SIAM J. CONTROL AND OPTIMIZATION
Vol. 14, No. 3, May 1976

WHEN IS A MULTISTAGE STOCHASTIC PROGRAMMING PROBLEM WELL-DEFINED?*

PAUL OLSEN†

Abstract. Certain measure-theoretic issues raise the possibility that (a) the optimal value of a multistage stochastic programming problem may be ill-defined and (b) the recursion defining the problem may fail, so the problem itself is not even defined. The first difficulty is illustrated by example. A rigorous definition of multistage stochastic programming with fixed linear recourse is shown to avoid this difficulty. In the context of the new definition, certain measurability, convexity, and lower-semicontinuity assumptions on the objective function preclude the second possibility.

1. A reformulation of the problem. Multistage stochastic programming with recourse corresponds to a situation in which information is revealed in stages and a decision is made at each stage based on the information revealed up to and including that stage and on the decisions already made. Number the stages $0, \cdots, K$ $(K < +\infty)$. Assume that the information revealed at stage k $(0 \leqq k \leqq K)$ is fully represented by the realization of a random vector X_k with values in R^{s_k}, where $X_0, \cdots, X_K$ are defined on the same sample space and have known joint distribution; assume that the decision at stage k is represented by a vector $u_k \in R^{n_k}$.

Let $S_k = \sum_{i=0}^{k} s_i$, let $S = S_K$, and let $\underline{X}_k$ denote the random vector $(X_0, \cdots, X_k)$. Let $F_{\underline{X}_k}$ denote the distribution function of $\underline{X}_k$. For each $1 \leqq k \leqq K$, there is a regular conditional distribution function (r.c.d.f.) for X_k given $\underline{X}_{k-1}$—i.e., a function $F_{X_k|\underline{X}_{k-1}} : R^{s_k} \times R^{S_{k-1}} \to [0, 1]$ such that:

(a) for each $\underline{x}_{k-1} \in R^{S_{k-1}}$, $F_{X_k|\underline{X}_{k-1}}(\cdot|\underline{x}_{k-1})$ is a proper distribution function on R^{s_k};

(b) for each $x_k \in R^{s_k}$,

$$F_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1}) = P\{X_k \leqq x_k | \underline{X}_{k-1} = \underline{x}_{k-1}\}$$

for almost every (a.e.) $\underline{x}_{k-1}$.

For convenience, also require that $F_{X_k|\underline{X}_{k-1}}(x_k|\cdot)$ be Borel measurable for each x_k. The existence of such a function follows from the fact that the values of X_k lie in a complete separable metric space—R^{s_k} [1, pp. 263–66].

For $0 \leqq i \leqq K$ and $0 \leqq j \leqq K$, let A_{ij} be a real $m_i \times n_j$ matrix (recall that n_j is the dimensionality of the stage j decision vector, u_j). Require that $A_{ij} = 0$ if $j > i$. Let $N_k = \sum_{j=0}^{k} n_j$, let $N = N_K$, and let $\underline{u}_k = (u_0, \cdots, u_k)$.

For $0 \leqq k \leqq K$, let

$$b_k : R^{S_k} \to R^{m_k}, \qquad c_k : R^{N_k} \times R^{S_k} \to [-\infty, +\infty].$$

The stochastic programming problem is defined recursively. Let $\bar{p}_{K+1}(\underline{u}_K; \underline{x}_K) \equiv 0$. Now let $1 \leqq k \leqq K$ and suppose that $\bar{p}_{k+1} : R^{N_k} \times R^{S_k} \to [-\infty, +\infty]$ has been defined. Let

$$r_k(\underline{u}_k; \underline{x}_k) \equiv c_k(\underline{u}_k, \underline{x}_k) + \bar{p}_{k+1}(\underline{u}_k; \underline{x}_k) + \psi(\underline{u}_k; \underline{x}_k),$$

* Received by the editors July 2, 1974, and in revised form March 6, 1975.
† Institute for Defense Analyses, Arlington, Virginia 22202.

where

$$\psi(\underline{u}_k; \underline{x}_k) = \begin{cases} 0, & \sum_{j=0}^{k} A_{kj} u_j = b_k(\underline{x}_k),\ u_k \geqq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

The convention $+\infty + (-\infty) = (-\infty) + (+\infty) = +\infty$ applies throughout.

Define the parameterized problem

$$P_k(\underline{u}_{k-1}; \underline{x}_k) : \underset{v \in R^{n_k}}{\text{minimize}}\ r_k(\underline{u}_{k-1}, v; \underline{x}_k).$$

Let $p_k(\underline{u}_{k-1}; \underline{x}_k) \equiv \inf(P_k(\underline{u}_{k-1}; \underline{x}_k))$, the problem's optimal value. Let

(1) $$\bar{p}_k(\underline{u}_{k-1}; \underline{x}_{k-1}) \equiv \int p_k(\underline{u}_{k-1}; \underline{x}_k)\, dF_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1}).^1$$

That completes the induction step. So that the integral is defined whenever the integrand is measurable, define (as in [9])

$$\int f\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu$$

for any measure μ and measurable, extended-real-valued function f, with the convention $+\infty - (+\infty) = -\infty + (+\infty) = +\infty$.

