Data and Risk Analytics for Production Planning

Suggested Citation: Liao Wang and David D. Yao (2018), "Data and Risk Analytics for Production Planning", : Vol. issue, No. XX, pp 1–14. DOI: XXX.

Liao Wang
lwang98@hku.hk

David D. Yao
yao@columbia.edu

Contents

Data and Risk Analytics for Production Planning

Liao Wang[1] and David D. Yao[2]

[1]*Faculty of Business and Economics, The University of Hong Kong; lwang98@hku.hk.*
[2]*Department of Industrial Engineering and Operations Research, Columbia University; yao@columbia.edu.*

ABSTRACT

We examine the classical productional planning model, where a capacity decision that has to be made at the beginning of the planning horizon is the primary means to protect against demand uncertainty. We provide a critique on the model focusing on its profit maximizing objective, its underlying assumptions on demand and related forecasting scheme, and its overall business relevance (or the lack thereof); and we do so in the context of data, risk and analytics. Specifically, we will consider minimizing a shortfall risk relative to a profit target, with a demand model that captures impacts from the financial market and can be learned from data sets that are application specific. With a jointly optimized production and hedging strategy, we show the new model outperforms traditional approaches in risk mitigation as well as in expected profit.

1 Motivation and Description of the Problem

Driven by the economy of scale, modern production typically runs in *batches*: to produce a sizable batch of goods is often cheaper and more

Liao Wang and David D. Yao (2018), “Data and Risk Analytics for Production Planning”, : Vol. issue, No. XX, pp 1–14. DOI: XXX.

profitable. This calls for a production quantity or *capacity* decision. Following this decision, it takes time to complete the production for goods due to many other subsequent actions (e.g. procuring raw materials, organizing the work force, arranging production lines as well as the physical production processes). This time span, usually called production lead-time, is referred to as the planning horizon below and denoted $[0, T]$.

The challenge is, when the capacity decision is made, at time 0, there will only be limited visibility about demand on the produced goods at T, as T can be several weeks or months away; whereas once the capacity decision is made, it is often impossible or too costly to change. For example, in the automotive industry, suppliers of certain key parts (with high costs and/or long lead-times) will require car makers to make firm commitments on order quantities. (Ernst & Young Advisory Services, *Automotive & Transportation Report*, "Shifting Gear — Capacity Management in the Automotive Industry.") Thus, the uncertainty involved in demand at the end of the horizon implies the possibility of over-production and under-production, both will negatively impact the profit.

Production planning teaches how to make the time-zero capacity decision so as to mitigate the negative impact of demand uncertainty. A representative approach is the newsvendor (NV) model. It makes this decision so as to maximize the expected terminal (time T) profit; and unsurprisingly, its starting point is the demand distribution, which can be estimated from some forecasting scheme.

The goal of this paper is to provide a critique on this classical approach to production planning, by examining both its profit-maximizing objective and its demand modeling (or forecast). We will bring forth the following important points, supported by analyses, results and insights:

- Maximizing profit is *not even wrong*; it's irrelevant, in the context of a firm's sales and operations planning (of which production planning is a component). What's relevant, and practical as well, in this context is to address a risk measure as objective; or even better, to characterize the entire *efficient frontier*, the risk-return profile (curve) associated with the decision.

- Demand modeling can be enhanced by bringing out explicitly its *functional* dependence on certain financial assets (such as commodities and exchange rates) and/or on the general economy (motivation on this below). To specify this functional dependence is quite doable via machine learning; and there's no reason why this should not be incorporated into and integrated with demand forecasting schemes.
- With the enhanced demand model, one can develop a risk-hedging strategy, a real-time control executed throughout the planning horizon $[0, T]$; and this hedging strategy can be jointly optimized with the time-zero capacity decision. The combined capacity and hedging decisions can be shown to both reduce the risk and increase the terminal profit from the classical model.

