The Slovak Society for Operations Research
Department of Operations Research and Econometrics
Faculty of Economic Informatics
University of Economics in Bratislava

Proceedings of the International Conference

Quantitative Methods in Economics

(MULTIPLE CRITERIA DECISION MAKING X)

November 30 – December 2, 2000
The High Tatras, Slovak Republic

November 30 – December 2, 2000
The High Tatras, Slovak Republic

Technical Editor: *Juraj Pekár, University of Economics, Bratislava*
Web Editor: *Martin Lukáčik, University of Economics, Bratislava*

ISBN 80-887 15-93-8

CONTENT

Round Trip Departures to Serve Customers under Risk

Ondrej Bartl
University of Žilina, Slovakia

1. Operations Management in Services

Service centres are expected to meet requirements of customers effectively. Operations management rules should accept the facts that service providing activities are made over time in an environment of risk. Thus, service centres belong to stochastic dynamic systems, where economic consequences of system operation must be taken into account. Compromises are required between serving customers promptly and letting them waiting for service. Set of prescriptions that show dispatchers in a service centre how to control service process over a planning horizon is a policy. The policy determines, whether financial outcome of the system operation is acceptable or not. If service centre has probabilistic laws of motion governed by the Markovian property, then the future depends only on the present situation in the system and not on the past history. In such a case the policy for cost effective decision making on providing services can be revealed by means of the theory of Markov decision processes.

2. A service centre with round trips to serve customers

There is a class of service systems, where batch service is practised. Here we deal with a system with one service car used for round trips to serve customers. Service centre of the system registers requirements of customers for service from an attraction area of the centre. According to the current situation in the system, dispatcher in the service centre decides when to depart the car with a serviceman to visit customers and provide them with a service required. Customer requirements for service arrive into the service centre as a (stationary) Poisson process with rate $l > 0$. Due to a limited service capacity, there is a limitation L for the number of customer requirements that the centre is able to accept. It means that the service requirement of a customer arriving into the service centre in the situation, when the current number of registered customer requirements is L, is rejected without any registration. It is obvious that the dispatcher in the service centre can decide on the start of a round trip to serve customers in some time points only. A return of the service car into the centre upon completion of a round trip is one occasion to decide. Each arrival of a customer requirement, while the service car is parking in the centre, is another possibility to make the decision. Points of time specified above are decision epochs. Time consumption to serve one customer is considered being a deterministic variable whose value is given by a technical norm. If $i > 0$ customers are to be served during a round trip, then the corresponding duration of the trip is $t(i) = r + si$. Time required to reach the area with customers from the service centre and to return back is roughly approximated by the constant quantity r. A long term non-stop operation of the service system in unchanged financial conditions is assumed. Costs associated with round trips and with waiting of customers for service, together with rejection costs, reflect economic consequences of the system activity. Decisions on departures of the service car are expected to minimize the total outcome of service centre expenses over the planning horizon. The rules for decision making can be found on the basis of Markov decision modelling.

3. A semi-Markov decision model for sequential decisions on starting round trips

The aforementioned system is a queueing system with Poisson arrivals and deterministic service times with a batch service pattern, where one server exists and the length of a queue is limited to L. Since decisions on service car departures are taken at points in time, when some important events occur, the service system considered is a discrete event dynamic system. Time intervals between consecutive decision epochs referred to as review intervals or stages are random. The decisions are made in an environment of risk, depending on the current system situation at decision epochs. The choice of a decision influences future decision making situations. Due to the memoryless property of the arrival Poisson process, the impact of the system history on future system evolution is realized only through a present system state and the subsequent decision on a control action without dependence on previous states and decisions. The controlled evolution of such a dynamic system with the Markovian property over an infinite planning horizon may be represented by a semi-Markov decision process. Building the corresponding decision model enables us to determine an optimal policy for taking control actions by means of the theory of Markov decision processes. A semi-Markov decision model is built up, when the state and action variables are defined, the state and action spaces are specified, the transition probabilities, the cost functions and the stage length functions are determined and the criterion of interest is formulated.

The time set of the system is $\boldsymbol{T} = [0, \infty)$. The set of stages is $\boldsymbol{S} = \{0,1,...\}$, where the starting stage is numbered by the numeral 0. Let, for $n \in \boldsymbol{S}$, T_n denote the time, at which the nth decision epoch occurs, with convention that $T_0 \equiv 0$. Quantities T_n, $n \in \boldsymbol{S}$, are random variables. The corresponding set of decision epochs is $\boldsymbol{T_S} = \{T_0, T_1, ...\}$, where $T_n \in \boldsymbol{T} = [0, \infty)$, $n \in \boldsymbol{S}$.

The situation in the system at the nth decision epoch T_n is the state of the system at time T_n. It is denoted by the random variable $X_n \equiv X_n(T_n)$. The state X_n represents the number of registered customer requirements for service at time T_n before making a decision on the control action. The set of possible values of states is the state space $\boldsymbol{X} = \{0,1,...,L\}$. The action $a_n \equiv a_n(T_n)$ chosen by a decision taken at decision epoch T_n is a technological operation that has to be realized during the stage $n \in \boldsymbol{S}$ between decision epochs T_n and T_{n+1}. The (control) action a_n is the mode to be applied for the service car in the centre from time T_n onward. In general, two modes are possible: to depart and ride (the action $a_n = 1$) or to wait and be idle (the action $a_n = 0$). If the former mode is chosen, the car with a serviceman will start a round trip to serve all the customers, whose requirements have been accepted, being included in the current registration list at the epoch T_n. Then the centre clears the list of entries. If the latter alternative is employed, the car will prolong its idle regime period, waiting in the centre for the next customer requirement arrival. The set of feasible actions in state $i \in \boldsymbol{X}$ is $\boldsymbol{A}(i)$, where

$$\boldsymbol{A}(i) = \begin{cases} 0 & \text{for} \quad i = 0 \\ \{0,1\} & \text{for} \quad i \in \{1,2,..., L-1\} \\ 1 & \text{for} \quad i = L \end{cases} \tag{1}$$

The action space $\boldsymbol{A}$ of the system is the union of all $\boldsymbol{A}(i)$, $i \in \boldsymbol{X}$. Hence, $\boldsymbol{A} = \{0,1\}$.

The length of the review interval between decision epochs T_n and T_{n+1} is represented by a non-negative random variable $t_n \equiv t_n(X_n,a_n)$. That is, t_n denotes the duration of the nth stage, $n \in \boldsymbol{S}$. Let V be a random variable with the exponential probability distribution function $G(t) = 1 - e^{-lt}$, $t \geq 0$ and the probability density function $g(t) = le^{-lt}$, $t \geq 0$. Then,

$$t_n \equiv t_n(X_n, a_n) = a_n t(X_n) + (1 - a_n)V, \ n \in \boldsymbol{S}. \tag{2}$$

Recall, that $t(X_n) = r + sX_n$ is the time consumption for a round trip to serve X_n customers and $a_n \in \boldsymbol{A}(X_n) \subset \boldsymbol{A} = \{0,1\}$ is the control action employed in the state X_n at the decision epoch T_n. The quantity V represents either an interoccurrence time between successive customer requirement arrivals or the time between a return of the service car to the centre and the next coming customer requirement. The same probability distribution function $G(t)$ is valid for both cases, which is justified by the memoryless property of the exponential distribution of interarrival times: $P\{V > t+z \mid V > t\} = P\{V > z\}$ for all $t \geq 0$, $z \geq 0$.

Let $Z(u)$ be the number of customer requirements that arrive into the service centre during a time interval with the length of u time units, $u \geq 0$. The behaviour of the service system over time is then represented by the following transition equation:

$$X_{n+1} = (1-a_n)(X_n+1) + a_n \min\{Z(t(X_n)), L\},\ n \in \boldsymbol{S}. \tag{3}$$

Letting $Z_n \equiv Z_n(X_n, a_n)$ denote the number of customer requirement arrivals within the review interval t_n, we find that $Z_n(X_n, a_n) = a_n Z(t(X_n)) + (1-a_n),\ n \in \boldsymbol{S}$. Note, that if at a decision epoch a 0-action not to depart the service car is taken, then the service centre waits for the next customer requirement arrival, when a new decision on the car departure is made. This results in the arrival of just one customer requirement during such a review interval, the length of which is exponentially distributed with the mean value of $1/l$ time units. The probability mass function for random variables Z_n, $n \in \boldsymbol{S}$, is stated as follows:

$$p_k(i,a) \equiv P\{Z_n(i,a) = k\} = \begin{cases} e^{-l t(i)} [l t(i)]^k / k! & \text{if} \quad a = 1, i \in \{1,2,\ldots,L\}, k \in \{0,1,\ldots\} \\ 1 & \text{if} \quad a = 0, i \in \{0,1,\ldots,L-1\}, k = 1 \\ 0 & \text{if} \quad a = 0, i \in \{0,1,\ldots,L-1\}, k \neq 1 \end{cases} \tag{4}$$

Probabilistic laws of system motion are expressed by stationary transition probabilities:

$$\begin{aligned} p(i,j,a) &= P\{X_{n+1} = j \,|\, X_n = i, a_n = a\} \\ &= \begin{cases} 1 & \text{if} \quad i \in \boldsymbol{X} - \{L\}, a = 0, j = i+1 \\ 0 & \text{if} \quad i \in \boldsymbol{X} - \{L\}, a = 0, j \in \boldsymbol{X} - \{i+1\} \\ p_j(i,1) & \text{if} \quad i \in \boldsymbol{X} - \{0\}, a = 1, j \in \boldsymbol{X} - \{L\} \\ 1 - \sum_{k=0}^{L-1} p_k(i,1) & \text{if} \quad i \in \boldsymbol{X} - \{0\}, a = 1, j = L \end{cases} \quad \text{for } n \in \boldsymbol{S}. \end{aligned} \tag{5}$$

Expected duration of review intervals between consecutive decision epochs is given by stationary stage length functions:

$$\begin{aligned} t(i,a) &= E[t(X_n,a_n) | X_n = i, a_n = a] = E[t_n(X_n,a_n) | X_n = i, a_n = a] \\ &= E[a_n t(X_n) + (1-a_n)V | X_n = i, a_n = a],\ i \in \boldsymbol{X}, a \in \boldsymbol{A}(i), n \in \boldsymbol{S}. \end{aligned} \tag{6}$$

We can specify, that

$$t(i,a) = \begin{cases} E[t(i)] = E[r+si] = r+si & \text{if} \quad i \in \boldsymbol{X} - \{0\}, a = 1 \\ E[V] = \int_0^\infty [1-G(t)]\,dt = 1/l & \text{if} \quad i \in \boldsymbol{X} - \{L\}, a = 0 \end{cases} \tag{7}$$

Let the random variable $C_n \equiv c_n(X_n, a_n)$ represent the cost associated with the system operation in the nth stage, which is incurred during the review interval t_n. Stationary cost functions described hereinafter express economic consequences of the service system activity:

$$\begin{aligned} c(i,a) &= E\left[c(X_n,a_n)\middle|X_n=i,a_n=a\right]=E\left[c_n(X_n,a_n)\middle|X_n=i,a_n=a\right] \\ &\equiv E\left[C_n\middle|X_n=i,a_n=a\right]=E\left[S(X_n,a_n)+W(X_n,a_n)+R(X_n,a_n)\middle|X_n=i,a_n=a\right] \\ &= E\left[S(X_n,a_n)\middle|X_n=i,a_n=a\right]+E\left[W(X_n,a_n)\middle|X_n=i,a_n=a\right]+E\left[R(X_n,a_n)\middle|X_n=i,a_n=a\right] \\ &= S(i,a)+W(i,a)+R(i,a), \quad i\in \boldsymbol{X}, a\in \boldsymbol{A}(i), n\in \boldsymbol{S}. \end{aligned} \tag{8}$$

The cost function $c(i,a)$ consists of a service cost function $S(i,a)$, a waiting cost function $W(i,a)$ and a rejection cost function $R(i,a)$.

The service cost function $S(i,a)$ represents an expected cost associated with a round trip to serve customers:

$$S(i,a)=E\left[a_n(f+mX_n)\middle|X_n=i,a_n=a\right]=E\left[a(f+mi)\right]=a(f+mi),\ i\in \boldsymbol{X}, a\in \boldsymbol{A}(i), n\in \boldsymbol{S}. \tag{9}$$

The cost coefficient m is a cost incurred to serve a customer and the constant f is a fixed cost spent by a round trip.

The waiting cost function $W(i,a)$ reflects (expected) expenses accompanying situations when customers are to wait for service. Let $x(t)$ be the number of customer requirements registered in the service centre at time t after the current decision epoch T_n, when $t = 0$ and $x(0) = (1 - a_n)X_n$. Thus, $x(0) = 0$ if the action taken in time T_n is $a_n = 1$. Considering w to express the cost incurred per customer waiting for service per unit time, we can state:

$$W(i,a)=E\left[a_n w\int_0^{t(X_n)} x(t)dt+(1-a_n)w\int_0^V X_n dt \mid X_n=i,a_n=a\right],\ i\in \boldsymbol{X}, a\in \boldsymbol{A}(i), n\in \boldsymbol{S}. \tag{10}$$

Then for $i\in \boldsymbol{X}-\{L\}$:

$$W(i,0)=E\left[w\int_0^V i dt\right]=\int_0^\infty E\left[w\int_0^V i dt\middle|V=u\right]g(u)du=\int_0^\infty E\left[w\int_0^u i dt\right]g(u)du=wil\int_0^\infty ue^{-lu}du=\frac{wi}{l}. \tag{11}$$

To derive $W(i,1)$ for $i\in \boldsymbol{X}-\{0\}$, we introduce non-negative random variables Y_k, $k = 1, 2, \ldots$. The quantity Y_k represents the point in time (measured since the current decision epoch), when the kth customer requirement after the current decision choice arrives into the service centre. This random variable has the Erlang-k distribution with the probability distribution function $G_k(t)$ and the probability density function $g_k(t)$:

$$G_k(t)=1-\sum_{i=0}^{k-1}e^{-lt}\frac{(lt)^i}{i!},\quad g_k(t)=le^{-lt}\frac{(lt)^{k-1}}{(k-1)!},\quad l>0,\ t\geq 0,\ k=1,2,\ldots \tag{12}$$

Take $I(y,t) = 1$ if $y \leq t$ and $I(y,t) = 0$ if $y > t$. Let

$$I_k^i\equiv I(Y_k,t(i))=\begin{cases}1 & \text{if } Y_k\leq t(i)\\ 0 & \text{if } Y_k>t(i)\end{cases}\quad \text{for } k\in\{1,2,\ldots\}, i\in \boldsymbol{X}-\{0\} \tag{13}$$

be an indicator random variable revealing, whether the time point Y_k of the kth customer requirement arrival after the service car departure has occurred before the time $t(i)$, when a round trip to serve i customers is completed. Recall that $x(t) \le L$ for $t \ge 0$. Then we can write:

$$
\begin{aligned}
W(i,1) &= E\left[w\int_0^{t(i)} x(t)dt\right] = E\left[\sum_{k=1}^{L} w(t(i)-Y_k)I_k^i\right] = \sum_{k=1}^{L}\int_0^{\infty} E\left[w(t(i)-Y_k)I(Y_k,t(i))\middle| Y_k = y\right]g_k(y)dy \\
&= \sum_{k=1}^{L}\left\{\int_0^{t(i)} E\left[w(t(i)-y)I(y,t(i))\right]g_k(y)dy + \int_{t(i)}^{\infty} E\left[w(t(i)-y)I(y,t(i))\right]g_k(y)dy\right\} \\
&= \sum_{k=1}^{L}\int_0^{t(i)} w(t(i)-y)g_k(y)dy = w\sum_{k=1}^{L}\left[t(i)G_k(t(i)) - \frac{k}{l}G_{k+1}(t(i))\right] \\
&= w\left\{Lt(i) - \frac{L(L+1)}{2l} + \sum_{j=0}^{L} p_j(i,1)\left[\frac{1}{2l}(L+j)(L-j+1) - (L-j)t(i)\right]\right\}, \quad i \in \boldsymbol{X} - \{0\}.
\end{aligned}
\tag{14}
$$

The rejection cost function $R(i,a)$ presents (expected) expenses associated with nonacceptance of customer requirements. Denoting by b a penalty cost incurred per customer requirement rejected due to the registration limit overrun, we state:

$$
R(i,a) = E\left[a_n b \max\{Z_n(X_n,a_n) - L, 0\}\middle| X_n = i, a_n = a\right], \quad i \in \boldsymbol{X}, a \in \boldsymbol{A}(i), n \in \boldsymbol{S}. \tag{15}
$$

We see that $R(i,0) = 0$ for $i \in \boldsymbol{X} - \{L\}$. If a 1-action to depart the service car is chosen, then:

$$
\begin{aligned}
R(i,1) &= E\left[b\max\{Z_n(i,1)-L,0\}\right] = \sum_{k=0}^{\infty} E\left[b\max\{Z_n(i,1)-L,0\}\middle| Z_n(i,1)=k\right]P\{Z_n(i,1)=k\} \\
&= \sum_{k=0}^{\infty} b\max\{k-L,0\}p_k(i,1) = \sum_{k=L+1}^{\infty} b(k-L)p_k(i,1) = \sum_{k=0}^{\infty} b(k-L)p_k(i,1) - \sum_{k=0}^{L} b(k-L)p_k(i,1) \\
&= bE[Z(t(i))] - bL - b\sum_{k=0}^{L}(k-L)p_k(i,1) = b\left[l\,t(i) - L + \sum_{k=0}^{L}(L-k)p_k(i,1)\right], \quad i \in \boldsymbol{X} - \{0\}.
\end{aligned}
\tag{16}
$$

The controlled stochastic process $\{X(t), t \ge 0\}$, where $X(t)$ is the state of the system at time t, characterized by transition probabilities (5), stage length functions (6), and cost functions (8), is a semi-Markov decision process satisfying the Markovian property:

$$
\begin{aligned}
&P\{X_{n+1} = j | X_n = i, a_n = u, X_{n-1} = i_{n-1}, a_{n-1} = u_{n-1}, \ldots, X_0 = i_0, a_0 = u_0\} \\
&= P\{X_{n+1} = j | X_n = i, a_n = u\} \\
&\text{for each } n \in \boldsymbol{S} = \{0,1,\ldots\}, \text{and for each } T_{n+1}, T_n, \ldots, T_1 \in \boldsymbol{T} = [0,\infty), \text{and } T_0 \equiv 0, \\
&\text{and for all } j, i, i_{n-1}, \ldots, i_1, i_0 \in \boldsymbol{X}, \text{and for all } u, u_{n-1}, \ldots, u_1, u_0 \in \boldsymbol{A}, \\
&\text{with notation } X_n \equiv X_n(T_n) = X(T_n),\ a_n \equiv a_n(T_n),\ T_n \in \boldsymbol{T},\ n \in \boldsymbol{S}
\end{aligned}
\tag{17}
$$

The process $\{X_n, n \in \boldsymbol{S}\}$ is the embedded Markov decision process corresponding to the system evolution described at decision epochs only.

A prescription for selecting actions in all states of the system at some decision epoch is the control rule. Let d_n denote a control rule for taking actions in stage $n \in \boldsymbol{S}$ at decision epoch $T_n \in \boldsymbol{T}$. The sequence $\boldsymbol{d} = \{d_n\}_{n \in \boldsymbol{S}}$ constitutes a (control) policy. The policy is a contingency plan prescribing control actions to be employed in the system states at any decision epoch over the planning horizon. Due to the Markovian property (17) and stationary nature of the system behaviour, a stationary Markov control policy

$$\boldsymbol{d} = \{d_n\}_{n \in \boldsymbol{S}} \text{ with } d_n = d \text{ for all } T_n \in \boldsymbol{T} = [0, \infty) \text{ and } n \in \boldsymbol{S} = \{0,1,\ldots\}, \text{ where } d : \boldsymbol{X} \to \boldsymbol{A}, \quad (18)$$

can be applied for sequential decision making on round trip departures to serve customers. The symbol d is often used to denote such a policy instead of $\boldsymbol{d}$.

The criterion of interest reflecting economic consequences of the policy application is stated in the form of conditional expectation:

$$A^d(i) = \lim_{t \to \infty} E\left[\frac{C(t)}{t} \middle| X_0 = i, d\right], \quad i \in \boldsymbol{X}, \quad (19)$$

where $C(t)$ is the total cost incurred up to time $t \geq 0$. The quantity $A^d(i)$ represents the long-run expected average cost per unit time given that the initial state is i and the stationary Markov policy d is applied. Using an alternative criterion formulation and noting that $C_n \equiv c_n(X_n, a_n)$, $t_n \equiv t_n(X_n, a_n)$, $a_n = d(X_n)$ and $p_i^d(j) = \lim_{m \to \infty} \left(1/m\right) \sum_{n=0}^{m-1} P\{X_n = j \mid X_0 = i, d\}$, we get:

$$A^d(i) = \lim_{m \to \infty} \frac{E\left[\sum_{n=0}^{m-1} C_n \middle| X_0 = i, d\right]}{E\left[\sum_{n=0}^{m-1} t_n \middle| X_0 = i, d\right]} = \frac{\sum_{j \in \boldsymbol{X}} c(j, d(j)) p_i^d(j)}{\sum_{j \in \boldsymbol{X}} t(j, d(j)) p_i^d(j)}, \quad i \in \boldsymbol{X}. \quad (20)$$

A stationary Markov policy d^* is said to be optimal, if

$$A^{d^*}(i) = \min_d A^d(i) \quad \text{for all } i \in \boldsymbol{X}. \quad (21)$$

The optimal policy d^* can be revealed by means of computational methods [1], [2] for semi-Markov decision process optimization.

References

1. Tijms, H. C.: *Stochastic Modelling and Analysis: A Computational Approach*. Wiley, Chichester, 1986
2. Puterman, M. L.: *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, New York, 1994
3.

Acknowledgement

This work was supported by Slovak Scientific Grant Agency under grant No. 1/7211/20.

Ondrej Bartl

Faculty of Management Science and Informatics

University of Žilina

Velký diel, 010 26 Žilina, Slovakia

tel.: +421-89-5134133, tel./fax: +421-89-5254613

e-mail: bartl@kst.fri.utc.sk

Does the SNA93 Subdivision of the Household Final Consumption into Consumption Expenditure and Actual Consumption Affect the CPI? The Case for Italy*

Buzzigoli L., Ferrari G.
Dipartimento di Statistica "G. Parenti", Università di Firenze

1 Introduction

The Consumer Price Index (CPI) is one of the most relevant tools for macro and micro economic analysis policy and decision. In macro economic analysis, its use as a measure of inflation is commonly accepted. Likewise, its relevance as one of the National Accounts (NA) deflators is widely recognised: on the one hand, policy makers greatly benefit from the information on inflation springing from the CPI figures when taking their decision; on the other hand, national accountants have at their disposal a powerful indicator inside the whole battery of deflators for the construction of constant price accounts. In micro economic ground, again the CPI plays a crucial role in inflation measurement, as it provides businessmen, enterprises, households, with a basic information for individual analysis and decisions.

Despite the many theoretical discussions, methodological controversies and analytic problems related to these uses and the consequent cautions one has to keep in mind when dealing with the above matter and when handling the CPI, there is no question about the substance of what outlined above: the central role held by the CPI in the whole system of Statistics for Economics.

In Italy, the National Statistics Institute (ISTAT) elaborates three CPIs, the so-called "System of consumption price indexes" (ISTAT, 1999):

(a) the "Indice nazionale dei prezzi al consumo per l'intera collettività" (NIC), which refers to generality of consumption of domestic household;

(b) the "Indice dei prezzi al consumo per le famiglie di operai ed impiegati non agricoli" (FOI), which refers to consumption of households whose head is a dependent non farmhand;

(c) the "Indice armonizzato dei prezzi al consumo per i paesi dell'Unione europea" (IPCA), which refers to the generality of domestic households, but restricts its field of observation to consumption of goods and services whose prices are comparable in the different countries of the UE; it covers about 94% of the NIC.

* This paper has benefited from a 40% MURST financial support.

All the above indexes are fixed base Laspeyres type indexes. The price relatives are ratios of monthly prices to base year average prices. The data on weights are represented by the consumption figures elaborated by the NA Division (NAD) of ISTAT and based on Family Budget Survey (FBS) data and on NA information. They are purposively modified and made more detailed by the Price Indexes Division (PID) on the basis of "external" information (Mostacci, 1999a). These weights do not take into account at all or, in some cases, in a very limited way, the distinction between *actual consumption* and *consumption expenditure*, now stated both by the SNA93 and the ESA95. This causes a likely significant bias in the estimates of the indexes at any aggregation level that cannot be disregarded, at least in its methodological aspects.

Yet that said, the aim of this paper is to check how big such a bias may be. This will be done by beginning with the analysis of the existing methodology and by introducing the modifications that allow to estimate the above bias. To this purpose, in paragraph 2 the state of the art of the indexes calculation will be analysed, whereas in paragraph 3 the suggested modifications will be discussed and the related indexes calculated. In paragraph 4 some summarising conclusions will be drawn.

2 The structure of the CPIs

In what follows, we will refer to the NIC. Things for the other two indexes being quite similar, except some details, not much important in themselves and however not influential for understanding of the whole methodological framework, the extension of any considerations to them is straightforward.

2.1 The methodology of construction

All the (a) - (c) above mentioned CPIs, are year *0* (now 1995=100) fixed base monthly Laspeyres type indexes, that is, of the form: ${}_0^pI_{s,t}^L = \sum_{h=1}^{k} \frac{p_{h,s,t}}{p_{h,o}}, w_{h,0}, s = 1,2,...,12; t = 1,..., T$, where $w_{h,0} = \frac{p_{h,0}q_{h,0}}{\sum_{h=1}^{k} p_{h,0}q_{h,0}}$.

The organisation of calculations, from products to national levels, is summarised in Tab. 1.

Starting from January 1999, a chain NIC, as well as a chain FOI, is being calculated by ISTAT, based on December of the previous year (therefore, now, December 1999), the so-called "calculation base", with 1995 as "reference base" (ISTAT, 1999; Ferrari, 1999). For timeliness reasons in data availability and of robustness in calculations, PID uses the weights of two years before. That is, now 1998, whereas the denominators of the price relatives are the prices of December of the previous year, that is, now December 1999.

Table 1 – Aggregation steps for national NIC calculation.

Prod-ucts	Regions										Totals
	1		2		......	i		......	20		
1	$p_{1,1,t}/p_{1,1,0}$	$p_{1,1,0}q_{1,1,0}$	.	.	......	.	.	......	.	.	$\sum_{i=1}^{20}\frac{p_{1,i,t}}{p_{0,i,t}}\frac{p_{1,i,0}q_{1,i,0}}{\sum_{i=1}^{20}p_{1,i,0}q_{1,i,0}}={}_0^pI_{h,t}^L$
.	.	.	.	.	......	.	.	......	.	.	
.	.	.	.	.	......	.	.	......	.	.	.
.	.	.	.	.	......	.	.	......	.	.	.
h	.	.	.	.	......	.	.	......	.	.	.
.	.	.	.	.	......	.	.	......	.	.	.
.	.	.	.	.	......	.	.	......	.	.	.
.	.	.	.	.	......	.	.	......	.	.	.
k	.	.	.	.	......	.	.	......	.	.	
											$\sum_{h=1}^{k}{}_0^pI_{h,t}^L={}_0^pI_t^L$

This circumstance does not allow to take into account the modifications in consumption occurred during the last 12 months, from year *t-2* to year *t-1*. Such a problem does not prevent to calculate the NIC, since the error is likely non significant, as one can reasonably assume that the consumption pattern has remained unchanged after one year. However, in order to partly correct this bias, PID makes resort to a price updating procedure consisting of inflating the quantities at year *t-2*, by evaluating them at the prices of December of year *t-1*. (Mostacci, 1999b)[1].

For 1999, the data on weights were represented by the 1997 provisional ones updated to 1998 by means of any other available sources. Thus, for the above indexes, at 1998, the weights were basically the structure: $w_{h,97}=\frac{p_{h,97}q_{h,97}}{\sum_{h=1}^{k}p_{h,97}q_{h,97}}$, price updated to get: $w'_{h,12,98}=\frac{p_{h,12,98}q_{h,97}}{\sum_{h=1}^{k}p_{h,12,98}q_{h,97}}$. The price relatives at

1 A further drawback arises as the weights are referred to a whole year, whereas the price base is a monthly one, i. e., December: therefore, at least in principle, the weights too should be those of December. Here, even more, for the reasons we have just mentioned, this inconvenience is more apparent than real. This applies also for fixed base NIC. This means that in 1996 the weights were given by the consumption expenditure pattern of 1994, price updated to 1995. But, as soon as the necessary NA and FBS information were available, that is, one year later, in 1997, the weights became those of 1995, the same as for the price relatives.

1999 were: $\frac{p_{h,s,99}}{p_{h,12,98}}$, whereas those at 2000 were: $\frac{p_{h,s,00}}{p_{h,12,99}}$. Therefore, the chained NICs for the 12 months of 1999 have been calculated as follows: $\sum_{h=1}^{k} \frac{p_{h,s,99}}{p_{h,12,98}} w'_{h,12,98}$.

2.2 The price survey

For price relatives calculation, the basket is composed by 558 products, even though roughly 930 products are priced, because of the complex nature of some of them, based on approximately 300,000 monthly quotations collected at municipal level, through a monthly survey conducted by the city statistical offices in no-sampling methods selected outlets. However, all kinds of outlets are included: big and small shops, markets, boutiques, supermarkets, hypermarkets, department stores, groceries, delicatessens, etc[2].

For the various levels of aggregation, that is, for 105 items, 38 groups and 12 categories of goods and services of the COICOP95 Rev.1 classification, the Laspeyres formula is used.

The price collection and the subsequent averaging procedure at the basic headings level do not present particular difficulties, except the customary ones of survey and synthesis, the latter being done now by means of a geometric mean of the price relatives of the quotations of each product. However, the problems that may arise in this context are no concern of this analysis.

2.3 The weighting structure

The weights are basically represented by NA data on annual final consumption expenditure of resident households supplied to the PID by the NAD. The first step performed by the PID is to manipulate them in order to conciliate the classifications adopted by NA, NIC and the Household Budget Survey (HBS), which represents one of the sources, accounting to 25% out of total, used by the NID to estimate the household annual final consumption expenditure, and to fit the coverage of NIC. Subsequently, for every item, the household expenditure relative to year *t-2* is price updated, either to year *t-1*: $p_{h,t-2} q_{h,t-2} \frac{p_{h,t-1}}{p_{h,t-2}} = p_{h,t-1} q_{h,t-2} = w'_{h,t-1}$, for fixed base NIC calculation, or to December of year *t-1*:

$p_{h,t-2} q_{h,t-2} \frac{p_{h,12,t-1}}{p_{h,t-2}} = p_{h,12,t-1} q_{h,t-2} = w'_{h,12,t-1}$, for chain NIC calculation.

The data supplied by the NAD only contains in a very limited way, 0.2% out of total expenditure, the transfers in kind from government to households, that is, all the expenses incurred by it for education,

[2] As a result, the coverage of the indexes does not include any rural areas and therefore they should more properly be viewed as urban indexes rather than as indexes covering the whole country.

health, social security and welfare, sport and recreation, culture. This means that the actual consumption of households for the corresponding goods and services is higher than the incurred expenditure and that, as a consequence, the weights of the related price items are underestimated. This causes a likely bias in NIC that should be estimated.

Before the adoption of the System of National Accounts 1993 (SNA93) and of the European System of Accounts 1995 (ESA95), the problem was, of course, present, but there were no standard indications on how to overcome it[3]. Now, the SNA93 and the ESA95 came into force. They have formally introduced and ruled the subdivision of final consumption into consumption expenditure and actual consumption. Consequently, the weights of the above goods and services can be accordingly revised.

The SNA93 and the ESA95 operate a subdivision within the household consumption concept, explicitly defining the household expenditure and the actual consumption and clearly state the distinctions between the two aggregates. Two concepts of final consumption are identified:

(a) *final consumption expenditure*, that refers to a sector's expenditure;

(b) *actual final consumption*, that refers to sector's acquisition of consumption goods and services.

The differences between these two concepts lies in the treatment of certain goods and services financed by the government or NPISHs but supplied to households as social transfers in kind. More precisely, *final consumption expenditure* consists of expenditure incurred by resident institutional units on goods or services that are used for the direct satisfaction of individual needs or wants or the collective needs of members of the community. *Actual final consumption* consists of the goods or services that are acquired by resident institutional units for the direct satisfaction of human needs, whether (i) *individual* or (ii) *collective*:

(i) goods and services for individual consumption ("individual goods and services") are acquired by a household and used to satisfy the needs and wants of members of that household;

(ii) services for collective consumption ("collective services") are provided simultaneously to all members of the community or all members of a particular section of the community, such as all households living in a particular region.

All household final consumption expenditure is *individual*. By convention, all goods and services provided by NPISHs are treated as *individual*. This means that both household final consumption expenditure and goods and services provided by NPISHs are considered by SNA93 and ESA95 as *actual consumption*. As for the goods and services provided by government units, the borderline between indi-

[3] In fact, both SNA68 and ESA70 did not mention the question, at least in the direction of trying to formally provide users with some indications on how to behave.

vidual and collective goods and services is drawn on the basis of the Classification of the Functions of Government (COFOG).

By convention, all government final consumption expenditures under each of the following headings should be treated as expenditures on *individual consumption services* – and therefore *actual consumption* - except for expenditure on general administration, regulation, research, etc. in each category:

04 Education; 05 Health; 06 Social Security and Welfare; 08.01 Sport and Recreation; 08.02 Culture.

In addition, expenditures under the following sub-headings should also be treated as *individual* when they are important:

07.11 (part of) the provision of housing; 07.31 (part of) the collection of household refuse; 12.12 (part of) the operation of transport system.

The *collective consumption expenditure* is the remainder of the government final consumption expenditure.

3 The suggested NIC's modifications

To estimate the bias in monthly NICs due to weights underestimation and to calculate the revised indexes, our reasoning has been as follows.

