Financial Contracts and Asset Selling under Ambiguity

Notation. Random variables are designated by tilde signs (*e.g.*, $\tilde{\Delta}$) and their realizations by the same symbols without tildes (e.g., Δ). For any Borel set $\mathcal{A} \in \mathcal{B}(\mathbb{R})$, we use $\mathcal{P}(\mathcal{A})$ to represent the set of all distributions on $\mathcal{A}$ and δ_Δ to denote the Dirac point mass at Δ. We denote by $\boldsymbol{e}_i$ the unit vector of appropriate length whose i^{th} component is 1.

1. Problem Setup

We consider the following dynamic asset selling problem. A seller wishes to sell a divisible asset over T period horizon. The asset is financed by the equity v and an initial debt b whose payment schedule is denoted by the vector $\boldsymbol{d} = (d_1, d_2, \ldots, d_T)$, where $d_t \geq 0$ represents the payment at period $t \in [T]$. In each period $t \in [T]$, the seller observes a stochastic price p_t and decides how much of the remaining asset x_t to sell. If she sells q_t amount, she receives revenue $p_t q_t$ and pays installment d_t. If the seller fails to cover the debt payment d_t, she goes bankrupt, and all remaining asset is liquidated immediately by the lender (bank).

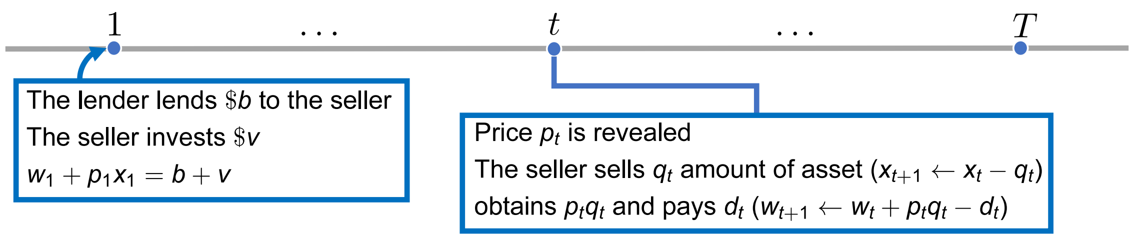
We model the stochastic price by $p_t(\tilde{\Delta}_t, p_{t-1}) = \tilde{\Delta}_t p_{t-1}$ for all $t \in \{2, \ldots, T\}$, where $\tilde{\Delta}_t$ is governed by a probability distribution. We assume that the seller knows a nominal distribution $\hat{\mathbb{P}}_t$ (*e.g.*, computed from historical price data) and that $\hat{\mathbb{P}}_t$ is a discrete distribution with N_t atoms, *i.e.*, $\hat{\mathbb{P}}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \delta_{\hat{\Delta}_t^i}$ where $\hat{\Delta}_t^i \in [\underline{\Delta}_t, \overline{\Delta}_t]$ for all $i \in [N_t]$. The seller does not have full confidence in the estimated distribution and believes that $\tilde{\Delta}_t \sim \mathbb{P}_t$ where $\mathbb{P}_t$ reside within the Wasserstein ball

$$\mathbb{B}_t(\hat{\mathbb{P}}_t, \epsilon_t) = \{\mathbb{P}_t \in \mathcal{P}([\underline{\Delta}_t, \overline{\Delta}_t]) : W_1(\mathbb{P}_t, \hat{\mathbb{P}}_t) \leq \epsilon_t\},$$

with radius $\epsilon_t \geq 0$, where W_1 denotes the type-1 Wasserstein distance

$$W_1(\mathbb{P}_t, \hat{\mathbb{P}}_t) = \inf_{\pi \in \Pi(\mathbb{P}_t, \hat{\mathbb{P}}_t)} \int |\Delta - \Delta'| \pi(\mathrm{d}\,\Delta, \mathrm{d}\,\Delta').$$

Figure 1 Timeline.



The set $\Pi(\mathbb{P}_t, \hat{\mathbb{P}}_t)$ is the set of all joint distributions of Δ and Δ' with marginals $\mathbb{P}_t$ and $\hat{\mathbb{P}}_t$, respectively. The decision variable π can be interpreted as a transportation plan for moving a mass distribution described by $\mathbb{P}_t$ to another described by $\hat{\mathbb{P}}_t$, where the norm $|\cdot|$ encodes the transportation costs. The Wasserstein ball $\mathbb{B}_t(\hat{\mathbb{P}}_t, \epsilon_t)$ therefore contains all distributions supported on $[\underline{\Delta}_t, \overline{\Delta}_t]$ that can be obtained by reshaping the nominal distribution $\hat{\mathbb{P}}_t$ at a transportation cost of at most ϵ_t. Intuitively, parameter ϵ_t quantifies the confidence of the seller in the nominal distribution, *e.g.*, $\epsilon_t = 0$ indicates that the seller fully trusts the nominal distribution whereas $\epsilon_t \to \infty$ indicates the opposite. For practical relevance, we henceforth assume that prices are always nonnegative, *i.e.*, $\underline{\Delta}_t \geq 0$ for all $t = 1, \dots, T$. Also, for the simplicity of notation we sometimes supress the dependence of the Wasserstein ball on $\hat{\mathbb{P}}_t$ and ϵ_t and write $\mathbb{B}_t$ when this dependence does not prevent the understanding of the arguments.

The seller is ambiguity averse in the sense that she wishes to maximize her worst-case expected residual revenue in view of all distributions in the ambiguity set. The respective dynamic program of the seller is formulated as

$$
\begin{aligned}
&V_t(x_t, w_t, p_t; \boldsymbol{d}) \\
&= \begin{cases} \displaystyle\sup_{\frac{(d_t - w_t)^+}{p_t} \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t - q_t, w_t + p_t q_t - d_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \text{ if } d_t - w_t \leq p_t x_t \\ 0 \;\; \text{otherwise,} \end{cases} \qquad \forall t \in [T] \\
&V_{T+1}(x_{T+1}, w_{T+1}, p_{T+1}; \boldsymbol{d}) = w_{T+1},
\end{aligned} \tag{1}
$$

where w_t denotes the capital that the seller holds at period $t \in [T]$. Note that the seller should sell at least $(d_t - w_t)^+/p_t$ amount of asset to pay the debt installment d_t. If $(d_t - w_t)^+/p_t \leq x_t$, the seller chooses $q_t \in [(d_t - w_t)^+/p_t, x_t]$ so as to maximize the worst-case expected payoff. Otherwise, the seller fails to meet the debt payment and goes bankrupt in which case her payoff vanishes to zero. Figure 2 visualizes the timeline and state transitions.

We assume that the bank encounters the same ambiguity as the seller. This assumption is realistic especially if the seller and bank have access to the same historical price data. Later in

Section 2.1.1, we also discuss the implications of information asymmetry. The bank is ambiguity averse and evaluates its worst-case expected payoff by

$$
\begin{aligned}
&B_t(x_t, w_t, p_t; \boldsymbol{d}) \\
&= \begin{cases} \inf\limits_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[B_{t+1}(x_t - q_t^\star, w_t + p_t q_t^\star - d_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] & \text{if } d_t - w_t \le p_t x_t \\ w_t + p_t x_t + \sum_{k=1}^{t-1} d_t & \text{otherwise,} \end{cases} \qquad \forall t \in [T] \\
&B_{T+1}(x_{T+1}, w_{T+1}, p_{T+1}; \boldsymbol{d}) = \sum_{t=1}^{T} d_t,
\end{aligned}
$$

where $q_t^\star$ denotes the seller's optimal policy.