1.1 Classical Production Planning—A Critique

The classical approach to production planning is the newsvendor (NV) model and its many variations. The basic model tries to find a production quantity Q at time $t = 0$ to supply a random demand D_T at time $T > 0$, the end of the planning horizon (or production lead-time) when the finished goods will supply demand; and the the objective is to maximize the expected profit, $\max_{Q \geq 0} \mathsf{E}[H_T(Q)]$, with

$$H_T(Q) := p(Q \wedge D_T) - c(Q - D_T)^+ = pQ - (p + c)(Q - D_T)^+, \quad (1)$$

where p is the unit profit (selling price minus cost), c is the net cost (cost minus salvage value) per unit, $\wedge$ denotes the min operator, and $(x)^+ := \max\{x, 0\}$. As motivated above, in the context of production planning, Q can be viewed as a proxy for *capacity*.

Taking derivative w.r.t. Q, and recognizing

$$\frac{d}{dQ} \mathsf{E}[(Q - D_T)^+] = \mathsf{P}(D_T \leq Q) := F(Q),$$

where $F(\cdot)$ denotes the distribution function of D_T, yields the optimal solution:

$$Q^{\mathrm{NV}} = F^{-1}\Big(\frac{p}{p + c}\Big). \quad (2)$$

Taking variance on $H_T(Q)$ in (1), we have

$$\mathsf{Var}[H_T(Q)] = (p+c)^2 \mathsf{Var}[(Q-D_T)^+]. \tag{3}$$

Direct derivation yields

$$\frac{d}{dQ}\mathsf{Var}[H_T(Q)] = 2(p+c)^2 \mathsf{E}[(Q-D_T)^+][1-\mathsf{P}(D_T \le Q)] \ge 0. \tag{4}$$

That is, $\mathsf{Var}[H_T(Q)]$ is increasing in Q, for all Q. Furthermore, from (1), we know $\mathsf{E}[H_T(Q)]$ is concave in Q; and specifically, it is increasing in $Q \in [0, Q^{\mathrm{NV}}]$ and decreasing in $Q > Q^{\mathrm{NV}}$.

Thus, we can conclude:

(i) For $Q \in [0, Q^{\mathrm{NV}}]$, where Q^{NV} follows (2), $\mathsf{E}[H_T(Q)]$ and $\mathsf{Var}[H_T(Q)]$ form an *efficient frontier*: they are both increasing in Q — a larger mean corresponds to a larger variance and vice versa.

In fact, more can be said about the efficient frontier; the following result can be directly verified:

(ii) Given m, let $Q_m (\le Q^{\mathrm{NV}})$ be the solution to $\mathsf{E}[H_T(Q)] = m$, and let $m^{\mathrm{NV}} := \mathsf{E}[H_T(Q^{\mathrm{NV}})]$. Then, $\mathsf{Var}[H_T(Q_m)]$ is increasing and convex in m, for $m \in [0, m^{\mathrm{NV}}]$.

Denote $v(m) := \mathsf{Var}[H_T(Q_m)]$. Then, the efficient frontier $(m, v(m))$ is an increasing and *convex* curve, meaning at a higher level of return (mean) any further increase is associated with a even steeper incremental increase in risk (variance). Thus, as we increase the production quantity towards the profit-maximizing Q^{NV}, the price we pay is the *maximal* incremental increase in risk.

Another weakness of the NV model is, it is out of context: who is exactly its intended user? Production planning is part of a firm's SOP (sales and operations planning) process, by which the executive in charge works out production decisions together with sales and operations managers. In that context, the profit/revenue target must have already been set by the firm's board, and it is imperative for the executive to meet (or beat) the target. In other words, maximizing profit is not even a relevant objective, not to add the enormous risk associated with such a pursuit as analyzed above.

With this application context in mind, it is useful to modify the classical NV model, replacing the profit-maximizing objective by minimizing a *risk measure*, the *shortfall*: $[m - H_T(Q)]^+$, where the constant $m \geq 0$ is a pre-specified profit target. Using the H_T expression in (1) and considering the two cases $pQ \leq m$ or $pQ \geq m$, we reach the following result:

(iii) Given m, the solution to $\min_{Q \geq 0} \mathsf{E}[m - H_T(Q)]^+ := s^{\text{NV}}(m)$ is $Q^{\text{NV}}(m) := \frac{m}{p} \wedge Q^{\text{NV}}$, with Q^{NV} following (2). Furthermore, $s^{\text{NV}}(m)$ is increasing (and convex) in m, which constitutes an efficient frontier.

