Some Remarks on Stochastic Versions of the Ramsey Growth Model*

Karel Sladký
Institute of Information Theory and Automation of the AS CR, Department of Econometrics, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic.
e-mail: sladky@utia.cas.cz

Abstract. In this note we focus attention on stochastic versions of the Ramsey growth model if either for a given time horizon expected value of the considered utility function should be maximized or if for infinite time horizon maximal average utility should be obtained. In contrast to the standard Ramsey economy growth model we assume that the production function considered in the economy model is influenced by some random factor with some specific properties. The aim is to discuss various approaches suitable for finding optimal policy of the "stochasticized" Ramsey model. To this end, we summarize basic features of multistage stochastic programming and stochastic dynamic programming – the two main methodologies that can be used to handle the above problem. Finally, we show how these approaches can be employed for finding optimal control policies for the "stochasticized" versions of the Ramsey problem if full or only partial information on the development of the economy over time is available.

Keywords: Economic dynamics, Ramsey growth model with disturbance, multistage stochastic programs, stochastic dynamic programming, finding optimal policies.

JEL classification: C61, E21, E22
AMS classification: 90C40, 91B15, 91B16

1. Ramsey growth model

The heart of the seminal paper of F. Ramsey [15] on mathematical theory of saving is an economy producing output from labour and capital and the task is to decide how to divide production between consumption and capital accumulation to maximize the global utility of the consumption. Ramsey's model is purely deterministic originally considered in continuous-time setting; Ramsey suggested some variational methods for finding an optimal policy how to divide the production between consumption and capital accumulation.

In the present section we formulate the Ramsey model in the discrete-time setting similarly as in the recent literature on economic growth models (see e.g. Le Van and Dana [4], Heer and Maußer [6], Majumdar, Mitra, and Nishimura [12], and Sladký [21, 22]). Moreover, in contrast to the standard Ramsey's model we assume

*This research was supported by the Grant Agency of the Czech Republic under Grants P402/11/0150, P402/10/0956 and P402/10/1610.

that the production function considered in the model is influenced by some random disturbances with specific properties. Finding optimal policy of the model can be formulated as finding optimal policy of a highly structured Markov decision process with complete or only partial information on the state of the economy over time, or as a specific multistage stochastic programming problem. We summarize basic features and similarities of the above two approaches and show how they can be employed for finding optimal control for the "stochasticized" versions of the Ramsey model if full or only partial information on the development of the economy over time is available.

1.1. Classical Ramsey growth model

We consider at discrete time points $t = 0, 1, \ldots$, an economy in which at each time t there are L_t (merely identical) consumers with consumption c_t per individual. The number of consumers grow very slowly in time, i.e. $L_t = L_0(1+n)^t$ for t with $\alpha := (1+n) \approx 1$. The economy produces at time t gross output Y_t using only two inputs: capital K_t and labour $L_t = L_0(1+n)^t$. A production function $F(K_t, L_t)$ relates input to output, i.e.

$$Y_t = F(K_t, L_t) \text{ with } K_0 > 0,\ L_0 > 0 \text{ given.} \tag{1}$$

We assume that $F(\cdot,\cdot)$ is a strictly increasing concave twice differentiable homogeneous function of degree one, i.e. $F(\theta K, \theta L) = \theta F(K, L)$ for any $\theta \in \mathbb{R}$.

The output must be split between consumption $C_t = c_t L_t$ and gross investment I_t, i.e.

$$C_t + I_t \le Y_t = F(K_t, L_t). \tag{2}$$

Investment I_t is used in whole (along with the depreciated capital K_t) for the capital K_{t+1} at the next time point $t+1$. In addition, capital is assumed to depreciate at a constant rate $\delta \in (0,1)$, so capital related to gross investment at time $t+1$ is equal to

$$K_{t+1} = (1-\delta)K_t + I_t. \tag{3}$$

In what follows let $k_t := K_t/L_t$ be the capital per consumer at time t, and similarly let $y_t := Y_t/L_t$ (resp. $c_t := C_t/L_t$) be the per capita output (resp. consumption) at time t. Recalling that the production function $F(\cdot,\cdot)$ is assumed to be homogeneous of degree one, then $f(k_t) := F(k_t, 1)$ denotes the per capita production at time t. In virtue of (2), (3) we get

$$c_t + (1+n)k_{t+1} - (1-\delta)k_t \le y_t = f(k_t), \tag{4}$$

and if we set for simplicity $\alpha \equiv (1+n) = 1$ then (4) can be written as

$$c_t + k_{t+1} - (1-\delta)k_t \le y_t = f(k_t). \tag{5}$$

The aim is to find an optimal control policy, i.e. a rule how to split at each time point production between consumption and capital accumulation, such that either (i) for the given time horizon T maximizes the utility function

$$U_{k_0}(c_0, \ldots, c_{T-1}) \text{ of a single consumer, or} \tag{6}$$

(ii) maximizes the mean utility of a single consumer for the infinite time horizon, i.e. the values

$$\lim_{T\to\infty} \frac{1}{T} U_{k_0}(c_0,\dots,c_{T-1}) \tag{7}$$

Recall that the utility function $U_{k_0}(c_0,\dots,c_T)$ is real, strictly increasing and concave function in all its arguments $c_0,\dots,c_T$ and if all $c_i \equiv 0$ then also $U_{k_0}(c_0,\dots,c_T) = 0$. As we shall see later, if the finite time horizon is considered optimal policy heavily depends on initial capital k_0; to this end this value appears in (6), (7) as a subscript.

