STABILITY OF A CLASS OF RISK-AVERSE MULTISTAGE STOCHASTIC PROGRAMS AND THEIR DISTRIBUTIONALLY ROBUST COUNTERPARTS

Jie Jiang, Zhiping Chen* and He Hu

School of Mathematics and Statistics, Xi’an Jiaotong University
Shaanxi, 710049, China
Center for Optimization Technique and Quantitative Finance
Xi’an International Academy for Mathematics and Mathematical Technology
Shaanxi, 710049, China

(Communicated by Ka Chun Cheung)

Abstract. In this paper, we consider the quantitative stability of a class of risk-averse multistage stochastic programs, whose objective functions are defined by multi-period $\mu$th order lower partial moments (LPM) with given targets, and their distributionally robust counterparts. We first derive the upper bounds of feasible solutions as preliminaries. Then, by employing calm modifications, the quantitative stability results are obtained under a special measurable perturbation of stochastic process, which extend the present results under risk-neutral cases to risk-averse ones. Moreover, we recast the risk-averse model by probability measures of stochastic process, and obtain new quantitative stability estimations on the basis of proper probability metrics under the general perturbation of stochastic process. Finally, motivated by the availability of only partial information about probability measures, we further consider the distributionally robust counterpart of our recasting model, and establish the discrepancy of optimal values with respect to the perturbation of ambiguity sets.

1. Introduction. Being able to describe complex decision-making problems under uncertainty, stochastic programs have become a very important branch in operations research, and have been applied extensively to many real-world problems, such as portfolio selection, power management and asset and liability management. Compared with two-stage stochastic programs, multistage stochastic programs are more attractive and important, especially, for the long-term decision-making problem.

Last decades have witnessed an intensive development about multistage stochastic programs, see excellent monographs [22, 29] and the references therein. When one solves a multistage stochastic programming problem numerically, very often, the original continuous data process is approximately replaced by a series of discrete...
probability distributions (or scenario tree) to avoid intractable high-dimensional integrals, see for example [11, 22]. A concerned question is: How the discrete problem approximates the original one? The quantitative stability analysis tries to settle this question. Therefore, the perturbation or stability analysis of multistage stochastic programs with respect to the change of the underlying stochastic process is critical for the development of reliable discretization techniques and numerical solution methods.

For two-stage stochastic programs, there have been many qualitative and quantitative stability results. We refer to [27] for a comprehensive review. The stability analysis of multistage stochastic programs is more intricate undoubtedly. Taking into account the perturbation of the filtration, Heitsch, Römisch and Strugarek discussed the quantitative stability of multistage stochastic linear programs in [12]. They established the Lipschitz continuity of the optimal value function by introducing the filtration distance for the first time. Under some recursive assumptions, Küchler proved in [16], by avoiding the filtration distance, that the optimal value function of multistage stochastic linear programming is Lipschitz continuous with respect to an $L_p$-distance. Statistical bounds for the optimal value function under the sample average approximation have been provided in [31]. The asymptotic stability of specific approximations for a class of convex multistage stochastic programming problems in terms of epi-convergence was established in [21]. Based on the barycentric approximation scheme, the continuity of the objective function of a convex multistage stochastic programming problem was proved in [17] under the compactness assumption for the stochastic variable space. More recently, Jiang and Chen discussed in [14] the quantitative stability of multistage stochastic linear programs by introducing two kinds of calm modifications, which simplifies and strengthens the results in [16].

All the above works concentrated on the risk-neutral case. In order to better reflect the decision-maker's risk preference, risk-averse multistage stochastic programs have attracted more and more attentions in recent years. In these models, the (linear) expected objective functional is replaced by a (nonlinear) risk functional, which brings much more difficulty for both theoretical study and numerical solution. The decomposition algorithm and dynamic sampling approach for Markov decision processes and multistage stochastic programs with risk aversion were proposed in [28] and [24], respectively. Meanwhile, Eichhorn and Römisch introduced the polyhedral risk measure in [6] which was defined as the optimal value of a two-stage linear stochastic program. Under this risk measure, the same authors, adopting the filtration distance in [12], gave the quantitative stability assertion of risk-averse multistage stochastic programs in [7]. Further, the extended multi-period polyhedral risk measure model was built in [9]. Philpott, Matos and Finardi derived in [25] the lower and upper bounds on the cost of an optimal policy for a class of multistage stochastic programs with coherent risk measures, by computing an inner approximation and an outer approximation to future cost functions, respectively. Dupačová and Kozmík studied the structural properties of risk-averse multistage stochastic programs with coherent risk measures in [5]. From a modeling and algorithmic perspective, Homem-de-Mello and Pagnoncelli discussed in [13] how to incorporate risk measures into multistage stochastic programs. More recently, Liu, Picher and Xu investigated in [19] the quantitative stability of multistage distributionally robust optimization problem, which can be viewed as a risk-averse model with coherent risk measure due to the robust reformulation of coherent risk measure.
We can see from the above review that there are relatively few contributions on the quantitative stability analysis of risk-averse multistage stochastic programs. In view of this, we focus on the quantitative stability analysis of a class of risk-averse multistage stochastic programs in this paper under some mild assumptions. More specifically, by adopting the framework of the separable expected conditional (SEC) multi-period risk measure in [23], we introduce the risk-averse multistage stochastic programming problem where the risk is measured by the multi-period \( p \)th order lower partial moment (LPM) risk measure. The LPM risk measure is a measurement of the dispersion of returns below a target return (or losses above a target loss), which can be used as a proxy for downside risk exposure. In this sense, the LPM risk measure is more plausible than variance as a measure of risk, in that it concerns only with adverse deviations. For more motivations about this kind of risk measures, we refer to the pioneering work [8] and excellent references [1, 3].

In traditional stochastic optimization problems, it is always assumed that the decision maker knows the exact probability measures of random variables. This hardly holds for many complex decision-making problems under uncertainty. In view of this, we see nowadays considerable works focusing on the situation that only partial information, such as historical data and moment information, is available. We can, according to these partial information, construct a set of probability measures which contain the true probability measure. That is the so-called ambiguity set. Then, to avoid the potential modeling risk as much as possible, the worst case over the ambiguity set is considered. This leads to the well-known distributionally robust optimization problem. For interested readers, we refer to, for example, [2, 4, 19, 20] and the references therein for further information. However, it seems to be difficult for us to extend the quantitative stability results of multistage stochastic programs to the distributionally robust ones by adopting the frameworks in existing works, such as [12, 14, 16].

All these considerations above lead to the following contributions of this paper. Firstly, we adopt the concept of calm modifications in [14] and extend the quantitative stability results about the optimal value under risk-neutral in [14] to multi-period \( p \)th order LPM risk-averse ones when the perturbed stochastic process is defined by a measurable stochastic process. Secondly, we recast multistage stochastic programming problems in terms of probability measures of stochastic processes. Thus, the perturbation of a stochastic process is equivalent to the perturbation of probability measures. After that, new quantitative stability results about the optimal value are derived with respect to some proper probability metrics under the general perturbation. Finally, under our new framework, we extend the above risk-averse multistage stochastic programming problem to its distributionally robust counterpart. The quantitative stability assertion is then given with respect to the discrepancy of ambiguity sets.

Throughout the paper, we adopt the following notations. Random variables are denoted by boldface letters, for example \( \mathbf{x} \) and \( \mathbf{\xi} \), and their realizations are denoted by \( x \) and \( \xi \), respectively. For a \( T \)-stage stochastic programming problem, we use \( \mathbf{\xi}_t \) to stand for the stochastic parameters in stage \( t \). \( \mathbf{\xi}^T \) denotes the block vector \( (\xi_1, \xi_2, \cdots, \xi_T) \). Specially, when \( t = T \), we usually omit the subscript \( T \), i.e., \( \mathbf{\xi} := \mathbf{\xi}^T \). \( \langle \cdot, \cdot \rangle \) denotes the inner product in finite dimensional Hilbert space. \( \|\cdot\| \) denotes the Euclidean norm in finite dimensional Euclidean space. For the block vector \( \mathbf{\xi}^T \), we define \( \|\mathbf{\xi}^T\| = \max_{1 \leq i \leq T} \|\xi_i\| \). For \( \bar{a}, \bar{b} \in \mathbb{R}^n, \mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n \), we use \( d(\bar{a}, \bar{b}) := \|\bar{a} - \bar{b}\|, d(\bar{a}, \mathcal{B}) := \inf_{b \in \mathcal{B}} \|\bar{a} - b\|, d(\mathcal{A}, \mathcal{B}) := \sup_{\bar{a} \in \mathcal{A}} \inf_{b \in \mathcal{B}} \|a - b\| \).
and \(d_H(A, B) := \max \{d(A, B), d(B, A)\}\) to denote the distance between \(\bar{a}\) and \(\bar{b}\), the distance from \(\bar{a}\) to \(B\), the deviation distance from \(A\) to \(B\) and the Pompeiu-Hausdorff distance between \(A\) and \(B\), respectively.

This paper is laid out as follows. In the next section, we introduce our models and state basic results that are useful in the sequel. In Section 3, we focus on the risk-averse multistage stochastic programming with multi-period \(p\)th order LPM risk measures. Different quantitative stability assertions of the optimal value function are established. In Section 4, we extend our quantitative stability results to the distributionally robust risk-averse multistage stochastic programming problem. Finally, we present some concluding remarks in Section 5.

2. Models and preliminaries.

2.1. Models. In order to describe a finite decision sequence under uncertainty, we consider an \(\mathbb{R}^s\)-valued discrete stochastic process \(\{\xi_t\}_{t=1}^T\) with time horizon \(2 \leq T \in \mathbb{N}\), where the random vector \(\xi_t\) is defined on some probability space \((\Omega, \mathcal{F}, \mathbb{P})\), that is, \(\xi_t : \Omega \rightarrow \Xi_t\), where \(\Xi_t\) is the support set of \(\xi_t\), i.e., the smallest closed subset of \(\mathbb{R}^s\) such that \(\mathbb{P}\{\omega \in \Omega : \xi_t(\omega) \in \Xi_t\} = 1\). Based on \(\mathbb{P}\), we can define a probability measure on the support set \(\Xi_t\) by \(\mathbb{P}_t := \mathbb{P} \circ (\xi_t)^{-1}\). We use \(\xi^t\) and \(\mathcal{F}_t\) to denote the random block vector \((\xi_1, \xi_2, \cdots, \xi_t)\) and the \(\sigma\)-field induced by \(\xi^t\), i.e., \(\mathcal{F}_t := \sigma(\xi^t)\). Thus, \(\xi^t\) is a random vector on the probability space \((\Omega, \mathcal{F}_t, \mathbb{P})\). We denote the support set of \(\xi^t\) by \(\Xi^t \subseteq \mathbb{R}^{Ts}\). Analogously, we denote by \(\mathbb{P}^t = \mathbb{P} \circ (\xi^t)^{-1}\) the induced probability measure on \(\Xi^t\) and \(\mathcal{P}(\Xi^t)\) the collection of all probability measures supported on \(\Xi^t\).

Moreover, we have \(\{\emptyset, \Omega\} = \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_T = \mathcal{F}\), which forms the so-called filtration. The stochastic decision \(x_t \in \mathbb{R}^n\) at time \(t\) is assumed to depend only on the available information until time \(t\), that is \(x_t = x_t(\xi^t)\). This is the well-known nonanticipativity condition, which is an essential property of multistage stochastic programming. This condition is equivalent to the measurability of \(x_t\) with respect to the \(\sigma\)-field \(\mathcal{F}_t\) [32]. Note that \(\mathcal{F}_1 = \{\emptyset, \Omega\}\), which implies that \(\xi_1\) and \(x_1\) are deterministic.