For stage 0 define

$$P_0(x_0): \text{minimize } c_0(u_0; x_0) + \bar{p}_1(u_0; x_0)$$
$$\text{subject to } A_{00} u_0 = b_0(x_0),$$
$$u_0 \geqq 0.$$

The preceding formulation of multistage stochastic programming with (fixed linear) recourse resembles that of Dantzig [3], who introduced the problem in [2]. Wets [10] presents results for a special case of the problem in which X_k is independent of $\underline{X}_{k-1}$ for each $1 \leqq k \leqq K$, but he also formulates the problem in the absence of stagewise independence. The Dantzig and Wets formulations differ from the one given here essentially in replacing the definition (1) above with

(1′) $$\bar{p}_k(\underline{u}_{k-1}; \underline{x}_{k-1}) \equiv E[p_k(\underline{u}_{k-1}; \underline{X}_k)|\underline{X}_{k-1} = \underline{x}_{k-1}].$$

Section 2 shows that $\bar{p}_k(\underline{u}_{k-1}; \underline{x}_{k-1})$ is indeed a version of the above conditional expectation, but Example 2.3 reveals that using (1′) instead of (1)—i.e., defining $\bar{p}_k(\underline{u}_{k-1}; \underline{X}_{k-1})$ to be any version of $E[p_k(\underline{u}_{k-1}; \underline{X}_k)|\underline{X}_{k-1}]$—can result in $E[\bar{P}_0(X_0)]$ (the optimal value of the stochastic programming problem) being ill-defined. According to Theorem 2.1, this cannot happen with the formulation presented here.

A separate issue is whether the function $p_k(\underline{u}_{k-1}; \cdot)$ has the measurability property that permits the integration in equation (1). (The same property must

[1] A more precise way of writing the integral is

$$\int p_k(\underline{u}_{k-1}; \underline{x}_{k-1}, y)\, dF_{X_k|\underline{X}_{k-1}}(y|\underline{x}_{k-1}).$$

The abbreviated notation will be used often.

hold or the conditional expectation in (1') is undefined.) Section 3 presents conditions under which the integrand in (1) has the requisite measurability property at each stage; it relies upon Theorem 2.2 (below) and Rockafellar's theory of normal convex integrands [6].

2. Uniqueness of the problem's optimal value. For now, measurability is assumed while attention focuses on the issue of whether P_0 is well-defined. Since the r.c.d.f. in (1) is not necessarily unique (there may be more than one function having properties (a) and (b) above), the definition of $\bar{p}_k$ in terms of an integral over "the" r.c.d.f. for X_k given $\underline{X}_{k-1}$ seems to raise the possibility that $\bar{p}_1$ can be essentially altered simply by using a different r.c.d.f. at each stage. In fact, that is not the case.

Suppose that for $1 \leqq k \leqq K$, $F_{X_k|\underline{X}_{k-1}}$ and $G_{X_k|\underline{X}_{k-1}}$ are both regular conditional distribution functions for X_k given $\underline{X}_{k-1}$. Using $F_{X_k|\underline{X}_{k-1}}$ at stage k for each k generates a sequence of problems $P'_K, \cdots, P'_0$ via the recursion

$$\bar{p}'_k(z ; \underline{x}_{k-1}) = \int p'_k(z ; \underline{x}_k)\, dF_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1}).$$

The functions $G_{X_k|\underline{X}_{k-1}}$, $1 \leqq k \leqq K$, generate a parallel sequence $P''_K, \cdots, P''_0$. (The data b_k and c_k are the same for P'_k and P''_k.)

THEOREM 2.1. *For each* $1 \leqq k \leqq K$, $\bar{p}'_k(\cdot ; \underline{x}_k) = \bar{p}''_k(\cdot ; \underline{x}_k)$ *for almost every* $\underline{x}_k$.[2] *Thus* P'_0 *is essentially the same as* P''_0.

Proof. Let $1 \leqq k \leqq K$. Assume that

$$\bar{p}'_{k+1}(\cdot ; \underline{x}_k) = \bar{p}''_{k+1}(\cdot ; \underline{x}_k) \quad \text{for a.e. } \underline{x}_k.$$

Then

$$p'_k(\cdot ; \underline{x}_k) = p''_k(\cdot ; \underline{x}_k) \quad \forall \underline{x}_k \in T,$$

where $P\{\underline{X}_k \in T\} = 1$. By the product measure theorem [1, p. 97],

$$\int \chi_T(\underline{x}_{k-1}, x_k)\, dG_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1}) = 1 \quad \forall \underline{x}_{k-1} \in \mathscr{S},$$

where $P\{\underline{X}_{k-1} \in \mathscr{S}\} = 1$.

Property (b) of the r.c.d.f. implies that for each fixed x_k,

$$F_{X_k|\underline{X}_{k-1}}(x_k|\cdot) = G_{X_k|\underline{X}_{k-1}}(x_k|\cdot) \quad \text{a.e.}$$

Since R^{N_k} is separable and distribution functions are right-continuous, there is a set $\mathscr{S}'$ of measure 1 such that if $\underline{x}_{k-1} \in \mathscr{S}'$,

$$F_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1}) = G_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1}) \quad \forall x_k.$$

If $\underline{x}_{k-1} \in \mathscr{S} \cap \mathscr{S}'$,

$$\int p'_k(z ; \underline{x}_k)\, dF_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1}) = \int p'_k(z ; \underline{x}_k)\, dG_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1})$$

$$= \int p''_k(z ; \underline{x}_k)\, dG_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1})$$

for every $z \in R^{N_{k-1}}$. That completes the induction step. The induction hypothesis is trivially true if $k = K$. □

[2] If $f: R^n \to R$ and $g: R^n \to R$, $f = g$ means $f(x) = g(x) \forall x \in R^n$.