As we shall see later since the development of the economy over time is governed by the recursive formulas (4), (5), the problem is much easier to solve if the utility function is additive, i.e. if for the considered time horizon T preferences for consumption of a single consumer are taken in the form (resp. in discounted form for discount factor $\beta \in (0,1)$)

$$U_{k_0}(c_0,\dots,c_T) = \sum_{t=0}^{T} u(c_t) \qquad (\text{resp.} \qquad U^{\beta}_{k_0}(c_0,\dots,c_T) = \sum_{t=0}^{T} \beta^t u(c_t)). \tag{8}$$

Observe that for $T \to \infty$ the value $U_{k_0}(c_0,\dots,c_T)$ is typically infinite. To this end we introduce mean global utility as

$$g := \liminf_{T\to\infty} (T)^{-1} \sum_{t=0}^{T-1} u(c_t) = \liminf_{T\to\infty} (T)^{-1} U(c_0,\dots,c_{T-1}) \tag{9}$$

and show that under some reasonable and realistic assumptions also

$$\lim_{T\to\infty} (T)^{-1} U(c_0,\dots,c_{T-1}) \tag{10}$$

exists and is independent of the initial capital k_0.

In the above formulation we assume that the per capita production function $f(k)$ and the utility function of the consumption $u(c)$ fulfil some standard assumptions on production and utility functions, in particular, that:

AS 1. The function $u(c) : \mathbb{R}^+ \to \mathbb{R}^+$ is twice continuously differentiable and satisfies $u(0) = 0$. Moreover, $u(c)$ is strictly increasing and concave (i.e., its derivatives satisfy $u'(\cdot) > 0$ and $u''(\cdot) < 0$) with $u'(0) = +\infty$ (so-called Inada Condition).

AS 2. The function $f(k) : \mathbb{R}^+ \to \mathbb{R}^+$ is twice continuously differentiable and satisfies $f(0) = 0$. Moreover, $f(k)$ is strictly increasing and concave (i.e., its derivatives satisfy $f'(\cdot) > 0$ and $f''(\cdot) < 0$) with $f'(0) = M < +\infty$, $\lim_{k\to\infty} f'(k) < 1$. Hence, if $f'(0) > 1$ there exists k^* such that $f(k^*) = k^*$.

Since $u(\cdot)$ is increasing (cf. assumption AS 1) in order to maximize global utility of the consumers is possible to replace (5) by the (nonlinear) difference equation

$$k_{t+1} - (1-\delta)k_t - f(k_t) = -c_t \quad \text{with } k_0 > 0 \text{ given} \tag{11}$$

or equivalently for $\tilde{f}(k) := f(k) + (1-\delta)k$ by

$$k_{t+1} - \tilde{f}(k_t) = -c_t \quad \text{with } k_0 \text{ given}, \tag{12}$$

where c_t $(t = 0, 1, \ldots)$ with $c_t \in [0, \tilde{f}(k_t)] = [0, f(k_t) + (1-\delta)k_t]$ is selected by the decision maker.

Up to now the system described above is purely deterministic; hence the initial capital k_0 along with the control policy c_t fully determines development of (k_t, c_t) over time. Recalling that in our model the utility functions $U_{k_0}(c_0, \ldots, c_T)$, $U^{\beta}_{k_0}(c_0, \ldots, c_T)$ are additive and the development over time is given by a recursive relation depending only on the current state of the system and the decision taken, optimal policy can be calculated using standard methods of dynamic programming.

1.2. Random shocks and imprecisions in the growth model

Unfortunately, in the real-life situations also some random shocks or imprecisions should be considered. For this reason, we shall assume that for a given value of k_t we obtain the output value y_t with some uncertainty; in particular we assume that $y_t \in [f_{\min}(k_t), f_{\max}(k_t)]$ (i.e. $f_{\min}(\cdot) \le f(\cdot) \le f_{\max}(\cdot)$; AS 2 also hold for $f_{\max}(k)$, $f_{\min}(k)$). Using this approach we can generate upper and lower bounds on optimal values on replacing in (11) $f(k_t)$ by $f_{\max}(k_t)$ and $f_{\min}(k_t)$ respectively.

Obviously, better results can be obtained if we replace the rough estimates of y_t generated by means of $f_{\max}(k_t)$ and $f_{\min}(k_t)$ by a more detailed information on the (random) output y_t generated by the capital k_t. To this end it is possible to assume that in (5) $y_t = f(k_t)$ is contaminated by adding random shocks ε_t, i.e.

$$\bar{f}(k_t) := f(k_t) + \varepsilon_t, \quad \text{where} \tag{13}$$

(i) either ε_t are i.i.d. random variables taking values in $[\underline{e}, \overline{e}] \in \mathbb{R}$ with known distribution, or

(ii) the sequence $\{\varepsilon_t, t = 0, 1, \ldots\}$ is a Markov sequence, i.e. distribution of ε_t depends on the value taken by ε_{t-1}.

Then by (12),(13) we can conclude that

$$k_{t+1} - \bar{f}(k_t) = -c_t \quad \text{with } k_0 \text{ given} \tag{14}$$

and $\bar{y}_t := \bar{f}(k_t)$ is he real output at time t (i.e. $\bar{y}_t$ is a random variable).

Such extensions well correspond to the models introduced and studied in [23] and also in [6, 12].