Generally, the multistage stochastic linear programming problem can be formulated as (see [22, 32]):

\[
\inf_{x_1 \in X_1} \left< c_1, x_1 \right> + E\left[ \inf_{x_2 \in X_2(x_1, \xi_2)} \left< c_2(\xi_2), x_2 \right> + E\left[ \inf_{x_3 \in X_3(x_2, \xi_3)} \left< c_3(\xi_3), x_3 \right> \right. \right. \\
\left. \left. \quad + \cdots + E\left[ \inf_{x_T \in X_T(x_{T-1}, \xi_T)} \left< c_T(\xi_T), x_T \right> \right| \mathcal{F}_{T-1} \right] \cdots \right| \mathcal{F}_2 \right],
\]

(1)

where \(E[\cdot]\) is the expectation operator with respect to \(\mathbb{P}\), and we neglect \(\mathbb{P}\) hereinafter without any confusion. \(X_1 \subseteq \mathbb{R}^n\) is a closed and bounded set and for \(t = 2, \cdots, T\),

\[
X_t(x_{t-1}, \xi_t) := \{x_t \in \mathbb{R}^n : A_{t,0}x_t + A_{t,1}(\xi_t)x_{t-1} \leq h_t(\xi_t)\}.
\]

(2)

To simplify notations, we also use \(X_1(x_0, \xi_1)\) to stand for \(X_1\) in the following. For \(t = 2, \cdots, T\), \(A_{t,1} : \Xi_t \rightarrow \mathbb{R}^{mxn}\), \(A_{t,0} \in \mathbb{R}^{mxn}\) and \(h_t : \Xi_t \rightarrow \mathbb{R}^m\) are called technology matrices, recourse matrices and right-hand side vectors, respectively. By convention, we assume that \(c_t(\xi_t), A_{t,1}(\xi_t)\) and \(h_t(\xi_t)\) depend affinely on \(\xi_t\) for \(t = 2, \cdots, T\). Thus, there exists a positive constant \(B\) such that

\[
\|\Upsilon(\xi_t)\| \leq B \max\{1, \|\xi_t\|\},
\]

\[
\|\Upsilon(\xi_t) - \Upsilon(\xi_t')\| \leq B \|\xi_t - \xi_t'\|
\]

(3)

for \(\xi_t, \xi_t' \in \Xi_t\) and \(t = 2, \cdots, T\), where \(\Upsilon(\cdot)\) can be \(c_t(\xi_t), A_{t,1}(\cdot)\) or \(h_t(\cdot)\).
Define \( \phi(x, \xi) = \sum_{t=1}^{T} \langle c_t (\xi_t), x_t \rangle \) and \( x = (x_1, x_2, \cdots, x_T) \). We can equivalently reformulate problem (1) as (see [32]):

\[
\inf \mathbb{E}[\phi(x, \xi)] \\
\text{s.t. } x_t \in X_t(x_{t-1}, \xi_t), \ t = 1, 2, \cdots, T.
\]

Model (1) is risk-neutral in the sense that it just cares the expectation of the random loss at each stage. To better reflect the risk-averse preference of a rational decision maker, the risk-averse ones have been recently proposed and extensively discussed, see for instance [9, 23, 30]. To derive the risk-averse multistage stochastic programming model in this paper, we first introduce the multi-period risk measure.

For notational simplicity, we use \( L_t \) to denote \( L_r(\Omega, \mathcal{F}_t, \mathbb{P}) \) which consists of random vectors with finite rth order absolute moment, i.e., \( \mathbb{E}[|\xi^t|'^r] < +\infty \) for some positive \( r \). In this paper, we always assume that the stochastic processes \( \xi^t \) for \( t = 2, 3, \cdots, T \) are contained in \( L_t \) for some \( r \geq pT + 1 \). Then, the risk of the \((t+1)\)th stage, which can be observed at stage \( t \), is defined by \( \rho_{t+1} : L_{t+1} \rightarrow L_t \).

Denote \( L_{t,T} := L_t \times L_{t+1} \times \cdots \times L_T \). Naturally, the risk of the random loss process between stage \( t+1 \) and stage \( T \) can be measured by a conditional risk mapping \( \rho_{t+1,T} : L_{t+1,T} \rightarrow L_t \), see [30]. Therefore, we can measure the risk during the whole \( T \) stages by \( \rho_{2,T} \). A typical multi-period risk measure induced in [23] is the so-called SEC mapping. To state the SEC risk measure, we use \( \{Z_t\}_{t=1}^{T} \) to denote a sequence of random losses in \( T \) stages, and we assume that, for \( t = 1, 2, \cdots, T \), \( Z_t \in L_t \). Analogously, we use \( Z_t,T \) to denote \( (Z_t, Z_{t+1}, \cdots, Z_T) \). Then the SEC risk measure between stage \( t+1 \) and stage \( T \) can be described as:

\[
\rho_{t+1,T}(Z_{t+1,T}) = \sum_{i=t+1}^{T} \mathbb{E}[\rho_i(Z_i)|\mathcal{F}_i].
\]

Time consistency has been a desirable property for multi-period risk measure. There are several ways to define time consistency. We adopt the following definition of time consistency (see also [33]): A multi-period risk measure \( \{\rho_t,T\}_{t=1}^{T} \) is called time consistency if for any \( 1 \leq \tau < \theta \leq T \) and two sequences of random losses \( \{Z_t\}_{t=1}^{\tau} \) and \( \{Y_t\}_{t=1}^{\tau} \) with \( Z_{\tau,T} \leq Y_{\tau,T} \) and \( \rho_{\tau+1,T}(Z_{\tau+1,T}) \leq \rho_{\tau+1,T}(Y_{\tau+1,T}) \) implies \( \rho_{\tau,T}(Z_{\tau,T}) \leq \rho_{\tau,T}(Y_{\tau,T}) \). It knows from [15] that the SEC mapping is time consistent under the above definition.

Of particular interest of this paper, we consider the \( p \)th \((p \geq 1)\) order LPM risk measure at each stage. That is, we take \( \rho_t[Z_t] = \mathbb{E}[ (Z_t - \eta_t)^p_+ | \mathcal{F}_{t-1} ] \), where \( \eta_t \) is the given target or threshold for stage \( t \) and \( p \) is a risk-aversion parameter. Here we assume that \( p \geq 1 \) because the decision maker becomes risk-preference when \( p < 1 \), which is not reasonable for a rational decision maker. \( \eta_t \) can be viewed as the maximum loss that the decision maker can endure in stage \( t \). As what we have stated in introduction, the LPM risk measure is a commonly used risk measure which controls decision makers’ downside risk. For two sequences of random losses \( \{Z_t\}_{t=1}^{T} \) and \( \{Y_t\}_{t=1}^{T} \) satisfying \( Z_{t+1,T} \leq Y_{t+1,T} \) and \( \rho_{\theta+1,T}(Z_{\theta+1,T}) \leq \rho_{\theta+1,T}(Y_{\theta+1,T}) \), we have \( Z_t \leq Y_t \) for \( \tau \leq t \leq \theta \) and

\[
\sum_{i=\theta+1}^{T} \mathbb{E}[(Z_i - \eta_i)^p_+ | \mathcal{F}_\theta] \leq \sum_{i=\theta+1}^{T} \mathbb{E}[(Y_i - \eta_i)^p_+ | \mathcal{F}_\theta]. \tag{4}
\]
Moreover, we know from the nonincreasing function $E \left[ (\cdot - \eta_t)_{+}^P \mid \mathcal{F}_{t-1} \right]$, $t = 2, \ldots, T$ that
\[
\sum_{j=t}^{\theta} E \left[ (Z_j - \eta_j)_{+}^P \mid \mathcal{F}_{\tau-1} \right] \leq \sum_{j=t}^{\theta} E \left[ (Y_j - \eta_j)_{+}^P \mid \mathcal{F}_{\tau-1} \right].
\] (5)

Note that (4) implies
\[
\sum_{i=\theta+1}^{T} E \left[ (Z_i - \eta_i)_{+}^P \mid \mathcal{F}_{\tau-1} \right] \leq \sum_{i=\theta+1}^{T} E \left[ (Y_i - \eta_i)_{+}^P \mid \mathcal{F}_{\tau-1} \right]
\] due to the linearity of conditional expectation operators. Combining (5) with (6), we obtain
\[
\sum_{i=\tau}^{T} E \left[ (Z_i - \eta_i)_{+}^P \mid \mathcal{F}_{\tau-1} \right] \leq \sum_{i=\tau}^{T} E \left[ (Y_i - \eta_i)_{+}^P \mid \mathcal{F}_{\tau-1} \right],
\] which implies $\rho_{\tau,T}(Z_{\tau,T}) \leq \rho_{\tau,T}(Y_{\tau,T})$. Therefore, our specific SEC risk measure is time consistent in the sense of above definition.

Under the above setting and taking $Z_t = \langle c_t(\xi_t), x_t \rangle$, we can measure the risk between stage $t$ and stage $T$ by
\[
\rho_{t,T} \left[ \sum_{i=t}^{T} \langle c_i(\xi_i), x_i \rangle \right] = \sum_{i=t}^{T} E \left[ \langle (c_i(\xi_i), x_i) - \eta_i \rangle_{+}^P \mid \mathcal{F}_{t-1} \right]
\] for $t = 2, 3, \ldots, T - 1$. Specially, we measure the risk during the horizon by
\[
\rho \left[ \sum_{i=1}^{T} \langle c_i(\xi_i), x_i \rangle \right] := \langle c_1, x_1 \rangle + \rho_{2,T} \left[ \sum_{i=2}^{T} \langle c_i(\xi_i), x_i \rangle \right]
\]
\[
= \langle c_1, x_1 \rangle + \sum_{i=2}^{T} E \left[ \langle (c_i(\xi_i), x_i) - \eta_i \rangle_{+}^P \right].
\] Then the risk-averse model under the multi-period LPM risk measure can be written as
\[
\inf_{x} \rho[\phi(x, \xi)]
\] s.t. $x_t \in X_t(x_{t-1}, \xi_t)$, $t = 1, 2, \ldots, T$. (7)

By considering the induced probability measures on corresponding support sets, we can rewrite problem (7) as the following equivalent reformulation:
\[
\inf_{x} \left( \langle c_1, x_1 \rangle + \sum_{i=2}^{T} \mathbb{E}_{\psi^t} \left[ \langle (c_i(\xi_i), x_i) - \eta_i \rangle_{+}^P \right] \right)
\] s.t. $x_t \in X_t(x_{t-1}, \xi_t)$, $t = 1, 2, \ldots, T$, (8)

where $\mathbb{P}^t = \mathbb{P} \circ (\xi^t)^{-1}$ is the probability measure of $\xi^t$ on its support set $\Xi^t$ for $t = 2, 3, \ldots, T$. We use the notation $\mathbb{E}_{\psi^t} [\cdot]$ to stress that the expectation is taken with respect to $\mathbb{P}^t$. In this way, we take the dynamics or information process into account. This ensures the well-definedness of the multi-period risk measure in problem (8).

As an illustration about the equivalence between problem (7) and problem (8), we consider the following example.
Example 1. Consider $T = 3$ and the scenario tree in Figure 1.

![Scenario tree for Example 1](image)

**Figure 1.** Scenario tree for Example 1

Denote the sample set by $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, assigned with the probability $p_1 = P(\omega_1) = 1/4$, $p_2 = P(\omega_2) = 1/2$, $p_3 = P(\omega_3) = 1/8$ and $p_4 = P(\omega_4) = 1/8$.