Note when $m \geq pQ^{\text{NV}}$, $Q^{\text{NV}}(m) = Q^{\text{NV}}$, and it is direct to verify:

$$s^{\text{NV}}(m) = m - \mathsf{E}[H_T(Q^{\text{NV}})]. \tag{5}$$

That is, the shortfall will grow linearly in m. In this case, a production-only decision becomes a handicap; and this motivates the risk-hedging strategy, to be detailed below in §2.2.

Last but not least, the primary concern of production planning is to deal with demand uncertainty. Specifically, the quantity decision Q in the NV model is to strike the right balance between over- and under-production against the uncertain demand D_T that will only realize at the end of the planning horizon. Thus, a central issue is demand forecast, which traditionally is carried out via some statistical analysis applied to past data (time series), so as to discern any trend (mean or rate) and to characterize the fluctuation (variance or "noise"). Yet, in this age of big data and analytics, one surely wonder, can't we do better?

2 Modeling Approach and Methodology

2.1 Machine Learning for Demand Modeling/Forecast

In Wang and Yao (2017a,b), we have developed a new demand model that allows the demand *rate* to be a function of certain financial assets:

$$dD_t = \mu(\mathbf{X}_t)dt + \sigma dB_t, \tag{6}$$

where $\mathbf{X}_t = (X_{kt})_{k=1}^{K}$ is a vector, with X_{kt} representing the price of k-th financial asset at time t. B_t is a Brownian motion that represents the noise independent from $\mathbf{X}_t$ and $\sigma > 0$ is the associated volatility parameter. To motivate, consider a couple of examples from corporate reports as well as business news:

- A major producer for farming equipment recognizes agricultural commodity prices as a significant risk factor affecting its sales. Note the agricultural commodities are tradable via futures. (*Deere & Company's 2017 10K Filing: Risk Factors*, December 18, 2017.)

- Besides lower oil prices, demand for cars also increases as the economic condition improves. A proxy for the economic condition can be a stock index such as S&P500, which is tradable in the market as an index fund or ETF. (*Wall Street Journal*, January 8, 2015, "GM CEO Sees U.S. Auto Industry Gains in 2015.")

- Baltic Dry Index (BDI), which tracks freight costs of around twenty routes, closely reflects the demand for dry-bulk shipping: falling BDI usually accompanies weak sales of the shipping firms. While BDI is known to be very volatile, it takes about two years to build a ship, which makes vessel investment a risky decision for the shipping companies. Recently, an ETF was created to track the daily changes in the prices of ocean freight futures. (*Forbes*, April 25, 2012, "High Ambitions, Rough Seas"; *Reuters*, March 8, 2018, "U.S. Fund Manager Breakwave to Launch First ETF for Shipping Futures.")

So far, no functional form is imposed on the rate function in (6), and we allow this to be learned from data specific to any particular firm in interest. To obtain $\mu(\cdot)$ from data, we apply machine learning (ML) approaches, and this involves: (i) variable selection — identifying, from a set of financial assets in the market, the ones exhibiting most relevance to the demand; and (ii) function learning — finding the functional form in the selected assets to describe how the demand rate moves with the selected variables. We detail the procedure below.

Learning $\boldsymbol{\mu}(\cdot)$ from Data Assume the pre-processed data are in the following format:

- Demand: $\{\Delta D_i,\ i = 1, \cdots, N\}$, where ΔD_i is the demand realized for the i-th period which starts on time t_i and ends on t_{i+1}. Typically each $[t_i, t_{i+1}]$ spans a month, a quarter or a year.
- Assets: The daily price data of asset k is in form of $\{X_{kt_{ij}}, i = 1, \cdots, N, j = 1, \cdots, N_i\}$. Here t_{ij} is the time of j-th trading day within $[t_i, t_{i+1}]$ (thus N_i is the total number of trading days within this period). Then, $\mathbf{X}_{t_{ij}} = (X_{kt_{ij}})_{k=1}^K$ is the vector of asset prices on t_{ij}.