In [11] it was assumed that (13),(14) are fulfilled only with a given probability $1-\gamma$, i.e. if Probab$\{\varepsilon_t \ne [\underline{e}, \overline{e}]\} = \gamma$. Then it is possible to verify (cf. [11], Prop. 1) that

$$\text{Probab}\{\varepsilon_t \in [\underline{e}, \overline{e}], t = 0, 1, 2, \ldots, T-1\} = (1-\gamma)^T.$$

In contrast to the deterministic model where the initial capital k_0 along with selected consumptions (i.e. the control policy) c_t $(t = 0, 1, \ldots)$ fully determines the development (k_t, c_t) over time, in the considered stochastic case (k_t, c_t) as well as the values of the considered utility functions $U_{k_0}(c_0, \ldots, c_t)$ resp. $U^{\beta}_{k_0}(c_0, \ldots, c_t)$ are random variables. To this end, quality of the considered control policy is evaluated according to the *expected values of* $U_{k_0}(c_0, \ldots, c_T)$ resp. $U^{\beta}_{k_0}(c_0, \ldots, c_T)$, say

$$\mathbb{E}\ U_{k_0}(c_0, \ldots, c_T), \quad \text{resp.} \ \ \mathbb{E}\ U^{\beta}_{k_0}(c_0, \ldots, c_T).$$

Moreover, in the stochastic case we may assume that either the current values of the total output $y_t,\ c_t,\ k_t$ are known to the decision maker, or except of the knowledge of initial capital k_0 no information on the current values of capital, output and consumption is available.

To analyze the stochasticized version at first we shall assume that the random disturbances ε_t take on only a finite number of values say $\underline{e} = e_1 < e_2 < \ldots < e_N = \bar{e}$. In particular,

if ε_t are i.i.d. random variables, we shall assume that Probab $\{\varepsilon_t = e_i\} = p_i$,

if $\{\varepsilon_t, t = 0, 1, \ldots\}$ is a Markov sequence then

Probab $\{\varepsilon_t = e_j\} = \sum_{j=1}^{N} p_{ij}$Probab $\{\varepsilon_{t-1} = e_i\}$ where

$P = [p_{ij}]$ is a given transition probability matrix.

Recalling that in our model the utility functions $U_{k_0}(c_0, \ldots, c_T)$, $U^{\beta}_{k_0}(c_0, \ldots, c_T)$ are additive, also $\mathbb{E}\, U_{k_0}(c_0, \ldots, c_T)$, and $\mathbb{E}\, U^{\beta}_{k_0}(c_0, \ldots, c_T)$ must be additive. Since the development of the economy over time is given by a recursive relation depending only on the current state of the system and the decision taken, optimal policy can be calculated using standard methods of stochastic dynamic programming or multistage programming.

In the next section we present and compare basic properties of these two methods.

2. Multistage stochastic programs and stochastic dynamic programming

Multistage stochastic programs and dynamic programming problems with discrete time parameter deal essentially with the same types of problems – the dynamic and stochastic decision processes. They were initiated approximately in the mid fifties, but they did follow an independent development and recognition of similarities and complementary features has been rare.

In this section we shall discuss similarities and differences of multistage stochastic programs with recourse and stochastic dynamic programs with discrete time parameter and for a fixed finite horizon T. The main distinction is in the decision concept, in different structures used in their formulation and, consequently, also in different solution methods. The material is adapted from [5, 17, 18, 24].

Multistage stochastic programming extends the two-stage stochastic programming problems to more steps. The problem originally stems from standard mathematical programs where the decision maker takes some action in the first stage, after which a random effect occurs affecting the outcome of the first stage decision. A recourse decision can then be made in the second stage that compensate for any bad effect that might have been experience as a result of the first stage decision. Multistage programs compensate bad effects of decision taken successively in a finite number of stages.

On the contrary to the multistage stochastic programs, most of the motivation for the research on dynamic programming models come from a class of operations research and engineering applications where dynamical properties of the considered system are mostly known. Here it is the decision rule that is primarily of interest and the horizon is very long, hence the insistence on finding a rule that depends on the observed state and not on the information we may infer about the underlying stochastic phenomena. An appropriate definition of state is then the central point of dynamic programming formulations whereas in the context of multistage stochastic programs states usually do not appear.

2.1. Multistage stochastic programs

In a generic form the general *T-stage stochastic program* can be written as

$$\text{maximize} \quad \mathbb{E}\left\{G(x_0, x_1(x_{[1]}, \omega_1), x_2(x_{[2]}, \omega_2), \ldots, x_{T-1}(x_{[T-1]}, \omega_{T-1}), \omega_{[T]})\right\} \tag{15}$$

$$\text{subject to} \;\; x_0 \in \mathcal{X}_0, \quad x_t \in \mathcal{X}_t(x_{t-1}(\omega_{[t-1]}), \omega_{[t]}) \;\; t = 1, \ldots, T-1. \tag{16}$$

Here $G(\cdot)$ is a measurable real function of its arguments, $\omega_{[T]} = (\omega_0, \omega_1, \ldots, \omega_{T-1})$ with $\omega_t \in \mathbb{R}^{r_t}$ is a random data process considered at time points $t = 0, 1, ..., T-1$; in what follows we assume that the probability distribution of $\omega_{[T]}$ is known. Similarly,

$$x_{[T]} = (x_0, \ldots, x_t, \ldots x_{T-1}) \;\text{ with }\; x_t \in \mathcal{X}_t \subset \mathbb{R}^{r_t} \;\text{ for }\; t = 0, 1, \ldots, T-1 \tag{17}$$

are the decision variables at time t. Observe that $x_{[1]} = x_0$, $\omega_{[1]} = \omega_0$, where the initial decision x_0 is independent of ω_0. Furthermore, the decision process is nonanticipative, i.e. decisions taken at any stage of the process do not depend on future realization of the data process or on future decisions and only the past information and process probabilistic specification is exploited.

Policy is a sequence of functions $x_0 \in \mathcal{X}_0$, along with $x_t \in \mathcal{X}_t(x_{t-1}(\omega_{[t-1]}), \omega_{[t]})$ for $t = 1, \ldots, T-1$, and is feasible if and only if it satisfies the feasibility condition (16) for almost every realization of the random data process. If the data process $\omega_{[T]}$ has a finite number of realization (called scenarios) the problem leads to a finite dimensional optimization.