Moreover, we define mappings

$$
\xi(\omega_1) = (\xi_1, \xi_2, \xi_3),
\xi(\omega_2) = (\xi_1, \xi_2, \xi_3),
\xi(\omega_3) = (\xi_1, \xi_2, \xi_3),
\xi(\omega_4) = (\xi_1, \xi_2, \xi_3).
$$

Then we have the probability space $(\Omega, \mathbb{F}, P)$, and also, we have the filtration for the above 3-stage problem: $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3$, where

$$
\mathcal{F}_1 = \{\emptyset, \Omega\},
\mathcal{F}_2 = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \Omega\},
\mathcal{F}_3 = 2^\Omega = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_3, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}, \Omega\}.
$$

Then, for the objective function of problem (7), we have

$$
\rho \left[ \sum_{t=1}^{T} \langle c_t(\xi_t), x_t \rangle \right] = \rho \left[ \sum_{t=1}^{3} \langle c_t(\xi_t), x_t \rangle \right] \\
= \langle c_1, x_1 \rangle + E \left[ (\langle c_2(\xi_2), x_2 \rangle - \eta_2)_+ \right] + E \left[ (\langle c_3(\xi_3), x_3 \rangle - \eta_3)_+ \right] \\
= \langle c_1, x_1 \rangle + \sum_{i=1}^{4} p_i \left( (\langle c_2(\xi_2(\omega_i)), x_2(\xi_2(\omega_i)) \rangle - \eta_2)_+ \right) \\
+ \sum_{i=1}^{4} p_i \left( (\langle c_3(\xi_3(\omega_i)), x_3(\xi_3(\omega_i)) \rangle - \eta_3)_+ \right) \\
= \langle c_1, x_1 \rangle + \frac{3}{4} \left( (\langle c_2(\xi_2(\omega_1)), x_2((\xi_1, \xi_2(\omega_1))) \rangle - \eta_2)_+ \right) + \frac{1}{4} \left( (\langle c_2(\xi_2(\omega_2)), x_2((\xi_1, \xi_2(\omega_2))) \rangle - \eta_2)_+ \right) \\
+ \frac{1}{4} \left( (\langle c_3(\xi_3(\omega)), x_3((\xi_1, \xi_3(\omega))) \rangle - \eta_3)_+ \right) + \frac{1}{2} \left( (\langle c_3(\xi_3(\omega)), x_3((\xi_1, \xi_3(\omega))) \rangle - \eta_3)_+ \right) \\
+ \frac{1}{8} \left( (\langle c_3(\xi_3(\omega)), x_3((\xi_1, \xi_3(\omega))) \rangle - \eta_3)_+ \right) + \frac{1}{8} \left( (\langle c_3(\xi_3(\omega)), x_3((\xi_1, \xi_3(\omega))) \rangle - \eta_3)_+ \right).
Instead, if we consider the induced probability distributions on support sets \( \Xi^2 = \{ (\xi_1, \xi_2), (\xi_1, \xi_2) \} \) and \( \Xi^3 = \{ (\xi_1, \xi_2, \xi_3), (\xi_1, \xi_2, \xi_3), (\xi_1, \xi_2, \xi_3) \} \) respectively, we have the induced probability

\[
P^2(\xi_1, \xi_2) = P(\xi_2)^{-1}(\xi_1, \xi_2) = P(\{\omega_1, \omega_2\}) = \frac{3}{4},
\]

\[
P^2(\xi_1, \xi_2) = P(\xi_2)^{-1}(\xi_1, \xi_2) = P(\{\omega_3, \omega_4\}) = \frac{1}{4}
\]

and similarly

\[
P^3(\xi_1, \xi_2, \xi_3) = 1, 4,
\]

\[
P^3(\xi_1, \xi_2, \xi_3) = 1, 2
\]

Taking these probabilities into the objective function of problem (8), we have

\[
\langle c_1, x_1 \rangle + \sum_{t=2}^{3} \mathbb{E}_{\bar{\mathcal{P}}^t}[\langle (c_t(\xi_t), x_t) - \eta_t \rangle_{+}]
\]

\[
= \langle c_1, x_1 \rangle + \frac{3}{4} (\langle c_2(\xi_2), x_2((\xi_1, \xi_2)) \rangle - \eta_2)_{+} + \frac{1}{4} (\langle c_2(\xi_2), x_2((\xi_1, \xi_2)) \rangle - \eta_2)_{+}
\]

\[
+ \frac{1}{4} (\langle c_3(\xi_3), x_3((\xi_1, \xi_2, \xi_3)) \rangle - \eta_3)_{+} + \frac{1}{4} (\langle c_3(\xi_3), x_3((\xi_1, \xi_2, \xi_3)) \rangle - \eta_3)_{+}
\]

\[
= c_t(\xi_t), x_t \rangle
\]

which verifies the consistency between problems (7) and (8).

Compared with problem (7), problem (8) has its computational advantage. It considers the problem with respect to the induced probability in the support space which, of course, is more concrete than that of the sample space. Moreover, usually, the probability space \( (\Omega, 2^\Omega, P) \) can not be directly observed in practical application. Instead, we may obtain some empirical data which is in the support set. In this case, it becomes more tractable for us to consider problem (8).

On the other hand, considering the incomplete information of \( \mathcal{P}^t \) for decision makers, we assume that one can construct an ambiguity set, denoted by \( \mathcal{P}^t \subseteq \mathcal{P}(\Xi^t) \), for the uncertain probability measure \( \mathcal{P}^t \) at each stage \( t \) for \( t = 2, 3, \cdots, T \). With this, we can deduce the following distributionally robust counterpart for model (8) (see also [18, 19]):

\[
\inf_{x} \langle c_1, x_1 \rangle + \sum_{t=2}^{T} \mathcal{P}^t \mathbb{E} \left[ (\langle c_t(\xi_t), x_t \rangle - \eta_t)_{+} \right]
\]

s.t. \( x_t \in X_t(x_{t-1}, \xi_t), \ t = 1, 2, \cdots, T \).

Further, we denote \( \mathcal{P} = (\mathcal{P}^2, \mathcal{P}^3, \cdots, \mathcal{P}^T) \). In our upcoming development, we use \( v(\xi) \) and \( v(\mathcal{P}) \) to denote the optimal value functions of problems (7) or (8) and
P to denote the total variation metric between $D$ the resulting pseudo metric is called the total variation metric. We use $G$ elements of a metric. When $G$ is called the pseudo metric with $D$ under the stochastic process $x_t$ analysis. In what follows, we present some pseudo metrics for further discussion.

As a result, the pseudo metric includes many widely used metrics in the area of stability $G$. That different pseudo metrics can be induced by choosing specific $P$. To describe the quantitative stability result, we need to introduce the concept of probability metrics. Probability metrics represent distance functions on the space of probability measures. Of particular interest of this paper, we need the following class of probability metrics, which are the so-called $ζ$-structure probability metrics, see for instance [27, 34].

Let $d$ be a set of real-valued measurable functions on $X$. For $t = 2, 3, \cdots, T$ and any two probability measures $P, Q^t \in P(X)$, $D_t(P^t, Q^t) = \sup_{g \in G} [E_P[g(ξ)] - E_Q^t[g(ξ)]]$ is called the pseudo metric with $ζ$-structure between $P^t$ and $Q^t$ induced by $G$.

We call $D_t(P^t, Q^t)$ the pseudo metric because it usually does not meet requirements of a metric. When $G$ is large enough such that $D_t(P^t, Q^t) = 0$ implies $P^t = Q^t$, the above pseudo metric becomes a metric. It is known from Definition 2.1 that different pseudo metrics can be induced by choosing specific $G$s. As a result, the pseudo metric includes many widely used metrics in the area of stability analysis. In what follows, we present some pseudo metrics for further discussion.

If we choose $G_{TV} = \{ g : X^t \to \mathbb{R} : g \text{ is measurable and } \sup_{ξ \in Ξ} |g(ξ)| \leq 1 \}$, the resulting pseudo metric is called the total variation metric. We use $D_{TV}(P^t, Q^t)$ to denote the total variation metric between $P^t$ and $Q^t$.

If $G$ is specified as $G_{FM}^q = \left\{ g : X^t \to \mathbb{R} : |g(ξ^t_1) - g(ξ^t_2)| \leq \max \{1, ||ξ^t_1||, ||ξ^t_2|| \} q^{-1} ||ξ^t_1 - ξ^t_2|| \right\}$, the corresponding pseudo metric is called the $q$th order Fortet-Mourier metric, denoted by $ζq(P^t, Q^t)$, which is often used in the stability analysis of two-stage stochastic programs, see [27].

The following proposition comes from Example 9.35 in [26].

**Proposition 1.** For $X_t(x_{t-1}, ξ_t)$, $t = 2, \cdots, T$, defined in (2), the following assertions hold:

$$d_H(X_t(x_{t-1}, ξ_t), X_t(ξ_{t-1}, ξ_t)) \leq M \max\{1, ||ξ_t|| \} ||x_{t-1} - x_{t-1}||,$$

$$d_H(X_t(x_{t-1}, ξ_t), X_t(x_{t-1}, ξ_t)) \leq M \max\{1, ||x_{t-1}|| \} ||ξ_t - ξ_t||$$

for some constant $M > 0$.

In order to derive an upper bound of feasible solutions, we need the following properties, see [10, Theorem 2.1].
Lemma 2.2. For $F(A,b) := \{ x \in \mathbb{R}^n : Ax \leq b \}$, where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, we have the following results.

(i) $F(A,b)$ is bounded if and only if $0 \in \text{int}(\text{conv}(A))$, where $\text{conv}(A)$ denotes the convex hull of $A$’s row vectors and $\text{int}(\text{conv}(A))$ denotes the interior points of $\text{conv}(A)$;

(ii) If $F(A,b)$ is bounded, then for every $x \in F(A,b)$, we have

$$
\|x\| \leq \frac{\max\{b_i, 1 \leq i \leq m\}}{d(0, \text{bd}(\text{conv}(A)))},
$$

where $\text{bd}(\text{conv}(A))$ denotes the boundary of $\text{conv}(A)$.

To proceed, we need the following technical assumption.

Assumption 2.2. $0 \in \text{int}(\text{conv}(A_{t,0}))$ for $t = 2, \cdots, T$.

Assumption 2.2 implies the boundedness of the feasible set $X_t(x_{t-1}, \xi_t), t = 2, \cdots, T$ for any fixed pair $(x_{t-1}, \xi_t)$. Then, based on Lemma 2.2, we have the following proposition.

Proposition 2. Under Assumption 2.2, there exists a positive constant $\hat{L}$ such that

$$
\|x_t\| \leq \hat{L} \max\{1, \|\xi_t\|\} \max\{1, \|x_{t-1}\|\}
$$

for any $x_t \in X_t(x_{t-1}, \xi_t), t = 1, 2, \cdots, T$ with $x_0 := 1$.

Proof. Due to Lemma 2.2, for $x_t \in X_t(x_{t-1}, \xi_t), t = 2, \cdots, T$, we have the estimate

$$
\|x_t\| \leq \frac{\max\{(h_t(\xi_t) - A_{t,1}(\xi_t)x_{t-1}), 1 \leq i \leq m\}}{d(0, \text{bd}(\text{conv}(A_{t,0})))} \leq \frac{\|h_t(\xi_t) - A_{t,1}(\xi_t)x_{t-1}\|}{d(0, \text{bd}(\text{conv}(A_{t,0})))}.
$$

Assumption 2.2 implies that $d(0, \text{bd}(\text{conv}(A_{t,0}))) > 0, t = 2, \cdots, T$. This together with the boundedness of $X_1$ ensures that

$$
\|x_t\| \leq \hat{L} \|h_t(\xi_t) - A_{t,1}(\xi_t)x_{t-1}\|
$$

holds for some constant $\hat{L} > 0$ and $t = 1, 2, \cdots, T$. Therefore, we have

$$
\|x_t\| \leq \hat{L}(\|h_t(\xi_t)\| + \|A_{t,1}(\xi_t)\| \|x_{t-1}\|) \leq 2\hat{L}B \max\{1, \|\xi_t\|\} \max\{1, \|x_{t-1}\|\}
$$

for $t = 1, 2, \cdots, T$. Setting $\hat{L} = 2\hat{L}B$ completes the proof.

It is noteworthy that an analogous setting can also be found in [16, Assumption 2.3], where the author directly assumed the boundedness property, i.e., Proposition 2. From Proposition 2, we can deduce the following corollary straightforwardly by induction on $t$.