THEOREM 2.2. *Let* $g: R^n \times R^{\bar{N}_k} \to [-\infty, +\infty]$ *be Borel measurable. Let* $F_{X_k|\underline{X}_{k-1}}$ *be any regular conditional distribution function for* X_k *given* $\underline{X}_{k-1}$ *such that* $F_{X_k|\underline{X}_{k-1}}(x_k|\cdot)$ *is Borel measurable for each fixed* x_k. *Let*

$$h(u, \underline{x}_{k-1}) = \int g(u, \underline{x}_{k-1}, y)\, dF_{X_k|\underline{X}_{k-1}}(y|\underline{x}_{k-1})$$

for every $u \in R^n$ *and every* $\underline{x}_{k-1}$. *Then h is Borel measurable, and*

$$\int h(u, \underline{x}_{k-1})\, dF_{\underline{X}_{k-1}}(\underline{x}_{k-1}) \leqq \int g(u, \underline{x}_k)\, dF_{\underline{X}_k}(\underline{x}_k) \tag{2}$$

for each $u \in R^n$. *If, for a given u,*

$$\int g^+(u, \underline{x}_k)\, dF_{\underline{X}_k}(\underline{x}_k) < +\infty \text{ or } \int g^-(u, \underline{x}_k)\, dF_{\underline{X}_k}(\underline{x}_k) < +\infty,$$

then (2) *holds as an equality, and*

$$h(u, \underline{x}_{k-1}) = E[g(u, \underline{X}_k)|\underline{X}_{k-1} = \underline{x}_{k-1}]$$

for a.e. $\underline{x}_{k-1}$.

The theorem extends Fubini's theorem [1, p. 101] to agree with the extended definition of integration given after equation (1). A proof appears in [4].

The first part of the theorem is used below to demonstrate measurability of $\bar{p}_k$ (whose role is played by h in the theorem) while the second part justifies the usual economic interpretation of the stochastic programming problem. The second part implies that if $Ep^-_{k+1}(\underline{u}_k ; \underline{X}_{k+1}) < +\infty$—for which the Appendix's Proposition A.1 gives a verifiable sufficient condition—then $E[p_{k+1}(\underline{u}_k ; \underline{X}_{k+1})|\underline{X}_k = \underline{x}_k]$ exists and equals $\bar{p}_{k+1}(\underline{u}_k ; \underline{x}_k)$ for almost every $\underline{x}_k$. Thus, $\bar{p}_{k+1}(\underline{u}_k ; \underline{x}_k)$ is the expected (minimum) cost of operations in stages $k+1$ through K given past decisions $u_0, \cdots, u_k$ and states of nature $x_0, \cdots, x_k$. The problem $P_k(\underline{u}_{k-1} ; \underline{x}_k)$ involves trying to choose u_k to minimize the sum of current costs, $c_k(\underline{u}_k ; \underline{x}_k)$, and expected future costs, $\bar{p}_{k+1}(\underline{u}_k ; \underline{x}_k)$. To be precise, at stage k the decision-maker observes x_k; knowing $\underline{x}_k$ and his past decisions, $\underline{u}_{k-1}$, he seeks a stage k decision u_k that minimizes current costs plus expected future costs subject to the constraints

$$A_{kk}u_k = b_k(\underline{x}_k) - \sum_{j=0}^{k-1} A_{kj}u_j$$

$$u_k \geqq 0.$$

Theorem 2.2 implies $\bar{p}_k(\underline{u}_{k-1} ; \underline{X}_{k-1})$ is a version of $E[p_k(\underline{u}_{k-1} ; \underline{X}_k)|\underline{X}_{k-1}]$. It does not follow that one could simply take $\bar{p}_k(\underline{u}_{k-1} ; \cdot)$ to be any version of the conditional expectation and still conclude (as in Theorem 2.1) that P_0 is well-defined. The example below illustrates how choosing a different version of the conditional expectation could make P_0 essentially different.

Example 2.3. Let $K = 2$, let $X_0 = 0$ almost surely, and let X_1 and X_2 be uniformly distributed on [0, 1]. Let

$$c_2(\underline{u}_2 ; \underline{x}_2) \equiv 0, \quad c_1(\underline{u}_1 ; \underline{x}_1) \equiv x_1 u_1, \quad c_0(u_0 ; 0) \equiv 0.$$

Let the constraints be

$$\begin{aligned} u_0 \quad &\leqq 1, \\ u_1 \quad &\leqq 1, \\ u_2 &\leqq 1, \\ u_0, u_1, u_2 &\geqq 0. \end{aligned}$$

Clearly, $\bar{p}_2$ is everywhere 0 or $+\infty$. Redefine it for each $\underline{u}_1$ on a set of measure 0:

$$\bar{p}_2'(\underline{u}_1; \underline{x}_1) = \begin{cases} 0, & u_1 \geqq 0, x_1 \neq u_1, \\ -1, & u_1 \geqq 0, x_1 = u_1, \\ +\infty, & \text{otherwise.} \end{cases}$$

For any fixed $\underline{u}_1$, $\bar{p}_2'(u_1; \underline{x}_1) = \bar{p}_2(\underline{u}_1; \underline{x}_1)$ for a.e. $\underline{x}_1$, so $\bar{p}_2(\underline{u}_1; \underline{X}_1)$ and $\bar{p}_2'(\underline{u}_1; \underline{X}_1)$ are two different versions of $E[p_2(\underline{u}_1; \underline{X}_2)|\underline{X}_1]$. Notice, however, that there is no regular conditional distribution function $F'_{\underline{X}_2|\underline{X}_1}$ such that

$$\bar{p}_2'(\underline{u}_1; \underline{x}_1) = \int p_2(\underline{u}_1; \underline{x}_2)\, dF'_{\underline{X}_2|\underline{X}_1}(\underline{x}_2|\underline{x}_1)$$

for every $\underline{u}_1$ and $\underline{x}_1$.