The first step being to impute the government final consumption expenditure that should be considered as individual consumption services to the concerned items, an analysis of the list of the 558 products of the basket showed that five of them would include some share of the above final consumption expenditure, i. e. (quoting, respectively, ISTAT code and COICOP95 Rev. 1 classification code): 313, 5950 Medicinal; 347, 111 Clinics; 480, 8010 Secondary school; 481, 8020 University; 546, 7610 Nurseries.

As far as the first item is concerned, the PID already adds the related government consumption expenditure (the social transfers in kind, as above said accounting for 0.2%), as a part of the category of "Social contribution in kind", supplied by the NAD, to the consumption expenditure, in order to get the actual final consumption. As for the remaining four items, things are more complicated, since the NAD provides the necessary information to convert the consumption expenditure into actual consumption – again through the imputation of shares of the above aggregate - only for the part of the expenditure incurred by the private bodies. As far as the expenditure incurred by the public bodies is concerned, by far more relevant than the previous one, no information is available to the PID. As a result, the weights of the above four items are underestimated and must be revised upwards. This has been done by making resort to the "Actual collective consumption" estimated by the NAD and accounting for 148,273 billions lira in 1998 (Collesi-Di Leo, 2000). No indications are available by the NAD that might allow to try to consistently impute the shares of it to the concerned items. Therefore, this amount has been imputed to the four items, by subdividing it among them according to assumptions based upon the general information about the actual amount of the government expenditure within the total expenditure of each of them one might have. Three sets of hypotheses of subdivision have been outlined, as shown in Tab.2.

The first hypothesis makes it operative a simple and quite rough statement: that the government social contributions to households are spread "quasi-uniformly" over the different social situations, basically represented by education and health/social activity. In other words, it postulates a sort of "neutral" policy of government.

Table 2 – Hypotheses of subdivision of the "Actual collective consumption" among the 4 items.

	347, 111 Clinics		480, 8010 Secon-		481, 8020 University		546, 7610 Nurs-		Total	
	%	Amount	%	Amount	%	Amount	%	Amount	%	Amount
Hypothe-	20	29,655	25	37,068	30	44,882	25	37,068	100	148,273
Hypothe-	10	14,827	40	59,309	40	59,309	10	14,828	100	148,273
Hypothe-	20	29,665	35	51,896	30	44,882	15	22,240	100	148,273

The second hypothesis is much more "Education" oriented, as it assumes that the most relevant share of the social contributions is allocated to this sector: in fact, 80% of them is equally (40% and 40%) attributed to Secondary school and University, whereas the remaining 20% is equally (10% and 10%) attributed to health/social activity. Finally, the third hypothesis is a refinement of the first one, giving more emphasis to the Secondary school (35%), to the prejudice of Nurseries (15%).

Needless to say, although appearing the most reasonable and, in a sense, paradigmatic ones, the above mentioned are not exhaustive but only a part of the many hypotheses one can propose.

These amounts have been added to the weights used by the PID (respectively, 15,663; 821; 3,681; 5,868) to obtain the revised weights. Then, the entire set of relative weights has been re-calculated, for each of the above hypotheses. Finally, the 1999 monthly NICs have been calculated, both in chained case, base December 1998 and in 1995 fixed base case. The results of these revisions are shown in Tab. 3 , together with the ISTAT estimates.

Tab. 3 – 1999 ISTAT and our own estimates of monthly NICs, both in calculation and reference base.

	January	Febru-	Marc	April	Ma	Jun	July	Au-	Sep-	Octo-	No-	Decem-	Aver-
Calculation base, December 1998=100													
ISTAT	100,1	100,3	100,5	100,7	10	100	101,2	101,3	101,5	101,8	102,0	102,1	101,1
Revised	100,1	100,3	100,5	100,7	10	100	101,2	101,2	101,5	101,8	102,0	102,2	101,1
Revised	100,1	100,3	100,5	100,8	10	101	101,5	101,6	102,0	102,4	102,6	102,7	101,4
Revised	100,1	100,3	100,5	100,8	10	101	101,5	101,6	102,0	102,4	102,6	102,7	101,4
Reference base, 1995=100													
ISTAT*	109,4	109,6	109,8	110,1	11	110	110,5	110,6	110,8	111,1	111,3	111,5	110,4

Revised	110.8	111.0	111.2	111.5	11	111	111.9	112.0	112.3	112.6	112.9	113.1	112.3
Revised	111,3	111,5	111,7	112,0	11	112	112,4	112,5	112,9	113,2	113,4	113,6	112,4
Revised	111.1	111.3	111.5	111.8	11	112	112.2	112.3	112.7	113.0	113.3	113.5	112.2

*These figures, calculated by us, are uniformly slightly higher (0.5-0.6 points) than those published by ISTAT, due to some missing price relatives and weights.

Let's first look at the calculation base indexes. The monthly NICs estimated under hypothesis 1 are uniformly equal to those calculated by ISTAT, except the August figures (the revised one being 0.1 percentage point smaller than the ISTAT one) and the December figures, with the revised estimation being 0.1 percentage point bigger than the ISTAT one. Revision 2 gives much different results: until March, the two sets of NICs are the same; from April to June, the revised NICs are 0.1 percentage points higher than the ISTAT ones. This difference, raises to 0.3 - 0.6 percentage points, quite progressively, from July to December. Exactly the same results as those of Revision 2 are provided by Revision 3.

As for the reference base NIC, the differences are much more evident. Revision 1 estimates are sensibly higher than ISTAT ones (1.4 to 1.6 percentage points). Revision 2 estimates emphasise this tendency, showing 1.9 to 2.1 percentage points higher figures. Unlike calculation base case, Revision 3 gives different, even though only slightly, results than those provided by Revision 2, as they exhibit 1.7 to 2.0 percentage points higher figures.

4 Conclusions

In the paper, we have analysed the state of art of calculation of the set of price indexes as carried out in Italy by ISTAT. Our initial assumption was that, due to the underestimation of the weights of the items whose expenditure is shared by government trough social contributions, the CPI might suffer from a bias and be either underestimated or overestimated, according to the relationship between price relatives and revised weights. Because of lack of information, and having at our disposal the total amount of the government actual collective consumption and all the data supplied by the NAD to the PID only, we have checked three alternative hypotheses of distribution of this aggregate over the four items whose household expenditure is supported by government under the form of social contributions in kind. We have shown that the monthly NICs, and therefore, the other two indexes, FOI and IPCA - the three CPIs elaborated by ISTAT - are actually biased downwards, both in calculation and in reference base. Some hypotheses have evidenced a bigger bias, some a lower one: but all of them have shown that the revised monthly NICs are higher or at least equal to those calculated by ISTAST without taking into account the government actual collective consumption. This is true both for the calculation base NICs and, even more notably, for the reference base NICs, where the differences between the revised figures and the official ones are very high, likely because of some accumulation effect.

To conclude, based upon the analysis performed in this paper, there is empirical evidence that the official CPIs are biased and that the size of the bias, and also its direction, depend upon the share of the amount of government final consumption expenditures that can be treated as actual consumption and upon its relevance inside the related goods and services. Of course, this is only a first and quite rough attempt to estimate the bias size. Further and more accurate analyses, based on more detailed official data are needed to get definitive and reliable results.

REFERENCES

1. Collesi, D., Di Leo, F. (2000), Spesa per consumi e consumi individuali, paper presented at the *Seminario "La nuova Contabilità Nazionale"*, ISTAT, Roma, 12-13 gennaio.
2. Ferrari, G. (1999), Chain Consumer Price Indices in the Measurement of Inflation: Theoretical Statements and Empirical Evidence, *Proceedings of the Measurement of Inflation Conference*, Eds. M. Silver and D. Fenwick, Cardiff, UK, August 31-September 1.
3. ISTAT (1999), Note Rapide, Roma, 15 marzo.
4. Mostacci, F. (1999a), How to Derive a System of Weights for CPI from National Accounts Estimates, *Proceedings of the Measurement of Inflation Conference*, Eds. M. Silver and D. Fenwick, Cardiff, UK, August 31-September 1.
5. Mostacci, F. (1999b), Aspetti teorici ed implicazioni pratiche nel passaggio dall'indice a base fissa all'indice a catena per i prezzi al consumo, internal paper, ISTAT, Roma, mimeo.

Buzzigoli L., Ferrari G.

Dipartimento di Statistica "G. Parenti",

Università di Firenze

Viale Morgagni, 59, 50134 Firenze, Italia

tel.: +39 055 4237258; fax: +39 055 4223560; E-mail: lucia@ds.unifi.it.

tel.: +39 055 4237221; fax: +39 055 4223560; E-mail: ferrari@ds.unifi.it

System of models for forecasting, optimisation and analysis of public funds revenues and expenditures

Michal Fendek
Department of Operations Research and Econometrics
University of Economics Bratislava

In the paper we analyse the up to the present development of revenues and expenditures of the individual types of public funds in the Slovak Republic during 1995-1999, then we identify and quantify its dependency on the macroeconomic indicators and on that base we forecast the development of public funds revenues and expenditures over the years 2000-2004. The stress is laid upon the methodological questions of described analyses and forecasts realisation.

1. Introduction

The Slovak economy development has run over the past few years in a quite complicated way and we cannot be satisfied with achieved results, occasionally. The Slovak government, experts on macroeconomic policy issues and also general public have been pushing forward the acceptance of such system measurements that will direct Slovak economy on the way of a long-run positive development.

In this connection we expect the improvement of control and management in public finance together with an increase of public finance tools effectiveness, because the development in this area of public finance has been one of the major sources of macroeconomic imbalance in the Slovak economy over the past years.

2. Models of public funds revenues and expenditures

System of models for forecasting, optimisation and analysis of public funds revenues and expenditures creates open three-level hierarchical structure of MS Excel models. It is a result of a nearly two-year co-operation of experts from Department of Operational Research and

Econometrics and Ministry of Finance and accounts for a modelling supporting apparatus assigned for a quantitative analysis of public funds management efficiency.

The first model designed for decision-making support in public funds revenues and expenditures analysis was developed at the Department of Operational Research and Econometrics in the year 1998. This early version of the model has constructed only short-run forecasts without any possibility to optimise the effects of used macroeconomic policy tools. The model has been later re-done as well as amended and its latest version "2000" comprises the following extensions and improvements:

a) the model enables to find optimal values of macroeconomic indicators on the basis of pre-defined goals for selected endogenous variables
b) a new sub-model for forecasting, analysis and optimisation of public funds expenditures, that analyses the structure of Pensions Security Fund expenditures, has been developed
c) the model extends the horizon of prognosis till the year 2004, i.e. the model performs a medium-term forecast of public funds revenues and expenditures

3. Forecasting, optimisation and analysis of public funds revenues

The base, or let us say, a starting point for public funds revenues and expenditures modelling is a forecast of elementary macroeconomic indicators and their subsequent optimisation accepting defined macroeconomic policy goals.

The historical development along with the prognosis of the Slovak Republic macroeconomic indicators are quoted in the model named *Macroeconomic indicators* that contains the following information:

— the historical development of indicators during the years 1993-1999
— the forecast of annual indicators for the years 2000-2004 that has been worked out on the basis of the forecast for quarterly periods
— the forecast of quarterly values of indicators for the years 2000-2004 that is a result of the prognostic application of the Slovak economy econometric model, resp. is an optimal solution of object programming task for given target values of selected macroeconomic indicators

The model approaches the following macroeconomic indicators:

HDP-P – gross domestic product (aggregate supply) forecasted through the use of production function

HDP – calculated as a sum of SS, VS, THK and SZO in billion SKK

SS – private consumption (consumption of households) in billion SKK

VS – public consumption (consumption of government) in billion SKK

THK – gross capital formation in billion SKK

VTS – exports of goods and services in billion SKK

DTS - imports of goods and services in billion SKK

SZO – foreign trade balance in billion SKK

PUMTV - average interest rate of time deposits, in %

PUMUSS - average interest rate of credits in private sector, in %

USDSK - USD/SKK exchange rate

DPR - import surcharge, in %

M1 - money supply M1, in billion SKK

ISC95 - consumer price index

CPO - overall incomes of households, in billion SKK

DIPO - disposable incomes of households, in billion SKK

D - total tax incomes, in billion SKK

ZAM - level of employment, in million of employees

NEZAM - number of unemployed, in million

EAOB - economic activity of the population, in million

OBYVAT - population, in million

W - average gross monthly wage, in SKK

TRW - rate of average gross monthly wage growth

PP - productivity of labour, in SKK

TRPP - growth rate of productivity of labour

MN - unemployment rate, in %

MINF - inflation rate measured by consumer prices index, in %

TRHDP - growth rate of GDP in fixed prices of the year 1995

The results of performed macroeconomic forecasting are shown in the Table 1. The modelling of revenues and a number of payers to public funds is ensured by the model FONDYPRJ that creates the forecasts of indicators of this particular part of public funds with a support of econometric models, that come out from the optimised values of macroeconomic indicators. At the same time they allow – by means of scenarios, whose parameters are set by users themselves - to create the alternative forecasts of indicators and accomplish a comparison analysis.

Submodels of system:

- Health insurance – ZP
- Social insurance – sickness – NP
- Social insurance –pension – DZ
- Unemployment insurance - FZ
- Public funds – revenues

Tab. 1

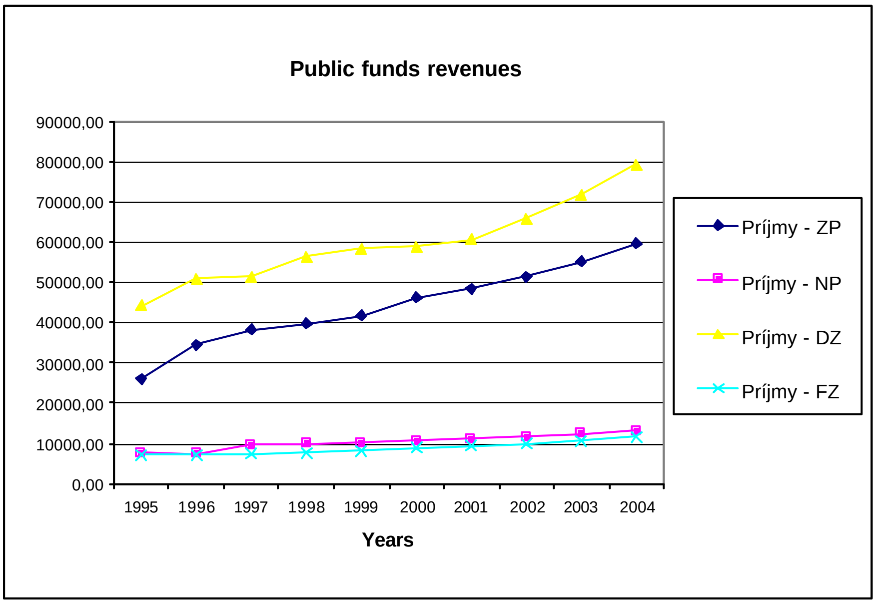
A. Development of basic macroeconomic indicators

						Annual forecast				
Indicators	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
PROGNOSIS										
HDP - P						908.19	955.61	1000.82	1044.47	1088.12
HDP	515.10	575.70	653.90	717.40	779.30	885.91	957.83	1046.80	1155.86	1288.23
SS	252.60	286.10	322.30	360.10	395.40	428.54	457.25	495.11	540.80	595.51
VS	104.00	129.70	143.50	151.80	156.70	176.85	154.56	173.77	184.82	207.45
THK	150.60	212.70	252.70	292.40	257.30	300.71	356.19	412.85	465.85	518.79
VTS	325.80	334.00	368.80	456.80	504.90	562.77	613.44	632.84	680.31	734.75
DTS	316.40	403.20	415.50	536.90	544.40	582.97	623.62	667.77	715.91	768.27
SZO	9.40	-69.20	-46.70	-80.10	-39.50	-20.20	-10.18	-34.93	-35.60	-33.52
PSR	163.14	166.33	180.83	177.83	171.11	184.61	196.00	212.27	228.59	248.52
VYSR	171.44	191.89	217.83	197.03	191.43	212.45	206.21	221.48	243.95	269.42
PUMTV	12.22	9.10	10.52	12.97	13.23	12.20	11.25	10.94	10.66	10.38
PUMUSS	16.22	14.42	15.56	17.78	17.92	18.15	17.40	17.32	17.74	18.08
USDSK	29.73	30.65	33.76	35.23	41.64	42.25	41.40	41.00	40.50	40.00
DPR	10.00	8.75	3.50	2.75	3.50	4.00	3.00	0.00	0.00	0.00
M1	128.73	154.65	158.75	150.20	162.84	184.67	197.80	211.79	226.57	244.60
ISC95	1.00	1.06	1.12	1.20	1.37	1.50	1.61	1.72	1.82	1.92
CPO	402.68	466.98	528.34	619.65	715.56	845.45	916.40	1020.13	1140.16	1289.46
DIPO	345.25	403.40	468.11	520.81	641.49	747.96	809.01	901.56	1001.56	1125.41
OW3	43.57	49.77	56.14	60.99	64.75	72.87	79.02	86.59	95.91	107.24
D	145.06	151.40	157.27	165.65	170.60	202.67	217.99	237.29	263.50	296.82
ZAM	2.02	2.04	2.03	2.03	2.01	2.04	2.07	2.11	2.15	2.20
NEZAM	0.35	0.32	0.34	0.38	0.44	0.43	0.40	0.39	0.37	0.34

EAOB	2.37	2.36	2.37	2.41	2.45	2.47	2.47	2.49	2.51	2.53
OBYVAT	5.37	5.39	5.42	5.44	5.45	5.45	5.48	5.50	5.52	5.54
W	7190.50	8147.25	9222.25	10003.75	10733.25	11895.54	12690.26	13704.50	14887.18	16266.14
TRW	1.143	1.133	1.132	1.085	1.073	1.290	1.269	1.277	1.251	1.282
PP	255018.94	282701.30	322253.1	353029.46	387531.47	434298.63	462353.01	497158.35	538306.74	586390.11
TRPP	1.14	1.11	1.14	1.10	1.10	1.35	1.31	1.28	1.24	1.27
MN	14.76%	13.74%	14.23%	15.85%	18.07%	17.43%	16.25%	15.57%	14.57%	13.25%
MINF	9.89%	5.78%	6.13%	6.72%	13.99%	9.57%	7.72%	6.94%	5.66%	5.64%
TRHDP	6.22%	5.66%	7.03%	2.81%	-4.70%	3.75%	0.37%	2.20%	4.51%	5.51%

The results of public funds revenues forecasting are shown in Table 2 and their illustration is in Fig. 1.

Fig.. 1



Tab. 2

E. Development of public funds revenuues – totally

						Annual Forecast				
Indicators	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
PROGNOSIS										
Public funds revenues										
Reven. – ZP	26031.44	34436.16	38322.92	39715.34	41690.32	46179.51	48422.97	51361.90	55100.02	59690.51
ZP/HDP	5.05%	5.98%	5.86%	5.54%	5.35%	5.21%	5.06%	4.91%	4.77%	4.63%
Reven. – NP	7693.10	7362.71	9575.58	9816.75	10128.24	10797.42	11119.50	11690.69	12327.04	13132.44
NP/HDP	1.49%	1.28%	1.46%	1.37%	1.30%	1.22%	1.16%	1.12%	1.07%	1.02%
Reven. – DZ	44242.51	50879.08	51510.57	56302.15	58352.57	58930.49	60712.92	65864.53	71830.61	79216.01
DZ/HDP	8.59%	8.84%	7.88%	7.85%	7.49%	6.65%	6.34%	6.29%	6.21%	6.15%
Reven. – FZ	7095.91	7184.96	7509.99	7645.32	8251.02	8840.10	9347.25	9982.79	10766.02	11718.44
FZ/HDP	1.38%	1.25%	1.15%	1.07%	1.06%	1.00%	0.98%	0.95%	0.93%	0.91%
Reven.	85062.95	99862.91	106919.0	113479.5	118422.1	124747.5	129602.6	138899.9	150023.6	163757.4

Reven./HDP	16.51%	17.35%	16.35%	15.82%	15.20%	14.08%	13.53%	13.27%	12.98%	12.71%
Collection to public funds										
Collec. – ZP	18726.89	23778.37	27584.52	28754.36	30081.52	34170.84	35660.26	37969.12	40886.63	44467.30
Collec. – NP	7587.70	7196.26	9367.44	9500.21	9737.61	10343.62	10572.50	11016.86	11560.99	12247.30
Collec. – DZ	39088.11	46853.17	50261.72	54391.95	56167.62	56405.81	57711.99	62396.83	67926.46	74754.47
Collec. – FZ	7095.91	7184.96	7509.99	7645.32	8251.02	8840.10	9347.25	9982.79	10766.02	11718.44
Collec.	72498.60	85012.75	94723.67	100291.8	104237.7	109760.3	113292.0	121365.6	131140.0	143187.5

3. Forecasting, optimisation and analysis of public funds expenditures

The modelling of public funds expenditures and a number of insurers is realised by the model FONDYVYDS that develops the forecasts of indicators describing this part of public funds with the use of econometric models resulting from the optimised macroeconomic indicators values. The constructed models enable – through scenarios whose parameters are set by the user himself – to create the alternative forecasts of indicators and perform various comparison analysis. This model consists of the following submodels as follow:

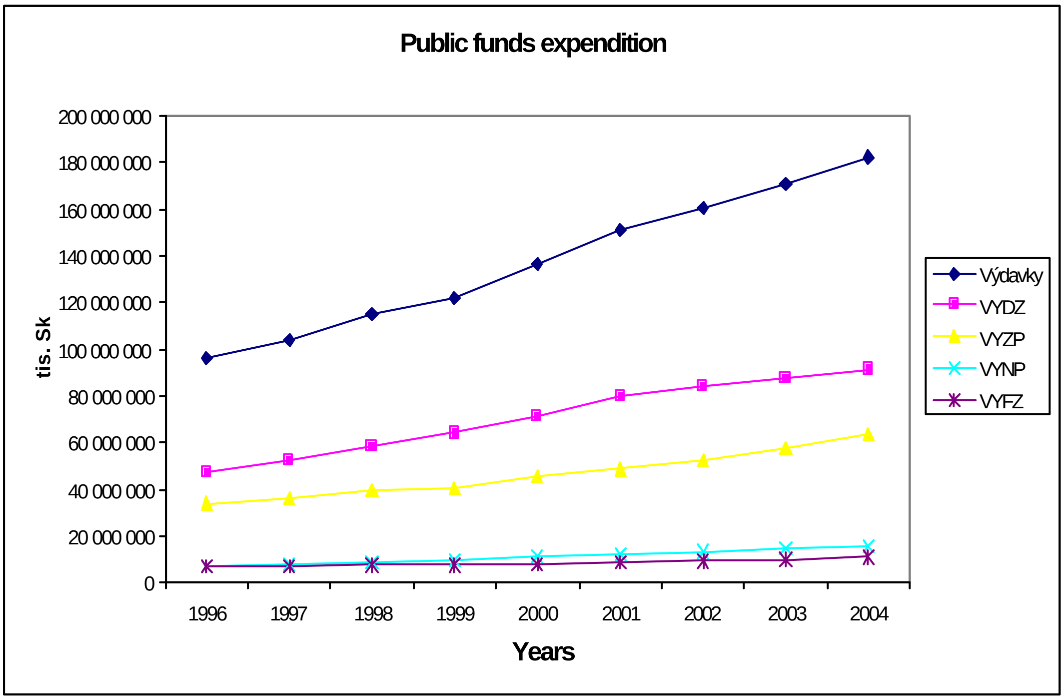
- Funds
- Pensions
- Pensions – Scenario

The achieved results of public funds expenditures prediction in the breakdown according to the individual types of public funds are shown in Table 3 and Fig. 2.

Tab. 3

B. Development of public funds expenditures									
					Annual Forecast				
Indicator	1996	1997	1998	1999	2000	2001	2002	2003	2004
Expenditures	96322949	104040463	114934239	122211358	136455046	150810211	160236802	170687195	182275418
VYDZ	47390593	52494639	58469694	64150874	71599903	80290449	84261713	88032427	91460326
VYZP	34138964	36429758	39723678	40797913	45577415	48878588	52849905	57720078	63612685
VYNP	7381026	8115437	8973745	9496270	11458139	12653170	13676510	14727995	16005926
VYFZ	7412366	7000629	7767122	7766301	7819590	8988004	9448674	10206695	11196480
Structure of public funds expenditures (share of total expenditures)									
VYDZ/VY	49.20%	50.46%	50.87%	52.49%	52.47%	53.24%	52.59%	51.58%	50.18%
VYZP/VY	35.44%	35.01%	34.56%	33.38%	33.40%	32.41%	32.98%	33.82%	34.90%
VYNP/VY	7.66%	7.80%	7.81%	7.77%	8.40%	8.39%	8.54%	8.63%	8.78%
VYFZ/VY	7.70%	6.73%	6.76%	6.35%	5.73%	5.96%	5.90%	5.98%	6.14%
Structure of public funfs expenditures (share of GDP)									
Výdavky/HDP	16.73%	15.91%	16.02%	15.68%	15.40%	15.74%	15.31%	14.77%	14.15%
VYDZ/HDP	8.23%	8.03%	8.15%	8.23%	8.08%	8.38%	8.05%	7.62%	7.10%
VYZP/HDP	5.93%	5.57%	5.54%	5.24%	5.14%	5.10%	5.05%	4.99%	4.94%
VYNP/HDP	1.28%	1.24%	1.25%	1.22%	1.29%	1.32%	1.31%	1.27%	1.24%
VYFZ/HDP	1.29%	1.07%	1.08%	1.00%	0.88%	0.94%	0.90%	0.88%	0.87%

Fig.2



5. Conclusion

The new system of models supporting the analysis, forecasting and optimisation of public funds revenues and expenditures has been presented in the paper and comprises the models of the following funds:

- Health Insurance Fund
- Pensions Security Fund
- Sickness Insurance Fund
- National Labour Office Fund,

In the first version of the above described model we have focused on the solution of the social security topic having been indeed an up-to-date and pressing issue under the current conditions in the Slovak Republic. In the next phase the system of models will be enriched about partial sub-models of expenditures of:

- Health Insurance Fund

- Sickness Insurance Fund
- Unemployment Insurance Fund

what enables the Treasury analysts to accomplish a comprehensive quantitative analysis of public funds revenues and expenditures.

6. References

1. Fendeková, E., Fendek, M.: Optimisation of fiscal policy instruments under regulation of natural monopoly**.** In: Zbornik z medzinárodnej konferencie MME 2000., University of economics, Praha 2000.
2. Fendek, M.: Využitie optimalizacných a ekonometrických modelov pri analýze efektívnosti nástrojov hospodárskej politiky. INFOMATIKA '97. IV. konferencia s medzinárodnou úcastou, Slovenská spolocnost aplikovanej informatiky a kybernetiky, Bratislava november 1997.
3. Fendek, M., M. Hatrák, V. Mlynarovic: Sústava modelov pre prognózovanie, optimalizáciu a analýzu príjmov a výdajov verejných fondov. Štúdia pre Ministerstvo financií SR, Bratislava, 2000.
4. Fendek, M. - Mlynarovic, V. - Hatrák, M.: Modelovanie príspevkov do poistných fondov. Ekonomický casopis - Journal of Economics, 1999, No. 3., Academic Press Bratislava
5. Feješová, V.: Systém zdravotného poistenia, sociálneho poistenia a poistenia v nezamest-nanosti a jeho väzby na ŠR. Ministerstvo financií SR, Bratislava 1997.
6. Návrh strednodobého financného výhladu na roky 2000 – 2003. Ministerstvo financií SR. Bratislava, apríl 2000.

Prof. Ing. Michal Fendek, CSc.
Katedra operacného výskumu a ekonometrie
Ekonomická univerzita
Dolnozemská 1
852 35 Bratislava
tel.: 07 67295 833
e-mail: fendek@dec.euba.sk

Estimating the Cost of Children for Divorced Parents in Tuscany

G. Ferrari, M. Maltagliati[*]

Dipartimento di Statistica "G. Parenti", Università di Firenze,

1. Introduction

This research sets in the framework of the studies whose object is the estimation of what is called the"*cost*" of children, which does not refer – as one might at first to deem – to an intrinsic concept, that only depends upon the characteristics of children, but expresses a conceptual category that regards in principle the characteristics of the household as well.

In this paper, we are specifically interested in the cost supported by the two parents, so that we would speak instead of *economic resources that the parents must devote to children support*. In general, rich parents, due to their higher resources, would spend, for supporting their own child, or children, a bigger amount of money than that spent, *ceteribus paribus*, by poor parents. In the continuation of this article, we will use the expression "cost of child (or children)", as it is now the one commonly accepted.

The knowledge of the cost of children is of a great utility in many concrete circumstances[1]. The objective of this research is to determine the *amount of money* ***judicially separated*** *or* ***divorced parents*** *must spend to support their children*, or, in other words, the *cost of children* for these kind of couples.

In case of judicial separation or divorce, the cost of children must be charged to the two parents proportionally to their means, that is, <u>the income of each of them</u> and to <u>the level of wealth one wants to secure to the child or the children</u>.

The breaking off of the couple intercourse and the subsequent separation and, eventually, divorce, in itself already a difficult moment in married life, may become a tearing event when there are some children. The pre-existent affective relation transfers into a judicial controversy and it becomes a judge's task to grant custody of children and the share of income the parent to whom custody of children is granted must be paid by the other parent. In these cases, it is a judge crucial need to know as more precisely and objectively as possible what the cost of children might be, which, as above outlined, depends upon many variables, like the <u>household wealth level</u> and the <u>age of children</u>. Thus, by taking advantage

[*] Financial support by the Italian MURST 40% funds is gratefully acknowledged.

[1] For example, when the government wants to carry out social policy measures in favour of less wealthy households, by means of some form of support to their expenses for children. Or when it wishes to carry into effect economic policy measures of re-distributive and/or fiscal type, aiming at reaching the equilibrium of household wealth.

of any available evidence, the judge will be in a position to share the children burden between the parents, with the utmost care, in proportion to the parents' wealth.

The estimate of the cost of children is not an easy task at all, as it is not possible to isolate the economic effect of the presence of children in a household[2]. Therefore, in order to get it, one must make resort to some model based technique. In this work, we will avail ourselves of a methodology that allows us to estimate the so-called Equivalence Scales (ES), on the basis of the *Price Scaled-Generalised Almost Ideal Demand System* (PS-GAIDS), belonging to the class of the so-called Demographically Extended Complete Demand Systems (DECDS).

The ESs so estimated will be used to calculate the income share the parent to whom custody of children has not been granted must transfer to the parent to whom custody has been granted to support children, taking into account the income of each of them under different hypotheses of combination of the two incomes as well as some different assumptions regarding the wealth we want to secure to the children and to the parent to whom custody is granted (however difficult, as we shall see, the distinction between the two wealths might be).

Thus, the opening of the paper will be devoted to an illustration of the problems of ES construction, together with a brief analysis of the PS-GAIDS used for their estimate, as well as an appraisal of the data used and of the results obtained (par. 2). Subsequently, we will move to the calculation of the "Equivalent Incomes" (EI) for each parents (par. 3). The paper will be concluded by a synthetic exam of the results of the research and by some general remarks about the most significant evidence (par. 4).

2. The Equivalence Scales: The model, the data and the parameters estimates

The ES are expenditure *deflators* that are used to calculate the relative amount of money two different types of households need to attain the same wealth level. Since they are essentially cost of living indexes, they are defined through the concept of *cost function*, which provides the minimum cost supported by a household to attain a given level of welfare (Muellbauer, 1977). Therefore, one can write $e = C_c / C_r$,where $\boldsymbol{e}$ indicates the ES and C_c and C_r are, respectively, the cost of the comparison household and that of the referring one.

[2] This is mainly due to several reasons like the <u>collective nature</u> of some expenses, the difficulties to attribute some particular expenses not of a collective type to the individual members of the household and therefore, to children and the modification in habits and tastes of the parents due to the presence of children.

One of the crucial points of this approach consists of understanding when two different households can be considered as sharing the same utility level[3]. As above said, a convincing way of constructing ES consists of deriving them from DECDS, that is, of inserting demographic variables – or, in a more restricted view, of incorporating the household composition effects – into a theoretically plausible complete demand system and therefore to derive the ES directly from the utility theory.

The DECDS allows to take into account *all* goods and services and express the expenditure shares as a function of the income of households, of relative prices and of the characteristics of households (number of components and their composition by age, sex, etc.).

Thus, an ES is an index of the following type $e = \frac{C_c}{C_r} = \frac{C_{0c}(U_0, \mathbf{p}, H_c)}{C_{0r}(U_0, \mathbf{p}, H_r)} = f(U_0, \mathbf{p}, H_c - H_r)$, that is, it depends upon utility level U_0, vector price $\mathbf{p}$ and demographic characteristics of household H_c as compared to those of household H_r.

In this paper, to introduce the demographic variables into the demand model, we will use the *Price Scaling* (PS) (Ray, 1983), based on a simplified version of the DS, the *Restricted Demographic Scaling* (RDS), sometimes also called *Engel Scaled*. In PS, a general form for the ES, m_c, is suggested

$$m_c(\mathbf{p}, H_c) = \left(1 + \sum_s q_{sc} h_{sc}\right) \prod_i p_i^{d_i h}, \quad \sum_i d_i = 0 \tag{1}$$

where h_{sc} is the number of comparison household components in age class s who exceeds the corresponding values in referring household and $h = \sum_s h_{sc}$ is the number of comparison household components less the number of components of referring household. This scale is a so-called Equivalent Scale Exact (ESE); in other words, it does not depend upon the household income and the utility level[4].

If the above PS model is inserted into the Generalised Almost Ideal Demand System (GAIDS), belonging to the Rank 3 class models (see Deaton-Muellbauer, 1980; Lewbel, 1991)

$$w_i = a_i + \sum_j g_{ij} \log p_j + b_i \log\left(\frac{x}{P}\right) + l_0 l_i \prod_k p_k^{(l_k - b_k)} \left[\log\left(\frac{x}{P}\right)\right]^2, \quad i, k = 1, \ldots, n, \tag{2}$$

after some algebra, one gets the PS-GAIDS, that is, the DECDS used in this paper

[3] To this purpose, many indicators have been suggested: Engel's view was to look at the food share; according to Rothbarth, one should look instead at the absolute amount of expenditure on selected goods, whereas the subjective approach claims to directly ask the respondents their perceived degree of wealth, and so on.