Corollary 1. Under Assumption 2.2, there exists a constant $\hat{L} > 0$, such that

$$
\|x_t\| \leq \hat{L} \max\{1, \|\xi_t\|\}^{t-1}
$$

for any $x_t \in X_t(x_{t-1}, \xi_t), t = 1, 2, \cdots, T$. 
To unify these upper bounds and simplify the following statement, we define

\[ L = \max\{L, L', 1\} \]

Then

\[ \|x_t\| \leq L \max\{1, \|\xi_t\|\} \max\{1, \|x_{t-1}\|\} \]

and

\[ \|x_t\| \leq L \max\{1, \|\xi_t\|\}^{t-1} \]

for \( t = 1, 2, \ldots, T \).

To describe the perturbation of stochastic processes and quantitatively estimate the distance between corresponding solutions under different stochastic processes, we introduce the following two definitions.

**Definition 2.3** (approximation of stochastic process, [16]). A stochastic process \( \xi \) on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) is called an approximation of \( \xi \), if there exist Borel-measurable mappings

\[ f_t : \Xi^t \to \Xi_t, \quad t = 1, 2, \ldots, T, \]

such that the following conditions are satisfied:

(a) \( \tilde{\xi}_t = f_t(\xi^t) \) for \( t = 1, 2, \ldots, T \);

(b) \( f^T(\Xi^T) \subseteq \Xi^T \);

(c) \( f_1(\xi_1) = \xi_1 \) for every \( \xi_1 \in \Xi_1 \);

(d) \( f^T(\xi^T) \in \mathcal{L}_r(\Omega, \mathcal{F}, \mathbb{P}) \) for \( r \geq pT \).

Here, \( f^t(\xi^t) := (f_1(\xi_1), f_2(\xi_2), \ldots, f_t(\xi_t)) \) for \( t = 1, 2, \ldots, T \).

According to the context, we sometimes also write \( f^T \) as \( f \) for the sake of simplicity.

**Definition 2.4** (calm modifications, [14]). For an optimal solution \( x^* \) under the stochastic process \( \xi \), we call

(i) \( \tilde{x}^*(\xi) = (\tilde{x}^*_1, \tilde{x}^*_2(\xi^2), \ldots, \tilde{x}^*_T(\xi^T)) \) the class I calm modification under the stochastic process \( \xi \), if it is defined by

\[ \tilde{x}^*_t = x^*_1, \quad \tilde{x}^*_t(\xi^t) \in \arg\min_{z \in X_t(\tilde{x}^*_{t-1}(\xi^{t-1}), \xi_t)} \|z - x^*_t(f^t(\xi^t))\| \]

for \( t = 2, \ldots, T \);

(ii) \( \tilde{x}^*(\xi) = (\tilde{x}^*_1, \tilde{x}^*_2(\xi^2), \ldots, \tilde{x}^*_T(\xi^T)) \) the class II calm modification under the stochastic process \( \xi \), if it is defined by

\[ \tilde{x}^*_1 = x^*_1, \quad \tilde{x}^*_t(\xi^t) \in \arg\min_{z \in X_t(\tilde{x}^*_{t-1}(\xi^{t-1}), \xi_t)} \|z - x^*_t(f^t(\xi^t))\| \]

for \( t = 2, \ldots, T \).

**Remark 1.** The class I, II calm modifications in Definition 2.4 was introduced in [14], which can be viewed as an extension of the calm modification in [16]. It is known from Assumption 2.1 that for any perturbation \( \xi \), the class I calm modification always exists, so does the class II calm modification. The measurability of class I, II calm modifications are ensured by [26, Theorem 14.37]. From the viewpoint of measurability, we know that \( \{f_t\}_{t=1}^T \) should be measurable too. If we consider a general perturbation to the stochastic process \( \xi \) as that in [12], it might be impossible to select a measurable class I calm modification and the complex filtration distance has to be adopted. In [16, Example A.3], the author illustrated that the measurability of \( f \) is indispensable for an approximating process. It is easy to see from Definition 2.4 that \( \tilde{x}^*(f^T(\xi)) = x^*(f^T(\xi)) \) and \( \tilde{x}^*(\xi) = x^*(\xi) \). Moreover,
both $\tilde{x}^*(\xi)$ and $\tilde{x}^*(\hat{\xi})$ are feasible solutions of problem (7) under the stochastic process $\xi$.

To illustrate the class I, II calm modifications under the LPM framework, we give the following numerical example.

**Example 2.** Consider the scenario trees corresponding to $\xi$, $f(\xi)$ and $\hat{\xi}$ respectively, as those shown in Figure 2.

![Scenario trees](image)

**Figure 2.** Scenario trees of $\xi$ (left), $f(\xi)$ (central), $\hat{\xi}$ (right) for Example 2

Here $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, assigned with probabilities $P(\omega_i) = 1/4$ for $i = 1, 2, 3, 4$, and we have the following mappings:

\[ \xi(\omega_1) = (1, 1, 1), \]
\[ \xi(\omega_2) = (1, 1, 2), \]
\[ \xi(\omega_3) = (1, 2, 2), \]
\[ \xi(\omega_4) = (1, 2, 3) \]

and

\[ \hat{\xi}(\omega_1) = (1, 1, 1), \]
\[ \hat{\xi}(\omega_2) = (1, 1, 2), \]
\[ \hat{\xi}(\omega_3) = (1, 2, 3), \]
\[ \hat{\xi}(\omega_4) = (1, 2, 3). \]

The measurable mapping $f$ is then defined as $f((1,1,1)) = (1,1,2)$, $f((1,1,2)) = (1,1,2)$, $f((1,2,2)) = (1,2,3)$ and $f((1,2,3)) = (1,2,3)$. Then, we obtain the filtration of $\xi$:

\[ \{\emptyset, \Omega\} \subseteq \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\} \subseteq 2^\Omega, \]

the filtration of $f(\xi)$:

\[ \{\emptyset, \Omega\} \subseteq \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\} \subseteq \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\} \]

and the filtration of $\hat{\xi}$:

\[ \{\emptyset, \Omega\} \subseteq \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\} \]
\[ \subseteq \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_3, \omega_4\}, \Omega\} \]

Moreover, we have the induced probability distributions of $\xi$, $f(\xi)$ and $\hat{\xi}$ as

\[
\text{Prob}(\xi = \xi) = \begin{cases} 
1/4 & \xi = (1, 1, 1) \\
1/4 & \xi = (1, 1, 2) \\
1/4 & \xi = (1, 2, 2) \\
1/4 & \xi = (1, 2, 3) 
\end{cases}, \quad \text{Prob}(f(\xi) = f(\xi)) = \begin{cases} 
1/2 & f(\xi) = (1, 1, 2) \\
1/2 & f(\xi) = (1, 2, 3) 
\end{cases}
\]
and
\[ \text{Prob}(\bar{\xi} = \xi) = \begin{cases} 
   1/4 & \xi = (1, 1, 1) \\
   1/4 & \xi = (1, 1, 2) \\
   1/2 & \xi = (1, 2, 3) 
\end{cases} \]

Notice that we use ‘Prob’ to distinguish with \( \mathbb{P} \), and in the rest of the paper, we use a superscript, like \( \mathbb{P}^t \), to stress that.

To simplify the illustrative numerical example, we choose the following specific values for of parameters. Let \( n = s = p = 1 \), \( c_1 = 1 \), \( c_2(\xi_2) = \xi_2 \), \( c_3(\xi_3) = \xi_3 \), \( \eta_2 = \eta_3 = 1/2 \), \( X_1 = [0, 1] \),
\[
\begin{align*}
X_2(x_1, \xi_2) &= \left\{ x_2 \in \mathbb{R} : \frac{1}{-1} x_2 + \left( \frac{\xi_2}{-\xi_2} \right) x_1 \leq \left( \frac{1 + \xi_2}{-1/2 - \xi_2} \right) \right\}, \\
X_3(x_2, \xi_3) &= \left\{ x_3 \in \mathbb{R} : \frac{1}{-1} x_3 \leq \left( \frac{\xi_3}{-\xi_3/2} \right) \right\}.
\end{align*}
\]
Then, under these specific settings, the original problem (7) or (8) can be formulated as
\[
\begin{align*}
\min_{x_1, x_2, x_3} & \quad x_1 + \mathbb{E}[\xi_2 x_2 - 1/2]_+ + \mathbb{E}[\xi_3 x_3 - 1/2]_+ \\
\text{s.t.} & \quad x_1 \in X_1, x_2 \in X_2(x_1, \xi_2), x_3 \in X_3(x_2, \xi_3).
\end{align*}
\]
Thus we can rewrite it as
\[
\begin{align*}
\min & \quad x_1 + \frac{1}{2} [x_2^1 - 1/2]_+ + \frac{1}{2} [2x_2^2 - 1/2]_+ + \frac{1}{2} [x_2^3 - 1/2]_+ \\
& \quad + \frac{1}{2} [x_3^1 - 1/2]_+ + \frac{1}{2} [2x_3^2 - 1/2]_+ + \frac{1}{2} [3x_3^3 - 1/2]_+ \\
\text{s.t.} & \quad x_1 \in [0, 1], 3/2 - x_1 \leq x_2^1 \leq 2 - x_1, 5/2 - 2x_1 \leq x_2^2 \leq 3 - 2x_1, \\
& \quad 1/2 \leq x_2^3 \leq 1, 1 \leq x_3^1 \leq 2, 1 \leq x_3^2 \leq 2, 3/2 \leq x_3^3 \leq 3,
\end{align*}
\]
where \( x_2^1 \) corresponds to \( x_2(\cdot) \) when \( (\xi_1, \xi_2) = (1, 1) \), \( x_2^2 \) corresponds to \( x_2(\cdot) \) when \( (\xi_1, \xi_2) = (1, 2) \), \( x_2^3 \) corresponds to \( x_3(\cdot) \) when \( (\xi_1, \xi_2, \xi_3) = (1, 1, 1) \), \( x_2^3 \) corresponds to \( x_3(\cdot) \) when \( (\xi_1, \xi_2, \xi_3) = (1, 1, 2) \), \( x_2^3 \) corresponds to \( x_3(\cdot) \) when \( (\xi_1, \xi_2, \xi_3) = (1, 2, 2) \), \( x_2^3 \) corresponds to \( x_3(\cdot) \) when \( (\xi_1, \xi_2, \xi_3) = (1, 2, 3) \). Then, it is not difficult to obtain the optimal solution
\[
x^*(\xi) = \begin{cases} 
(1, 1/2, 1/2) & \xi = (1, 1, 1), \\
(1, 1/2, 1/2) & \xi = (1, 1, 2) \text{ or } (1, 2, 2), \\
(1, 1/2, 3/2) & \xi = (1, 2, 3).
\end{cases}
\]
Then the class I calm modification \( \bar{x}^*(\hat{\xi}) \) under the stochastic process \( \hat{\xi} \) is defined by
\[
\bar{x}_1^* = x_1^* = 1
\]
and
\[
\bar{x}_2^*(\xi_2) = \arg\min_{z \in X_2(x_1^*, \xi_2)} \| z - x^*_2(f^*(\xi_2)) \|
= \arg\min_{z \in X_2(x_1^*, \xi_2)} \| z - 1/2 \|.
\]
Since \( X_2(x_1^*, \xi_2) \) is always \([1/2, 1]\) for any realization of \( \hat{\xi}_2 \), we have \( \bar{x}_2^*(\xi_2) = 1/2 \). Further, we consider
\[
\bar{x}_3^*(\hat{\xi}) = \arg\min_{z \in X_3(x_2^*(\xi_2), \xi_3)} \| z - x^*_3(f(\xi)) \|
\]
and obtain
\[
\bar{x}_3^*(\hat{\xi}) = \begin{cases} 
1 & \hat{\xi} = (1, 1, 1) \text{ or } (1, 1, 2), \\
3/2 & \hat{\xi} = (1, 2, 2) \text{ or } (1, 2, 3).
\end{cases}
\]
Finally, we have
\[ \bar{x}^*(\hat{\xi}) = \begin{cases} (1, 1/2, 1) & \hat{\xi} = (1, 1, 1) \text{ or } (1, 1, 2), \\ (1, 1/2, 3/2) & \hat{\xi} = (1, 2, 2) \text{ or } (1, 2, 3). \end{cases} \]