Recall that the asset is financed by the seller's equity v and an initial debt b borrowed from the bank. A debt contract $(b, \boldsymbol{d})$, which consists of a debt value b and its repayment plan $\boldsymbol{d}$, is called feasible if the following two conditions are satisfied.

$$
\begin{aligned}
&\mathcal{V}_1(x_1, w_1, p_1; \boldsymbol{d}) \ge v \quad \text{(individual rationality of the seller)} \\
&\mathcal{B}_1(x_1, w_1, p_1; \boldsymbol{d}) \ge b \quad \text{(individual rationality of the lender)}
\end{aligned} \tag{2}
$$

The first condition ensures that the seller's worst-case expected revenue exceeds her investment v, and the second condition ensures that the bank's worst-case payoff exceeds the debt value b. The seller and bank agree on the terms of the debt contract $(b, \boldsymbol{d})$ only if inequalities (2) hold.

In the following sections, we will investigate the optimal policy of the seller and the worst-case payoff of the bank as well as how these change with respect to the ambiguity level. We moreover will discuss implications of ambiguity on the terms of debt contracts. Before that, we present the following preliminary result that will be used throughout the paper. We drop the index t in the following proposition as this is not relevant.

PROPOSITION 1 **(Worst-Case Distribution)**. *If $f : \mathbb{R} \to \mathbb{R}$ is nonincreasing and concave then*

$$
\sup_{\mathbb{P} \in \mathbb{B}(\hat{\mathbb{P}}, \epsilon)} \mathbb{E}_{\mathbb{P}}[f(\tilde{\Delta})] = \frac{1}{N} \left(\sum_{i : \hat{\Delta}^i \ge \Delta^\star} f(\Delta^\star) + \sum_{i : \hat{\Delta}^i < \Delta^\star} f(\hat{\Delta}^i) \right),
$$

where $\Delta^\star = \max\{\underline{\Delta}, \min\{\Delta : \frac{1}{N} \sum_{i=1}^{N} (\hat{\Delta}^i - \Delta)^+ \le \epsilon\}\}$. The optimal value is attained by the extremal distribution

$$
\mathbb{P}^\star = \frac{1}{N} \sum_{i : \Delta^i < \Delta^\star} \delta_{\Delta^i} + \frac{1}{N} \sum_{i : \Delta^i \ge \Delta^\star} \delta_{\Delta^\star}.
$$

2. Optimal Asset Selling Policy and Its Implications on the Terms of Debt Contract

In this section, we show that seller's optimal selling policy is a threshold policy and characterize it analytically. We moreover discuss its implications on the terms of the debt contract. We first present the following structural results that will serve to our purpose.

Lemma 1. *The value function given in* (1) *satisfies the following.*

(i) $V_T(x_T, w_T, p_T; \boldsymbol{d}) = (w_T + p_T x_T - d_T)^+$

(ii) $V_t(0, w_t, p_t; \boldsymbol{d}) = (w_t - \sum_{j=t}^{T} d_j)^+$ *for all* $t \in [T]$

(iii) $V_t(x_t, w_t, p_t; \boldsymbol{d}) = V_t(cx_t, w_t, p_t/c; \boldsymbol{d})$ *for all* $c \geq 0$ *and* $t \in [T]$

(iv) $V_t(x_t, w_t, p_t; \boldsymbol{d})$ *is jointly nondecreasing in* (w_t, p_t).

Assertion (i) encodes that the seller sells all remaining asset at period T. Assertion (ii) shows that the cash is carried over all periods after the debt installments are discounted provided that there is no leftover asset. Assertion (iii) implies that scaling the asset quantity and the price by the same constant does not change the value. Assertion (iv) formalizes that the seller's revenue is nondecreasing in cash and price as one would expect. We emphasize that Lemma 1 holds not only for Wasserstein ambiguity but for any ambiguity set.

Next, we characterize the optimal policy of the seller.

Theorem 1. *There exists an optimal solution to the dynamic program* (1) *that satisfies*

$$q_t^\star \in \{x_t\} \cup \left\{ \frac{(\sum_{j=t}^{\ell} d_j - w_t)^+}{p_t} : \ell = t, \dots, \bar{t} \right\}, \tag{3}$$

where $\bar{t} = \max\{\ell \in \{t, \dots, T\} : (\sum_{j=t}^{\ell} d_j - w_t)/p_t < x_t\}$ *for all* $t \in [T]$*. In other words, the seller sells either all asset or enough to pay all consecutive debt installments up to period* $\ell \in \{t, \dots, \bar{t}\}$.

Unlike Lemma 1, the proof of Theorem 1 relies on the choice of Wasserstein ball as the ambiguity set. Even though we were not able to prove Theorem 1 under general ambiguity sets, we discuss that Theorem 1 remains to hold under other reasonable ambiguity sets and limitations of generalizing it to any ambiguity set in Remark 2 after the proof in appendix.

In the proof of Theorem 1, we also proved that the value function $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is convex in p_t. We formalize this result with the following lemma.

Lemma 2. *The value function* $V_t(x_t, w_t, p_t; \boldsymbol{d})$ *is convex in* p_t.

Remark 1. The proof of Theorem 1 is inspired by the proof of Proposition 6 in Ahn et al. (2019). Our proof similarly relies on showing by induction that the value function $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is jointly piece-wise convex in (x_t, w_t) for any $t \in [T]$. However, we additionally need to show that $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is convex in p_t during the induction, which was not needed in Ahn et al. (2019) due to absence of ambiguity.

2.1. Single Debt Payment

We now focus on the case where there is a single debt payment at the end of time horizon in which case (3) simplifies to

$$q_t^\star \in \begin{cases} \{0, x_t\} \cup \left\{ \frac{(d_T - w_t)^+}{p_t} \right\} & \text{if } \frac{(d_T - w_t)^+}{p_t} \le x_t \\ \{0, x_t\} & \text{otherwise.} \end{cases}$$

The following proposition shows that the optimal quantity always amounts to either 0 or x_t.

PROPOSITION 2. *If $\boldsymbol{d} = d_T \boldsymbol{e}_T$, then an optimal solution to the dynamic program* (1) *is given by*

$$q_t^\star = \begin{cases} x_t & \textit{if } w_t + p_t x_t - d_T \ge \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\ 0 & \textit{otherwise.} \end{cases}$$

We emphasize that the proof of Proposition 2, unlike Theorem 1, does not rely on the choice of Wasserstein ambiguity and that Proposition 2 holds for any ambiguity set. As the value function is nondecreasing (Lemma 1(*iv*)) and convex (Lemma 2) in p_t and by Proposition 1, we have

$$\inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; d_T)] = \mathbb{E}_{\mathbb{P}^\star_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; d_T)],$$

where $\mathbb{P}^\star_{t+1} = \frac{1}{N} \sum_{i: \Delta^i_{t+1} < \Delta^\star_{t+1}} \delta_{\Delta^i_{t+1}} + \frac{1}{N} \sum_{i: \Delta^i_{t+1} \ge \Delta^\star_{t+1}} \delta_{\Delta^\star_{t+1}}$. However, we avoid writing this equality in Proposition 2 as the proposition does not depend on the specifics of Wasserstein ball but holds for any ambiguity set.