[4] In other words, one in which the relative increase in expenditure caused by a child is the same both for rich and poor households and the utility level U_0 is unimportant.

$$w_i = d_i h + a_i + \sum_j g_{ij} \log p_j + b_i \log\left(\frac{x}{P}\right) + l_0 l_i \prod_k p_k^{(l_k - b_k)} \left[\log\left(\frac{x}{P}\right)\right]^2 ; \; i,k = 1,..., n \tag{3}$$

where $\log P = a_0 + \sum_i a_i \log p_i + \frac{1}{2}\sum_i \sum_j g_{ij} \log p_i \log p_j + \log\left(1 + \sum_{sc} q_s h_{sc}\right) + \sum_i d_i h \log p_i$.

In our case, we have introduced some simplifications. We consider one period only, the biennium 1995-96, within which the prices **p** do not vary. Moreover, the referring household is the one composed by a 39 year old single, with age selected on the basis of the observed average age in Tuscany.

We have specified the age class of the children, as follows. Number of components: in age class 0-2 years = η_1; in age class 3-5 years = η_2; in age class 6-10 years = η_3; in age class 11-14 years = η_4; in age class 15-18 years = η_5; in age class 19-25 years = η_6.

Thus, the final version of the PS-GAIDS we have estimated is the following

$$w_i = a_i + b_i \log\left(\frac{x}{e}\right) + g_i \left[\log\left(\frac{x}{e}\right)\right]^2 , \quad i,k = 1,..., n . \tag{4}$$

The data base used in this analysis is the pooled cross section-time series unit record data taken from the Italian Family Budget Survey (FBS) published by the Bureau of Statistics (ISTAT)[5]. The consumption items, originally 75, have been grouped in 9 commodity groups: *food, tobacco, apparel, housing and energy, furniture, sanitary services, transport, entertainment* and *other goods and services*. The estimates of the parameters of the equation (4) on the above 9 groups of goods and services are Full Information Maximum Likelihood (FIML) estimates, performed on the first 8 equation only. The parameters of the 9th equation (other goods and services), has been calculated by making resort to the constraints of the GAIDS.

3. The cost of children and the EI

To begin with, let's introduce the following notations:

MI_{FM} = monetary income of a household composed by father and mother. As we assume that the child does not have income of his own, this is also the monetary income of a household composed by father, mother and son, MI_{FMS}; MI_S = monetary income of a household composed by a single; ${}_S e_{FMC}$ = equivalence scale which shows how much a household composed by father, mother and child must spend to be as well-off as a household composed by a single.

[5] These records contain a huge amount of information on income, demographic variables and labour supply for more than 33,000 households per year. In our case, as the interest was limited to Tuscan households only, it has been necessary to restrict to them the field of observation and, in order not to dramatically reduce the number of observed cases, it has been decided to consider the Tuscan household surveyed in two consecutive years, 1995 and 1996. This has been done by inflating the 1995 data to 1996 via the CPI. In this way, we were able to estimate the model on about 2,500 households.

Definition 1: the Equivalent Income (EI) of a household is the ratio between its monetary income and the ES, that is, $EI = MI/h$.

Therefore, the income of a comparison household is equivalent to that of the referring one if and only if ${}_R e_C = 1$. Then, the EI is the income the comparison household needs to be as well-off as the reference household.

The above ratio, despite not representing a cardinal measure of wealth, allows the ordering of households according to the degree of wealth attained: the higher the EI is, the higher the wealth will be and, similarly, two household with the same EI will be equally well-off.

We must now state how much a child does cost to *both* parents, in the cases of: *(a) united household* and *(b) separation or divorce of the parents*. We can reasonably agree on the following

Definition 2: The cost of a child, CC, is the difference between the monetary income of a household composed by father, mother and child, MI_{FM}, and the income that would allow the household without the child to be as well-off as the household with the child, M^*_{FM}.

Let's first approach case *(a)*, the united family. In the case of a household composed by father, mother and child, its EI will be: $\frac{MI_{FM}}{{}_S e_{FMC}}$, with $MI_{FM} = MI_F + MI_M$. If one looks at the same household without the child, the same EI would be attained at a lower monetary income, $MI_{FN} - CC$, so that:

$$\frac{MI_{FM} - CC}{{}_S e_{FM}} = \frac{MI_{FM}}{{}_S e_{FMC}} \quad \Rightarrow \quad CC = MI_{FM}\left(1 - \left({}_S e_{FM} / {}_S e_{FMC}\right)\right) = MI_{FM} - MI_{FM} \frac{{}_S e_{FM}}{{}_S e_{FMC}} = MI_{FM} - MI^*_{FM}.$$

(5)

This implies transitivity of the ES, as $\left({}_S e_{FM} / {}_S e_{FMC}\right) = {}_{FMC} e_{FM}$.

Now, let's consider case *(b)*, the parents do separate or divorce. In these case, we must define: (i) the child's cost supported by the father, CC_F, and (ii) that supported by the mother, CC_M.

(i) to determine the amount of money the father must spend to support the child is trivial: this is simply equal to the amount of money S he must pay for supporting him.

Definition 3: From father's point of view, the cost of the child is $CC_F = S$.

(ii) it is not as simple as in previous case to determine the money the mother must spend to maintain the child. The EI of *a mother who lives with the child* and receives from the father an amount S is $EI_M = \frac{MI_M + S}{{}_S e_{MC}}$. Without the child, the mother would get the same EI if she earn an income MI^*_M such that $\frac{MI_M + S}{{}_S e_{MC}} = \frac{MI^*_M}{{}_S e_M}$. Hence, one has $CC_M = MI_M - MI^*_M$ and, therefore

$CC_M = MI_M - (MI_M + S)\dfrac{{}_S e_M}{{}_S e_{MC}} = MI_M - (MI_M + S)\,{}_{MC}e_M$. To conclude, the global cost of the child is

$$CC = CC_F + CC_M = S + MI_M - (MI_M + S)\,{}_{MC}e_M = (MI_M + S)(1 - {}_{MC}e_M)\,. \qquad (6)$$

If we assume now that the household will separate or divorce, giving life to two new households, father on the one hand and mother and child on the other hand, we ask to ourselves what the amount S the father should provide for supporting the child. According to the principles one wants to follow and to the specific household background, one may get different solutions, inspired to as many "philosophies". In this work, we opt for the approach that, despite some forcing, has the advantage of allowing to *separate the mother's welfare from the child's one* and, consequently, of allowing to *precisely calculate the expenditure share the father must support directly for the child*. This option can be called *philosophy of separation of mother's wealth from that of the child.*

Such a way of considering the calculation of the amount S consists of assuming that each component of the same household may have his own wealth. This assumption allows us to imagine securing the child a level of wealth different from that of the cohabiting parent, sharing its burden between the two parents according to proportions to be defined. For example, we may establish that after the divorce of parents, the child is secured the same level of wealth enjoyed before separation or divorce. To this purpose, we must first determine the amount of means necessary to secure the child such a wealth and then, to decide, on the basis of suitable criteria, how the burden must be shared in between the two parents.

Before the divorce, the child was enjoying a wealth level measured by the equivalent income $\dfrac{MI_{FM}}{{}_S e_{FMC}}$. After the divorce, in the household to which the child is allocated there must be an amount of resources such that the child wealth level remains unchanged. Of such an amount of resources, a part is enjoyed by the mother and a part by the child. The part imputable to the child represents the burden to be divided between the parents.

First of all, we must make a multiple comparison among the EIs of the three household types, that is: 1) united household; 2) household composed by mother and child; 3) household composed by the mother alone. This is as follows $\dfrac{MI_{FM}}{{}_S e_{FMC}} = \dfrac{MI_M^{**}}{{}_S e_{MC}} = \dfrac{MI_M^{***}}{{}_S e_C}$ and the cost of the child is

$$\mathrm{CC} = MI_M^{**} - MI_M^{***} = \frac{MI_{FM}}{{}_S e_{FMC}}({}_S e_{MC} - {}_S e_M) = \frac{MI_F + MI_M}{{}_S e_{FMC}}({}_S e_{MC} - {}_S e_M) \qquad (7)$$

which provides the amount of means necessary within the household mother-child to keep the child at the same wealth level he was enjoyng before the household would divorce. Such an amount of means must be charged to father (the amount S_F) and mother (the amount S_M).

This subdivision can take place according to different criteria: 1) equal shares: $S_F = S_M$; 2) proportionally to the income of the two parents: $S_F = b_1 \cdot MI_F$, $S_M = b_2 \cdot MI_M$[6]; 3) according to a progressively higher proportion for the parent who earns the highest income. In such a case, we may imagine to burden the father's income with a share $MI_F^r \big/ MI_{FM}^r$ of the total cost of child, with $r>1$. The remaining $MI_M^r \big/ MI_{FM}^r$ will be charged to the mother[7]. The amounts of money charged to the two parents are: $S_F = MI_F^r \cdot \frac{{}_S e_{MC} - {}_S e_M}{{}_S e_{FMC}}$ and $S_M = MI_M^r \cdot \frac{{}_S e_{MC} - {}_S e_M}{{}_S e_{FMC}}$.

For the calculation of the amounts due by the two parents for children support, we have chosen the option whose at point 3). After choosing progressivenesss in sharing, we have thought that the judge might greatly benefit from the knowledge of the composition of children cost in terms of goods and services necessary for their maintenance, according to consumption categories. In other words, to know what the expenditure for child's foods, apparel, transportation, and so on might be. Actually, such an information can eventually allow the judge to charge directly to the divorced parents the cost of at least some groups of consumption, so avoiding to pay to the other parent the amount of money related to those expenditure groups and to pay instead directly "at the source"[8].

Consequently, we have decided to calculate both the global cost of child for father and mother, and its subdivision among the same 9 commodity groups as used for the estimation of the ES. Evidently, as both income and the *r* parameter are continuous variables, it is possible to calculate as many table as many pairs of father and mother income and *r* values one may imagine.

[6] For example, if $b_1 = b_2 = 1$, the share of the total child's cost charged to the father would be MI_F / MI_{FM}, whereas the remaining share MI_M / MI_{FM} would be charged to the mother. Hence: $S_F = MI_F \frac{{}_S e_{MC} - {}_S e_M}{{}_S e_{FMC}}$ and $S_M = MI_M \frac{{}_S e_{MC} - {}_S e_M}{{}_S e_{FMC}}$.

[7] Obviously, point 2) is a special case of 3), i. e., $r=1$. It must be stressed that, with $r=2$, there is a remarkable progression in cost share of the richest parent.

[8] For example, by paying directly the medical expenses of the child, his transportation costs, and so on. Obviously, the remaining share of money to be payed to the other parent will be calculated as the difference between the global charge the parent must pay for child's maintenance and the consumption directly payed. To this purpose, if we put in equation (4), for *x*, the income of the mother inclusive of the amount payed by the divorced husband (for the different hypotheses) and, for ε, the estimate of the ES for the household composed by mother and child, we obtain the expenditure shares. From equation (7), we can thus calculate the cost of the child for groups of goods and services.

In this paper, just for the sake of discussing the features and the reliability of the income distribution models here outlined, we refer to some tables only, with significant parent and children ages. The tables, along with any others that may interest the readers, are available upon request.

4. Comments on some results and conclusions

Let's consider first the case of a household composed by a 38 year old father, a 35 year old mother and a 5 year old child[9]. We have calculated father's and mother's maintenance burdens for 3 significant values of *r*: *r*=0,75; 1,00; 1,50. As for each spouse's income, we have considered 9 income round figures expressed in millions of Italian Lira, 0 to 8 millions for the father and 1 to 9 millions for the mother.

Then, one can see that, with *r=0,75*, when the mother's income is, as an example, 0, she obviously must not pay anything, whereas the father must disburse, for child's support, an amount of 188,067 Lira.[10]. With *r=1* and *r=1,5* and the same combination of spouses' incomes, the father must disburse the same amount, that is, 188,067 Lira. This equality always holds, independently on *r*; no matter what the father's income may be, when the mother's income is 0. Evidently, this is due to the fact that the mother has no income: therefore, for the father, the amount to pay remains the same, independently on *r*.

This specific case represents, nevertheless, a quite unusual situation. Let's take a more "normal" case, for example one in which the father's income accounts for 3 million Lira and the mather's income accounts for 1 million Lira, with, obviously, the 5 year old child. If we observe the matrix where the *r=0,75* option is reported, we can see that the amount to be disbursed by the father accounts for 522,883 Lira, as compared to that to be disbursed by the mother, accounting for 229,384. It should be said at once that these seem reasonable amounts: for this type of household, with a household income accounting for 4 millions, it seems to us that a global burden for supporting a 5 year old child accounting for 652,269 Lira (incidentally, 18.8% out of the household income) is in line at all to the individual's expectations formed according the common logic sense based on the daily life evidence. If we take the *r=1* option, we observe that the amount the father must disburse increases to 564,201 Lira and that the one to be disbursed by the mother reduces to 188,067 Lira[11]. If we consider the *r=1,5,* option, we check that the

[9] *r*=0,75 states a subdivision of burden between the married couple with a progressiveness less than proportional for the spouse with the highest income; *r*=1 states a subdivision of burden between the married couple with a progressiveness proportional between the two incomes; *r*=1,5 states a subdivision of burden between the married couple with a progressiveness more than proportional for the spouse with the highest income.

[10] It is worth reminding that this figure, as all the others contained in the tables, represents the amount of means that are necessary, inside the household mother-child, to support only the child at the same level of wealth as before divorce.

[11] We should point out that, just in view of the proportionality of progressiveness option implied by this hypothesis on *r*, the money to be payd by the father remains the same by the whole row, that is, independently on the amount to be disbursed by the mother and that the money to be payd by the mother remains unchanged by column, that is, independently on the money to be payd by the father.

amount to be disbursed by the father does increase furtherly, up to 630,859 Lira, whereas the one to be payd by the mother decreases again decidedly enough, up to 121,408 Lira[12].

Subsequently, the total costs have been disaggregated according to the 9 commodity categories. Looking at them and considering again the pair with father's income accounting for 3 millions and mother's income accounting for 1 million, one can see that 176,275 Lira, out of 752,269 forming the total cost of the child, are spent for foods, 47,080 Lira for apparel, 164,875 Lira for housing, 40,335 for electricity, 29,680 for furniture, 20,109 for medical services, 110,404 for transports, 50,634 for entertainment and 112,857 for other goods and services. Taking into account the child's age, i.e., 5 years and the returns to scale that take place in supporting to him household collective services, for example, electricity and housing, they appear, as the previous ones, rather reliable figures.

Let's now pass to analyse the case of a household composed by father and mother in the same ages as before, but with two children, respectively, 10 year and 7 year olds. Let's refer again to the same pair of father-mother incomes. For *r=0,75* option, the cost of the two children goes, for the father, up to 907,312 Lira (from 522,883 Lira of the cost for the father of a 5 year old child) and for the mother, up to 398,030 Lira (from 229,384 Lira of the cost of a 5 year old child). In all, the family couple burden for the two children support accounts for 1,305,343 Lira (little less than 33% out of household's income).

It is not twice as much as the same couple did spend to support a 5 year old child and this is reasonable at all, both as evidently it is not automatic that two children cost the double as much as one does (all the more if they have different ages in beween them and as regards to an only child), and because of the returns to scale that occur when passing from the first child to the second one, particularly in cases like this, where the age difference between the two children is only three years (both as regards to household collective services and as regards to single goods: let's think of the re-use, for the second child, of underwear, toys, furniture, etc., belonged to the first child). In fact, while the total cost of the two children increases by about 73%, the cost for housing only increases by about 13%.

References

1. Deaton, A.S., Muellbauer, J., (1980a), An Almost Ideal Demand System, *American Economic Review*, Vol.70, 312-336.
2. Lewebel, A., (1991), The Rank of Demand Systems: Theory and Non-parametric Estimation, *Econometrica*, vol.59, no.3, 711-730.

[12] The three matrices, except for the column corresponding to 0 income of the mother, are symmetric, with the two figures contained in each cell put in the reverse order into the two triangles of the matrices themselves. The diagonals at the crossing of the

3. Muellbauer, J., (1977), Testing the Barten Model for Household Composition Effects and the Cost of Children, *The Economic Journal*, vol.87, 460-487.
4. Ray, R., (1983), Measuring the Cost of Children: an Alternative Approach, *Journal of Public Economics*, vol.22, 89-102.

G. Ferrari, M. Maltagliati*

Dipartimento di Statistica "G. Parenti", Università di Firenze,

Viale Morgagni, 59, 50134 Firenze

tel.: +39 055 4237221/+39 055 4237269;

fax: +39 055 4223560;

E-mail: ferrari@ds.unifi.it/mmaltag@ds.unifi.it.

same father's and mother's incomes contain pairs of equal figures and are they themselves equal in the three matrices. Obviously, the sums of the two figures into each cell are tidily equal, that is, for the same cell, for the three matrices.

* Financial support by the Italian MURST 40% funds is gratefully acknowledged.

BOUNDS FOR THE M|M|¥ QUEUEING SYSTEM BUSY PERIOD DISTRIBUTION FUNCTION

MANUEL ALBERTO M. FERREIRA
I.S.C.T.E./U.N.I.D.E.

1. INTRODUCTION

In a M|G|∞ queue system

- λ is the Poisson process arrival rate,
- α is the mean service time,
- $G(\cdot)$ is the service time d.f. and, so,

$$\alpha = \int_0^\infty [1 - G(t)]dt \qquad \textbf{(1.1),}$$

- $F(\cdot)$ is the service time equilibrium d.f. whose expression is

$$F(t) = \frac{1}{\alpha}\int_0^t [1 - G(x)]dx \qquad \textbf{(1.2),}$$

- $\rho = \lambda\alpha$ is the traffic intensity,
- B is the busy period length.

Note the importance of the busy period study, for this queueing system, because in it any customer when arrives finds a server available. So the problem is for how long the servers must be available that is how long is a busy period length.

The B d.f. has not a simple form and it can be written as (Stadje (1985))

$$P(B \le t) = 1 - \lambda^{-1}\sum_{n=1}^{\infty} c^{*n}(t) \qquad \textbf{(1.3)}$$

where c^{*n} is the *nth* convolution of c with itself being

$$c(t) = \lambda(1 - G(t))e^{-\lambda\int_0^t [1-G(x)]dx} \quad \textbf{(1.4).}$$

Only for the service time d.f. collection given for (Ferreira (1998))

$$G(t) = 1 - \frac{(1-e^{-\rho})(\lambda+\beta)}{\lambda e^{-\rho}\left(e^{(\lambda+\beta)t}-1\right)+\lambda},\ t \geq 0,\ -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho}-1} \quad \textbf{(1.5)}$$

the expression **(1.3)** becomes simple:

$$P(B \leq t) = 1 - \frac{\lambda+\beta}{\lambda}\left(1-e^{-\rho}\right)e^{-\rho(\lambda+\beta)t},\ t \geq 0,\ -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho}-1} \quad \textbf{(1.6).}$$

This does not happen for the M|M|∞ queueing systems (service time exponential) and so we will give in this paper some simple bounds for $P(B \leq t)$ in that case.

Finally note that if the p.d.f. allows the study of the distribution structure only the d.f. allows the probabilities calculation.

2. BOUNDS FOR THE M|M|¥ QUEUEING SYSTEM $P(B \leq t)$

We can write $c(t)$ as

$$c(t) = \rho f(t)e^{-\rho F(t)} \quad \textbf{(2.1)}$$

where $f(t) = \dfrac{dF(t)}{dt}$. So,

$$c(t) \geq \rho f(t)e^{-\rho} \quad \textbf{(2.2)}$$

and

$$\boxed{P(B \le t) \le 1 - \lambda^{-1} \sum_{n=1}^{\infty} f^{*n}(t)\rho^n e^{-n\rho}} \qquad \textbf{(2.3)}$$

or

$$\boxed{c(t) \le \rho f(t)} \qquad \textbf{(2.4)}$$

and

$$\boxed{P(B \le t) \ge 1 - \lambda^{-1} \sum_{n=1}^{\infty} f^{*n}(t)\rho^n} \qquad \textbf{(2.5)}$$

f^{*n} is the p.d.f. of the sum of n i.i.d. r.v. whose d.f. is given by **(1.2)**. So the bounds given by **(2.3)** and **(2.5)** depend only on ρ, λ and $F(\cdot)$.

For the M|M|∞ queue, $G(t) = 1 - e^{-t/\alpha}$ and, so, $f(t) = \dfrac{1 - 1 + e^{-t/\alpha}}{\alpha} = \dfrac{1}{\alpha} e^{-t/\alpha}$. Then, $f^{*n}(t) = \left(\dfrac{e^{-t/\alpha}}{\alpha}\right)^{*n} = \dfrac{t^{n-1}}{\alpha^n (n-1)!} e^{-t/\alpha}$. And $\sum_{n=1}^{\infty} f^{*n}(t)\rho^n e^{-n\rho} = \sum_{n=1}^{\infty} \dfrac{t^{n-1}}{\alpha^n (n-1)!} e^{-t/\alpha} \rho^n e^{-n\rho} =$
$= \dfrac{\rho}{\alpha} e^{-\rho} e^{-t/\alpha} \cdot \sum_{n=1}^{\infty} \dfrac{1}{(n-1)!} \left(\dfrac{\rho}{\alpha} e^{-\rho} t\right)^{n-1} = \lambda e^{-\rho - \frac{t}{\alpha}} \cdot e^{\lambda e^{-\rho} t} = \lambda e^{-\rho + \left(\lambda e^{-\rho} - \frac{1}{\alpha}\right)t} = \lambda e^{-\rho + \frac{\rho e^{-\rho} - 1}{\alpha} t}$. So

$$\boxed{P(B^M \le t) \le 1 - e^{-\rho - \frac{1 - \rho e^{-\rho}}{\alpha} t}} \qquad \textbf{(2.6)}$$

after **(2.3)**. From **(2.5)**, as $\sum_{n=1}^{\infty} f^{*n}(t)\rho^n = \sum_{n=1}^{\infty} \dfrac{t^{n-1}}{\alpha^n (n-1)!} e^{-t/\alpha} \rho^n = \dfrac{\rho}{\alpha} e^{-t/\alpha} \sum_{n=1}^{\infty} \dfrac{1}{(n-1)!} \left(\dfrac{\rho}{\alpha} t\right)^{n-1} =$
$= \lambda e^{-t/\alpha} e^{-\lambda t}$ we conclude that

$$\boxed{P(B^M \le t) \ge 1 - e^{-\frac{1-\rho}{\alpha} t}} \qquad \textbf{(2.7).}$$

The bound given by **(2.6)** is always less than **1.**. The one given by **(2.7)** is positive only for $\rho < 1$.

Otherwise $1-e^{-\rho-\frac{1-\rho e^{-\rho}}{\alpha}t} \geq 1-e^{-\frac{1-\rho}{\alpha}t} \Leftrightarrow -\rho-\frac{1-\rho e^{-\rho}}{\alpha}t \leq -\frac{1-\rho}{\alpha}t \Leftrightarrow \frac{1-\rho-1+\rho e^{-\rho}}{\alpha}t \leq \rho \Leftrightarrow$ $\Leftrightarrow \frac{\rho\left(e^{-\rho}-1\right)}{\alpha}t \leq \rho \Leftrightarrow t \geq \frac{\alpha}{e^{-\rho}-1} = -\frac{1}{1-e^{-\rho}} < 0,\ \rho > 0$ and the bound given by **(2.6)** is always greater than the one given by **(2.7)**.

In Ferreira e Ramalhoto (1994) we proved that

$$\boxed{G(t)e^{-\rho} \leq P(B \leq t) \leq G(t)} \qquad \textbf{(2.8)}$$

Consequentely

$$\boxed{\left(1-e^{-\frac{t}{\alpha}}\right)e^{-\rho} \leq P\left(B^{M} \leq t\right) \leq 1-e^{-t/\alpha}} \qquad \textbf{(2.9)}$$

The lower bound given in **(2.9)** is always positive, but for $\rho < 1$ the one given for **(2.7)** is better.

As $1-e^{-\rho-\frac{1-\rho e^{-\rho}}{\alpha}t} \leq 1-e^{-t/\alpha} \Leftrightarrow -\rho-\frac{1-\rho e^{-\rho}}{\alpha}t \geq -\frac{t}{\alpha} \Leftrightarrow \frac{1-1+\rho e^{-\rho}}{\alpha}t \geq \rho \Leftrightarrow t \geq \alpha e^{\rho}$ we conclude that for $t \geq \alpha e^{\rho}$ the bound given by **(2.6)** is better than the one given by **(2.9)**.

3. CONCLUSIONS

We presented bounds for $P(B \leq t)$ that can be used for any service time distribution. But they give simple expressions for exponential service times. It is even possible to compare them in order to make the better option in their use through very simple rules.

4. REFERENCES

1. FERREIRA, M.A.M., "Aplicação da Equação de Ricatti ao Estudo do Período de Ocupação do Sistema M|G|∞". Revista de Estatística. Vol. I. I.N.E.. 1998.

2. FERREIRA, M.A.M. e RAMALHOTO, M.F., "Estudo dos parâmetros básicos do período de ocupação da fila de espera M|G|∞". A Estatística e o Futuro e o Futuro da Estatística. Actas do I Congresso Anual da S.P.E.. Lisboa. 1994.

3. STADJE, W., "The busy period of the queueing systems M|G|∞", J.A.P., 22. 697-704.1985.

MANUEL ALBERTO M. FERREIRA
I.S.C.T.E./U.N.I.D.E.
AVª DAS FORÇAS ARMADAS
1649-026 LISBOA
PORTUGAL

Modeling of supplier-customer relations in supply chains[1]

Petr Fiala
University of Economics, Prague

1. Supply chain management

Supply chain management is now seen as a governing element in strategy and as an effective way of creating value for customers (see [5], [9])). An increasing number of companies in the world subscribe to the idea that developing long-term coordination and cooperation can significantly improve the efficiency of supply chains and provide a way to ensure competitive advantage. The expanding importance of supply chain integration presents a challenge to operations research to focus more attention on supply chain modeling (see [6], [7], [8]).

Supply chain is defined as a system of suppliers, manufacturers, distributors, retailers and customers where material, financial and information flows connect participants in both directions (see Fig.1).

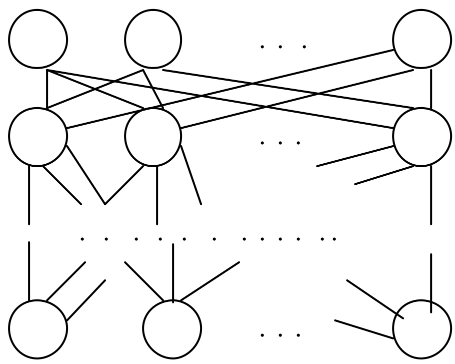


Fig. 1. Structure of the supply chain

Most supply chains are composed of independent agents with individual preferences. It is expected that no single agent has the power to optimize the supply chain. Each agent will attempt to

[1] The research project was supported by Grant No. 402/ 99/ 0852 from the Grant Agency of Czech Republic „Modeling and Analysis of Production Systems“ and CEZ: J 18/98: 311401001 from the University of Economics „Models and Methods for Economic Decisions“.

optimize his own preference, knowing that all of the other agents will do the same. This competitive behavior does not lead the agents to choose polices that optimize overall supply chain performance due to supply chain externalities. The agents can benefit from coordination and the end customer will receive a higher quality, cost effective value package in a shorter amount of time. Supply chain partnership (see [5]) leads to increased information flows, reduced uncertainty, and a more profitable supply chain.

The supplier-customer relations in supply chain can be taken as centralized or decentralized (see Fig.2).

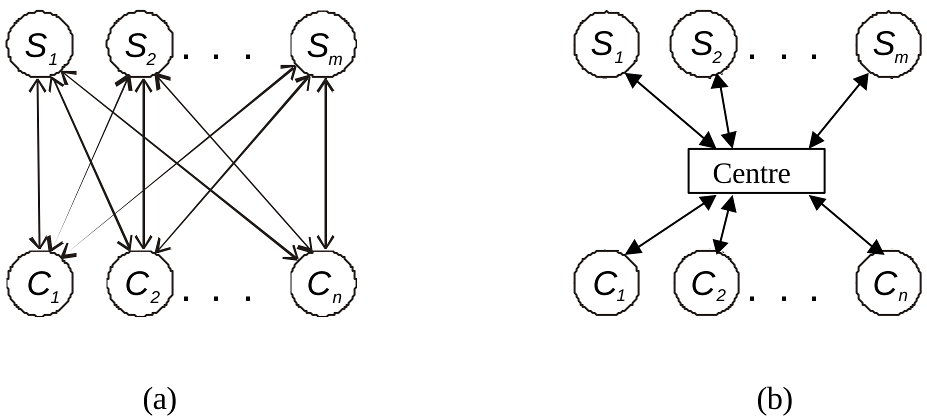


Fig. 2. Decentralized (a) and centralized (b) supplier-customer relations

The decentralized system causes some inefficiency in supply chains. The fully centralized system can be taken as a benchmark situation.

2. An example of inefficiency in supply chains

The so called bullwhip effect (see [4], [6]), describing growing variation upstream in a supply chain, is probably the most famous demonstration that decentralized decision making can lead to poor supply chain performance. The basic phenomenon is not new and has been recognized by Forrester. But the analyses of causes and suggestions for reducing the bullwhip effect in supply chains are challenges to modeling techniques.

There are some known causes (see [4],[6]), of the bullwhip effect as information asymmetry, demand forecasting, lead-times, batch ordering, supply shortages and price variations.

We use a simple two stages supply chain model with a single retailer and a single manufacturer. The customer demands D_t are independent and identically distributed random variables. The retailer observes customer demand D_t and places an order q_t to the manufacturer. The order is received with lead-time L, where L is the time it takes an order placed by the retailer to be received at that retailer. The retailer uses the moving average forecast model with p observations. To quantify increase in variability, it is necessary to determine the variance of orders q_t relative to the variance of demands D_t . It can be shown (see [6]):

$$\frac{\mathrm{Var(q)}}{\mathrm{Var(D)}} \geq 1 + \frac{2L}{p} + \frac{2L^2}{p^2} \qquad (1)$$

Information sharing of customer demand has an impact on the bullwhip effect. We consider now a k-stages supply chain with decentralized information and lead-times L_i between stages i and i+1. The variance increase is multiplicative at each stage of the supply chain.

In the case of centralized information, i.e. the retailer provides every stage of the supply chain with complete information on customer demand, the variance increase is additive. The centralized solution can be used as a benchmark, but the bullwhip effect is not completely eliminated.

The analysis of causes of the bullwhip effect has lead to suggestions for reducing the bullwhip effect in supply chains: reducing uncertainty, reducing variability, lead-time reduction and strategic partnership.

The strategic partnership means cooperation and coordination of actions through the supply chain The expected result is a mutually beneficial, win-win partnership that creates a synergistic supply chain in which the entire chain is more effective than the sum of its individual parts.

The strategic partnerships change material, financial and information flows among participants in the supply chain. The way of information sharing is changed by information centralizing using information technology. The material flows are managed within the supply chain. In vendor managed inventory the manufacturer manages the inventory of the product at the retailer and does not rely on the orders by the retailer, and thus avoids the bullwhip effect entirely. The financial flows are changed also. The agents can benefit from coordination. The typical solution is for the agents to agree to a set of transfer payments that modifies their incentives, and hence modifies their behavior. Many types of transfer payments are possible.

The partnership relations are based on supply contracts. The contracts are negotiated by multiple criteria as time, quantity, quality and costs. There are different approaches to modeling multicriteria negotiation processes to reach a consensus among participants (see [3]).

3. Modeling of negotiation process

Negotiation is a process by which a joint decision is made by two or more parties. It is a communication designed to reach an agreement when the parties have some interests that are shared and others that are opposed. Both the communication and cooperation are required in the problem solving process of negotiation, and there are some important relationships between the communication and the cooperative behavior. The very act of communication in negotiating involves some cooperation between the negotiators. The more cooperative they are the greater is the possibility of reaching an agreement.

The problem of coordination in supply chains involves multiple agents with multiple goals. Goals can be divided into two types, goals that are mutual for all the agents and goals that are different and require cooperation of multiple agents to achieve a consensus. There are two very important aspects of group decision making: assertiveness and cooperativeness. Assertiveness is satisfaction of one´s own concerns and cooperativeness is a tendency to satisfy others. A cooperative decision making requires free communication among agents and gives synergical effects in a conflict resolution. The basic trend in the cooperative decision making is to transform a possible conflict to a joint problem.

The problem solving process is an extensive organization of information and attitudes surrounding a problem situation. The process involves understanding the goals, setting criteria for evaluating possible solutions, the exploration of alternative solutions, selection of the solution that best fits the criteria.

Some basic ideas of formal approaches of the problem solving can be introduced to cooperative decision making. There are two aspects of the problem solving - representation and searching. The state space representation introduces the concepts of states and operators. An operator transforms one state into another state. A solution could be obtained by a search process, first applies operators to the initial state to produce new states and so on, until the goal state is produced. Communication between suppliers and customers can be provided through information sharing (schematically see Fig. 2).