Accordingly, (6) is discretized as the following:

$$\Delta D_i = \sum_{j=1}^{N_i} \delta_{ij} \mu(\mathbf{X}_{t_{ij}}) + \sigma\sqrt{\delta_i}\xi_i, \quad i = 1, \cdots, N. \tag{7}$$

On the right hand side, the first term (the sum) is the quadrature approximation for the time-integral in (6), and δ_{ij} are the associated weights which only depend on the time points t_{ij}. The second term uses the distribution of Brownian increments, and ξ_i are i.i.d. standard normal random variables (which are also independent from $\mathbf{X}_{t_{ij}}$); $\delta_i = t_{i+1} - t_i$ are also constants.

Variable Selection. There is a rich set of variable selection methods; refer to Guyon and Elisseeff (2003). Among others, linear predictor with a variable selection feature is immediately applicable. Specifically, this means imposing $\mu(\mathbf{x}) = \beta_0 + \sum_{k=1}^K \beta_k x_k$ (for variable selection purpose) and using algorithms to reduce some of the β_k to 0; then, variables associated with non-zero β_k are selected. One commonly used approach is ℓ_1-regularization, and here it amounts to solving:

$$\min_{\beta_0,\beta_1,\cdots,\beta_K} \sum_{i=1}^N \frac{1}{\delta_i} \Big[\Delta D_i - \beta_0 \Big(\sum_{j=1}^{N_i} \delta_{ij}\Big) - \sum_{k=1}^K \beta_k \Big(\sum_{j=1}^{N_i} \delta_{ij} X_{kt_{ij}}\Big)\Big]^2 + \lambda \sum_{k=1}^K |\beta_k|;$$

$\lambda > 0$ is a chosen tuning parameter that can be set to control the number of selected assets. The formulation above is essentially a Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani (1996)) with $\sum_{j=1}^{N_i} \delta_{ij} X_{kt_{ij}}$, the daily average asset price over the i-th period, as the k-th predictor.

Function Learning. Based on (7), finding the functional form of $\mu(\cdot)$ in the selected assets can be cast into the following generic least squares problem:

$$\min_{\mu \in \boldsymbol{\mathcal{U}}} \quad \sum_{i=1}^{N} \frac{1}{\delta_i} \Big[\Delta D_i - \sum_{j=1}^{N_i} \delta_{ij} \mu(\mathbf{X}_{t_{ij}}) \Big]^2; \tag{8}$$

where $\boldsymbol{\mathcal{U}}$ is the function space of candidates of $\mu(\cdot)$. The problem above can be further reduced to structured regression by imposing restrictions on $\boldsymbol{\mathcal{U}}$. Specifically, $\mu(\cdot)$ is represented by a linear combination of basis functions (Hastie *et al.* (2009)):

$$\mu(\mathbf{x}) = \sum_{m=1}^{M} \beta_m b_m(\mathbf{x}; \theta). \tag{9}$$

Each $b_m(\mathbf{x}; \theta)$ is a basis function with a specified functional form parameterized by θ. Rearranging the terms, (8) reduces to:

$$\min_{\beta_m, \theta} \quad \sum_{i=1}^{N} \frac{1}{\delta_i} \Big[\Delta D_i - \sum_{m=1}^{M} \beta_m \Big(\sum_{j=1}^{N_i} \delta_{ij} b_m(\mathbf{X}_{t_{ij}}; \theta) \Big) \Big]^2. \tag{10}$$

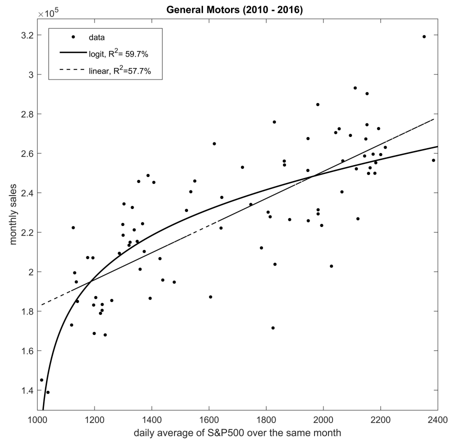
Common choices of the basis functions include polynomials, logarithm, and logit functions.

In Figure 1, we illustrate a preliminary study of General Motors's sales versus S&P500. In the graph on the left side of the figure, we examine the relationship between the firm's monthly sales (y-axis) and the daily (average) S&P500 within the same month (x-axis). A linear fit appears to capture the relationship quite well, and the learning is further improved by using a logit function. The plot presented in the graph slightly deviates from the demand model in (6) since the predictor is $\mu(\int_0^T X_t dt)$ as opposed to $\int_0^T \mu(X_t) dt$. Results for implementing (6) (with discretization in (7)) are collected in the table next to the graph. Note, it is obvious that the linear model in the table coincides with that in the plot. For logit model, there is a slight modification to the functional form but the improvement from the linear fit is not altered.