Note that the optimal quantity $q_t^\star$ is a function of ambiguity level ϵ_t even though we suppress this dependence in the exposition for ease of notation. The Wasserstein ambiguity set $\mathbb{B}_{t+1}(\hat{\mathbb{P}}_{t+1}, \epsilon_{t+1})$ contains all distributions whose Wasserstein distance from the nominal distribution $\hat{\mathbb{P}}_{t+1}$ does not exceed the ambiguity level ϵ_{t+1}, and $\mathbb{B}_{t+1}(\hat{\mathbb{P}}_{t+1}, \epsilon_{t+1}) \subseteq \mathbb{B}_{t+1}(\hat{\mathbb{P}}_{t+1}, \epsilon'_{t+1})$ if $\epsilon_{t+1} \le \epsilon'_{t+1}$. Thus, for all $\epsilon_{t+1} \le \epsilon'_{t+1}$,

$$\inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}(\hat{\mathbb{P}}_{t+1}, \epsilon_{t+1})} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \ge \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}(\hat{\mathbb{P}}_{t+1}, \epsilon'_{t+1})} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})].$$

We therefore have the following corollary of Proposition 2.

COROLLARY 1. *If $\boldsymbol{d} = d_T \boldsymbol{e}_T$, then the optimal quantity $q_t^\star$ is nondecreasing in ϵ_{t+1}.*

Corollary 1 implies that the seller tends to sell the asset earlier if she encounters more ambiguity. Proposition 2 also allows us to re-express the bank's worst-case expected payoff.

COROLLARY 2. *If $\boldsymbol{d} = d_T \boldsymbol{e}_T$, the bank's worst-case expected payoff simplifies to*

$$\begin{aligned} & B_t(x_t, w_t, p_t; \boldsymbol{d}) \\ & = \begin{cases} d_T & \textit{if } w_t + p_t x_t - d_T \ge \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\ \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[B_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] & \textit{otherwise} \end{cases} \quad \forall t \in [T-1] \qquad (4) \\ & B_T(x_T, w_T, p_T; \boldsymbol{d}) = \min\{w_T + p_T x_T, d_T\}. \end{aligned}$$

Corollary 2 implies the following. If the seller cannot meet the debt payment d_T by selling all asset at time period t, *i.e.*, $w_t + p_t x_t < d_T$, then the bank's worst-case payoff

$$B_t(x_t, w_t, p_t; \boldsymbol{d}) = \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}(\hat{\mathbb{P}}_{t+1}, \epsilon_{t+1})} \mathbb{E}_{\mathbb{P}_{t+1}}[B_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]$$

and is therefore nonincreasing in the ambiguity level. If the seller can meet the debt payment d_T by selling all asset at time period t, *i.e.*, $w_t + p_t x_t < d_T$, then the bank's worst-case payoff is nonincreasing up to a threshold ambiguity level

$$\bar{\epsilon}_{t+1} = \max\left\{0, \min\left\{\epsilon : w_t + p_t x_t - d_T \geq \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}(\hat{\mathbb{P}}_{t+1}, \epsilon_{t+1})} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})\right\}\right\}$$

where the seller changes her optimal policy $q_t^\star$ from 0 to x_t and then amounts to d_T.

This observation can perhaps be interpreted as follows. A larger ambiguity naturally means less understanding of the future prices and creates risk for both the seller and the bank. The bank encounters this risk arised due to ambiguity only if the debt repayment relies on the future prices, *i.e.*, if the success of the seller's repayment depends on her revenues in the future. Sufficiently large ambiguity can force the seller to adopt riskless decisions. Encountering too much ambiguity, the seller in fact might prefer to sell everything today to ensure the debt repayment. In this case, ambiguity plays no role for the bank's payoff anymore as the payment is guaranteed.

2.1.1. Information Asymmetry. We now investigate the information asymmetry among the seller and bank, *i.e.*, the seller's and bank's ambiguity are modeled by the potentially different ambiguity sets $\mathbb{B}_t^{\mathrm{s}}$ and $\mathbb{B}_t^{\mathrm{b}}$, respectively for all $t \in \{2, \ldots, T\}$. We assume that the agents do not know each other's ambiguity set.

The seller's worst-case expected residual revenue does not depend on the bank's ambiguity set whereas the bank's worst-case expected payoff depends on the seller's optimal sales quantity decision and therefore her ambiguity set. If the bank only knows that the seller's ambiguity set contains the true distribution, we model the worst-case payoff of the bank as

$$\begin{aligned}
&B_t(x_t, w_t, p_t; \boldsymbol{d}) \\
&= \inf_{\substack{\mathbb{B}^{\mathrm{s}}_{t+1} \subseteq \mathcal{P}([\underline{\Delta}_t, \overline{\Delta}_t]): \\ \mathbb{B}^{\mathrm{s}}_{t+1} \cap \mathbb{B}^{\mathrm{b}}_{t+1} \neq \emptyset}} \begin{cases} d_T & \text{if } w_t + p_t x_t - d_T \geq \inf_{\mathbb{P}_{t+1} \in \mathbb{B}^{\mathrm{s}}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\ \inf_{\mathbb{P}_{t+1} \in \mathbb{B}^{\mathrm{b}}_{t+1} \cap \mathbb{B}^{\mathrm{s}}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[B_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] & \text{otherwise} \end{cases} \quad \forall t \in [T-1] \\
&B_T(x_T, w_T, p_T; \boldsymbol{d}) = \min\{w_T + p_T x_T, d_T\}.
\end{aligned} \tag{5}$$

Specifically, the bank evaluates its worst-case payoff in view of all ambiguity sets the seller can have. The seller's ambiguity set should have a nonempty intersection with the one of the bank as both contains the true distribution. Moreover, the bank knows that the true value distribution lies in this intersection so can evaluate its worst-case expected future payoff only over the intersection.

Note that the bank's worst-case payoff can only be worse under information asymmetry than under information symmetry, *i.e.*, when $\mathbb{B}_t^{\mathrm{s}} = \mathbb{B}_t^{\mathrm{b}}$ for all $t \in \{2, \dots, T\}$ as this constitutes a feasible solution to the infimum taken over all ambiguity sets the seller can have and yields to the same payoff value as (4).

...

As $B_t(x_t, w_t, p_t; \boldsymbol{d}) \le d_T$, we can reformulate this worst-case payoff as

$$
\begin{aligned}
&B_t(x_t, w_t, p_t; \boldsymbol{d}) \\
&= \begin{cases} d_T \quad \text{if } w_t + p_t x_t - d_T \ge \sup\limits_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}^{\mathrm{b}}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\ \begin{cases} \inf\limits_{\mathbb{B}_{t+1}^{\mathrm{s}} \subseteq \mathbb{B}_{t+1}^{\mathrm{b}}} \inf\limits_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}^{\mathrm{s}}} \mathbb{E}_{\mathbb{P}_{t+1}}[B_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\ \text{s.t.} \quad w_t + p_t x_t - d_T < \inf\limits_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}^{\mathrm{s}}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \end{cases} \quad \text{otherwise} \end{cases} \qquad \forall t \in [T-1] \\
&B_T(x_T, w_T, p_T; \boldsymbol{d}) = \min\{w_T + p_T x_T, d_T\}.
\end{aligned}
$$

Insights.

- Seller's worst-case expected residual revenue is not affected by the information asymmetry.
- Bank's worst-case payoff is worse under information asymmetry than information symmetry.
- Both the bank's and seller's payoffs are nonincreasing in the level of their own ambiguity and not affected by each other's ambiguity level (as the agents do not have the information of the each other's ambiguity as well as ambiguity level).
- There will not occur information share as both bank and seller have incentive to lie to affect the terms of the debt contract. Specifically, it is better for both agents if the opponent has a higher payoff and, the agents can manipulate each other in this direction by lying.