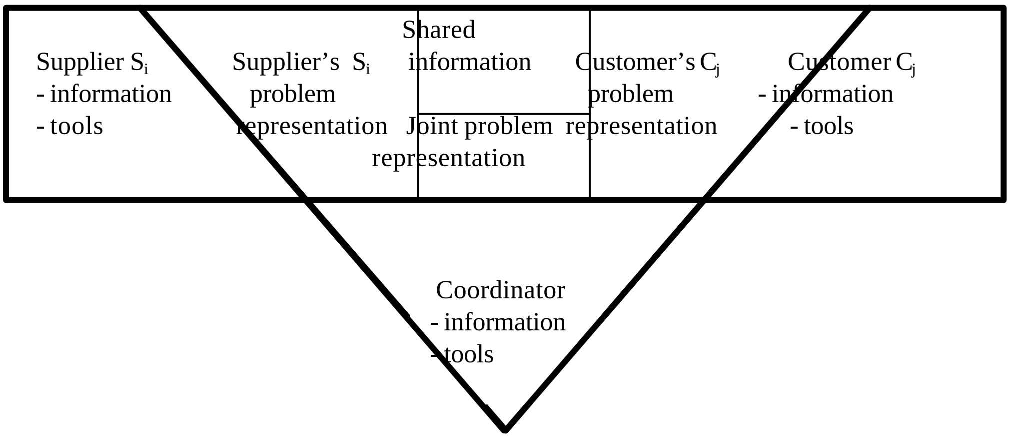


Fig. 2. Communication through information sharing

We propose a two phases' interactive approach for solving cooperative decision making problems (see [2]):

1. Finding the ideal solution for individual agents.
2. Finding a consensus for all the agents.

In the first phase every decision maker search the ideal alternative by the assertivity principle.

The general formulation of a multicriteria decision problem for an individual unit is expressed as follows

$$\mathbf{z(x)} = (z_1(\mathbf{x}), z_2(\mathbf{x}), ..., z_k(\mathbf{x})) \rightarrow \text{"max"}$$

$$\mathbf{x} \in X ,$$

where X is a decision space, x is a decision alternative and $z_1, z_2, ..., z_k$ are the criteria. The decision space is defined by objective restrictions and by mutual goals of all the decision makers in the aspiration level formulation. The decision alternative **x** is transformed by the criteria to criteria values $\mathbf{z} \in Z$, where Z is a criteria space. Every decision making units has its own criteria. People appear to satisfy rather than attempting to optimize. That means substituting goals of reaching specified aspiration levels for goals of maximizing.

We denote $\mathbf{y}^{(s)}$ aspiration levels of the criteria and $\Delta\mathbf{y}^{(s)}$ changes of aspiration levels in the step s. We search alternatives for which it holds

$$\mathbf{z(x)} \geq \mathbf{y}^{(s)}$$

$$\mathbf{x} \in X .$$

According to heuristic information from results of the previous condition the decision making unit changes the aspiration levels of criteria for step s+1 :

$$\mathbf{y}^{(s+1)} = \mathbf{y}^{(s)} + \Delta\mathbf{y}^{(s)} .$$

We can formulate the multicriteria decision problem as a state space representation. The state space corresponds with the criteria space Z, where the states are the aspiration levels of the criteria $\mathbf{y}^{(s)}$ and the operators are changes of the aspiration levels $\Delta\mathbf{y}^{(s)}$. The start state is a vector of the initial aspiration levels and the goal state is a vector of the criteria levels for the best alternative. For finding the ideal alternative we use the depth-first search method with backtracking procedure. The heuristic information is distance between an arbitrary state and the goal state.

In the second phase a consensus could be obtained by the search process and the principle of cooperativeness is applied. The heuristic information for the decision making unit is the distance between his proposal and the opponent's proposal.

For simplicity we assume the model with one supplier and one customer

$$z^1(\mathbf{x}) \to \text{"max"}$$

$$z^2(\mathbf{x}) \to \text{"max"}$$

$$\mathbf{x} \in X .$$

The decision making units search a consensus on a common decision space X. The decision making units change aspiration levels of the criteria y^1, y^2. The sets of feasible alternatives for the aspiration levels y^1 and y^2 are X^1 and X^2.

$$z^1(\mathbf{x}) \geq y^1 \qquad\qquad z^2(\mathbf{x}) \geq y^2$$

$$\mathbf{x} \in X \qquad\qquad\quad \mathbf{x} \in X .$$

The consensus set S of the negotiations is the intersection of sets X^1 and X^2

$$S = X^1 \cap X^2 .$$

By changes of the aspiration levels the consensus set S is changed too. The decision making units search one element consensus set S by alternating of the consensus proposals. Independently on the meaning of local symbols, the cooperative decision making model of negotiation is represent by the state vector **z**.

5 Conclusions

In the paper we analyze supplier-customer relations in supply chains. Information asymmetry is one of the most powerful sources of inefficiency. Information technology has lead to communication, coordination and cooperation among units in supply chains. Building of different types of strategic partnerships and different type of contracts among participants can significantly reduce or eliminate inefficiency in supply chains.

References

[1] Browne, J., Sackett, P.J., Wortmann, J.C.: Future manufacturing systems - toward the extended enterprise. Computers in Industry 25 (1995), s. 235-254.

[2] Fiala, P.: Models of cooperative decision making. In: Fandel, G., Gal, T.: Multiple Criteria Decision Making. Springer, Berlin 1997, s. 128-136.

[3] Fiala, P.: Modelování vyjednávacího procesu s více hodnotícími kritérii. Politická ekonomie 47 (1999), s. 253-268.

[4] Lee, H.L., Padmanabhan, Whang, S.: Information Distortion in a Supply Chain: The Bullwhip Effect. Management Science 43, 1997, 546-558.

[5] MacBeth, D.K., Ferguson,N.: Partnership Sourcing: An Integrated Supply Chain Management Approach. Pitman Publishing, London 1994.

[6] Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E.: Designing and Managing the Supply Chain: Concepts, Strategies and Case studies. Irwin/ Mc Graw-Hill, 1999.

[7] Slats, P.A., Bhola, B., Evers, J.J.M., Dijkhuizen, G.: Logistics Chain Modelling. European Journal of Operational Research 87, 1995, 1-20.

[8] Tayur, S., Ganeshan, R., Magazine, M.: Quantitative models for supply chain management, Kluwer, Boston 1999.

[9] Thomas, D.J., Griffin, P.M.: Coordinated Supply Chain Management. European Journal of Operational Research 94, 1996, 1-15.

Doc. RNDr. Ing. Petr Fiala, CSc., MBA,

Department of Econometrics, University of Economics,
W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic

A multivariate labour market model in the Czech Republic[1]

Jana Hanclová

Faculty of Economics, VŠB-Technical University Ostrava

Abstract: The paper deals with an existence of an equilibrium unemployment-vacancy rate relationship in the Czech Republic. We formulate a basic disequilibrium growth model of employment. Furthermore we study a long-run steady state of this bivariate labour market model and we try to introduce a random disturbance term. The next extension of this model makes (un)employment flows endogenously using macroeconomic cyclical variables as well as variables associated with the composition of the unemployment. We try to show an empirical study with monthly time series over the period 1993-1999 for the Czech labour market.

Keywords: Unemployment-vacancies relationship, cointegrating analysis, equilibrium models, labour market model

1. Introduction

Many economic studies deal with the inverse relationship between the unemployment rate and vacancy rate (*uv*). The basic bivariate labour market model proposed by Beveridge is possible to extend by introducing endogeneity in employment search. This modification gives rise to study a multivariate framework of the (dis)equilibrium relationship.

The objective of this study is to examine the existence of the *uv* relationship in the Czech Republic using more recent data, from January 1993 to December 1999. We use appropriate econometric time series methods to synthesize past studies with an equilibrium or disequilibrium approach. Cyclical macroeconomic variables and variables associated with the composition of the unemployment pool are examined whether they are related with *u* and *v* in the long run using cointegration analysis.

---

[1] The paper is supported by the Grant Agency of the Czech Republic – grant. No. 402/00/1165 and corresponds to the research programme of the Faculty of Economics VŠB-Technical University no. CEZ:J17/98:275100015

2. The basic and expanded (dis)equilibrium models

A disequilibrium model of outflows from and inflows to employment can be constructed:

$$\Delta E_t = E_t - E_{t-1} \equiv H_t - Q_t \tag{E1}$$

where

H_t – hires at period t,

Q_t – quits lay-offs at period t.

A hiring function can be specified by the Cobb-Douglas function with constant returns to scale. The flow of the number of hires and rehires at period *t* depends on the stocks of unemployed people (*U*) and vacancies (*V*) at the end of period *(t-1)*, and the constant rate job search parameter b_0 according to a following equation:

$$H_t = b_0 \cdot \left[U_{t-1}^{a} \cdot V_{t-1}^{a}\right], \tag{E2}$$

where

$b_0 -$ describes the efficiency of job search behaviour.

Hiring function H_t is assumed to be homogeneous of degree one in *U* and *V*. It is also assumed that the turnover (quit) rate is proportional to the level of employment with the constant rate d_0. The net change in employment can be expressed using **a basic disequilibrium labour market model:**

$$\Delta E_t = b_0\left[U_{t-1}^{a} \cdot V_{t-1}^{1-a}\right] - d_0 \cdot E_{t-1} \tag{E3}$$

or

$$e_t = b_0\left[ue_{t-1}^{a} \cdot ve_{t-1}^{1-a}\right] - d_0, \tag{E4}$$

where

$$e_t = \Delta E_t / E_{t-1} \qquad ue_{t-1} = U_{t-1} / E_{t-1} \qquad ve_{t-1} = V_{t-1} / E_{t-1}. \tag{E5}$$

In a **long-run equilibrium labour market model** the job stock and labour force are constant and there is equality between the flows of accessions and separations on the labour market (i.e. $e_t = 0$) :

$$b_0\left[ue_{t-1}^{a} \cdot ve_{t-1}^{1-a}\right] = d_0 \quad \text{or} \quad \ln(ue_{t-1}) = -\frac{(1-a)}{a} \cdot \ln(ve_{t-1}) + \frac{1}{a} \cdot \ln(d_0 / b_0), \tag{E6}$$

which is the hyperbolic *uv* relationship.

Anthony (1999) **introduced a random disturbance term (e_t)** into E3 as a multiplicative term in both the hiring and quit function. In this case we can express the equation E3 as follows:

$$\Delta E_t = b_0 \left[U_{t-1}^{a} \cdot V_{t-1}^{1-a}\right]\left\{e^{m_{t-1}}\right\} - d_0 \cdot \left\{e^{n_{t-1}}\right\} \cdot E_{t-1} \qquad \text{(E7)}$$

or

$$\ln(ue_t) = -\frac{1-a}{a}\ln(ve_t) + \frac{1}{a}\ln(d_0 / b_0) + \frac{1}{a}e_t \qquad \text{(E8)}$$

where

$$e_t = n_t - m_t, \qquad m_t \sim N(0, s_m^2) \qquad n_t \sim N(0, s_n^2) \qquad \text{(E9)}$$

This framework allows cointegration tests to be conducted using the disturbance term in equation E8. If it can be shown that logged ue_t and ve_t are I(1) and e_t is I(0), then ue_t and ve_t are cointegrated and there is long-run equilibrium.

The bivariate labour market model is restrictive as b_0 and d_0 were specified as fixed parameters, implying that un(employment) flows are exogenous. As a consequence, this model will fail to identify an equilibrium relationship if additional variables constitute part of the long-run relationship. Therefore we modify our model by allowing (un)employment flows to be endogenously determined. These flows can be dependent on cyclical macroeconomic variables as well as factors associated with structural unemployment. Next we can test whether this **multivariate specification of the Beveridge Curve** does in fact represent a long-run or cointegration relationship using our data sample in the Czech Republic.

We try to introduce following **cyclical macroeconomic variables** :

i_t – interest rate (PRIBOR 3M) during month *t*,

rwi_t – real wages in industry during month *t*,

$subsr_t$ – replacement ratio for period *t* [=100*(unemployment benefits/nominal wages in industry)].

Factors associated with the **structural composition of unemployment** may also be relevant in influencing *uv* as they would partially capture costs of labour search. We consider to introduce:

l_t - number of long-term unemployed people (> 1 year) at the period *t*,

f_t - number of unemployed females at the period *t*.

The variables to be expressed as functions of unemployment rate and vacancy rate in the long run. If we the quit rate is defined as

$$d_t = d_0 \cdot rwi_t^n \cdot subsr_t^j \cdot e^{l \cdot i_t} \qquad \text{(E10)}$$

then **a multivariate disequilibrium labour market model** can be expressed:

$$\Delta E_t = b_0 \cdot l_{t-1}^{b_1} \cdot f_{t-1}^{b_2} \left[U_{t-1}^a \cdot V_{t-1}^{1-a}\right]\left\{e^{m_{t-1}}\right\} - d_0 \cdot rwi_{t-1}^g \cdot subsr_{t-1}^j \cdot e^{l \cdot i_{t-1}} \cdot \left\{e^{n_{t-1}}\right\} \cdot E_{t-1} \qquad \text{(E11)}$$

and **a multivariate long-run equilibrium model** is

$$\ln ue_t = -\frac{(1-a)}{a} \cdot \ln ve_t + \left(\frac{g}{a}\right) \cdot \ln rwi_t + \left(\frac{j}{a}\right) \cdot \ln subsr_t + \left(\frac{l}{a}\right) \cdot i_t - \left(\frac{b_1}{a}\right) \cdot \ln le_t - \left(\frac{b_2}{a}\right) \cdot \ln fe_t + \frac{1}{a} \cdot \ln\left(\frac{d_0}{b_0}\right) + \frac{1}{a} e_t \qquad \text{(E12)}$$

where

$$e_t = n_t - m_t, \qquad b_1 + b_2 = 0, \qquad le_t = \frac{l_t}{E_t}, \qquad fe_t = \frac{f_t}{E_t}, \qquad l, j > 0 \qquad \text{(E13)}$$

Equation E12 provides a richer model structure than the conventional model E8.

3. Empirical study

3.1. Time series data

Monthly time series data are used for our empirical study in the Czech labour market from the January 1993 to the December 1999 (i.e. 84 observations excluding lags in modelling). All series are seasonally adjusted with the exception of interest rates and unemployment benefits. Where quarterly series are available, monthly series are obtained by linear interpolation.

Unemployment rate (ue_t) and vacancy rate (ve_t) are shown in Fig. 1.

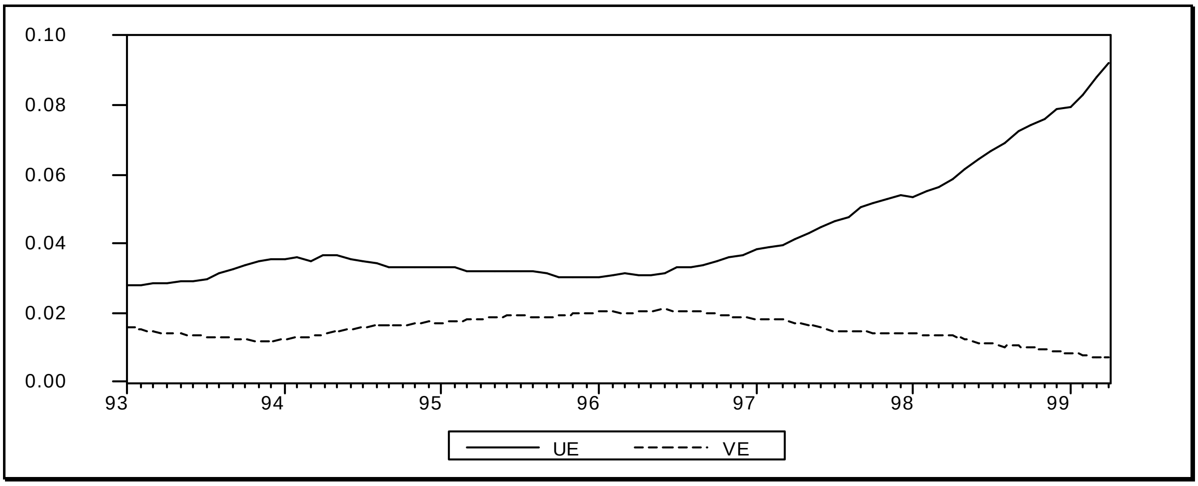


Fig. 1 : Unemployment rate and vacancy rate development

In determinig the (non)stationarity properties of the data we tested on a truncated sample between January 1994 and December 1999 to allow for possible leads and lags estimation. We used the PP (Phillips–Perron) tests for unit roots with results reported in the Table 1.

Test results for the **logged unemployment rate and vacancy rate** are marginal yet indicative of **nonstacionarity** for PP tests support the same results**.** The PP test makes a correction to the t-statistic of the coefficient of $ln(variable)_{t-1}$ to account for serial correlation in random term.

Furthemore PP tests indicate that employment growth and marginal real wages in industry are stationary or I(0). All other data series (**yue, yve, ysubsr, yle, yfe**) **are nonstationary or I(1).**. Supplementary testing on the I(1) series fond none of them to be I(2). Logged **interest rate data series was I(1)**. Since the unemployment rate and vacancy rate are found to exhibit nonstationary behaviour using PP tests, it seems appropriate to model the *uv* relationship using cointegrating analysis.

Variable	Description	PPstatistics	Lag lenghts
e	Employment growth	**-8,760***	0
yue	ln(unemployment rate)	1,495	0
yve	ln(vacancy rate)	2,224*	0
i	interest rate	-2,169*	0
ysrwi	ln(real wage)	**-4,720***	0
ysubsr	ln(replacement rate)	-1,917*	0
yle	ln(long-term unemployment rate)	-1,018	0
yfe	ln(female unemployment rate)	0,076	0
due	Δ yue	**-5,357***	0
dve	Δ uve	**-5,263***	0
di	Δ i	**-7,411***	0
dsubsr	Δ subsr	**-3,920***	0
dle	Δ l	**-5,585***	0
dfe	Δ f	**-8,773***	0
Δ = difference of variable			
Critical values (including intercept and trend)			
1% (-4.085) **5%** (-3.470) **10%** (-3.162)			

Table 1: Tests for the null of nonstationarity and stationarity

3.2 Cointegrating analysis for the basic model_1 (*yue, yve*)

Then we try to test for cointegration (i.e. that linear combination of nonstacionary time series yue_t and yve_t is I(0)). If an long-run equilibrium relationship exists we estimate the cointegration equation. We applied 3 cointegration tests for **model_1 *(yue, yve)***:

a) Linear regression_1 of $ln(ue_t)$ on $ln(ve_t)$ and testing for nonstationarity of residua time series (without trend and intercept) and also linear reggresion_2 of $ln(ve_t)$ on $ln(ue_t)$ and testing for nonstationarity of the second residua time series,
b) Engle-Granger test,
c) Johansen test.

Results are reported in Table 2.

ADF (Augmented Dickey-Fuller) tests for residua_01 and residua_02 estimated by OLS method (since results can vary with the ordering of variables in regression) indicated cointegration *yue* and *yve* but with slow DW value for the model estimating residua.

The **Granger(1969)** approach is to see how much the current *yue* can be explained by past values of *yue* and then to see whether adding lagged values of *yve* can improve the explanation (and vice versa). The null hypothesis is therefore that *yve* does not Granger-cause *yve* in the first regression and that *yue* does not Granger-cause *yve* in the second regression. For our results **we cannot reject the hypothesis that *yve* does not Granger cause *yue*** but **we do reject the hypothesis that *yue* does not Granger cause *yve*** on 5% level of significance.

<table>
<tr><th></th><th>Test for cointegration</th><th>Results</th></tr>
<tr><td rowspan="2">a)</td><td>OLS:
$E(yue) = -7.279^{*} - 0{,}964^{*} \cdot yve \quad R^2 = 0.614^{*} \quad DW = 0.019 \quad sign.F = 0.00$
Resid_01: Observed ADF statistic =-2,582* Critical value(5%)=-1,946
$\Rightarrow$ Resid_01 ~ I(0)</td><td rowspan="2">Cointegration yue, yve</td></tr>
<tr><td>OLS:
$E(yve) = -6.254^{*} - 0{,}635^{*} \cdot yue \quad R^2 = 0.612 \quad DW = 0.033 \quad sign.F = 0.00$
Resid_02: Observed ADF statistic =-2,296* Critical value(5%)=-1,946
$\Rightarrow$ Resid_02 ~ I(0)</td></tr>
</table>

b)	**Engle-Granger** (76 obs.): **Null hypothesis:** **F-Statistic** **Probability** *YVE* does not Granger Cause *YUE* 1.43761 0.18001 ***YUE* does not Granger Cause *YVE*** **2.16005** **0.02859**	*yue -> yve*

c)	**Johansen test:** Test assumption: Linear deterministic trend in the data **Series: YUE YVE** **Eigenvalue** **Likelihood** **Ratio** **5 %** **Critical Value** **1% Critical Value** **Hypothesized** **No. of CE(s)** 0.249926 **29.54842** **> 25.32** 30.45 None* 0.100931 **7.979640** **< 12.25** 16.26 At most 1 * (**) denotes rejection of the hypothesis at 5% (1%) significance level **L.R. test indicates 1 cointegrating equation at 5%** **Normalized Cointegrating Coefficients: 1 Cointegrating Equation** ***YUE*** ***YVE*** ***@TREND(92:02)*** ***C*** 1.0 0.672935 -0.008472 6.478017	**1 Cointegrating equation at 5% significance level**

Table 2: Various tests for cointegration for model_1 (*yue, yve*)

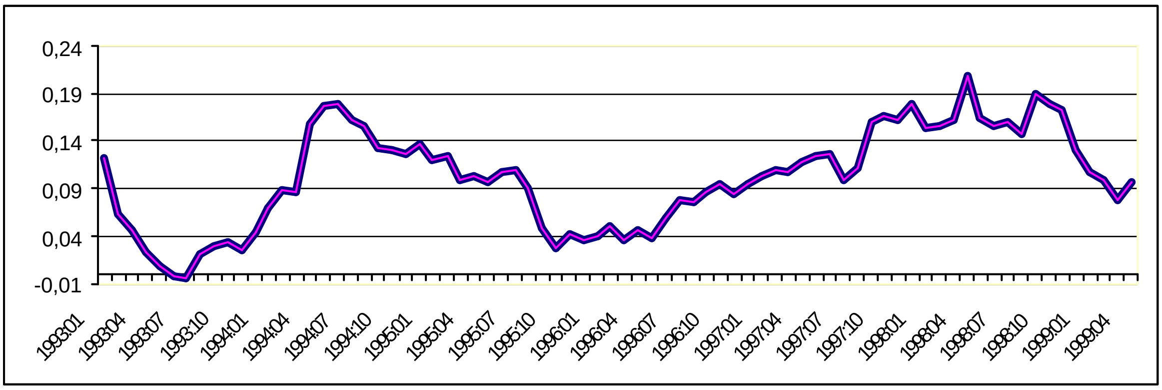


Figure 1: Normalized cointegrating equation for model_1 (*yve,yue*)

Johansen (1991, 1995) developed the methodology for implementation VAR-based cointegration tests. Johansen's method is to test the restrictions imposed by cointegration on the unresctricted VAR involving the series. In the second column (Likelihood Ratio) gives *trace statistic*. To determine the **number of cointegrating relations** (***k***), we proceeded sequentially from $k=0$ to (number of non-stationary variables - 1) until we fail to reject hypothesis. For our model_1 (yue,yve) Johansen's test indicated **1 cointegrating eaquation at 5% significance level**. The figure 1 shows normalized cointegrating equations for model_1 - (*yve,yue*).

3.3 Cointegrating analysis for extended multivariate models

Furthermore we applied cointegration analysis for other 3 models using three above-mentioned methods:

- Model_2 (*yue, yve, i),*
- Model_3 (*yue, yve, ysubsr*),
- Model_4 (*yue, yve, i, ysubsr*).

The Table 3 includes results only for Johansen's test. Johansen's trace statistic detected:

- **2 cointegrating equation** at 5% significance level for **model_2 (*yue, yve, i),***
- **2 cointegrating equations** at 5% significance level for **model_3 (*yue, yve, ysubsr*),**
- **3 cointegrating equations** at 5 % significance level for **model_4 (*yue, yve, ysubsr, i*).**

Model	Johansen cointegrating tests
Model_2 ***(yue,yve,i)***	Series: **YUE YVE I** (Test assumption: No deterministic trend in the data) Likelihood 5 Percent 1 Percent Hypothesized Eigenvalue Ratio Critical Value Critical Value No. of CE(s) 0.307474 **48.88154** **> 34.91** 41.07 None ** 0.157188 **21.32584** **> 19.96** 24.60 At most 1 * 0.107147 **8.499966** **< 9.24** 12.97 At most 2 **L.R. test indicates 2 cointegrating equation(s) at 5% significance level**
Model_3 ***(yue, yve, ysubsr)***	Series: **YUE YVE YSUBSR** (Test assumption: No deterministic trend in the data) Likelihood 5 Percent 1 Percent Hypothesized Eigenvalue Ratio Critical Value Critical Value No. of CE(s)

	0.288928 **38.11028** **> 24.31** 29.75 None ** 0.124556 **12.53662** **> 12.53** 16.31 At most 1 * 0.033555 **2.559848** **< 3.84** 6.51 At most 2 **L.R. test indicates 2 cointegrating equation(s) at 5% significance level**
Model_4 (*yue*, *yve*, *i*, *ysubsr*)	Series: **YUE YVE YSUBSR I** (Test assumption: No deterministic trend in the data) Likelihood 5 Percent 1 Percent Hypothesized Eigenvalue Ratio Critical Value Critical Value No. of CE(s) 0.451889 **84.97240** **> 39.89** 45.58 None ** 0.270473 **39.87666** **> 24.31** 29.75 At most 1 ** 0.194510 **16.22470**

	> 12.53
	16.31
	At most 2 *
	2.55E-05
	0.001912
	< 3.84
	6.51
	At most 3
	L.R. test indicates 3 cointegrating equation(s) at 5% significance level

Table 3: Results for Johansen's cointegrating test

(*[**] denotes rejection of the hypothesis at 5%[1%] significance level)

4. Summary and conclusions

The aim of this paper was to investigate existence of the *uv* relationship for the labour market in the Czech Republic using cointegrating analysis.

Cointegrating test results suggest that **bivariate long-run *uv* relationship exists** (model_1) and it is depicted in Fig. 1. These results support the basic labour market search model which assumes exogenity of (un)employment flows.

A modified Beveridge curve was used to examine the existence multivariate equilibrium relationship. This was achieved by making (un)emloyment flows endogenously determined by macroeconomic cyclical variables (*i, subsr*). The **Johansen's cointegrating tests indicated the existence of equilibrium relationship between unemployment, vacancies and the replacement ratio and interest rate.** This suggests that cyclical variables are influential in determining the stability of Beveridge curve. The empirical results show that unemployment rises when the ratio of uneployment benefits increases to work income. Also the interest rate was cointegrated with unemployment and vacancies rate and move positively with unemployment.

References

1. Anthony J.D.F. (1999): The relationship between unemployment and vacancies in Australia. Applied Economics, No. 31, pp. 641-652.
2. Eviews User's Guide (1998). Quantitative Micro Software.

3. Granger,C.W.J (1969): Investigating causal relations by econometric models and cross-spectral methods. Econometrica, No. 37, pp.424-438.
4. Jablonský,J. (1993).: Multicriteria evaluations of urban regions in the Czech Republic. In: Applying Multicriteria Methods for Environmental Management. Ispra 1993, s. 34.
5. Johansen,S.(1991): Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. Econometrica, No. 59, pp. 1551-1580.
6. Johansen,S.(1995): Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press.

Jana Hanclová,
Faculty of Economics VŠB-
Technical University Ostrava,
Czech Republic
jana.hanclova@vsb.cz

Factors related to competitiveness in the Portuguese manufacturing industry

Manuela Magalhães Hill
Instituto Superior de Ciências do Trabalho e da Empresa
PORTUGAL

INTRODUCTION

The study analysed questionnaire data obtained from a sample of 678 companies based in the two metropolitan areas of Portugal – Lisbon and Oporto. The sample was representative in terms of size of company (number of employees) and Sector of activity. Data were obtained on a wide range of variables covering six areas; Characteristics of the company, Research and Development activities, Human resources policies and practices, Company and work organization, Unionization and worker representation, Characteristics of managers.

The index of competitiveness

In the absence of data on market share and similar variables Cashflow per employee for the year 1998 was used as an index of competitiveness. Strictly speaking this is a measure of company performance, but companies that perform well in this sense are likely to be more competitive. If they are not, their market share will decrease and performance will decline. Thus Cashflow per employee represents a correlate of competitiveness which may plausibly be used as an indirect Index. Four other indicators were developed as weighted combinations of questionaire variables using Non-linear Principal Component analysis. These were:

1. **Inovation** (in the previous five years).
2. **Level of information technology**
3. **Organizational change** (in the previous five years).
4. **Changes in human resources management** (in the previous two years).

THE RESEARCH STRATEGY

It is recognized in Portugal that the two metropolitan areas differ somewhat in their 'industrial cultures' and in view of this it was thought likely that there may be regional differences in the factors associated with competitiveness. Thus, the overall strategy of the research was to investigate factors associated with competitiveness in three separate analyses – the first dealing with the two metropolitan areas combined, the second dealing only with companies in the Lisbon area and the third dealing only with companies in the Oporto area.

METHODS OF RESEARCH

Using cut-offs at the 47^{th} and 53^{rd} percentiles the Index of competitiveness was dichotomized to form two groups of companies - a HIGH competitive group and a LOW competitive group. In each of the three separate analyses the dichotomized Index of competitiveness was used as the dependent variable in a logistic regression (Forward LR method) the objective of which was to discover the main factors associated with competitiveness. The independent variables in each regression were the four indicators and a set of variables drawn from each of the six areas covered by the questionaire. These variables were selected by prior univariate analysis as significantly related to the Index of competitiveness.

RESULTS

The analysis for the two metropolitan areas combined

Five factors were found to be related to competitiveness and the regression model gave a good fit to the data according to the Hosmer and Lemeshow procedure (χ^2 = 8.567, d.f. = 8, P = 0,380) and showed a good Holdout cross-validation. The regression correctly classified the group membership (HIGH/LOW competitiveness) of 72.8% of all companies (74.2% for the LOW group and 71.5% for the HIGH group). The five factors associated with competitiveness were; level of inovation in the previous five years, level of information technology, use of the MRP method for planning production needs, Salary costs as a percentage of Total costs, use of Universities (other than for recruiting graduate staff).

Further analysis showed that within each of the competitiveness groups (High/Low) there were no significant differences between Lisbon and Oporto companies on four of the five factors. On the fifth factor, Salary costs as a percentage of total costs, ANOVA showed a significant interaction between Competitiveness and Metropolitan area.

Inovation

Two types of inovation were related to competitiveness. Product inovation and inovation in the production process. Significantly more Highly competitive companies had introduced new products and /or improved products in the previous five years, and in the case of improved products the improvements were in terms of quality rather than materials or design. Significantly more competitive companies had also changed the manufacturing process .

Information technology

Highly competitive companies had much more information technology than was the case with the low competitive companies. But competitiveness was associated with the possession of seven specific types of technology, **four of which were concerned with the automation, or automatic control, of the production process**. The seven types were; mainframe computers, CAD, CAD/CAM, communication network between work stations, robotic components, automated control of sections of the factory and production management assisted by computer. These seven types of technology were related to competitiveness **in all sizes of company** (less than 50 employees, 50 – 100, 101 – 200 and greater than 200.

Use of the MRP method for planning the needs of production

The percentage of highly competitive companies using this method (39,4%) was approximately twice that of the low competitive (20,6%). Moreover the use of this method interacted with the use of information technology to influence competitiveness. 64,5% of all companies above the median on the Information technology indicator were in the High competitive group. But of the companies in this 'high tec.' group that ***also*** used MRP 73.3% were in the High competitive group. Of the companies in the 'high tec.' group that did ***not*** use MRP only 59.4% were in the High competitive group.

Salary costs as a percentage of total costs

Average Salary costs as a percentage of total costs was significantly lower in the High competitive group (22.5%) than in the Low group 35.7%. In the two metropolitan areas combined there was no significant effect of company size or level of information technology on salary costs as a percentage of total costs.

Use of Universities

Although comparatively few (20.8%) of the highly competitive companies made use of university services the percentage of low competitive companies doing so was significantly less (9.0%). The kind of service principally used was assistance in solving specific problems, including environmental problems, and this was mainly on an *ad hoc* basis.

The analysis for the Lisbon metropolitan area

Five factors were found to be associated with competitiveness and again the regression model gave a good fit to the data (χ^2 = 6.480, d.f. = 8, P = 0,594). It correctly classsified the group membership of 71.4% of companies (70.0% in the Low competitive group and 72.6% in the High). The five factors associated with competitiveness were: Organizational change, acquisition of services from external agencies, use of individual pay increases based on merit/worker performance, salary costs as a percentage of total costs, perceived importance of product price as a competitive advantage.

Organizational changes

The main organizational changes related to competitiveness were made by small or medium sized companies (fewer than 200 employees). 34.7% of the highly competitive companies of this size had decentralized decision making compared to 16.7% of the low competitive. 43.9 % of the High had improved the flow of information in the company compared to 20.5% of the Low. In the High group 24.5% of companies had created working groups/teams but only 15.4% in the Low group, and 45.9% of the High group companies had created new methods for planning, controlling and managing production (28.2% Low). This last change was also related to competitiveness in large companies (more than 200 employees) where 48.3% of the High group had created new methods compared to 28.9% of the Low group. Large competitive companies were also more likely to have created new departments/sections (36.7% compared to 13.2% in the Low group).

Acquisition of services from external agencies

Table 1 shows that in small and medium sized companies competitiveness was significantly associated with the “buying in” of financial/auditing services and training courses but not accountancy services. In large companies competitiveness was not related to the buying in of any of these services. It is not clear why the small/medium sized competitive companies were more likely buy in financial/auditing and training services (but not accountancy services) though the reason ***may*** be connected with the fact that 77% of these companies had graduate CEO’s while the corresponding figure for the low competitive companies of the same size was 47%. More highly educated CEO’s may consider that “in house” accountancy is essential for constant monitoring of the firm’s financial position, but that it is more cost-effective to buy in services that are only required periodically.

Table 1. Percentage of companies acquiring various types of financial and training services from external sources

Type of service aquired	Size of company < 200 empregados			Size of company > 200 empregados		
	Competitiveness			Competitivieness		
	Low	High	Sig.	Low	High	Sig.
Accountancy	52,6	27,6	0,001	10,0	10,3	N.S.
Financial/auditing services	28,2	41,8	0,05	80,0	74,4	N.S.
Technical consultancy services	29,5	38,8	N.S.	65,0	64,1	N.S.
Training	28,2	51,0	0,002	70,0	74,4	N.S.