2.2 Incorporating the Hedging Strategy

With the new demand model, in addition to the one-time capacity decision that must be made at $t = 0$, we can carry out a real-time hedging strategy throughout the planning horizon $[0, T]$, by taking

Figure 1: Monthly sales of General Motors versus S&P500. In the graph, y-axis represents monthly sales (in units) and x-axis represents daily average of S&P500 within the same month. For the graph, "linear" stands for $\mu(x) = \beta_0 + \beta_1 x$ (x is the daily average of S&P500 over a month); "logit" stands for $\mu(\hat{x}) = \beta_0 + \beta_1 \log(\hat{x}/(1-\hat{x}))$, where $0 \leq \hat{x} \leq 1$ is the transformed x: $\hat{x} = (x-L)/(U-L)$ with $(L, U) = (1000, 5000)$. For the table, "linear" stands for $\mu(x) = \beta_0 + \beta_1 x$ (x is the daily quote of S&P500); "logit" stands for $\mu(\hat{x}) = \beta_0 + \beta_1 \log(\hat{x}/(1-\hat{x}))$ with $\hat{x}$ defined in the same way as above.



model	R^2
linear	57.7%
logit	60.8%

positions on the underlying asset(s) involved in the rate function. So, the new problem formulation is as follows, where to facilitate exposition, we consider a single asset X_t (instead of the vector $\mathbf{X}_t$):

$$\inf_{Q\geq 0,\vartheta} \mathsf{E}\Big\{ \big[m - H_T(Q) - \chi_T(\vartheta)\big]^+ \Big\} \tag{11}$$
$$\text{s.t.} \quad \chi_t := \int_0^t \theta_s dX_s \geq -C, \quad \theta_t \in \mathcal{F}_t^X, \qquad t \in [0, T].$$

Here, $m > 0$ is the given target and $H_T(Q)$ is the payoff due to production, both following what have been specified in §1.1; $\chi_T(\vartheta)$ is the terminal wealth from the hedging strategy $\vartheta := \{\theta_t,\ t \in [0, T]\}$, where θ_t denotes the position taken at time t on the underlying asset with price X_t. For simplicity, assume X_t follows the geometric Brownian motion, the standard asset price model; with additional complexity, X_t can be allowed to follow a general diffusion process.

Note, by restricting the hedging strategy to $\{\mathcal{F}_t^X\}$, the filtration generated by $\{X_t\}$, we are pursuing a formulation with "partial information". This is often a good representation of reality: in practice the hedging/trading decision is often a real-time decision taking input from the financial market, as embodied in the filtration $\{\mathcal{F}_t^X\}$. It would be unrealistic to assume that one could simultaneously also keep track of demand projection, which typically involves piecing together disperse information garnered from polling the sales force, and is hence updated much less frequently, at much longer time scales. Thus, in the research literature partial information is an important model in its own right; see, e.g., Caldentey and Haugh (2006).

Optimal Hedging Strategy To solve the above problem, we first assume Q is given. Then, the hedging problem is solved through a convex duality approach (refer to, e.g., Pham (2009)). Applying Jensen's inequality, along with conditional expectation on X_T and $A_T := \int_0^T \mu(X_t)dt$, we first turn the real-time hedging problem into a *static* optimization problem:

$$\min_{V_T} \mathsf{E}\big[(m - H_T - V_T)^+\big] \quad \text{s.t.} \quad V_T \geq -C, \quad \mathsf{E}^M(V_T) \leq 0\,. \tag{12}$$

Note the objective function in the above optimization problem can be shown to be a lower bound of the original objective in (11). On the other hand, the constraint $\mathsf{E}^M(V_T) \leq 0$ was not present in the original problem. It follows from χ_t being a P^M-supermartingale. (Here P^M denotes the probability under the risk-neutral measure — as opposed to the physical measure, without the superscript M; and E^M denotes the expectation under the risk-neutral measure.) This additional constraint serves the purpose of closing the duality gap: the problem in (12) can be shown to be equivalent to the original problem in (11).

