_t = \Gamma_t dB_t,$$

where $\{\Gamma_t\}$ is a process adapted to $\mathcal{G}_t$. Now, apply Itô's Lemma to $\tilde{M}_t$:

$$d\tilde{M}_t = d\Big(\frac{1}{Z_t}M_t\Big) = \frac{1}{Z_t}dM_t + M_t d\frac{1}{Z_t} + dM_t d\frac{1}{Z_t}$$

$$
\begin{aligned}
&= \frac{1}{Z_t}\Gamma_t dB_t + M_t \frac{1}{Z_t}[\eta_t^2 dt + \eta_t dB_t] + \frac{\Gamma_t}{Z_t}\eta_t dt \\
&= \frac{M_t \eta_t}{Z_t}[\eta_t dt + dB_t] + \frac{\Gamma_t}{Z_t}[\eta_t dt + dB_t] \\
&= \Big[\frac{M_t \eta_t + \Gamma_t}{Z_t}\Big]dB_t^M \\
&= \Big[\frac{M_t \eta_t + \Gamma_t}{Z_t \sigma_t X_t}dX_t^M\Big]
\end{aligned}
$$

Since $\{M_t\}$, $\{\Gamma_t\}$, $\{\eta_t\}$, $\{Z_t\}$, $\{\sigma_t\}$ and $\{X_t\}$ are all adapted to $\mathcal{G}_t$, this completes the proof. $\qquad\square$

REMARK EC.2. If $\mathcal{F}_t$ is generated by $\{B_t\}$ and $\{\tilde{\mathbf{B}}_t\}$, where the latter is a multi-dimensional Brownian motion that is independent from $\{B_t\}$, (EC.7.1) in Part (i) of Proposition EC.3 still holds and the only modification is to replace $b_s d\tilde{B}_s$ with $\mathbf{b}_s \cdot d\tilde{\mathbf{B}}_s$ for some adapted (vector) process $\{\mathbf{b}_t\}$. The proof is completely analogous.

COROLLARY EC.3. *For any nonnegative and P^M-integrable random variable $M_T \geq 0$ that is measurable with respect to $\mathcal{G}_T$,*

$$
\mathsf{E}^M(M_T\,|\,\mathcal{F}_t) = \mathsf{E}^M(M_T\,|\,\mathcal{G}_t) = M_0 + \int_0^t \gamma_s dX_s, \quad \forall 0 \leq t \leq T, \tag{EC.7.4}
$$

where $\{\gamma_t\}$ is a process that is adapted to $\mathcal{G}_t$. In particular, the above holds for $M_T = Z_T$ defined in (12).

Proof. Define $Y_t := \mathsf{E}^M(M_T\,|\,\mathcal{G}_t)$, which is clearly a nonnegative martingale under P^M with respect to $\mathcal{G}_t$. Since $M_T \in \mathcal{G}_T$, we have $Y_T = M_T$. By Proposition EC.3 Part (ii), Y_t admits the martingale representation (recall, $\{B_t^M = B_t + \int_0^t \eta_s ds\}$ is a Brownian motion under P^M):

$$
dY_t = \gamma_t dX_t = \gamma_t \sigma_t X_t dB_t^M
$$

for some process $\{\gamma_t\}$ adapted to $\mathcal{G}_t$; recall, $\{\sigma_t\}$ and $\{X_t\}$ are also adapted to $\mathcal{G}_t$. So, Y_t is a local martingale under P^M and, being nonnegative, a supermartingale under P^M (Revuz and Yor 1999). Next, for any $0 \leq t \leq T$,

$$
\mathsf{E}^M(Y_t) = \mathsf{E}^M[\mathsf{E}^M(M_T\,|\,\mathcal{G}_t)] = \mathsf{E}^M(M_T),
$$

which is a constant. Therefore, being a supermartingale with constant expectation, $\{Y_t\}$ is a martingale under P^M (Revuz and Yor 1999). Then, we have

$$
Y_t = \mathsf{E}^M(Y_T\,|\,\mathcal{F}_t) = \mathsf{E}^M(M_T\,|\,\mathcal{F}_t).
$$

This establishes (EC.7.4). Since $Z_T > 0$ and $Z_T \in \mathcal{G}_T$, (EC.7.4) holds for $M_T = Z_T$. This completes the proof. $\qquad\square$

EC.7.2.   Proof of Lemma 1

Recall, X_t follows the dynamics in (EC.7.3), which we copy below for easy reference, and hence is a (local) martingale under P^M:

$$dX_t = \sigma_t X_t dB_t^M.$$

According to the definition of A_t, for all $u \geq t \geq 0$,

$$A_u = C_u + \tilde{\sigma}\tilde{B}_u = C_t + \tilde{\sigma}\tilde{B}_t + C_u - C_t + \tilde{\sigma}(\tilde{B}_u - \tilde{B}_t) = A_t + C_u - C_t + \tilde{\sigma}(\tilde{B}_u - \tilde{B}_t). \quad \text{(EC.7.5)}$$

By the assumption that $C_u - C_t \in \sigma(\{X_s, t \leq s \leq u\})$ and (X_t, σ_t) is Markovian under P^M, the distribution of $C_u - C_t$ (conditional on $\mathcal{F}_t$) is in turn determined by (X_t, σ_t). Then, as already argued, (A_t, X_t, σ_t) is Markovian under P^M.

We reiterate the projected production payoff process in (17):

$$
\begin{aligned}
V_t(P, R) &= \mathsf{E}^M[H_T(P, R) \,|\, \mathcal{F}_t] \\
&= (R - bP)(P - c) - (P - s)\mathsf{E}^M[(R - A_T)^+ \,|\, \mathcal{F}_t] \\
&= (R - bP)(P - c) - (P - s)\mathsf{E}^M[(R - A_T)^+ \,|\, X_t, A_t, \sigma_t] \\
&= (R - bP)(P - c) - (P - s)f(t, X_t, A_t, \sigma_t),
\end{aligned}
$$

where the third equality follows from the Markovian property of (A_t, X_t, σ_t) and the function in the last line is defined as follows:

$$f(t, x, a, \sigma) := \mathsf{E}^M[(R - A_T)^+ \,|\, X_t = x, A_t = a, \sigma_t = \sigma].$$

Clearly, V_t is a martingale under P^M. By Proposition EC.3 and accounting for (EC.7.3), $V_t(P, R)$ has the following representation, with $V_0(P, R) = \mathsf{E}^M[H_T(P, R)]$:

$$V_t(P, R) = V_0(P, R) + \int_0^t \psi_s(P, R) dB_s^M + \int_0^t \delta_s(P, R) d\tilde{B}_s.$$

ψ_t and δ_t are processes adapted to $\mathcal{F}_t$ and they lead to (18) by applying Itô's lemma to $V_t(P, R)$:

$$
\begin{aligned}
dV_t(P, R) &= -(P - s)df(t, x, a, \sigma) \\
&= -(P - s)\Big\{ f_t dt + f_x dX_t + f_a dA_t + f_\sigma d\sigma_t + \frac{1}{2}f_{xx}x^2\sigma_t^2 dt + \frac{1}{2}f_{aa}\tilde{\sigma}^2 dt + \frac{1}{2}f_{\sigma\sigma}(d\sigma_t)^2 \\
&\quad + f_{xa}(dX_t dA_t) + f_{a\sigma}(dA_t d\sigma_t) + f_{x\sigma}(dX_t d\sigma_t) \Big\} \\
&= -(P - s)\{ [\ldots]dt + \psi_t dB_t^M + f_a \tilde{\sigma} d\tilde{B}_t \} \\
&= -(P - s)\Big\{ \frac{\psi_t}{\sigma_t X_t} dX_t + f_a \tilde{\sigma} d\tilde{B}_t \Big\}.
\end{aligned}
$$

In the last line, the dt-term vanishes due to martingale property of V_t. Furthermore, dB_t^M is related to dX_t based on the dynamics of X_t in P^M; see (EC.7.3). Comparing the last line with the martingale representation of $V_t(P, R)$ above and matching the coefficient of the term $d\tilde{B}_t$ leads to:

$$\delta_t(P, R) = -\tilde{\sigma}(P - s)f_a(t, X_t, A_t, \sigma_t).$$

(Matching the dX_t-term leads to $\xi_t = \psi_t/(\sigma_t X_t)$).

Applying (EC.7.5),

$$
\begin{aligned}
f(t, x, a, \sigma) &= \mathsf{E}^M[(R - A_T)^+ \mid X_t = x, A_t = a, \sigma_t = \sigma] \\
&= \mathsf{E}^M\left[(R - (a + C_T - C_t + \tilde{\sigma}(\tilde{B}_T - \tilde{B}_t)))^+ \mid X_t = x, A_t = a, \sigma_t = \sigma\right] \\
&= \mathsf{E}^M\left[(R - (a + C_T - C_t + \tilde{\sigma}\sqrt{T - t}Z))^+ \mid X_t = x, A_t = a, \sigma_t = \sigma\right] \\
&= \mathsf{E}^M\left[(R - (a + C_T - C_t + \tilde{\sigma}\sqrt{T - t}Z))^+ \mid X_t = x, \sigma_t = \sigma\right]
\end{aligned}
$$

where $Z = (\tilde{B}_T - \tilde{B}_t)/\sqrt{T - t}$ follows standard normal distribution independent of all other random variables. For the last equality, note that as argued above, the distribution of $C_T - C_t$ (given $\mathcal{F}_t$) depends on realization of (X_t, σ_t) only. In other words, the conditional distribution $[C_T - C_t \mid X_t = x, A_t = a, \sigma_t = \sigma]$ is the same as $[C_T - C_t \mid X_t = x, \sigma_t = \sigma]$. Therefore, we can take first derivative of $f(t, x, a, \sigma)$ with to a as follows:

$$
\begin{aligned}
f_a(t, x, a, \sigma) &= -\mathsf{E}^M\left[\mathbf{1}\{a + C_T - C_t + \tilde{\sigma}\sqrt{T - t}Z \le R\} \mid X_t = x, \sigma_t = \sigma\right] \\
&= -\mathsf{P}^M(A_T \le R \mid X_t = x, A_t = a, \sigma_t = \sigma) \\
&= -\mathsf{P}^M(A_T \le R \mid \mathcal{F}_t).
\end{aligned}
$$

Therefore, we have:

$$\delta_t(P, R) = \tilde{\sigma}(P - s)\mathsf{P}^M(A_T \le R \mid \mathcal{F}_t).$$

Clearly, $\delta_t(P, R)$ increases in both P and R. □

EC.7.3. Proof of Theorem 1

In this part, we apply the quadratic hedging technique in Gourieroux et al. (1998) to solve the hedging problem. The setup in Gourieroux et al. (1998) is semimartingale-based and abstract, hence the solution is not as explicit as ours since our setup is based on Brownian motions. In Section EC.7.3.1, we lay out technical preparations that are needed in the proof of Theorem 1 in Section EC.7.3.

EC.7.3.1.   Technical Preparation Recall, the risk-neutral measure P^M is defined via the associated Radon-Nikodym (R-N) derivative:

$$Z_T = \frac{d\mathsf{P}^M}{d\mathsf{P}} := e^{-\int_0^T \eta_t dB_t - \frac{1}{2}\int_0^T \eta_t^2 dt},$$

and $Z_t^M := \mathsf{E}^M(Z_T \,|\, \mathcal{F}_t) = \mathsf{E}(Z_T^2 \,|\, \mathcal{F}_t)/Z_t$ in (16), and thus $Z_0^M = \mathsf{E}(Z_T^2)$. Note that $Z_t^M > 0$ since $Z_T > 0$. We introduce a process N_t and another probability measure P^R. With Assumption 5, P^R below is well-defined:

$$\frac{d\mathsf{P}^R}{d\mathsf{P}} = \frac{\left(Z_T^M\right)^2}{Z_0^M}, \qquad \text{thus} \quad \frac{d\mathsf{P}^R}{d\mathsf{P}^M} = \frac{Z_T^M}{Z_0^M}. \tag{EC.7.6}$$

Note that Z_t^M is a positive P^M-martingale, and by Corollary EC.3, it admits the following martingale representation:

$$dZ_t^M = \zeta_t dX_t^M = \zeta_t \sigma_t X_t dB_t^M, \tag{EC.7.7}$$

where ζ_t is an adapted process to $\mathcal{G}_t$ (recall, $\mathcal{G}_t$ is the filtration generated by B_t, hence independent of $\tilde{B}_t$). $B_t^M = \int_0^t \eta_s dt + B_t$, which is a Brownian motion under P^M. The second equality in (EC.7.7) follows from (EC.7.3). Define:

$$\psi_t := \sigma_t X_t \zeta_t. \tag{EC.7.8}$$

Now, from (EC.7.6), Z_t^M/Z_0^M is the density process associated with $d\mathsf{P}^R/d\mathsf{P}^M$. Then, applying Girsanov's Theorem and accounting for (EC.7.7), the MPR process associated with $d\mathsf{P}^R/d\mathsf{P}^M$ is

$$\eta_t^M := -\frac{\psi_t}{Z_t^M}. \tag{EC.7.9}$$

In addition, B_t^R defined below is a Brownian motion under P^R:

$$dB_t^R := dB_t^M + \eta_t^M dt = dB_t^M - \frac{\psi_t}{Z_t^M} dt. \tag{EC.7.10}$$

Next, we introduce $N_t = (N_t^0, N_t^1)$ as follows:

$$N_t^0 := \frac{1}{Z_t^M}, \qquad N_t^1 := \frac{X_t}{Z_t^M}. \tag{EC.7.11}$$

Recall, N_t^0 and N_t^1 are interpreted as the original assets, dollar (i.e. 1) and financial asset (X_t), *denominated* in Z_t^M. By change of measure and Jensen's inequality,

$$Z_t^M = \frac{1}{Z_t}\mathsf{E}(Z_T^2 \,|\, \mathcal{F}_t) \geq \frac{1}{Z_t}\mathsf{E}^2(Z_T \,|\, \mathcal{F}_t) = Z_t > 0,$$

with Z_t defined in (12). By Assumption (4), $X_t > 0$. Therefore, N_t^0 and N_t^1 are well-defined, and both are strictly positive.

Applying Itô's lemma and accounting for (EC.7.10):

$$dN_t^0 = -(N_t^0)^2 \psi_t [dB_t^M - \psi_t N_t^0 dt] = -(N_t^0)^2 \psi_t dB_t^R,$$

$$dN_t^1 = -N_t^0[N_t^1\psi_t - \sigma_t X_t][[dB_t^M - \psi_t N_t^0 dt]] = -N_t^0[N_t^1\psi_t - \sigma_t X_t]dB_t^R. \qquad \text{(EC.7.12)}$$

Clearly, both N_t^0 and N_t^1 are *local martingales* under P^R; being nonnegative, they are also *super-martingales*. It is easy to verify that they are indeed P^R-martingales by having constant means (and being supermartingales):

$$\mathsf{E}^R(N_t^0) = \frac{1}{Z_0^M}\mathsf{E}^M\left[Z_T^M\frac{1}{Z_t^M}\right] = \frac{1}{Z_0^M}\mathsf{E}^M\left[\frac{1}{Z_t^M}\mathsf{E}^M(Z_T^M|\mathcal{F}_t)\right] = \frac{1}{Z_0^M};$$

the first equality is change of measure using (EC.7.6), the second one uses iterated conditioning on $\mathcal{F}_t$ and the fact that Z_t^M is adapted to $\mathcal{F}_t$. The last equality uses the definition of Z_t^M in (16), noting $Z_T = Z_T^M$. Analogous verification can be applied to N_t^1:

$$\mathsf{E}^R(N_t^1) = \mathsf{E}^M\left[\frac{Z_t^M}{Z_0^M}\frac{X_t}{Z_t^M}\right] = \frac{X_0}{Z_0^M} = N_0^1;$$

the first equality uses the fact that Z_t^M/Z_0^M is the density process for $d\mathsf{P}^R/d\mathsf{P}^M$ and N_t^1 is adapted to $\mathcal{F}_t$, and the second equality is by martingale property of X_t. We summarize the analysis above into the following lemma.

LEMMA EC.1. *N_t^0 and N_t^1 in (EC.7.11) are both martingales under P^R.*

Now, we are ready to define $\mathcal{M}_X$, $\mathcal{M}_N$, $\mathcal{A}_X$ and $\mathcal{A}_N$, all of which are technically crucial in defining *admissible class* of hedging strategies. ($\sim$ stands for the equivalence between probability measures.)

$$\mathcal{M}_X := \left\{\mathsf{P}^{\bar{M}} \sim \mathsf{P} : \frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \in L_2(\mathsf{P}),\ X_t \text{ is a } \mathsf{P}^{\bar{M}}\text{-martingale}\right\}. \qquad \text{(EC.7.13)}$$

$\mathcal{M}_X$ contains the equivalent martingale measures that have square-integrable R-N derivatives. By Assumption 4, $\mathsf{P}^M \in \mathcal{M}_X$, hence $\mathcal{M}_X \neq \emptyset$. Similarly for N_t, we define

$$\mathcal{M}_N := \left\{\mathsf{P}^{\bar{R}} \sim P : \frac{1}{Z_T^M}\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \in L_2(\mathsf{P}),\ N_t^0 \text{ and } N_t^1 \text{ are } \mathsf{P}^{\bar{R}}\text{-martingales}\right\}. \qquad \text{(EC.7.14)}$$

It is straightforward to verify that for P^R defined in (EC.7.6), $\frac{1}{Z_T^M}\frac{d\mathsf{P}^R}{d\mathsf{P}} \in L_2(\mathsf{P})$, and together with Lemma EC.1 this implies $\mathsf{P}^R \in \mathcal{M}_N$, hence $\mathcal{M}_N \neq \emptyset$.

Based on $\mathcal{M}_X$ and $\mathcal{M}_N$, we define admissible classes of trading strategies. We start with $\mathcal{A}_X$, the admissible class of hedging strategies in (14). An adapted process $\vartheta = \{\theta_t, t \in [t, T]\}$ is admissible by belonging to the following set:

$$\mathcal{A}_X := \{\vartheta : \vartheta \text{ is } X_t\text{-integrable};\ \chi_T(\vartheta) \in L_2(\mathsf{P}); \forall \mathsf{P}^{\bar{M}} \in \mathcal{M}_X,\ \{\chi_t(\vartheta), t \in [0, T]\} \text{ is a } \mathsf{P}^{\bar{M}}\text{-martingale}\}.$$
$$\text{(EC.7.15)}$$

(Recall, $\vartheta = \{\theta_t, t \in [0, T]\}$ and $\chi_t(\vartheta) = \int_0^t \theta_s dX_s$.) Next, we define the set of all terminal wealth attainable by admissible trading strategies:

$$\chi_T(\mathcal{A}_X) := \{\chi_T(\vartheta) \,|\, \vartheta \in \mathcal{A}_X\}. \tag{EC.7.16}$$

We remark that $\chi_T(\mathcal{A}_X)$ is closed in $L^2(\mathsf{P})$; see Lemma 2.6 and Theorem 2.8 of Černỳ and Kallsen (2008); and for a brief review on this, see Theorem A.1 of Wang and Wissel (2013). This property of $\chi_T(\mathcal{A}_X)$ allows us to establish the following technical result, with proof collected in Section EC.7.3.2.

LEMMA EC.2. *Let Z_t^M be defined in (16), with dynamics specified in (16) which is reiterated below:*

$$dZ_t^M = \zeta_t dX_t.$$

Under Assumptions 4 and 5, $\zeta_t \in \mathcal{A}_X$; in other words, ζ_t is an admissible hedging strategy with respect to X_t. Hence, by definition of the set of admissible strategies $\mathcal{A}_X$ in (EC.7.15), Z_t^M is a $\mathsf{P}^{\bar{M}}$-martingale for each $\mathsf{P}^{\bar{M}}$ in $\mathcal{M}_X$.

It will become clear later that Lemma EC.2 is crucial in establishing connection between $\mathcal{M}_X$ and $\mathcal{M}_N$, which plays a key role in solving the quadratic hedging problem.

Next, recall that N_t in (EC.7.11) can be viewed as asset prices *denominated* in Z_t^M, hence we can also define admissible trading strategies with respect to N_t. An adapted process $\varphi = \{\phi_t = (\phi_t^0, \phi_t^1), t \in [0, T]\}$ is admissible if it belongs to the following set:

$$\mathcal{A}_N := \Big\{ \varphi : \varphi \text{ is } N_t\text{-integrable and}$$
$$Z_T^M \pi_T(\varphi) \in L_2(\mathsf{P}), \forall \mathsf{P}^{\bar{R}} \in \mathcal{M}_N, \{\pi_t(\varphi), t \in [0, T]\} \text{ is a } \mathsf{P}^{\bar{R}}\text{-martingale}\Big\}, \tag{EC.7.17}$$

where the notation parallels those for $\mathcal{A}_X$: $\varphi = \{\phi_t, t \in [0, T]\}$, with $\phi_t = (\phi_t^0, \phi_t^1)$ being a two-dimensional adapted process. And

$$\pi_t(\varphi) := \int_0^t \phi_s \cdot dN_s = \int_0^t \phi_s^0 dN_s^0 + \int_0^t \phi_s^1 dN_s^1. \tag{EC.7.18}$$

Similar to (EC.7.16), we define $\pi_T(\mathcal{A}_N) := \{\pi_T(\varphi) | \varphi \in \mathcal{A}_N\}$ to be the attainable terminal wealth by admissible strategies in $\mathcal{A}_N$.

Now, we establish bijection between $\mathcal{M}_X$ in (EC.7.13) and $\mathcal{M}_N$ in (EC.7.14), which will be used later in proving the key lemma of this section. The following lemma is a special case of Proposition 3.1 in Gourieroux et al. (1998), and our proof here will make explicit uses of Bayes formula based on Doob's martingale.

LEMMA EC.3. *Recall P^M and P^R are defined in (12) and (EC.7.6), respectively.*

(i) $\forall \mathsf{P}^{\bar{M}} \in \mathcal{M}_X$, *the probability measure defined below is in* $\mathcal{M}_N$.

$$\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} := \frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \cdot \frac{Z_T^M}{Z_0^M} = \frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \cdot \frac{d\mathsf{P}^R}{d\mathsf{P}^M}.$$

(ii) $\forall \mathsf{P}^{\bar{R}} \in \mathcal{M}_N$, *the probability measure defined below is in* $\mathcal{M}_X$.

$$\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} := \frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \cdot \frac{Z_0^M}{Z_T^M} = \frac{d\mathsf{P}^{\bar{R}}}{dP} \cdot \frac{1}{\frac{d\mathsf{P}^R}{d\mathsf{P}^M}}$$

The proof is to check the conditions specified in (16) and (EC.7.13) for each case respectively, and we collect the details in Section EC.7.3.2.