Use of $\bar{p}_2'$ in place of $\bar{p}_2$ in the stage 1 objective function results in the sequence of problems

$$P_1'(u_0; \underline{x}_1)\text{: minimize } x_1 u_1 - \chi_{\{x_1\}}(u_1)$$

$$\text{subject to } u_1 \leqq 1$$

$$u_0, u_1 \geqq 0$$

$$P_0'(x_0)\text{: minimize } -\tfrac{2}{3}$$

$$\text{subject to } 0 \leqq u_0 \leqq 1.$$

Thus $\inf(P_0'(0)) = -\frac{2}{3} < 0 = \inf(P_0(0))$.

Although $\bar{p}_2'(\underline{u}_1; \cdot)$ was created by redefining $\bar{p}_2(\underline{u}_1; \cdot)$ on a set of measure 0 for each fixed $\underline{u}_1$, the set of every $\underline{x}_1$ such that $\bar{p}_2'(\underline{u}_1; \underline{x}_1) \neq \bar{p}_2(\underline{u}_1; \underline{x}_1)$ for *some* $\underline{u}_1$ has positive measure; that is the source of the discrepancy. In general, $p_k(\cdot; \underline{x}_k)$ depends (at least potentially) on the whole of the function $\bar{p}_{k+1}(\cdot; \underline{x}_k)$. Altering $\bar{p}_{k+1}(\cdot; \underline{x}_k)$ on a set of measure 0 alters $p_k(\cdot; \underline{x}_k)$ on a set of measure 0 and therefore leaves $\bar{p}_1$ essentially unchanged. Altering the function on a set of positive measure can essentially alter P_0 and destroy the original problem.

3. Measurability of each stage's objective function. Let $\mu_k(\cdot|\underline{x}_{k-1})$ denote the Lebesgue–Stieltjes measure determined by $F_{\underline{X}_k|\underline{X}_{k-1}}(\cdot|\underline{x}_{k-1})$. The recursive definition of the stochastic programming problem tacitly assumes that at each stage, for a.e. $\underline{x}_{k-1}$, $p_k(\underline{u}_{k-1}; \underline{x}_{k-1}, \cdot)$ is $\mathscr{T}$-measurable for every $\underline{u}_{k-1}$, where $\mathscr{T}$ is the completion of the Borel sets in R^{s_k} with respect to $\mu_k(\cdot|\underline{x}_{k-1})$; if that is not the case, the function $\bar{p}_k(\cdot; \underline{x}_{k-1})$ is undefined on a set of $\underline{x}_{k-1}$'s of positive measure, and the recursion cannot continue.

DEFINITION 3.1. $\bar{p}_{K+1}$ is said to be essentially defined. Let $1 \leqq k \leqq K$. One says that $\bar{p}_k$ is *essentially defined* if and only if: (i) $\bar{p}_{k+1}$ is essentially defined; and (ii) for a.e. $\underline{x}_{k-1}$, $p_k(\underline{u}_{k-1}; \underline{x}_{k-1}, \cdot)$ is measurable with respect to the completion of the σ-algebra of Borel sets in R^{s_k} under the measure $\mu_k(\cdot|\underline{x}_{k-1})$ for each $\underline{u}_{k-1}$.

PROPOSITION 3.2. *Let* $1 \leqq k \leqq K$. *Assume*:

(a) $\bar{p}_{k+1}$ *is essentially defined, and there is a Borel measurable function* $\bar{p}^0_{k+1}$ *such that* $\bar{p}_{k+1}(\cdot; \underline{x}_k) = \bar{p}^0_{k+1}(\cdot; \underline{x}_k)$ *for a.e.* $\underline{x}_k$.

(b) $\bar{p}^0_{k+1}(\cdot; \underline{x}_k)$ *is convex for a.e.* $\underline{x}_k$.

(c) c_k *and* b_k *are Borel measurable.*

(d) $c_k(\cdot; \underline{x}_k)$ *is convex for a.e.* $\underline{x}_k$.

(e) *there is a countable collection* V *of Borel measurable functions* $v: R^{N_{k-1}} \times R^{S_k} \to R^{n_k}$ *such that for a.e.* $\underline{x}_k$, $V(\underline{u}_{k-1}, \underline{x}_k) \cap \operatorname{dom} r_k(\underline{u}_{k-1}, \cdot; \underline{x}_k)$ *is dense in* $\operatorname{dom} r_k(\underline{u}_{k-1}, \cdot; \underline{x}_k)$ *for every* $\underline{u}_{k-1}$, *where*

$$V(\underline{u}_{k-1}, \underline{x}_k) = \{v(\underline{u}_{k-1}, \underline{x}_k) : v \in V\}.[3]$$

Then $\bar{p}_k$ *is essentially defined, and there is a Borel measurable function* $\bar{p}^0_k$ *such that: (i)* $\bar{p}^0_k(\cdot; \underline{x}_{k-1}) = \bar{p}_k(\cdot; \underline{x}_{k-1})$ *for a.e.* $\underline{x}_{k-1}$; *and (ii)* $\bar{p}^0_k(\cdot; \underline{x}_{k-1})$ *is convex for a.e.* $\underline{x}_{k-1}$.