Use of individual pay increases based on merit/jobperformance

60.9% of highly competitive companies used this type of pay increase compared to 42.2% of the low competitive. Its use was more common in highly competitive 'hi-tec' companies (69.3%) and less common in highly competitive 'low tec' companies (37.8%). Possibly 'hi-tec.' competitive companies use this type of increase because man-computer systems confer little advantage unless the human component performs well. Thus in such companies high quality individual performance is likely to be highly valued and rewarded.

Ssalary costs as a percentage of total costs

Average salary costs as a percentage of total costs were 24.1% in the High competitive group and 33.0% in the Low group. Average salary costs were not influenced by level of information technology in either group, but in the Low group salary costs as a percentage of total costs was influenced by size of company (31.5% for companies with less than 200 employees and 38.8% for those with more than 200).

Perceived importance of product price as a competitive advantage

The majority of companies in both the High and Low competitive groups perceived the price of their product to be a competitive advantage, but whereas 39.3% of the Low group believed it was a very important advantage only 26.8% of the High group believed this.

The analysis for the Oporto metropolitan area

Six factors were found to be associated with competitiveness and the regression model gave an acceptable fit to the data (χ^2 = 13.227, d.f. = 8, P = 0,104). It correctly classified the group membership of 72.9% of all companies (72.5% in the High competitive group and 73.3% in the Low group). The six factors were: maintenance of contacts /relations with Centres of R&D, use of the MRP method for planning production needs, changes in human resources management, salary costs as a percentage of total costs, use of individual pay increases based on contribution to the objectives of the company and use of general pay increases made in accordance with collective agreements.

Maintainance of contacts /relations with Centres of R&D

36.4% of companies in the Oporto area carried out their own R&D and 26.1% of the High competitive group maintained contact/relations with centres of R&D (8.9% in the Low group). But the tendency to keep links with R&D centres was influenced by size of company. 66.7% of large highly competitive companies (more than 200 employees) kept such links and 47.6% used Centres of R&D for assistance in solving production problems. In contrast only 13.4% of small highly competitive companies maintained links with Centres of R&D.

Use of the MRP method for planning production needs

The proportion of companies in the High competitive group that used MRP was twice that in the Low group (High 38.6% Low 17.7%). But use of MRP was especially strongly related to competitiveness in the group of ‘low tec.’ companies where 41.9% of the highly competitive used it compared to only 15.0% of the low competitive.

Change in Human resources management

The changes associated with competitiveness were, with one exception, made only by small and medium sized companies and concerned improved recruitment and selection (29.5% High group, 17.7% Low group) and improved conditions of hygiene and safety (45.5% High group and 29.0% Low group). The exception concerned change in the ways information was spread through the company, which was a change made by 47,6% of large highly competitive companies and 11.1% of large low competitive companies. For small and medium size enterprises the corresponding percentages were 25.0% (High) and 6.5% (Low).

S*alary costs as a percentage of total costs*

For companies in the High competitive group average Salary costs as a percentage of total costs was 20.2% while for those in the Low group it was 38.0%. The average salary costs/total cost ratio was not significantly influenced by size of company in either group, though in the Low competitive group it was influenced by the extent to which companies possessed information technology. Companies in the Low group having low levels of information technology had an average salary costs/total costs ratio of 38.9% while for those having high levels of information technology the ratio was 32.3%.

Use of individual pay increases based on contribution to the objectives of the company

In the High competitive group 43.2% of companies used this form of pay increase compared to 21.8% in the Low group. The use of this type of pay increase was not significantly influenced by size of company but was influenced by the level of information technology. In the group of 'low tec' companies 58.1% of the highly competitive used this type of pay increase compared to 17.5% of those that were low in competitiveness. Competitiveness was **not** associated with the use of individual pay increases based on the individual's contribution to the objectives of the company in the case of companies that used a great deal of information technology.

Use of general pay increases made in accordance with collective agreements.

The use of this type of pay increase was **negatively** related to competitiveness irrespective of size of company and level of information technology. In the Low competitive group 68.5% of companies used this form of pay increase compared to 45.5% in the High competitive group.

DISCUSSION

As expected, two types of factor were related to competitiveness -- **General Factors** which applied to companies in both metropolitan areas, and **Regional Factors** which were related to competitiveness in one area but not the other. An exception to this occurred in the case of Salary costs as a percentage of total costs. This appeared as both a General factor and a Regional factor in each area. The reason for this was that, unlike the other General factors, in the analysis for both metropolitan areas combined average Salary costs not only differed significantly for high and low competitive groups but also showed a significant interaction between competitiveness and metropolitan area.

Further analysis of the General factors revealed that the likelihood of a company being highly competitive depended not only upon inovation in terms of the introduction of new products or products improved in quality, but depended also upon ***how*** the inovation was achieved. For example, the conditional probability of a company being in the High competitive group given that it had introduced new product(s) in the previous five years was 0.615. If such a company were above the median in its level of information technology this probability increased to 0.716. If the company also carried out R&D the probability rose further to 0.766 and if, in addition, it had introduced new methods for planning, controlling and managing production the conditional probability of being in the High competitive group became 0.857. Similar increases occurred in the case of improved products (from 0.567 to 0.800).

The importance of ***how*** inovation was achieved can be seen in the fact that, for a company that had introduced new product(s) but had low information technology, did not carry out R&D and had not introduced new methods for planning, controlling and managing production, the conditional probability of being in the High competitive group was 0.313.

These findings suggest that competitiveness in the two metropolitan areas is strongly related to product inovation brought about by the use of R&D and the use of information technology for planning, controlling and carrying out production of new products and products improved in quality.

Competitiveness was also strongly related to reduced Salary costs as a percentage of total costs. In the Lisbon area the average salary costs/total costs ratio for highly competitive companies was 73% that of the low competitive companies, and in the Oporto area this figure was even lower (53%). These percentages indicate that highly competitive companies have a substantial advantage in terms of salary costs in relation to total costs when compared to the low competitive companies, though it was not clear from the research how this advantage was achieved. Clearly this is a point that needs a good deal more investigation.

Competitiveness was significantly associated with the acquisition of various types of specialized expertise from external agencies. These included (in both metropolitan areas) university services for assistance in solving specific types of problem, the use of financial /auditing services and training courses (in small/medium sized companies in the Lisbon area) and services of Centres of R&D (in large companies in the Oporto area). Perhaps this is not surprising. To be highly competitive companies need to use the best technical services they can and (especially for small/medium sized companies) it may not be economically appropriate for periodically required services to be furnished by full time 'in-house' experts.

Competitiveness was associated with the use of individual pay increases based upon some criterion of job performance. This type of pay increase tends to be perceived by employees as recognition of their individual value to the company and it has long been known that recognition of this sort increases worker motivation -- and to be highly competitive a company needs a well motivated workforce. In contrast, the use of general pay increases made in acccordance with collective agreements was **negatively** related to competitiveness. This type of pay increase is seldom perceived by employees as recognizing their individual value to the company, and although such pay increases may operate to prevent or reduce dissatisfaction they tend not to have a positive motivating effect.

In the Lisbon area certain organizational changes in small/medium sized companies were associated with competitiveness; in particular, the creation of working groups/teams, decentralization of decision making and improvement in the flow of information in the company. These seem to represent at least a partial shift towards improved use of human resources through job enlargement/enrichment. In the Oporto area more basic changes in human resource management were associated with competitiveness – improved recruitment/selection of employees, improvements in hygiene and safety and changes in the way information was spread through the company.

Foreign Trade, Devaluation and Elasticities: A Model Approach

prof. Ing. Jaroslav Husár, CSc. Mgr.
ek. University of Economics, Bratislava

Introduction

From the theory of the international economics it is knowwn that the expenditure-switching policies primarily work by changing relative prices. The main form of such a policy is a change in exchange rates. i.e. a devaluation or a revaluation of the domestic currency. Direct controls can also be classified under this heading and are usually applied to restrict imports. Consumers will then try to buy domestic goods instead of imported goods, and hence direct controls can be viewed as a switching device.

We shall, however, concentrate on a discussion of devaluation (Slovakia has experienced it in 1993) to show how important is the theory to indicate the desired economic policy.

Trade balance and the devaluation: the model

The traditional approach to the effects of devaluation on the balance of trade runs in terms of elasticities (Bo Sodersten). The core of the traditional approach is contained in so-called Marshall-Lerner condition, which states that the sum of elasticities of demand for a country's export and of its demand for imports has to be greater than unity for a devaluation to have a positive effect on a country's trade balance (A.P. Lerner). If the sum of these elasticities is smaller than ubity, a country can instead improve its balance of trade by revaluation.

Bo Sodersten showed us that if we want to express this condition in terms of formula, it can be set out as follows:

$$dB = kX_{f}(e_{1m} + e_{2m} - 1) \qquad (1)$$

where dB is the change in the trade balance, k the devaluation in percentage, X_f the value of exports expressed in foreign currency, e_{1m} the devaluing country's demand elasticity for imports (first country in the next text, reason for 1), and e_{2m} the second country's (say the rest of the world) demand elsticity for export from the devaluing country (Slovakia).

It is easy to see from the expression (1) that the sum of the two critical elasticities has to be larger than unity for the trade balance to improve because of a devaluation. If the sum is less than unity, an appreciation should instead be used to cure a deficit in the trade balance.

We in Slovakia have experienced that the devaluation leads to an increase in the price of imports. What the effect of this price increase will be depends on the elasticity of demand for imports. The larger it is, the greater will be the fall in the volume of imports. The value of the demand elasticity of imports depends, of course, on what type of goods the devaluing country imports. If a country primarily imports necessities, raw materials and goods needed as inputs for its industries (Slovakia is a typical country), the demand elasticity of imports may be very low, and a devaluation may not be a very efficient means of correcting deficit.

We want to mention that the Marshall-Lerner condition set out in formula (1) is built on some strict simplifications. It assumes, roughly, that the supply elasticities are large and that the trade balance is in equilibrium when devaluation takes place.

Let us start by setting out the following equation for the trade balance:

$$B_{1f} = x_1 P_{2m} - m_1 P_{2x} = X_{1f} - M_{1f} \qquad (2)$$

where B_{1f} denotes the devaluing country's trade balance in foreign currency, where x_1 and m_1 are devaluing country's volume of exports and imports, respectively; P_{2m} and P_{2x} are the prices of imports and exports in second country (rest of the world); and X_{1f} and M_{1f} are the value of exports and imports in country one (devaluing), both denoted in foreign currency.[1]

Next we need to get is the differentiating of equation (2). Differentiating gives

[1] The principle of this approach was first developed by S.S. Alexander

$$dB_{1f} = dx_1P_{2m} - dP_{2m}x_1 - dm_1P_{2x} - dP_{2x}m_1$$

$$= X_{1f}\left(\frac{dx_1}{x_1} + \frac{dP_{2m}}{P_{2m}}\right) + M_{1f}\left(-\frac{dm_1}{m_1} - \frac{dP_{2x}}{P_{2x}}\right) \tag{3}$$

Now we define these four elasticities:

$$s_{1x} = \frac{dx_1}{dP_{1x}}\frac{P_{1x}}{x_1}\text{, the elasticity of home export supply} \tag{4}$$

$$e_{2m} = -\frac{dx_1}{-d_{2m}}\frac{P_{2m}}{x_1}\text{, the elasticity of foreign demand for exports} \tag{5}$$

$$s_{2m} = \frac{dm_1}{dP_{2x}}\frac{P_{2x}}{m_1}\text{, the elasticity of foreign supply of imports} \tag{6}$$

$$e_{1m} = -\frac{dm_1}{dP_{1m}}\frac{P_{1m}}{m_1}\text{, the elasticity of home demand for imports} \tag{7}$$

We observe from the way in which these four elasticities have been defined that that they will all be positive – barring Giffen goods.

The theory assumes that we have price equalization between the two countries - expected case in market economies - through the exchange rate, r, so that we get

$$P_{2x} = P_{1m}r \tag{8}$$

Differentiaiting equation (8) totally and adding in equation (8) gives

$$P_{2x} + dP_{2x} = P_{1m}r + drP_{1m}$$

$$= (P_{1m} + dP_{1m})r - k(P_{1m} + dP_{1m})r \tag{9}$$
$$= (P_{1m} + dP_{1m})r(1 - k)$$

In equation (9) we have introduced the devaluation coefficient k, which shows the relative change in the exchange rate. We can define k in the following way:

$$k = -\frac{P_{1m}}{P_{1m} + dP_{1m}}\frac{dr}{r} = -\frac{dr}{r}\frac{1}{1+\frac{dP_{1m}}{P_{1m}}}$$

$$\approx -\frac{dr}{r}\left(1 - \frac{dP_{1m}}{P_{1m}}\right) \approx -\frac{dr}{r} \tag{10}$$

From equation (9) we can get

$$\frac{dP_{2x}}{P_{2x}} = -k + \frac{dP_{1m}}{P_{1m}}(1-k) \tag{11}$$

In a completely analogous way we deduce that

$$\frac{dP_{2m}}{P_{2m}} = -k + \frac{dP_{1x}}{P_{1x}}(1-k) \tag{12}$$

The relative changes in volumes and prices can now be expressed in terms of elasticities and the devaluation coefficient k. Using equations (5) and (12) we get

$$\frac{dx_1}{x_1} = -e_{2m}\frac{dP_{2m}}{P_{2m}} = -e_{2m}\left[-k + \frac{dP_{1x}}{P_{1x}}(1-k)\right] \tag{13}$$

But $dx_1/x_1 = s_{1x}(dP_{1x}/P_{1x})$. By substitution we get

$$\frac{dx_1}{x_1} = e_{2m}k - \frac{e_{2m}}{s_{1x}}(1-k)\frac{dx_1}{x_1}$$

From this follows

$$\frac{dx_1}{x_1} = \frac{e_{2m}k}{1+(e_{2m}/s_{1x})(1-k)} = \frac{s_{1x}e_{2m}k}{s_{1x}+e_{2m}(1-k)} \tag{14}$$

In an analogous way we can derive

$$\frac{dP_{2m}}{P_{2m}} = -\frac{ks_{1x}}{s_{1x}+e_{2m}(1-k)} \tag{15}$$

$$\frac{dm_1}{m_1} = -\frac{ks_{2m}e_{1m}}{e_{1m}+s_{2m}(1-k)} \tag{16}$$

$$\frac{dP_{2x}}{P_{2x}} = -\frac{ke_{1m}}{e_{1m}+s_{2m}(1-k)} \tag{17}$$

Now using the last four expressions we get the effect of a devaluation on the trade balance:

$$dB_{1f} = k\left[X_{1f}\,\frac{s_{1x}(e_{2m}-1)}{s_{1x}+e_{2m}(1-k)} + M_{1f}\,\frac{e_{1m}(s_{2m}+1)}{e_{1m}+s_{2m}(1-k)}\right] \tag{18}$$

What does the (18) says? Expression (18) shows that the effects of the devaluation are somewhat more complicated than shown in equation (1); i.e. if we do not assume that supply elasticities are infinitely large the situation becomes *somewhat more complex*. If, to take an extreme example, we assumed that the supply elasticities were equal to zero there would be no improvement in the trade balance because of increasing exports but some improvement because of fall in demand for imports.

Generally speaking, we can say that if the elasticities are larger than unity, then the larger they are, both on the supply side and the demand side, the larger will be the improvement in trade balance.

Sindy Alexander showed the way to arrive at formula (1) from (18). It is as follows: if supply elasticities tend to infinity, then

$$\frac{e_{2m} - 1}{1 + (e_{2m} / s_{1x})(1-k)} \to e_{2m} - 1$$

If, furthemore, k is small, we get

$$\frac{e_{1m}\left[1 + (1/s_{2m})\right]}{(e_{1m} / s_{2m}) + 1 - k} \to \frac{e_{1m}}{1-k}$$

But if k is *small* and if we assume that trade is balanced before the devaluation, we get

$$dB_{1f} = kM_{1f}(e_{2m} + e_{1m} - 1) \qquad (1)$$

We have now set out the main parts of the elasticity approach to devaluation. The dubious aspect of this approach is that it is built on a parital type of theorizing and that it does not take into account consideration of general equilibrium.

Conclusions

Demand and supply elasticities are conventionally defined *ceteris paribus*, i.e. other prices and incomes are supposed to be constant, but in devaluation prices and incomes will certainly change. Therefore, the use of partial elasticities in connection with devaluation can easily be misleading. What one would like to know is the value of the "total" elasticities, i.e. the value of an elasticity when all the factors involved in the devaluation change. Such a total elasticity measures how quantities are affected by price changes when everything likely to change has done so. This is, however, not an operational concept, as it will never be possible to know in advance the values of such elasticities.

The result of devaluation depends not only on partial elasticities but also on the aggregate behavior of the economic system. In the literature is known an alternative approach to the effects of devaluation in macro terms. It is known as the *absorption approach*.

Literature

1. Baumol B. W. and Blinder A.S: Economics, Harcourt Brace J. New York, 1994
2. Husar J.: Makroekonomia, Kartprint, Bratislava 1998
3. Sondersen Bo: International Economics, Macmillan, London, 1980

Jaroslav Husár
Department of Operations Research and Econometrics,
University of Economics
Dolnozemská 1/b Bratislava, Slovak Republic

A Comparison of Two Solution Approaches to the Analytic Hierarchy Process

Michal Chobot, Lea Škrovánková
University of Economics, Bratislava, Slovak Republic

1. Introduction

The Analytic Hierarchy Process (AHP) is according to Saaty [5] a general theory of measurement applied within elaborate ordered structures to derive ratio scales from paired comparisons on a group of homogeneous elements with respect to common attributes and synthesize these scales to a unidimensional ratio scale. Such a scale enables one to rank the activities involved in the structures, make choices and allocate resoursec to these choices in proportion to their derived ranks. Many problems involve the assignment of priorities to a set of objects, projects, products, etc.

Preferences are expressed in a form of pair-wise comparisons, rather than as weights. The AHP has been used in numerous applications in decision making, planning, benefit / cost and risk analysis and many other multicriteria problems. Our special interest was caused by the work of Gass [3], who applied this technique to determine priorities and weights for large scale linear goal programming. However, when we studied the AHP problem in detail, we found out that different solutions can lead to different results. In this note we want to demonstrate this by a comparison of two solution approache to the AHP problem.

2. Formulation of the AHP problem and description of solutions

The AHP problem can consist of several subproblems. For our analysis we will use only the following basic formulation. Let $A = (a_{ij})$ is *n* x *n* matrix satisfying:

$$0 < a_{ij} = \frac{1}{a_{ji}}.$$

These are approximate judgments of the ratios of the underlying weights associated with objects *i* and *j*. The purpose is to retrieve weights w_i from A such that w_i / w_j will approximate a_{ij} in some sense.

Let $w = (w_k)$ is an *n* – dimensional vector satisfying

$$w_k > 0, \quad \prod_{k=1}^{n} w_k = 1 \tag{1}$$

and $V = (v_{ij})$ is any *n* x *n* matrix satisfying $v_{ij} = w_i / w_j$ for some vector w. The expression (1) represents multiplicative normalization induced by the multiplicative nature of the problem. The data are given in terms of ratios and A is consistent if and only if for all *i, j* and *k;* $a_{ij} = a_{ik} \cdot a_{kj}$.

For our comparison we are interested in two solutions of our AHP problem. The first approach to this problem was proposed by Saaty. It consists in taking as weights the components of the (right) eigenvector of the matrix A. In our notation the eigenvector is defined by

$$\bar{e}(A) = w, \quad Aw = \lambda_{max} w\,, \tag{2}$$

where λ_{max} is the largest eigenvalue of A. It must be noted that this eigenvector solution is normalized additively, i. e.

$$\bar{e}(A) = w, \quad Aw = \lambda_{max} w, \quad \sum_{k=1}^{n} w_k = 1. \tag{3}$$

A number of authors have suggested alternatives to this approach. For example, the eigenvector method was criticized in [2]. The authors pointed out that there is a little justification for the use of eigenvectors to derive the requisite weights. There are also other mathematical arguments against using the eigenvector to estimate the underlying scale for this problem. In [1] the authors indicate that the eigenvector solution violates the sense of consistency and does not take into account the multiplicative nature of the problem.

The second solution to the AHP problem is the geometric mean defined by

$$g(A) = w, \quad w_i = \Big(\prod_{j=1}^{n} a_{ij} \Big)^{\frac{1}{n}} . \tag{4}$$

The authors of [1] proposed an axiomatic approach of this solution. They have shown that this solution satisfied a set of basic consistency axioms. Moreover, in [2] it is shown that the geometric mean can be derived from the solution of two optimization problems. The geometric mean solution can be derived as the solution of the following optimization problem:

$$\min_{w} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\ln a_{ij} - \ln\left(\frac{w_i}{w_j}\right)\right]^2, \tag{5}$$

subject to

$$\prod_{i=1}^{n} w_i = 1, \quad w_i > 0, \quad i = 1, 2, \ldots, n.$$

The geometric mean can be also derived with the help of an information theory approach. If one normalizes **w** then it can be viewed as a probability measure. In an information – theoretic sense the problem is to determine the vector **w** that is as close as possible to the collumns of A. For that purpose one can use the minimum discrimination technique (basic notions of MDI measures we described in [4]). The corresponding optimization problem becomes

$$\min_{w} f(w) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \ln\left(\frac{w_i}{a_{ij}}\right) \tag{6}$$

subject to

$$\sum_{i=1}^{n} w_i = 1.$$

We will not pay attention here to a detail analysis of these problems, instead of this we present an illustrative example.

3. An Example

For our purpose we selected the school selection example given in [6]. The author assumes that the decision maker is faced with comparison of high schools **A**, **B** and **C** according to their desirability. The comparison is based on six criteria or characteristics: learning (L), friends (F), school life (S), vocational training (V), college preparation (C) and music classes (M). The pairwise judgment matrix comparing the characteristics with respect of overall satisfaction with school is described in Table 1. It is labeled accordingly with **O** in its left-top corner. Table 2 contains the six 3x3 judgment matrices corresponding to these six criteria, each labeled in its left-top corner.

O	L	F	S	V	C	M
L	1	4	3	1	3	4
F	1/4	1	7	3	1/5	1
S	1/3	1/7	1	1/5	1/5	1/6
V	1	1/3	5	1	1	1/3
C	1/3	5	5	1	1	3
M	1/4	1	6	3	1/3	1

Table 1.

L	A	B	C	F	A	B	C	S	A	B	C
A	1	1/3	1/2	A	1	1	1	A	1	5	1
B	3	1	3	B	1	1	1	B	1/5	1	1/5
C	2	1/3	1	C	1	1	1	C	1	5	1
V	**A**	**B**	**C**	**C**	**A**	**B**	**C**	**M**	**A**	**B**	**C**
A	1	9	7	A	1	1/2	1	A	1	6	4
B	1/9	1	1/5	B	2	1	2	B	1/6	1	1/3
C	1/7	5	1	C	1	1/2	1	C	1/4	3	1

Table 2.

The eigenvector solution computed by Saaty for matrices in Table 1 and Table 2 is given in Table 3 (labeled as $\bar{e}$). Here the first three entries in each row are components of the eigenvector of the corresponding matrix in Table 2. The six number in the fourth column labeled as **O** are the components of the eigenvector of the matrix in Table 1 (here the results are rounded from the original table). All eigenvectors are normalized additively. The combined overall solution vector is given by

$$\mathbf{w} = (0.37,\ 0.38,\ 0.25).$$

ē	**A**	**B**	**C**	**O**
L	0.16	0.59	0.25	0.32
F	0.33	0.33	0.33	0.14
S	0.455	0.09	0.455	0.03
V	0.77	0.05	0.18	0.13
C	0.25	0.5	0.25	0.24
M	0.69	0.09	0.22	0.14

Table 3.

The first component of **w** is the weighted arithmetic mean of the column **A** in the Table 3, with the column **O** entries serving as weights. Similarly, the two other components of **w** are computed. The optimal choice by this solution approach is thus high school **B** with $w_2 / w_1 =$ 1.03. Now we will use the geometric mean solution to the same problem. It is given in the same format in Table 4. The solution to the matrix in Table 1 is given in the column **O**. It is normalized additively (the sum of its components is 1) since it is not in the lowest level of the hierarchy. Solutions to the matrices in Table 2 are normalized multiplicatively – the product of component is 1, since it is in the lowest level of the hierarchy. The combined overall solution vector **w** is the geometric mean of these vectors with the components of **O** vector serving as weights

$$w = w_L^{O_1} \cdot w_F^{O_2} \cdot w_S^{O_3} \cdot w_V^{O_4} \cdot w_C^{O_5} \cdot w_M^{O_6} .$$

All operations are componentwise. All six vectors are 3-dimensional with the components of w_L being in first three entries of the row **L** of Table 4 , etc. The solution is also 3-dimensional and is given by **w** = (1.11, 1.001, 0.899). It clear that by construction vector **w** satisfied the multiplicative normalization. The optimal choice in this case is high school **A**, with w_1 / w_2 = 1.11.

w	A	B	C	O
L	0.55	2.08	0.87	0.32
F	1	1	1	0.14
S	1.71	034	1.71	0.03
V	3.98	0.28	0.89	0.12
C	0.79	1.59	0.79	0.24
M	2.88	0.38	0.91	0.15

Table 4.

4. Concluding remarks

In this note we compared two solution approaches to the AHP problem. We emphasized that a number of authors provided many arguments against using the eigenvector as a solution of the AHP problem. The eigenvector can lead to a wrong solution because it violates basic sense of consistency. Generalization of the eigenvector method can lead to nonuniqueness, which produces different solutions depending on whether one uses the right or left eigenvector.

The alternative solution obtained by the geometric mean eliminates these shortages. Besides the geometric mean can be derived by several approaches. We point out here derivations of the geometric mean as the solution of the optimization problems (5) and (6). These approaches also allow interpretations of the dual solutions. It will also interesting to compare solutions of the AHP problem obtained from these optimization problems with previous solutions.

References

[1] Barzilai J., Cook W. D., Golany B.: Consistent Weights for Judgments Matrices of the Relative Importance of Alternatives. Operation Research Letters 6, 1987.

[2] Crawford G., Williams C.: A Note on the Analysis of Subjective Judgment Matrices. Journal of Mathematical Psychology 29, 1985.

[3] Gass S.: A Process to Determine Priorities and Weights for Large-Scale Linear Goal Programming. The 12th. International Symposium on Mathematical Programming. Boston, 1985.

[4] Chobot M., Škrovánková L.: The Use of Information theory in Marketing models. The 6th. International conference Quantitative Methods in Business and Management. Bratislava, 1999.

[5] Saaty T. L.: The Analytic Hierarchy Process. McGrow-Hill, New York, 1980.

[6] Saaty T. L.: The Analytic Hierarchy Process: A 1993 Overview. The International workshop on multicriteria decision making. Methods, Algorithms, Applications. Celákovice, 1993.

Michal Chobot
Department of Operations Research and Econometrics,
University of Economics
Dolnozemská 1/b Bratislava, Slovak Republic

Date of Speculative Attack-Crises of Exchange Rates

Ivanicová Zlatica, University of Economics Bratislava

A fundamental proposition of the open economy macroeconomics is that viability of a fixed exchange rate regime requires maintaining long-run consistency between monetary, fiscal and exchange rate policies. Excessive domestic credit growth leads to a gradual loss of foreign reserves and ultimately to an abandonment of the fixed exchange rate. The central bank is enable of defending the parity any longer.

In paper, Krugman (1979) showed that under a fixed exchange rate regime domestic credit creation in excess of money demand growth leads to a sudden speculative attack against the currency that forces the abandonment of the fixed exchange rate and the adoption of a flexible rate regimes.

During the year 1998 the Slovak currency was overvalue, the purchasing power parity and interest rate parity, respectively did not valid. In August 1998 reserves of the foreign currencies in the Slovak National Bank fall to the lower bound. The result was the devaluation of the currency and from the October 1998 the fixed exchange rate was changed to the flexible one.

Over past year, a large theoretical approaches has focused on the short- and long-run consequences of incompatible macroeconomics policies for the balance of payments of a small open economy in which agents are able to anticipate future decision by policymakers.

In our paper we present a single-good full employment small open economy model that specifies the basic theoretical framework used for analyzing balance of payments crises. The model provides an explicit calculation of the time of occurrence of the crisis by assuming that the exchange rate in the post-collapse regime is allowed to float permanently.

A model of a small open economy

Consider a small open economy whose residents consume a single tradable good. Domestic supply of the good is exogenous, and its foreign-currency price is fixed. The domestic-price level is equal, through a purchasing power parity condition, to the nominal exchange rate. Agents hold three categories of assets: domestic money (which is not held abroad) and domestic and foreign bonds, which are perfectly substitutable. There are no private banks, so that the money stocks is equal to the sum of domestic credit issued by the central bank and the domestic-currency value of foreign reserves held by the central bank. Foreign reserves earn no interest, and domestic credit expands at a constant growth rate, agents are endowed with perfect foresight.

The model is defined by the following set of equations:

$$m_t - p_t = \varphi y - \alpha i_t \qquad \varphi, \alpha > 0 \qquad (1)$$

$$m_t = \gamma D_t + (1 - \gamma) R_t \qquad 0 < \gamma < 1 \qquad (2)$$

$$D_t' = \mu \qquad \mu > 0 \qquad (3)$$

$$p_t = e_t \qquad (4)$$

$$i_t = i^* + e_t' \qquad (5)$$

Where: m_t is the nominal money stock,

D_t is domestic credit,

R_t is the domestic currency value of foreign reserves held by the central bank,

e_t is the spot exchange rate,

p_t is the price level,

y is exogenous output,

$i*$ is the foreign interest rate (assumed constant),

i_t is the domestic interest rate.

All variables except interest rates are measured in logarithms.

Equation (1) defines relation between the real demand for money and the domestic interest rate. Equation (2) denotes a logarithm-linear approximation of the money stock as the stock of reserves and domestic credit. The money stocks grow at the rate μ, equation (3). Equations (4) and (5) relate, respectively, purchasing power parity and uncovered interest rate parity.

Assume that $\delta = \varphi y - \alpha i^*$ and combining equations (1), (4) and (5) yields

$$m_t - e_t = \delta - a e_t^*, \qquad \delta > 0. \qquad (6)$$

Under a fixed nominal exchange rate regime, $e_t = e$ and $e_t' = 0$ so that

$$m_t - e = \delta, \qquad (6')$$

which indicates that the central bank accommodates any change in domestic money demand through the purchase or sale of foreign reserves to the public. Using equation (2) and (6') yields

and using (3)

$$R_t = \frac{d + e - gD_t}{1 - g} \qquad (7)$$

$$R_t' = -\frac{m}{q}, \qquad q = \frac{1-g}{g} \tag{8}$$

Equation (8) indicates that if domestic credit expansion is excessive, reserves are run down at the rate proportional to the rate of credit expansion. Any finite stock of foreign reserves will therefore be exhausted in finite of time.

Suppose that the central bank announces at time t that it will stop defending the current fixed exchange rate after reserves reach a low bound R_D, at which point it will withdraw from the foreign exchange market and allow the exchange rate to float freely. With a positive rate of domestic credit growth, rational agents will anticipate that without speculation reserves will eventually fall to the lower bound (zero) and will foresee therefore the ultimate collapse of the system. To avoid losses at the time of collapse, speculators will force a crisis before this point is reached. Then problem is to determine the exact moment at which the fixed exchange rate regime is abandoned or, equivalently, the time of transition to a floating rate regime.

The shadow floating exchange rate

The shadow floating rate is the exchange rate that would achieve if reserves had fallen to minimum level and the exchange rate were allowed to float freely. From this point of view, the time of collapse is found at the point where the shadow floating rate, which reflects market fundamentals, is equal to the prevailing fixed rate. If depreciated fixed exchange rate exceeds the shadow-floating rate, the fixed rate is not viable.

In equilibrium, under perfect foresight, agents can never expect a discrete jump in the level of the exchange rate since a jump would provide them with profitable arbitrage opportunities. If the shadow floating rate falls below the prevailing fixed rate, speculators would not profit from purchasing the government's reserves stock and precipitating the adoption of a floating rate regime, since they would experience an instantaneous capital loss on their purchases. Symmetrically, if the shadow-floating rate is above the fixed rate, speculators would experience an instantaneous capital gain. Competition of the speculators eliminate such opportunities and their behaviour leads to an equilibrium attack, which incorporates the arbitrage condition that the pre-attack fixed rate should equal the post-attack floating rate.

The shadows floating exchange rate is achieved through following considerations:

1. The floating exchange rate is equal

$$e_t = k_0 + k_t m_t , \tag{9}$$

where k_0 and k_t are undetermined parameters.

2. Nominal level stock, when reserves reach their lower level R_D, is equal

$m_t = \gamma D_t + (1 - \gamma) R_D$, and under floating rate regime $m_t' = \gamma D_t'$ where $D_t' = \mu$.

From this equations yields

$$e_t' = k_1 gm \tag{10}$$

In the post-collapse regime, therefore, the exchange rate depreciates steadily and proportionally to the rate of growth of domestic credit. Substituting (10) in (6) yields, with $d=0$ for simplicity,

$$e_t = m_t + \alpha k_1 \gamma\mu. \tag{11}$$

Comparing equation (11) and (9) yields

$$k_0 = \alpha \gamma\mu. \qquad k_1 = 1.$$

From equation (3), $D_t = D_0 + \mu t$. Using the definition of m_t given above and substituting in equation (11) yields

$$e_t = g (D_0 + am) + (1 - g) \underline{R} + gmt. \tag{12}$$

Date of speculative attack

The fixed exchange rate regime collapses when the prevailing parity e equals the shadow floating rate e_t . From (12) the exact time of collapse T is obtained by setting $e = e_t$, to that

$$T = \frac{e - gD_0 - (1-g) R_D}{gm} - a$$

or, since from equation (2) and (6') $e = -gD_0 + (1 - g)R_0$,

$$T = \frac{q (R_0 - R_D)}{m} - a \tag{13}$$

where R_0 denotes the initial stock of reserves.