To this point, we are ready to present the key lemma of this section, which spells out an one-to-one relationship between the two admissible classes $\mathcal{A}_X$ in (EC.7.15) and $\mathcal{A}_N$ in (EC.7.17).

LEMMA EC.4. *(i) For any given* X_t*-admissible trading strategy* $\vartheta = \{\theta_t, t \in [0, T]\} \in \mathcal{A}_X$, *there exists an* N_t*-admissible strategy* $\varphi = \{\phi_t = (\phi_t^0, \phi_t^1), t \in [0, T]\} \in \mathcal{A}_N$, *such that* $\forall t \in [0, T]$,

$$\frac{\chi_t(\vartheta)}{Z_t^M} = \pi_t(\varphi), \quad \text{and} \quad \phi_t = (\chi_t(\vartheta) - \theta_t X_t, \theta_t).$$

(ii) Conversely, given any N_t*-admissible strategy* $\varphi = \{\phi_t = (\phi_t^0, \phi_t^1), t \in [0, T]\} \in \mathcal{A}_N$, *there exists an* X_t*-admissible trading strategy* $\vartheta = \{\theta_t, t \in [0, T]\} \in \mathcal{A}_X$, *such that* $\forall t \in [0, T]$,

$$\frac{\chi_t(\vartheta)}{Z_t^M} = \pi_t(\varphi), \quad \text{and} \quad \theta_t = \zeta_t(\pi_t(\varphi) - \phi_t \cdot N_t) + \phi_t^1 \quad \text{with } \zeta_t \text{ defined in (16)}.$$

(iii) Combining (i) and (ii), we have:

$$\pi_T(\mathcal{A}_N) = \frac{\chi_T(\mathcal{A}_X)}{Z_T^M} := \left\{ \frac{\chi_T(\vartheta)}{Z_T^M} : \vartheta \in \mathcal{A}_X \right\};$$

recall, $\chi_T(\mathcal{A}_X)$ *is the set of attainable wealth defined in (EC.7.16) and* $\pi_T(\mathcal{A}_N)$ *is similarly defined.*

Proof. We first show Part (i) of this lemma. Given $\vartheta = \{\theta_t, t \in [0, T]\}$, write $\chi_t = \chi_t(\vartheta) = \int_0^t \theta_s dX_s$ for lighter notation. Apply Itô's lemma on $\frac{\chi_t}{Z_t^M}$, accounting for the dynamics of Z_t^M, N_t^0 and N_t^1 in, respectively, (EC.7.7) and (EC.7.12):

$$
\begin{aligned}
d\chi_t N_t^0 &= \chi_t dN_t^0 + N_t^0 d\chi_t + d\chi_t dN_t^0 \\
&= \chi_t dN_t^0 + \theta_t N_t^0 \sigma_t X_t dB_t^M - \theta_t \sigma_t X_t N_t^0 (N_t^0 \psi_t) dt \\
&= (\chi_t - \theta_t X_t) dN_t^0 + \left[\theta_t X_t dN_t^0 + \theta_t N_t^0 \sigma_t X_t dB_t^M - \theta_t \sigma_t X_t N_t^0 (N_t^0 \psi_t) dt \right] \\
&= (\chi_t - \theta_t X_t) dN_t^0 + \left[(\theta_t X_t)[-N_t^0 (N_t^0 \psi_t) dB_t^R] + \theta_t \sigma_t X_t N_t^0 [dB_t^M - (N_t^0 \psi_t) dt] \right] \\
&= (\chi_t - \theta_t X_t) dN_t^0 + \theta_t \left[-(N_t^0 \psi_t) N_t^1 + N_t^0 \sigma_t X_t \right] dB_t^R \\
&= (\chi_t - \theta_t X_t) dN_t^0 + \theta_t dN_t^1 \\
&= \phi_t \cdot dN_t,
\end{aligned}
\tag{EC.7.19}
$$

and this leads to the equality stated in (i).

The rest is to show $\varphi \in \mathcal{A}_N$. First, $Z_T^M \int_0^T \phi_t \cdot dN_t = \chi_T(\vartheta) \in L^2(P)$, since $\vartheta \in \mathcal{A}_X$. Next, $\forall \mathsf{P}^{\bar{R}} \in \mathcal{M}_N$, we have $\mathsf{E}^{\bar{R}}\big[|\frac{\chi_T(\vartheta)}{Z_T^M}|\big] = \mathsf{E}\Big[|\chi_T(\vartheta)|(\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}/Z_T^M)\Big] < \infty$, by Cauchy-Schwarz inequality, the fact $\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}/Z_T^M \in L^2(\mathsf{P})$ (since $\mathsf{P}^{\bar{R}} \in \mathcal{M}_N$) and $\chi_T(\vartheta) \in L^2(\mathsf{P})$ (by $\vartheta \in \mathcal{A}_X$). Then $\pi_T(\varphi) \in L^1(\mathsf{P}^{\bar{R}})$ and we can take conditional expectation:

$$\mathsf{E}^{\bar{R}}\Big[\frac{\chi_T(\vartheta)}{Z_T^M}\Big|\mathcal{F}_t\Big] = \mathsf{E}\Big[\chi_T(\vartheta)\frac{\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}}{(Z_T^M/Z_0^M)}\Big|\mathcal{F}_t\Big]\frac{1}{\mathsf{E}[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}|\mathcal{F}_t]Z_0^M}.$$

Applying Bayes formula:

$$\mathsf{E}\Big[\frac{\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}}{(Z_T^M/Z_0^M)}\Big|\mathcal{F}_t\Big] = Z_0^M\Big[\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}N_T^0\Big|\mathcal{F}_t\Big]\frac{1}{\mathsf{E}[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}|\mathcal{F}_t]}\Big]\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big|\mathcal{F}_t\Big]$$

$$= Z_0^M\mathsf{E}^{\bar{R}}[N_T^0|\mathcal{F}_t]\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big|\mathcal{F}_t\Big]$$

$$= Z_0^M N_t^0\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big|\mathcal{F}_t\Big];$$

the above accounts for the fact that N_t^0 is a $\mathsf{P}^{\bar{R}}$-martingale.

Combining the above, we derive:

$$\mathsf{E}^{\bar{R}}\Big[\frac{\chi_T(\vartheta)}{Z_T^M}\Big|\mathcal{F}_t\Big] = \Big[\mathsf{E}\Big[\chi_T(\vartheta)\frac{\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}}{(Z_T^M/Z_0^M)}\Big|\mathcal{F}_t\Big]\frac{1}{\mathsf{E}\Big[\frac{\frac{d\mathsf{P}^{\hat{R}}}{d\mathsf{P}}}{Z_T^M/Z_0^M}\Big|\mathcal{F}_t\Big]}\Big]N_t^0$$

$$= \mathsf{E}^{\bar{M}}[\chi_T(\vartheta)|\mathcal{F}_t]N_t^0$$

$$= \frac{\chi_t(\vartheta)}{Z_t^M};$$

where $\mathsf{P}^{\bar{M}}$ is defined by:

$$\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} := \frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\frac{Z_0^M}{Z_T^M},$$

and by Lemma EC.3, $\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \in \mathcal{M}_X$. Then, since $\vartheta \in \mathcal{A}_X$, $\chi_t(\vartheta)$ must be a martingale under $\mathsf{P}^{\bar{M}}$, which gives the last equality in the derivation above. Then, the above implies that $\pi_t(\varphi) = \chi_t(\vartheta)/Z_t^M$ is a $\mathsf{P}^{\bar{R}}$-martingale $\forall \mathsf{P}^{\bar{R}} \in \mathcal{M}_N$, hence $\varphi \in \mathcal{A}_N$ and (i) is proved.

Now we prove Part (ii). Given $\varphi = \{\phi_t = (\phi_t^0, \phi_t^1), t \in [0,T]\} \in \mathcal{A}_N$, we apply Itô's lemma on $Z_t^M \int_0^t \phi_s \cdot dN_s$:

$$dZ_t^M \int_0^t \phi_s \cdot dN_s = Z_t^M(\phi_t^0 dN_t^0 + \phi_t^1 dN_t^1) + \Big(\int_0^t \phi_s \cdot dN_s\Big)\psi_t dB_t^M + \phi_t^0 dZ_t^M dN_t^0 + \phi_t^1 dZ_t^M dN_t^1$$

$$= Z_t^M\phi_t^0 dN_t^0 + Z_t^M\phi_t^1 dN_t^1 + \Big(\int_0^t \phi_s \cdot dN_s\Big)\psi_t dB_t^M - \phi_t^0(\psi_t N_t^0)^2 dt$$

$$- \phi_t^1(\psi_t N_t^0)[\psi_t N_t^1 - \sigma_t X_t]dt. \tag{EC.7.20}$$

Now, use (EC.7.12) to express dN_t^0 and dN_t^1 (and choose the representation involving dB_t^M), then it is straightforward to verify that dt-term vanishes and the equation above reduces to:

$$
\begin{aligned}
dZ_t^M \int_0^t \phi_s \cdot dN_S &= \Big[-\phi_t^0 N_t^0 \psi_t - \phi_t^1 (N_t^0 X_t \psi_t - \sigma_t) + \psi_t \int_0^t \phi_s \cdot dN_s \Big] dB_t^M \\
&= \Big[-\phi_t^0 N_t^0 \zeta_t - \phi_t^1 (N_t^0 X_t \zeta_t - 1) + \zeta_t \int_0^t \phi_s \cdot dN_s \Big] \sigma_t dB_t^M \\
&= \Big[\zeta_t \Big(\int_0^t \phi_s \cdot dN_s - \phi_t^0 N_t^0 - \phi_t^1 N_t^1 \Big) + \phi_t^1 \Big] dX_t,
\end{aligned} \tag{EC.7.21}
$$

where the second equality uses $N_t^1 = X_t N_t^0$, as well as the relation between ζ_t and ψ_t in (EC.7.8); the third equality uses the P^M-dynamics of X_t: $dX_t = \sigma_t X_t dB_t^M$. The integrand with respect to dX_t in the last line above gives the expression for θ_t specified in (ii).

What remains is to show ϑ stated in (ii) is in $\mathcal{A}_X$. First, note $\chi_T(\vartheta) = Z_T^M \pi_T(\varphi)) \in L^2(P)$ follows from $\varphi \in \mathcal{A}_N$. Next, $\forall \mathsf{P}^{\bar{M}} \in \mathcal{M}_X$, define $\mathsf{P}^{\bar{R}}$ by

$$
\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} := \frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \frac{Z_T^M}{Z_0^M}.
$$

By Lemma EC.3, $\mathsf{P}^{\bar{R}} \in \mathcal{M}_N$. Note $\mathsf{E}^{\bar{M}}[|\chi_T(\vartheta)|] = \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \Big| \pi_T(\varphi) Z_T^M \Big| \Big] < \infty$ is easily verified by using Cauchy-Schwarz inequality, $\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \in L^2(\mathsf{P})$ and $\pi_T(\varphi) Z_T^M \in L^2(\mathsf{P})$ (since $\varphi \in \mathcal{A}_N$). Hence we compute the following conditional expectation under $\mathsf{P}^{\bar{M}}$ and apply Bayes formula to change the measure:

$$
\mathsf{E}^{\bar{M}}[Z_T^M \pi_T(\varphi) | \mathcal{F}_t] = \mathsf{E}\Big[(Z_T^M \pi_T(\varphi) \frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}) \Big| \mathcal{F}_t \Big] \frac{1}{\mathsf{E}[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} | \mathcal{F}_t]}
$$

By Bayes formula,

$$
\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \Big| \mathcal{F}_t \Big] = \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \frac{Z_T^M}{Z_0^M} \Big| \mathcal{F}_t \Big] = \mathsf{E}^{\bar{M}}\Big[\frac{Z_T^M}{Z_0^M} \Big| \mathcal{F}_t \Big] \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \Big| \mathcal{F}_t \Big] = \frac{Z_t^M}{Z_0^M} \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \Big| \mathcal{F}_t \Big];
$$

where the first quality uses definition of $\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}$ above; the second equality switches measure between $\mathsf{P}^{\bar{M}}$ and P; and the last equality is by the fact that Z_t^M is a martingale under $\mathsf{P}^{\bar{M}}$ as implied by Lemma EC.2.

Combining all above,

$$
\begin{aligned}
\mathsf{E}^{\bar{M}}[Z_T^M \pi_T(\varphi) | \mathcal{F}_t] &= \mathsf{E}\Big[Z_T^M \pi_T(\varphi) \frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \Big| \mathcal{F}_t \Big] \frac{1}{\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \Big| \mathcal{F}_t \Big]} \\
&= \mathsf{E}\Big[Z_T^M \pi_T(\varphi) \frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \Big| \mathcal{F}_t \Big] \frac{1}{\mathsf{E}[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} | \mathcal{F}_t]} \frac{Z_t^M}{Z_0^M} \\
&= \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \pi_T(\varphi) \Big| \mathcal{F}_t \Big] \frac{1}{\mathsf{E}[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} | \mathcal{F}_t]} Z_t^M \\
&= \mathsf{E}^{\bar{R}}[\pi_T(\varphi) | \mathcal{G}_t] Z_t^M \\
&= \pi_t(\varphi) Z_t^M
\end{aligned}
$$

where the last equality is due to $\pi_t(\varphi)$ is a $\mathsf{P}^{\bar{R}}$-martingale implied by $\varphi \in \mathcal{A}_N$ (recall, $\mathsf{P}^{\bar{R}} \in \mathcal{M}_N$ by Lemma EC.3). This concludes that $\chi_t(\vartheta) = \pi_t(\varphi)Z_t^M$ is a $\mathsf{P}^{\bar{M}}$-martingale $\forall \mathsf{P}^{\bar{M}} \in \mathcal{M}_X$; hence ϑ defined in (ii) is in $\mathcal{A}_X$. This completes the proof of (ii). $\qquad\qquad\qquad\square$

EC.7.3.2.   Proof of Lemma EC.2 and and EC.3. We first prove Lemma EC.2 and here is an outline of the proof. Recall, $\chi_T(\mathcal{A}_X)$, the set of wealth attainable by admissible strategies with X_t is introduced in (EC.7.16), which is a nonempty set closed in $L^2(\mathsf{P})$. Our approach is to first show $Z_T^M - Z_0^M \in \chi_T(\mathcal{A}_X)$. Once this is established, then $\exists\{\theta_t, t \in [0,T]\} \in \mathcal{A}_X$ such that $Z_T^M - Z_0^M = \int_0^T \theta_t dX_t$, i.e., θ_t is an admissible strategy that attains $Z_T^M - Z_0^M$. Then, we have

$$\mathsf{E}^M(Z_T^M - Z_0^M \,|\, \mathcal{F}_t) = Z_t^M - Z_0^M = \int_0^t \zeta_s dX_s = \mathsf{E}^M\left(\int_0^T \theta_t dX_t \,\Big|\, \mathcal{F}_t\right) = \int_0^t \theta_s dX_s;$$

the first equality follows the definition of Z_t^M in (16) (note $Z_T = Z_T^M$), the second equality uses the dynamics of Z_t^M in (16); the third equality uses definition of θ_t: an admissible strategy attaining $Z_T^M - Z_0^M$; the last equality uses the admissibility of θ_t: the induced wealth process is an martingale under any measure from $\mathcal{M}_X$ (defined in (EC.7.13)), and in particular, recall that $\mathsf{P}^M \in \mathcal{M}_X$. In this way, we establish that $\int_0^t \zeta_s dX_s = \int_0^t \theta_s dX_s$; note both integrals are continuous martingales under P^M, hence $\zeta_t = \theta_t$. Thus, we can establish $\zeta_t \in \mathcal{A}_X$, which in turn implies $Z_t^M = Z_M^0 + \int_0^t \zeta_s dX_s$ is an martingale under any measure from $\mathcal{M}_X$, and the lemma is proved.

Now we proceed with showing $Z_T^M - Z_0^M \in \chi_T(\mathcal{A}_X)$. Recall, $\chi_T(\mathcal{A}_X)$ is closed in $L^2(\mathsf{P})$, hence it is sufficient to find a sequence of elements in this set, with limit (in $L^2(\mathsf{P})$) as $Z_T^M - Z_0^M$, then the desired result will follow from the closedness. To do this, we follow the approach similar to that in the proof of Theorem 3.5 in Wang and Wissel (2013).

Define the sequence of stopping times with respect to $\mathcal{G}_t$:

$$\tau_k := \inf\{t \geq 0 : |\eta_t| \geq k\} \wedge T; \; k \in \mathbb{N}; \tag{EC.7.22}$$

recall η_t is the MPR process defined in (16). Since η_t is continuous, we have $\tau_k \uparrow T$ as $k \to \infty$. Clearly,

$$Z_{\tau_k} \to Z_T \quad a.s. \tag{EC.7.23}$$

Recall Z_t is assumed to be a continuous square-integrable martingale under P (by Assumption 5), hence Doob's L^p inequality implies:

$$\mathsf{E}\left[\sup_{t \in [0,T]} Z_t^2\right] \leq 2\mathsf{E}(Z_T^2) < \infty; \tag{EC.7.24}$$

Clearly, $\sup_{k \in \mathbb{N}} Z_{\tau_k}^2 \leq \sup_{t \in [0,T]} Z_t^2$ (note Z_t is a positive process), hence (EC.7.24) implies

$$\mathsf{E}\left[\sup_{k \in \mathbb{N}} Z_{\tau_k}^2\right] < \infty.$$

The above invokes dominated convergence in (EC.7.23), and we establish

$$Z_{\tau_k} \to Z_T \quad \text{in} \quad L^2(\mathsf{P}). \tag{EC.7.25}$$

Clearly, $Z_{\tau_k} \in L^2(\mathsf{P})$, hence also in $L^1(\mathsf{P}^M)$. So, for each $k \in \mathbb{N}$, define

$$M_t^{(k)} := \mathsf{E}^M(Z_{\tau_k}|\mathcal{F}_t) = \frac{1}{Z_t}\mathsf{E}(Z_T Z_{\tau_k}\,|\,\mathcal{F}_t) = M_0^{(k)} + \int_0^t \theta_s^{(k)} dX_s. \tag{EC.7.26}$$

the first equality is the change of measure formula, and the second equality is martingale representation with $\theta_t^{(k)}$ being a process adapted to $\mathcal{G}_t$ (see Corollary EC.3, accounting for the fact that $Z_{\tau_k} \geq 0$ and $Z_{\tau_k} \in \mathcal{G}_T$). In particular, note $M_T^{(k)} = Z_{\tau_k}$, and also $M_0^{(k)} = \mathsf{E}(Z_T Z_{\tau_k}) \to \mathsf{E}(Z_T^2) = Z_0^M$ by (EC.7.25); so we have

$$M_T^{(k)} - M_0^{(k)} \to Z_T - Z_0^M = Z_T^M - Z_0^M \quad \text{in} \quad L^2(\mathsf{P});$$

Now we have found a sequence of elements, $M_T^{(k)} - M_0^{(k)}$, converging to $Z_T^M - Z_0^M$ in $L^2(\mathsf{P})$. As outlined, the next step is to show $M_T^{(k)} - M_0^{(k)} \in \chi_T(\mathcal{A}_X)$ for each k. To this end, fix k and $\mathsf{P}^{\bar{M}} \in \mathcal{M}_X$, we will show $M_t^{(k)} - M_0^{(k)} = \int_0^t \theta_s^{(k)} dX_s$ defined in (EC.7.26) is a martingale in $\mathsf{P}^{\bar{M}}$, as follows. Clearly, $\int_0^t \theta_s^{(k)} dX_s$ is a $\mathsf{P}^{\bar{M}}$-local-martingale, since X_t is a martingale under this probability measure. To proceed, the crux is to examine the following. Recall Z_t has the exponential form as defined in (12), hence

$$
\begin{aligned}
Z_{t\wedge\tau_k} &= \exp\left\{-\int_0^{t\wedge\tau_k} \eta_s dB_s - \frac{1}{2}\int_0^{t\wedge\tau_k} \eta_s^2 ds\right\} \\
&= \exp\left\{-\int_0^t \eta_s \mathbf{1}\{s \leq \tau_k\} dB_s - \frac{1}{2}\int_0^t \eta_s^2 \mathbf{1}\{s \leq \tau_k\} ds\right\} \\
&= \exp\left\{-\int_0^t \hat{\eta}_s(dB_s^M - \eta_s ds) - \frac{1}{2}\int_0^t \hat{\eta}_s^2 ds\right\} \\
&= \exp\left\{-\int_0^t \hat{\eta}_s dB_s^M - \frac{1}{2}\int_0^t \hat{\eta}_s^2 ds\right\}\exp\left\{\int_0^t \hat{\eta}_s^2 ds\right\}
\end{aligned}
$$

where B_t^M is the P^M-Brownian-motion $(B_t^M = \eta_t dt + dB_t)$, and $\hat{\eta}_s$ is defined as $\hat{\eta}_s = \eta_s \mathbf{1}\{s \leq \tau_k\}$ (note $\hat{\eta}_s \eta_s = \hat{\eta}_s^2$); also note $\hat{\eta}_s \leq k$ by definition of τ_k in (EC.7.22). Denote

$$W_t = \exp\left\{-\int_0^t \hat{\eta}_s dB_s^M - \frac{1}{2}\int_0^t \hat{\eta}_s^2 ds\right\} \quad \text{and} \quad C_t = \exp\left\{\int_0^t \hat{\eta}_s^2 ds\right\};$$

then the expression above for $Z_{t\wedge\tau_k}$ becomes

$$Z_{t\wedge\tau_k} = W_t C_t.$$

For C_t, since each $\hat{\eta}_s$ is bounded by k, we have

$$1 \leq C_t \leq e^{k^2 T}, \forall t \in [0, T].$$

The above immediately implies $W_t \leq Z_{t \wedge \tau_k}$.

Next, clearly W_t is a local martingale undre P^M, and since $\hat{\eta}_t \leq k$, we have

$$\mathsf{E}^M\Big[\frac{1}{2}\exp\Big\{\int_0^T \hat{\eta}_t^2 dt\Big\}\Big] \leq e^{\frac{1}{2}k^2 T} < \infty;$$

in other words, Novikov's condition holds, indicating that W_t is a P^M-martingale.