Calculating maximum in (15) subject to (16) is a very difficult problem. To this end, we make

Assumption 1. The objective function $G(\cdot)$ occurring in (15) is separable with respect to the stage index t. In particular, we assume that $G(\cdot)$ is additive, i.e. in (15) we try to find maximum of

$$\mathbb{E}\left\{G_0(x_0) + G_1(x_1(\omega_{[1]}), \omega_1) + G_2(x_2(\omega_{[2]}), \omega_2) \ldots + G_{T-1}(x_{T-1}(\omega_{[T-1]}), \omega_{T-1})\right\} \tag{18}$$

where $G_t : \mathbb{R}^{n_t} \times \mathbb{R}^{r_t} \to \mathbb{R}$ are measurable functions and $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \mathbb{R}^{r_t} \to \mathbb{R}^{n_t}$ are measurable multifunctions (cf. e.g. [17] for precise definition). In the first stage the function $G_0 : \mathbb{R}^{n_0} \to \mathbb{R}$ and the set $\mathcal{X}_0$ are deterministic.

Remark. In particular, the multistage program is *linear* if in (15)–(16) the objective function and the constraints are linear, that is if

$$G_t(x_t, \omega_t) = c_t^{\mathrm{T}} x_t, \quad \mathcal{X}_0 = \{x_0 : A x_0 = b_0, x_0 \geq 0\}$$

$$\mathcal{X}_t(x_{t-1}, \omega_t) = \{x_t : B_t x_{t-1} + A_t x_t = b_t, x_t \geq 0\} \;\text{ for }\; t = 1, 2, \ldots, T-1$$

where $\omega_0 = (c_0, A_0, b_0)$ and for $t = 1, 2, \ldots, T-1$

$\omega_t = (c_t, B_t, A_t, b_t) \in \mathbb{R}^{r_t}$ are data vectors with possibly random elements.

Here in $\mathcal{X}_t(x_{t-1}, \omega_t)$ only one step dependence on x_{t-1} is considered; in general we might extend the linear dependence on all $x_0, x_1, \dots x_{t-1}$.

On introducing conditional expectation, say $\mathbb{E}_{|\xi_1}[g(\xi_1, \xi_2)] := \mathbb{E}[g(\xi_1, \xi_2)|\xi_1]$ for the real function $g(\cdot)$, we can conclude that by (18)

$$\mathbb{E}\left\{G_0(x_0) + G_1(x_1(\omega_{[1]}), \omega_1) + G_2(x_2(\omega_{[2]}), \omega_2) \dots + G_{T-1}(x_{T-1}(\omega_{[T-1]}), \omega_{T-1})\right\} \quad (19)$$

$$= G_0(x_0) + \mathbb{E}_{|\omega_0}\left[G_1(x_1(\omega_{[1]}), \omega_1) + \mathbb{E}_{|\omega_1}\left[G_2(x_2(\omega_{[2]}), \omega_2)\right.\right.$$

$$\left.\left. + \mathbb{E}_{|\omega_3}\left[\dots + \mathbb{E}_{|\omega_{T-2}}\left[G_{T-1}(x_{T-1}(\omega_{[T-1]}), \omega_{T-1})\right]\right]\right]\right]. \quad (20)$$

This, together with an interchangeability property of the expectation a maximization operators (trivially fulfilled for a finite number of realizations), leads to the following nested formulation of the multistage problem (18)

$$\max_{x_0 \in \mathcal{X}_0} \{G(x_0) + \mathbb{E}_{|\omega_0}\Big[\sup_{x_1 \in \mathcal{X}_1(x_0, \omega_1)} G_1(x_1, \omega_1) + \mathbb{E}_{|\omega_{[1]}}\Big[$$

$$\dots + \mathbb{E}_{|\omega_{[T-2]}} \sup_{x_{T-1} \in \mathcal{X}_{T-1}(x_{T-2}, \omega_{T-1})} G_{T-1}(x_{T-1}, \omega_{T-1})\Big]\Big]\}. \quad (21)$$

This decomposition property of the expectation operator is a basis for deriving the dynamic programming equations, i.e. finding maximum of the objective function sequentially by going backward in time. To this end, it is necessary to assume stagewise independence of elements ω_t's of the random data process $\omega_{[T]}$, in the multistage stochastic programs given by (15)–(16).

Assumption 2. (Interstage independence assumption.) The random process $\omega_1, \dots, \omega_T$ is *stagewise independent* if random variable ω_{t+1} is independent of $\omega_{[t]}$ for $t = 0, 1, \dots, T-1$.

Under Assumptions 1 and 2, going backward in time the so-called *reward-to-go* (also called *value*) *functions* are defined recursively for $t = T-1, \dots, 1, o$ as follows

$$V_t(x_{t-1}, \omega_{[t]}) = \sup_{x_t \in \mathcal{X}_t(x_{t-1}, \omega_t)} \{G_t(x_t, \omega_t) + \mathcal{V}_{t+1}(x_t, \omega_{[t]})\} \quad (22)$$

where

$$\mathcal{V}_{t+1}(x_t, \omega_{[t]}) = \mathbb{E}\{V_{t+1}(x_t, \omega_{[t+1]})|\omega_{[t]}\} \quad (23)$$

with $\mathcal{V}_T(\cdot, \cdot) = 0$ by definition. At the first stage the following problem should be solved

$$\max_{x_0 \in \mathcal{X}_0} \{G_0(x_0) + \mathbb{E}[V_1(x_0, \omega_1]\}. \quad (24)$$

A policy $\hat{x}_t(\omega_{[t]})$ $t = 0, 1, \dots, T-1$ is called *optimal* if $\hat{x}_0$ is an optimal solution of the first stage problem (24) and for $t = 1, \dots, T-1$

$$\hat{x}_t(\omega_{[t]}) \in \arg \max_{x_t \in \mathcal{X}_t(\hat{x}_{t-1}(\omega_{[t-1]}), \omega_t)} \{G_t(x_t, \omega_t) + \mathcal{V}_{t+1}(x_t, \omega_{[t]})\}, \quad \text{w.p.1.} \quad (25)$$