2.2. Debt Contract Terms

In this section, we investigate how the debt contract terms, *i.e.*, debt value b and the payment schedule $\boldsymbol{d}$ are affected by the ambiguity.

First, suppose that debt value b as well as the total payment d^{tot} are fixed but the agents can freely choose how to schedule the payments. In other words, the seller and the bank are allowed to choose any $\boldsymbol{d}$ satisfying $\sum_{t=1}^{T} d_t = d^{\mathrm{tot}}$. Consider $T = 2$ for simplicity. By Theorem 1, we know that the seller sells either all asset or enough to pay the debt installment at period $T-1$. Thus the seller's worst-case residual revenue for any $\boldsymbol{d}$ satisfying $\sum_{t=1}^{T} d_t = d^{\text{tot}}$ and $p_{T-1}x_{T-1} \geq d_{T-1} - w_{T-1}$ amounts to

$$V_{T-1}(x_{T-1}, w_{T-1}, p_{T-1}; \boldsymbol{d}) = \max\Big\{ \underbrace{(w_{T-1} + p_{T-1}x_{T-1} - d^{\text{tot}})^+}_{\text{if } q^\star_{T-1} = x_{T-1}}, \underbrace{\inf_{\mathbb{P}_T \in \mathbb{B}_T} \mathbb{E}_{\mathbb{P}_T}[(w_{T-1} + (d_{T-1} - w_{T-1})^+ + \tilde{\Delta}_T(p_{T-1}x_{T-1} - (d_{T-1} - w_{T-1})^+) - d^{\text{tot}})^+]}_{\text{if } q^\star_{T-1} = (d_{T-1} - w_{T-1})^+/p_{T-1}} \Big\} \tag{6}$$

Note that this is the seller's worst-case residual revenue given that she can meet the debt installment at period $T-1$. Otherwise, her worst-case residual revenue amounts to zero.

Proposition 3. *The seller's worst-case residual revenue in* (6) *satisfies the following.*

(i) For $d_{T-1} \leq w_{T-1}$, $V_{T-1}(x_{T-1}, w_{T-1}, p_{T-1}; \boldsymbol{d})$ is constant in d_{T-1}.

(ii) For $d_{T-1} > w_{T-1}$, $V_{T-1}(x_{T-1}, w_{T-1}, p_{T-1}; \boldsymbol{d})$ is nonincreasing in d_{T-1}.

(iii) $V_{T-1}(x_{T-1}, w_{T-1}, p_{T-1}; \boldsymbol{d})$ is nonincreasing in d_{T-1}.

(iv) Any $d_{T-1} \in [0, w_{T-1}]$ maximizes $V_{T-1}(x_{T-1}, w_{T-1}, p_{T-1}; \boldsymbol{d})$.

Proposition 3 implies that if given the right the seller selects a payment schedule $\boldsymbol{d}$ such that $d_{T-1} \in [0, w_{T-1}]$. This choice prevents the early bankruptcy.

Proposition 4. *The seller's optimal choice between $q^\star_{T-1} = x_{T-1}$ and $q^\star_{T-1} = (d_{T-1} - w_{T-1})^+/p_{T-1}$ remains the same for all $\boldsymbol{d}$ satisfying $\sum_{t=1}^{T} d_t = d^{tot}$. In other words, the seller either sells all asset today or enough to pay the debt installment at period $T-1$ irrespective of d_{T-1}.*

Unlike seller, the bank can benefit from selecting $d_{T-1} > w_{T-1}$. For any $\boldsymbol{d}$ satisfying $\sum_{t=1}^{T} d_t = d^{\text{tot}}$, if $q^\star_{T-1} = (d_{T-1} - w_{T-1})^+/p_{T-1}$ then the bank's worst-case payoff is given by

$$B_{T-1}(x_{T-1}, w_{T-1}, p_{T-1}; \boldsymbol{d}) = \begin{cases} \inf_{\mathbb{P}_T \in \mathbb{B}_T} \mathbb{E}_{\mathbb{P}_T}[\min\{d^{\text{tot}}, w_{T-1} + (d_{T-1} - w_{T-1})^+ + \tilde{\Delta}_T(p_{T-1}x_{T-1} - (d_{T-1} - w_{T-1})^+)\}] & \text{if } p_{T-1}x_{T-1} \geq d_{T-1} - w_{T-1} \\ w_{T-1} + p_{T-1}x_{T-1} & \text{otherwise.} \end{cases} \tag{7}$$

If $q^\star_{T-1} = x_{T-1}$ then the bank's worst-case payoff amounts to $\min\{d^{\text{tot}}, w_{T-1} + p_{T-1}x_{T-1}\}$.

Proposition 5. *The bank's worst-case payoff ... Missing structural results for bank's payoff and optimal choice of d_{T-1}*

References

Ahn HS, Wang DD, Wu OQ (2019) Asset selling under debt obligations, forthcoming in Operations Research.

Jiang R, Guan Y (2018) Risk-averse two-stage stochastic program with distributional ambiguity. *Operations Research* .

Kuhn D, Esfahani PM, Nguyen VA, Shafieezadeh-Abadeh S (2019) Wasserstein distributionally robust optimization: Theory and applications in machine learning. *Operations Research & Management Science in the Age of Analytics*, 130–166 (INFORMS).

Postek K, Ben-Tal A, Den Hertog D, Melenberg B (2018) Robust optimization with ambiguous stochastic constraints under mean and dispersion information. *Operations Research* 66(3):814–833.

Appendix. Proofs

Proof of Proposition 1. If $\frac{1}{N}\sum_{i=1}^N(\hat{\Delta}^i-\underline{\Delta})^+=\frac{1}{N}\sum_{i=1}^N(\hat{\Delta}^i-\underline{\Delta})\leq\epsilon$, the ambiguity set $\mathbb{B}(\hat{\mathbb{P}},\epsilon)$ simplifies to a support-only ambiguity set. In this case, $\Delta^\star=\underline{\Delta}$, and the claim holds because the worst-case distribution is a Dirac point mass at $\underline{\Delta}$. We now focus on the non-straightforward case, *i.e.*, when $\frac{1}{N}\sum_{i=1}^N(\hat{\Delta}^i-\underline{\Delta})^+=\frac{1}{N}\sum_{i=1}^N(\hat{\Delta}^i-\underline{\Delta})>\epsilon$ in which case $\Delta^\star=\min\{\Delta:\frac{1}{N}\sum_{i=1}^N(\hat{\Delta}^i-\Delta)^+\leq\epsilon\}$. We have

$$\sup_{\mathbb{P}\in\mathbb{B}(\hat{\mathbb{P}},\epsilon)}\mathbb{E}_{\mathbb{P}}[f(\tilde{\Delta})]\leq\inf_{\gamma\geq 0}\gamma\epsilon+\frac{1}{N}\sum_{i=1}^N\sup_{\Delta\in[\underline{\Delta},\overline{\Delta}]}f(\Delta)-\gamma|\Delta-\hat{\Delta}^i| \tag{8}$$

where the inequality follows from weak duality, see, *e.g.*, Theorem 7 in Kuhn et al. (2019). To prove the claim, we show that the there exist feasible solutions to the optimization problems on both sides of the inequality (8) that result in the identical objective value $\frac{1}{N}\left(\sum_{i:\hat{\Delta}^i\geq\Delta^\star}f(\Delta^\star)+\sum_{i:\hat{\Delta}^i<\Delta^\star}f(\hat{\Delta}^i)\right)$. For ease of exposition, we assume that f is differentiable throughout this proof, which can be generalized to the non-differentiable f by interchanging derivatives by sub-differentials.