Combining the above, we have

$$
\begin{aligned}
0 \leq M_0^{(k)} + \int_0^t \theta_s^{(k)} dX_s &= M_t^{(k)} \\
&= \mathsf{E}^M(M_T^{(k)} \mid \mathcal{F}_t) = \mathsf{E}^M(Z_{\tau_k} \mid \mathcal{F}_t) \\
&= \mathsf{E}^M(Z_{T \wedge \tau_k} \mid \mathcal{F}_t) = \mathsf{E}^M(W_T C_T \mid \mathcal{F}_t) \\
&\leq e^{k^2 T} W_t \leq e^{k^2 T} Z_{t \wedge \tau_k} \\
&\leq e^{k^2 T} \sup_{t \in [0,T]} Z_t
\end{aligned}
\tag{EC.7.27}
$$

the first line is just the definition in (EC.7.26), and the second line uses $M_T^{(k)} = Z_{\tau_k}$; the first equality on the third line uses the obvious fact $\tau_k \wedge T = \tau_k$, and the second equality makes use of the representation of $Z_{t \wedge \tau_k}$ established above; the fourth line is based on the bound on C_T and the martingale property of W_t established above, as well as $W_t \leq Z_{t \wedge \tau_k}$ as shown above.

Now, combining (EC.7.27) and (EC.7.24) implies $\sup_{t \in [0,T]} \int_0^t \theta_s^{(k)} dX_s \in L^2(\mathsf{P})$. Finally, by Cauchy-Schwarz inequality and $d\mathsf{P}^{\bar{M}}/d\mathsf{P} \in L^2(\mathsf{P})$

$$\mathsf{E}^{\bar{M}}\Big[\Big|\sup_{t \in [0,T]}\int_0^t \theta_s^{(k)} dX_s\Big|\Big] = \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\Big|\sup_{t \in [0,T]}\int_0^t \theta_s^{(k)} dX_s\Big|\Big] \leq \mathsf{E}\Big[\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\Big)^2\Big]\mathsf{E}\Big[\Big(\sup_{t \in [0,T]}\int_0^t \theta_s^{(k)} dX_s\Big)^2\Big] < \infty.$$

Therefore, $\int_0^t \theta_s^{(k)} dX_s = M_t^{(k)} - M_0^k$ is a $\mathsf{P}^{\bar{M}}$-martingale (Pham 2009). This establishes the desired result, and proves Lemma EC.2.

Based on Lemma EC.2, we are now ready to prove Lemma EC.3.

For Part (i), suppose $\mathsf{P}^{\bar{M}} \in \mathcal{M}_X$ is given and $\mathsf{P}^{\bar{R}}$ follows the stated definition. First note that clearly $\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} > 0$ almost surely since $\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} > 0$ and $Z_T^M = Z_T$ takes exponential form (see (12)); hence $\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \sim \mathsf{P}$. Next, we have

$$\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big] = \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\Big] = \mathsf{E}^{\bar{M}}\Big[\frac{Z_T^M}{Z_0^M}\Big] = 1.$$

The last equality follows from Lemma EC.2, which indicates that Z_t^M is a $\mathsf{P}^{\bar{M}}$-martingale since $\mathsf{P}^{\bar{M}} \in \mathcal{M}_X$ (see Lemma EC.2). Next, derive

$$\mathsf{E}\Big[\Big(\frac{1}{Z_T^M}\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big)^2\Big] = \mathsf{E}\Big[\Big(\frac{1}{Z_T^M}\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\Big)^2\Big] = \Big(\frac{1}{Z_0^M}\Big)^2 \mathsf{E}\Big[\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\Big)^2\Big] < \infty.$$

The $<$ follows from the fact that $\mathsf{P}^{\bar{M}} \in \mathcal{M}_X$; hence the above implies $\frac{1}{Z_T^M}\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \in L^2(\mathsf{P})$.

What remains is to show that N_t^0 and N_t^1 are $\mathsf{P}^{\bar{R}}$-martingales. First note

$$\mathsf{E}^{\bar{R}}(N_T^0) = \mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\frac{1}{Z_T^M}\Big) = \frac{1}{Z_0^M} < \infty;$$

hence we can apply conditional expectation and compute

$$\mathsf{E}^{\bar{R}}(N_T^0\,|\,\mathcal{F}_t) = \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\frac{1}{Z_T^M}\Big|\mathcal{F}_t\Big]\frac{1}{\mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\,\Big|\,\mathcal{F}_t\Big)}$$

$$= \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\Big|\mathcal{F}_t\Big]\frac{1}{\mathsf{E}^{\bar{M}}(Z_T^M\,|\,\mathcal{F}_t)\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\,\Big|\,\mathcal{F}_t\Big]}$$

$$= \frac{1}{Z_t^M} = N_t^0;$$

the first equality applies change of measure from $\mathsf{P}^{\bar{R}}$ to P, and the second equality follows from changing measure from P to $\mathsf{P}^{\bar{M}}$ on the term $\mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\,\Big|\,\mathcal{F}_t\Big)$; the third equality again uses the fact that Z_t^M is $\mathsf{P}^{\bar{M}}$-martingale based on Lemma EC.2. From above, we can conclude that N_t^0 is a martingale under $\mathsf{P}^{\bar{R}}$. Similar derivation applies to N_t^1 as follows. First note that

$$\mathsf{E}^{\bar{R}}(N_T^1) = \mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\frac{X_T}{Z_T^M}\Big) = \frac{1}{Z_0^M}\mathsf{E}^{\bar{M}}(X_T) = \frac{X_0}{Z_0^M} < \infty;$$

the last equality accounts for the fact that $\mathsf{P}^{\bar{M}}$ is a martingale measure with respect to X_t. Now we can compute the conditional expectation:

$$\mathsf{E}^{\bar{R}}(N_T^1\,|\,\mathcal{F}_t) = \mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\frac{X_T}{Z_T^M}\,\Big|\,\mathcal{F}_t\Big)\frac{1}{\mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\frac{Z_T^M}{Z_0^M}\Big|\mathcal{F}_t\Big)}$$

$$= \mathsf{E}^{\bar{M}}(X_T\,|\,\mathcal{F}_t)\mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\,\Big|\,\mathcal{F}_t\Big)\frac{1}{\mathsf{E}^{\bar{M}}(Z_T^M\,|\,\mathcal{F}_t)\mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\,\Big|\,\mathcal{F}_t\Big)}$$

$$= \frac{X_t}{Z_t^M} = N_t^1; \tag{EC.7.28}$$

the first equality applies change of measure formula on $\mathsf{E}^{\bar{R}}(N_T^1\,|\,\mathcal{F}_t)$; the second equality does the same for $\mathsf{E}^{\bar{M}}(X_T\,|\,\mathcal{F}_t)$ and $\mathsf{E}^{\bar{M}}(Z_T^M\,|\,\mathcal{F}_t)$, respectively; the third equality uses the fact that Z_t^M is a $\mathsf{P}^{\bar{M}}$-martingale (by Lemma EC.2). Hence, N_t^1 is also a martingale under $\mathsf{P}^{\bar{M}}$ by the derivation above. To this point, we have checked that $\mathsf{P}^{\bar{R}}$ satisfies conditions specified in $\mathcal{M}_N$, hence belongs to this set; this proves (i).

For Part (ii), the proof is analogous. Suppose $\mathsf{P}^{\bar{R}} \in \mathcal{M}_N$ is given, and define $\mathsf{P}^{\bar{M}}$ as stated. First note by the same argument as that for Part (i), $\mathsf{P}^{\bar{M}} > 0$ almost surely, hence equivalent to P. And,

$$\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\Big] = \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\frac{Z_0^M}{Z_T^M}\Big] = Z_0^M\mathsf{E}^{\bar{R}}(N_T^0) = Z_0^M N_0^0 = 1;$$

the third equality uses the fact that $\mathsf{P}^{\bar{R}}$ is a martingale measure for N_t^0 and N_t^1. Next, check

$$\mathsf{E}\Big[\Big(\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}}\Big)^2\Big] = (Z_0^M)^2 \mathsf{E}\Big[\Big(\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \cdot \frac{1}{Z_T^M}\Big)^2\Big] < \infty;$$

$<$ follows from $\frac{1}{Z_T^M} \frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \in L^2(\mathsf{P})$; hence the above implies $\frac{d\mathsf{P}^{\bar{M}}}{d\mathsf{P}} \in L^($$\mathsf{P})$.

Then, the rest is to show X_t is a martingale under $\mathsf{P}^{\bar{M}}$. We start with checking the integrability condition:

$$\mathsf{E}^{\bar{M}}(X_T) = \mathsf{E}\Big(\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \frac{Z_0^M}{Z_T^M} X_T\Big) = Z_0^M \mathsf{E}^{\bar{R}}(N_T^1) = Z_0^M N_0^1 = X_0 < \infty;$$

the third equality follows from that $\mathsf{P}^{\bar{R}} \in \mathcal{M}_N$ is a martingale measure for N_t^0 and N_t^1. Next, compute the conditional expectation

$$\begin{aligned}
\mathsf{E}^{\bar{M}}(X_T|\mathcal{F}_t) &= \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \frac{Z_0^M}{Z_T^M} X_T \Big| \mathcal{F}_t\Big] \frac{1}{\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} \frac{Z_0^M}{Z_T^M}\Big|\mathcal{F}_t\Big]} \\
&= \mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} N_T^1 \Big| \mathcal{F}_t\Big] \frac{1}{\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}} N_T^0 \Big|\mathcal{F}_t\Big]} \\
&= \mathsf{E}^{\bar{R}}(N_T^1|\mathcal{F}_t)\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big|\mathcal{F}_t\Big] \frac{1}{\mathsf{E}\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big|\mathcal{F}_t\Big]\mathsf{E}^{\bar{R}}(N_T^0|\mathcal{F}_t)} \\
&= \frac{N_t^1}{N_t^0} = X_t; \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (\text{EC.7.29})
\end{aligned}$$

the first equality applies change of measure formula on $\mathsf{E}^{\bar{M}}(X_T|\mathcal{F}_t)$; the second equality does the same, respectively, for $\mathsf{E}^{\bar{R}}(N_T^1|\mathcal{F}_t)$ and $\mathsf{E}^{\bar{R}}(N_T^0|\mathcal{F}_t)$; the third equality recognizes that N_t^0 and N_t^1 are martingales under $\mathsf{P}^{\bar{R}}$. This concludes that X_t is a $\mathsf{P}^{\bar{M}}$-martingale and proves Part (ii). $\quad\square$

EC.7.3.3. Proof of Theorem 1 In this section, we derive the optimal hedging strategy, θ_t^*, of (20) as well as the associated minimum variance function $B(m,P,R)$ in (21), and thereby provide proofs to Theorem 1 ; all derivations are based on results established in Section EC.7.3.1.

We will first transform the equality-constrained problem in (14) to an equivalent unconstrained *quadratic hedging* problem, and then solve the latter by applying the numeraire-based technique. Define

$$A(\lambda) := \inf_{\vartheta \in \mathcal{A}_X} \mathsf{E}\Big[\Big(\lambda - H_T(P,R) - \chi_T(\vartheta)\Big)^2\Big], \quad\quad\quad\quad (\text{EC.7.30})$$

where $A(\lambda)$ relates to $B(m,P,R)$ by (see Proposition 6.6.5 in Pham (2009)):

$$B(m,P,R) = \max_{\lambda}\big[A(\lambda) - (m-\lambda)^2\big]. \qu\quad\quad\quad\quad (\text{EC.7.31})$$

In addition, the optimal hedging strategy induced by the problem in (EC.7.30), with λ being set as the optimal solution to the maximization problem (EC.7.31), is also optimal to the problem (14). We will show $A(\lambda)$ takes the following expression:

$$A(\lambda) = \frac{[\lambda - V_0(P,R)]^2}{Z_0^M} + \int_0^T \mathsf{E}\Big[\frac{Z_t}{Z_t^M}\delta_t^2(P,R)\Big]dt, \qu\quad\quad\quad (\text{EC.7.32})$$

where $V_0(P, R)$ and $\delta_t(P, R)$ are terms involved in the martingale representation of $V_t(P, R)$ in (17); in particular, $V_0(P, R) = \mathsf{E}^M[H_T(P, R)]$, and $\delta_t(P, R)$ is defined in (19) . Z_t and Z_t^M follow (12) and (16), respectively.

In (EC.7.32), λ only enters the first component as a quadratic term; the second component is independent of λ. Thus, the minimization problem in (EC.7.31) has a quadratic objective function, since both $A(\lambda)$ expressed in (EC.7.32) and $(m - \lambda)^2$ are quadratic functions in λ, and it is straightforward to verify that the λ specified in (20) solves the right hand side of (EC.7.31) and gives the expression of $B(m, P, R)$ in (21).

Starting from here, we begin to prove (EC.7.32) by deriving solution to the hedging problem in (EC.7.30). Write $\hat{H}_T(\lambda) := \lambda - H_T$, and start with definition of $A(\lambda)$ in (EC.7.30):

$$
\begin{aligned}
A(\lambda) &= \inf_{\vartheta \in \mathcal{A}_X} \mathsf{E}\left[\left(\hat{H}_T(\lambda) - \chi_T(\vartheta)\right)^2\right] \\
&= \inf_{\vartheta \in \mathcal{A}_X} Z_0^M \mathsf{E}\left[\frac{(Z_T^M)^2}{Z_0^M}\left(\frac{\hat{H}_T(\lambda)}{Z_T^M} - \frac{\chi_T(\vartheta)}{Z_T^M}\right)^2\right] \\
&= Z_0^M \inf_{\vartheta \in \mathcal{A}_X} \mathsf{E}^R\left[\left(\frac{\hat{H}_T(\lambda)}{Z_T^M} - \frac{\chi_T(\vartheta)}{Z_T^M}\right)^2\right];
\end{aligned}
\tag{EC.7.33}
$$

recall the probability measure P is defined in (EC.7.6).

Continue with (EC.7.33), by Lemma EC.4 Part (iii),

$$
A(\lambda) = Z_0^M \inf_{\varphi \in \mathcal{A}_N} \mathsf{E}^R\left[\left(\frac{\hat{H}_T(\lambda)}{Z_T^M} - \int_0^T \phi_t \cdot dN_t\right)^2\right].
\tag{EC.7.34}
$$

A natural next step is to express $\frac{\hat{H}_T(\lambda)}{Z_T^M}$ as a stochastic integral with respect N_t, so that we can choose the optimal ϕ_t based on this representation. By the fact that $\hat{H}_T(\lambda)$ is bounded hence has finite second moment in P, it is easy to check that $\hat{H}_T(\lambda)/Z_T^M$ has finite second moment in P^R. Hence, we are able to define the Doob's martingale:

$$
\hat{M}_t := \mathsf{E}^R\left[\frac{\hat{H}_T(\lambda)}{Z_T^M} \,\middle|\, \mathcal{F}_t\right];
\tag{EC.7.35}
$$

note $\frac{\hat{H}_T(\lambda)}{Z_T^M} = \hat{M}_T$. Furthermore, $\hat{M}_t$ is a square-integrable martingale under P^R. Applying Bayes formula, we have

$$
\hat{M}_t = \mathsf{E}^M\left[\frac{\hat{H}_T(\lambda)}{Z_T^M}\frac{Z_T^M}{Z_t^M} \,\middle|\, \mathcal{F}_t\right] = N_t^0 \mathsf{E}^M[\hat{H}_T(\lambda) \,|\, \mathcal{F}_t].
$$

Let $M_t := \mathsf{E}^M[\hat{H}_T(\lambda) \,|\, \mathcal{F}_t]$, which is clearly a martingale under P^M. By Proposition EC.3 and (EC.7.3), M_t admits the martingale representation:

$$
dM_t = a_t dB_t^M + b_t d\tilde{B}_t
$$

for some adapted processes $\{a_t\}$ and $\{b_t\}$. Next, apply Itô's Lemma on $\hat{M}_t$, accounting for the dynamics of N_t^0 and N_t^1 in (EC.7.12) and the definition of η_t^M in (EC.7.9):

$$
\begin{aligned}
d\hat{M}_t = dN_t^0 M_t &= N_t^0 dM_t + M_t dN_t^0 + dM_t dN_t^0 \\
&= N_t^0 a_t dB_t^M + N_t^0 b_t d\tilde{B}_t + M_t dN_t^0 - (N_t^0)^2 a_t \psi_t dt \\
&= N_t^0 a_t [dB_t^M - N_t^0 \psi_t dt] + N_t^0 b_t d\tilde{B}_t + M_t dN_t^0 \\
&= N_t^0 a_t [dB_t^M + \eta_t^M dt] + N_t^0 b_t d\tilde{B}_t + M_t dN_t^0 \\
&= N_t^0 a_t dB_t^R + N_t^0 b_t d\tilde{B}_t + M_t dN_t^0.
\end{aligned}
$$

From the dynamics of N_t^0 and N_t^1 in (EC.7.12) and accounting for the fact that $N_t^1 = N_t^0 X_t$, it is straightforward to verify the following:

$$
\sigma_t N_t^1 dB_t^R = dN_t^1 - X_t dN_t^0.
$$

In particular, note that $\sigma_t > 0$ and $N_t^1 = X_t/Z_t^M > 0$. Substitute this equation to the dynamics of $\hat{M}_t$ above, we have:

$$
d\hat{M}_t = N_t^0 b_t d\tilde{B}_t + M_t dN_t^0 + \frac{N_t^0 a_t}{\sigma_t N_t^1}(dN_t^1 - X_t dN_t^0) = \left(M_t - \frac{a_t}{\sigma_t}\right) dN_t^0 + \frac{a_t}{\sigma_t X_t} dN_t^1 + N_t^0 b_t d\tilde{B}_t.
$$

Therefore, $\hat{M}_t$ admits the following representation:

$$
\hat{M}_t = \mathsf{E}^R \left[\frac{\hat{H}_T(\lambda)}{Z_T^M} \right] + \int_0^t \phi_s^H \cdot dN_s + \int_0^t \gamma_s d\tilde{B}_s, \tag{EC.7.36}
$$

where ϕ_t^H and γ_t are some adapted processes. In particular, note that $\tilde{B}_t$ is a Brownian motion under P^R since the market risk of price process associated with it is 0.

Substitute (EC.7.36) to (EC.7.34) and take into the consideration that N_t and $\tilde{B}_t$ are independent, we can further expand $A(\lambda)$:

$$
\begin{aligned}
A(\lambda) = Z_0^M \inf_{\varphi \in \mathcal{A}_N} \mathsf{E}^R & \left[\left(\mathsf{E}^R \left[\frac{\hat{H}_T(\lambda)}{Z_T^M} \right] + \int_0^T \gamma_t d\tilde{B}_t + \int_0^T \phi_t^H \cdot dN_t - \int_0^T \phi_t \cdot dN_t \right)^2 \right] \\
= Z_0^M \inf_{\varphi \in \mathcal{A}_N} \mathsf{E}^R & \left[\left(\mathsf{E}^R \left[\frac{\hat{H}_T(\lambda)}{Z_T^M} \right] + \int_0^T \gamma_t d\tilde{B}_t \right)^2 \right] + \mathsf{E}^R \left[\left(\int_0^T \phi_t^H \cdot dN_t - \int_0^T \phi_t \cdot dN_t \right)^2 \right].
\end{aligned}
$$
$$\tag{EC.7.37}$$

To reach the second line of above, note the cross term is zero, following the fact that $\int_0^t \gamma_s d\tilde{B}_s$, $\int_0^t \phi_s^H \cdot dN_s$ and $\int_0^t \phi_s \cdot dN_s$ are all square-integrable martingales under P^R. Specifically, for $\int_0^t \gamma_s d\tilde{B}_s$ and $\int_0^t \phi_s^H \cdot dN_s$, this follows from (EC.7.36) and that $\hat{M}_t$ is a P^R-square-integrable martingale. For $\int_0^t \phi_s \cdot dN_s$, this follows from the N_t-admissibility of ϕ_t: $\int_0^t \phi_s \cdot dN_s$ is P^R-martingale and $\int_0^T \phi_t \cdot dN_t \in L^2(\mathsf{P}^R)$ (since $Z_T^M \int_0^T \phi_t \cdot dN_t \in L^2(\mathsf{P})$). Hence, $\int_0^t \gamma_s d\tilde{B}_s$ and $\int_0^t (\phi_s^H - \phi_s) \cdot dN_s$ are two

square-integrable martingales under P^R, and they are independent since N_t is adapted to $\sigma(B_t)$, which is independent of $\tilde{B}_t$. This makes the cross term vanish.

Two results follow (EC.7.37). First, it is obvious that the optimal ϕ_t of this problem, denoted by $\phi_t^* = (\phi_t^{*0}, \phi_t^{*1})$, should be set as:

$$\phi_t^* = \phi_t^H; \tag{EC.7.38}$$

substituting this to (EC.7.37), apply Itô's isometry and switch measure from P^R to P, we have:

$$
\begin{aligned}
A(\lambda) &= Z_0^M \mathsf{E}^R \left[\left(\mathsf{E}^R \left[\frac{\hat{H}_T(\lambda)}{Z_T^M} \right] + \int_0^T \gamma_t d\tilde{B}_t \right)^2 \right] \\
&= Z_0^M \left(\mathsf{E}^M \left[\frac{\hat{H}_T(\lambda)}{Z_T^M} \frac{Z_T^M}{Z_0^M} \right] \right)^2 + Z_0^M \int_0^T \mathsf{E}^M \left[\gamma_t^2 \cdot \frac{Z_t^M}{Z_0^M} \right] dt \\
&= \frac{(\lambda - V_0)^2}{Z_0^M} + \int_0^T \mathsf{E} \left(\gamma_t^2 Z_t Z_t^M \right) dt. \tag{EC.7.39}
\end{aligned}
$$

The expression above for $A(\lambda)$ does not coincide with (EC.7.32) yet (the second term takes a different form); we will come back to this later.

The other result from (EC.7.37) is the expression for the optimal hedging strategy θ_t^*. With (EC.7.38), invoking Part (ii) of Lemma EC.4, we have the expression for θ_t^*:

$$\theta_t^* = \zeta_t \left(\int_0^t \phi_s^* \cdot dN_s - \phi_t^* \cdot N_t \right) + \phi_t^{*1}, \tag{EC.7.40}$$

and recall, ζ_t is defined in (16).