In what follows, we consider the class II calm modification under the stochastic process \( \hat{\xi} \). According to the definition, we have
\[ \bar{x}^*_1 = x^*_1 = 1, \]
\[ \bar{x}^*_2(\hat{\xi}^2) = \arg\min_{z \in X_2(\bar{x}_1, \hat{\xi}_2)} \| z - \bar{x}^*_2(\hat{\xi}^2) \| \]
\[ = \arg\min_{z \in [1/2, 1]} \| z - 1/2 \| = 1/2. \]

and
\[ \bar{x}^*_3(\hat{\xi}) = \arg\min_{z \in X_3(\bar{x}_2, \hat{\xi}_3)} \| z - \bar{x}^*_3(\hat{\xi}) \|, \]

which implies
\[ \bar{x}^*_3(\hat{\xi}) = \begin{cases} 1/2 & \hat{\xi} = (1, 1, 1), \\ 1 & \hat{\xi} = (1, 1, 2), \\ 3/2 & \hat{\xi} = (1, 2, 3). \end{cases} \]

Finally, we obtain \( \bar{x}^*(\hat{\xi}) \) as follows:
\[ \bar{x}^*(\hat{\xi}) = \begin{cases} (1, 1/2, 1/2) & \hat{\xi} = (1, 1, 1), \\ (1, 1/2, 1) & \hat{\xi} = (1, 1, 2), \\ (1, 1/2, 3/2) & \hat{\xi} = (1, 2, 3). \end{cases} \]

We can recursively deduce from Proposition 1 and Definition 2.4 the following estimation.

**Proposition 3.** Let Assumptions 2.1 and 2.2 hold. Then there exists a constant \( G > 0 \), such that
\[ \| \bar{x}^*_t(\xi^t) - \bar{x}^*_t(f^t(\xi^t)) \| \leq G \max\left\{ 1, \| \xi^t \|, \| f^t(\xi^t) \| \right\}^{t-2} \| f^t(\xi^t) - \xi^t \|, \quad (11a) \]
\[ \| \bar{x}^*_t(\xi^t) - \bar{x}^*_t(f^t(\xi^t)) \| \leq G \max\left\{ 1, \| \xi^t \|, \| f^t(\xi^t) \| \right\}^{t-2} \| f^t(\xi^t) - \xi^t \| \quad (11b) \]

for \( t = 2, \cdots, T \).

We omit the proof of Proposition 3. (11a) is similar to that in [16, Proposition 4.1] and one can refer to [14, Proposition 2.2] for (11b).

With these prerequisites, we can now study the quantitative stability of models (7) and (9).

3. **Stability of model (7).** We consider the stability of model (7) or (8) under different frameworks. To make our presentation clear, we divide this section into two parts. The first part studies the quantitative stability of model (7) under the approximation of stochastic process (see Definition 2.3). The second part discusses the quantitative stability under a general perturbation.

3.1. Stability analysis under the approximation of stochastic process. In this subsection, we first consider the following perturbed problem of model (7)

\[
\inf_{x_t} \rho[\phi(x_t, \tilde{\xi})] \\
\text{s.t. } x_t \in X_t(x_{t-1}, \tilde{\xi}_t), \ t = 1, 2, \cdots, T,
\]

where \(\tilde{\xi}\) is an approximation to the stochastic process \(\xi\) in terms of Definition 2.3. Denote by \(\tilde{x}^*(\tilde{\xi})\) and \(x^*(\xi)\) the optimal solutions of problems (12) and (7), respectively. \(\tilde{x}^*(\xi)\) and \(\tilde{x}^*(\xi)\) are the corresponding class I and class II calm modifications of \(\tilde{x}^*\) and \(x^*\), that is

\[
\tilde{x}_1^* := \tilde{x}_1^*, \ \tilde{x}_t^*(\xi^t) \in \arg\min_{z \in X_t(\tilde{x}_{t-1}^*(\xi^t-1), \tilde{\xi}_t)} \|\tilde{x}_t^*(\xi^t) - z\|, \ t = 2, 3, \cdots, T
\]

and

\[
\tilde{x}_1^* := x_1^*, \ \tilde{x}_t^*(\xi^t) \in \arg\min_{z \in X_t(\tilde{x}_{t-1}^*(\xi^t-1), \tilde{\xi}_t)} \|x_t^*(\xi^t) - z\|, \ t = 2, 3, \cdots, T.
\]

In addition, for any \(z, z_1, z_2 \in \mathbb{R}\) and \(t = 2, 3, \cdots, T\), we have the following two facts:

\[
(z - \eta)^p_+ = (\max\{z, \eta\} - \eta)^p
\leq 2^p |\max\{z, \eta\}|^p + 2^p |\eta|^p
\leq 2^p (|z| + |\eta|)^p + 2^p |\eta|^p
\leq 2^{2^p} |z|^p + (2^{2^p} + 2^p) |\eta|^p
\leq C_1 \max\{1, |z|^p\},
\]

where \(C_1 := 2^{2^p} + (2^{2^p} + 2^p) \max_{2 \leq t \leq T} \{|\eta|^p\}\), and

\[
|(z_1 - \eta)^p_+ - (z_2 - \eta)^p_+| \leq |(\max\{z_1, \eta\} - \eta)^p - (\max\{z_2, \eta\} - \eta)^p| \leq C_2 \max\{1, |z_1|^{p-1}, |z_2|^{p-1}\} |z_1 - z_2|,
\]

where \(C_2 := p \left(2^{2^p-2} + (2^{2^p-2} + 2^p-1) \max_{2 \leq t \leq T} \{|\eta_t|^{p-1}\}\right)\).

With the above preparations, we have the following two propositions.

Proposition 4. Suppose that Assumptions 2.1 and 2.2 hold, and \(\tilde{\xi}\) is defined in Definition 2.3. Then there exists a positive constant \(D\) such that

\[
\left|\rho[\phi(\tilde{x}^*(\tilde{\xi}), \tilde{\xi})] - \rho[\phi(x^*(\xi), \xi)]\right| \leq D \sum_{t=2}^{T} E \left[\max\left\{1, \|\xi^t\|, \|\tilde{\xi}^t\|\right\}^{p-1} \left\|\tilde{\xi}^t - \xi^t\right\|\right],
\]

\[
\left|\rho[\phi(\tilde{x}^*(\tilde{\xi}), \tilde{\xi})] - \rho[\phi(\tilde{x}^*(\tilde{\xi}), \tilde{\xi})]\right| \leq D \sum_{t=2}^{T} E \left[\max\left\{1, \|\xi^t\|, \|\tilde{\xi}^t\|\right\}^{p-1} \left\|\tilde{\xi}^t - \xi^t\right\|\right].
\]

Proof. We only prove the first assertion, and the second one can be verified similarly. According to the definitions of \(\rho\) and the class II calm modification, we have

\[
\rho[\phi(\tilde{x}^*(\tilde{\xi}), \tilde{\xi})] - \rho[\phi(x^*(\xi), \xi)] = \\
\sum_{t=2}^{T} E \left[\left(\langle c_t(\xi_t), \tilde{x}_t^*(\xi^t)\rangle - \eta_t\right)^p_+\right] - \sum_{t=2}^{T} E \left[\left(\langle c_t(\xi_t), \tilde{x}_t^*(\xi^t)\rangle - \eta_t\right)^p_+\right].
\]
Proposition 5.

Proof. For the first assertion, we have from the definition of (15)

\[ \left| \rho[\phi(x^*(\xi)), \xi] - \rho[\phi(x^*(\xi), \xi)] \right| \]
\[ \leq \sum_{t=2}^{T} E \left[ \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)^p - \left( \langle c_t(\xi_t), \tilde{x}_t^*(\xi_t) \rangle - \eta_t \right)^p \right] \]
\[ \leq C_2 \sum_{t=2}^{T} E \left[ \max \left\{ 1, \left| \langle c_t(\xi_t), x_t^*(\xi^t) \rangle \right|^{p-1}, \left| \langle c_t(\xi_t), \tilde{x}_t^*(\xi^t) \rangle \right|^{p-1} \right\} \cdot \left| \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \langle c_t(\xi_t), \tilde{x}_t^*(\xi^t) \rangle \right| \right]. \]

We know from (3) and (10) that

\[ \left| \langle c_t(\xi_t), x_t^*(\xi^t) \rangle \right| \leq BL \max \left\{ 1, \left\| \xi^t \right\| \right\}^t, \]
\[ \left| \langle c_t(\xi_t), \tilde{x}_t^*(\xi^t) \rangle \right| \leq BL \max \left\{ 1, \left\| \xi_t \right\| \right\}^{t-1}, \]
\[ \left| \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \langle c_t(\xi_t), \tilde{x}_t^*(\xi^t) \rangle \right| \leq BL \left\| \xi_t - \xi_t^t \right\| \max \left\{ 1, \left\| \xi_t^t \right\| \right\}^{t-1}. \]

Therefore, we have

\[ \left| \rho[\phi(x^*(\xi), \xi)] - \rho[\phi(x^*(\xi), \xi)] \right| \leq C_2(BL)^p \sum_{t=2}^{T} E \left[ \max \left\{ 1, \left\| \xi^t \right\|, \left\| \xi^t \right\| \right\}^{p-1} \left\| \xi^t - \xi^t \right\| \right]. \]

Letting \( D = C_2(BL)^p \) completes the proof of the first assertion.

Proposition 5. Suppose that Assumptions 2.1 and 2.2 hold, and \( \xi \) is defined in Definition 2.3. Then there exists a positive constant \( F \), such that the following assertions hold:

\[ \left| \rho[\phi(x^*(\xi), \xi)] - \rho[\phi(x^*(\xi), \xi)] \right| \leq F \sum_{t=2}^{T} E \left[ \max \left\{ 1, \left\| \xi^t \right\|, \left\| \xi^t \right\| \right\}^{p-1} \left\| \xi^t - \xi^t \right\| \right], \]
\[ \left| \rho[\phi(x^*(\xi), \xi)] - \rho[\phi(x^*(\xi), \xi)] \right| \leq F \sum_{t=2}^{T} E \left[ \max \left\{ 1, \left\| \xi^t \right\|, \left\| \xi^t \right\| \right\}^{p-1} \left\| \xi^t - \xi^t \right\| \right]. \]

Proof. For the first assertion, we have from the definition of \( \rho \) that

\[ \left| \rho[\phi(x^*(\xi), \xi)] - \rho[\phi(x^*(\xi), \xi)] \right| = \]
\[ \sum_{t=2}^{T} E \left[ \left( \langle c_t(\xi_t), x_t^*(\xi_t) \rangle - \eta_t \right)^p \right] - \sum_{t=2}^{T} E \left[ \left( \langle c_t(\xi_t), x_t^*(\xi_t) \rangle - \eta_t \right)^p \right]. \]  

Note that \( x_t^*(\xi_t^t) = \tilde{x}_t^*(\xi_t^t) \). Thus, we have the following upper bound estimation of (15)

\[ \sum_{t=2}^{T} E \left[ \left( \langle c_t(\xi_t), x_t^*(\xi_t) \rangle - \eta_t \right)^p - \left( \langle c_t(\xi_t), \tilde{x}_t^*(\xi_t^t) \rangle - \eta_t \right)^p \right]. \]
Based on (14), we obtain
\[
\sum_{t=2}^{T} \mathbb{E} \left[ \left( \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle - \eta_t \right)^p - \left( \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle - \eta_t \right)^p \right] 
\leq C_2 \sum_{t=2}^{T} \mathbb{E} \left[ \max \left\{ 1, \left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right|^{p-1}, \left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right|^{p-1} \right\} \cdot \left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle - \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right| \right].
\]

Analogous to that in Proposition 4, we have
\[
\left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right| \leq BL \max \left\{ 1, \| \xi_t^t \| \right\} \max \left\{ 1, \| \xi_t^t \| \right\}^{t-1},
\]
\[
\| c_t(\xi_t), \bar{x}_t^*(\xi_t') \| \leq BL \max \left\{ 1, \| \xi_t^t \| \right\}^{t},
\]
\[
\left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle - \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right| \leq BG \max \left\{ 1, \| \xi_t^t \|, \| \xi_t^t \| \right\}^{t-1} \| \xi_t^t - \xi_t^t \|,
\]
where the last inequality follows from (11b) in Proposition 3. Therefore, we obtain that
\[
\left| \rho[\phi(\bar{x}^*(\xi), \xi)] - \rho[\phi(x^*(\xi), \xi)] \right| \leq F \sum_{t=2}^{T} \mathbb{E} \left[ \max \left\{ 1, \| \xi_t^t \|, \| \xi_t^t \| \right\}^{p-1} \| \xi_t^t - \xi_t^t \| \right],
\]
where \( F = C_2 B^p L^{p-1} G \).