Step 1. Consider $\mathbb{P}^\star$ as a solution to $\sup_{\mathbb{P}\in\mathbb{B}(\hat{\mathbb{P}},\epsilon)}\mathbb{E}_{\mathbb{P}}[f(\tilde{\Delta})]$. This solution is feasible because its support is a subset of $[\underline{\Delta},\overline{\Delta}]$, and

$$W_1(\mathbb{P}^\star,\hat{\mathbb{P}}_t)=\inf_{\pi\in\Pi(\mathbb{P}^\star,\hat{\mathbb{P}}_t)}\int|\Delta-\Delta'|\pi(\mathrm{d}\,\Delta,\mathrm{d}\,\Delta')\leq\frac{1}{N}\sum_{i:\hat{\Delta}^i\geq\Delta^\star}|\Delta^\star-\hat{\Delta}^i|=\frac{1}{N}\sum_{i=1}^N(\hat{\Delta}^i-\Delta^\star)^+\leq\epsilon.$$

The first inequality follows from that the joint distribution

$$\pi=\frac{1}{N}\sum_{i=\hat{\Delta}^i<\Delta^\star}\delta_{(\hat{\Delta}^i,\hat{\Delta}^i)}+\frac{1}{N}\sum_{i=\hat{\Delta}^i\geq\Delta^\star}\delta_{(\Delta^\star,\hat{\Delta}^i)}$$

has marginals $\mathbb{P}^\star$ and $\hat{\mathbb{P}}_t$, and the second inequality follows from the definition of $\Delta^\star$. Note that the objective value of $\mathbb{P}^\star$ amounts to $\frac{1}{N}\left(\sum_{i:\hat{\Delta}^i\geq\Delta^\star}f(\Delta^\star)+\sum_{i:\hat{\Delta}^i<\Delta^\star}f(\hat{\Delta}^i)\right)$.

Step 2. Consider the solution $\gamma^\star=-f'(\Delta^\star)$ to the inf problem on the right-hand side of (8). This solution is feasible because the function f is nonincreasing, *i.e.*, $f'(\Delta)\leq 0$ for all Δ. Note that as f is a nonincreasing function and $\gamma^\star\geq 0$, we have

$$\sup_{\Delta\in[\underline{\Delta},\overline{\Delta}]}f(\Delta)-\gamma^\star|\Delta-\hat{\Delta}^i|=\sup_{\Delta\in[\underline{\Delta},\hat{\Delta}^i]}f(\Delta)-\gamma^\star(\hat{\Delta}^i-\Delta).$$

For any $i \in [N]$ such that $\hat{\Delta}^i < \Delta^\star$ and for any $\Delta \in [\underline{\Delta}, \hat{\Delta}^i]$, we have

$$f(\Delta) + f'(\Delta^\star)(\hat{\Delta}^i - \Delta) \le f(\Delta) + f'(\hat{\Delta}^i)(\hat{\Delta}^i - \Delta) \le f(\hat{\Delta}^i), \tag{9}$$

where the first inequality follows from the assumption that f is nonincreasing and concave, *i.e.*, f' is nonincreasing, and the second inequality again follows from the concavity of f, *i.e.*, $f(\Delta) - f(\hat{\Delta}^i) \le f'(\hat{\Delta}^i)(\Delta - \hat{\Delta}^i)$. Inequality (9) implies that

$$\sup_{\Delta \in [\underline{\Delta}, \hat{\Delta}^i]} f(\Delta) - \gamma^\star(\hat{\Delta}^i - \Delta) = f(\hat{\Delta}^i) \ \ \forall i \in [N] : \hat{\Delta}^i < \Delta^\star. \tag{10}$$

For any $i \in [N]$ such that $\hat{\Delta}^i \ge \Delta^\star$ and for any $\Delta \in [\underline{\Delta}, \hat{\Delta}^i]$, we have

$$f(\Delta) + f'(\Delta^\star)(\hat{\Delta}^i - \Delta) \le f(\Delta^\star) + f'(\Delta^\star)(\hat{\Delta}^i - \Delta^\star), \tag{11}$$

where the inequality follows from the concavity of f, *i.e.*, $f(\Delta) \le f(\Delta^\star) + f'(\Delta^\star)(\Delta - \Delta^\star)$. Inequality (11) implies that

$$\sup_{\Delta \in [\underline{\Delta}, \hat{\Delta}^i]} f(\Delta) - \gamma^\star(\hat{\Delta}^i - \Delta) = f(\Delta^\star) + f'(\Delta^\star)(\hat{\Delta}^i - \Delta^\star) \ \ \forall i \in [N] : \hat{\Delta}^i \ge \Delta^\star. \tag{12}$$

By (10) and (12), the objective value of $\gamma^\star$ in (P1) amounts to

$$\begin{aligned}
& -f'(\Delta^\star)\epsilon + \frac{1}{N}\left(\sum_{i:\hat{\Delta}^i \ge \Delta^\star} (f(\Delta^\star) + f'(\Delta^\star)(\hat{\Delta}^i - \Delta^\star)) + \sum_{i:\hat{\Delta}^i < \Delta^\star} f(\hat{\Delta}^i) \right) \\
& = -f'(\Delta^\star)\epsilon + \frac{1}{N} \sum_{i:\hat{\Delta}^i \ge \Delta^\star} f'(\Delta^\star)(\hat{\Delta}^i - \Delta^\star) + \frac{1}{N}\left(\sum_{i:\hat{\Delta}^i \ge \Delta^\star} f(\Delta^\star) + \sum_{i:\hat{\Delta}^i < \Delta^\star} f(\hat{\Delta}^i) \right) \\
& = \frac{1}{N}\left(\sum_{i:\hat{\Delta}^i \ge \Delta^\star} f(\Delta^\star) + \sum_{i:\hat{\Delta}^i < \Delta^\star} f(\hat{\Delta}^i) \right),
\end{aligned}$$

where the second equality holds because $\frac{1}{N}\sum_{i:\hat{\Delta}^i \ge \Delta^\star}(\hat{\Delta}^i - \Delta^\star) = \epsilon$ by definition of $\Delta^\star$. The claim thus follows. $\square$

Proof of Lemma 1. *Proof of assertion (i).* As $V_{T+1}(x_{T+1}, w_{T+1}, p_{T+1}; \boldsymbol{d}) = w_{T+1}$, we have

$$\begin{aligned}
V_T(x_T, w_T, p_T; \boldsymbol{d}) &= \begin{cases} \sup_{\frac{(d_T - w_T)^+}{p_T} \le q_T \le x_T} w_T + p_T q_T - d_T & \text{if } d_T - w_T \le p_T x_T, \\ 0 \ \ \text{otherwise} \end{cases} \\
&= \begin{cases} w_T + p_T x_T - d_T & \text{if } d_T - w_T \le p_T x_T, \\ 0 \ \ \text{otherwise.} \end{cases}
\end{aligned}$$

Note that $w_T + p_T x_T - d_T$ is always nonnegative under the condition $d_T - w_T \le p_T x_T$ and negative otherwise. We thus have $V_T(x_T, w_T, p_T; \boldsymbol{d}) = (w_T + p_T x_T - d_T)^+$.