In the dynamic programming formulation the problem is reduced to solving a family of finite dimensional problems (22)–(23). In particular, under Assumptions 1 and 2 for the expected value of the value function, say $\mathcal{V}_t(x_{t-1}, \omega_{[t-1]}) := \mathbb{E}\{V_t(x_{t-1}, \omega_t|\omega_{[t-1]})\}$, we have

$$\mathcal{V}_{T-1}(x_{T-2}, \omega_{[T-2]}) = \mathbb{E}\{V_{T-1}(x_{T-2}, \omega_{T-1})|\omega_{[T-2]}\} \tag{26}$$

does not depend on $\omega_{[T-2]}$. By induction on t going backward in time, it can be shown that:

Under Assumptions 1 and 2 the (expected) value of $\mathcal{V}_t(x_t, \omega_{[t]})$, say $V_t(x_t)$ for $t = 1, \ldots, T-1$ do not depend on the data process and equations (22) take the form

$$V_t(x_{t-1}, \omega_t) = \sup_{x_t \in \mathcal{X}_t(x_{t-1}, \omega_t)} \{G_t(x_t, \omega_t) + \mathcal{V}_{t+1}(x_t)\} \tag{27}$$

Using the above method under Assumptions 1 and 2 the general T-stage stochastic programming problem (15)–(16) reduces to an optimization problem of a finite system of parametric (one-stage) optimization problems with an inner type of dependence. This can be useful for many problem arising in multistage programming, see e.g. [9], [10]. However, even this approach is based on backward recursive formulas of dynamic programming, it is limited only to *finite horizon* models and does not employ explicitly the dynamical properties of the considered model. Having reasonable information concerning the behaviour of the considered model over time, it is possible to introduce the state space characterization of the model and extend the analysis also to infinite time horizon.

2.2. Connections with stochastic dynamic programming

To show the connections between multistage stochastic programs and stochastic dynamic programs, let us consider again the sequence (17) assuming that the next stage of the considered process is entirely determined by the state, decision and random data occurring at the current stage. In particular, we make the following assumptions concerning the dynamics of the system:

Assumption 3. For every stage $t = 0, 1, \ldots, T-1$

$$x_t = (s_t, d_t) \quad \text{with} \quad s_{t+1} = F_t(s_t, d_t, \omega_t), \quad \text{and} \quad x_0 = (s_0, d_0) \ \text{ with } s_0 \ \text{ given,} \tag{28}$$

where

$s_t \in \mathcal{S} \subset \mathbb{R}$, $d_t \in \mathcal{D}(s_t) \subset \mathbb{R}$, $\omega_t \in \Omega_t \subset \mathbb{R}$, and
$F_t(\cdot, \cdot, \cdot)$ is a mapping from $\mathcal{S} \times \mathcal{D}(s_t) \times \Omega_t$ onto $\mathcal{S}$.

The variable s_t is called the *state* of the considered random process at stage t and $\mathcal{D}(s)$ is the *decision or action*) taken at stage t if the process is in state $s \in \mathcal{S}$. Similarly, $\mathcal{S}$ (resp. $\mathcal{D}(s)$) is the *state space* (resp. *action set*) at stage t. Let $d^t := (d_0, d_1, \ldots, d_t)$ be a sequence of decisions (called also *policy*) controlling the considered process. Observe that in virtue of (28) the sequence of the states $s_0, s_1, \ldots, ...s_t, \ldots$ is a Markov process, and that (28) in fact replaces the more general conditions (16) by describing the development of the system over time.

The "Markovian" property for generating the sequences of states by (28) of Assumption 3 along with mutual independence of ω_t (interstage independence by Assumption 2) yields the following important implication:

$$\text{Prob}\,[s_{t+1} = s|s_1, \ldots, s_t; d_1, \ldots, d_t] = \text{Prob}\,[s_{t+1} = s|s_t,\, d_t] := p_t(s_t, s; d_t) \qquad (29)$$

and the form of objective function (18) follows from Assumption 3:

$$\mathbb{E}\,\{G_0(s_0) + \sum_{t=1}^{T-1} G_t(s_t; d_t)\} \qquad (30)$$

with $G_t(s_t; d_t) = G_t(s_t) + \bar{G}_t(s_t; d_t)$ for $t = 1, \ldots, T-1$.

In the literature (cf. [6, 12] or the monograph [23]) it is usually assumed that the sequence of states $\{s_0, s_1, \ldots s_t, \ldots\}$ is a Markov process in general with state space $\mathbb{R}$. Moreover, usually we assume that *the decision maker can observe the current values of the total output* y_t *and then select the value of* k_{t+1}. Unfortunately, assuming that the state space of the considered Markov process is an compact set of $\mathbb{R}$ then a rigorous treatment of the model requires a very sophisticated mathematics (see e.g. [7] or [23]) and is not suitable for numerical computation.

3. Approximations of the stochasticized growth model

To make the model computationally tractable we shall approximate our model by a discretized model with finite state space (see [20]).

To this end, we shall assume that the values of c_t, k_t, and y_t take on only a finite number of discrete values. In particular, we assume that for sufficiently small $\Delta > 0$ there exists nonnegative integers $\bar{c}_t$, $\bar{k}_t$, and $\bar{y}_t$ such that for every $t = 0, 1, \ldots$ it holds:

$\bar{c}_t\Delta = c_t$, $\bar{k}_t\Delta = k_t$, and $\bar{y}_t\Delta = y_t$ with $\bar{k}_t \le K := k_{\max}/\Delta$
and similarly $\bar{y}_t \le Y := y_{\max}/\Delta$.