The above dual problem is solved by a standard Lagrangian multiplier approach, and the solution is:

$$V_T^* = (p+c)(Q - \hat{D}_T^+)^+ + (m - pQ + C)\mathbf{1}\{\lambda^* Z_T \leq 1\} - C, \tag{13}$$

where λ^* is the (positive) Lagrangian multiplier, $Z_T := \frac{d\mathsf{P}^M}{d\mathsf{P}} = e^{-\eta B_T - \eta^2 T/2}$ is the Radon-Nikodym derivative, and $\hat{D}_T := A_T + \sigma\sqrt{T}\Phi^{-1}(\lambda^* Z_T)$ is the "proxy" for D_T (as the latter is not accessible due to partial information) with Φ^{-1} being the inverse distribution function of the standard normal random variable. Once $V_T^*(=\chi_T^*)$ is derived, the optimal hedging strategy θ_t^* follows from Itô's Lemma along with martingale representation theorem.

Optimal Production Quantity Under the optimal hedging strategy corresponding to V_T^* in (13), the minimized shortfall (with Q given) can be expressed as

$$s(m,Q) = (p+c)\mathsf{E}\big[(Q \wedge \hat{D}_T^+ - D_T^+)^+\big] + (m - pQ + C)\mathsf{P}(\lambda^* Z_T \geq 1). \quad (14)$$

It can be shown (with considerable effort, as both λ^* and $\hat{D}_T$ also depend on Q) that the above is a convex function of Q. Thus, finding the optimal solution $Q^*(m)$, *jointly* with the optimal hedging, is a readily solved convex minimization problem. To make it even better, a universal upper bound (i.e., independent of m) on the optimal Q can be explicitly identified, and this further facilitates the line search for $Q^*(m)$. Furthermore, it can be shown that $s(m, Q^*(m))$ is increasing in m, hence constitutes an efficient frontier — setting a higher target will lead to a higher shortfall.

3 Results and Insights

From the V_T^* expression in (13), observe the following:

- $m - H_T(Q) = (p+c)(Q - D_T^+)^+ + (m - pQ)$ is the remaining gap (from the target) after the payoff from production.
- The first term of V_T^*, $(p+c)(Q - \hat{D}_T^+)^+$, can be viewed as a "put option". It tries to close the first part of the gap, $(p+c)(Q - D_T^+)^+$, but needs to use $\hat{D}_T$ as a surrogate for D_T due to partial information.
- The second term of V_T^* is a "digital option", $(m - pQ + C)\mathbf{1}\{\lambda^* Z_T \leq 1\}$. (Note $\lambda^* Z_T \leq 1$ if and only if $\hat{D}_T < \infty$.) It aims to close the other part of the gap (after subtracting C).

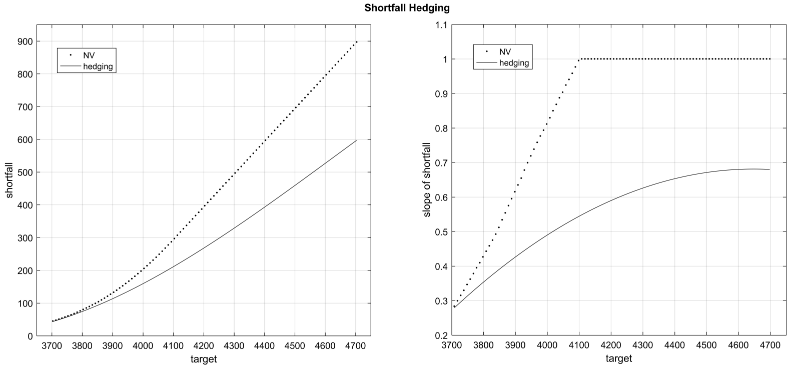
In summary, the shortfall hedging strategy takes the form of two options, a put option and a digital option, and the underlying for both is the "surrogate demand" $\hat{D}_T$, necessitated by the partial information on the real demand. Since the risk measure is shortfall, both options are designed to contribute more terminal wealth, so as to supplement production and help closing the gap from the target.