Proof of assertion (ii). Note that $x_t = 0$ implies that $q_t = 0$ is the only feasible solution and, therefore, $x_\ell = 0$ and $q_\ell = 0$ for all $\ell \ge t$. The uncertain price thus plays no role. If the debt installment can be met then $w_{t+1} = w_t - d_t$ is carried to the next period whereas, otherwise, the seller goes bankrupt and $V_t(0, w_t, p_t; \boldsymbol{d}) = 0$. As this argument remains true for all $\ell \ge t$, we have $V_t(0, w_t, p_t; \boldsymbol{d}) = w_t - \sum_{j=t}^T d_j$ if $w_t \ge \sum_{j=t}^T d_j$ and $V_t(0, w_t, p_t; \boldsymbol{d}) = 0$ otherwise.

Proof of assertion (iii). We prove this assertion by induction. Note that the claim holds for $t = T$ because of assertion (i). Suppose now that the claim holds for $t+1$ and fix an arbitrary $c \geq 0$. Then, by the induction assumption, $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is equal to

$$\sup_{\frac{(d_t - w_t)^+}{p_t} \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(c(x_t - q_t), w_t + p_t q_t - d_t, \tilde{\Delta}_{t+1} p_t / c; \boldsymbol{d})]$$

if $d_t - w_t \leq p_t x_t$ and to 0 otherwise. Letting $x_t' = c x_t$, $q_t' = c q_t$ and $p_t' = p_t / c$, we have

$$\begin{aligned} &V_t(x_t, w_t, p_t; \boldsymbol{d}) \\ &= \begin{cases} \sup\limits_{\frac{(d_t - w_t)^+}{p_t'} \leq q_t' \leq x_t'} \inf\limits_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t' - q_t', w_t + p_t' q_t' - d_t, \tilde{\Delta}_{t+1} p_t'; \boldsymbol{d})] \text{ if } d_t - w_t \leq p_t' x_t', \\ 0 \text{ otherwise.} \end{cases} \end{aligned}$$

This implies that $V_t(x_t, w_t, p_t; \boldsymbol{d}) = V_t(x_t', w_t, p_t'; \boldsymbol{d}) = V_t(c x_t, w_t, p_t / c; \boldsymbol{d})$.

Proof of assertion (iv). We prove this assertion by induction. By assertion (i), $V_T(x_T, w_T, p_T; \boldsymbol{d})$ is jointly nondecreasing in (w_T, p_T). Suppose now that the claim holds for $t+1$ and recall the definition of V_t given in (1). The induction assumption implies that $V_{t+1}(x_t - q_t, w_t + p_t q_t - d_t, \Delta_{t+1} p_t; \boldsymbol{d})$ is jointly nondecreasing in (w_t, p_t) for all $\Delta_{t+1} \in [\underline{\Delta}_{t+1}, \overline{\Delta}_{t+1}]] \subseteq \mathbb{R}_{\geq 0}$. The expectation under $\mathbb{P}_{t+1}$ is a linear operation and preserves monotonicity. Moreover, the inf over the Wasserstein ball also preserves monotonicity as for all $w_t \geq w_t'$ and $p_t \geq p_t'$

$$\begin{aligned} f_{\mathbb{P}_{t+1}}(w_t, p_t) \geq f_{\mathbb{P}_{t+1}}(w_t', p_t') \ \forall \mathbb{P}_{t+1} \in \mathbb{B}_{t+1} \quad &\Longrightarrow f_{\mathbb{P}_{t+1}}(w_t, p_t) \geq \inf_{\mathbb{P}_{t+1}' \in \mathbb{B}_{t+1}} f_{\mathbb{P}_{t+1}'}(w_t', p_t') \ \forall \mathbb{P}_{t+1} \in \mathbb{B}_{t+1} \\ &\Longrightarrow \inf_{\mathbb{P}_{t+1}' \in \mathbb{B}_{t+1}} f_{\mathbb{P}_{t+1}'}(w_t, p_t) \geq \inf_{\mathbb{P}_{t+1}' \in \mathbb{B}_{t+1}} f_{\mathbb{P}_{t+1}'}(w_t', p_t'). \end{aligned}$$

A similar argument holds for the sup over feasible q_t. Thus,

$$\sup_{\frac{(d_t - w_t)^+}{p_t} \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t - q_t, w_t + p_t q_t - d_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]$$

is jointly nondecreasing in (w_t, p_t). It is easy to verify by an induction hypothesis that this function is always nonnegative under the condition $d_t - w_t \leq p_t x_t$. The observation that $p_t x_t + w_t \geq p_t' x_t + w_t'$ if $w_t \geq w_t'$ and $p_t \geq p_t'$ completes the proof. $\square$

Proof of Theorem 1. The proof relies on showing by induction that, for any $t \in [T]$, $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is convex in p_t, and it is jointly convex in (x_t, w_t) when w_t is in one of the following intervals.

$$\mathcal{W}_t^t = [0, d_t], \ \mathcal{W}_t^\ell = \left[\sum_{j=t}^{\ell-1} d_j, \sum_{j=t}^{\ell} d_j\right] \ \forall \ell = t+1, \ldots, T, \ \mathcal{W}_t^{T+1} = \left[\sum_{j=t}^{T} d_j, +\infty\right)$$

Note that $\mathcal{W}_t^\ell$, $\ell = t, \ldots, T+1$, can overlap only at the boundaries and that $\bigcup_{\ell=t}^{T+1} \mathcal{W}_t^\ell = [0, +\infty)$.

The claim holds for T because of assertion (i) of Lemma 1. Suppose now that the claim holds for $t+1$. Suppose also that $d_t - w_t \leq p_t x_t$. We then have

$$V_t(x_t, w_t, p_t; \boldsymbol{d}) = \sup_{\frac{(d_t - w_t)^+}{p_t} \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t - q_t, w_t + p_t q_t - d_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})].$$

V_{t+1} is convex and nondecreasing in Δ_{t+1} by the induction assumption and assertion (iv) of Lemma 1, respectively. Using Proposition 1, we can therefore write

$$V_t(x_t, w_t, p_t; \boldsymbol{d}) = \sup_{\frac{(d_t - w_t)^+}{p_t} \le q_t \le x_t} \mathbb{E}_{\mathbb{P}^\star_{t+1}}[V_{t+1}(x_t - q_t, w_t + p_t q_t - d_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})], \tag{13}$$

where $\mathbb{P}^\star_{t+1} = \frac{1}{N}\sum_{i:\Delta^i_{t+1} < \Delta^\star_{t+1}} \delta_{\Delta^i_{t+1}} + \frac{1}{N}\sum_{i:\Delta^i_{t+1} \ge \Delta^\star_{t+1}} \delta_{\Delta^\star_{t+1}}$ and

$$\Delta^\star_{t+1} = \max\{\underline{\Delta}_{t+1}, \min\{\Delta : \tfrac{1}{N}\textstyle\sum_{i=1}^N (\hat{\Delta}^i_{t+1} - \Delta)^+ \le \epsilon\}\}.$$