The second assertion can be proved similarly, thus we omit it here.

With the above preparatory work, we now give the quantitative stability conclusion of problem (7).

**Theorem 3.1.** Let Assumptions 2.1 and 2.2 hold and \( \tilde{\xi} \) be defined in Definition 2.3. Then there exists a constant \( H > 0 \) such that
\[
\left| v(\xi) - v(\tilde{\xi}) \right| \leq H \sum_{t=2}^{T} \mathbb{E} \left[ \max \left\{ 1, \| \xi_t^t \|, \| \xi_t^t \| \right\}^{p-1} \| \xi_t^t - \xi_t^t \| \right].
\]

**Proof.** If
\[
\left| v(\xi) - v(\tilde{\xi}) \right| = v(\tilde{\xi}) - v(\xi),
\]
we have that
\[
v(\tilde{\xi}) - v(\xi) = \rho[\phi(\bar{x}^*(\tilde{\xi}), \tilde{\xi})] - \rho[\phi(x^*(\xi), \xi)]
\leq \rho[\phi(\bar{x}^*(\tilde{\xi}), \tilde{\xi})] - \rho[\phi(x^*(\xi), \xi)]
= \rho[\phi(\bar{x}^*(\tilde{\xi}), \tilde{\xi})] - \rho[\phi(\bar{x}^*(\tilde{\xi}), \tilde{\xi})] + \rho[\phi(\bar{x}^*(\tilde{\xi}), \tilde{\xi})] - \rho[\phi(x^*(\xi), \xi)].
\]
Then, we get from Proposition 4 and Proposition 5 that
\[
v(\tilde{\xi}) - v(\xi) \leq (D + F) \sum_{t=2}^{T} \mathbb{E} \left[ \max \left\{ 1, \| \xi_t^t \|, \| \xi_t^t \| \right\}^{p-1} \| \xi_t^t - \xi_t^t \| \right].
\]
If
\[
\left| v(\xi) - v(\tilde{\xi}) \right| = v(\xi) - v(\tilde{\xi}),
\]
then

\[
\sum_{t=2}^{T} \mathbb{E} \left[ \left( \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle - \eta_t \right)^p - \left( \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle - \eta_t \right)^p \right] \leq C_2 \sum_{t=2}^{T} \mathbb{E} \left[ \max \left\{ 1, \left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right|^{p-1}, \left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right|^{p-1} \right\} \cdot \left| \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle - \langle c_t(\xi_t), \bar{x}_t^*(\xi_t') \rangle \right| \right].
\]
we can similarly derive from Proposition 4 and Proposition 5 that
\[ v(\xi) - v(\hat{\xi}) = \rho[\phi(x^*(\xi), \xi)] - \rho[\phi(x^*(\hat{\xi}), \hat{\xi})] \]
\[ \leq \rho[\phi(\tilde{x}^*(\xi), \xi)] - \rho[\phi(\tilde{x}^*(\hat{\xi}), \hat{\xi})] \]
\[ = \rho[\phi(\tilde{x}^*(\xi), \xi)] - \rho[\phi(\tilde{x}^*(\hat{\xi}), \xi)] + \rho[\phi(\tilde{x}^*(\hat{\xi}), \xi)] - \rho[\phi(\tilde{x}^*(\hat{\xi}), \hat{\xi})] \]
\[ \leq (D + F) \sum_{t=2}^{T} \mathbb{E} \left[ \max \left\{ 1, \|\xi^t\|, \|\xi^t\| \right\} t^{\rho t - 1} \|\tilde{\xi}^t - \xi^t\| \right]. \]

The proof completes by letting \( H = D + F. \)

**Remark 2.** Theorem 3.1 gives a quantitative stability assertion when the perturbed stochastic process \( \hat{\xi} \), defined in Definition 2.3, is measurable. This kind of perturbation, as far as we know, was first investigated in [16] for the multistage case, where the author studied a risk-neutral model under some recursive assumptions, see [16, Assumption 2.6]. Heitsch et al considered in [12] a general perturbation for the risk-neutral model, that is, they did not require the measurability of the perturbed stochastic process \( \hat{\xi} \) with respect to \( \xi \). As a result, they employed the so-called filtration distance. The quantitative stability conclusion in Theorem 3.1 can be simplified by adopting a common relaxed term
\[ \left| v(\xi) - v(\hat{\xi}) \right| \leq H(T - 1) \mathbb{E} \left[ \max \left\{ 1, \|\xi\|, \|\hat{\xi}\| \right\} T^{\rho t - 1} \|\tilde{\xi} - \xi\| \right]. \]

Specially, for the risk-neutral model (1), one can verify that the difference of optimal values under stochastic processes \( \xi \) and \( \hat{\xi} \) can be bounded above by
\[ \text{Constant} \cdot \mathbb{E} \left[ \max \left\{ 1, \|\xi\|, \|\hat{\xi}\| \right\} T^{\rho t - 1} \|\tilde{\xi} - \xi\| \right]. \]

More details can be found in [14].

### 3.2. Stability analysis under a general perturbation

We examined in Theorem 3.1 the quantitative stability under the measurable perturbation (see Definition 2.3). To extend this result to the general perturbation (9), such as that in [12], as well as to the distributionally robust counterpart, we consider the equivalent formulation (8) and its perturbed problem under the general stochastic process \( \hat{\xi} \):
\[
\inf_{x} \langle c_{1}, x_{1} \rangle + \sum_{t=2}^{T} \mathbb{E}_{Q^t} \left[ \left( (c_{t}(\xi_{t}), x_{t}) - \eta_{t} \right)_{+} \right] \tag{16}
\]
\[ \text{s.t. } x_{t} \in X_{t}(x_{t-1}, \xi_{t}), \ t = 1, 2, \cdots, T. \]

Similar to what we have stated in Section 2, \( Q^t := \mathbb{P} \circ (\hat{\xi}^t)^{-1} \) is the general perturbation probability measure of \( \mathbb{P}^t \) for \( t = 2, 3, \cdots, T \) under the perturbed stochastic process \( \hat{\xi} \). In particular, if \( \xi = \hat{\xi} \), problem (16) is equivalent to problem (12). Then we have the following general stability result.

**Theorem 3.2.** Suppose that Assumptions 2.1 and 2.2 hold and \( \hat{\xi} \) is a general perturbation stochastic process of \( \xi \) with \( \hat{\xi} \in L_{r}(\Omega, \mathcal{F}, \mathbb{P}) \) and \( r \geq p + 1 \). Then there exists a constant \( E > 0 \), such that
\[ \left| v(\xi) - v(\hat{\xi}) \right| \leq E \sum_{t=2}^{T} \mathbb{D}_{t}(\mathbb{P}^t, Q^t), \]
where \( P^t := P \circ (\xi^t)^{-1} \), \( Q^t := P \circ (\xi^t)^{-1} \) and

\[
\mathbb{D}_t(P^t, Q^t) := \mathbb{D}_{TV}(P^t, Q^t)^{|P^t|} \cdot \left( \zeta_1(P^t, Q^t) + (pt + 1)\zeta_{pt+1}(P^t, Q^t) + 2\mathbb{E}_{P^t} [||\xi^t||] + 2\mathbb{E}_{P^t} [||\xi^t||^{pt+1}] + 1 \right)
\]

for \( t = 2, 3, \cdots, T \).

**Proof.** Denote by \( x^*(\xi) \) and \( \hat{x}^*(\hat{\xi}) \) the optimal solutions of problems (8) and (16), respectively. Then we have

\[
v(\xi) - v(\hat{\xi}) = \langle c_1(\xi_1), x_1^* \rangle + \sum_{t=2}^T \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)_+ P^t(d\xi^t) \\
- \langle c_1(\xi_1), \hat{x}_1^* \rangle - \sum_{t=2}^T \int_{\Xi^t} \left( \langle c_t(\xi_t), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ Q^t(d\xi^t).
\]

To simplify the notation, we define

\[
\int_{\Xi^1} \left( \langle c_1(\xi_1), x_1^*(\xi^1) \rangle - \eta_1 \right)_+ P^1(d\xi^1) = \langle c_1(\xi_1), x_1^* \rangle \\
\text{and}
\int_{\Xi^1} \left( \langle c_1(\xi_1), \hat{x}_1^*(\xi^1) \rangle - \eta_1 \right)_+ Q^1(d\xi^1) = \langle c_1(\xi_1), \hat{x}_1^* \rangle.
\]

Moreover, when

\[
\sum_{t=1}^T \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)_+ P^t(d\xi^t) \geq \sum_{t=1}^T \int_{\Xi^t} \left( \langle c_t(\xi_t), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ Q^t(d\xi^t),
\]

we have

\[
\left| \sum_{t=1}^T \left( \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)_+ P^t(d\xi^t) - \int_{\Xi^t} \left( \langle c_t(\xi_t), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ Q^t(d\xi^t) \right) \right| \\
= \sum_{t=1}^T \left( \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)_+ P^t(d\xi^t) - \int_{\Xi^t} \left( \langle c_t(\xi_t), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ Q^t(d\xi^t) \right) \\
\leq \sum_{t=1}^T \int_{\Xi^t} \left( \langle c_t(\xi_t), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t).
\]

Otherwise, we have

\[
\left| \sum_{t=1}^T \left( \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)_+ P^t(d\xi^t) - \int_{\Xi^t} \left( \langle c_t(\xi_t), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ Q^t(d\xi^t) \right) \right| \\
= \sum_{t=1}^T \left( \int_{\Xi^t} \left( \langle c_t(\xi_t), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ Q^t(d\xi^t) - \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)_+ P^t(d\xi^t) \right) \\
\leq \sum_{t=1}^T \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^*(\xi^t) \rangle - \eta_t \right)_+ (Q^t - P^t)(d\xi^t).
\]
Therefore, we obtain

\[
|v(\xi) - v(\hat{\xi})| \leq \max \left\{ \sum_{t=1}^{T} \int_{\Xi} \left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t), \right. \\
\sum_{t=1}^{T} \int_{\Xi} \left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (Q^t - P^t)(d\xi^t) \left. \right\}
\]

Then, for \( t = 2, 3, \ldots, T \), we consider

\[
\int_{\Xi} \left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t) = \\
\int_{\Xi} \left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t) + \\
\int_{\Xi} \left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t),
\]

(18)

where \( R_t > 0 \) and \( \Xi_{R_t} := \{ \xi^t \in \Xi : \|\xi^t\| \leq R_t \} \).