To this end, we have obtained the expressions for both $A(\lambda)$ and θ_t^*, in (EC.7.39) and (EC.7.40) respectively, and both expressions involve terms related to N_t. Next, we will replace such terms by terms associated with X_t and V_t only. To do so, the crux is to compare (17) and (EC.7.36) and match integrands for dt, dB_t^M and $d\tilde{B}_t$. Start with

$$\hat{M}_t = \frac{\lambda - V_0}{Z_0^M} + \int_0^t \phi_s^* \cdot dN_s + \int_0^t \gamma_s d\tilde{B}_s,$$

and this equation comes directly from (EC.7.36), changing measure from P^R to P^M for the first term, and accounting for $\phi_t^* = \phi^H$. Alternatively, $\hat{M}_t$ can also be represented as the following by applying change of measure formula:

$$
\begin{aligned}
\hat{M}_t &= \mathsf{E}^M \left[\frac{\lambda - H_T}{Z_T^M} \cdot \frac{Z_T^M}{Z_0^M} \,\Big|\, \mathcal{F}_t \right] \cdot \frac{1}{\mathsf{E}^M \left[\frac{Z_T^M}{Z_0^M} \,\big|\, \mathcal{F}_t \right]} = (\lambda - V_t) N_t^0 \\
&= \left[\lambda - V_0 - \int_0^t \xi_s dX_s - \int_0^t \delta_s d\tilde{B}_s \right] N_t^0, \tag{EC.7.41}
\end{aligned}
$$

where the first equality switches the measure from P^R to P^M; the second equality accounts for the fact $Z_T = Z_T^M$ and definition of N_t^0 in (EC.7.12); the third equality makes use of the martingale

representation of V_t in (17). Now, apply Itô's lemma on both (EC.7.41) and (EC.7.41), use the dynamics for N_t in (EC.7.12)), and match the $d\tilde{B}_t$ term, we have:

$$\gamma_t = -\frac{\delta_t}{Z_t^M}. \tag{EC.7.42}$$

Substituting (EC.7.42) in (EC.7.39) gives (EC.7.32), as desired.

Next, match dB_t^M (matching for dt term gives the same result) and obtain:

$$\xi_t + (\lambda - V_t)N_t^0 \zeta_t = \zeta_t(\phi_t^* \cdot N_t) - \phi_t^{*1}. \tag{EC.7.43}$$

Recall, ζ_t and ψ_t are defined in (16) and (EC.7.7) respectively, and to reach (EC.7.43) the relation in (EC.7.8) is used. Substituting (EC.7.43) to (EC.7.40) gives (20), taking into the following fact implied by Lemma EC.4 Part (ii):

$$\frac{\int_0^t \theta_s^* dX_s}{Z_t^M} = \int_0^t \phi_s^* \cdot dN_s.$$

Now, we show that θ_t^* above is an admissible trading strategy. By Part (ii) of Lemma EC.4, it is sufficient to show ϕ_t^* in (EC.7.38) is in $\mathcal{A}_N$. Note $\phi_t^* = \phi_t^H$. Denote $\hat{M}_t^\circ = \int_0^t \phi_s^H \cdot dN_s$; we already showed that (see the arguments below (EC.7.37)) $\hat{M}_t^\circ$ is a square-integrable martingale under P^R, and this implies it has finite expected quadratic variation under P^R:

$$\mathsf{E}^R([\hat{M}^\circ, \hat{M}^\circ]_t) < \infty. \tag{EC.7.44}$$

Then, for any $\mathsf{P}^{\bar{R}} \in \mathcal{M}_N$, we have

$$\mathsf{E}^{\bar{R}}(\sqrt{[\hat{M}^\circ, \hat{M}^\circ]_t}) = \mathsf{E}^R\Big[\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}^R}\sqrt{[\hat{M}^\circ, \hat{M}^\circ]_t}\Big] \leq \mathsf{E}^R\Big[\Big(\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}^R}\Big)^2\Big]\mathsf{E}^R([\hat{M}^\circ, \hat{M}^\circ]_t) < \infty; \tag{EC.7.45}$$

where the $\leq$ follows Cauchy–Schwarz inequality; and the $<$ follows (EC.7.44) and the following

$$\mathsf{E}^R\Big[\Big(\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}^R}\Big)^2\Big] = \mathsf{E}\Big[\Big(\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\frac{d\mathsf{P}}{d\mathsf{P}^R}\Big)^2\frac{(Z_T^M)^2}{Z_0^M}\Big] = \mathsf{E}\Big[\Big(\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big)^2\frac{(Z_0^M)^2}{(Z_T^M)^4}\frac{(Z_T^M)^2}{Z_0^M}\Big] = Z_0^M\mathsf{E}\Big[\Big(\frac{1}{Z_T^M}\frac{d\mathsf{P}^{\bar{R}}}{d\mathsf{P}}\Big)^2\Big] < \infty; \tag{EC.7.46}$$

where the $<$ follows from the definition of $\mathcal{M}_N$; see (EC.7.14). Now by Burkholder-Davis-Gundy inequality, $\hat{M}_t^\circ$ is an martingale under $\mathsf{P}^{\bar{R}}$. Last, $Z_T^M \hat{M}_T^\circ \in L^2(P)$ easily follows from that $\hat{M}_t^\circ$ is square-integrable under P^R. Combining the arguments above, we have $\phi_t^* = \phi_t^H \in \mathcal{A}_N$, hence θ_t^* is admissible as argued. This proves Part (i) of Theorem 1.

Now, we show that $\mathsf{Var}(H_T + \chi_T^*)$ has the expression specified in (21) of Part (ii) of this theorem. First note that from (EC.7.39), $A(\lambda)$ is a quadratic function in λ and its second term (the integral) does not involve λ. Then, by (EC.7.31), clearly the optimal λ should solve

$$\max_\lambda \frac{(\lambda - V_0(P, R))^2}{Z_0^M} - (m - \lambda)^2,$$

and it is easy to verify that the solution has the expression in (20). Substituting (EC.7.42) to the second term of $A(\lambda)$ in (EC.7.39) equates this expression to (EC.7.32). Then, using the expression of λ_m in (20) and applying (EC.7.31), the expression of $B(m, P, R)$ in (21) immediately follows.

What remains is to show Part (iii). Recall, the set of all wealth attainable by admissible trading strategies is defined in (EC.7.16). For simpler notation, $\chi_T(\mathcal{A}_X)$ is written as $\chi_{\mathcal{A}}$ below. As noted following (EC.7.16), $\chi_{\mathcal{A}}$ is *closed.* So, being a subspace of $L^2(\mathsf{P})$ [15], $\chi_{\mathcal{A}}$ is a Hilbert space (with the inner product being that of $L^2(\mathsf{P})$).

Next, note that the problem in (EC.7.30), which is equivalent to the original hedging problem in (14) with λ taking the value in (20), is equivalent (in the sense of for any $\chi_T \in \chi_{\mathcal{A}}$, it is attained by some trading strategy in $\mathcal{A}_X$):

$$\min_{\chi_T \in \chi_{\mathcal{A}}} \mathsf{E}[(\lambda - H_T - \chi_T)^2] \tag{EC.7.47}$$

Then, by Projection Theorem, $\forall \chi_T \in \chi_{\mathcal{A}}$, the following must hold:

$$\mathsf{E}[(\lambda - H_T - \chi_T^*)\chi_T] = 0. \tag{EC.7.48}$$

Note that due to the fact that H_T is bounded, $\{-\xi_t, 0 \le t \le T\} \in \mathcal{A}_X$. So, $\{\iota_t = \theta_t^* + \xi_t, 0 \le t \le T\} \in \chi_{\mathcal{A}}$ and thus $\chi_T^{\mathrm{iv}} \in \chi_{\mathcal{A}}$. Then, let $\chi_T = \chi_T^{\mathrm{iv}}$ in (EC.7.48), rearrange the terms and apply the definition of H_T^{h} in (23), we have (accounting for that $\chi_T^* = \chi_T^{\mathrm{rm}} + \chi_T^{\mathrm{iv}}$):

$$\mathsf{E}(\chi_T^{\mathrm{iv}} H_T^{\mathrm{h}}) = \lambda \mathsf{E}(\chi_T^{\mathrm{iv}}) - \mathsf{E}[(\chi_T^{\mathrm{iv}})^2] = \lambda(m - V_0) - \mathsf{E}[(\chi_T^{\mathrm{iv}})^2], \tag{EC.7.49}$$

where the second equality is due to the following facts:

$$\mathsf{E}(H_T + \chi_T^*) = \mathsf{E}(H_T^{\mathrm{h}} + \chi_T^{\mathrm{iv}}) = m$$

and

$$\mathsf{E}(H_T^{\mathrm{h}}) = \mathsf{E}\left(H_T + \chi_T^{\mathrm{rm}}\right) = \mathsf{E}\left(V_0 + \int_0^T \xi_t dX_t + \int_0^T \delta_t d\tilde{B}_t + \int_0^T (-\xi_t) dX_t\right) = V_0 + \mathsf{E}\left(\int_0^T \delta_t d\tilde{B}_t\right) = V_0.$$

The second equality is due to $H_T = V_T$ and the representation of V_t of (18) in Lemma 1. Then, we can derive the left hand side of (24), using (EC.7.49) and the expression of λ in (20):

$$\mathsf{Var}(\chi_T^{\mathrm{iv}}) + \mathsf{Cov}(\chi_T^{\mathrm{iv}}, H_T^{\mathrm{h}}) = \mathsf{E}[(\chi_T^{\mathrm{iv}})^2] - [\mathsf{E}(\chi_T^{\mathrm{iv}})]^2 + \lambda(m - V_0) - \mathsf{E}[(\chi_T^{\mathrm{iv}})^2] - \mathsf{E}(\chi_T^{\mathrm{iv}})\mathsf{E}(H_T^{\mathrm{h}})$$

$$= -(m - V_0)^2 + \lambda(m - V_0) - (m - V_0)V_0$$

$$= \frac{1}{Z_0^M - 1}(m - V_0)^2,$$

and this establishes (24).

Now, note that

$$B(m, P, R) = \mathsf{Var}(H_T + \chi_T^*) = \mathsf{Var}(H_T^{\mathrm{h}} + \chi_T^{\mathrm{iv}}) = \mathsf{Var}(\chi_T^{\mathrm{iv}}) + 2\mathsf{Cov}(H_T^{\mathrm{h}}, \chi_T^{\mathrm{iv}}) + \mathsf{Var}(H_T^{\mathrm{h}}).$$

Then, combining (24) and (21) immediately gives (25). This completes the proof of Theorem 1. □

[15] It is straightforward to verify that any linear combination of admissible strategies is also admissible

EC.7.4.   Proof of Proposition 3

EC.7.4.1.   Verification of Assumption 5 First, we prove that under the parametric condition in (31), Assumption 5 holds.

LEMMA EC.5. *For X_t specified in (29), let Z_t follow the definition in (12) (then η_t involved in Z_t is defined in (30)). Under the parametric condition in (31) holds, i.e.,*

$$\kappa T < \frac{\pi}{4}.$$

Assumption 5 holds—that is, Z_t is a square-integrable martingale under P.

Proof. By definition of Z_t in (12), we have

$$dZ_t = -\eta_t Z_t dB_t = -Z_t \frac{\kappa}{\sigma}\Big(\alpha + \frac{\sigma^2}{2\kappa} - Y_t\Big)dB_t; \tag{EC.7.50}$$

the second equality uses the expression of η_t in (30).

Note that (EC.7.50) implies that Z_t is a local martingale under P. We will prove that it is a martingale as follows. Introduce the following process, which is a candidate for $\mathsf{E}(Z_T^2 \mid \mathcal{F}_t)$.

$$J_t := Z_t^2 f(t, Y_t) = Z_t^2 \exp\{f_0(T-t) + f_1(T-t)Y_t + f_2(T-t)Y_t^2\}; \tag{EC.7.51}$$

where f_i, $i = 0, 1, 2$ are functions defined in (34); and

$$f(t, y) := \exp\{f_0(T-t) + f_1(T-t)y + f_2(T-t)y^2\}. \tag{EC.7.52}$$

Since $f_i(0) = 0$ for all $i = 0, 1, 2$, we have $f(T, Y_t) = 0$ and thus $J_T = Z_T^2$. Below we show that properties of J_t guarantee Z_t to be a square-integrable martingale.

Applying Itô's lemma, it is straightforward to obtain the following dynamics:

$$\begin{aligned}
d\Big(\frac{J_t}{Z_t}\Big) &= dZ_t f(t, Y_t) = f(t, Y_t)dZ_t + Z_t df(t, Y_t) + dZ_t df(t, Y_t) \\
&= Z_t\Big(f_t - \frac{1}{2}\sigma^2 f_y + \frac{1}{2}\sigma^2 f_{yy}\Big)dt + Z_t(\sigma f_y - \eta_t f(t, Y_t))dB_t \\
&= Z_t\Big[f_t - \kappa\Big(\alpha + \frac{\sigma^2}{\kappa} - Y_t\Big)f_y + \frac{1}{2}\sigma^2 f_{yy} + \eta_t^2 f(t, Y_t)\Big]dt \\
&\quad - Z_t\Big[\eta_t f(t, Y_t) - \sigma f_y\Big]d\Big(dB_t + \eta_t dt\Big); \tag{EC.7.53}
\end{aligned}$$

f_t, f_y and f_{yy} are usual notations for partial derivatives; note they all depend on (t, Y_t), but for simplicity the arguments are dropped. The second line uses the expression of η_t in (30).

Next, applying Itô's lemma on J_t, it is straightforward to obtain:

$$dJ_t = Z_t d\Big(\frac{J_t}{Z_t}\Big) + \frac{J_t}{Z_t}dZ_t + d\Big(\frac{J_t}{Z_t}\Big)dZ_t$$

$$= Z_t^2 \Big[f_t - \kappa \big(\alpha + \frac{\sigma^2}{\kappa} - Y_t \big) f_y + \frac{1}{2} \sigma^2 f_{yy} + \eta_t^2 f(t, Y_t) \Big] dt$$
$$+ \Big(\sigma f_y - \eta_t f(t, Y_t) - \eta_t Z_t^2 f(t, Y_t) \Big) dB_t \tag{EC.7.54}$$

It can be verified that the function $f(t, y)$ defined in (EC.7.52) solves the following partial differential equation:

$$f_t - \kappa \big(\alpha + \frac{\sigma^2}{\kappa} - y \big) f_y + \frac{1}{2} \sigma^2 f_{yy} + \big(\frac{\kappa}{\sigma} \big)^2 \big(\alpha + \frac{\sigma^2}{2\kappa} - y \big)^2 f(t, y) = 0$$
$$\text{s.t.} \quad f(T, y) = 1, \quad \forall y \in \mathbb{R}. \tag{EC.7.55}$$

To check this, take derivatives of $f(t, y)$ using (EC.7.52). Then, collect the coefficients for y^2, y and the term independent of y, and set these coefficients to 0, then (EC.7.55) reduces to an ordinary differential equation system:

$$-f_2' + (2\kappa) f_2 + (2\sigma^2) f_2^2 + \big(\frac{\kappa}{\sigma} \big)^2 = 0$$
$$-f_1' + \kappa [f_1 - 2 \big(\alpha + \frac{\sigma^2}{\kappa} \big) f_2] + (2\sigma^2 f_2) f_1 - 2 \big(\frac{\kappa}{\sigma} \big)^2 \big(\alpha + \frac{\sigma^2}{2\kappa} \big) = 0$$
$$-f_0' - \kappa \big(\alpha + \frac{\sigma^2}{\kappa} \big) f_1 + \frac{1}{2} \sigma^2 (f_1^2 + 2 f_2) + \big(\frac{\kappa}{\sigma} \big)^2 \big(\alpha + \frac{\sigma^2}{2\kappa} \big)^2 = 0$$
$$\text{s.t.} \quad f_i(0) = 0, i = 0, 1, 2. \tag{EC.7.56}$$

It is straightforward to verify that under the parameter condition in (31), f_i, $i = 0, 1, 2$ specified in (34), solve this ODE system. So, $f(t, y)$ solves the PDE above, and this makes the dt-term of dJ_t in (EC.7.54) vanish, reducing (EC.7.54) to:

$$dJ_t = \Big[\sigma f_y - \eta_t f(t, Y_t) - \eta_t Z_t^2 f(t, Y_t) \Big] dB_t. \tag{EC.7.57}$$

And, (EC.7.53) reduces to (note the PDE expression is also involved in the dt-term on the second line of (EC.7.53)):

$$d \Big(\frac{J_t}{Z_t} \Big) = -Z_t \Big[\eta_t f(t, Y_t) - \sigma f_y \Big] d \big(dB_t + \eta_t dt \big). \tag{EC.7.58}$$

Clearly, by (EC.7.57), J_t is a local martingale under P.

Now, observe that the term inside the exponential of $f(t, Y_t)$ is a quadratic function in Y_t at each time t. Under (31), $f_2(T - t) \geq 0$ for all $t \in [0, T]$, hence

$$f(t, Y_t) \geq \exp \Big\{ f_0(T - t) - \frac{f_1^2(T - t)}{4 f_2(T - t)} \Big\}, \quad t \in [0, T);$$

It is easy to verify that $f_1^2(T - t) / f_2(T - t) \to 0$ as $t \to T$. Clearly, the function $f_0(T - t) - \frac{f_1^2(T-t)}{4f_2(T-t)}$ is continuous, hence admits a minimum on $[0, T]$; so, there exists a positive number, $c > 0$, such that $f(t, Y_t) \geq c > 0$ for all $t \in [0, T]$. Therefore,

$$J_t = Z_t^2 f(t, Y_t) \geq c Z_t^2 \Rightarrow Z_t \leq \sqrt{\frac{1}{c} J_t}. \tag{EC.7.59}$$

Next, we will make use of (EC.7.59) to prove Z_t is a square-integrable martingale under P. Define the following sequence of stopping times with respect to $\mathcal{G}_t$:

$$\tau_k := \inf\{t \in [0, T] \,|\, J_t \geq k\} \wedge T, \quad k \in \mathbb{N}. \tag{EC.7.60}$$

Clearly, J_t is a continuous process, hence $\tau_k \uparrow T$ as $k \to \infty$. Since each τ_k bounds the stopped version of J_t, hence also bounds the stopped version of Z_t via (EC.7.59). So, both $Z_{t \wedge \tau_k}$ and $J_{t \wedge \tau_k}$ are bounded P-martingales, and we can apply Doob's inequality:

$$\mathsf{E}\Big[\sup_{t \in [0,T]} Z^2_{t \wedge \tau_k}\Big] \leq 2\mathsf{E}(Z^2_{T \wedge \tau_k}) = 2\mathsf{E}(Z^2_{\tau_k}) \leq \frac{2}{c}\mathsf{E}(J_{\tau_k}) = \frac{2}{c}J_0. \tag{EC.7.61}$$

the first $\leq$ is application of Doob's inequality, and the following equality uses the obvious fact $\tau_k \leq T$. The second $\leq$ uses (EC.7.59) and the $=$ is application of optional stopping theorem on the bounded martingale $J_{t \wedge \tau_k}$. So, (EC.7.61) implies

$$\mathsf{E}\Big[\sup_{t \in [0,T]} Z^2_{t \wedge \tau_k}\Big] \leq \frac{2}{c}J_0;$$

clearly, because τ_k increases in k, so does $\sup_{t \in [0,T]} Z^2_{t \wedge \tau_k}$ (since the sup is taken on a longer time interval for larger τ_k), so we can let $k \to \infty$ and apply monotone convergence to above to reach the following, accounting for $\tau_k \uparrow T$:

$$\mathsf{E}\Big[\sup_{t \in [0,T]} Z^2_t\Big] \leq \frac{2}{c}J_0;$$

this is sufficient to establish that Z_t is a square-integrable martingale under P, and completes the proof of Lemma EC.5. $\square$

EC.7.4.2.   Proof of Proposition 3 Now that we know Z_t is a P-martingale, P^M specified in (12) is well-defined. Then, Girsanov Theorem applies and we have the P^M-Brownian-motion:

$$dB^M_t = dB_t - \eta_t dt.$$

Then, (EC.7.58) becomes

$$d\Big(\frac{J_t}{Z_t}\Big) = -Z_t\Big[\eta_t f(t, Y_t) - \sigma f_y\Big]dB^M_t; \tag{EC.7.62}$$

using the expression of $f(t, y)$ in (EC.7.52), it is straightforward to derive an explicit expression for f_y. Then, arrange the terms to explicitly write (EC.7.62) as:

$$d\Big(\frac{J_t}{Z_t}\Big) = -\sigma\Big(\frac{J_t}{Z_t}\Big)\Big[\frac{\kappa}{\sigma^2}b(T-t)(\alpha - Y_t) + a(T-t)\Big]dB^M_t; \tag{EC.7.63}$$

with the two deterministic functions $a(\cdot)$ and $b(\cdot)$ specified in (32); note $b(T-t) > 0$ for all $t \in [0, T]$. Next, observe that $V_t := \frac{\kappa}{\sigma^2}b(T-t)(\alpha - Y_t) + a(T-t)$ follows a linear stochastic differential equation of the following form:

$$dV_t = (x_t + y_t V_t)dt + z_t dB^M_t;$$

where x_t, y_t and z_t above are deterministic functions; the linearity comes from Y_t, which also follows a linear SDE (see (29)), as well as from η_t in (30), which is linear in Y_t. Now, Lemma A.4 in Wang and Wissel (2013) immediately applies, and we conclude that J_t/Z_t is a martingale under P^M. Next,

$$Z_t^M := \mathsf{E}^M(Z_T \,|\, \mathcal{F}_t) = \mathsf{E}^M \left(\frac{J_T}{Z_T} \,\Big|\, \mathcal{F}_t \right) = \frac{J_t}{Z_t} = Z_t f(t, Y_t). \tag{EC.7.64}$$

The first $=$ follows from $J_T = Z_T^2$, and the second one uses the established fact that J_t/Z_t is a martingale under P^M. The above can be written as:

$$\frac{Z_t}{Z_t^M} = \frac{1}{f(t, Y_t)} = \exp\{-f_0(T-t) - f_1(T-t)Y_t - f_2(T-t)Y_t^2\}. \tag{EC.7.65}$$

Now, (EC.7.63) can be written as:

$$\begin{aligned}
dZ_t^M &= -Z_t^M \left[\frac{\kappa}{\sigma^2} b(T-t)(\alpha - Y_t) + a(T-t) \right] \sigma dB_t^M \\
&= -\frac{Z_t^M}{X_t} \left[\frac{\kappa}{\sigma^2} b(T-t)(\alpha - Y_t) + a(T-t) \right] dX_t;
\end{aligned} \tag{EC.7.66}$$

where the second line uses $dX_t = \sigma X_t dB_t^M$. Hence, here the quantity ζ_t defined in (16) has the expression

$$\zeta_t = -\frac{Z_t^M}{X_t} \left[\frac{\kappa}{\sigma^2} b(T-t)(\alpha - Y_t) + a(T-t) \right]. \tag{EC.7.67}$$

Now we can apply Theorem 1 to establish Proposition 3. In particular, Z_t/Z_t^M involved in $B(m, P, R)$ expressed in (21) follows (EC.7.65), and ζ_t involved in θ_t^* specified in (20) follows (EC.7.67). This completes the proof. $\square$

EC.8. Proofs for Section 5
EC.8.1. Proof of Propositions 4 and 6

It is clear that both Propositions 4 and 6 immediately follow from the following lemma. Specifically, for Proposition 4 (resp., Proposition 6), $C_T = C_T^{(2)}$ and $C_T^M = C_T^{(1)}$ (resp., $C_T = C_T^{(1)}$ and $C_T^M = C_T^{(2)}$), where $C_T^{(1)}$ and $C_T^{(2)}$ are defined in Lemma EC.6 below.