As the expectation preserves convexity and by the induction assumption, the objective function of the right-hand side in (13) is convex in q_t when $w_t + p_t q_t - d_t \in \mathcal{W}^\ell_{t+1}$, $\ell = t+1, \dots, T+1$, *i.e.*, when

$$q_t \in \mathcal{Q}^\ell_t = \left[\frac{(d_t - w_t)^+}{p_t}, x_t\right] \cap \left[\frac{(\sum_{j=t}^{\ell-1} d_j - w_t)}{p_t}, \frac{(\sum_{j=t}^{\ell} d_j - w_t)}{p_t}\right] \text{ for } \ell = t+1, \dots, T$$

and $q_t \in \mathcal{Q}^{T+1}_t = [(d_t - w_t)^+/p_t, x_t] \cap [(\sum_{j=t}^T d_j - w_t)/p_t, +\infty)$ for $\ell = T+1$. Note that $\mathcal{Q}^\ell_t$, $\ell = t+1, \dots, T+1$, can overlap only at the boundaries and that $\cup_{\ell=1}^{T+1} \mathcal{Q}^\ell_t = [(d_t - w_t)^+/p_t, x_t]$. Note also that $\mathcal{Q}^\ell_t$ can be empty for some ℓ in which case no feasible q_t can result in $w_t + p_t q_t - d_t \in \mathcal{W}^\ell_{t+1}$. Thus, the objective value is piecewise convex in q_t. This implies that there exists an optimal solution to problem (13) that is at the boundary of the set $\mathcal{Q}^\ell_t$ for some $\ell = t+1, \dots, T+1$. We thus have

$$q^\star_t \in \begin{cases} \{0, x_t\} \cup \left\{(\sum_{j=t}^\ell d_j - w_t)/p_t : \ell = \tau, \dots, \bar{t}\right\} & \text{if } w_t \in \mathcal{W}^\tau_t \text{ with } \tau \ge t+1 \\ \{x_t\} \cup \left\{(\sum_{j=t}^\ell d_j - w_t)/p_t : \ell = \tau, \dots, \bar{t}\right\} & \text{if } w_t \in \mathcal{W}^\tau_t \text{ with } \tau = t, \end{cases} \tag{14}$$

where $\bar{t} = \max\{\ell \in \{t, \dots, T\} : (\sum_{j=t}^\ell d_j - w_t)/p_t < x_t\}$ because of the following. If $w_t \in \mathcal{W}^\tau_t$ with $\tau \ge t+1$, then $d_\ell - w_t \le 0$ for all $\ell = t, \dots, \tau - 1$. This implies that if $w_t \in \mathcal{W}^\tau_t$ with $\tau \ge t+1$, then the set $\mathcal{Q}^\ell_t$ is either empty or contains only 0 for all $\ell < \tau$. Moreover, 0 constitutes a boundary point of the set $\mathcal{Q}^\tau_t$ and therefore is always a candidate solution. Note moreover that if $w_t \in \mathcal{W}^t_t$, then $(\sum_{j=t}^\ell d_j - w_t)/p_t \ge 0$ for all $\ell = t, \dots, \bar{t}$. It now suffices to show that (14) holds true by finalizing the induction step as (14) can be re-expressed in the form of (3). We continue by investigating the cases $w_t \in \mathcal{W}^\tau_t$ with $\tau \ge t+1$ (Case 1) and with $\tau = t$ (Case 2), separately.

Case 1 ($\tau \ge t+1$): Firstly, note that our initial assumption, *i.e.*, $d_t - w_t \le p_t x_t$, always holds in this case as $\tau \ge t+1$ implies $w_t \ge d_t$. Using (14), we can rewrite the right-hand side of (13) as

$$\begin{aligned} V_t(x_t, w_t, p_t; \boldsymbol{d}) = \max\Big\{&\mathbb{E}_{\mathbb{P}^\star_{t+1}}[V_{t+1}(x_t, w_t - d_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})], (p_t x_t + w_t - \textstyle\sum_{j=t}^T d_j)^+, \\ &\max_{\ell = \tau, \dots, \bar{t}} \mathbb{E}_{\mathbb{P}^\star_{t+1}}[V_{t+1}(x_t - (\textstyle\sum_{j=t}^\ell d_j - w_t)/p_t, \textstyle\sum_{j=t+1}^\ell d_j, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]\Big\}. \end{aligned} \tag{15}$$

The first term in the right-hand side of (15) is convex in p_t by the induction assumption, *i.e.*, $V_{t+1}(x_{t+1}, w_{t+1}, p_{t+1}; \boldsymbol{d})$ is convex in p_{t+1}, and by that the expectation preserves convexity. To see that it is also jointly convex in (x_t, w_t) when $w_t \in \mathcal{W}^\tau_t$, note that if $\tau \le T$, then

$$w_t \in \mathcal{W}^\tau_t = \left[\sum_{j=t}^{\tau-1} d_j, \sum_{j=t}^{\tau} d_j\right] \implies w_t - d_t = \left[\sum_{j=t+1}^{\tau-1} d_j, \sum_{j=t+1}^{\tau} d_j\right] = \mathcal{W}^\tau_{t+1}.$$

The induction assumption, *i.e.*, $V_{t+1}(x_{t+1}, w_{t+1}, p_{t+1}; \boldsymbol{d})$ is jointly convex in (x_{t+1}, w_{t+1}) when $w_{t+1} \in \mathcal{W}^\tau_{t+1}$, now implies that $V_{t+1}(x_t, w_t - d_t, \Delta_{t+1} p_t; \boldsymbol{d})]$ is also jointly convex in (x_t, w_t) when $w_t \in \mathcal{W}^\tau_t$. We can show that a similar argument holds for $\tau = T+1$. The second term in (15) is clearly convex in p_t and jointly convex in (x_t, w_t). We can replace the last term with the following using assertion (*iii*) of Lemma 1 (by selecting $c = p_t \geq 0$).

$$\max_{\ell=\tau,\ldots,\bar{t}} \mathbb{E}_{\mathbb{P}^\star_{t+1}}[V_{t+1}(p_t x_t - (\textstyle\sum_{j=t}^{\ell} d_j - w_t), \sum_{j=t+1}^{\ell} d_j, \tilde{\Delta}_{t+1}; \boldsymbol{d})]$$

This equivalent representation depends on p_t and (x_t, w_t) only through x_{t+1}. By the induction assumption and as expectation and maximization preserve convexity, it thus is convex in p_t and jointly convex in (x_t, w_t).

Recalling again that the maximum operator preserves convexity, we thus conclude that $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is convex in p_t and jointly convex in (x_t, w_t) when $w_t \in \mathcal{W}_t^\tau$ with $\tau \geq t+1$.

Case 2 ($\tau = t$): Unlike the first case, note that the initial assumption $d_t - w_t \leq p_t x_t$ may not hold in this case as $w_t \in \mathcal{W}_t^t = [0, d_t]$. We first discuss that the claim holds under the assumption that $d_t - w_t \leq p_t x_t$. Using (14), we can write (13) as

$$\begin{aligned} &V_t(x_t, w_t, p_t; \boldsymbol{d}) \\ &= \max\Big\{(p_t x_t + w_t - \textstyle\sum_{j=t}^{T} d_j)^+, \max_{\ell=\tau,\ldots,\bar{t}} \mathbb{E}_{\mathbb{P}^\star_{t+1}}[V_{t+1}(x_t - (\sum_{j=t}^{\ell} d_j - w_t)/p_t, \sum_{j=t+1}^{\ell} d_j, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]\Big\}. \end{aligned}$$

Following very similar steps to the proof of the first case, we can show that $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is convex in p_t and jointly convex in (x_t, w_t) when $w_t \in \mathcal{W}_t^t$.