We know from (13) that

\[
\left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ \leq C_1 \max \left\{ 1, |\langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle|^p \right\} \leq C_1 BL \left( 1 + \|\xi^t\|^{pt} \right).
\]

For the first term on the right-hand side of (18), we have

\[
\int_{\Xi} \left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t) \leq \frac{1}{C_1 BL (1 + R_t^{pt})} \mathbb{D}_{TV}(P^t, Q^t) \leq \frac{1}{C_1 BL R_t^{pt}} \mathbb{D}_{TV}(P^t, Q^t).
\]

For the second term on the right-hand side of (18), we have

\[
\int_{\Xi} \left( \langle c_t(\xi_t), \mathbf{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t + Q^t)(d\xi^t) \leq \frac{1}{R_t} \int_{\Xi} \left. \right. \left. \right. C_1 BL \left( \|\xi^t\| + \|\xi^t\|^{pt+1} \right) (P^t + Q^t)(d\xi^t)
\]

\[
\leq \frac{C_1 BL R_t^{pt+1}}{R_t} \int_{\Xi} \left( \|\xi^t\| + \|\xi^t\|^{pt+1} \right) (P^t + Q^t)(d\xi^t).
\]
Meanwhile, it is known from [27, Page 490] that
\[
\int_{\mathbb{E}} \left\| \xi^t \right\|^2 (Q^t - P^t)(d\xi^t) \leq \zeta_1(P^t, Q^t),
\]
\[
\int_{\mathbb{E}} \left\| \xi^t \right\|^{pt+1} (Q^t - P^t)(d\xi^t) \leq (pt + 1)\zeta_{pt+1}(P^t, Q^t).
\]
Thus, we obtain
\[
\int_{\mathbb{E}} \left( \left\| \xi^t \right\| + \left\| \xi^t \right\|^{pt+1} \right) (P^t + Q^t)(d\xi^t) \leq \\
\zeta_1(P^t, Q^t) + (pt + 1)\zeta_{pt+1}(P^t, Q^t) + 2E_{P^t} \left[ \left\| \xi^t \right\| \right] + 2E_{P^t} \left[ \left\| \xi^t \right\|^{pt+1} \right].
\]
Then, we can continue that
\[
\int_{\mathbb{E}\setminus E_{\mathcal{H}}} \left( \langle c_t(\xi), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t) \leq \\
\frac{C_t BL}{R_t} \left( \zeta_1(P^t, Q^t) + (pt + 1)\zeta_{pt+1}(P^t, Q^t) + 2E_{P^t} \left[ \left\| \xi^t \right\| \right] + 2E_{P^t} \left[ \left\| \xi^t \right\|^{pt+1} \right] \right).
\]
Combining the above estimations for the two terms on the right-hand side of (18), we obtain that
\[
\int_{\mathbb{E}} \left( \langle c_t(\xi), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right)_+ (P^t - Q^t)(d\xi^t) \leq \\
\frac{C_t BL}{R_t} \frac{D_{TV}(P^t, Q^t) + 1}{C_t BL R_t^{pt}} \left( \zeta_1(P^t, Q^t) + (pt + 1)\zeta_{pt+1}(P^t, Q^t) + 2E_{P^t} \left[ \left\| \xi^t \right\| \right] + 2E_{P^t} \left[ \left\| \xi^t \right\|^{pt+1} \right] \right).
\]
By the same argument, we can derive the same upper bound for
\[
\int_{\mathbb{E}} \left( \langle c_t(\xi), \hat{x}_t^*(\xi^t) \rangle - \eta_t \right) (Q^t - P^t)(d\xi^t).
\]
If $D_{TV}(P^t, Q^t) \neq 0$, let $R_t = D_{TV}(P^t, Q^t)^{\frac{1}{m-1}}$, and we obtain
\[
\left| v(\xi) - v(\hat{\xi}) \right| \leq \sum_{t=2}^{T} \left( \frac{D_{TV}(P^t, Q^t)}{C_t BL R_t^{pt}} + \frac{C_t BL}{R_t} \left( \zeta_1(P^t, Q^t) + (pt + 1)\zeta_{pt+1}(P^t, Q^t) + 2E_{P^t} \left[ \left\| \xi^t \right\| \right] + 2E_{P^t} \left[ \left\| \xi^t \right\|^{pt+1} \right] \right) \right) \\
\leq \sum_{t=2}^{T} D_{TV}(P^t, Q^t)^{\frac{1}{m-1}} \left( \zeta_1(P^t, Q^t) + (pt + 1)\zeta_{pt+1}(P^t, Q^t) + 2E_{P^t} \left[ \left\| \xi^t \right\| \right] + 2E_{P^t} \left[ \left\| \xi^t \right\|^{pt+1} \right] + 1 \right) \\
= \sum_{t=2}^{T} D_{t}(P^t, Q^t),
\]
where $E := \max \left\{ \frac{1}{C_t BL}, C_t BL \right\}$.
If $D_{TV}(P^t, Q^t) = 0$, letting $R_t \to +\infty$, we obtain $\left| v(\xi) - v(\hat{\xi}) \right| = 0$. To sum up, we have
\[
\left| v(\xi) - v(\hat{\xi}) \right| \leq E \sum_{t=2}^{T} D_{t}(P^t, Q^t),
\]
which completes the proof.

**Remark 3.** We have two explanations about Theorem 3.1 and Theorem 3.2 as follows.

(i) Compared with a measurable perturbation in Theorem 3.1, Theorem 3.2 discusses a general perturbation from the viewpoint of probability measures over the support set. Unfortunately, it seems to be impossible for us to derive a quantitative comparison between the upper bounds in Theorem 3.1 and Theorem 3.2. Theorem 3.1 has a restrictive perturbation of $\xi$, i.e., the approximation of stochastic process in terms of Definition 2.3. The expectation is taken with respect to the probability $P$ in probability space $(\Omega, \mathcal{F}, P)$, which is numerically tractable. Theorem 3.2 is derived under a new framework, whose perturbation can be a more general one. Moreover, it is more friendly for us to extend it to the distributionally robust counterpart (see Section 4). However, to some extent, it might be numerically intractable due to the adoption of $\zeta$-structure probability metrics.

(ii) If we have an upper bound estimation for $\zeta_{pt+1}(P_t, Q_t)$, denoted by $\varrho$, the quantitative stability results can be simplified greatly. Note the fact that $\zeta_1(P_t, Q_t) \leq \zeta_{pt+1}(P_t, Q_t)$. Concretely, by letting $R_t = \mathbb{D}_{TV}(P_t, Q_t)^{\frac{1}{pt+1}}$, we have

$$
\left| v(\xi) - v(\hat{\xi}) \right| \leq \sum_{t=2}^{T} \left( \frac{1}{C_1 BL R_t^p} \mathbb{D}_{TV}(P_t, Q_t) \right.
+ \left. \frac{C_1 BL}{R_t} \left( (pt + 2) \varrho + 2 E_{P_t} \left[ ||\xi^t|| \right] + 2 E_{P_t} \left[ ||\xi^{t+1}||^{pt+1} \right] \right) \right)
\leq \sum_{t=2}^{T} \lambda_t \mathbb{D}_{TV}(P_t, Q_t)^{\frac{1}{pt+1}},
$$

where $\lambda_t = \frac{1}{C_1 BL} + C_1 BL \left( (pt + 2) \varrho + E_{P_t} \left[ ||\xi^t|| \right] + E_{P_t} \left[ ||\xi^{t+1}||^{pt+1} \right] \right)$.

With the reformulation (8) and Theorem 3.2, it provides us an avenue to extend the quantitative stability conclusion to the distributionally robust case.

4. **Stability analysis of model (9).** In this section, we turn to the perturbation analysis of problem (9). Let the perturbed ambiguity set be $Q := (Q^2, Q^3, \cdots, Q^T)$. Then the perturbed problem of model (9) can be formulated as

$$
\inf_{x} \langle c_1, x_1 \rangle + \sum_{t=2}^{T} \sup_{Q_t \in Q^t} E_{Q_t} \left[ \langle c_t(\xi_t), x_t \rangle - \eta_t \right]_{P_t}^{P}
\text{ s.t. } x_t \in X_t(x_{t-1}, \xi_t), \ t = 1, 2, \cdots, T.
$$

(19)

Following the notation introduced before, we use $\nu(Q)$ to denote the optimal value of problem (19). To continue the discussion and simplify notations in what follows, we define by convention that

$$\mathbb{D}_t(P^t, Q^t) = \inf_{Q^t \in Q^t} \mathbb{D}_t(P^t, Q^t),$$
$$\mathbb{D}_t(P^t, Q^t) = \sup_{P^t \in P^t} \mathbb{D}_t(P^t, Q^t),$$
$$\mathbb{H}_t(P^t, Q^t) = \max \left\{ \mathbb{D}_t(P^t, Q^t), \mathbb{D}_t(Q^t, P^t) \right\}.$$
Then, we have the following discrepancy estimation between problems (9) and (19).

**Theorem 4.1.** Suppose that: (i) Assumptions 2.1 and 2.2 hold; (ii) \( \mathcal{P} \) and \( \mathcal{Q} \) are the ambiguity sets of the original problem and the perturbed problem with solutions of problems (9) and (19), respectively. Then, we have the following estimate

\[
\sup_{\mathcal{P}^t \in \mathcal{P}_t} \mathbb{E}_{\mathbf{P}^t}[||\mathbf{Q}^t||^{p+1}] < +\infty
\]

and

\[
\mathbb{E}_{\mathbf{Q}^t}[||\mathbf{Q}^t||^{p+1}] < +\infty
\]

for \( \mathbf{Q}^t \in \mathcal{Q}^t \) and \( t = 2, \ldots, T \). Then

\[
|\nu(\mathcal{Q}) - \nu(\mathcal{P})| \leq E \sum_{t=2}^{T} H_t(\mathcal{P}^t, \mathcal{Q}^t),
\]

where \( E \) is the positive constant defined in Theorem 3.2.

**Proof.** Let \( \mathbf{x}^P := (x_1^P, x_2^P, \ldots, x_T^P) \) and \( \mathbf{x}^Q := (x_1^Q, x_2^Q, \ldots, x_T^Q) \) be any optimal solutions of problems (9) and (19), respectively. Then, we have the following estimation:

\[
|\nu(\mathcal{Q}) - \nu(\mathcal{P})| = \left| \langle c_1, x_1^Q \rangle + \sum_{t=2}^{T} \sup_{\mathcal{Q}^t \in \mathcal{Q}_t} \mathbb{E}_{\mathbf{Q}^t} \left[ \left( \langle c_t(\xi_t), x_t^Q(\xi_t) \rangle - \eta_t \right)_+ \right] 
- \left( \langle c_1, x_1^P \rangle + \sum_{t=2}^{T} \sup_{\mathcal{P}^t \in \mathcal{P}_t} \mathbb{E}_{\mathbf{P}^t} \left[ \left( \langle c_t(\xi_t), x_t^P(\xi_t) \rangle - \eta_t \right)_+ \right] \right) \right|
\]

By adopting the similar argument as that in Theorem 3.2, we have

\[
|\nu(\mathcal{Q}) - \nu(\mathcal{P})| 
\leq \max \left\{ \sum_{t=2}^{T} \sup_{\mathcal{Q}^t \in \mathcal{Q}_t} \inf_{\mathbf{P}^t \in \mathcal{P}_t} \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^P(\xi_t) \rangle - \eta_t \right)_+ (\mathbf{Q}^t - \mathbf{P}^t)(d\xi_t), \right. 
\sum_{t=2}^{T} \sup_{\mathcal{P}^t \in \mathcal{P}_t} \inf_{\mathcal{Q}^t \in \mathcal{Q}_t} \int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^Q(\xi_t) \rangle - \eta_t \right)_+ (\mathbf{P}^t - \mathbf{Q}^t)(d\xi_t) \left. \right\}
\]

and

\[
\int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^P(\xi_t) \rangle - \eta_t \right)_+ (\mathbf{Q}^t - \mathbf{P}^t)(d\xi_t) \leq E D_t(\mathcal{P}^t, \mathcal{Q}^t),
\]

\[
\int_{\Xi^t} \left( \langle c_t(\xi_t), x_t^Q(\xi_t) \rangle - \eta_t \right)_+ (\mathbf{P}^t - \mathbf{Q}^t)(d\xi_t) \leq E D_t(\mathcal{P}^t, \mathcal{Q}^t).
\]

for any \( \mathbf{P}^t \in \mathcal{P}_t \) and \( \mathbf{Q}^t \in \mathcal{Q}_t \). Therefore, we obtain that

\[
|\nu(\mathcal{Q}) - \nu(\mathcal{P})| \leq E \sum_{t=2}^{T} \max \left\{ \sup_{\mathcal{Q}^t \in \mathcal{Q}_t} \inf_{\mathbf{P}^t \in \mathcal{P}_t} D_t(\mathcal{P}^t, \mathcal{Q}^t), \sup_{\mathcal{P}^t \in \mathcal{P}_t} \inf_{\mathcal{Q}^t \in \mathcal{Q}_t} D_t(\mathcal{P}^t, \mathcal{Q}^t) \right\}
\]

Then we complete the proof.