Proof. Let S be a Borel set of measure 1 such that $\underline{x}_k \in S$ implies:

(i) $c_k(\cdot; \underline{x}_k)$ and $\bar{p}^0_{k+1}(\cdot; \underline{x}_k)$ are convex;

(ii) $V(\underline{u}_{k-1}, \underline{x}_k) \cap \operatorname{dom} r_k(\underline{u}_{k-1}, \cdot; \underline{x}_k)$ is dense in $\operatorname{dom} r_k(\underline{u}_{k-1}, \cdot; \underline{x}_k)$ for every $\underline{u}_{k-1}$;

(iii) $\bar{p}_{k+1}(\cdot; \underline{x}_k) = \bar{p}^0_{k+1}(\cdot; \underline{x}_k)$.

Redefine $\bar{p}^0_{k+1}(\cdot; \underline{x}_k)$ outside S to be identically $+\infty$. Let P^0_k be the (parameterized) problem formed from P_k by using $\bar{p}^0_{k+1}$ in place of $\bar{p}_{k+1}$. Denote its return function by r^0_k and its perturbation function by p^0_k. By [1, 1.5.6], r^0_k is Borel measurable.[4]

The first step in the proof is to show that

$$p^0_k(z; \underline{x}_k) \equiv \inf\{r^0_k(z, v(z, \underline{x}_k); \underline{x}_k) : v \in V\}. \tag{3}$$

Let $z \in R^{N_{k-1}}$, $\underline{x}_k \in R^{S_k}$.

If $p^0_k(z; \underline{x}_k) = +\infty$, $r^0_k(z, u_k; \underline{x}_k) = +\infty$ for every $u_k \in R^{n_k}$, and (3) holds trivially.

Alternatively, suppose $p^0_k(z; \underline{x}_k) < +\infty$.

Choose a sequence $\{u^n_k\}$ such that

$$r^0_k(z, u^n_k; \underline{x}_k) < +\infty \quad \text{for every } n$$

and

$$r^0_k(z, u^n_k; \underline{x}_k) \searrow p^0_k(z; \underline{x}_k).$$

Choose $\{\alpha_n\} \subset R$ such that $r^0_k(z, u^n_k; \underline{x}_k) < \alpha_n$ for every n and $\alpha_n \searrow p^0_k(z; \underline{x}_k)$.

For any given n, suppose that u^n_k lies in the relative interior of $\operatorname{dom} r^0_k(z, \cdot; \underline{x}_k) \triangleq \{u_k : r^0_k(z, u_k; x_k) < +\infty\}$.[5] Then by [7, Thm. 10.1] and property (ii), there is a function $v \in V$ such that

$$r^0_k(z, v(z, \underline{x}_k); \underline{x}_k) < \alpha_n + (1/n).$$

---

[3] Cf. the definition of "normal convex integrand" in [6].

[4] Note that Ash does not use the term "Borel measurable" in the same sense as this paper, where it means "measurable with respect to the σ-algebra of Borel sets." Also, the theorem in [1] must be extended slightly since $r_k(\underline{u}_k; \underline{x}_k)$ may involve a sum of the form $+\infty + (-\infty)$, which Ash does not define.

[5] See [7] for the definitions of "relative interior" and "relative boundary".

Alternatively, if u_k^n lies in the relative boundary of dom $r_k^0(z,\cdot\,;\underline{x}_k)$, Corollary 7.3.1 of [7] reveals the existence of a point w in the relative interior of dom $r_k^0(z,\cdot\,;\underline{x}_k)$ such that $r_k^0(z,w;\underline{x}_k)<\alpha_n$, and the preceding argument can be applied to w.

Thus, in any case, there is a sequence $\{v^n\}\subset V$ such that

$$r_k^0(z,v^n(z,\underline{x}_k);\underline{x}_k)\to p_k^0(z;\underline{x}_k).$$

That demonstrates (3).

Since r_k^0 and v are Borel measurable, $r_k^0(z,v(z,\underline{x}_k);\underline{x}_k)$ is a Borel measurable function of $(z,\underline{x}_k)$ for each $v\in V$. As the infimum of a countable collection of Borel measurable functions, p_k^0 is Borel measurable. Then by Theorem 2.2,

$$\bar{p}_k^0(\underline{u}_{k-1};\underline{x}_{k-1})\triangleq\int p_k^0(\underline{u}_{k-1};\underline{x}_k)\,dF_{X_k|\underline{X}_{k-1}}(x_k|\underline{x}_{k-1})$$

is defined for each $\underline{u}_{k-1}$ and $\underline{x}_{k-1}$, and $\bar{p}_k^0$ is Borel measurable. By Proposition A.2 (Appendix), $\bar{p}_k^0(\cdot\,;\underline{x}_{k-1})$ is convex for every $\underline{x}_{k-1}$. Since $p_k(\cdot\,;\underline{x}_k)=p_k^0(\cdot\,;\underline{x}_k)$ for a.e. $\underline{x}_k$, Theorem 2.2 reveals that for a.e. $\underline{x}_{k-1}$, there is a Borel set B in R^{s_k} such that $\mu_k(B|\underline{x}_{k-1})=1$ and $p_k(\cdot\,;\underline{x}_{k-1},y)=p_k^0(\cdot\,;\underline{x}_{k-1},y)$ if $y\in B$. Consequently, $\bar{p}_k$ is essentially defined. □

The preceding proposition's assumption (e) is appealing but not directly verifiable. The next proposition gives a sufficient condition for it. The hypothesis of Proposition 3.3 is the conclusion of Theorems 5.5, 5.9, and 5.12 of [5]. Consequently, one can merge the hypothesis of any of those theorems with assumptions (c) and (d) of Proposition 3.2 and prove inductively that $\bar{p}_1$ is essentially defined.