Cases 1 and 2 prove the claim for $V_t(x_t, w_t, p_t; \boldsymbol{d})$ given in (13) and whose definition relies on the assumption that $d_t - w_t \leq p_t x_t$. Unlike Case 1, this assumption does not always hold for Case 2. If $d_t - w_t > p_t x_t$, we have $V_t(x_t, w_t, p_t; \boldsymbol{d}) = 0$ by definition (1). Note that when $d_t - w_t = p_t x_t$, $V_t(x_t, w_t, p_t; \boldsymbol{d}) = 0$ by assertion (*ii*) of Lemma 1 because $q_t = \frac{d_t - w_t}{p_t} = x_t$ is the only feasible decision and, consequently both x_{t+1} and w_{t+1} diminish to 0. Moreover, $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is always nonnegative by definition. These observations complete the proof for that $V_t(x_t, w_t, p_t; \boldsymbol{d})$ defined in (1) is convex in p_t and jointly convex in (x_t, w_t) when $w_t \in \mathcal{W}_t^t$. The proof thus completes. □

Remark 2. The proof of Theorem 1 uses Proposition 1 and thus relies on the choice of Wasserstein ball as the ambiguity set. Specifically, Proposition 1 allows us to replace inf over the Wasserstein ball by the expectation with respect to the worst-case distribution because the value function $V_t(x_t, w_t, p_t; \boldsymbol{d})$ is convex and nondecreasing in p_t. For general ambiguity sets, this would not be possible. We thus were not able to generalize Theorem 1 to any ambiguity set as the inf operator does not preserve convexity, which results in the failure of the induction hypotheses. On the other hand, Theorem 1 remains to hold for ambiguity sets consisting of distributions having given mean and dispersion measures (see, *e.g.*, Postek et al. (2018)) and for ambiguity sets consisting of all distributions whose total variation distance to the nominal distribution is limited above by a level of ambiguity (see, *e.g.*, Jiang and Guan (2018)). This is because the worst-case distribution can be identified in closed form for the expectation of convex functions.

The proof of Proposition 2 relies on the following technical lemma.

Lemma 3. *If $\boldsymbol{d} = d_T \boldsymbol{e}_T$, then the following hold.*

(i) $aV_t(x_t, w_t, p_t; \boldsymbol{d}) = V_t(ax_t, aw_t, p_t; a\boldsymbol{d})$ *for all* $a \geq 0$

(ii) $V_t(x_t, w_t, p_t; \boldsymbol{d}) + (w' - d')^+ \geq V_t(x_t, w_t + w', p_t; \boldsymbol{d} + d'\boldsymbol{e}_T)$ *for all* $w', d' \geq 0$

Proof. Proof of assertion (i). We prove (i) by induction. For T and for any $a \geq 0$, we have

$$aV_T(x_T, w_T, p_T; \boldsymbol{d}) = a(w_T + p_T x_T - d_T)^+ = (aw_T + ap_T x_T - ad_T)^+ = V_T(ax_T, aw_T, p_T; a\boldsymbol{d}),$$

where the first and last equalities follow from Lemma 1 (i). Suppose now that the claim holds for $t+1$. We then have

$$\begin{aligned}
aV_t(x_t, w_t, p_t; \boldsymbol{d}) &= \sup_{0 \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[aV_{t+1}(x_t - q_t, w_t + p_t q_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\
&= \sup_{0 \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(ax_t - aq_t, aw_t + ap_t q_t, \tilde{\Delta}_{t+1} p_t; a\boldsymbol{d})] \\
&= \sup_{0 \leq q_t' \leq ax_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(ax_t - q_t', aw_t + p_t q_t', \tilde{\Delta}_{t+1} p_t; a\boldsymbol{d})] = V_t(ax_t, aw_t, p_t; a\boldsymbol{d}),
\end{aligned}$$

where the second equality follows from the induction assumption, and the third equality follows from the variable transformation $q_t' \leftarrow aq_t$.

Proof of assertion (ii). We prove (ii) by induction. For T and for any $w', d' \geq 0$, we have

$$\begin{aligned}
&V_T(x_T, w_T, p_T; \boldsymbol{d}) + (w' - d')^+ = (w_T + p_T x_T - d_T)^+ + (w' - d')^+ \\
&\geq (w_T + w' + p_T x_T - d_T - d')^+ = V_T(x_T, w_T + w', p_T; \boldsymbol{d} + d'\boldsymbol{e}_T),
\end{aligned}$$

where the equalities follow from Lemma 1 (i). Suppose now that the claim holds for $t+1$. We then have

$$\begin{aligned}
&V_t(x_t, w_t, p_t; \boldsymbol{d}) + (w' - d')^+ \\
&= \sup_{0 \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t - q_t, w_t + p_t q_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d}) + (w' - d')^+] \\
&\geq \sup_{0 \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t - q_t, w_t + w' + p_t q_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d} + d'\boldsymbol{e}_T)] \\
&= V_t(x_t, w_t + w', p_t; \boldsymbol{d} + d'\boldsymbol{e}_T),
\end{aligned}$$

where the inequality follows from the induction assumption. □

Proof of Proposition 2. By Lemma 1 (i), we know that $V_T(x_T, w_T, p_T; \boldsymbol{d}) = (w_T + p_T x_T - d_T)^+$. As $d_t = 0$ for all $t < T$, we can write the value function at period $t < T$ as

$$V_t(x_t, w_t, p_t; \boldsymbol{d}) = \sup_{0 \leq q_t \leq x_t} \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t - q_t, w_t + p_t q_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})].$$

If $q_t = x_t$, the objective function amounts to $(w_t + p_t x_t - d_T)^+$ by Lemma 1 (ii). On the other hand, if $q_t = 0$, the objective function amounts to $\inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]$. We next prove the inequality

$$\begin{aligned}
&\max\left\{(w_t + p_t x_t - d_T)^+, \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]\right\} \\
&\geq \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t - q_t, w_t + p_t q_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \quad \forall q_t \in [0, x_t]
\end{aligned}$$

that immediately implies that $q_t^\star$ is equal to either x_t or 0. We have

$$\begin{aligned}
&\max\left\{(w_t + p_t x_t - d_T)^+, \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]\right\} \\
&\geq \lambda(w_t + p_t x_t - d_T)^+ + (1 - \lambda) \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \quad \forall \lambda \in [0, 1].
\end{aligned} \tag{16}$$

Moreover, for any $\lambda \in [0,1]$, we have

$$
\begin{aligned}
&\lambda(w_t + p_t x_t - d_T)^+ + (1-\lambda) \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\
&= \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[\lambda(w_t + p_t x_t - d_T)^+ + (1-\lambda) V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})] \\
&= \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[(\lambda w_t + \lambda p_t x_t - \lambda d_T)^+ + V_{t+1}((1-\lambda) x_t, (1-\lambda) w_t, \tilde{\Delta}_{t+1} p_t; (1-\lambda)\boldsymbol{d})] \\
&\geq \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}((1-\lambda) x_t, w_t + \lambda p_t x_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]
\end{aligned}
$$

where the second equality follows from Lemma 3 (*i*), and the inequality follows from Lemma 3 (*ii*). Letting $q_t = \lambda x_t$ proves the desired inequality (16).

Note that the second term on the left-hand side of (16) is always nonnegative by definition. Thus, we can drop the operator $(\cdot)^+$ in the first term. We thus have $q_t^\star = x_t$ if $w_t + p_t x_t - d_T \geq \inf_{\mathbb{P}_{t+1} \in \mathbb{B}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}[V_{t+1}(x_t, w_t, \tilde{\Delta}_{t+1} p_t; \boldsymbol{d})]$ as under this condition the first term, *i.e.*, the objective value of $q_t = x_t$, exceeds the second term, *i.e.*, the objective value of $q_t = 0$. $\square$