**Remark 4.** As for Theorem 4.1, we have the following remarks.
(i) In Theorem 4.1, we give a quantitative stability assertion for the multistage distributionally robust optimization problem by using the well-known total variation metric. It is noteworthy that in [19], the authors also studied the quantitative stability of the multistage distributionally robust optimization problem by using the nested distance under some assumptions of continuity and convexity. In Theorem 4.1, we only require some boundedness assumption. Moreover, we consider the risk-averse model under the SEC risk measure framework, which is also different from that in [19].

(ii) We require \( \sup_{P_t \in \mathcal{P}_t} E_{P_t} [ \| \xi_t \|_{p_t + 1} ] < +\infty \). The main reason is that (see (17))

\[
\mathbb{D}_t(P^t, Q^t) = \mathbb{D}_{TV}(P^t, Q^t)^{1/p_t+1} \cdot \left( \zeta_1(P^t, Q^t) + (p_t + 1) \zeta_{p_t+1}(P^t, Q^t) \right) + 2E_{P_t} \left( \| \xi_t \|_{p_t + 1} \right) + 1
\]

is well-defined only when \( \sup_{P_t \in \mathcal{P}_t} E_{P_t} [ \| \xi_t \|_{p_t + 1} ] < +\infty \). Under this circumstance, we surely have that \( \sup_{P_t \in \mathcal{P}_t} E_{P_t} [ \| \xi_t \|_{p_t + 1} ] < +\infty \). It should be pointed out that \( \xi_t \in \mathcal{L}(\Omega, F_t, \mathbb{P}) \) for some \( r \geq p_t + 1 \) in Theorem 3.2, which is clarified in Section 2. \( E_{Q_t} [ \| \xi_t \|_{p_t + 1} ] < +\infty \) for \( Q^t \in \mathcal{Q}^t \) means that the perturbation should be contained in \( \mathcal{P}_{p_t+1}(\mathbb{Z}) \), which ensures the well-definedness of \( \zeta_{p_t+1}(P^t, Q^t) \), as well as \( \zeta_1(P^t, Q^t) \).

**Remark 5.** There are only a few quantitative stability results for multistage stochastic programming problems. For the risk-neutral ones, the pioneering work is [12], where the difference between optimal values under different stochastic processes is bounded by the following term:

\[
\text{Constant} \cdot \left( \sum_{t=1}^{T} E[\| \xi_t - \xi_t' \|_r^r] \right)^{1/r} + \mathbb{D}_f(\xi, \xi')
\]

for some positive scalar \( r \), where \( \xi' \) is a general perturbation of \( \xi \) and \( \mathbb{D}_f \) is the filtration distance (see [12, (2.4)]). By noticing that the filtration distance is not intuitional numerically, Küchler presented in [16] the following upper bound for the difference between optimal values under different stochastic processes:

\[
\text{Constant} \cdot E \left[ 1, \| \xi \|, \| \xi' \| \right]^{m_1} \| \xi - \xi' \|
\]

for some positive constant \( m_1 \). This upper bound is more numerical friendly but it requires the perturbed stochastic process \( \tilde{\xi} \) is measurable with respect to \( \xi \) (see [16, Definition 3.1]), and, moreover, some strong recursive assumptions were imposed to derive the stability result.

In view of these limitations, Jiang and Chen reconsidered the quantitative stability of multistage linear stochastic programs in [14]. By introducing two types of calm modifications, they avoided those recursive assumptions and obtained a tighter upper bound when the perturbed process is measurable with regard to \( \xi \).

For the risk-averse ones, the only work [7] studied the quantitative stability of multistage stochastic programs with polyhedral risk measures in the objective. Analogous to the analysis in [12], the quantitative stability result relied on the filtration distance, see [7, Theorem 3.2].

This paper has generalized the above works. Firstly, under the measurable perturbation, we established the quantitative stability of multistage stochastic programs with separable expected conditional LPM objective functions by adopting
two kinds of calm modifications in [14], which avoided recursive assumptions in [16] and extended the quantitative stability results in [14] to the risk-averse case. Furthermore, to extend the measurable perturbation to the general perturbation (such as [12]) as well as to avoid the filtration distance, we proposed a new framework. This framework paved the way to the quantitative stability analysis of the distributionally robust counterpart of risk-averse multistage stochastic programs.

5. Conclusions. In this article, we have studied the quantitative stability of the risk-averse multistage stochastic programs with the $p$th order LPM risk measure. Based on the approximation of stochastic process and calm modifications, we first derive the quantitative stability results under the approximation of stochastic process. Then, we extend this result to the general perturbation case. Finally, we consider the distributionally robust counterpart of the risk-averse model, and establish the stability conclusion under the perturbation of the ambiguity set.

Our results improve the current results for risk-neutral multistage stochastic programs in [14], and extend them to risk-averse multistage stochastic programs with nonlinear risk measures. Moreover, we establish the quantitative stability results under a new framework, which also lays a foundation for the stability analysis of its distributionally robust counterpart.

There are still several relevant issues waiting to be settled. For example, the quantitative stability analysis of risk-averse multistage stochastic programs induced by other multi-period risk measures; the construction of the ambiguity sets of the multistage distributionally robust optimization problem in a dynamic form, and its theoretical analysis. These topics are left for future research.

Acknowledgments. The authors are grateful to the editor and the anonymous referee for their detailed and valuable comments and suggestions, which have helped us to significantly improve the presentation and quality of this paper.

REFERENCES

[1] S. A. Anthonisz, Asset pricing with partial-moments, *J. Bank. Financ.*, 36 (2012), 2122–2135.
[2] A. Ben-Tal, L. El Ghaoui and A. Nemirovski, *Robust Optimization*, Princeton Series in Applied Mathematics, 28, Princeton University Press, Princeton, NJ, 2009.
[3] A. J. Broganan and S. Stidham, Non-separation in the mean–lower-partial-moment portfolio optimization problem, *European J. Oper. Res.*, 184 (2008), 701–710.
[4] Z. Chen and J. Jiang, Stability analysis of optimization problems with $k$th order stochastic and distributionally robust dominance constraints induced by full random recourse, *SIAM J. Optim.*, 28 (2018), 1396–1419.
[5] J. Dupačová and V. Kozmík, Structure of risk-averse multistage stochastic programs, *OR Spectr.*, 37 (2015), 559–582.
[6] A. Eichhorn and W. Römisch, Polyhedral risk measures in stochastic programming, *SIAM J. Optim.*, 16 (2005), 69–95.
[7] A. Eichhorn and W. Römisch, Stability of multistage stochastic programs incorporating polyhedral risk measures, *Optimization*, 57 (2008), 295–318.
[8] P. C. Fishburn, Mean-risk analysis with risk associated with below-target returns, *Amer. Econ. Rev.*, 67 (1977), 116–126.
[9] V. Guigues and W. Römisch, Sampling-based decomposition methods for multistage stochastic programs based on extended polyhedral risk measures, *SIAM J. Optim.*, 22 (2012), 286–312.
[10] Y. Han and Z. Chen, Quantitative stability of full random two-stage stochastic programs with recourse, *Optim. Lett.*, 9 (2015), 1075–1090.
[11] H. Heitsch and W. Römisch, Scenario tree modeling for multistage stochastic programs, *Math. Program.*, 118 (2009), 371–406.
[12] H. Heitsch, W. Römisch and C. Strugarek, Stability of multistage stochastic programs, *SIAM J. Optim.*, 17 (2006), 511–525.

[13] T. Homem-de Mello and B. K. Pagnoncelli, Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective, *European J. Oper. Res.*, 249 (2016), 188–199.

[14] J. Jiang and Z. Chen, Quantitative stability of multistage stochastic programs via calm modifications, *Oper. Res. Lett.*, 46 (2018), 543–547.

[15] R. Kovacevic and G. C. Pflug, Time consistency and information monotonicity of multiperiod acceptability functionals, in *Advanced Financial Modelling*, Radon Ser. Comput. Appl. Math., 8, Walter de Gruyter, Berlin, 2009, 347–369.

[16] C. Küchler, On stability of multistage stochastic programs, *SIAM J. Optim.*, 19 (2008), 952–968.

[17] D. Kuhn, *Generalized Bounds for Convex Multistage Stochastic Programs*, Lecture Notes in Economics and Mathematical Systems, 548, Springer-Verlag, Berlin, 2006.

[18] J. Liu and Z. Chen, Time consistent multi-period robust risk measures and portfolio selection models with regime-switching, *European J. Oper. Res.*, 268 (2018), 373–385.

[19] Y. Liu, A. Pichler and H. Xu, Discrete approximation and quantification in distributionally robust optimization, *Math. Oper. Res.*, 44 (2019), 19–37.

[20] P. Mohajerin Esfahani and D. Kuhn, Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations, *Math. Program.*, 171 (2018), 115–166.

[21] T. Pennanen, Epi-convergent discretizations of multistage stochastic programs via integration quadratures, *Math. Program.*, 116 (2009), 461–479.

[22] G. C. Pflug and A. Pichler, *Multistage Stochastic Optimization*, Springer Series in Operations Research and Financial Engineering, Springer, Cham, 2014.

[23] G. C. Pflug and W. Römisch, *Modeling, Measuring and Managing Risk*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007.

[24] A. B. Philpott and V. L. De Matos, Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion, *European J. Oper. Res.*, 218 (2012), 470–483.

[25] A. Philpott, V. de Matos and E. Finardi, On solving multistage stochastic programs with coherent risk measures, *Oper. Res.*, 61 (2013), 957–970.

[26] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren der mathematischen Wissenschaften, 317, Springer, Berlin, Heidelberg, 2009.

[27] W. Römisch, Stability of stochastic programming problems, in *Stochastic Programming*, Handbooks Oper. Res. Management Sci., 10, Elsevier Sci. B. V., Amsterdam, 2003, 483–554.

[28] A. Ruszczyński, Risk-averse dynamic programming for Markov decision processes, *Math. Program.*, 125 (2010), 235–261.

[29] A. Ruszczyński and A. Shapiro, *Stochastic Programming*, Elsevier, 2003.

[30] A. Ruszczyński and A. Shapiro, Conditional risk mappings, *Math. Oper. Res.*, 31 (2006), 544–561.

[31] A. Shapiro, Inference of statistical bounds for multistage stochastic programming problems, *Math. Methods Oper. Res.*, 58 (2003), 57–68.

[32] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, MOS-SIAM Series on Optimization, SIAM, 2014.

[33] T. Wang, A class of dynamic risk measures, *Univ. British Columbia*, 21 (1999).

[34] C. Zhao and Y. Guan, Data-driven risk-averse two-stage stochastic program with $\zeta$-structure probability metrics, 2015. Available from: http://www.optimization-online.org/DB_FILE/2015/07/5014.pdf.

Received February 2019; revised January 2020.

E-mail address: jiangjiecq@163.com
E-mail address: zchen@mail.xjtu.edu.cn
E-mail address: sxytjxyhh@stu.xjtu.edu.cn