Let elements of $\bar{k}_t$ be labelled by integers from $\mathcal{I}_K = \{0, 1, \ldots, K\}$ and elements of $\bar{y}_t$ by integers from $\mathcal{I}_Y = \{0, 1, \ldots, Y\}$. Hence for the total output y_t generated by the "randomized" production function we get for $\ell = 0, 1, 2, \ldots, L$ and $\bar{k}_t = 0, 1, \ldots, K$

$$\bar{y}_t = f(\bar{k}_t\Delta)/\Delta - \ell \;\text{ with probability } p(\bar{k}_t; \ell) \text{ such that} \qquad (31)$$

$$p(\bar{k}_t; 0) \gg p(\bar{k}_t; \ell) \;\text{ for any } \ell \neq 0;\text{ obviously, } \sum_{\ell=0}^{L} p(\bar{k}_t; \ell) = 1.$$

If the (random) total output at time t $\bar{y}_t = \bar{y}$ then the decision maker have option to invest for the next time point the capital $k_{t+1} = \bar{k}_{t+1}\Delta$ where $\bar{k}_{t+1} = 0, 1, \ldots, f_{\max}(\bar{k}_t)$, and hence $u((\bar{y}_t - \bar{k}_{t+1})\Delta)$ is the instantaneous utility accrued at time t to the global utility.

Using the above discretization and taking decisions with respect to the current states, the development of the economy over time can be well described by a (structured) Markov reward chain $X = \{X_\tau, \tau = 0, 1, \ldots\}$ with finite state space $\mathcal{I} = \mathcal{I}_K \cup \mathcal{I}_Y$ (with $\mathcal{I}_K \cap \mathcal{I}_Y = \emptyset$), transition probabilities $p(\bar{k}_t; \bar{y}_t) = p_{ij}$, for $i = \bar{k}_t \in \mathcal{I}_K$, $j = \bar{y}_t \in \mathcal{I}_Y$, and a "non-random" transition from state $j = \bar{y}_t \in \mathcal{I}_Y$ to state $\ell = \bar{k}_t \in \mathcal{I}_K$ *associated with one-stage reward* $r_{j\ell} = u((\bar{y}_t - \bar{k}_{t+1})\Delta)$. Observe that actually "two transitions" of the Markov chain $X = \{X_\tau, \tau = 0, 1, \ldots\}$ occur within one-time period of the considered economy model and the *one-stage reward is accrued only in even transitions.* Hence the global utility (i.e. the total reward of the Markov chain) $U_{k_0}(T) = \mathbb{E}\{\sum_{t=1}^{T} r_{X_{2t-1},X_{2t}} | X_0 = \bar{k}_0\}$, resp. $U^{\beta}_{k_0}(T) = \mathbb{E}\{\sum_{t=1}^{T} \beta^{k-1} r_{X_{2t-1},X_{2t}} | X_0 = \bar{k}_0\}$ if discounted reward is considered

4. Formulation in terms of dynamic programming

The above model can be treated as a highly structured Markov decision chain with finite state space $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ (with $\mathcal{I}_1 \cap \mathcal{I}_2 = \emptyset$), finite set $\mathcal{D}_i = \{0, 1, \ldots, d(i)\}$ of possible decisions (actions) in state $i \in \mathcal{I}$ and the following transition and reward structure:

$p_{ij}(a)$: transition probability from $i \to j$ $(i, j \in \mathcal{I})$ if action $a \in \mathcal{D}_i$ is selected,

r_{ij} : one-stage reward for a transition from $i \to j$, with $r_{ij} = u((i-j)\Delta)$ if $i \in \mathcal{I}_2$ and $j \in \mathcal{I}_1$,

$r_i(a)$: expected value of the one-stage rewards incurred in state i if decision $a \in \mathcal{D}_i$ is selected in state i; in particular $r_i(a) = \sum_{j \in \mathcal{I}} p_{ij}(a) \cdot r_{ij}$.

Policy controlling the chain, say π, is a rule how to select actions in each state. Policy π is then fully identified by a sequence $\{d_\tau, \tau = 0, 1, \ldots\}$ of decision vectors (of dimension K and Y in odd and even steps respectively) whose ith element $d_\tau(i) \in \mathcal{D}_i$ identifies the action taken if $X_\tau = i$.

Let the vector $U^\pi(\tau)$ (with elements $U^\pi_i(\tau)$) denote expectation of the (random) global utility ξ_τ received in the τ next transitions of the considered Markov chain X if policy $\pi = (d_\tau)$ is followed, given the initial state $X_0 = i$, i.e., for the elements of $U^\pi(\tau)$ we have $U^\pi_i(\tau) = \mathbb{E}^\pi_i[\xi_\tau]$ where $\xi_\tau = \sum_{k=0}^{\tau-1} r_{X_k, X_{k+1}}$ and $\mathbb{E}^\pi_i$ is the expectation if $X_0 = i$ and policy $\pi = (d_\tau)$ is followed. Then obviously

$$U^\pi_i(\tau+1) = r_i(d_\tau(i)) + \sum_{j \in \mathcal{I}} p_{ij}(d_\tau(i)) \cdot U^\pi_j(\tau), \quad i \in \mathcal{I} \tag{32}$$

and in the case with discounting we have

$$U^{\beta,\pi}_i(\tau+1) = r_i(d_\tau(i)) + \beta \sum_{j \in \mathcal{I}} p_{ij}(d_\tau(i)) \cdot U^{\beta,\pi}_j(\tau), \quad i \in \mathcal{I}. \tag{33}$$

If we restrict on stationary policies, i.e. the rules selecting actions only with respect to the current state of Markov chain X, then policy π is fully determined by $d_t \equiv d$. Observe that decision vector d then completely identifies the transition probability matrix $P(d)$ and the ith row of $P(d)$ has elements $p_{i1}(d(i)), \ldots, p_{iN}(d(i))$. Similarly, $r(d)$ is a (column) vector of one-stage expected rewards (i.e. i-th element of $r(d)$ is equal to $r_i(d(i))$). Observe that if the decision maker's policy it to *select* *the same decision (i.e. to apply the same action) in any state no information on the current state of the system is necessary* (cf. [1, 2]).