PROPOSITION 3.3. *Let $1\leqq k\leqq K$. Suppose there are matrices D_1 and D_2 and a Borel measurable function d: $R^{S_k}\to R^m$ such that*

$$\text{dom } r_k(\cdot\,;\underline{x}_k)=\{\underline{u}_k:D_1\underline{u}_{k-1}+D_2u_k\geqq d(\underline{x}_k),\ \underline{u}_k\geqq 0\}$$

for a.e. $\underline{x}_k$. Then assumption (e) of Proposition 3.2 holds.[6]

Proof. Let $W=[D_2\,|-I]$. There are square matrices $M_1,\cdots,M_m$ such that, for any given $a\in R^m$, w is an extreme point of $\{w:Ww=a,\ w\geqq 0\}$ if and only if $w\geqq 0$ and $w=M_ia$ for some i. Let $a_1,\cdots,a_s$ be the extreme directions of $\{w:Ww=0,\ w\geqq 0\}$. Now redefine $M_1,\cdots,M_m$ by deleting all but the first n_k rows of each matrix, and redefine $a_1,\cdots,a_s$ by deleting all but the first n_k components of each vector. For any $z\in R^{N_{k-1}}$ and any $\underline{x}_k$,

$$\text{dom } r_k(z,\cdot\,;\underline{x}_k)=\{y:y=\sum_{i=1}^m\lambda_iM_i(d(\underline{x}_k)-D_1z)+\sum_{i=1}^s\gamma_ia_i,$$

$$\sum_{i=1}^m\lambda_i=1,\ \lambda_i\geqq 0\,\forall i,\ \gamma_i\geqq 0\,\forall i,\ y\geqq 0\}.$$

Let V be the collection of functions v such that

$$v(z,\underline{x}_k)\equiv\sum_{i=1}^m\lambda_iM_i(d(\underline{x}_k)-D_1z)+\sum_{i=1}^s\lambda_ia_i,$$

where $\sum_{i=1}^m\lambda_i=1$, $\lambda_i\geqq 0\,\forall i$, $\gamma_i\geqq 0\,\forall i$, and λ_i and γ_i are rational numbers. □

[6] Actually, the linear mappings D_1 and D_2 could be replaced by a Borel measurable mapping and a continuous mapping, respectively. See the proof of Proposition 4.6 in [4].

THEOREM 3.4 *Assume for each $1\leqq k\leqq K$:*

(a) *c_k and b_k are Borel measurable;*

(b) *$c_k(\cdot\,;\underline{x}_k)$ is l.s.c. and convex for a.e. $\underline{x}_k$;*

(c) *There are functions β_k and α_k in $L_1(R^{S_k})$ with values in $[-\infty,0]$ such that*

$$c_k(\underline{u}_k;\underline{x}_k)\geqq\beta_k(\underline{x}_k)\|\underline{u}_k\|_2+\alpha_k(\underline{x}_k)$$

for every u_k and every x_k, and

$$E|\beta_k(\underline{X}_k)b_i(\underline{X}_i)|<+\infty\quad\text{if } 1\leqq i\leqq k;^7$$

(d) $\{w: A_{kk}w=0,\ w\geqq 0\}=\{0\}$.

Then $\bar{p}_1$ is essentially defined, and for each $1\leqq k\leqq K$, there is a Borel measurable function $\bar{p}_k^0$ such that $\bar{p}_k(\cdot\,;\underline{x}_{k-1})=\bar{p}_k^0(\cdot\,;\underline{x}_{k-1})$ for a.e. $\underline{x}_{k-1}$ and $\bar{p}_k^0(\cdot\,;\underline{x}_{k-1})$ is l.s.c. and convex for a.e. $\underline{x}_{k-1}$.

The theorem's proof uses the following lemma, which extends [6, Thm. 5].

LEMMA 3.5. *Let f: $R^n\times R^m\to[-\infty,+\infty]$ be Borel measurable. Let μ be a Borel probability measure on R^m. Assume that $f(\cdot\,;x)$ is l.s.c. and convex for a.e. $x\in R^m$. Then there is a countable collection U of Borel measurable functions u: $R^m\to R^n$ such that $U(x)\cap$dom $f(\cdot\,;x)$ is dense in dom $f(\cdot\,;x)$ for a.e. x, where*

$$U(x)=\{u(x):u\in U\}.$$

Proof. Redefine $f(\cdot\,;x)$ to be identically $+\infty$ on the set of measure 0 where it is not lower-semicontinuous and convex.

Define a multifunction $K_1: R^m\to R^n$ by

$$K_1(x)\equiv\{u\in R^n: f(\cdot\,;x)\in R\}.$$

Since f is Borel measurable, the graph of K_1 is a Borel set, and therefore by [6, Thm. 2], K_1 is measurable with respect to $\mathcal{B}(R^m)^*$, the completion of the σ-algebra of Borel sets under the measure μ. Consequently,

$$\begin{aligned}S_1&\triangleq\{x: f(u;x)\in R\text{ for some }u\}\\&=\{x:K_1(x)\neq\varnothing\}\\&=K_1^{-1}(R^n)\in\mathcal{B}(R^m)^*.\end{aligned}$$

Define a multifunction K_2 by

$$K_2(x)\equiv\{u\in R^n: f(u;x)=-\infty\}.$$

An argument parallel to the preceding one reveals that

$$\begin{aligned}S_2&\triangleq\{x: f(u;x)=-\infty\text{ for some }u\}\\&=\{x:K_2(x)\neq\varnothing\}\in\mathcal{B}(R^m)^*.\end{aligned}$$

Let

$$g(u;x)=\begin{cases}0, & x\in S_2\text{ and }f(u;x)<+\infty,\\ f(u;x) & \text{otherwise.}\end{cases}$$

[7] The notation signifies that $\underline{X}_i$ is a subvector of $\underline{X}_k$.