Note here any loss from the hedging strategy is limited by C, in a pathwise (i.e., almost sure) sense. On the other hand, sensitivity/asymptotic analysis will show that the upside of hedging is closely tied to m, which in our application setting is usually set at no greater than pQ^{NV}. This is reassuring as it guarantees that the hedging strategy will not result in any extreme swings on profit/revenue (in either direction).

Next, we compare the efficient frontiers of the NV model (with a shortfall objective) and our new model highlight above. In Figure 2, the left panel, we plot the two efficient frontiers. Clearly the new model results in a significant improvement: for any target value (x-axis), the corresponding shortfall (y-axis) is substantially lower. Of particular interest is the range of target values (moderately) above 4100, which coincides with $m = pQ^{\mathrm{NV}}$ in this problem instance. Recall from (5), for $m \geq pQ^{\mathrm{NV}}$, the shortfall of the NV model will increasing linearly in m, at a slope of 1; and this is confirmed here in the figure. In particular, note the right panel, which plots the rate of increase (first derivative) of the frontier curves. In contrast, with the addition of the hedging strategy, the increase of the shortfall is at a substantially lower slope (peaked at slightly below 0.7, and this is achieved via a modest hedging budget of $C = 0.1m$).

Another insightful result is this. One might be concerned that by minimizing a shortfall objective, we may do poorly in the expected profit at T, the end of the horizon. It turns out this concern is unwarranted. It can be shown (Proposition 9 in Wang and Yao (2017b)) that the shortfall-minimizing objective leads to a jointly optimized production and risk-hedging decision that will contribute *more profit* than the profit-maximizing production-only decision. In other words, even if one chooses to do profit maximization, what can be achieved (via a production-only decision) will be inferior to minimizing a shortfall risk measure along with a hedging strategy. Indeed, the improvement in expected total terminal wealth (at T), above and beyond the NV model,

Figure 2: Efficient Frontiers. The left panel shows the shortfall as an increasing function of the target level, for both the NV model and the new model with risk hedging. The right panel shows the derivatives (w.r.t. the target) of the two curves.



can be lower-bounded as follows:

$$\mathsf{E}[H_T(Q^*(m)) + V_T^*] - \mathsf{E}[H_T(Q^{\mathrm{NV}}(m))] \geq \beta(m - pQ^{\mathrm{NV}})^+ + C\psi > 0,$$

where $\beta \in [0,1]$ and $\psi \in [0,\lambda]$ are two positive parameters that only depend on given data, via λ (recall λ is the Lagrangian multiplier from solving the hedging problem).

4 Further Research

So far, risk originating from the supply side is not considered — we have assumed a constant production cost c (per unit). In reality, this cost may also depend on certain financial assets with volatile prices. Examples include commodities, taking into account that a bulk of cost is due to raw materials; another example is exchange rates, given the global nature of supply chains. Thus, a better model is to replace c with $c(\mathbf{X}_t)$, in the same spirit as the demand rate function $\mu(\mathbf{X}_t)$ in (6). (Without loss of generality, $\mathbf{X}_t$ can be a vector that collects the union of assets that impact cost and demand respectively.) This way, certain functional dependence as well as statistical correlation can be built into the cost and demand models, and will be reflected accordingly in the hedging strategy. We are currently working on this more general model, and expect to report results in the near future.

References

Caldentey, R. and Haugh, M., Optimal Control and Hedging of Operations in the Presence of Financial Markets. *Math. Opns. Res.* **31** (2006), 285-304.

Guyon, I. and Elisseeff, A., An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **3** (2003), 1157-1182.

Hastie, T., Tibshirani, R. and Friedman., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* (2009) Springer-Verlag New York.

Pham, H., *Continuous-time Stochastic Control and Optimization with Financial Applications.* (2009) Springer-Verlag, Berlin/Heidelberg.

Tibshirani, R., Regression Shrinkage and Selection via the Lasso. *J. Royal. Statist. Soc B.* **58(1)** (1996), 267-288.

Wang, L. and Yao, D.D., Production with Risk Hedging: Optimal Policy and Efficient Frontier. *Opns. Res.* **65** (2017a), 1095-1113.

Wang, L. and Yao, D.D., Production Planning with Shortfall Hedging under Partial Information and Budget Constraint. *Preprint* (2017b).