In particular, for the discounted model with T tending to infinity, i.e. when $\lim_{T\to\infty} U_i^{\beta,\pi}(T) = U_i^{\beta,\hat{\pi}}$, then by (33) can conclude that

$$U_i^{\beta,\hat{\pi}} = r_i(d(i)) + \beta \sum_{j\in\mathcal{I}} p_{ij}(d(i)) \cdot U_j^{\beta,\hat{\pi}}, \quad i \in \mathcal{I} \tag{34}$$

Furthermore, (stationary) policy $\hat{\pi}$ maximizing total expected rewards can be found as a solution to

$$U_i^{\beta,\hat{\pi}} = \max_{d\in\mathcal{D}_i} [r_i(d(i)) + \beta \sum_{j\in\mathcal{I}} p_{ij}(d(i)) \cdot U_j^{\beta,\hat{\pi}}], \quad i \in \mathcal{I} \tag{35}$$

In what follows we assume that for an arbitrary policy the considered Markov chain contains a single class of recurrent states guaranteed by the following:

Assumption 4. There exists state $i_0 \in \mathcal{I}$ that is accessible from any state $i \in \mathcal{I}$ for any stationary policy.

In contrast to (33),(35) the undiscounted values $U_i^{\pi}(T)$ can be typically infinite. However, under Assumption 4 it can be shown that for any stationary policy the long range average reward is independent of the starting state, in particular, there exist w_i^{π} and number $g^{\pi} = \lim_{T\to\infty} \frac{1}{T} U_i^{\pi}(T)$ such that

$$g^{\pi} + w_i^{\pi} = r_i(d(i)) + \sum_{j\in\mathcal{I}} p_{ij}(d(i)) \cdot w_j^{\pi}, \quad i \in \mathcal{I}. \tag{36}$$

where g^{π} is unique and w_i^{π}'s (for $i \in \mathcal{I}$) are unique up to an additive constant depending on the selected initial conditions (for details see e.g. [14, 16]).

If π^T is (in general nonstationary) policy maximizing the values $U_i^{\pi^T}(T)$ for the fixed time horizon T then

$$U_i^{\hat{\pi}^{\tau}}(\tau) = \max_{d\in\mathcal{D}_i} [r_i(d_\tau(i)) + \sum_{j\in\mathcal{I}} p_{ij}(d_\tau(i)) \cdot U_j^{\hat{\pi}^{\tau-1}}(\tau-1)], \quad \text{for} \quad \tau = T, T-1, \ldots, 1, 0. \tag{37}$$

Furthermore, for T tending to infinity then by (36), (37) we can conclude that the maximum average reward $g^{\hat{\pi}}$ can be found as a solution to

$$g^{\hat{\pi}} + U_i^{\hat{\pi}} = \max_{d\in\mathcal{D}_i} [r_i(d(i)) + \sum_{j\in\mathcal{I}} p_{ij}(d(i)) \cdot U_j^{\hat{\pi}}], \quad i \in \mathcal{I}. \tag{38}$$

where $g^{\hat{\pi}}$ is unique and $U_i^{\hat{\pi}}$'s (for $i \in \mathcal{I}$) are unique up to an additive constant depending on the selected initial conditions (for details see e.g. [14, 16]).

5. Computation of optimal policies

In case that the time horizon T is finite, it is necessary to calculate (backwards) the dynamic programming recursion according to (37). Considering the infinite time horizon (i.e. if $T \to \infty$), finding a solution of (37) is in some aspects much easier. Optimal policy can be found in the class of stationary policies (i.e. policies selecting actions only with respect to the current state of Markov chain) and can be performed either by value iterations (successive approximations) or by policy iterations. For details see e.g. [14] or [16].

Algorithm 1 (Policy iterations – Howard [8].)
Step 0. Select arbitrary policy, say $d^{(0)}$.
Step 1 – Policy evaluation. For stationary policy $d^{(n)}$ find $v = v(d^{(n)})$ as the solution of

$$g\, e + v = r(d^{(n)}) + P(d^{(n)})$$

(e denotes a unit vector).

Step 2 – Policy improvement. For a given $v(d^{(n)})$ find policy $d^{(n+1)}$ such that

$$r(d^{(n+1)}) + P(d^{(n+1)})v(d^{(n)}) = \max{}_{d\in D}[r(f) + P(d)v(d^{(n)})].$$

If there exists $d^{(n+1)} = d^{(n)}$, then stop and policy $d^{(n)}$ is an optimal policy, else go to Step 1.

Algorithm 2 (Value iteration – Bellman [3].)
Select $v^{(0)} = 0$, choose some (sufficiently small) $\varepsilon > 0$, and iterate

$$v^{(n+1)} := \max_{d\in D}[r(d) + P(d)v^{(n)}] \quad \text{for} \quad n = 0, 1, \ldots .$$

If $\|v^{(n+1)} - v^{(n)}\| < \varepsilon$ then stop, and $g^{(n)} = n^{-1}v^{(n)}$ is a very good approximation of the mean reward.

Remark. Observe that $v^{(n)}$'s are identical with $U^{\pi}(n)$ if policy π is identified by the decision vectors generated by Algorithm 2.