LEMMA EC.6. *Let* $(P^{\mathrm{NV}}, R^{\mathrm{NV}})$ *be the solution to problem (2), i.e.,* $(P^{\mathrm{NV}}, R^{\mathrm{NV}}) = \arg\max_{P,R} \mathsf{E}[H_T(P, R)]$. *Then, as* C_T *increases stochastically, both* P^{NV} *and* R^{NV} *increase. Specifically, for any* $C_T^{(1)} \preceq C_T^{(2)}$, *the corresponding solutions to problem (2), denoted* $(P_i^{\mathrm{NV}}, R_i^{\mathrm{NV}})$ *for* $i = 1, 2$, *satisfy* $P_1^{\mathrm{NV}} \leq P_2^{\mathrm{NV}}$ *and* $R_1^{\mathrm{NV}} \leq R_2^{\mathrm{NV}}$. *In addition, let* $H_T^{(i)}$ *denote the production payoff with* $C_T = C_T^{(i)}$. *Then,* $\mathsf{E}[H_T^{(1)}(P, R)] \leq \mathsf{E}[H_T^{(2)}(P, R)]$ *for all* (P, R).

Proof. Let the distribution functions of $A_T^{(i)}$ be F_i. For a given P, let $R_i^{\mathrm{NV}}(P)$ be the corresponding profit-maximizing decision, i.e.,

$$R_i^{\mathrm{NV}}(P) := \arg\max_R \mathsf{E}[H_T(P, R)]$$

$$= \arg\max_R (P-c)(R-bP) - (P-s)\mathsf{E}[(R-A_T^{(i)})^+] = F_i^{-1}\left(\frac{P-c}{P-s}\right).$$

Clearly, $R_1^{\mathrm{NV}}(P) \le R_2^{\mathrm{NV}}(P)$ due to $A_T^{(1)} \preceq A_T^{(2)}$ (refer to Remark 1).

Define

$$H_i(P) = (P-c)(R-bP) - (P-s)\mathsf{E}[(R_i^{\mathrm{NV}}(P) - A_T^{(i)})^+], \quad i=1,2.$$

Then, the first-order derivative with respect to P is:

$$H_i'(P) = R_i^{\mathrm{NV}}(P) - 2bP + bc - \mathsf{E}[(R_i^{\mathrm{NV}}(P) - A_T^{(i)})^+], \quad i=1,2$$

For any given P, we show $H_1'(P) \le H_2'(P)$:

$$\begin{aligned}
H_1'(P) &= R_1^{\mathrm{NV}}(P) - 2bP + bc - \mathsf{E}[(R_1^{\mathrm{NV}}(P) - A_T^{(1)})^+] \\
&\le R_1^{\mathrm{NV}}(P) - 2bP + bc - \mathsf{E}[(R_1^{\mathrm{NV}}(P) - A_T^{(2)})^+] \\
&= \mathsf{E}[R_1^{\mathrm{NV}}(P) \wedge A_T^{(2)}] - 2bP + bc \\
&\le \mathsf{E}[R_2^{\mathrm{NV}}(P) \wedge A_T^{(2)}] - 2bP + bc \\
&= R_2^{\mathrm{NV}}(P) - 2bP + bc - \mathsf{E}[(R_2^{\mathrm{NV}}(P) - A_T^{(2)})^+] = H_2'(P).
\end{aligned}$$

The first inequality is due to $A_T^{(1)} \preceq A_T^{(2)}$ and the second one follows from $R_1^{\mathrm{NV}}(P) \le R_2^{\mathrm{NV}}(P)$. Suppose P_i^{NV} is the one of the larger zeros to $H_i'(P) = 0$. (Recall, $H_i'(P) = 0$ has two solutions and the lager one is the newsvendor's solution; see Section EC.6.1.) Then,

$$0 = H_1'(P_1^{\mathrm{NV}}) \le H_2'(P_1^{\mathrm{NV}}),$$

thus $H_2'(P)$ stays nonnegative before P exceeds P_1^{NV}, indicating $P_2^{\mathrm{NV}} \ge P_1^{\mathrm{NV}}$. Furthermore,

$$R_1^{\mathrm{NV}}(P_1^{\mathrm{NV}}) = F_1^{-1}\left(\frac{P_1^{\mathrm{NV}}-c}{P_1^{\mathrm{NV}}-s}\right) \le F_2^{-1}\left(\frac{P_1^{\mathrm{NV}}-c}{P_1^{\mathrm{NV}}-s}\right) \le F_2^{-1}\left(\frac{P_2^{\mathrm{NV}}-c}{P_2^{\mathrm{NV}}-s}\right) = R_2^{\mathrm{NV}}(P_2^{\mathrm{NV}}) = R_2^{\mathrm{NV}}.$$

The "in addition" part follows immediately from the fact that $H_T(P,R)$ is an increasing function in A_T for any (P,R), and $A_T^{(1)} \preceq A_T^{(2)}$. This completes the proof. $\qquad\square$

EC.8.2. Proof of Propositions 5 and 7

We first prove Proposition 7 and, as an intermediate step, establish Lemma EC.7. Based on this lemma, we further prove Proposition 5.

Proof of Proposition 7 Denote the second term of $B(m,P,R)$ as $\Psi(P,R)$. Clearly, $\Psi(P,R)$ strictly increases in both P and R. Then, write

$$B(m,P,R) = C[(m - V_0(P,R))]^2 + \Psi(P,R),$$

where $C = 1/(Z_0^M - 1) > 0$. (Below, we refer to $C[(m - V_0(P, R))]^2$ (resp., $\Psi(P, R)$) as the "first term" (resp., "second term") of $B(m, P, R)$.) Note that (P_m^h, R_m^h) satisfy the optimality equations

$$2C(V_0 - m)\frac{\partial V_0(P_m^h, R_m^h)}{\partial P} + \frac{\partial \Psi(P_m^h, R_m^h)}{\partial P} = 0, \quad 2C(V_0 - m)\frac{\partial V_0(P_m^h, R_m^h)}{\partial R} + \frac{\partial \Psi(P_m^h, R_m^h)}{\partial R} = 0.$$

Since $\Psi(P, R)$ is strictly increasing in both P and R, we must have

$$(V_0 - m)\frac{\partial V_0(P_m^h, R_m^h)}{\partial P} < 0, \quad (V_0 - m)\frac{\partial V_0(P_m^h, R_m^h)}{\partial R} < 0. \tag{EC.8.1}$$

Recall, $V_0(P, R)$ is concave in P (resp., R) for any given R (resp., P). We use the notation in (EC.8.9):

$$P^{\mathrm{NV(M)}}(R) := \arg\max_P V_0(P, R), \quad R^{\mathrm{NV(M)}}(P) := \arg\max_R V_0(P, R),$$

and it is straightforward to derive:

$$P^{\mathrm{NV(M)}}(R) = \frac{bc + \mathsf{E}^M(R \wedge A_T)}{2b}, \quad R^{\mathrm{NV(M)}}(P) = F_M^{-1}\left(\frac{P - c}{P - s}\right); \tag{EC.8.2}$$

where F_M is the distribution function of A_T under P^M. Clearly, $P^{\mathrm{NV(M)}}(R)$ (resp., $R^{\mathrm{NV(M)}}(P)$) increases in R (resp., P). Moreover, by its partial concavity in P (resp., R), for given R (resp., P), $V_0(P, R)$ increases in P (resp., R) for $P \leq P^{\mathrm{NV(M)}}(R)$ (resp., $R \leq R^{\mathrm{NV(M)}}(P)$) and then decreases. Thus, for any given R, the corresponding P that minimizes $B(m, P, R)$ cannot exceed $P^M(R)$. Otherwise, a P smaller than $P^{\mathrm{NV(M)}}(R)$, P', can be found to satisfy $V_0(P, R) = V_0(P', R)$. Then, $B(m, P', R) < B(m, P, R)$, contradicting with P minimizing $B(m, P, R)$. Completely analogous argument can be applied to R. Applying the above to optimality of (P_m^h, R_h^m), we must have:

LEMMA EC.7. $P_m^h \leq P^{\mathrm{NV(M)}}(R_m^h), \quad R_m^h \leq R^{\mathrm{NV(M)}}(P_m^h)$.

Clearly, Lemma EC.7 implies:

$$\frac{\partial V_0(P_m^h, R_m^h)}{\partial P} \geq 0, \quad \frac{\partial V_0(P_m^h, R_m^h)}{\partial R} \geq 0. \tag{EC.8.3}$$

Combining (EC.8.3) with (EC.8.1), we conclude $V_0(P_m^h, R_m^h) \leq m$. This proves Proposition 7.

Proof of Proposition 5 $P_m^h \leq P^{\mathrm{NV(M)}}$ and $R_m^h \leq R^{\mathrm{NV(M)}}$. First, we consider the case of $m \geq V_0(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$. For this case, P_m^h and R_m^h cannot both exceed $P^{\mathrm{NV(M)}}$ and $R^{\mathrm{NV(M)}}$, respectively; otherwise both terms of $B(m, P_m^h, R_m^h)$ will exceed those of $B(m, P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$, contradicting optimality of (P_m^h, R_m^h). Then, if $P_m^h \leq P^{\mathrm{NV(M)}}$, by Lemma EC.7,

$$R_m^h \leq R^{\mathrm{NV(M)}}(P_m^h) \leq R^{\mathrm{NV(M)}}(P^{\mathrm{NV(M)}}) = R^{\mathrm{NV(M)}}.$$

If $R_m^h \leq R^{\mathrm{NV(M)}}$, again by Lemma EC.7,

$$P_m^h \leq P^{\mathrm{NV(M)}}(R_m^h) \leq P^{\mathrm{NV(M)}}(R^{\mathrm{NV(M)}}) = P^{\mathrm{NV(M)}}.$$

Hence, for both cases, we must have $P_m^h \leq P^{\mathrm{NV(M)}}$ and $R_m^h \leq R^{\mathrm{NV(M)}}$.

Now, we consider the other case, $m < V_0(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}})$. Let $v(P) := V_0(P, R^{\mathrm{NV(M)}}(P))$. Clearly, $v(c) \leq 0 \leq m$. On the other hand,

$$v(P^{\mathrm{NV(M)}}) = V_0(P^{\mathrm{NV(M)}}, R^M(P^{\mathrm{NV(M)}})) = V_0(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}}) > m.$$

Thus, $\exists P_1 \in [c, P^{\mathrm{NV(M)}}]$ such that $v(P_1) = m$; let $R_1 = R^{\mathrm{NV(M)}}(P_1)$. Note, $R_1 = R^{\mathrm{NV(M)}}(P_1) \leq R^{\mathrm{NV(M)}}(P^{\mathrm{NV(M)}}) = R^{\mathrm{NV(M)}}$. Then, P_m^h and R_m^h cannot both exceed, respectively, P_1 and R_1, otherwise both terms of $B(m, P_m^h, R_m^h)$ are larger than those of $B(m, P_1, R_1)$ (the first term of which is zero). Then, if $P_m^h \leq P_1 (\leq P^{\mathrm{NV(M)}})$, applying Lemma EC.7,

$$R_m^h \leq R^{\mathrm{NV(M)}}(P_m^h) \leq R^{\mathrm{NV(M)}}(P_1) \leq R^{\mathrm{NV(M)}}(P^{\mathrm{NV(M)}}) = R^{\mathrm{NV(M)}}.$$

If $R_m^h \leq R_1 (\leq R^{\mathrm{NV(M)}})$, then again by Lemma EC.7,

$$P_m^h \leq P^{\mathrm{NV(M)}}(R_m^h) \leq P^{\mathrm{NV(M)}}(R_1) \leq P^{\mathrm{NV(M)}}(R^{\mathrm{NV(M)}}) = P^{\mathrm{NV(M)}}.$$

Then, for both cases, we must also have $P_m^h \leq P^{\mathrm{NV(M)}}$ and $R_m^h \leq R^{\mathrm{NV(M)}}$. Summarizing all above leads to $P_m^h \leq P^{\mathrm{NV(M)}}$ and $R_m^h \leq R^{\mathrm{NV(M)}}$. This proves Proposition 5.

EC.8.3.   Proof of Theorem 3

We layout key definitions and technical preparations in Section EC.8.3.1 and then prove Theorem 3 in Section EC.8.3.

EC.8.3.1.   Definitions and Technical Preparations Throughout the proof, m denotes the newsvendor's maximum expected profit, i.e.,

$$m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] > 0.$$

(The > 0 is due to Assumption 2.) Also recall,

$$P^{\mathrm{NV(M)}}(R) := \arg\max_P V_0(P, R) = \frac{\mathsf{E}^M(R \wedge A_T) + bc}{2b},$$
$$R^{\mathrm{NV(M)}}(P) := \arg\max_R V_0(P, R) = F_M^{-1}\left(\frac{P-c}{P-s}\right).$$

Related to m, below we define three critical values.

DEFINITION EC.1. (i) P_1 is the smallest solution to $\mathsf{E}^M[H_T(P_1, R^{\mathrm{NV(M)}}(P_1))] = m$ and $R_1 =: R^{\mathrm{NV(M)}}(P_1)$. (ii) P_2 is the solution to $R^{\mathrm{NV(M)}}(P_2) = R^{\mathrm{NV}}$. (iii) P_3 is the smallest solution to $V_0(P_3, R^{\mathrm{NV}}) = m$.

The lemma below collects properties of P_i, $i = 1, 2, 3$.

LEMMA EC.8. *All P_1, P_2 and P_3 exist, hence by definition they are unique. Furthermore,*

$$P_2 = \frac{c - s\mathsf{P}^M(A_T \leq R^{\mathrm{NV}})}{1 - P^M(A_T \leq R^{\mathrm{NV}})} = s + \frac{c - s}{\mathsf{P}^M(A_T \geq R^{\mathrm{NV}})}. \tag{EC.8.4}$$

And

$$P_3 = P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) - \sqrt{(P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) - P^{\mathrm{NV}})(P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) + P^{\mathrm{NV}} - 2s)}. \tag{EC.8.5}$$

Moreover,

$$P_i \leq P^{\mathrm{NV}}, \quad i = 1, 2, 3. \tag{EC.8.6}$$

Proof. We first show P_1 exists. (Existence of P_2 and P_3 will be clear after (EC.8.4) and (EC.8.5) are proved.) Recall,

$$V_0(P, R) = (P - c)(R - bP) - (P - s)\mathsf{E}^M[(R - A_T)^+].$$

As $P \to c$, $R^{\mathrm{NV(M)}}(P) = F_M^{-1}((P - c)/(P - s)) \to -\infty$. Thus, as $P \to c$, $(P - c)(R - bP)$ is eventually nonpositive (hence < m) and so is $V_0(P, R^{\mathrm{NV(M)}}(P))$. Now, we check $V_0(P^{\mathrm{NV}}, R^{\mathrm{NV(M)}}(P^{\mathrm{NV}}))$:

$$\begin{aligned}
V_0(P^{\mathrm{NV}}, R^{\mathrm{NV(M)}}(P^{\mathrm{NV}})) &= \max_R V_0(P^{\mathrm{NV}}, R) \\
&\geq V_0(P^{\mathrm{NV}}, R^{\mathrm{NV}}) \\
&= (P^{\mathrm{NV}} - c)(R^{\mathrm{NV}} - bP^{\mathrm{NV}}) - (P^{\mathrm{NV}} - s)\mathsf{E}^M[(R^{\mathrm{NV}} - A_T)^+] \\
&\geq (P^{\mathrm{NV}} - c)(R^{\mathrm{NV}} - bP^{\mathrm{NV}}) - (P^{\mathrm{NV}} - s)\mathsf{E}[(R^{\mathrm{NV}} - A_T)^+] \\
&= \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] = m.
\end{aligned}$$

The second $\leq$ is due to $A_T \preceq A_T^M$. Summarizing the above, there exists value(s) of P in $[c, P^{\mathrm{NV}}]$ such that $V_0(P, R^{\mathrm{NV(M)}}(P)) = m$, hence P_1 uniquely exists and in particular, $P_1 \leq P^{\mathrm{NV}}$.

For P_2, by definition:

$$R^{\mathrm{NV(M)}}(P_2) = F_M^{-1}\left(\frac{P_2 - c}{P_2 - s}\right) = R^{\mathrm{NV}},$$

which immediately leads to the expression in (EC.8.4). Furthermore,

$$P_2 = s + \frac{c - s}{\mathsf{P}^M(A_T \geq R^{NV})} \leq s + \frac{c - s}{\mathsf{P}(A_T \geq R^{NV})} = P^{\mathrm{NV}}(R^{\mathrm{NV}}) = P^{\mathrm{NV}}.$$

The $\leq$ is due to $A_T \preceq A_T^M$.

For P_3, note that given $R = R^{\mathrm{NV}}$, $V_0(P, R^{\mathrm{NV}}) = m$ is a quadratic equation in P. Rearranging the terms and accounting for the following:

$$P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) = \frac{\mathsf{E}^M(R^{\mathrm{NV}} \wedge A_T) + bc}{2b},$$

the equation becomes

$$-bP^2 + 2bP^{\mathrm{NV(M)}}(R^{\mathrm{NV}})P - (c - s)R^{\mathrm{NV}} - 2bsP^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) + bcs = m. \tag{EC.8.7}$$

Using the optimality equations specified in Proposition 1, it is straightforward to derive the following:

$$m = b(P^{\mathrm{NV}})^2 - (c-s)R^{\mathrm{NV}} - 2bsP^{\mathrm{NV}} + bcs.$$

Substituting this expression of m to (EC.8.7), the equation becomes:

$$-P^2 + 2P^{\mathrm{NV(M)}}(R^{\mathrm{NV}})P - [(P^{\mathrm{NV}})^2 + 2sP^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) - 2sP^{\mathrm{NV}}] = 0. \tag{EC.8.8}$$

Then, it is straightforward to verify that (EC.8.5) is the smallest solution to (EC.8.8). In particular, by $A_T \preceq A_T^M$,

$$P^{\mathrm{NV}} = \frac{1}{2b}[\mathsf{E}(R^{\mathrm{NV}} \wedge A_T) + bc] \le \frac{1}{2b}[\mathsf{E}^M(R^{\mathrm{NV}} \wedge A_T) + bc] \le = P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}).$$

In addition, by Assumption 2, $P^{\mathrm{NV}} > c > s$, hence $P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) > c > s$. Therefore, P_3 in (EC.8.5) is well-defined (i.e., real-valued). Next, using the fact that the function $x - \sqrt{(x-y)(x+y-2s)}$ decreases in x for $x \ge y \ge s$, $P_3 \le P^{\mathrm{NV}}$ immediately follows from $P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) \ge P^{\mathrm{NV}}$. $\qquad\square$

The next result indicates that R_1 upper bounds R_m^h.

LEMMA EC.9. $R_m^h \le R_1$.

Proof. If $R_m^h > R_1$ and $P_m^h > P_1$, then $B(m, P_1, R_1) < B(m, P_m^h, R_m^h)$ since the both the first term (which is zero) and the second term of the former are smaller than the latter. This contradicts with optimality of (P_m^h, R_m^h). Therefore $R_m^h > R_1$ and $P_m^h > P_1$ cannot hold at the same time. If $R_m^h > R_1$, we must have $P_m^h \le P_1$, but this introduces contradiction because by Lemma EC.7, we must also have

$$R_m^h \le R^{\mathrm{NV(M)}}(P_m^h) \le R^{\mathrm{NV(M)}}(P_1) = R_1.$$

Therefore, $R_m^h \le R_1$ must hold. $\qquad\square$

Finally, we need the following definition and lemma.

$$P^{\mathrm{NV(M)}}(R) := \arg\max_P V_0(P,R), \quad R^{\mathrm{NV(M)}}(P) := \arg\max_R V_0(P,R). \tag{EC.8.9}$$

Based on (EC.8.9), we define a critical point, (P^*, R^*):

$$R^* \text{ is the smallest root of } V_0(P^{\mathrm{NV(M)}}(R), R) = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})], \quad P^* := P^{\mathrm{NV(M)}}(R^*). \tag{EC.8.10}$$

LEMMA EC.10. *Suppose Assumptions 1–3 hold and $C_T \preceq C_T^M$. Then, (P^*, R^*) defined in (EC.8.10) exists and satisfies*

$$P^* \le P^{\mathrm{NV}} \quad and \quad R^* \le R^{\mathrm{NV}}.$$

Proof. Let $f(R) := V_0(P^{\mathrm{NV(M)}}(R), R)$. For notational simplicity, let $P(R) = P^{\mathrm{NV(M)}}(R)$ and $m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})]$. Given R, $V_0(P, R)$ is concave in P, hence by setting $\partial V_0/\partial P$ to zero, $P(R)$ satisfies the following optimality equation:

$$2bP(R) = \mathsf{E}^M[R \wedge A_T] + bc. \tag{EC.8.11}$$

Then, it is straightforward to verify:

$$f(R) = b(P(R))^2 - 2bsP(R) - (c-s)R + bcs.$$

Next, we will show $f(R^{\mathrm{NV}}) \geq m$ and $f(bc) \leq m$, which in turn indicates that $f(R)$ has at least one root within $[bc, R^{\mathrm{NV}}]$ and thus proves the result. First, examine $f(R^{\mathrm{NV}})$. With $C_T \preceq C_T^M$, by independence of $\tilde{B}_T$ from $\{B_t, 0 \leq t \leq T\}$, by Lemma EC.6 we have $A_T \preceq A_T^M$ (A_T^M is the version of A_T under P^M: $\mathsf{P}(A_T^M \leq a) = \mathsf{P}^M(A_T \leq a)$ for all a.). Then,

$$2bP(R^{\mathrm{NV}}) = \mathsf{E}^M[R^{\mathrm{NV}} \wedge A_T] + bc \geq \mathsf{E}[R^{\mathrm{NV}} \wedge A_T] + bc = 2bP^{\mathrm{NV}},$$

where the last equality follows from (4), leading to $P(R^{\mathrm{NV}}) \geq P^{\mathrm{NV}}$. Then, by the first optimality equation in (4), it is straightforward to verify the following:

$$m = b(P^{\mathrm{NV}})^2 - 2bsP^{\mathrm{NV}} - (c-s)R^{\mathrm{NV}} + bcs.$$

Analogously, it can be verified that

$$f(R^{\mathrm{NV}}) = b(P(R^{\mathrm{NV}}))^2 - 2bsP(R^{\mathrm{NV}}) - (c-s)R^{\mathrm{NV}} + bcs.$$

Since $P(R^{\mathrm{NV}}) \geq P^{\mathrm{NV}} \geq s$, we have $f(R^{\mathrm{NV}}) \geq m$.