Equation (13) indicates that, the higher the initial stock of reserves is, the lower the critical level, or the lower the rate of credit expansion is, the longer it will take before the collapse occurs. Without speculation, α=0 and the collapse occurs when reserves are run down to the minimum level. Setting α=0 in equation (13) yields the time of ”natural collapse”. The interest rate elasticity of money demand determines the size of the downward shift in money demand and reserves that takes place when the fixed exchange rate regime collapses and the nominal interest rate jumps to reflect an expected depreciation of the domestic currency. The larger α is, the earlier is the crisis.

The analysis therefore implies that the speculative attack always occurs before the central bank would have reached the minimum level of reserves in the absence of speculation. To determine the stock of reserves just before ate attack (Th - marginal) use equation (7) to obtain, with δ=0,

$$R_{Th} \equiv \lim R_t = \frac{e - gD_{Th}}{1-g}, t \to T_{Th}$$

where $D_{Th} = D_0 + mTh$, so that

$$R_{Th} = \frac{e - g\,(D_0 + mTh)}{1-g}. \tag{14}$$

Using equation (13) yields

$$e - gD_0 = gm(Th + a) + (1 - g)R_D. \tag{15}$$

Combining (14) and (15) finally yields

$$R_{Th} = R_D + \mu\alpha/\theta. \tag{16}$$

References:

[1] Ingram J.C., Dunn R.M.: International Economics, John Willey & Sons, Inc. New York, 1993

[2] Jong de Eelke.: Exchange rate Determination: Is There a Role for Macroeconomic Fundamentals? De Economist 145, No 4, 1997

[3] Bacer J.C.: Selected International Investment Portfolios, Amstedram, Elsevier Science 1998

Zlatica Ivanicová
Department of Operations Research and Econometrics,
University of Economics
Dolnozemská 1/b Bratislava, Slovak Republic

MS Excel based system for multicriteria evaluation of alternatives[1]

Josef Jablonský, Pavel Urban
University of Economics Prague, Department of Econometrics

1 Introduction

Multicriteria evaluation of alternatives belongs among the basic decision problems of multicriteria decision making with very large possibilities of real applications (evaluation of investment alternatives, evaluation of credibility of bank clients, rating of companies, consumer goods evaluation and many others). The theory of multicriteria evaluation of alternatives is very good established and there are available many different methods for this kind of problems. From the decision maker's point of view the wider utilisation of multicriteria evaluation techniques depends on the availability of the powerful and user-friendly oriented software tools. The paper describes simple application for multicriteria evaluation of alternatives developed under MS Excel environment as standard add-in application written in Visual Basic for Application (VBA). The name of the system is SANNA (System for ANalysis of Alternatives).

The multicriteria evaluation of alternatives problem is usually defined by criterion matrix as follows:

$$\begin{array}{c} \\ X_1 \\ X_2 \\ \vdots \\ X_n \end{array} \begin{array}{c} \begin{array}{cccc} Y_1 & Y_2 & \cdots & Y_k \end{array} \\ \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1k} \\ y_{21} & y_{22} & \cdots & y_{2k} \\ \vdots & & \ddots & \\ y_{n1} & y_{n2} & \cdots & y_{nk} \end{bmatrix} \end{array}$$

where $X_1, X_2, ..., X_n$ is the set of n alternatives to be evaluated by k criteria $Y_1, Y_2, ..., Y_k$ and y_{ij}, i=1,2,...,n, j=1,2,...,k is the criterion value of the alternative X_i with respect to its evaluation by criterion Y_j. The criterion values are the basic input values for solving multicriteria evaluation of alternatives problems.

Multicriteria decision problems must always be solved with respect to the final objective of the evaluation. The objective of the evaluation can influence the selection of the solving

[1] The paper is supported by the Grant Agency of Czech Republic – grant no. 402/98/1488 and corresponds to the research program of the Faculty of Informatics and Statistics no. CEZ:J18/98:311401001.

technique and information that must be specified before or during the evaluation process. The basic objectives of the evaluation of alternatives are:

- selection of the “best” alternative,
- complete ranking of alternatives,
- classification of the alternatives (splitting of the alternatives into several classes, e.g. ABC analysis),
- selection of the subset of “good” alternatives.

The SANNA system does not contain an expert system that could help to decision makers to select the best solution technique for solving their problems depending on the objectives of the evaluation. However, the evaluation methods containing in the SANNA system belong among the standard and most often used techniques and can usually be used in most cases.

All the decision methods supported by SANNA system expect the knowledge of the information about the decision maker’s preferences of the criteria in the quantified form. This means that the decision maker must input the weights of the criteria – w_1, w_2, …, w_k. The specification of the weights is facilitated by the special module of SANNA that helps to estimate the weights by simple well-known methods.

2 General structure of the application

SANNA is the standard MS Excel add-in application. After the user opens this application the new SANNA menu item is added into the basic Excel menu bar. Simultaneously the SANNA toolbar with buttons for easy working with the application is opened (see Fig. 2). The structure of the SANNA menu corresponds to the structure of the application. The SANNA system consists of five basic modules plus simple help tool (each module corresponds to one item of SANNA menu):

1. *Data management module*.

   SANNA works with standard .xls worksheets. Input data information of the decision problem with the criterion matrix is placed in this worksheet in the sheet named “data”. Data management module either opens new data sheet (if does not already exists) based on specification of the basic parameters of the problem (the number of alternatives and criteria) or deals with the current data sheet. The basic functions for dealing with the current data set are adding of the new alternative or criteria or removing the alternative or criteria marked by the active cell of the sheet.

2. *Non-dominance filter*.

   This module makes it possible to test the non-dominance of alternatives in the current data set. The alternatives are marked according to the results of the test as *dominated* or *non-dominated*. Another feature of the module is the possibility to remove the dominated alternatives from the data set. The removing of dominated alternatives is non-reversible.

3. *Estimation of weights module*.

   The decision maker can specify the weights of the criteria either directly by their setting in the data sheet or by means of three simple well-known methods supported by estimation of weights module. The decision maker can use the following techniques:

   - scale method (weights are set in any numerical scale and normalised),
   - Fuller method (based on paired comparisons of the criteria and selection one of the following: one of the criteria is more important or both the criteria are equally important),
   - AHP (based on the standard Saaty's scale) – uses either eigenvector calculation or logarithmic least square method (approximation of the eigenvector); the consistence of the comparison matrix is checked.

   The estimation weights module contains the function that transmits the weights calculated by above mentioned methods into the data sheet and the function than graphically presents the weights.

4. *Methods for multicriteria evaluation of alternatives*.

   The current version of the SANNA system supports five basic methods for multicriteria evaluation of alternatives (WSA, TOPSIS, ELECTRE I, PROMETHEE I, II and MAPPAC). All the methods need the knowledge of the weights of the criteria. ELECTRE and PROMETHEE class methods require specification of some additional parameters. A brief characteristics of the available methods is following:

   <u>*WSA*</u>

   The WSA method is based on linear utility function. This method computes the global utility of the alternatives as the weighted sum of normalised criterion values. The method provides complete ranking of alternatives according to their global utilities.

TOPSIS

The TOPSIS method is based on the computation of global utilities of alternatives according to their closeness to the ideal criteria values and distance from the nadir criteria values. Analogous to the WSA the TOPSIS method provides complete ranking of alternatives according to their global utilities.

ELECTRE I

The basic output of the ELECTRE I method is the split of the set of alternatives into two disjoint subsets of efficient and inefficient alternatives. Except the criterion matrix and weights of the criteria the user must specify a concordance and discordance thresholds. The final result of the evaluation (subset of efficient and inefficient alternatives) is often very sensitive on the above mentioned thresholds. The user can repeat the calculation with different sets of thresholds and compare the results.

PROMETHEE I and II

The PROMETHEE methods need the knowledge of criterion matrix, weights of the criteria and preference functions of criteria with their parameters for measuring the strength of the preference of the pairs of alternatives with respect to the given criterion. SANNA uses the same types of preference functions as proposed by authors of the method in [1]. PROMETHEE I method provides partial ranking of alternatives, PROMETHEE II method offers complete ranking according to the net flow values.

MAPPAC

The MAPPAC method works with the criterion matrix and weights of the criteria only. The method splits the alternatives into several preference groups. The evaluation of alternatives by this method can be very long for problems with several tens of alternatives comparing to other supported approaches.

5. *Report wizard.*

The successfully completed evaluation by any of the methods adds a new sheet into the active worksheet with basic information about the results. By report wizard the user of the application can build his own output sheet with the results in the user-defined form. The user selects required items for the report from the set of items associated to the current method (including simple graph presentations of results). The basic window of the report definition for TOPSIS method is presented on Fig. 3.

6. *Help*.

   SANNA contains simple help window that can be activated from the SANNA menu or by the appropriated button on the SANNA toolbar.

3 Using the application

The application can solve multicriteria decision problems with maximum 100 alternatives and 50 evaluation criteria. Typical sequence of steps in processing a decision problem by SANNA is as follows:

1. Building the new data set. The user works with any existing spreadsheet file and calls the command *Data-New data set*. This command displays the dialogue window *Parameters of the problem* (Fig. 1) and the user must specify number of alternatives and criteria. Then a new sheet “data” containing the table for input/edit of the problem is added into the active worksheet. The “data” sheet for a problem with 6 alternatives and criteria is presented on Fig.2. Except the basic form of the “data” sheet fig.2 shows the basic menu of SANNA and its set is control buttons. Once the worksheet with the “data” sheet is created it can be saved and used at any time.

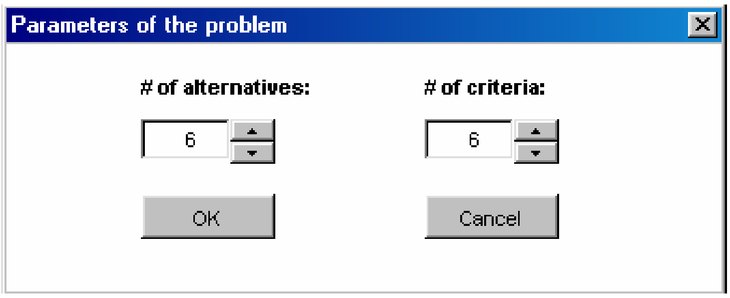


Fig. 1 – SANNA – Specification of parameters of the problem.

2. Test of non-dominance of the set of alternatives and possibly reduction of the set of alternatives by removing of dominated ones.

3. The weights of criteria can be inserted either directly into the “data” sheet or by means of one of the available methods. The results of calculation of weights are placed into new

sheet that is added in to the worksheet (this sheet is named in accordance with the name of the method used). After the weights are calculated they can be transmitted into the "data" sheet by command *Weights-Apply weights*. For this command the sheet with results of weights calculation must be active. Weights contained in the "data" sheet are used as inputs for all of the available multicriteria evaluation methods.

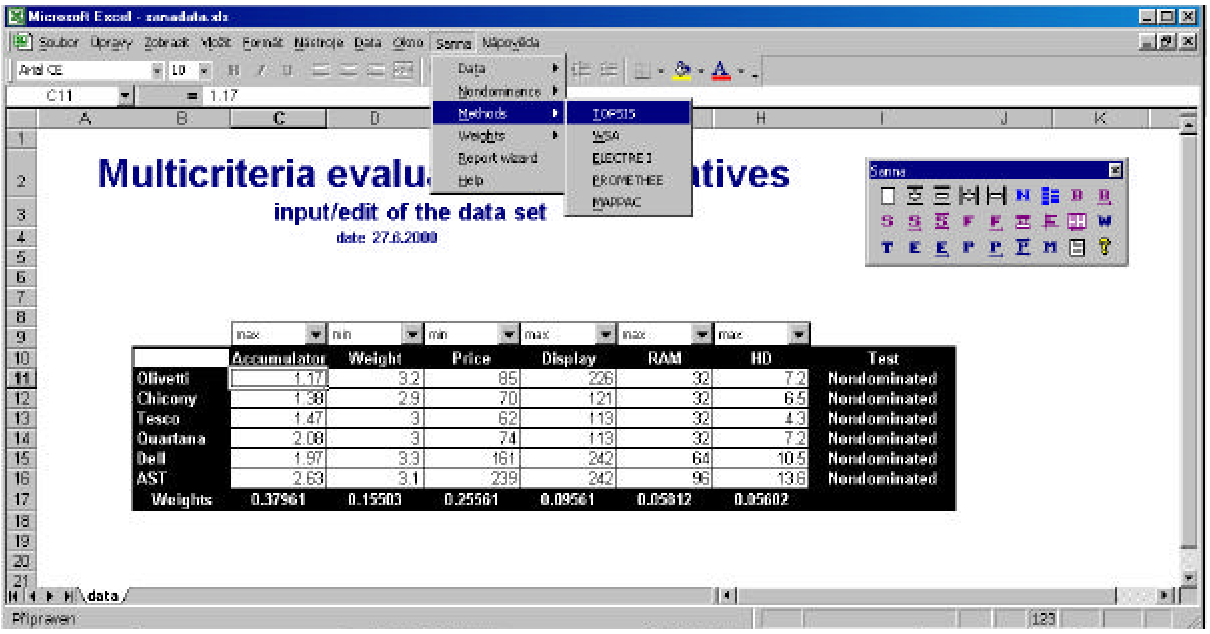


Fig. 2 – SANNA – Data sheet.

4. Application of any of the available multicriteria evaluation methods. WSA, TOPSIS and MAPPAC methods need the knowledge of weights of criteria only. In addition ELECTRE I requires specification of concordance and discordance thresholds and PROMETHEE method need input of types of preference functions and their parameters. The basic results of calculations are placed in new sheets named according to the appropriate method.
5. By the command *Report wizard* the report definition window is displayed and the user can define the form of the created report. The standard form of the report definition window is presented on Fig.3 (the example is for PROMETHEE method). Before the Report wizard is activated any sheet with results of the evaluation must be active. The created report is placed into the new separate sheet.

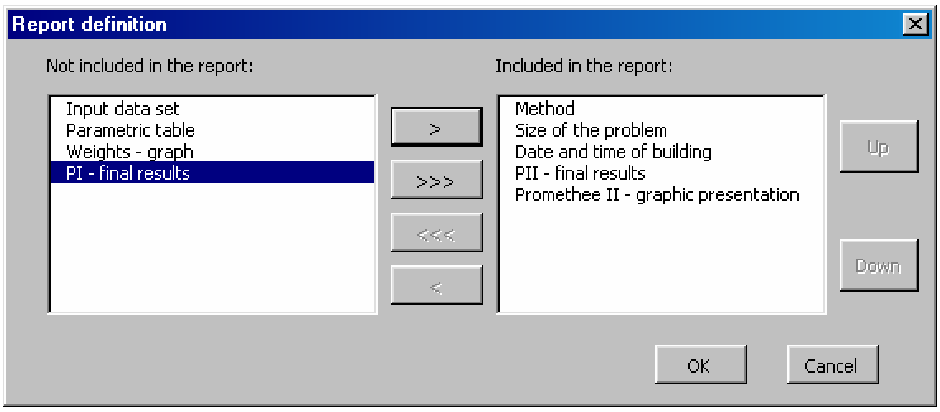


Fig. 3 – SANNA – Report wizard.

The advantage of SANNA consists in the possibility to solve one problem (defined in the “data” sheet) by several methods or by one method by using several weights vectors. All the results for one problem are placed into separate sheets of the active worksheet and in this way they can be very simply compared. SANNA is building in user-friendly MS Excel environment. That is why it can be used without problems on all computers (with Excel) and no special installation is required.

The time of evaluation of a multicriteria decision problem of the limiting size (100 alternatives and 50 criteria) by all the methods except MAPPAC is in seconds (several tens of seconds for PROMETHEE class methods). For the MAPPAC method the same time in in several minutes. Naturally the time depends on the computer performance – the presented values hold for PIII/450MHz.

SANNA is created as the modular system. That is why it should not be difficult to extend its possibilities e.g. by adding new methods for multicriteria evaluation or improving outputs (graphical presentation of results). It will be our effort in the future.

References

1. Brans,J.P, Mareschal,B.: The PROMCALC and GAIA decision support system for multicriteria decision aid. University of Brussels, 1991.
2. Fiala,P., Fábry,J.: SYMCA – System for Multicriteria Analysis. CEJORE 2(1993), no.2, pp.177-183.
3. Jablonský,J.: Multicriteria modelling by means of the LINGO system. In: Methods and applications of operational research II, Katowice 1998, pp. 63-71.

4. Jablonský,J.:Mathematical Programming Modelling and Optimization Systems. CEJORE 6(1998), no. 3-4, pp.279-288.

Doc. Ing. Josef Jablonský, CSc.
University of Economics Prague,
Department of Econometrics,
W. Churchill Sq. 4, 130 67 Praha 3, Czech Republic,
e-mail: jablon@vse.cz, www: http://nb.vse.cz/~jablon/

Forming of Transport-Optimal Region with Uncertain Weights of Regional Centres

Jaroslav Janácek

Department of Transportation Networks, Faculty of Management and Informatics,

University of Žilina, Slovak Republic

1. Introduction

When an administration system of a region is changed, a repartition of the region to subregions is performed. The repartition is usually accompanied by location of subregion centres. These both activities, the location and subregion forming, have direct impact on public service accessibility. Considering the fact that public services are located in a centre of subregion, the accessibility can be easily evaluated as an average distance of a subregion inhabitant from the subregion centre. This parameter is possible to compute for the whole region, when a centre location and a dwelling place assignment to the individual subregions is known. Furthermore, it is possible to formulate an associated model of mathematical programming and to solve the problem of optimal centre location and optimal assignment exactly. However the above quantified accessibility does not cover the whole quality of the region partitioning, because it has also some economical aspects. These aspects do not enable to create arbitrary number of subregions nor to locate a subregion centre at arbitrary place. This is the reason, why number p of subregions and list I of possible centre locations are given in the above mentioned problem. But not even this inputs let us express the fact that the possible locations differ in infrastructure and other equipment necessary for the location of a subregion centre. To evaluate the cost of the missing equipment is very hard task and furthermore it would be very intricated to express the accessibility in the same financial units. That is why the lack of equipment in the individual possible locations is estimated by experts using some weights. These estimated weights however must be taken as uncertain values. In the following sections we show an approach based on the theory of fuzzy sets, which enables us to handle the uncertainty in this location and region partitioning problem.

2. Classical approach to region partitioning

The classical approach [1] starts from the presumption that there are known fixed charges f_i, which express the lack of equipment at a possible centre location $i \in I$. The public service accessibility is evaluated with costs $c_{ij} = d_{ij}b_j$ for dwelling place $j \in J$ and possible centre location $i \in I$ where, d_{ij} denotes the distance between places i and j and b_j is number of inhabitants of dwelling place j. Having introduced *0-1* variables $y_i \in \{0,1\}$ for each $i \in I$, which describe the decision if a centre is or is not located at place i and variables $x_{ij} \in \{0,1\}$ for each pair i, j, $i \in I$, $j \in J$, which assign dwelling place j to centre location i, we can set the following model of the total accessibility maximization with regard to the weights.

$$\textit{minimize} \quad \sum_{i\in I} f_i\, y_i + \sum_{i\in I} \sum_{j\in J} c_{ij}\, x_{ij} \tag{1}$$

$$\textit{subject to} \quad \sum_{i\in I} x_{ij} = 1 \qquad \textit{for } j \in J \tag{2}$$

$$x_{ij} \le y_i \qquad \textit{for } i \in I \textit{ and } j \in J \tag{3}$$

$$\sum_{i\in I} y_i \le p \tag{4}$$

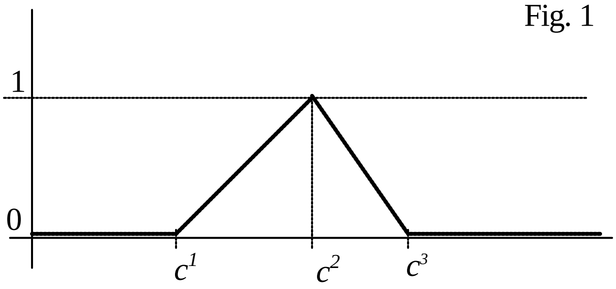
In this model, constraints *(2)* ensure that each dwelling place must be assigned to exactly one subregion and constraints *(3)* force placement of a centre at location i whenever a dwelling place is assigned to centre location i. Constraint *(4)* ensures that the total number of the placed centres does not exceed the given number p.

The problem *(1)-(3)* is known as the uncapacitated location problem and it can be effectively solved making use of implementation of the branch and bound method with Erlenkotter's lower bounding [2]. Computation behaviour of the technique was broadly examined in [5] and it was shown that this approach is able to manage large size problems of practice. The whole problem *(1) - (4)* can be easily transformed to *(1) - (3)* using Lagrangean relaxation for the constraint *(4)* and using algorithm from [4].

When uncertainty of the coefficients f_i is taken into account, sensitivity analysis is usually performed. Due the large size of problem it is impossible to investigate dependence of the optimal result on a change of each coefficient. That is why the coefficients are simultaneously changed. The fixed charges are modified according to the terms $f_i = \alpha\, f_i$ for $i \in I$. The parameter α is changed over an interval with some finite step. A sequence of optimal results is obtained this way and influence of uncertainty on region partitioning and subregion centres location can be studied.

3. Fuzzy approach to the region partitioning

Fig. 1

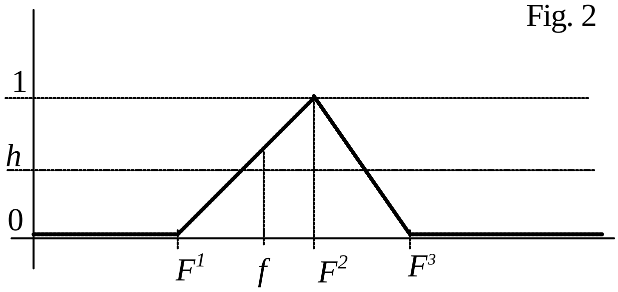


In the theory of fuzzy sets [7] an uncertain value is described by so called triangular fuzzy number. When triangular form of fuzzy numbers is used, then uncertain value $\boldsymbol{c}$ is described by triple < c^1, c^2, c^3> and the membership function in Fig. 1 describes power of assertion that a value belongs to $\boldsymbol{c}$.

Fuzzy numbers can be added, subtracted, multiplied and divided. When coefficients $\boldsymbol{c}_j$ of an objective function $F = \boldsymbol{c}_1x_1 + \boldsymbol{c}_2x_2 +...+\boldsymbol{c}_nx_n$ are fuzzy numbers, then value of the objective function is also a triangular fuzzy number for a given set of variable values. This fuzzy number can be written as

$$\boldsymbol{F}(\mathbf{x}) = <F^1(\mathbf{x}),\ F^2(\mathbf{x}),\ F^3(\mathbf{x})> = \left\langle \sum_{j=1}^{n} c_j^1 x_j, \sum_{j=1}^{n} c_j^2 x_j, \sum_{j=1}^{n} c_j^3 x_j \right\rangle .$$

Fig. 2

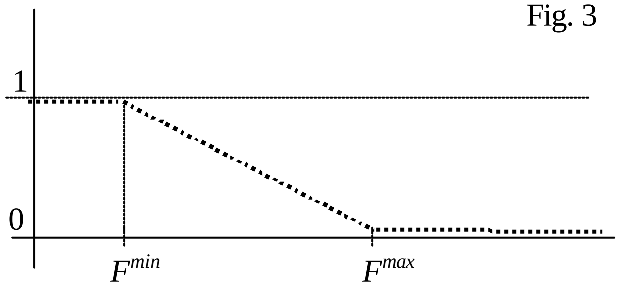


To evaluate power of value membership, a term called„level of satisfaction" is introduced. It is said that value f belongs to fuzzy number $\boldsymbol{F}$ at level of satisfaction h if membership function of the fuzzy number $\boldsymbol{F}$ has value greater or equal to h for argument f (see Fig. 2).

So the minimal value of $\boldsymbol{F}(\mathbf{x})$ at level of satisfaction h is $F^1(\mathbf{x}) + (F^2(\mathbf{x})-F^1(\mathbf{x}))h$. Standard way of mathematical problem solution with fuzzy coefficients consists in previously defining desired level of satisfaction h^* and in standard linear programming problem solution, which objective function $f(\mathbf{x})$ is given by the following term [7].

$$f(\mathbf{x}) = \sum_{j=1}^{n} \left[(h^* - 1)c_j^1 + h^* c_j^2 \right] x_j \tag{5}$$

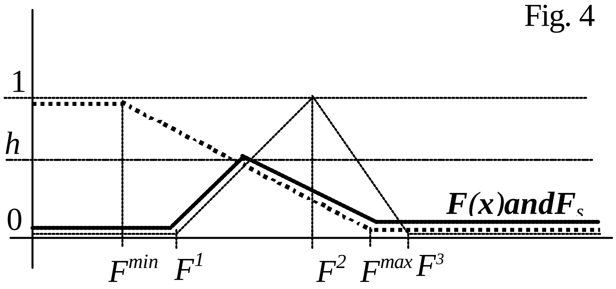
Fig. 3



Credibility of the associated result of this approach depends on expert's experience in determining suitable level of satisfaction.

We suggested other approach [3] providing crisp values of the decision variables. The approach is based on fuzzy set $\boldsymbol{F_s}$, which expresses that „value of F is small" with membership function shown in Fig. 3, where F^{min} and F^{max} denote minimal values of $F^1(\boldsymbol{x})$ and $F^2(\boldsymbol{x})$ respectively over set of feasible solutions.

Fig. 4



In our approach we search for feasible solution $\boldsymbol{x}^*$, for which membership function of fuzzy set „ $\boldsymbol{F(x)andF_s}$" obtains maximal value h (see Fig. 4).

The maximal value h of membership function of fuzzy set $\boldsymbol{F(x)andF_s}$ for a given $\boldsymbol{x}$ has to satisfy following equality $F^1(\boldsymbol{x})+(F^2(\boldsymbol{x})-F^1(\boldsymbol{x}))h=F^{max}-(F^{max}-F^{min})h$ in the cases when $F^1(\boldsymbol{x})\le F^{max}$ holds and in the opposite cases h can be set to zero. For the former case we get

$$h(\boldsymbol{x}) = \frac{F^{max} - F^1(\boldsymbol{x})}{F^2(\boldsymbol{x}) - F^1(\boldsymbol{x}) + F^{max} - F^{min}} \qquad (6)$$

and we seek for $\boldsymbol{x}^*$ maximizing $h(\boldsymbol{x})$, what is a non-linear programming problem. The approximate solution of the problem can be obtained by the following numerical process.

(i) Set $\underline{h}$ at an initial positive value near zero.

(ii) Minimize following objective function $F^1(\boldsymbol{x})+(F^2(\boldsymbol{x})-F^1(\boldsymbol{x}))\underline{h}$ over set of feasible $\boldsymbol{x}$ and denote $\boldsymbol{x}(\underline{h})$ as the associated optimal solution.

(iii) Compute $h(\boldsymbol{x}(\underline{h}))$ according to *(6)*.

(iv) If $|\underline{h}-h(\boldsymbol{x}^*(\underline{h}))| < \varepsilon$ then stop else set $\underline{h}=h(\boldsymbol{x}^*(\underline{h}))$ and go to (ii).

4. An estimation of the city weights

In general, it is almost unsolvable problem to determine the weights so that they express the city inconvenience for subregion centre location and do not suppress value of the accessibility. However we try to establish several useful rules how to determine an interval for the weights not to harm the objective function. The resulting value of the Lagrangean multiplier introduced for constraint *(4)* in the problem *(1) - (4)* with zero charges f_i may be used as an upper bound on these weights. If greater values are used, then the optimal solution *(1) -(4)* could contain less subregions than p. As an general lower bound on the weights once again a resulting value of the Lagrangean multiplier can be used, but this time it should be obtained from the problem where $p+1$ subregions are formed. This way we can get upper and lower bound on the weights. Having determined this basic interval $<f^1, f^3>$ for the weights, two types of the fuzzy weights can be introduced. The first type of fuzzy weight $<f^1, f^w, f^3>$ may serve as a weight of a city, which used to be a centre and is well equipped for new centre location without any additional investment. In this case, value f^w is placed approximately in the first third of interval. The second type of fuzzy weight $<f^1, f^u, f^3>$ is used to appreciate unsatisfactory equipped cities. In this case, value f^u is chosen from the third third of the interval. The individual city weight of the both types may by modified in accordance to some parameter of a city. One of the relevant city parameter could be number n of city inhabitants. Then the above mentioned weights can take form of $<f^1, f^3 -(f^3 -f^w)\text{Ö}(n/n_{max}), f^3>$ or the form of $<f^1, f^3 -(f^3 -f^u)\text{Ö}(n/n_{max}), f^3>$ respectively, where n_{max} denotes number of inhabitants of the most populated city of the region.

5. Numerical experiments

Numerical experiments were focused on computational time consumption of the algorithm (i) - (iv). To test the algorithm on real sized region, road network of whole Slovak Republic was used. The network has semieuclidean form [6] and the associated data were taken from [2]. There was extracted capital Bratislava and its surroundings from the network. After this operation the network had 2889 dwelling places. The set of possible locations I was formed of seventy centres of current districts. In the experiments, there was studied the case, when Slovak Republic should be decomposed into twelve counties including one county formed by Bratislava and its surroundings. It means that during our experiments was value of p set at 11. To construct the weights of the seventy cities, set I was split to two subsets. The first subset was created of seven cities - centres of the current counties and the second set included the rest of set I.

The basic interval $<f^1, f^3>$ was obtained from the experiments made in [4] and it was found that Lagrangean multiplier f^3 for $p=11$ took value 90000 and multiplier f^1 for $p = 12$ took value 73000. The initial values of f^w and f^u were 79000 and 84000 respectively. It was found that within considered set of cities I the maximal number of inhabitants was $n_{max} = 125348$. The individual city weights were determined in accordance to the terms designed in section 4. There were performed six experiments for different initial values f^w, which was changed from 79000 to 84000 with step 1000. The results are summarized in table 1, where d^* denotes an optimal value of the accessibility, t is the total computational time in seconds, $\boldsymbol{F}(\boldsymbol{x}^*)=<F^1(\boldsymbol{x}^*), F^2(\boldsymbol{x}^*), F^3(\boldsymbol{x}^*)>$ denotes fuzzy objective function value of resulting solution $\boldsymbol{x}^*$ and h^* denotes resulting value of satisfaction level. In addition F^{min} and F^{max} denote values of objective function value for the cases, where all weights are set at f^1 or f^3 respectively. There was claimed that in our case $F^{min} = 2224397$and $F^{max} = 2411397$. Letter i in Tab. 1 denotes number of iteration of algorithm (i) - (iv).

Tab. 1

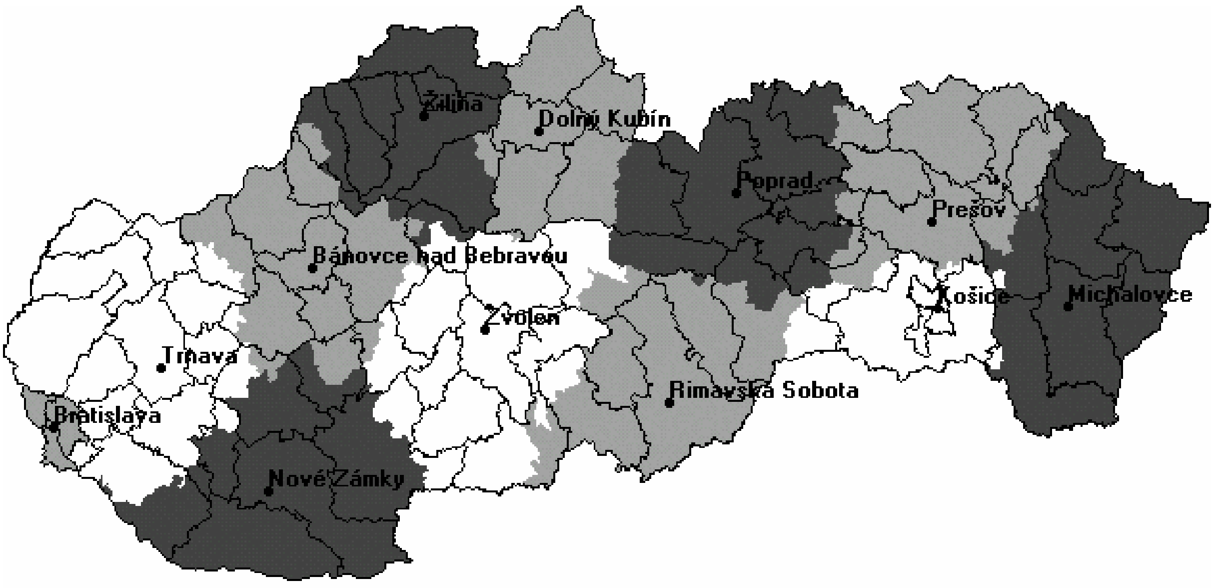
f^w	79000	80000	81000	82000	83000	84000
d^* [km]	28.85	28.85	28.85	28.85	28.85	28.85
t [s]	115	102	88	71	111	67
$F^1(\boldsymbol{x}^*)$	2224397	2224397	2224397	2224397	2224397	2224397
$F^2(\boldsymbol{x}^*)$	2351466	2354909	2358352	2361795	2365237	2368679
$F^3(\boldsymbol{x}^*)$	2411397	2411397	2411397	2411397	2411397	2411397
i	2	2	2	2	2	2
h^*	0.6	0.59	0.58	0.58	0.57	0.56

In all experiments the centres of subregions were set at the same places, which were Trnava, Nové Zámky, Bánovce nad Bebravou, Žilina, Zvolen, Poprad, Rimavská Sobota, Košice - Staré Mesto, Prešov, Michalovce and Dolný Kubín.

6. Conclusions

It was found that suggested algorithm can handle large problems of described type in a sensible time. Furthermore, it follows from the reported experiments that the location are stable with respect to changing weight f^w. The absence of the three cities which represent former county centres in the list of the location may be caused by their unsatisfactory high weights or by their unsuitable location. The possible repartition of Slovak Republic excluding county Bratislava is shown in Fig. 5 for f^w =79000.