Since $(R^n \times S_2) \cap \{(u, x): f(u; x) < +\infty\} \in \mathscr{B}(R^m)^*$, g is $\mathscr{B}(R^m)^*$-measurable. If $x \in S_2$, $f(\cdot\,; x)$ is identically $-\infty$ on its effective domain; hence, for each $x \in R^m$, $g(\cdot\,; x)$ is a lower-semicontinuous convex function with the same effective domain as $f(\cdot\,; x)$.

Let $T = S_1 \cup S_2$. By [6, Thm. 5], there is a countable collection U of $\mathscr{B}(R^m)^*$-measurable functions $u: T \to R^n$ such that $U(x) \cap \operatorname{dom} g(\cdot\,; x)$ is dense in $\operatorname{dom} g(\cdot\,; x)$ for every $x \in T$. Extend each function in U to a function on R^m by making it identically 0 outside T. If $u \in U$, there is a Borel measurable function $\hat{u}: R^m \to R^n$ such that $u = \hat{u}$ a.e. (μ) [8, p. 145]; let $\hat{U}$ be the collection consisting of one such $\hat{u}$ for each $u \in U$. Then $\hat{U}(x) = U(x)$ for a.e. $x \in R^m$ since U has countably many elements. Hence, there is a Borel subset T^0 of T such that $\mu(T^0) = \mu(T)$ and $\hat{U}(x) \cap \operatorname{dom} g(\cdot\,; x)$ is dense in $\operatorname{dom} g(\cdot\,; x)$ for each $x \in T^0$. Since $\operatorname{dom} f(\cdot\,; x) = \operatorname{dom} g(\cdot\,; x)$ for any $x \in R^m$ and $\operatorname{dom} f(\cdot\,; x) = \varnothing$ if $x \notin T$, the collection $\hat{U}$ has the desired property. □

Proof of Theorem 3.4. Let $1 \leqq k \leqq K$. Assume that $\bar{p}_{k+1}$ is essentially defined and there is a Borel measurable function $\bar{p}^0_{k+1}$ such that $\bar{p}_{k+1}(\cdot\,; \underline{x}_k) = \bar{p}^0_{k+1}(\cdot\,; \underline{x}_k)$ for a.e. $\underline{x}_k$ and $\bar{p}^0_{k+1}(\cdot\,; \underline{x}_k)$ is l.s.c. and convex for a.e. $\underline{x}_k$. Also assume that

$$\bar{p}^0_{k+1}(\underline{u}_k; \underline{x}_k) \geqq \delta(\underline{x}_k)\|\underline{u}_k\|_2 + \gamma(\underline{x}_k)$$

for every $\underline{u}_k$ and every $\underline{x}_k$, where δ and γ have the nonpositivity, mean, and covariance properties of β_k and α_k, respectively, in (c).

Let P^0_k be the problem formed from P_k by replacing $\bar{p}_{k+1}$ with $\bar{p}^0_{k+1}$. Proposition A.1 (Appendix) implies that $p^0_k(\cdot\,; \underline{x}_k) > -\infty$ for a.e. $\underline{x}_k$. Then by Proposition A.3, $p^0_k(\cdot\,; \underline{x}_k)$ is l.s.c. and convex for a.e. $\underline{x}_k$. Since $\bar{p}^0_{k+1}(\cdot\,; \underline{x}_k) > -\infty$ and $c_k(\cdot\,; \underline{x}_k) > -\infty$ for a.e. $\underline{x}_k$, $r^0_k(\cdot\,; \underline{x}_k)$ is l.s.c. and convex for a.e. $\underline{x}_k$ by [7, Thm. 9.3]. As a finite sum of Borel measurable functions, it is Borel measurable. Lemma 3.5 reveals the existence of a countable collection U of Borel measurable functions $u: R^{S_k} \to R^{N_k}$ such that $U(\underline{x}_k) \cap \operatorname{dom} r^0_k(\cdot\,; \underline{x}_k)$ is dense in $\operatorname{dom} r^0_k(\cdot\,; \underline{x}_k)$ for every $\underline{x}_k \in S$, where S is a Borel set of measure 1. Let T be a Borel subset of S of measure 1 such that $\underline{x}_k \in T$ implies $r^0_k(\cdot\,; \underline{x}_k)$ and $p^0_k(\cdot\,; \underline{x}_k)$ are l.s.c. and convex. Redefine r^0_k and p^0_k to be identically $+\infty$ outside T. Then for any $z \in R^{N_{k-1}}$ and $\underline{x}_k \in R^{S_k}$,

$$p^0_k(z; \underline{x}_k) = \liminf_{z' \to z} \left(\inf_{u_k} r^0_k(z, u_k; \underline{x}_k)\right)$$

$$= \lim_{n \to \infty} \inf_{\underline{u}_k} \{r^0_k(\underline{u}_k; \underline{x}_k) : \|\underline{u}_{k-1} - z\|_2 < 1/n\}$$

$$= \lim_{n \to \infty} \inf \{r^0_k(\underline{u}_k(\underline{x}_k); \underline{x}_k) : \|\underline{u}_{k-1}(\underline{x}_k) - z\|_2 < 1/n,\ \underline{u}_k \in U\}.$$

It follows that p^0_k is Borel measurable. An argument identical to one in the proof of Proposition 3.2 shows that $\bar{p}_k$ is essentially defined and $\bar{p}_k(\cdot\,; \underline{x}_{k-1}) = \bar{p}^0_k(\cdot\,; \underline{x}_{k-1})$ for a.e. $\underline{x}_{k-1}$.