Next, we check $f(bc)$. Let $\epsilon = \mathsf{E}^M[(bc - A_T)^+]$, then clearly $\epsilon \leq [\mathsf{E}[(bc - A_T)^+]]$. From Assumption 3, we have

$$\frac{[\mathsf{E}[(bc - A_T)^+]]^2}{4bm} \leq 1.$$

Note $P(bc) = [\mathsf{E}^M(bc \wedge A_T) + bc]/2b = c - \epsilon/2b$, then it is straightforward to verify that:

$$f(bc) < \frac{\epsilon^2}{4b}.$$

Thus,

$$\frac{f(bc)}{m} < \frac{\epsilon^2}{4bm} \leq \frac{[\mathsf{E}[(bc - A_T)^+]]^2}{4bm} \leq 1 \quad \Rightarrow \quad f(bc) < m.$$

Combining the above, $f(R)$ must have root(s) within $[bc, R^{\mathrm{NV}}]$, hence by definition of R^*, $R^* \leq R^{\mathrm{NV}}$.

Next, we proceed with showing $P^* \leq P^{\mathrm{NV}}$. Using (EC.8.11), it is easy to verify that

$$(m =) \quad f(R^*) = b(P^*)^2 - 2bsP^* - (c-s)R^* + bcs.$$

Comparing with the expression of m above and taking into account $R^* \leq R^{\mathrm{NV}}$, clearly, $P^* \leq P^{\mathrm{NV}}$ must hold. This completes the proof. □

EC.8.3.2. **Proof of Theorem 3** Now we are ready to prove Theorem 3.

Proof of Part (i): By Lemma EC.10, $P^* \leq P^{\mathrm{NV}}$, so it is sufficient to prove $P_m^h \leq P^*$. To this end, we consider two cases, $R_m^h \leq R^*$ or $R_m^h \geq R^*$. For the first case, apply Lemma EC.7, we have

$$P_m^h \leq P^{\mathrm{NV(M)}}(R_m^h) \leq P^{\mathrm{NV(M)}}(R^*) = P^*.$$

Now, consider the other case, $R_m^h \geq R^*$. Note that P_m^h and R_m^h cannot exceed P^* and R^* at the same time: the first term of $B(m, P^*, R^*)$ is zero (since $V_0(P^*, R^*) = m$) and thus increasing both P and R beyond P^* and R^* increases both terms $B(m, P, R)$. Therefore, $R_m^h \geq R^*$ must lead to $P_m^h \leq P^*$.

Proof of Part (ii): Applying Lemma EC.8, it is straightforward to verify that the stated inequality in (36) is equivalent to $P_3 \leq P_2$. By Lemma EC.9, it is sufficient to prove $R_1 \leq R^{\mathrm{NV}}$ under this condition. To this end, first note that

$$P^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) = \frac{1}{2b}[\mathsf{E}^M(R^{\mathrm{NV}} \wedge A_T) + bc] \geq \frac{1}{2b}[\mathsf{E}(R^{\mathrm{NV}} \wedge A_T) + bc] = P^{\mathrm{NV}}.$$

The $\geq$ is due to $A_T \preceq A_T^M$. Combining this with (EC.8.6) and accounting for the concavity of $V_0(P, R)$ in P with a given R, $P_3 \leq P_2$ leads to

$$V_0(P_1, R_1) = m = V_0(P_3, R^{\mathrm{NV}}) \leq V_0(P_2, R^{\mathrm{NV}}).$$

By $V_0(P_1, R_1) = \max_R V_0(P_1, R) \geq V_0(P_1, R^{\mathrm{NV}})$, we have:

$$V_0(P_1, R^{\mathrm{NV}}) \leq V_0(P_2, R^{\mathrm{NV}}).$$

Applying $P_1, P_2 \leq P^{\mathrm{NV(M)}}(R^{\mathrm{NV}})$ and concavity of $V_0(P, R)$ in P with given R again, we have

$$P_1 \leq P_2.$$

Therefore,

$$R_1 = R^{\mathrm{NV(M)}}(P_1) \leq R^{\mathrm{NV(M)}}(P_2) = R^{\mathrm{NV}},$$

which completes the proof of Part (ii).

Proof of Part (iii): Violation of (36) is equivalent to $P_2 < P_3$. Reversing the argument in proof of Part (ii) above, it is easy to obtain $P_1 > P_2$, hence $R_1 = R^{\mathrm{NV(M)}}(P_1) > R^{\mathrm{NV(M)}}(P_2) = R^{\mathrm{NV}}$. Next,

$$V_0(P_1, R_1) = \max_R V_0(P_1, R) = m = V_0(P_3, R^{\mathrm{NV}}) \quad \Rightarrow \quad V_0(P_1, R^{\mathrm{NV}}) \leq V_0(P_3, R^{\mathrm{NV}}).$$

Again, by concavity of $V_0(P, R^{\mathrm{NV}})$ in P and $P_1, P_3 \leq P^{\mathrm{NV}} \leq P^{\mathrm{NV(M)}}(R^{\mathrm{NV}})$, we have $P_1 \leq P_3$. In summary, we have

$$c < P_2 < P_1 \leq P_3, \quad R_1 > R^{\mathrm{NV}}. \tag{EC.8.12}$$

$(P_2 > c$ is obvious from (EC.8.4.)

Next, rearranging the terms of $V_0(P,R)$ leads to:

$$V_0(P,R) = (-bP^2 + bcP) + (P-s)\mathsf{E}^M(R \wedge A_T) - (c-s)R.$$

Then, $m = V_0(P_1, R_1) = V_0(P_3, R^{\mathrm{NV}})$ leads to:

$$(-bP_1^2 + bcP_1) - (-bP_3^2 + bcP_3) + (P_1 - s)\mathsf{E}^M[R_1 \wedge A_T] - (c-s)R_1$$
$$= (P_3 - s)\mathsf{E}^M[R^{\mathrm{NV}} \wedge A_T] - (c-s)R^{\mathrm{NV}}.$$

Since the function $-bP^2 + bcP$ decreases in $P \geq c/2$ and $c < P_1 \leq P_3$, we have $(-bP_1^2 + bcP_1) - (-bP_3^2 + bcP_3) \geq 0$. Then, the equality above implies:

$$(P_1 - s)\mathsf{E}^M[R_1 \wedge A_T] - (c-s)R_1 \leq (P_3 - s)\mathsf{E}^M[R^{\mathrm{NV}} \wedge A_T] - (c-s)R^{\mathrm{NV}}. \qquad \text{(EC.8.13)}$$

The LHS of (EC.8.13) is:

$$\mathsf{E}^M(R_1 \wedge A_T) = R_1\mathsf{P}^M(A_T \geq R_1) + \mathsf{E}^M[A_T\mathbf{1}\{A_T \leq R_1\}] = R_1\frac{c-s}{P_1 - s} + \mathsf{E}^M[A_T\mathbf{1}\{A_T \leq R_1\}],$$

where the equality is due to definition of R_1: $\mathsf{P}^M(A_T \leq R_1) = (P_1 - c)/(P_1 - s)$, substituting the above expression of $\mathsf{E}^M[R_1 \wedge A_T]$ to the LHS of (EC.8.13), we have:

$$\text{LHS of (EC.8.13)} = R_1(c-s) + (P_1 - s)\mathsf{E}^M[A_T\mathbf{1}\{A_T \leq R_1\}] - (c-s)R_1$$
$$= (P_1 - s)\mathsf{E}^M[A_T\mathbf{1}\{A_T \leq R_1\}]$$
$$= \mathsf{E}^M[A_T\mathbf{1}\{A_T \leq R_1\}]\frac{c-s}{\mathsf{P}^M(A_T \geq R_1)},$$

where the last equality is due to: $\mathsf{P}^M(A_T \geq R_1) = (c-s)/(P_1 - s)$.

Now, we focus on the RHS of (EC.8.13). We first show that it is positive. Using $V_0(P_3, R^{\mathrm{NV}}) = m > 0$ and rearranging terms of $V_0(P_3, R^{\mathrm{NV}})$, we have:

$$0 < m = V_0(P_3, R^{\mathrm{NV}}) = (-bP_3)(P_3 - c) + (P_3 - s)\mathsf{E}^M(R^{\mathrm{NV}} \wedge A_T) - (c-s)R^{\mathrm{NV}}.$$

Note $P_3 > P_2 > c$ and thus $(-bP_3)(P_3 - c) < 0$, which implies

$$\text{RHS of (EC.8.13)} = (P_3 - s)\mathsf{E}^M(R^{\mathrm{NV}} \wedge A_T) - (c-s)R^{\mathrm{NV}} > 0.$$

Next, we have:

$$\mathsf{E}^M[R^{\mathrm{NV}} \wedge A_T] = 2bP^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) - bc = 2b\bar{P}^{\mathrm{NV(M)}}(R^{\mathrm{NV}}) + 2bs - bc = 2b\bar{P}^{\circ} + 2bs - bc,$$

where the last equality follows from the definition of $\bar{P}^\circ$ in (35). By (EC.8.5) and definition of r° in (35), it is straightforward to verify the following:

$$\bar{P}_3 \equiv P_3 - s = \frac{\bar{P}^{\text{NV}}}{r^\circ + \sqrt{(r^\circ)^2 - 1}}.$$

Therefore, the RHS of (EC.8.13) is rewritten as:

$$\text{RHS of (EC.8.13)} = \frac{\bar{P}^{\text{NV}}(2b\bar{P}^\circ + 2bs - bc)}{r^\circ + \sqrt{(r^\circ)^2 - 1}} - (c - s)R^{\text{NV}} \ (> 0).$$

Combining the above and dividing both sides by $c - s$, the inequality (EC.8.13) is equivalent to:

$$\frac{\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R_1\}]}{\mathsf{P}^M(A_T \geq R_1)} \leq \frac{\bar{P}^{\text{NV}}(2b\bar{P}^\circ + 2bs - bc)}{(c - s)(r^\circ + \sqrt{(r^\circ)^2 - 1})} - R^{\text{NV}}. \tag{EC.8.14}$$

In particular, the RHS of (EC.8.14) is positive.

Next, we show that (EC.8.14) produces an upper bound on R_1, R°, hence by Lemma EC.9 also bounds R_m^h. Define the following function in R:

$$T(R) = \frac{\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R\}]}{\mathsf{P}^M(A_T \geq R)}.$$

Taking derivative of $T(R)$ leads to:

$$T'(R) = \frac{f^M(R)}{[\mathsf{P}^M(A_T \geq R)]^2}\mathsf{E}^M(R \wedge A_T),$$

where $f^M(r)$ is the probability density function of A_T under P^M. Note, by $A_T \succeq A_T^M$ and Assumption 2,

$$\mathsf{E}^M(R^{\text{NV}} \wedge A_T) \geq \mathsf{E}(R^{\text{NV}} \wedge A_T) = 2bP^{\text{NV}} - bc > 2bc - bc = bc > 0.$$

So, clearly $T'(R) > 0$ for $R \geq R^{\text{NV}}$. Thus, $T(R)$ strictly increases in R for $R \in [R^{\text{NV}}, \infty)$.

Next, we show $R(R^{\text{NV}})$ is smaller than RHS of (EC.8.14). There are two cases: $\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\text{NV}}\}] < 0$ or $\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\text{NV}}\}] \geq 0$. For the first case, clearly $T(R^{\text{NV}}) < 0$ and thus smaller than RHS of (EC.8.14). Now, suppose the other case holds, i.e., $\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\text{NV}}\}] \geq 0$, then we have:

$$\begin{aligned}
T(R^{\text{NV}}) &= \frac{\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\text{NV}}\}]}{\mathsf{P}^M(A_T \geq R^{\text{NV}})} \leq \frac{P_3 - s}{c - s}\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\text{NV}}\}] \\
&= \frac{P_3 - s}{c - s}\left(\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\text{NV}}\}] + R^{\text{NV}}\frac{c - s}{P_3 - s}\right) - R^{\text{NV}} \\
&\leq \frac{P_3 - s}{c - s}\left(\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\text{NV}}\}] + \mathsf{P}^M(A_T \geq R^{\text{NV}})R^{\text{NV}}\right) - R^{\text{NV}} \\
&= \frac{P_3 - s}{c - s}\mathsf{E}^M[R^{\text{NV}} \wedge A_T] - R^{\text{NV}} = \frac{\bar{P}^{\text{NV}}(2b\bar{P}^\circ + 2bs - bc)}{(c - s)(r^\circ + \sqrt{(r^\circ)^2 - 1})} - R^{\text{NV}} \\
&= \text{RHS of (EC.8.14)}.
\end{aligned}$$

The two $\leq$ involved in the derivations above uses the fact $\mathsf{P}^M(A_T \leq R^{\mathrm{NV}}) \leq (P_3 - c)/(P_3 - s)$, which follows from (based on the proved fact that $R_1 \geq R^{\mathrm{NV}}$ and $P_3 \geq P_1$):

$$\mathsf{P}^M(A_T \leq R^{\mathrm{NV}}) \leq \mathsf{P}^M(A_T \leq R_1) = \frac{P_1 - c}{P_1 - s} \leq \frac{P_3 - c}{P_3 - s}, \quad \text{and thus} \quad \mathsf{P}^M(A_T \geq R^{\mathrm{NV}}) \geq \frac{c - s}{P_3 - s}.$$

Combining both cases, we have:

$$T(R^{\mathrm{NV}}) \leq \text{RHS of (EC.8.14)}.$$

Furthermore, since $\mathsf{E}^M(A_T) \geq \mathsf{E}^M(R^{\mathrm{NV}} \wedge A_T) \geq \mathsf{E}(R^{\mathrm{NV}} \wedge A_T) = 2bP^{\mathrm{NV}} - bc > 0$, clearly $T(R) \to \infty$ as $R \to \infty$. Recall, $T(R)$ strictly increases in R for $R \geq R^{\mathrm{NV}}$, the analysis above indicates that there exist a unique $R^\circ \geq R^{\mathrm{NV}}$ such that:

$$T(R^\circ) = \frac{\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^\circ\}]}{\mathsf{P}^M(A_T \geq R^\circ)} = \text{RHS of (EC.8.14)} = \frac{P^{\bar{\mathrm{N}}\mathrm{V}}(2b\bar{P}^\circ + 2bs - bc)}{(c - s)(r^\circ + \sqrt{(r^\circ)^2 - 1})} - R^{\mathrm{NV}}.$$

Combining the proved fact that $R_1 \geq R^{\mathrm{NV}}$ with $T(R_1) \leq \text{RHS of (EC.8.14)}$, $R_1 \leq R^\circ$. Applying Lemma EC.9, $R_m^h \leq R_1 \leq R^\circ$ immediately follows.

What remains is to show $R^\circ \leq R^{\mathrm{NV(M)}}$. Let $R = R^{\mathrm{NV(M)}}$ and use $\mathsf{P}^M(A_T \geq R^{\mathrm{NV(M)}}) = (c - s)/(P^{\mathrm{NV(M)}} - s)$:

$$T(R^{\mathrm{NV(M)}}) = \frac{\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\mathrm{NV(M)}}\}]}{\mathsf{P}^M(A_T \geq R^{\mathrm{NV(M)}})} = \frac{P^{\mathrm{NV(M)}} - s}{c - s} \mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\mathrm{NV(M)}}\}]. \tag{EC.8.15}$$

Here we show $\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\mathrm{NV(M)}}\}] > 0$ (hence $T(R^{\mathrm{NV(M)}}) > 0$) as follows. First note that

$$0 < m = \mathsf{E}[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] \leq \mathsf{E}^M[H_T(P^{\mathrm{NV}}, R^{\mathrm{NV}})] \leq \max_{P,R} V_0(P, R) = V_0(P^{\mathrm{NV}}, R^{\mathrm{NV}})$$

(the first $\leq$ is due to $A_T \succeq A_T^M$). Next,

$$\begin{aligned}
V_0(P^{\mathrm{NV(M)}}, R^{\mathrm{NV(M)}}) &= (P^{\mathrm{NV(M)}} - c)(R^{\mathrm{NV(M)}} - bP^{\mathrm{NV(M)}}) - (P^{\mathrm{NV(M)}} - s)\mathsf{E}^M[(R^{\mathrm{NV(M)}} - A_T)^+] \\
&= (P^{\mathrm{NV(M)}} - c)(R^{\mathrm{NV(M)}} - bP^{\mathrm{NV(M)}}) - (P^{\mathrm{NV(M)}} - s)R^{\mathrm{NV(M)}}\mathsf{P}^M(A_T \leq R^{\mathrm{NV(M)}}) \\
&\quad + (P^{\mathrm{NV(M)}} - s)\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\mathrm{NV(M)}}\}] \\
&= (-bP^{\mathrm{NV(M)}})(P^{\mathrm{NV(M)}} - c) + (P^{\mathrm{NV(M)}} - s)\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\mathrm{NV(M)}}\}] > 0.
\end{aligned}$$

The third equality uses the optimality equation $\mathsf{P}^M(A_T \leq R^{\mathrm{NV(M)}}) = (P^{\mathrm{NV(M)}} - c)/(P^{\mathrm{NV(M)}} - s)$. Since $(-bP^{\mathrm{NV(M)}})(P^{\mathrm{NV(M)}} - c) < 0$ (by $P^{\mathrm{NV(M)}} > c > s$), we must have $\mathsf{E}^M[A_T \mathbf{1}\{A_T \leq R^{\mathrm{NV(M)}}\}] > 0$.

Let $R_3 := F_M^{-1}((P_3 - c)/(P_3 - s))$, i.e., $R_3 = R^{\mathrm{NV(M)}}(P_3)$. As $P_2 < P_3$,

$$R^{\mathrm{NV}} = R^{\mathrm{NV(M)}}(P_2) \leq R^{\mathrm{NV(M)}}(P_3) = R_3.$$

By $P_3 \leq P^{\mathrm{NV}}$,

$$R_3 = R^{\mathrm{NV(M)}}(P_3) \leq R^{\mathrm{NV(M)}}(P^{\mathrm{NV}}) \leq R^{\mathrm{NV(M)}}.$$

Combining the above, we have $R^{\mathrm{NV}} \le R_3 \le R^{\mathrm{NV(M)}}(P^{\mathrm{NV}}) \le R^{\mathrm{NV(M)}}$. Let

$$g(R) := \frac{P_3 - s}{c - s} \mathsf{E}^M[R \wedge A_T] - R.$$

Clearly, $g(R^{\mathrm{NV}}) = $ RHS of (EC.8.14). Taking derivative of $g(R)$, we have

$$g'(R) = \frac{P_3 - s}{c - s} \mathsf{P}^M(A_T \ge R) - 1 \ge 0, \quad \text{for any} \quad R \le R_3,$$

since $\mathsf{P}^M(A_T \ge R) \ge (c - s)/(P_3 - s)$ for $R \le R_3$. Thus, $g(R^{\mathrm{NV}}) \le g(R_3)$. Now we check $g(R_3)$:

$$\begin{aligned}
g(R_3) &= \frac{P_3 - s}{c - s} \mathsf{E}^M[R_3 \wedge A_T] - R_3 \\
&= \frac{P_3 - s}{c - s} \mathsf{E}^M[A_T \mathbf{1}\{A_T \le R_3\}] + \frac{P_3 - s}{c - s} R_3 \mathsf{P}^M(R_3 \le A_T) - R_3 \\
&= \frac{P_3 - s}{c - s} \mathsf{E}^M[A_T \mathbf{1}\{A_T \le R_3\}] \\
&\le \frac{P^{\mathrm{NV(M)}} - s}{c - s} \mathsf{E}^M[A_T \mathbf{1}\{A_T \le R^{\mathrm{NV(M)}}\}] \\
&= T(R^{\mathrm{NV(M)}}).
\end{aligned}$$

The third $=$ is due to the definition of R_3: $\mathsf{P}^M(R_3 \le A_T) = (c - s)/(P_3 - s)$ and the last $=$ is due to (EC.8.15). The $\le$ is due to $P^{\mathrm{NV(M)}} \ge P^{\mathrm{NV}} \ge P_3 > c > s$, $\mathsf{E}^M[A_T \mathbf{1}\{A_T \le R^{\mathrm{NV(M)}}\}] > 0$ and $R^{\mathrm{NV(M)}} \ge R_3$. Therefore,

$$T(R^\circ) = \text{RHS of (EC.8.14)} = g(R^{\mathrm{NV}}) \le g(R_3) \le T(R^{\mathrm{NV(M)}}),$$

which leads to $R^\circ \le R^{\mathrm{NV(M)}}$. This completes the proof. $\qquad\square$

EC.8.4. Proof of Proposition 8

To establish the efficient frontier, direct differentiation using the expression of $B(m, P, R)$ in (21) yields (with $C = 1/(Z_0^M - 1) > 0$):

$$\begin{aligned}
\frac{dB(m, P_m^h, R_m^h)}{dm} &= 2C[m - V_0(P_m^h, R_m^h)] + \frac{\partial B(m, P_m^h, R_m^h)}{\partial P} \frac{dP_m^h}{dm} + \frac{\partial B(m, P_m^h, R_m^h)}{\partial R} \frac{dR_m^h}{dm} \\
&= 2C[m - V_0(P_m^h, R_m^h)] \ge 0,
\end{aligned}$$

where the second equality follows from optimality of P_m^h and R_m^h (which makes the partial derivatives vanish) and the ≥ 0 follows from Proposition 7. $\qquad\square$

EC.9.   Proofs for Section EC.4
EC.9.1.   Proof of Proposition EC.2

For any given R, $V_0(P, R)$ is concave in P; in particular, it increases in P up to $P^{\mathrm{NV(M)}}(R)$ and then decreases. First, suppose $m > \max_P V_0(P, R)$. Then, increasing P beyond $P^{\mathrm{NV(M)}}(R)$ increases both terms of $B(m, P, R)$, thus is not optimal. Therefore, for this case the optimal P is bounded by

$P^{\mathrm{NV(M)}}(R)$. Now consider the other case, $m \le \max_P V_0(P, R)$. Then, increasing P beyond $\bar{P}(R)$ (i.e. the smaller root of $V_0(P, R) = m$) is not optimal as it also increases both terms of $B(m, P, R)$.