Fig. 5



References

1. Cenek, P., Jánošíková, L.: Model dopravnej obsluhy regiónu. Horizonty dopravy, No. 1, 2000, pp. 10-12
2. Erlenkotter, D.: A Dual-Based Procedure for Uncapacitated Facility Location. Operations Research, Vol. 26, No 6, November-December 1978, pp. 992-1009
3. Janácek, J.: Location Problem under Uncertain Costs and Charges. In Proceedings of the 18 th Conference „Mathematical Methods in Economics", 13.-15.Sept.2000, Prague, to appear
4. Janácek, J.: Tvorba dopravnì optimálních regionù. In: Proceedings of the International Conference „Riadenie a informatika v novom tisícroèí", 12.-13.Sept. 2000, Žilina, to appear
5. Janácek J., Kovaèiková J.: Exact Solution Techniques for Large Location Problems. In: Proceedings of the Mathematical Methods in Economics, Ostrava,Sept.9-11, 1997,pp.80-84
6. Janácková, M.: Problémy adekvátnosti a dekompozície semieuklidovskej siete. Horizonty dopravy, No. 4, 1991, pp. 1-10
7. Teodoroviè, D., Vukadinoviè, K.: Traffic Control and Transport Planning: A Fuzzy Sets and Neural Networks Approach. Boston:Kluwer Academic Publishers, (1998), 387 p

Acknowledgement

This work has been supported by research project VEGA 1/7211/20

Prof. RNDr. Jaroslav Janácek CSc.,
Katedra dopravných sietí, Fakulta riadenia a informatiky,
Žilinská univerzita v Žiline, 01026,
e-mail: jardo@frdsa.utc.sk

Stochastic Programming Approach to Multiobjective Optimization Problems with Random Element

Vlasta Kanková[1]

Institute of Information Theory and Automation

Academy of Sciences of the Czech Republic

1. Introduction

It happens rather often that it is suitable or even necessary to evaluate simultaneously an economic process by a few objective functions. If, moreover, there exist a random element and a "parameter" (not determined completely but whose value must only fulfil some conditions), then such situation usually leads (from the mathematical point of view) to a multiobjective optimization problem with a random factor. We introduce this type of problems in a rather general form as the following problem.

Find

$$\min g_i(x, x),\ i = 1,2,\ldots,l_1 \tag{1.1}$$

subject to

$$x \in X \quad \text{such that} \quad g_i(x, x) \le 0,\ i = l_1 + 1,\ldots,l, \tag{1.2}$$

where $g_i(x,z),\ i = 1,\ldots,l$ are real-valued functions defined on $E_n \times E_s; x := x(w)$ is an *s*-dimensional random vector defined on a probability space (W, S, P); $X \subset E_n$ is a nonempty set. ($E_n,\ n \ge 1$ denotes *n*-dimensional Euclidean space.)

If the solution x can depend on the random element x realization, then (1.1), (1.2) is the problem of deterministic multiobjective optimization. If it has to be determined without knowing the random element realization, then at first some deterministic decision rule must be selected, i.e. a deterministic optimization problem must be assigned to the original problem with the random element. To this end the theory of the stochastic programming can be employed (for details see e.g. [6]). The new problem (obtained by this approach) depends on the "underlying" distribution function $\mathrm{F}^x(\cdot)$, corresponding to the random element x and on some additional new parameters. The aim of the contribution is to investigate its stability with respect to the distribution functions space. To achieve new stability assertions we employ the results of [1], [4] and [5]. In the literature, the stability of the multiobjective stochastic programming problems were already investigated (see e.g. [2], [3], [8]).

[1] This research was supported by the Grant Agency of the Czech Republic under Grants 402/98/0742 and 402/99/1136

2. Stochastic Programming Approach

To analyse the original problem (1.1), (1.2) we follow the philosophy employed in [6]. To this end let us assume.

A.1. It is "reasonable" to prefer a few of the objective functions, say

$g_i(x, x),\ i = 1, \ldots, l_2,\ 1 \le l_2 \le l_1,$

A.2. there exist real-valued constants $k_{l_2+1}, \ldots, k_{l_1}$ such that fulfilling of the relations

$$g_i(x, x) \le k_i,\ i = l_2 + 1, \ldots, l_1, \quad \text{is "very" acceptable,} \tag{2.1}$$

A.3. the assumptions A.1 and A.2 can reduce (in the above mentioned sense) the multiobjective problem introduced by the relations (1.1) and (1.2) to

a. the multiobjective optimization problem:

Find

$$\min g_i(x, x), \quad i = 1, 2, \ldots, l_2 \quad \text{over } \{x \in X : \hat{g}_i(x, x) \le 0, \quad i = l_2 + 1, \ldots, l\}, \tag{2.2}$$

b. one -objective optimization problem:

Find

$$\min \sum_{i=1}^{l_2} l_i g_i(x, x) \quad \text{over} \quad \{x \in X : \hat{g}_i(x, x) \le 0, \quad i = l_2 + 1, \ldots, l\} \tag{2.3}$$

with $l_i \in \langle 0, 1 \rangle,\ i = 1, \ldots, l_2,\ \sum_{i=1}^{l_2} l_i = 1$.

$\hat{g}_i(x, z),\ i = l_2 + 1, \ldots, l$ are defined by the relations

$$\begin{aligned} \hat{g}_i(x, z) &= g_i(x, z) - k_i, \quad i = l_2 + 1, \ldots, l_1, x \in E_n, \quad z \in E_s, \\ \hat{g}_i(x, z) &= g_i(x, z), \qquad i = l_l + 1, \ldots, l, \ x \in E_n, \quad z \in E_s, \end{aligned} \tag{2.4}$$

The problem (2.2) is a multiobjective optimization problem with the random element that tries to guarantee the "acceptable" criteria value for $i = l_2 + 1, \ldots, l_1$ as well as the fulfilling of the constraints with random elements while (2.3) is one-objective optimization problem that tries to guarantee the same value of "unicluded" criteria and the constraints.

Evidently, under conditions A.1, A.2 and A.3a usually the following multiobjective deterministic problem can be considered.

P.1. Find

$$\inf \left\{ \mathrm{E}_{F^x} g_i(x, x) \mid x \in \overline{X}_{F^x}(\overline{a}) \right\}, i = 1, \ldots, l_2 \tag{2.5}$$

with

$$\begin{aligned} \overline{X}_{F^x}(\overline{a}) &= \bigcap_{i=l_2+1}^{l} \overline{X}^i_{F^x}(a_i), \\ \overline{X}^i_{F^x}(a_i) &= \left\{ x \in X : P_{F^x}\{\hat{g}_i(x, x) \le 0\} \ge a_i \right\}, i = l_2 + 1, \ldots, l \end{aligned} \tag{2.6}$$

P.2. Find

$$\inf \left\{ \mathrm{E}_{F^x} g_i(x,x) \mid x \in X_{F^x}(a) \right\}, i = 1,\ldots, l_2 \tag{2.7}$$

with

$$X_{F^x}(a) = \left\{ x \in X : P_{F^x} \left\{ \hat{g}_i(x,x) \le 0, i = l_2 + 1,\ldots, l \right\} \ge a \right\}, \tag{2.8}$$

$\overline{a} = \left(a_{l_2+1},\ldots, a_l\right), a_i \in \rangle 0,1\langle, i = l_2 + 1,\ldots, l; a \in \rangle 0,1\langle$; $P_{F^x}(\cdot), \mathrm{E}_{F^x}$ denote the probability measure and mathematical expectation corresponding to $F^x(\cdot)$. If A.3a can be replaced by A.3b, then the following one -objective deterministic optimization problems can be (very often) considered.

I. Find

$$\inf \left\{ \mathrm{E}_{F^x} \sum_{i=1}^{l_2} l_i g_i(x,x) \middle| x \in \overline{X}_{F^x}(\overline{a}) \right\} \tag{2.9}$$

II. Find

$$\inf \left\{ \mathrm{E}_{F^x} \sum_{i=1}^{l_2} l_i g_i(x,x) \middle| x \in X_{F^x}(a) \right\} \tag{2.10}$$

The problems I and II are well -known from the stochastic programming literature (see e.g. [7]). To solve exactly the problems P.1, P.2, I and II the complete information about $P_{F^x}(\cdot)$ must be known. Evidently, this assumption is fulfilled in the real -life situations very seldom. According to these facts, the investigation of the stability with respect to the “measure parameter” is very suitable. To this end we restrict to the case when

B.1. $s \ge l - l_2$ and, moreover, there exist continuous functions $h_i(x), i = l_2 + 1,\ldots, l$ such that

$$g_i(x,z) = h_i(x) - z_i, i = l_2 + 1,\ldots, l\, x \in E_n, z = (z_1,\ldots, z_s) \in E_s. \tag{2.11}$$

We define the quantiles $k_{F_i^x}(a_i)$, the multifunctions $\overline{K}_i(z_i), i = l_2 + 1,\ldots, l, \overline{K}(z), \overline{K}(k, k_F(\overline{a}))$, $k = \left(k_{l_2+1},\ldots, k_{l_1}\right), k_{F^x}(\overline{a}) = \left(k_{F^x_{l_2+1}}\left(a_{l_2+1}\right),\ldots, k_{F_l^x}(a_l) \right)$ by the relations

$$k_{F_i^x}(a_i) = \sup \left\{ z_i : P_{F^x}\left\{ z_i \le x_i \right\} \ge a_i \right\},$$

$$\overline{K}_i(z_i) = \left\{ x \in X : h_i(x) \le z_i \right\}, \overline{K}(z) = \bigcap_{i=l_2+1}^{l} \overline{K}_i(z_i), \qquad . \tag{2.12}$$

$$\overline{K}\left(k, k_{F^x}(\overline{a})\right) = \left(\bigcap_{i=l_2+1}^{l_1} \overline{K}_i\left(k_i + k_{F_i^x}(a_i)\right) \right) \bigcap \left(\bigcap_{i=l_1+1}^{l} \overline{K}_i\left(k_{F_i^x}(a_i)\right) \right),$$

$F_i^x(\cdot), i = 1,\ldots, s$ denote one –dimensional marginals corresponding to $F^x(\cdot)$. We shall introduce the following assumption.

B.2 $\overline{K}(z), \overline{K}(z')$ are nonempty sets, $d > 0$ and, moreover, there exists a constant $C > 0$ such that

$$\Delta\left[\bar{K}(z), \bar{K}(z')\right] \le C\|z - z'\|^{*}_{l-l_2}; \|z - z'\| = \left(\sum_{i=l_2+1}^{l}(z_i - z_i')^2\right)^{\frac{1}{2}}$$

$$z = (z_1,\dots,z_s),\ z' = (z_1',\dots,z_s') \text{such that } z_i \in \bar{Z}_i,\ i = l_2+1,\dots,l.$$

$$\bar{Z}_i = \left(k_{F_i^x}(a_i) + k_i - 2d,\ k_{F_i^x}(a_i) + k_i + 2d\right),\ i = l_2+1,\dots,l_1, \tag{2.14}$$

$$\left(k_{F_i^x}(a_i) - 2d,\ k_{F_i^x}(a_i) + 2d\right),\ i = l_1+1,\dots,l.$$

3. Some Definitions and Auxiliary Assertions

A multiobjective deterministic optimization problem can be defined as the problem. Find

$$\min\ f_i(x), i = 1,\dots,r \text{subject to } x \in K. \tag{3.1}$$

$f_i(x), i = 1,\dots,r$ are real -valued functions defined on $E_n, K \subset E_n$ is a nonempty set.

Definition 1. The vector x^* is an efficient solution of the problem (3.1) if and only if there exists no $x \in K$such that $f_i(x) \le f_i(x^*)$for $i = 1,\dots,r$ and such that for at least one i_0 one has $f_{i_0}(x) < f_{i_0}(x^*)$.

Definition 2. The vector $x*$ is a properly efficient solution of the multiobjective optimization problem (3.1) if and only if it is efficient and if there exists a scalar $\bar{M} > 0$ such that for each i and each $x \in K$ satisfying $f_i(x) < f_i(x^*)$ there exists at least one j such that $f_j(x^*) < f_j(x\ )$ and

$$\frac{f_i(x^*) - f_i(x\ )}{f_j(x\ ) - f_j(x^*)} \le \bar{M}\ . \tag{3.2}$$

Proposition 1. [1] Let K be convex set and let $f_i(.), i = 1,\dots,r$ be convex functions on K. Then x^0 is a properly efficient solution of the problem (3.1) if and only if x^0 is optimal in

$$\min_{x\in K}\sum_{i=1}^{r} l_i f_i(x) \text{for some } l_1,\dots, l_r > 0; \sum_{i=1}^{r} l_i = 1.$$

To recall auxiliary assertions on the stability of one-objective stochastic programming problems we define functions $F_d(z), F^d(z), F_{i,d}(z_i), F^{i,d}(z_i), i = 1,\dots,s, z = (z_1,\dots,z_s) \in E_s$, $d > 0$ and the sets G_F, X_F by the relations

$$\begin{aligned} F_d(z) &= F^x(z) - \frac{1}{2^{s+2}}\inf_{z'\in B(z,2\sqrt{sd})} P_{F^x}\left\{Z\left[z', \frac{d}{2}\right]\right\}, \\ F^d(z) &= F^x(z) + \frac{1}{2^{s+2}}\inf_{z'\in B(z,2\sqrt{sd})} P_{F^x}\left\{Z\left[z', \frac{d}{2}\right]\right\}, \end{aligned} \tag{3.3}$$

$$F_{i,d}(z_i) = F_i^x(z_i - d), F^{i,d}(z_i) = F_i^x(z_i + d), \tag{3.4}$$

$$G_F = \{y \in E_{l_2} : y_j = \mathrm{E}_{F^x} g_j(x,x), \, j = 1,\dots, l_2 \text{ for some } x \in \overline{X}_{F^x}(\overline{a}); y = (y_1,\dots, y_{l_2})\}$$

$$\overline{X}_F = \{x \in X_F : x \text{ is a properly efficient point of the problem } P.1\}.$$

$$Z[z,d] = \{z' \in E_s : z_i' \in \rangle z_i, z_i + d), i = 1,2,\dots, s, z = (z_1,\dots,z_s), z' = (z_1',\dots, z_s')\}, \tag{3.5}$$

$B(z,d)$ denotes the d - neighbourhood of the point $z \in Z_{F^x}$.

If we replace in (3.5) $F^x(\cdot)$ by some another s - dimensional distribution function $G(\cdot)$, then we obtain the sets $G_G, \overline{X}_G$ instead of the original $G_F, \overline{X}_F$.

Proposition 2. [5] Let $a_i \in (0,1), i = l_2 + 1,\dots, l, d, e :> 0$. Let, moreover, X be a nonempty, compact set. If

1. $g(x)$ is a real -valued, Lipschitz function on X with the Lipschitz constant L_g,
2. $G(z)$ is an arbitrary s-dimensional distribution function such that its one-dimensional marginal distribution functions $G_i(z_i), i = l_2 + 1,\dots, l$ fulfil the relation

   $G_i(z_i) \in \rangle F_{i,d}(z_i), F^{i,d}(z_i)\langle, z_i \in (k_{F_i}(a_i) - d - e, k_{F_i}(a_i) + d + e),$
3. the assumptions B.1, B.2 are fulfilled, then

$$\left| \min_{x \in \overline{X}_F(\overline{a})} g(x) - \min_{x \in \overline{X}_G(\overline{a})} g(x) \right| \le C L_g s d.$$

Proposition 3. [4] Let $d > 0, X \subset E_n$ be nonempty, compact set. If

1. $g_1(x,z)$ is a uniformly continuous on $X \times E_s$ and, simultaneously, for every $x \in X$ a Lipschitz function on $Z_{F^x}(sd)$ with the Lipschitz constant L_1,
2. $G(z)$ is an arbitrary l-dimensional distribution function such that

$$G(z) \in \rangle F_d(z), F^d(z)\langle, z \in E_s,$$

3. for every $x \in X$ there exists finite $E_{F^x} g_1(x,x), E_G g_1(x,x)$, then

$$\left| \inf_X \mathrm{E}_F g_1(x,x) - \inf_X \mathrm{E}_G g_1(x,x) \right| \le 4L\sqrt{sd}.$$

$Z_{F^x}(e), e > 0$ denotes the $e -$neighbourhood of the support Z_{F^x} of $P_{F^x}(\cdot)$.

4. Main Results

To introduce the main results let us assume

B.3 for every $x \in X$ there exist finite $E_{F^x} g_i(x,x), E_G g_i(x,x), i = 1,\dots,l_2$.

Theorem 1. Let $a_i \in (0,1), i = l_2+1,\dots,l, d, e > 0, X$ be a compact set. If

1. $g_i(x,z), i = 1,\dots,l_2$ are Lipschitz functions on $X \times Z_F(sd)$ with the Lipschitz constants L_i,
2. $G(z)$ is an arbitrary s-dimensional distribution function such that

$G(z) \in \rangle F_d(z), F^d(z)\langle, z \in E_s$ and simultaneo usly

$G_i(z_i) \in \rangle F_{i,d}(z_i), F^{i,d}(z_i)\langle, z_i \in (k_{F_i}(a_i) - d - e, k_{F_i}(a_i) + d + e), i = l_2+1,\dots,l$

3. the assumptions B.1, B.2 and B.3 are fulfilled, then

$$\Delta[\overline{G}_F, \overline{G}_G] \le 4[C+1]sd \sum_{i=1}^{l_2} L_i$$

$D[\cdot,\cdot]$ denotes the Hausdorff distance in the space of nonempty, closed subsets of E_n (for the definition of the Hausdorff distance see e.g. [9]).

Theorem 2. Let $a_i \in (0,1), i = 1,2,\dots,l_2, d, e > 0, X$ be a compact set. If

1. $g_i(x,z), i = 1,\dots,l_2$ are
   a. Lipschitz functions on $X(e) \times Z_{F^x}(e)$ with the Lipschitz constants L_i,
   b. for every $z \in Z_{F^x}(e)$, convex functions.
2. there exists $j \in \{1,\dots,l_2\}$ such that $g_j(x,z)$ is a strongly convex function with a parameter r on X,
3. $G(z)$ is an arbitrary s-dimensional distribution function such that
   $G(z) \in \rangle F_d(z), F^d(z)\langle, z \in E_s$ and simultaneo usly
   $G_i(z_i) \in \rangle F_{i,d}(z_i), F^{i,d}(z_i)\langle, z_i \in (k_{F_i}(a_i) - d - e, k_{F_i}(a_i) + d + e),, i = l_2+1,\dots,l$
4. $h_j(x), j = l_2+1,\dots,l$ are convex functions on X,
5. the assumptions B.1, B.2 and B.3 are fulfilled,
   then $[\Delta[\overline{X}_F, \overline{X}_G]]^2 \le [C+1]sd \dfrac{16}{r} \sum_{i=1}^{l_2} L_i$.

For definition strongly convex functions see e.g. [3].

Sketch of the proofs. Employing the proofs technique of [5] we can see that the assertions of Theorem 1 and Theorem 2 follow from the triangular inequality, the assertions of Proposition 1, Proposiotion 2 and Proposition 3.

5. Conclusion

In the paper we introduced stability results concerning the problem P.1. Evidently, the similar results can be obtain for the problems P.2. However, to introduce the corresponding results is over the possibility of this paper.

References

[1] A.M. Geoffrion: Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. *22*, (1968), 3, 618 –630

[2] G.-M. Cho: Stability of the multiple objective linear stochastic programming problems. Bull. Korean Math. Soc. *32* (1995), 2, 287 –296.

[3] V. Kanková: A note on stability of multiobjective stochastic programming (in Czech).In: Kvantitatívne metódy v ekonomike. (Viacekriteriálna optimizácia VII). Ekonomická univerzita v Bratislave, Bratislava 1994, 97 -103.

[4] V. Kanková: On Distribution Sensitivity in Stochastic Programming. Research Report ÚTIA 1994, No. 1826

[5] V. Kanková: On the stability in stochastic programming: the case of individual probability constraints. Kybernetika *33* (1997),525 -546.

[6] V. Kanková: A note on analysis of economic activities with random elements. In: Internat. Conference Mathematical Methods in Economy, Cheb 1998.(M.Plevný and V. Friedrich, eds.) University of West Bohemia and Czech Society for Operations Research, Cheb 1999, 53 -58.

[7] A. Prékopa: Stochastic Programming. Akadémiai Kiadó, Budapest and Kluwer Publisher, Dordrecht 1995

[8] S.Vogel:On stability in multiobjective programming- a stochastic approach. Mathematical programming *50* (1992), 197 -236.

[9] G. Salinetti, and R.J.-B. Wets: On the convergence of convex sets in finite dimensions. SIAM Review 21 (1979), 18-33.

RNDr. Vlasta Kanková, CSc.
Department of Econometrics
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodárenskou veží 4, 182 08 Praha 8

RECURRENT UNEMPLOYMENT AND RESIDUAL HETEROGENEITY: ARE UNEMPLOYMENT SPELLS INDEPENDENT EVENTS?

Rui Menezes

Instituto Superior de Ciências do Trabalho e da Empresa, Lisboa

Key words: Independent events, logit models, multiple spells, residual heterogeneity, scale parameter, time-varying covariates, unemployment spells.

I. INTRODUCTION

Analyses of worker movement between several employment statuses have received increasing attention on the part of labour economists over the last two decades. In many studies, the theoretical basis of analysis has a semi-Markov framework, and unemployment duration relies on data for single spells of unemployment [see, *inter alia*, Arulampalam and Stewart (1995), Bosworth (1992), and Narendranathan and Stewart (1993)]. Typically, these studies have addressed the presence and sign of state duration dependence within unemployment [e.g. van den Berg and van Ours (1994)], but claims exist that other forms of state dependence may also be of relevance [Heckman and Borjas (1980)].

One problem associated with Markov renewal models concerns the assumption of independence between multiple spells in and between states for the same person [Blanchard and Diamond (1990), Theeuwes *et al.* (1990)]. This means that the length of the current spell in a state does not depend on the occurrence and length of a previous spell in any state, and every new spell is treated as a new observation. Such an assumption, however, may not be suitable in a number of situations. For example, some workers, namely unemployed workers, may accept a stopgap job while still searching for a better alternative. If this job is temporary a second unemployment spell may follow, which is not independent from the first spell; nor it is a possible second employment spell relatively to the first one. It is therefore important to account for the problem of non-independent multiple spells in one state of the labour market when analysing recurrent event models.

This paper is concerned with the analysis of binary recurrent events for unemployment, within a framework that allows for time-varying explanatory variables and residual heterogeneity. Binary recurrent events are outcomes (y_{it}) over time (t) for individuals or cases (i) that take values of 1 if the event happens and 0 if it does not. A binary sequence of zeros and ones can thus be obtained for each individual or case. The sequences may record an outcome of interest, such as being unemployed, at discrete, regular intervals.[1] For instance, in collecting work-history data retrospectively, one may be able to obtain monthly or weekly data on general employment status, but it is usually very difficult to know precisely the day of every employment status change. Moreover, this is a simple device to allow for incorporation of time-varying covariates in the model as explanatory variables [Chamberlain (1979)]. Time-varying covariates are variables that change value over time, eventually at regular intervals. Continuous-time duration models typically assume that explanatory variables do not change between events. However, the breakdown of this assumption in conventional models can lead to serious bias in parameter estimation [see, *e.g.*, Flinn and Heckman (1982)].

The remainder of the paper is organised as follows. In Section II we present the basic issues concerning the modelling of recurrent events. This involves the use of standard logit as well as some extensions. The next Section presents a short description of the data used to empirically estimate the models. The data used here rely on a large-scale redundancy occurred at Michelin headquarters in the UK in the mid eighties. Section IV presents the results obtained from the empirical implementation of the models. Three model specifications were tested in order to learn the effect of standard variables, as well as time-varying and residual heterogeneity effects. Finally, some concluding remarks are presented in Section V.

II. MODELLING RECURRENT EVENTS

In many situations, standard regression models with fairly attractive statistical and computational properties provide effective and robust representations of the phenomena under analysis. This allows the systematic variation attributable to explanatory variables to be distinguished from the random variation due to other effects. Typically, the assumption of a normal random error term is sufficient for many purposes. However, the statistical modelling of recurrent events is complicated by the dependence of current behaviour on past behaviour (feedback effect), time-varying exogenous variables (non-stationary) and the variation between individuals due to unmeasured and possibly immeasurable variables (residual heterogeneity). Ignoring any of these phenomena can result in seriously misleading conclusions about the others and the effects of explanatory variables included in the model variables [Davies and Pickles (1985)]. Our modelling of recurrent events will be based on the well-known logit model, where residuals are logistic-normally distributed.

Consider for the moment a sequence of recurrent outcomes $(y_{i1}, y_{i2}, \cdots)$ for individual or case i. The simplest version of the logit model would then imply that the contribution to the likelihood by the i^{th} individual or case and t^{th} event, is given by

$$g\left(y_{it} \mid b; x_{it}\right) = \frac{\left[\exp\left(b' x_{it}\right)\right]^{y_{it}}}{1+\exp\left(b' x_{it}\right)} \qquad (1),$$

where x_{it} is a vector of explanatory variables and b is a vector of parameter estimates. For the whole sequence of recurrent outcomes for individual or case i one has

$$g\left((y_{i1}, y_{i2}, \cdots) \mid b; x_{it}\right) = \prod_{t=1}^{T_i} \frac{\left[\exp\left(b' x_{it}\right)\right]^{y_{it}}}{1+\exp\left(b' x_{it}\right)} \qquad (2),$$

where T_i is the sequence length.

An alternative formulation of the model just outlined includes an explicit representation of the effect of omitted variables, *i.e.*, a case-specific random error term. This random effects model is known as a residual or internal heterogeneity [Allison (1987)] formulation of the logit model, where the error term e_i is normally distributed. The density of the observed sequence for case i in the logistic-normal mixture model can be written as

$$g\left((y_{i1}, y_{i2}, \cdots) \mid b; x_{it}; e_i\right) = \prod_{t=1}^{T_i} \frac{\left[\exp\left(b' x_{it} + e_i\right)\right]^{y_{it}}}{1+\exp\left(b' x_{it} + e_i\right)} \qquad (3),$$

[1] Alternatively, the outcomes may be binary responses at specific, possibly irregular, occasions in time, such as, for example, voting or abstaining in a General Election.

where e_i is a vector of individual-specific error terms.

The main problem associated with the estimation of this individual-specific random error term model is that nuisance parameters cannot be estimated simultaneously with the structural parameters. Since the number of nuisance parameters increases with the number of cases, the limit conditions required for asymptotic theory would be violated, and the structural parameters b would be inconsistent [Neyman and Scott (1948)]. This problem may be overcome by eliminating the error terms before estimating the model, using either conditional or marginal likelihood methods.

Conditional likelihood methods can only be applied to a restricted class of models, namely for small datasets and when all covariates are time varying. Marginal (or integrated) methods, on the other hand, provide a useful framework for estimating individual-specific random error term models, in which the error terms are assumed to be independent of included explanatory variables, and are integrated out of the likelihood. The density function unconditional on the e_i is given by

$$g\left((y_{i1}, y_{i2}, \cdots) \,|\, q; x_{it}\right) = \int g\left((y_{i1}, y_{i2}, \cdots) \,|\, b; x_{it}; e_i\right) f(e)\, de \qquad (4),$$

where $f(e)$ is the pdf (or mixing distribution) of the error term, and the parameter vector q consists of *ß* and the parameters of $f(e)$.[2] For a sample of N individuals, the sequence or marginal likelihood is

$$L_i = \sum_{j=1}^{N} \left[\prod_{t=1}^{T_i} \frac{\left[\exp\left(b' x_{it} + x z_j\right)\right]^{y_{it}}}{1 + \exp\left(b' x_{it} + x z_j\right)} \right] p_j \qquad (5),$$

where z_j and p_j are, respectively, the fixed quadrature locations and probabilities and the scale parameter x is the unknown standard deviation of the mixing distribution. It should be noted that the error terms are assumed to be iid and are therefore uncorrelated with the explanatory variables included in the model [Chesher and Lancaster (1983)].[3]

---

[2] In our case, *f*(.) is assumed to be Normal.

[3] For further details on these models see Barry *et al.* (1990).

III. THE MICHELIN DATA

The data used in this paper concern a longitudinal survey of 230 Michelin workers made redundant in 1985, from the Michelin plant in Stoke-on-Trent, UK. These workers were among the 2137 made compulsorily redundant at that time.

The sampled workers were followed over a period of more than two years after May 1985, and some 227 of the original sample of 230 people were still part of the final group interviewed. The attrition rate for the Michelin longitudinal survey was, therefore, lower than that found in other redundancy longitudinal studies. Some ex-Michelin employees had gone away to other regions of the UK, but none went further than the far southwest of England. Nevertheless, the vast majority remained in area, and the Stoke TTWA encompasses the bulk of their post-redundancy labour market experiences. The ex-Michelin workers in the sample proved to be a group of mature, well-qualified, and in general eminently “employable” workers. They are mainly male prime-aged workers, and relatively well paid for the region. The vast majority of workers in the sample were in semi-skilled manual occupations at Michelin. Their unemployment experiences prior to working at Michelin were very limited.[4]

For each of the 227 ex-Michelin workers in the sample, there is data on their post-redundancy labour market experiences on a weekly basis. There is also available data on worker personal characteristics and local labour market (Stoke TTWA) unemployment rates.[5] Local unemployment rates will be used in the logit-normal model as a time-varying explanatory variable, in order to capture the influence of local economic conditions on the probability of individual *i* being found in, *e.g.*, employment, rather than unemployment, at a given date. The labour market statuses of interest to our study are unemployment and employment, irrespective of the actual employer.[6]

The information available was collected in three main groups: employment, unemployment and not in the labour force. Were taken as employed at some given specific date those individuals who stated they had any kind of paid job at that time. This includes all full-time and part-time jobs, and self-

---

[4] For further details on the sample and local labour market characteristics see Menezes (1995).

[5] These are official unemployment rates published by the Employment Gazette.

employment. As unemployed were considered those individuals who stated that they were out of employment and looking for a job in the labour market. All other statuses, including those who stated retirement, sickness, further education, and others, were classified as not in the labour force. Some 15 observations on individuals not in the labour force for more than two thirds of the two-year period in analysis were dropped from the sample. This leaves a final sample of 212 individuals.

For our purposes, we chose to consider five points in time to collect cross-section measurements on the employment/unemployment status for our group of ex-Michelin workers: 3, 6, 12, 18 and 24 months after May 1985, the date where the main redundancies occurred.[7] A binary sequence of five elements is thus available for each individual *i*. This leads to a total of 1038 measurements of the binary dependent variable y_{it}.[8] It should be noted that the choice of these dates is rather arbitrary (and illustrative in our context); it could be possible to consider any other set of points in time within the period to perform this analysis.

Table 1 describes the set of explanatory variables used in this analysis. These variables aim to capture the *ceteris paribus* influence of worker personal characteristics, previous labour market history, and local economic conditions on the log of the odds of being in one state rather than another. The personal characteristics included in the model are apprenticeship, gender, age, final weekly wage and final job classification at Michelin. Local economic conditions have changed over the period, so that it is included as a time-varying covariate in the logit-normal mixture formulation. Continuous variables such as age, wage, and unemployment rate were factorised and then converted into dummies. All categorical variables were also turned into dummies, so that all variables in the model are binary in form.

---

[6] However, longitudinal surveys of this kind typically allow for much more detailed information than just these labour market statuses. From the original questionnaires of the Michelin survey, it is possible to distinguish from at least 11 different labour market statuses following displacement.

[7] A total of 40 individuals left Michelin before (33) or after (7) May 1985, but none after September 1985.

[8] This binary variable takes the value one if the individual is classified as employed at time *t*, and zero if he is classified as unemployed.

TABLE 1: EXPLANATORY VARIABLES

Individual and Labour Market Characteristics

NAMES	DESCRIPTION
Apprenticeship	Served apprenticeship dummy. 1=yes
Gender	Sex dummy. 1=female
Age 35-45	Age left Michelin. 1=between 35 and 45
Age 45-55	Age left Michelin. 1=between 45 and 55
Age 55Plus	Age left Michelin. 1=over 55
Mid Wage	Final weekly wage at Michelin. 1=between £120 and £160
High Wage	Final weekly wage at Michelin. 1=over £160
Occupation 2	Final job classification at Michelin. 1=skilled manual worker
Occupation 3	Final job classification at Michelin. 1=semi-skilled manual worker
Occupation 4	Final job classification at Michelin. 1=unskilled manual worker
UHistory	Unemployment experience prior to Michelin. 1=greater than 12 weeks
URate 11-12	Unemployment rate in Stoke TTWA. 1=between 11% and 12%
URate 12-13	Unemployment rate in Stoke TTWA. 1=between 12% and 13%

With the above configuration of the independent variables, the base-case against which we stress the parameter estimates of our models can be stated as being a male, without a served apprenticeship, under 35 years old, low-waged, non-manual and without previous unemployment longer than 12 weeks.[9] The base-case individual searches for new employment under local economic conditions where the unemployment rate lies between 10 and 11 percent.

IV. EMPIRICAL RESULTS

We now turn to consider the results obtained from estimating the logit and logit-normal models described in Section II. The models were estimated by maximum likelihood using Sabre's routines.[10] Sabre uses Gaussian quadrature for the computation of each likelihood integral, assuming a Normal

[9] Previous unemployment here means unemployment prior to the Michelin job.

[10] The reader is referred to Sabre manual for extensive details about the models and related computational issues.

parametric form for the density. Three different specifications were fitted: (1) a standard logit where all the 1038 measurements are assumed to be independent observations, (2) a logistic-normal mixture model that accounts for residual heterogeneity due to omitted variables, and (3) a logistic-normal mixture model that accounts for residual heterogeneity and time-varying local economic conditions. The dependent variable measures whether the individual was employed (1) or unemployed (0) at given dates, as described earlier.