By Proposition A.1, there functions $\hat{\delta}$ and $\hat{\gamma}$ in $L_1(R^{S_{k-1}})$ with values in $[-\infty, 0]$ such that

$$\bar{p}^0_k(\underline{u}_{k-1}; \underline{x}_{k-1}) \geqq \hat{\delta}(\underline{x}_{k-1})\|\underline{u}_{k-1}\|_2 + \hat{\gamma}(\underline{x}_{k-1})$$

for every $\underline{u}_{k-1}$ and every $\underline{x}_{k-1}$; also, $E|\hat{\delta}((\underline{X}_{k-1})b_i(\underline{X}_i)| < +\infty$ if $1 \leqq i \leqq k$[8]. Then by Proposition A.4, $\bar{p}^0_k(\cdot\,; \underline{x}_{k-1})$ is l.s.c. and convex for a.e. $\underline{x}_{k-1}$.

That completes the induction step. The induction hypothesis is obviously true if $k = K$. □

Appendix. These propositions are restatements of some propositions in [5], whose numbers are displayed in parentheses.

PROPOSITION A.1 (2.2). *Assume for each* $0 \leqq k \leqq K$,

(a) *For some selected functions* β_k *and* α_k *in* $L_1(R^{S_k})$ *with values in* $[-\infty, 0]$,

$$c_k(\underline{u}_k; \underline{x}_k) \geqq \beta_k(\underline{x}_k)\|\underline{u}_k\|_2 + \alpha_k(\underline{x}_k)$$

for every $\underline{u}_k$ *and every* $\underline{x}_k$.

(b) $E|\beta_k(\underline{X}_k)b_i(\underline{X}_i)| < +\infty$ *if* $0 \leqq i \leqq k$. ($\underline{X}_i$ *is a subvector of* $\underline{X}_k$.)

(c) $\{w: A_{kk}w = 0,\ w \geqq 0\} = \{0\}$.

Then, for each $0 \leqq k \leqq K$, $E\bar{p}_k(\underline{u}_{k-1}; \underline{X}_k) < +\infty$, *and there are functions* δ_k *and* γ_k *in* $L_1(R^{S_k})$, *with values in* $[-\infty, 0]$, *such that*

$$\bar{p}_{k+1}(\underline{u}_k; \underline{x}_k) \geqq \delta_k(\underline{x}_k)\|\underline{u}_k\|_2 + \gamma_k(\underline{x}_k)$$

for every $\underline{u}_k$ *and every* $\underline{x}_k$.

PROPOSITION A.2 (3.2, 3.3). *Let* $1 \leqq k \leqq K$. *If* $c_k(\cdot\,; \underline{x}_k)$ *and* $\bar{p}_{k+1}(\cdot\,; \underline{x}_k)$ *are convex for every* $\underline{x}_k$, $\bar{p}_k(\cdot\,; \underline{x}_{k-1})$ *is convex for every* $\underline{x}_{k-1}$.

PROPOSITION A.3 (3.4). *Let* $1 \leqq k \leqq K$. *Assume that* $\{w: A_{kk}w = 0, w \geqq 0\} = \{0\}$. *Fix* $\underline{x}_k$. *If* $c_k(\cdot\,; \underline{x}_k)$ *and* $\bar{p}_{k+1}(\cdot\, \underline{x}_k)$ *are l.s.c. convex functions and* $p_k(\cdot\,; \underline{x}_k) > -\infty$, *then* $p_k(\cdot\,; \underline{x}_k)$ *is l.s.c. and convex.*

PROPOSITION A.4 (3.5). *Let* $1 \leqq k \leqq K$. *Fix* $\underline{x}_{k-1}$. *If* $p_k(\cdot\,; \underline{x}_{k-1}, \hat{x}_k)$ *is l.s.c. and convex for every* $\hat{x}_k$ *and if* $\bar{p}_k(\cdot\,; \underline{x}_{k-1}) > -\infty$, *then* $\bar{p}_k(\cdot\,; \underline{x}_{k-1})$ *is l.s.c. and convex.*

REFERENCES

[1] R. B. ASH (1972), *Real Analysis and Probability*, Academic Press, New York.

[2] G. B. DANTZIG (1955), *Linear programming under uncertainty*, Management Sci., 1, pp. 197–206.

[3] ——— (1963), *Linear Programming and Extensions*, Princeton University Press, Princeton, N. J.

[4] P. OLSEN (1973), *Measurability in stochastic programming*, Tech. Rep. 196, Dept. of Operations Research, Cornell Univ., Ithaca, N. Y.

[5] ——— (1974), *Multistage stochastic programming with recourse: The equivalent deterministic problem*, this Journal, 14 (1976), pp. 495–517.

[6] R. T. ROCKAFELLAR (1969), *Measurable dependence of convex sets and functions on parameters*, J. Math. Anal. Appl., 28, pp. 4–25.

[7] ——— (1970), *Convex Analysis*, Princeton University Press, Princeton, N. J.

[8] W. RUDIN (1966), *Real and Complex Analysis*, McGraw-Hill, New York.

[9] D. W. WALKUP AND R. WETS (1967), *Stochastic programs with recourse*, SIAM J. Appl. Math., 15, pp. 1299–1314.

[10] R. WETS (1972), *Stochastic programs with recourse: A basic theorem for multistage problems*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 21, pp. 201–206.

[8] Strictly speaking, the argument is circular since Proposition A.1 makes no sense unless $\bar{p}_K, \cdots, \bar{p}_1$ are essentially defined. But the proposition's inductive proof can be woven into the theorem's (inductive) proof to avoid the circularity.