In summary, for a given R, the optimal P satisfies $V_0(P, R) \le m$ over $P \le \bar{P}(R)$, and thus, by concavity of $V_0(P, R)$, the first term of $B(m, P, R)$ is convex in P. The second term of $B(m, P, R)$ is a convex quadratic function in P, and this concludes the proof. $\square$

EC.10. Proofs for Section EC.5

Here we spell out the solution to the hedging problem in (EC.5.4). Similar to the price-setting newsvendor model, the solution approach relies on the *projected* production payoff process:

$$V_t(\mathbf{q}) := \mathsf{E}^M[H_T(\mathbf{q}, \mathbf{A}_T) \,|\, \mathcal{F}_t]. \tag{EC.10.1}$$

Clearly, V_t is a martingale under P^M. Then, we apply the martingale representation theorem to decompose V_t and we summarize the decomposition in the following lemma. The proof is collected in Section EC.10.1.

LEMMA EC.11. *$V_t(\mathbf{q})$ defined in (EC.10.1) has the following representation:*

$$V_t(\mathbf{q}) = V_0(\mathbf{q}) + \int_0^t \xi_s(\mathbf{q})dX_s + \int_0^t \boldsymbol{\delta}_s(\mathbf{q})d\tilde{\mathbf{B}}_s, \tag{EC.10.2}$$

with $V_0(\mathbf{q}) = \mathsf{E}^M[H_T(\mathbf{q}, A_T)]$. ξ_t and $\boldsymbol{\delta}_t$ are processes adapted to $\mathcal{F}_t$. In particular,

$$\boldsymbol{\delta}_t(\mathbf{q}) = \nabla_{\mathbf{a}}' f(t, X_t, \mathbf{A}_t, \sigma_t)\tilde{\boldsymbol{\Sigma}}, \tag{EC.10.3}$$

where $f(t, x, \mathbf{a}, \sigma) := \mathsf{E}^M[H_T(\mathbf{q}, \mathbf{A}_T) \,|\, X_t = x, \mathbf{A}_t = \mathbf{a}, \sigma_t = \sigma]$.

The expression of $\xi_t(\mathbf{q})$ needs further to specify the general processes $\mathbf{C}_t$ and σ_t.

The optimal solution to the hedging problem (EC.5.4) is summarized by following theorem, with proof detailed in Section EC.10.2.

THEOREM EC.2. *(i) The optimal trading strategy is*

$$\theta_t^*(\mathbf{q}) = \underbrace{-\xi_t(\mathbf{q})}_{\text{risk-mitigation position}} \underbrace{-\frac{\zeta_t}{Z_t^M}[\lambda_m - V_t(\mathbf{q}) - \chi_t^*]}_{\text{investment position}}, \tag{EC.10.4}$$

$$\lambda_m = \frac{mZ_0^M - V_0(\mathbf{q})}{Z_0^M - 1}. \tag{EC.10.5}$$

(Note: $Z_0^M = \mathsf{E}(Z_T^2) > 1$.)

(ii) The optimal objective function value has the following expression:

$$B(m, \mathbf{q}) = \underbrace{\frac{[m - V_0(\mathbf{q})]^2}{Z_0^M - 1}}_{\text{squared investment risk}} + \underbrace{\int_0^T \mathsf{E}\Big[\frac{Z_t}{Z_t^M}||\boldsymbol{\delta}_t(\mathbf{q})||^2\Big]dt}_{\text{squared unhedgeable risk}}, \tag{EC.10.6}$$

In particular, $Z_0^M = \mathsf{E}(Z_T^2) > 1$, and $0 \le Z_t/Z_t^M \le 1$.

EC.10.1.   Proof of Lemma EC.11

Recall that X_t is a (local) martingale under P^M and admits the following dynamics:

$$dX_t^M = \sigma_t X_t dB_t^M.$$

According to the definition of $\mathbf{A}_t$, for all $u \geq t \geq 0$,

$$\mathbf{A}_u = \mathbf{C}_u + \tilde{\boldsymbol{\Sigma}}\tilde{\mathbf{B}}_u = \mathbf{A}_t + \mathbf{C}_u - \mathbf{C}_t + \tilde{\boldsymbol{\Sigma}}(\tilde{\mathbf{B}}_u - \tilde{\mathbf{B}}_t). \qquad (\text{EC.10.7})$$

By the assumption that $\mathbf{C}_u - \mathbf{C}_t \in \sigma(\{X_s, t \leq s \leq u\})$ and (X_t, σ_t) is Markovian under P^M, the distribution of $\mathbf{C}_u - \mathbf{C}_t$ (conditional on $\mathcal{F}_t$) is in turn determined by (X_t, σ_t). Then, clearly $(\mathbf{A}_t, X_t, \sigma_t)$ is also Markovian under P^M. Next, express $V_t(\mathbf{q})$ as:

$$V_t(\mathbf{q}) := \mathsf{E}^M[H_T(\mathbf{q}, \mathbf{A}_T) \,|\, \mathcal{F}_t] = \mathsf{E}^M[H_T(\mathbf{q}, \mathbf{A}_T) \,|\, X_t, \mathbf{A}_t, \sigma_t] = f(t, \mathbf{q}, X_t, \mathbf{A}_t, \sigma_t),$$

where the second equality follows from the Markovian property of $(\mathbf{A}_t, X_t, \sigma_t)$ and the function $f(\cdot)$ in the last line is defined as follows:

$$f(t, \mathbf{q}, x, \mathbf{a}, \sigma) := \mathsf{E}^M[H_T(\mathbf{q}, \mathbf{A}_T) \,|\, X_t = x, \mathbf{A}_t = \mathbf{a}, \sigma_t = \sigma].$$

Clearly, V_t is a martingale under P^M. By Proposition EC.3 (see Remark EC.2) and (EC.7.3), $V_t(\mathbf{q})$ has the following representation with $V_0(\mathbf{q}) = \mathsf{E}^M[H_T(\mathbf{q}, \mathbf{A}_T)]$ ($\mathbf{q}$ is dropped from the notation since it is given):

$$V_t = V_0 + \int_0^t \psi_s dB_s^M + \int_0^t \boldsymbol{\delta}_s \cdot d\tilde{\mathbf{B}}_s.$$

ψ_t and $\boldsymbol{\delta}_t$ are processes adapted to $\mathcal{F}_t$. To show that it leads to (EC.10.2), apply Itô's lemma:

$$
\begin{aligned}
dV_t &= df(t, \mathbf{q}, X_t, \mathbf{A}_t, \sigma_t) \\
&= f_x dX_t + \sum_{j=1}^J f_{a_j} dA_{jt} + f_\sigma d\sigma_t + \left(f_t + \frac{1}{2}f_{xx}\sigma_t^2 X_t^2 + \frac{1}{2}\sum_{i=1}^J\sum_{j=1}^J \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \frac{\partial^2 f}{\partial a_i \partial a_j}\right)dt \\
&\quad + \sum_{j=1}^J f_{xa_j}(dX_t dA_{jt}) + \sum_{j=1}^J f_{\sigma a_j}(d\sigma_t dA_{jt}) + f_{x\sigma}(dX_t d\sigma_t) \\
&= f_x dX_t + \nabla_{\mathbf{a}}' f \tilde{\boldsymbol{\Sigma}} d\tilde{\mathbf{B}}_t + \sum_{j=1}^J f_{a_j} dC_{jt} + f_\sigma d\sigma_t + \left(f_t + \frac{1}{2}f_{xx}\sigma_t^2 x^2 + \frac{1}{2}\sum_{i=1}^J\sum_{j=1}^J \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \frac{\partial^2 f}{\partial a_i \partial a_j}\right)dt \\
&\quad + \sum_{j=1}^J f_{xa_j}(dX_t dA_{jt}) + \sum_{j=1}^J f_{\sigma a_j}(d\sigma_t dA_{jt}) + f_{x\sigma}(dX_t d\sigma_t) \\
&= [...]dt + \psi_t dB_t^M + \nabla_{\mathbf{a}}' f \tilde{\boldsymbol{\Sigma}} d\tilde{\mathbf{B}}_t \\
&= \frac{\psi_t}{\sigma_t X_t} dX_t + \nabla_{\mathbf{a}}' f \boldsymbol{\Sigma} d\tilde{\mathbf{B}}_t
\end{aligned}
$$

In the last line, the dt-term vanishes due to martingale property of V_t. Furthermore, dB_t^M is related to dX_t based on the dynamics of X_t in P^M; see (EC.7.3). Comparing the last line with the martingale representation of $V_t(\mathbf{q})$ above and matching the coefficient of the term $d\tilde{\mathbf{B}}_t$ leads to:

$$\boldsymbol{\delta}_t(\mathbf{q}) = \nabla_{\mathbf{a}}' f(t, \mathbf{q}, X_t, \mathbf{A}_t, \sigma_t) \tilde{\boldsymbol{\Sigma}}.$$

(Matching the dX_t-term leads to $\xi_t = \psi_t/(\sigma_t X_t)$.) $\qquad\square$

EC.10.2. Proof of Theorem EC.2

The methodology specified in Section EC.7.3 leading to Theorem 1 directly applies to prove Theorem EC.2, because it does not assume any specific form of the payoff function and the asset price model is the same. The only difference here is that the δ_t^2 involved in $B(\cdot)$ in (21) is replaced by $\|\boldsymbol{\delta}_t\|^2$ in (EC.10.6). Specifically, the only modification needed is to replace the one-dimensional δ_t in (18) with the vector $\boldsymbol{\delta}_t$ in (EC.10.3), hence the one-dimensional γ_t in (EC.7.42) is replaced by a vector $\boldsymbol{\gamma}_t = -\boldsymbol{\delta}_t/Z_t^M$, accordingly. Then, when applying Itô's isometry in (EC.7.39), we have

$$Z_0^M \mathsf{E}^R\left[\left(\int_0^T \boldsymbol{\gamma}_t \cdot d\tilde{\mathbf{B}}_t\right)^2\right] = Z_0^M \int_0^T \mathsf{E}^M\left[\|\boldsymbol{\gamma}_t\|^2 \frac{Z_t^M}{Z_0^M}\right] dt = \int_0^T \mathsf{E}\left[\|\boldsymbol{\gamma}_t\|^2 Z_t^M Z_t\right] dt = \int_0^T \mathsf{E}\left[\frac{Z_t}{Z_t^M}\|\boldsymbol{\delta}_t\|^2\right] dt.$$

The rest of the proof is the same as that in Section EC.7.3.3, which leads to the expression of $B(m, \mathbf{q})$ in (EC.10.6). This completes the proof. $\qquad\square$

EC.10.3. Proof of Theorem EC.1

EC.10.3.1. Technical Preparation
In this section, we present a collection of technical results that will be used later in the proof of Theorem EC.1.

LEMMA EC.12. *For any given* $\mathbf{A}_1 \preceq \mathbf{A}_2$ *(defined as* $\mathsf{P}(\mathbf{A}_1 \in U) \leq \mathsf{P}(\mathbf{A}_2 \in U)$ *for any upper set* $U \subseteq \mathbb{R}^J$*), suppose* $H_T(\mathbf{q}, \mathbf{a})$ *has increasing differences in* $(\mathbf{q}, \mathbf{a})$. *Then,* $\mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_2)] - \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_1)]$ *is entry-wise increasing in* $\mathbf{q}$.

Proof. For any $\mathbf{q}_2 \geq \mathbf{q}_1 \geq \mathbf{0}$, we want to prove:

$$\mathsf{E}[H_T(\mathbf{q}_2, \mathbf{A}_2) - H_T(\mathbf{q}_2, \mathbf{A}_1)] \geq \mathsf{E}[H_T(\mathbf{q}_1, \mathbf{A}_2) - H_T(\mathbf{q}_1, \mathbf{A}_1)]$$

which is equivalent to

$$\mathsf{E}[H_T(\mathbf{q}_2, \mathbf{A}_2) - H_T(\mathbf{q}_1, \mathbf{A}_2)] \geq \mathsf{E}[H_T(\mathbf{q}_2, \mathbf{A}_1) - H_T(\mathbf{q}_1, \mathbf{A}_1)].$$

To this end, let $\phi(\mathbf{a}) = H_T(\mathbf{q}_2, \mathbf{a}) - H_T(\mathbf{q}_1, \mathbf{a})$, then $\phi(\mathbf{a})$ is entry-wise increasing in $\mathbf{a}$ as $H_T(\mathbf{q}, \mathbf{a})$ has increasing differences in $(\mathbf{q}, \mathbf{a})$. Then, by $\mathbf{A}_1 \preceq \mathbf{A}_2$, we have:

$$\mathsf{E}[\phi(\mathbf{A}_2)] \geq \mathsf{E}[\phi(\mathbf{A}_1)],$$

which is equivalent to $\mathsf{E}[H_T(\mathbf{q}_2, \mathbf{A}_2)] - H_T(\mathbf{q}_1, \mathbf{A}_2)] \geq \mathsf{E}[H_T(\mathbf{q}_2, \mathbf{A}_1)] - H_T(\mathbf{q}_1, \mathbf{A}_1)]$. This completes the proof. $\qquad\square$

Note that Lemma EC.12 implies:

$$\frac{\partial \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_2)]}{\partial q_i} \geq \frac{\partial \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_1)]}{\partial q_i} \quad \text{for all } i.$$

Next, we present an extension of Topkis Theorem (Topkis (1978)) to the payoff function $H_T(\mathbf{q}, \mathbf{A}_T)$ described in Section EC.5.

THEOREM EC.3 (**Extension of Topkis Theorem**). *Suppose $H_T(\mathbf{q}, \mathbf{a})$ satisfies Assumption EC.1. Let*

$$\mathbf{q}_i^* = \min \arg \max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_i)], \quad i = 1, 2.$$

If $\mathbf{A}_1 \preceq \mathbf{A}_2$, then $\mathbf{q}_1^* \leq \mathbf{q}_2^*$.

Proof. We prove this result by mathematical induction on the dimension of $\mathbf{q}$, denoted by N. For $N = 1$, let

$$q_i^* = \min \arg \max_{q \geq 0} \mathsf{E}[H_T(q, \mathbf{A}_i)], \quad i = 1, 2.$$

The optimality condition yields:

$$\frac{\partial \mathsf{E}[H_T(q_i^*, \mathbf{A}_i)]}{\partial q} \leq 0, \quad i = 1, 2.$$

(Specifically, the above holds as equation for $q_i^* > 0$ and may hold as a strict inequality of $q_i^* = 0$.) Then, by Lemma EC.12,

$$\frac{\partial \mathsf{E}[H_T(q_2^*, \mathbf{A}_1)]}{\partial q} \leq \frac{\partial \mathsf{E}[H_T(q_2^*, \mathbf{A}_2)]}{\partial q} \leq 0.$$

If $q_1^* = 0$, then $q_1^* \leq q_2^*$ clearly holds. Otherwise, $q_1^* > 0$ and we have

$$\frac{\partial \mathsf{E}[H_T(q_2^*, \mathbf{A}_1)]}{\partial q} \leq 0 = \frac{\partial \mathsf{E}[H_T(q_1^*, \mathbf{A}_1)]}{\partial q}.$$

By concavity of $\mathsf{E}[H_T(q, \mathbf{A}_1)]$ in q, we have $q_1^* \leq q_2^*$. Combining both cases, $q_1^* \leq q_2^*$ holds for $N = 1$

Now, suppose that the result holds for N and below we prove that it also holds for $N + 1$. Let

$$\mathbf{q}_i^* = \min \arg \max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_i)], \quad \text{where } \mathbf{q} = (q_1, q_2, ..., q_{N+1}), \quad i = 1, 2.$$

For any given $q_j \geq 0$, $j \in \{1, 2, ..., N+1\}$, let

$$\mathbf{q}_{i,-j}^*(q_j) = \min \arg \max_{\mathbf{q}_{-j} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_i)], \quad i = 1, 2,$$

where $\mathbf{q}_{-j} = (q_1, q_2, ..., q_{j-1}, q_{j+1}, ..., q_{N+1})$. Clearly, for any $j \in \{1, 2, ..., N+1\}$, $\mathbf{q}_{i,-j}^*(\mathbf{q}_{i,j}^*)$ is just $\mathbf{q}_i^*$ with the j-th entry excluded. The argument $\mathbf{q}_{-j}$ is an N-dimensional vector. Then, by induction hypothesis,

$$\mathbf{q}_{1,-j}^*(q_j) \leq \mathbf{q}_{2,-j}^*(q_j) \tag{EC.10.8}$$

for any given $q_j \geq 0$, $j \in \{1, 2, ..., N+1\}$. Furthermore, by supermodularity of H_T in $\mathbf{q}$ and Topkis Theorem, $\mathbf{q}^*_{i,-j}(q_j)$ is increasing in q_j for both $i = 1, 2$.

Suppose there exists a $j \in \{1, 2, ..., N+1\}$, such that $\mathbf{q}^*_{1,j} \leq \mathbf{q}^*_{2,j}$. From (EC.10.8) and that $\mathbf{q}^*_{i,-j}(q_j)$ increases in q_j, we must have

$$\mathbf{q}^*_{1,-j}(\mathbf{q}^*_{1,j}) \leq \mathbf{q}^*_{2,-j}(\mathbf{q}^*_{1,j}) \leq \mathbf{q}^*_{2,-j}(\mathbf{q}^*_{2,j})$$

which leads to

$$\mathbf{q}^*_1 \leq \mathbf{q}^*_2.$$

In other words, if there exists some j such that $\mathbf{q}^*_{1,j} \leq \mathbf{q}^*_{2,j}$, then $\mathbf{q}^*_1 \leq \mathbf{q}^*_2$ must hold. Otherwise, suppose there does not exist such j, i.e., $\mathbf{q}^*_{1,j} > \mathbf{q}^*_{2,j}$ for all $j \in \{1, 2, ..., N+1\}$; in other words, $\mathbf{q}^*_1 > \mathbf{q}^*_2$. By definition of $\mathbf{q}^*_1$, we have

$$\mathsf{E}[H_T(\mathbf{q}^*_1, \mathbf{A}_1)] - \mathsf{E}[H_T(\mathbf{q}^*_2, \mathbf{A}_1)] > 0.$$

This is because that by the assumed $\mathbf{q}^*_1 > \mathbf{q}^*_2$, $\mathbf{q}^*_2$ can not be the optimal solution to $\max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_1)]$ due to $\mathbf{q}^*_1$ being defined as the smallest optimal solution. Then, by Lemma EC.12,

$$\mathsf{E}[H_T(\mathbf{q}^*_1, \mathbf{A}_2)] - \mathsf{E}[H_T(\mathbf{q}^*_2, \mathbf{A}_2)] \geq \mathsf{E}[H_T(\mathbf{q}^*_1, \mathbf{A}_1)] - \mathsf{E}[H_T(\mathbf{q}^*_2, \mathbf{A}_1)] > 0$$

and this contradicts the optimality of $\mathbf{q}^*_2$ to $\max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_2)]$. By contradiction, there must exists $j \in \{1, 2, ..., N+1\}$ such that $\mathbf{q}^*_{1,j} \leq \mathbf{q}^*_{2,j}$ and as already proved above, this leads to $\mathbf{q}^*_1 \leq \mathbf{q}^*_2$. This completes the proof. □

LEMMA EC.13. *Let*

$$\mathbf{q}^{*M} := \min \arg \max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}^M[H_T(\mathbf{q}, \mathbf{A}_T)] = \min \arg \max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_T^M)], \tag{EC.10.9}$$

and

$$\mathcal{S}_m^h = \arg \min_{\mathbf{q} \geq \mathbf{0}} B(m, \mathbf{q}). \tag{EC.10.10}$$

Then, for any $m \geq \mathsf{E}^M[H_T(\mathbf{0}, \mathbf{A}_T)]$, it holds that:

$$\mathbf{q}_m^h \leq \mathbf{q}^{*M}, \quad \forall \mathbf{q}_m^h \in \mathcal{S}_m^h.$$

Proof. Recall the notation $V_0 := \mathsf{E}^M(H_T)$.

We prove the stated result by mathematical induction on the dimension of $\mathbf{q}$, denoted by N. First, consider $N = 1$. If $q_m^h = 0$, clearly $q_m^h \leq q^{*M}$ holds. Otherwise, $q_m^h > 0$ and the following optimality equation $\partial B(m, q_m^h)/\partial q = 0$ must hold, which is expressed as:

$$2C(V_0(q_m^h) - m)\frac{\partial V_0(q_m^h)}{\partial q} + \frac{\partial \Psi(q_m^h)}{\partial q} = 0.$$

By Assumption EC.2, $\partial\Psi(q_m^h)/\partial q \geq 0$, hence $(V_0(q_m^h) - m)(\partial V_0(q_m^h)/\partial q) \leq 0$ must hold. If $m \geq V_0(q^{*M})$, then clearly $V_0(q_m^h) - m \leq 0$ and $\partial V_0(q_m^h)/\partial q \geq 0$. By concavity of V_0 in q, this indicates $q_m^h \leq q^{*M}$. Now consider the other case, $m < V_0(q^{*M})$. Since $V_0(0) = 0 \leq m$, there exists $\tilde{q} \in [0, q^{*M}]$ such that $V_0(\tilde{q}) = m$, and thus the first term of $B(m, \tilde{q})$, $C[m - V_0(\tilde{q})]^2$, is zero. Therefore, any q with $q \geq \tilde{q}$ can not be optimal since for such q, $B(m, q)$ has both terms larger than those of $B(m, \tilde{q})$. Hence, $q_m^h \leq \tilde{q} \leq q^{*M}$ must hold. In summary, the result stated in this theorem holds for $N = 1$.

Now, suppose the stated result holds for $N \geq 1$. For an $(N + 1)$-dimensional decision vector $\mathbf{q}$, write $\mathbf{q} = (q_j, \mathbf{q}_{-j})$, where $\mathbf{q}_{-j}$ is $\mathbf{q}$ excluding q_j for some $j \in \{1, 2, ..., N + 1\}$; clearly, $\mathbf{q}_{-j}$ is of dimension N. Following this notation, for the problem of dimension-$(N + 1)$, express $\mathbf{q}_m^h$ as $\mathbf{q}_m^h = (q_{m,j}^h, \mathbf{q}_{m,-j}^h)$. In the following, we will prove two results. First, we prove that if there exists a $j \in \{1, 2, ..., N + 1\}$ such that $q_{m,j}^h \leq q_j^{*M}$, then $\mathbf{q}_{m,-j}^h \leq \mathbf{q}_{-j}^{*M}$. Second, we prove that it is impossible that $q_{m,j}^h \geq q_j^{*M}$ for all $j \in \{1, 2, ..., N + 1\}$; in other words, $\mathbf{q}_m^h \geq \mathbf{q}^{*M}$ can not hold. Combining these two results immediately leads to $\mathbf{q}_m^h \leq \mathbf{q}^{*M}$ for the dimension-$(N + 1)$ problem, thereby proves the stated result stated by this lemma.

Next, we prove the two results stated above.