Algorithm 3 (Value iteration (modified) – White [25], Odoni [13], Sladký [19].)
Select $w^{(0)} = 0$, choose some (sufficiently small) $\varepsilon > 0$, set $w_N^{(n)} \equiv 0$ for $n = 0, 1, \ldots$, and iterate

$$w^{(n+1)} := \max_{d\in D}\,[r(d) + P(d)w^{(n)}].$$

Then

$$\max_{i\in\mathcal{I}}\,[w_i^{(n+1)} - w_i^{(n)}], \qquad \text{resp.} \quad \min_{i\in\mathcal{I}}[w_i^{(n+1)} - w_i^{(n)}]$$

is an upper (resp. lower) bound on g converging monotoneously to g.
If $\|w^{(n+1)} - w^{(n)}\| < \varepsilon$ then stop.

6. Conclusions

The paper is devoted to finding optimal control policy of a stochasticized version of the Ramsey growth model. The problem can be treated by two approaches for handling multistage stochastic models: multistage stochastic programming and stochastic dynamic programming. It is shown that under some simplifying assumptions on the general model of stochastic programming we can arrive at the same formulas as obtained by stochastic dynamic programming approach. Using a reasonable approximation of the stochasticized model we are able to formulate the problem as finding optimal control of a Markov decision chain with finite state space and action sets. For the infinite horizon models we can restrict on the class of stationary policies and relabel actions in each state such that in each state optimal action has the same label. This enables to apply optimal control even for problems where the decision maker has no information on the current state of the economy system if the long run optimality criteria are considered.

References

[1] Assaf, D.: Invariant problems in discounted dynamic programming. *Adv. Appl. Probab.* **10** (1978), 472–490.

[2] Assaf, D.: Invariant problems in dynamic programming– average reward criterion. *Stochastic Processes and their Applications* **10** (1980), 313–322.

[3] Bellman, R.: *Dynamic Programming.* Princenton Univ. Press, Princenton, NJ. 1957.

[4] Dana, R.-A. and Le Van, C: *Dynamic Programming in Economics.* Kluwer, Dordrecht 2003.

[5] Dupačová, J. and Sladký, K.: *Comparison of Multistage Stochastic Programming and Stochastic Dynamic Programming with Discrete Time.* Zeitschrift fuer Angewandete Mathematik und Mechanik **82** (2002), No. 1112, 753-765.

[6] Heer, B. and Maußer, A.: *Dynamic General Equilibrium Modelling.* Springer, Berlin 2005.

[7] Hernández-Lerma, O. and Lasserre J.: *Discrete-Time Markov Control Processes. Basic Optimality Criteria.* Springer, New York 1995.

[8] Howard, R. A.: *Dynamic Programming and Markov Processes.* MIT Press, Cambridge, Mass. 1960.

[9] Kaňková, V. and Šmíd, M.: A Remark on Approximation in Multistage Stochastic Programs: Markov Depen- dence. *Kybernetika* **40** (2004), 5, 625–638.

[10] Kaňková, V.: Multistage Stochastic Programs via Autoregressive Sequences and Individual Probability Constraints. *Kybernetika* **44** (2008), 2, 151–170.

[11] Kaňková, V.: Ramsey stochastic model via multistage stochastic programming. In: *Proc. 28th Internat. Conf. Mathematical Methods in Economics 2009*, Univ. South Bohemia, České Budějovice 2010, pp. 328–333.

[12] Majamdar, M., Mitra, T., and Nishinmura, K. (eds.): *Optimization and Chaos.* Springer, Berlin 2000.

[13] Odoni, A. R.: On finding the maximal gain for Markov decision processes. *Operations Research* **17** (1969), 857–860.

[14] Puterman, M. L.: *Markov Decision Processes – Discrete Stochastic Dynamic Programming.* Wiley, New York 1994.

[15] Ramsey, F. P.: A mathematical theory of saving. *Economic Journal* **38** (1928), 543–559.

[16] Ross, S. M.: *Applied Probability Models with Optimization Applications.* Holden–Day, San Francisco, Calif. 1970.

[17] Shapiro, A., Dentcheva, D. and Ruszynski, A.: *Lectures on Stochastic Programming: Modeling and Theory.* SIAM, Philadelphia 2009.

[18] Shapiro, A.: Minimax and risk averse multistage stochastic programming. *European Journal of Operational Research* **219** (2012), 719–726.

[19] Sladký, K.: O metodě postupných aproximací pro nalezení optimálního řízení markovského řetězce. *Kybernetika* **4** (1969), 167–176.

[20] Sladký, K.: Approximations in stochastic growth models. In: *Proceedings of the 24th Internat. Conference Mathematical Methods in Economics 2006*, University of West Bohemia, Pilsen 2006, pp. 465–470.

[21] Sladký, K.: Ramsey growth model under uncertainty. In: *Proc. 27th Internat. Conf. Mathematical Methods in Economics 2009*, Faculty of Economics and Management, Czech University of Life Sciences Prague, Prague 2009, pp. 296–300.

[22] Sladký, K.: Risk-sensitive Ramsey growth model. In: *Proc. 28th Internat. Conf. Mathematical Methods in Economics 2009*, Univ. South Bohemia, České Budějovice 2010, pp. 560–565.

[23] Stokey, N.-L. and Lukas, R. E. Jr.: *Recursive Methods in Economic Dynamics.* Harvard Univ. Press, Cambridge, Mass. 1989.

[24] Varaiya, P. and Wets, R. J.-B.: Stochastic dynamic optimization approaches and computation. In: Iri, M. and Tanabe, K. (eds.): *Mathematical programming.* KTKSci Publ., Tokyo 1990, pp. 309–332. See also WP-88-87, IIASA Laxenburg.

[25] White, D.J.: Dynamic programming, Markov chains, and the method of successive approximations. *J. Math. Anal. Appl.* **6** (1963), 373–376.