First, we prove that if there exists a $j \in \{1, 2, ..., N + 1\}$ such that $q_{m,j}^h \leq q_j^{*M}$, then $\mathbf{q}_{m,-j}^h \leq \mathbf{q}_{-j}^{*M}$. Suppose for some $j \in \{1, 2, ..., N + 1\}$, $q_{m,j}^h \leq q_j^{*M}$ holds. Next, fix $q_j = q_{m,j}^h$ in the optimization problems (EC.10.9) and (EC.10.10) so that the decision variable is just $\mathbf{q}_{-j}$ for both problems. Then, for any $\tilde{\mathbf{q}}_{m,-j}^h \in \arg\min_{\mathbf{q}_{-j} \geq \mathbf{0}} B(m, \mathbf{q}_{-j}, q_{m,j}^h)$, it holds that

$$\tilde{\mathbf{q}}_{m,-j}^h \leq \min\arg\max_{\mathbf{q}_{-j} \geq \mathbf{0}} V_0(\mathbf{q}_{-j}, q_{m,j}^h) \leq \min\arg\max_{\mathbf{q}_{-j} \geq \mathbf{0}} V_0(\mathbf{q}_{-j}, q_j^{*M}) = \mathbf{q}_{-j}^{*M}.$$

The first $\leq$ is by induction hypothesis (note that $\mathbf{q}_{-j}$ is of dimension N) and the second $\leq$ is by Topkis Theorem (recall, V_0 is supermodular in all q_j's). Recognizing that $\mathbf{q}_{m,-j}^h \in \arg\min_{\mathbf{q}_{-j}} B(m, \mathbf{q}_{-j}, q_{m,j}^h)$, $\mathbf{q}_{m,-j}^h \leq \mathbf{q}_{-j}^{*M}$ immediately follows.

Now, we prove the second result: it is impossible that $q_{m,j}^h \geq q_j^{*M}$ for all $j \in \{1, 2, ..., N + 1\}$. First, consider the case of $m \geq V_0(\mathbf{q}^{*M})$. Then, any $\mathbf{q}$ with $\mathbf{q} \geq \mathbf{q}^{*M}$ can not be optimal since it induces a higher variance than $\mathbf{q}^{*M}$ does: the latter has both terms of $B(m, \mathbf{q})$ smaller than the former. Hence, for this case the stated result holds. Now we focus on the other case, $m < V_0(\mathbf{q}^{*M})$. Suppose $\mathbf{q}_m^h \geq \mathbf{q}^{*M}$. Let $V_{00} = V_0(\mathbf{0})$; note $V_{00} = \mathsf{E}^M[H_T(\mathbf{0}, \mathbf{A}_T)] \leq m$. In addition, let $\mathbf{q}_{[n]} = (q_j)_{j=1}^n$ for $n = 1, 2, \cdots, N + 1$, i.e., $\mathbf{q}_{[n]}$ is the vector of the first n entries of $\mathbf{q}$. For each $n \in \{1, 2, \cdots, N + 1\}$, define problem n as:

$$V_{0n} := \max_{\mathbf{q}_{[n]} \geq \mathbf{0}} V_0(\mathbf{q}_{[n]}, 0, \cdots, 0), \quad \mathbf{q}^{(n)} := \min\arg\max_{\mathbf{q}_{[n]} \geq \mathbf{0}} V_0(\mathbf{q}_{[n]}, 0, \cdots, 0). \quad (\text{EC.10.11})$$

Clearly, $\mathbf{q}^{(N+1)} = \mathbf{q}^{*M}$ and thus $V_{0(N+1)} = V_0(\mathbf{q}^{*M}) > m$. So, V_{0n} increases from $V_{00} \leq m$ to $V_{0(N+1)} > m$ as n increases from 0 to $N + 1$. Therefore, $\exists \tilde{n} \leq N - 1$ such that

$$V_{0\tilde{n}} \leq m \leq V_{0(\tilde{n}+1)}.$$

Now consider the problem $\tilde{n} + 1$, which produces $\mathbf{q}^{(\tilde{n}+1)}$. Consider the vector $\mathbf{q}(q) = (\mathbf{q}_{[\tilde{n}]}^{(\tilde{n}+1)}, q, 0, \cdots, 0)$. First, $V_0(\mathbf{q}(q_{\tilde{n}+1}^{(\tilde{n}+1)})) = V_0((\mathbf{q}^{(\tilde{n}+1)}, 0, \cdots, 0)) = V_{0(\tilde{n}+1)} \geq m$. Next, note

$$V_0(\mathbf{q}(0)) = V_0((\mathbf{q}_{[\tilde{n}]}^{(\tilde{n}+1)}, 0, \cdots, 0)) \leq \max_{\mathbf{q}_{[\tilde{n}]} \geq \mathbf{0}} V_0(\mathbf{q}_{[\tilde{n}]}, 0, \cdots, 0) = V_{0\tilde{n}} \leq m.$$

Therefore, there exists $q^\circ \in [0, q_{\tilde{n}+1}^{(\tilde{n}+1)}]$ such that

$$V_0(\mathbf{q}(q^\circ)) = V_0((\mathbf{q}_{[\tilde{n}]}^{(\tilde{n}+1)}, q^\circ, 0, \cdots, 0)) = m.$$

Clearly, $\mathbf{q}(q^\circ) \leq \mathbf{q}(q_{\tilde{n}+1}^{(\tilde{n}+1)})$. Now, with $\mathbf{q}^{*M} \geq \mathbf{0}$ and Topkis Theorem:

$$\mathbf{q}^{(\tilde{n}+1)} = \min \arg \max_{\mathbf{q}_{[\tilde{n}+1]} \geq \mathbf{0}} V_0(\mathbf{q}_{[\tilde{n}+1]}, 0, \cdots, 0) \leq \min \arg \max_{\mathbf{q}_{[\tilde{n}+1]} \geq \mathbf{0}} V_0(\mathbf{q}_{[\tilde{n}+1]}, q_{\tilde{n}+2}^{*M}, \cdots, q_{N+1}^{*M}) = \mathbf{q}_{[\tilde{n}+1]}^{*M},$$

so,

$$\mathbf{q}_{[\tilde{n}]}^{(\tilde{n}+1)} \leq \mathbf{q}_{[\tilde{n}]}^{*M} \quad \text{and} \quad q^\circ \leq q_{\tilde{n}+1}^{(\tilde{n}+1)} \leq q_{\tilde{n}+1}^{*M}.$$

Therefore,

$$\mathbf{q}(q^\circ) \leq \mathbf{q}(q_{\tilde{n}+1}^{(\tilde{n}+1)}) = (\mathbf{q}^{(\tilde{n}+1)}, 0, ..., 0) \leq \mathbf{q}^{*M}$$

In summary, there exists $\mathbf{q}(q^\circ) \leq \mathbf{q}^{*M} \leq \mathbf{q}_m^h$ such that the first term of $B(m, \mathbf{q}(q^\circ))$ (which is zero) is smaller than that of $B(m, \mathbf{q}_m^h)$ and so is its second term, this leads to $B(m, \mathbf{q}(q^\circ)) \leq B(m, \mathbf{q}_m^h)$. Thus, $\mathbf{q}_m^h$ cannot be optimal. By contradiction, it is impossible that $\mathbf{q}_m^h \geq \mathbf{q}^{*M}$. $\qquad \square$

EC.10.3.2. Proof of Theorem EC.1. When the asset price trend positively impacts the stochastic operational factors, i.e., $\mathbf{C}_T^M \preceq \mathbf{C}_T$, we have $\mathbf{A}_T^M \preceq \mathbf{A}_T$. From Theorem EC.3, we have:

$$\mathbf{q}^{*M} \leq \mathbf{q}^*.$$

By $\mathbf{A}_T^M \preceq \mathbf{A}_T$, $\mathsf{E}[H_T(\mathbf{0}, \mathbf{A}_T)] \geq \mathsf{E}^M[H_T(\mathbf{0}, \mathbf{A}_T)]$. So, $m = \max_{\mathbf{q} \geq \mathbf{0}} \mathsf{E}[H_T(\mathbf{q}, \mathbf{A}_T)] \geq \mathsf{E}[H_T(\mathbf{0}, \mathbf{A}_T)] \geq \mathsf{E}^M[H_T(\mathbf{0}, \mathbf{A}_T)]$ and thus Lemma EC.13 immediately leads to

$$\mathbf{q}_m^h \leq \mathbf{q}^{*M} \leq \mathbf{q}^*.$$

This completes the proof. $\qquad \square$

EC.10.4. Proof of Corollary EC.1

We begin with verifying that H_T satisfies all conditions specified in Assumption EC.1. Recall,

$$H_T((\mathbf{P}, \mathbf{R}), \mathbf{A}_T) = \sum_{j=1}^J \left[P_j \left(R_j - \sum_{i=1}^J \gamma_{ji} P_i \right) - (P_j + c_j)(R_j - A_{jT})^+ \right].$$

For each R_j, the term in H_T that involves this variable is $P_j(R_j - \sum_{i=1}^J \gamma_{ij} P_i) - (P_j + c_j)(R_j - A_{jT})^+$, which is a linear function minus a convex function of R_j, hence concave in R_j. For each P_j, the

nonlinear term in H_T involving this variable is: $-\gamma_{jj}P_j^2$, which is concave in P_j. Hence, H_T is concave in each R_j and P_j. Next, clearly, H_T is increasing in A_{jT}. Now, we check the supermodularity of H_T in the decision variables, and this is equivalent to verifying this property for each pair of variables. For any $i \neq j$, there is no interaction term involving R_i and R_j, or involving P_i and R_j. So, we only need to consider pairs of (P_i, P_j) for $i \neq j$ and (P_j, R_j). It is straightforward to obtain:

$$\frac{\partial^2 H_T}{\partial P_i \partial P_j} = -(\gamma_{ji} + \gamma_{ij}) > 0,$$

hence H_T's supermodularity in (P_i, P_j) holds. For any P_j, it is easy to derive:

$$\frac{\partial H_T}{\partial P_j} = R_j \wedge A_{jT} - 2\sum_{i=1}^{J} \gamma_{ji} P_i,$$

which increases in R_j and leads to supermodularity of H_T in P_j and R_j. Hence, H_T is supermodular in all the decision variables. Last, we verify the increasing difference property. For $\mathbf{A}_T \leq \mathbf{A}'_T$, it is straight forward to derive

$$H_T((\mathbf{P}, \mathbf{R}), \mathbf{A}_T) - H_T((\mathbf{P}, \mathbf{R}), \mathbf{A}'_T) = \sum_{j=1}^{J} (P_j + c_j)(R_j \wedge A'_{jT} - A_{jT})^+,$$

which clearly increases in all P_j and R_j and this confirms the increasing difference property. In summary, H_T satisfies Assumption EC.1.

Now we turn to verify Assumption EC.2. Recall, by Theorem EC.2,

$$\Psi(\mathbf{P}, \mathbf{R}) = \int_0^T \mathsf{E}\Big[\frac{Z_t}{Z_t^M} \|\boldsymbol{\delta}_t^2(\mathbf{P}, \mathbf{R})\| \Big] dt,$$

where

$$\boldsymbol{\delta}_t(\mathbf{P}, \mathbf{R}) = \nabla' f \tilde{\boldsymbol{\Sigma}} = \sum_{j=1}^{J} f_{a_j} \tilde{\boldsymbol{\sigma}}'_j, \quad f(t, x, \mathbf{a}, \sigma) = \mathsf{E}^M(H_T \mid X_t = x, \mathbf{A}_t = \mathbf{a}, \sigma_t = \sigma).$$

To obtain the expression of $\boldsymbol{\delta}_t$, we derive f_{a_j} as follows. Note

$$
\begin{aligned}
f(t, x, \mathbf{a}, \sigma) &= \mathsf{E}^M(H_T \mid X_t = x, \mathbf{A}_t = \mathbf{a}, \sigma_t = \sigma) \\
&= \sum_{j=1}^{J} P_j \Big(R_j - \sum_{i=1}^{J} \gamma_{ji} P_i\Big) - \sum_{j=1}^{J} (P_j + c_j)\mathsf{E}^M[(R_j - A_{jT})^+ \mid X_t = x, \mathbf{A}_t = \mathbf{a}, \sigma_t = \sigma] \\
&= \sum_{j=1}^{J} P_j \Big(R_j - \sum_{i=1}^{J} \gamma_{ji} P_i\Big) \\
&\quad - \sum_{j=1}^{J} (P_j + c_j)\mathsf{E}^M[(R_j - A_{jt} - (C_{jT} - C_{jt}) - \tilde{\boldsymbol{\sigma}}_j \cdot (\tilde{\mathbf{B}}_T - \tilde{\mathbf{B}}_t))^+ \mid X_t = x, \mathbf{A}_t = \mathbf{a}, \sigma_t = \sigma] \\
&= \sum_{j=1}^{J} P_j \Big(R_j - \sum_{i=1}^{J} \gamma_{ji} P_i\Big)
\end{aligned}
$$

$$- \sum_{j=1}^{J} (P_j + c_j) \mathsf{E}^M [(R_j - a_j - (C_{jT} - C_{jt}) - \tilde{\boldsymbol{\sigma}}_j \cdot (\tilde{\mathbf{B}}_T - \tilde{\mathbf{B}}_t))^+ \mid X_t = x, \sigma_t = \sigma],$$

where the last $=$ is follows from $C_{jT} - C_{jt} \in \sigma(\{X_u, t \le u \le T\})$, Markovian property of (X_t, σ_t) and independence of $\tilde{\mathbf{B}}_T - \tilde{\mathbf{B}}_t$ from $\mathcal{F}_t$. Then, for each $j = 1, \cdots, J$, we have:

$$f_{a_j}(t, x, \mathbf{a}, \sigma) = \sum_{i=1}^{J} (P_i + c_i) \mathsf{P}^M (a_i + (C_{iT} - C_{it}) + \tilde{\boldsymbol{\sigma}}_i \cdot (\tilde{\mathbf{B}}_T - \tilde{\mathbf{B}}_t) \le R_i \mid X_t = x, \sigma_t = \sigma),$$

hence

$$
\begin{aligned}
f_{a_j}(t, X_t, \mathbf{A}_t, \sigma_t) &= \sum_{i=1}^{J} (P_i + c_i) \mathsf{P}^M (A_{it} + (C_{iT} - C_{it}) + \tilde{\boldsymbol{\sigma}}_i \cdot (\tilde{\mathbf{B}}_T - \tilde{\mathbf{B}}_t) \le R_i \mid X_t, \sigma_t) \\
&= \sum_{i=1}^{J} (P_i + c_i) \mathsf{P}^M (A_{it} + (C_{iT} - C_{it}) + \tilde{\boldsymbol{\sigma}}_i \cdot (\tilde{\mathbf{B}}_T - \tilde{\mathbf{B}}_t) \le R_i \mid \mathcal{F}_t) \\
&= \sum_{i=1}^{J} (P_i + c_i) \mathsf{P}^M (A_{iT} \le R_i \mid \mathcal{F}_t),
\end{aligned}
$$

the second $=$ follows from, again, $C_{jT} - C_{jt} \in \sigma(\{X_u, t \le u \le T\})$, Markovian property of (X_t, σ_t) and independence of $\tilde{\mathbf{B}}_T - \tilde{\mathbf{B}}_t$ from $\mathcal{F}_t$. Then, we can derive:

$$\|\boldsymbol{\delta}_t(\mathbf{P}, \mathbf{R})\|^2 = \|\sum_{j=1}^{J} f_{a_j} \tilde{\boldsymbol{\sigma}}_j'\|^2 = \sum_{i=1}^{J} \sum_{j=1}^{J} (P_i + c_i)(P_j + c_j) \mathsf{P}^M (A_{iT} \le R_i \mid \mathcal{F}_t) \mathsf{P}^M (A_{jT} \le R_j \mid \mathcal{F}_t)(\tilde{\boldsymbol{\sigma}}_i \cdot \tilde{\boldsymbol{\sigma}}_j).$$

Clearly, when $\tilde{\boldsymbol{\sigma}}_i \cdot \tilde{\boldsymbol{\sigma}}_j \ge 0$ for all i, j, $\|\boldsymbol{\delta}_t^2(\mathbf{P}, \mathbf{R})\|^2$ increases in all P_j and R_j, hence so does $\Psi(\mathbf{P}, \mathbf{R})$. This completes the proof. $\qquad\square$

EC.10.5. Proof of Corollary EC.2

Recall that the profit function is

$$H_T((P, \gamma), A_T) = Pd(A_T, P, \gamma) - k(\gamma).$$

To prove this corollary, it suffices to check that the operational payoff function $H_T((P, \gamma), A_T)$ above satisfies Assumptions EC.1 and EC.2, then Theorem EC.1 applies. To this point, we verify each of the conditions specified in these two assumptions.

We start with Assumption EC.1. First, concavity of H_T in P and γ. For γ, the concavity immediately follows from the concavity of $d(\cdot)$ and convexity of $k(\cdot)$ in this variable. For P, it is straightforward to obtain:

$$\frac{\partial^2 H_T}{\partial P^2} = 2 \frac{\partial d}{\partial P} + P \frac{\partial^2 d}{\partial P^2}.$$

Since $d(\cdot)$ is decreasing and concave in P, the derivative above is negative, which leads to the concavity in P. Next, since $d(\cdot)$ is increasing in the market size A_T, clearly so does H_T. The supermodularity of H_T in (P, γ) follows from:

$$\frac{\partial^2 H_T}{\partial P \partial \gamma} = \frac{\partial d}{\partial \gamma} + P \frac{\partial^2 d}{\partial P \partial \gamma} \ge 0.$$

The ≥ 0 follows from $d(\cdot)$ increasing in γ and being supermodular in (P, γ). Last, for the increasing difference property of H_T in $((P, \gamma), A_T)$, we show that H_T is supermodular in (P, γ, A_T), which is a sufficient condition for this property. To this end, it is sufficient to show that H_T is supermodular in all three pairs: (P, γ), (γ, A_T) and (P, A_T). Supermodularity in the first pair is already shown above and that in the second pair follows immediately from supermodularity of $d(\cdot)$ in this pair. Now, supermodularity of H_T in (P, A_T) is a result of:

$$\frac{\partial^2 H_T}{\partial P \partial A_T} = \frac{\partial d}{\partial A_T} + P \frac{\partial^2 d}{\partial P \partial A_T} \geq 0,$$

where the ≥ 0 follows from $d(\cdot)$ being increasing in the market size A_T and supermodular in (P, A_T). In summary, H_T specified in this corollary satisfies Assumption EC.1.

Next, we verify that Assumption EC.2 holds. Recall, by Theorem EC.1, the squared unhedgeable risk has the following expression:

$$\Psi(P, \gamma) = \int_0^T \mathsf{E}\Big[\frac{Z_t}{Z_t^M} \delta_t^2(P, \gamma)\Big] dt,$$

where

$$\delta_t(P, \gamma) = \tilde{\sigma} \frac{\partial f(t, X_t, A_t, \sigma_t)}{\partial a}, \quad \text{with} \quad f(t, x, a, \sigma) = \mathsf{E}^M(H_T \mid X_t = x, A_t = a, \sigma_t = \sigma).$$

With $H_T = Pd(A_T, P, \gamma) - k(\gamma)$, we have

$$
\begin{aligned}
f(t, x, a, \sigma) &= \mathsf{E}^M[Pd(A_T, P, \gamma) - k(\gamma) \mid X_t = x, A_t = a, \sigma_t = \sigma] \\
&= \mathsf{E}^M[Pd(a + C_T - C_t + \tilde{\sigma}(\tilde{B}_T - \tilde{B}_t), P, \gamma) - k(\gamma) \mid X_t = x, \sigma_t = \sigma],
\end{aligned}
$$

where the second $=$ follows from $C_T - C_t \in \sigma(\{X_u, t \leq u \leq T\})$, Markovian property of (X_t, σ_t) and independence of $\tilde{B}_T - \tilde{B}_t$ from $\mathcal{F}_t$. This gives

$$\frac{\partial f}{\partial a} = P\mathsf{E}^M\Big[\frac{\partial d(a + C_T - C_t + \tilde{\sigma}(\tilde{B}_T - \tilde{B}_t), P, \gamma)}{\partial A_T}\Big| X_t = x, \sigma_t = \sigma\Big].$$

Hence, $\partial f/\partial a$ is nonnegative (since $d(\cdot)$ in increases in A_T) and is increasing in both P and γ (by supermodularity of $d(\cdot)$ in (A_T, P, γ)), and this leads to the fact that $\delta_t^2(P, \gamma)$ increases in P and γ. Since $Z_t/Z_t^M > 0$, this further implies that $\Psi(P, \gamma)$ increases in P and γ. This completes the proof. $\square$

References

Aydin G, Porteus EL (2008) Joint inventory and pricing decisions for an assortment. *Oper. Res.* 56(5):1247–1255.

Bagwell K (2007) The economic analysis of advertising. Armstrong M, Porter R, eds., *Handbook of Industrial Organization* (Amsterdam: North Holland).

Černỳ A, Kallsen J (2008) Mean–variance hedging and optimal investment in Heston's model with correlation. *Math. Finance* 18(3):473–492.

Dong L, Kouvelis P, Tian Z (2009) Dynamic pricing and inventory control of substitute products. *Manuf. Serv. Oper. Manag.* 11(2):317–339.

Dorfman R, Steiner P (1954) Optimal advertising and optimal quality. *Am. Econ. Rev.* 44(5):826–836.

Dunn WE, Vine DJ (2016) Why are inventory-sales ratios at us auto dealerships so high? *Finance and Economics Discussion Series 2016 Washington: Board of Governors of the Federal Reserve System* .

Gourieroux C, Laurent JP, Pham H (1998) Mean-variance hedging and numéraire. *Math. Finance* 8(3):179–200.

Hamilton J (1994) *Time Series Analysis* (Princeton University Press).

Jones DR, Schonlau M, Welch W (1998) Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* 13(4):455–492.

Mann H, Whitney D (1947) On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18(1):50–60.

Markowitz H (1952) Portfolio selection. *J. Finance* 7(1):77–91.

Markowitz H (2014) Mean–variance approximations to expected utility. *Eur. J. Oper. Res.* 234(2):346–355.

Pham H (2009) *Continuous-time Stochastic Control and Optimization with Financial Applications* (Springer-Verlag, Berlin/Heidelberg).

Revuz D, Yor M (1999) *Continuous Martingales and Brownian Motion* (Springer-Verlag Berlin Heidelberg New York).

Topkis D (1978) Minimizing a submodular function on a lattice. *Oper. Res.* 26(2):305–321.

Topkis D (1998) *Supermodularity and Complementarity* (Princeton University Press).

Wang L, Wissel J (2013) Mean-variance hedging with oil futures. *Financ. Stoch.* 17(4):641–683.