The results for the three models are presented in Table 2. All the variables are taken as time-invariant except local unemployment rate (URate) that enters in the 3rd model. The variable URate was measured at the same time points for which we have measurements on the dependent variable. Notice that the coefficients of the model measure the effect of gaining the underlying attribute on the logarithm of the odds of being employed rather than unemployed, at given specific dates, harassed against the general profile of the base-case individual.

TABLE 2: LOGIT AND LOGIT-NORMAL REGRESSIONS

Dependent variable is whether employed at given date

VARIABLE	**STANDARD LOGIT**	**LOGIT-NORMAL**	**LOGIT-NORMAL WITH TIME-VARYING COVARIATES**
	(1)	**(2)**	**(3)**
Constant	3.013**	3.088**	5.202**
	(8.61)	(6.33)	(6.86)
Apprenticeship	-0.002	0.351	0.374
	(0.01)	(1.23)	(1.04)
Gender	-0.621*	0.089	-0.027
	(2.06)	(0.21)	(0.05)
Age 35-45	-0.654**	-0.397	-0.474
	(2.89)	(1.26)	(1.23)
Age 45-55	-1.020**	-1.023**	-1.289**
	(4.19)	(3.05)	(3.00)

Age 55Plus	-1.994**	-1.877**	-2.387**
	(6.57)	(4.85)	(4.43)
Mid Wage	0.200	-0.664*	0.775
	(0.92)	(2.02)	(1.85)
High Wage	-0.995**	-1.784**	-2.166**
	(2.91)	(3.75)	(3.61)
Occupation 2	-0.625	-0.996*	-1.092
	(1.78)	(2.08)	(1.85)
Occupation 3	-1.058**	-1.625**	-1.936**
	(3.61)	(4.02)	(3.92)
Occupation 4	-1.916**	-3.221**	-3.974**
	(5.10)	(4.98)	(4.67)
UHistory	-0.885**	-0.961*	-1.198*
	(2.67)	(2.34)	(2.21)
URate 11-12			-0.144
			(0.34)
URate 12-13			-2.200**
			(6.21)
x		0.013	0.605
		(0.01)	(1.74)
Log-Likelihood	-483.737	-430.928	-389.304
Number of cases		212	
Number of measurements		1038	

t-statistics in parentheses

* *Significantly different from zero at the 5% level*

** *Significantly different from zero at the 1% level*

Most parameter estimates are significant at the 5% level of better,[11] the results being in line with previous findings and our prior expectations. In fact, being female implies lower chances of finding new employment following redundancy, but this is only apparent in the standard logit. Older workers face lower prospects of finding new employment after redundancy than younger workers (the older the worker the lower the chance). Being a high-waged worker at Michelin implies a lower probability of being employed following redundancy than the counterpart. Manual workers face lower prospects of finding new employment following redundancy than non-manual workers (the lower the skills the worse the prospects). Unemployment experience prior to Michelin exerts a negative effect on the probability of finding new employment after redundancy. Finally, it is found that adverse local economic conditions (evaluated by the unemployment rate) exert a negative influence on the probability of being found employed following redundancy (column 3).

The results for the wage variables are consistent with job-search theory. Job-search theory postulates that relatively well-paid workers (in the sense that they have invested more in human capital), once displaced, tend to spend more time searching for a better job opportunity. This is so because these workers establish initial reservation wages at high levels. As time passes, the reservation wage reduces, because the expected future returns from searching, say, another week, decrease under a realistic finite horizon context.

The time-varying variable URate appears significantly, which means that the model is sensitive to the specification of non-stationarity. But, what can we say about residual heterogeneity? Recall that the residual heterogeneity is captured by the scale parameter (x), which measures the unknown standard deviation of the mixing distribution. Residual heterogeneity arises because of differences between individuals due to omitted variables. In our case, residual heterogeneity does not seem to be a big problem, since x is only significant at the 10% level in the 3rd model. This is consistent with our data, since the ex-Michelin workers are a somewhat homogeneous workforce in terms of work attributes, and, since they were all made redundant at the same time, they faced similar conditions in the labour market following redundancy. However, one should be aware of the potential bias caused by residual heterogeneity. A final note about the goodness of the fitted models: the log-L statistic shows that allowing for residual heterogeneity and non-stationarity actually improves the power of the fit relative to the standard logistic model.

Finally, Table 3 presents the estimated probabilities of being employed and unemployed, from our models, for the base-case individual in the sample. Comparative changes in the probability that an individual is found employed given the acquisition of certain characteristics relative to the base-case worker, are shown in Table 4. These changes measure the deviation, in percentage terms, relative to the base-case individual.

TABLE 3: ESTIMATED PROBABILITIES FOR EMPLOYMENT AND UNEMPLOYMENT

Base-Case Individual

	EMPLOYMENT	UNEMPLOYMENT
Standard Logit	95.32%	4.68%
Logit-Normal	95.64%	4.36%
Logit-Normal with T. V. Covariates	99.45%	0.55%

TABLE 4: COMPARATIVE RESULTS AGAINST THE BASE-CASE

Employment Status

VARIABLE	STANDARD LOGIT	LOGIT-NORMAL	LOGIT-NORMAL WITH T. V. COVARIATES
Female	-3.69%	-	-
Age 35-45	-3.95%	-	-
Age 45-55	-7.30%	-6.89%	-1.41%
Age 55PLUS	-21.82%	-18.60%	-5.10%
Mid Wage	-	-3.78%	-
High Wage	-7.05%	-16.99%	-4.03%
Occupation 2	-	-6.63%	-
Occupation 3	-7.72%	-14.28%	-3.13%
Occupation 4	-20.33%	-48.96%	-22.10%
UHistory	-5.95%	-6.29%	-1.24%

[11] Nearly two thirds of the estimates are significant at the 5% level or better, of which 4/5 are significant at least at the 1% level.

Urate 12-13	-	-	-4.18%

V. CONCLUDING REMARKS

This paper seeks to show that recurrent unemployment can be modelled by case-specific random error term techniques, within a standard logit framework. The trick is to consider a random disturbance with normal distribution added to the linear part of the model. This is conceptually straightforward, but raises several problems of estimation. Sabre's computer program offers a convenient way for dealing with this type of issues, based on marginal likelihood methods. This model is a transformation of the standard logit model, and is known as the logistic-normal mixture model. This is a useful device for dealing with non-stationarity (time-varying covariates), and residual heterogeneity (omitted variables). Residual heterogeneity is measured by the unknown standard deviation of the mixing distribution x. In this way, one needs not assume that multiple spells of unemployment (employment) are independent events and treated as separate observations.

The database used in this paper concerns a large-scale redundancy occurred at the Michelin tyre plant in Stoke-on-Trent, UK, in May 1985. The Michelin workers are locally regarded as the pick of the local workforce. The redundant Michelin workers sampled were a long-serving workforce, often possessing firm-specific training and skills, which may have made their worth to other potential employers much less than their worth to Michelin. For other characteristics of the sample, it was found that the ex-Michelin workers were mainly male, prime-aged, and well paid for the local area. Most of these workers were semi-skilled manual workers, and had limited unemployment experience prior to working at Michelin.

In general, our results are compatible with the following:

- The probability of being found employed after redundancy is affected by time-invariant worker personal characteristics;
- Time-varying covariates aiming to capture the effect of local economic conditions appear significantly in the model, and improve the goodness of fit of the regression;

- Residual heterogeneity due to omitted variables does not seem to be a severe problem in our context, probably because the Michelin workers are fairly homogeneous in terms of attributes and work experiences; however, accounting for this improves the general goodness of fit of the model, and refines the estimates of other included explanatory variables.

Finally, a word about the computer software. The software used in this paper is suitable for analyses of recurrent unemployment, recurrent employment, etc (or for analyses where multiple spells are treated as independent), but not for analyses of labour market histories where spells in state j may be related to subsequent spells in state k. A natural updating of this software might allow for multiple choices in recurrent events, in a manner "similar" to the conventional multinomial approach. It would prove highly interesting to test the results obtained by applying a methodology similar to ours with those obtainable by using such software, was it to be available to the researchers. Another possible extension could permit for functional forms other than the typical logistic function, so as to enable alternative experiments in the field of data modelling.

REFERENCES

1. Allison, P.D. (1987): "Introducing a Disturbance into Logit and Probit Regression Models." *Sociological Methods and Research*, 15, 355-374.
2. Arulampalam, W., and M.B. Stewart (1995): "The Determinants of Individual Unemployment Durations in an Era of High Unemployment." *Economic Journal*, 105 (429), March, 321-332.
3. Barry, J., B. Francis, R. Davies (1990): *SABRE: Software for the Analysis of Binary Recurrent Events. A Guide for Users*. Centre for Applied Statistics, Lancaster University.
4. Blanchard, O.J. and P. Diamond (1990): "Unemployment and Wages: What Have we Learned from the European Experience?" Employment Institute, April.
5. Bosworth, D.L. (1992): "Duration of Unemployment: An Analysis of the Labour Force Survey." In C.H.A. Verhaar and L.G. Jansma *et al.* (eds.), *On the Mysteries of Unemployment*. Kluwer Academic Publishers, Dordrecht, 129-144.
6. Chamberlain, G. (1979): "Heterogeneity, Omitted Variable Bias, and Duration Dependence." DP691 Harvard Institute of Economic Research, Cambridge, New York.

7. Chesher, A., and T. Lancaster (1983): "The Estimation of Models of Labour Market Behaviour." *Review of Economic Studies*, 50, 609-624.
8. Davies, R.B., and A.R. Pickles (1983): "The Estimation of Duration-of-Residence Effects: A Stochastic Modelling Approach." *Geographical Analysis*, 15, 305-317.
9. Flinn, C., and J.J. Heckman (1982): "Models for the Analysis of Labor Force Dynamics." *Advances in Econometrics*, 1, 35-95.
10. Heckman, J.J. and G.J. Borjas (1980): "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence." *Economica*, 47, August, 247-283.
11. Menezes, R.M. (1995): "Patterns of Post-Displacement Adjustment in the Labour Market: Evidence from a Longitudinal Redundancy Study." Unpublished Ph.D. thesis, Department of Economics, University of Keele.
12. Narendranathan, W., and M.B. Stewart (1993): "How Does the Benefit Effect Vary as Unemployment Spells Lengthen?" *Journal of Applied Econometrics*, 8, 361-381.
13. Neyman, J., and E. Scott (1948): "Consistent Estimates Based on Partially Consistent Observations." *Econometrica*, 16, 1-32.
14. Theeuwes, J., M. Kerkhofs, and M. Lindeboom (1990): "Transition Intensities in the Dutch Labour Market 1980–85." *Applied Economics*, 22, 1043-1061.
15. van den Berg, G.J. and J.C. van Ours (1994): "Unemployment Dynamics and Duration Dependence in France, The Netherlands and The United Kingdom." *Economic Journal*, 104, March, 432-443.

Rui Menezes

Instituto Superior de Ciências do Trabalho e da Empresa,

Av. das Forças Armadas, 1600 Lisboa

Generalized Efficiency in Portfolio Selection Models

Vladimir Mlynarovic
University of Economics in Bratislava,Slovak Republic

The last fifteen years have seen a revolution in the way financial economists understand the investment world. Until the mid – 1980, financial economists' view of the investment world was based on the three pillars:

1. The CAPM is a good explanation of the fact that some assets (stocks, portfolios, strategies, or mutual funds) earn higher average return than others. The CAPM states that assets can only earn a high average return if they have a high "beta", which measures the tendency of the individual assets to move up or down with the market as a whole.
2. Returns are unpredictable, like a coin flip. Technical analysis that tries to divine future returns from patterns of past returns and prices is nearly useless. In addition, stock market volatility does not change much in time. Not only returns are close to unpredictable, they are nearly identically distributed, as well.
3. Professional managers do not reliably outperform simple indices and passive portfolios once one corrects for risk (beta). The average actively managed fund performs about 1 percent worse then the market index. The more actively a fund trades, the lower the returns to investor.

Together, these views reflect a guiding principle that asset markets are, to a good approximation, *informationally efficient* [Fama, 1970, 1991]. The only way to earn large returns is by taking on additional risk.

These views are not ideological or doctrinaire beliefs. Rather, they summarize the findings of a quarter century of careful empirical work. However, every one of them has now been extensively revised by a new generation of empirical research. The new findings need not overturn the cherished view that markets are reasonably competitive and, therefore, reasonably efficient. However, they substantially enlarge financial economists' view of what activities provide rewards for holding risks, and they challenge our understanding of those risk premiums. Now, financial economists know that:

1. There are assets whose average returns can not be explained by their beta. Multifactor extensions of the CAPM that associate high average returns with a tendency to move with other risk factors in addition to movements in the market as a whole dominate the description, performance attribution, and explanation of average returns.
2. Returns are predictable. This phenomenon occurs over business cycle and longer horizons. These predictions are not guarantees, there is a substantial risk - but the tendency is discernible.
3. Some mutual funds seem to outperform simple indices, even after controlling for risk through market betas. Fund returns are also slightly predictable. Multifactor models explain most fund persistence: funds earn persistent return by following fairly mechanical *styles*, not by persistent skill at stock selection.

Again, these statements are not dogma, but a caution summary of a large body of careful empirical work. The strength an usefulness of many results are hotly debated, as are the underlying reasons for many of these new facts.

2. The classical modern portfolio theory

The traditional academic portfolio theory, starting from Markowitz [1952] and expounded in every finance textbook, remains one of the most useful and enduring bits of economists developed in the last 50 years.

The traditional advice is to split your investments between a money-market fund and broad-based, passively managed stock fund. That fund should concentrate on minimizing fees and transaction costs. It should avoid the temptation to actively manage its portfolio, trying to chase the latest hot stock or trying to foresee market movements. An index fund or other approximation to the *market portfolio* that passively holds a bit of every stock is ideal. In the mean – variance space the capital market line describes the efficient portfolios and every portfolio on this efficient frontier can be formed as a combination of a risk-free money-market and the *market portfolio* of all risky assets. Therefore, every investor need only hold different proportions of these *two funds*.

This advice has had a sizable impact on portfolio practice. Before this advice was widely popularized in the early 1970s, the proposition that professional active management and stock selection could outperform blindly holding an index seemed self-evident, and passively managed index were practically unknown. They have exploded in size since then. The remaining actively managed funds clearly feel the need to defend active management in the face of the advice to hold passive index funds and the fact that active managers selected on any ex-ante basis under-perform indices ex-post.

2. The new modern portfolio theory

Mean – variance portfolio theory addresses the investor's asset selection problem for an investment horizon of one period. Progress in portfolio theory came as financial economists relaxed this restrictive assumption. The relaxation of the single-period assumption proceeded along two lines: first, in discrete time multiperiod models by Samuelason [1969], Rubinstein [1976] and others, and second, in continuos time models by Merton [1973] an others.

2.1 Merton's Intertemporal Capital Asset Pricing Model

Merton in his continuos-time intertemporal portfolio selection model [Merton, 1973] to derive the optimal consumption and investment policies $J(W, x, t)$, where $W(t)$ is investor's wealth at time t, $x(t)$ is a random variable, as a function of a consumption strategy and the investment strategy $\mathbf{w} = (w_1, ..., w_n)$ with $w_i(t)$ as a invest fraction of the wealth in risky asset i, i = 1, ..., n, Obtained

$$\mathbf{w} = \left(\frac{-J_W}{W\ J_{WW}} \right) \mathbf{C}^{-1}(\mathbf{E} - R_f \mathbf{e}) - \frac{J_{Wx}}{W\ J_{WW}} \mathbf{C}^{-1} s_x \tag{1}$$

where **C** is the covariance matrix with i ´ j element s_{ij} and s_x is a vector with ith element s_{ix}, **E** is the vector of expected returns of risky assets, R_f is the riskless rate of interest and **e** is te vector of ones.

If we define

$$\mathbf{w}_H = \frac{\mathbf{C}^{-1} s_x}{\mathbf{e}^T \mathbf{C}^{-1} s_x}$$

then the equation (1) can be written in the form

$$\mathbf{w} = \left(\frac{-J_W}{W\, J_{WW}} \right) \left[\mathbf{e}^T \mathbf{C}^{-1} (\mathbf{E} - R_f \mathbf{e}) \right] \mathbf{w}_T + \left(\frac{-J_{Wx}}{W\, J_{WW}} \right) \left[\mathbf{e}^T \mathbf{C}^{-1} s_x \right] \mathbf{w}_H \tag{2}$$

where $\mathbf{w}_T$ is the vector of portfolio weights of the tangency portfolio on the frontier of minimum variance portfolios generated by the n risky assets. This result defines a three-fund separation. The three-fund portfolio separation obtains: The investor invests in the riskless asset, the tangency portfolio $\mathbf{w}_T$ and the hedging portfolio $\mathbf{w}_{H.}$ The weights which the investor assigns to each portfolio depend on his preferences and are, therefore, investor-specific. We may interpret the three fund separation result as follow: The investor invests in the riskless asset an in the tangency portfolio, as in the mean-variance case, but modifies his portfolio investing in (or selling short) a third portfolio which has returns maximally correkated with changes in the state variable(s) x which represents shifts in the investment opportunity set and tastes. In general, if we have m state variables we obtain (m + 2)-fund separation where the investor invests in the riskless asset, the tangency portfolio and the m hedging portfolios.

2.2 Multifactor Portfolio Efficiency and Multifactor Asset Pricing

In the presented *Merton* approach an intertemporal model that uses utility maximization to get exact multifactor predictions of expected security returns was developed. This concept of multifactor portfolio efficiency plays a role in *Merton's* intertemporal CAPM (the ICAPM), like that of mean-variance efficienncy in the *Sharpe – Lintner* CAPM. The drawback of *Metron's* approach is degree of difficulty. His continuos time methods do not yield easy insight.

On the other hand, the powerful intuition of the CAPM centers on *Markowitz's* concept of mean variance efficiency. The CAPN starts with assumptions that inply that investors hold mean-variance-efficient portfolios. When there is a risk-free asset *f*, mean-variance-efficient combine *f* with one mean-variance-efficient portfolio of risky securities, the tangency portfolio *T*. Since the *T* is the risky component of all mean-variance-efficient portfolios, market-clearing security prices require that *T* is the value-weight market portfolio *M*. The familiar CAPM relation between the expected return $E(r_i)$ on any security *i*, and its market risk b_{iM} (the slope in the regression of r_i a r_M)

$$E(r_i) - R_f = b_{iM}\left\lfloor E(r_M) - R_f \right\rfloor \tag{3}$$

is then just the application to *M* of the condition on security weights that holds in any mean-variance-efficient portfolio.

Fama [Fama, 1996] has showed that *Merton's* ICAPM can be built on similar intuition. ICAPM investors hold multifactor-efficient portfolios that generalize the notion of portfolio efficiency. Like CAPM investors, ICAPM investors dislike wealth uncertainty. Byt ICAPM investors are also concerned with hedging more specific aspects of future consumption-investment opportunities, such as relative prices of consumption goods and the risk-return tradeoffs they will face in capital market. As a result, the typical ICAPM multifactor-efficient portfolio combines one of *Markowitz's* mean-variance-efficient portfolios with hedging portfolios that mimic uncertainty about the *m* future consumption-investment state variables of concern to investors.

The ICAPM risk-return relation is then a natural generalization of (3). One siply adds risk premiums for the sensitivities of r_i to the returns r_s, $s = 1, ..., m$, on the state-variable mimicking portfolios

$$E(r_i) - R_f = b_{iM}\left[E(r_M) - R_f\right] + \sum_{s=1}^{m} b_{is}\left[E(r_s) - R_f\right] \tag{4}$$

where b_{iM} and b_{is}, $s = 1, ..., m$, are the slopes from the multiple regression of r_i on r_M and r_s.

As in the CAPM, the ICAPM relation (4) between expected return and multifactor risks is the condition on the weights for securities that holds in any multifactor-efficient portfolio, applied to the market portfolio *M*. And just as market equilibrium in the CAPM requires that *M* is mean-variance efficient, in the ICAPM market-clearing price imply that *M* is multifactor- efficient. All these results one can derive from the detailed analysis of the following mathematical programming problem:

$$\min \quad o^2(r) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j s_{ij}$$

subject to

$$\sum_{i=1}^{n} w_i b_{is} = b_s, \quad s = 1, ..., m$$

$$\sum_{i=1}^{n} w_i E(r_i) = E(r)$$

$$\sum_{i=1}^{n} w_i = 1$$

where b_s, $s = 1, ...m$, are target values of state variables for multifactor-efficient portfolio, and the $E(r)$ is the target expected return of the portfolio.

References

1. Cochrane, J. H.,1999: New fact in finance. Economic Perspctives.Federal Reserve Bank of Chicago. Vol. 23, pp. 36 – 58.
2. Cochrane, J. H.,1999: Portfolio advice for a multifactor world. Economic Perspctives.Federal Reserve Bank of Chicago. Vol. 23, pp. 59 – 78.

3. Fama, E. F., 1996: Multifactor Portfolio Efficiency and Multifactor Asset Pricing. Journal of Financial and Quantitative Analysis, Vol. 3, No. 4, pp. 441 – 465.
4. Markowitz, H., 1952: Portfolio Selection. Journal of Finance, Vol. 7, No. 1, pp. 77 - 91
5. Merton, R. C., 1973: An Intertemporal Capital Asset Pricing Model. Econometrica, Vol. 41, No. 5, pp. 867 – 887.
6. Rubinstein, M., 1976: The Valuation of Uncertain Income streams and the Pricing of Option. Bell J. Econ. Manage. Sci., 7, 407 - 425
7. Samuelson, P., 1969: Lifetime Portfolio Selection by Dynamic Stochastic Programming. Rev. Econ. Stat. 57, pp.239 – 246.

Vladimír Mlynarovic

Department of Operations Research and Econometrics, University of Economics

Dolnozemská 1/b Bratislava, Slovak Republic

e-mail: mlynar@dec.euba.sk

OPTIMIZATION OF THE SYSTEM RELIABILITY

Jan Pelikán, VŠE Praha

The article deals with the problem of doubling the part of the system in order to increase the reliability of this system. The task is to find out which part should be doubled if it is not possible to double all of them for economic reasons. The goal is to maximize the system's reliability characterized by the probability of the faultless functioning or to minimize the losses caused by the system's failure. The article proposes a procedure based on linear programming methods that find the optimal decision, which of the parts should be doubled.

AMS classification: 90B25, 90C10

Key words: reliability, integer programming

INTRODUCTION

An optimization model proposed in the article is solving the question of the reserves for the functional components-parts of a mechanism in order to increase its reliability. It presupposes
the knowledge of the probabilities of these parts' failure and the estimation of the losses caused by this failure. The model is a problem with 0-1 variables [1]. The solution of the model divides the parts into those which are to be doubled and those which are not.

The following example illustrates the model. Large and complex mechanisms are composed of a great number of components, aggregates, partial machinery's. Each of the parts is responsible for the right functioning of the whole system and vice versa, each part's failure can disturb the system or completely put it out of operation and cause damages in its effect.

One of the possibility of eliminating or at least diminishing these damages is the doubling of some important parts. Having these parts doubled, there is a possibility to replace immediately a non-functioning part by a functioning one (or in other words, the failure is reduced only to a necessary time of a switch-over).

On the other hand, when the part is not doubled, it has to be removed from the system and then replaced by a new one. The example can be a power-distribution network composed of the electric line, switches, fuses, transformers and other parts. If, for example, a transformer fails out, the consumers dependent on this particular transformer are without the power supply for a certain time and the losses as an effect are obvious.

This period depends on the time of removing the transformer and replacing it by another one. If there is, at the same location, another transformer as a reserve, then the period of switching over the reserve is much shorter than the period of the transformer's replacement.

Similar problems can emerge in projects of a regulating system or a communication network and so on.

On the other side there are costs of doubling, that is the price of a doubled part. For that reason not all of the parts can be doubled, especially the expensive ones and also those whose failure does not bring so expensive damages.

RELIABILITY MODEL

If we want to know which parts are to be doubled, the following optimization model can be used. First we introduce the assumptions of the model.

Let us consider n parts of the system (aggregates, components) Z_1, $Z_2,...,Z_n$. Each of these parts is characterized by:

p_i probability of the failure-free run of Z_i without a reserve,

$\overline{p_i}$ probability of the failure-free run of Z_i with a reserve,

q_i the mean value of losses caused by Z_i's disorders without a reserve,

$\overline{q_i}$ the mean value of losses caused by Z_i 's disorders with a reserve,

c_i costs of the purchase and maintenance of the reserve for Z_i .

Obviously: $p_i \le \overline{p_i}$ and $q_i \ge \overline{q_i}$.

Next we assume:

- statistical independence of the failures of parts,
- the costs of the parts' doubling are limited by the amount K.

Let us introduce 0-1 variables $x_1, x_2, \ldots x_n$, the variable x_i involves the decision between the doubling of Z_i (x_i=1) or not-doubling (x_i=0) . Total costs of the reserves for the parts are $\sum_{i=1}^{n} c_i x_i$ and since the resources for reserves are limited by K, so it has to be valid $\sum_{i=1}^{n} c_i x_i \le K$.

Under that conditions we can:

(a) maximize the relialibility of the system, that is the failure-free run,

(b) minimize the mean value of the sum of losses caused by the parts' disorders.

In the case (a) the probability of the failure-free state of the system is the product of the probabilities of the failure-free states of all the parts.

The part Z_i will be failure-free with the probability $\overline{p_i}$, if it has a reserve (x_i = 1). If the part Z_i is without reserve(x_i = 0) then the failure-free probability is p_i . Altogether the probability of the part Z_i 's failure-free state can be put in the form $p_i + (\overline{p_i} - p_i)x_i$. Hence the total probability of the failure-free state is $\Pi = \prod_{i=1}^{n} [p_i + (\overline{p_i} - p_i)x_i]$.

After taking logarithm in order to make the objective function linear we get the objective function

in the form $z(x) = \log(\Pi) = \sum_{i=1}^{n} \log[p_i + (\overline{p_i} - p_i)x_i]$. This function will be maximized.

Since the expression $\log[p_i + (\overline{p_i} - p_i)x_i]$ for x_i=0 equals $\log[p_i]$ and for x_i=1 equals $\log[\overline{p_i}]$,

we can write the expression $\log[p_i + (\overline{p_i} - p_i)x_i]$ in the form $(1 - x_i)\log(p_i) + x_i \log(\overline{p_i})$.

The function

$$z(x) = \sum_{i=1}^{n} [(1 - x_i)\log(p_i) + x_i \log(\overline{p_i})] =$$

$$\sum_{i=1}^{n}[\log(p_i) + x_i \log(\overline{p_i}/p_i)] = \sum_{i=1}^{n}\log(p_i) + \sum_{i=1}^{n} x_i \log(\overline{p_i}/p_i)$$

expresses the logarithm of the whole system's reliability.

Maximizing reliability model is:

$$z(x) = \sum_{i=1}^{n}\log(p_i) + \sum_{i=1}^{n} x_i \log(\overline{p_i}/p_i) \to \max,$$

$$\sum_{i=1}^{n} c_i x_i \le K,$$

$$x_i \in \{0,1\}, i = 1,2,\ldots,n.$$

In the case (b) the mean value of the losses caused by the part Z_i's failure without the reserve is q_i and with the reserve $\overline{q_i}$. The mean value of the total losses is then:

$z(x) = \sum_{i=1}^{n}(1 - x_i)q_i + x_i \overline{q_i} = \sum_{i=1}^{n} q_i - \sum_{i=1}^{n} x_i \Delta q_i$, where $\Delta = q_i - \overline{q_i}$. This function will be minimized.

The minimal losses model is:

$$z(x) = \sum_{i=1}^{n} q_i - \sum_{i=1}^{n} x_i \Delta q_i \to \min,$$

$$\sum_{i=1}^{n} c_i x_i \le K,$$

$$x_i \in \{0,1\}, i = 1,2,\ldots,n.$$

Now, two models can be distinguished: The first one is a one-case situation, when the function operates for a short period - the static case. The second one is a long-time model for a longer time period - the dynamic model.

STATIC MODEL

The probability of the components' failure Z_i will be denoted as p_i. Consequently the probability of the failure-free run of the part without reserve is $p_i = (1 - p_i)$ and the probability of the failure-free run of the part with a reserve is $\overline{p}_i = (1 - p^2{}_i)$.

If the loss caused by one failure of the part Z_i is denoted by Q_i, the mean value of the losses is:

$q_i = p_i Q_i$ in case when there is no reserve for the part Z_i, $\overline{q}_i = p^2{}_i Q_i$ in case when there is a reserve.

Example.

Let us have parts Z_1, Z_2, Z_3 ,Z_4 ,Z_5. Table 1. contains their main characteristics.

The costs of the parts' doubling are limited by the amount K=100.

	Z_1	Z_2	Z_3	Z_4	Z_5
p_i	0.9	0.8	0.9	0.93	0.91
$p_i = (1 - p_i)$	0.1	0.2	0.1	0.07	0.09
p_i^2	0.1	0.04	0.01	0.0049	0.0081
$\overline{p_i} = 1 - p_i^2$	0.99	0.96	0.99	0.9951	0.9919
c_i	80	30	35	50	20
Q_i	1666	250	333	1613	989
$q_i = Q_i p_i$	167	50	33	113	89
$\overline{q}_i = Q_i p_i^2$	16.6	10	3.3	8	8
$\Delta q_i = q_i - \overline{q}_i$	150	40	30	105	81

Table 1.

Reliability model which maximizes the failure-free probability
is:

$$z(x) = \log(0.548402) + x_1 \log(0.99/0.9) + x_2 \log(0.96/0.8) + x_3 \log(0.99/0.9) + x_4 \log(0.9951/0.93) + \\ + x_4 \log(0.9951/0.93) + x_5 \log(0.9919/0.91) \to \max$$

$$80x_1 + 30x_2 + 35x_3 + 50x_4 + 20x_5 \le 100,$$

$$x_i \in \{0,1\}, i = 1,2,...,5 .$$

By using standard software LINGO [2] we get the optimal solution x=(0,1,1,0,1) with the failure-free probability equal to 0.789041, which is maximal. From the result follows that we have to double Z_2,Z_3,Z_5.

solution x	reliability	losses	doubling costs
(0,0,0,0,0)	0.548402	452	0
(0,1,1,0,1)	**0.789041**	301	85
(1,0,0,0,1)	0.657534	**221**	100
(1,1,1,1,1)	0.934507	46	215

Table 2.

Model which minimizes the mean
value of the total losses is :

$$z(x) = 452 - 150x_1 - 40x_2 - 30x_3 - 105x_4 - 81x_5 \to \min$$

$$80x_1 + 30x_2 + 35x_3 + 50x_4 + 20x_5 \le 100,$$

$$x_i \in \{0,1\}, i = 1,2,...,5 .$$

When we use again LINGO system, we get the optimal solution x=(1,0,0,0,1) with the minimal value of losses 221 in the mean value. According to this solution only Z_1 and Z_5 will be doubled.

The differences in the solutions obtained above can be explained by great influence of the amount of the losses in the optimal solution in the second model. First solution x=(0,1,1,0,1) means the most reliable system, but the losses are not minimal. Second solution x=(1,0,0,0,1) gives us less reliable system, but the losses are minimal.

We can observe the values of reliability, mean losses and costs of reserves for different solutions in the Table 2. The values with a bullet are optimal for K=100.

DYNAMIC MODEL

Let us suppose that the system's reliability should be optimized within a period <0,T> and the periods between failures of the parts are exponentially distributed. Let the mean value of the period between two failures of the component Z_i be $1/?_i$.

If the Z_i is without a reserve, then the probability of the failure-free run is $p_i = \exp(-l_i T)$. Since number of failures within the period <0,T> is Poisson distributed with the mean value $?_i$ T, the mean value of the losses in the case $x_i=0$ is equal to $q_i = l_i T Q_i$,where Q_i is the loss caused by the part Z_i 's failure.

In case that component Z_i has a reserve $(x_i=1)$ then the failed component is replaced by its reserve immediately; however, the installation of a new reserve takes a fixed time t_i, where t_i <<T. During the period t_i the failure-free probability is $p_i = \exp(-l_i t_i)$.
Since the mean value of the number of the component Z_i 's failures within <0,T> is $l_i T$, we get an approximate formula describing the reliability in the form

$$\overline{p}_i = \exp(-l_i T) + (1 - \exp(-l_i T)) \exp(-t_i l_i T).$$

When the part is doubled, the expected number of its double-failures is l_i times the total period of installations, which is approximately $(l_i T) t_i$. Hence the mean value of losses is $\overline{q}_i = l_i^2 T t_i Q_i$.

The reduction of time period T by reserve parts' installations was not taken into account in the formula for $\overline{q}_i$.

* The research was supported by the grant no.402/00/0458 of the Grant Agency of the Czech Republic.

BIBLIOGRAPHY

[1] Nemhauser,G.L.,Wolsey,L.A.: Integer and Combinatorical Optimization. J.Wiley, New York 1988.

[2] Solver Suite LINDO Systems Inc., Chicago 1995

Jan Pelikán,
University of Economics,
13067 Praque, Czech Republic

Some queue models with different service rates

Július REBO, Žilinská univerzita, DP Prievidza

1 Introduction: They are well known models in the queue theory in Kendall's classification denoted as M/M/n/N with equal rates of service in each service point. The first *M* describes an arrival assumption of requirement for service and follows the Poisson distribution with a finite rate $l > 0$. Each service point of n disposes of an exponential distributed service time with probability density function $f(t) = me^{-mt}$, $m > 0$ (the second *M*). An integer *N* denotes a number of common places reserved for requirements in the queue (n parallel service points) for their service and in a waiting line (N–n places). In the next parts we shall consider that number *N* can take countable value (either finite or infinite). Queues with different rates of service are studied as special cases of the models with one service point and a value of rate usually depends on some activity indicators of the systems. In this paper we shall be interested in a general method to derive characteristics of those systems when an assumption of the different rates represents a solitary property of system. We shall apply above assumptions in detail for a closed model marked M/M/n/m and accomplished results we shall use for a generalisation of two models M/M/n/N and M/M/n/∞. As a common foundation of solving method for considered models is an application of a general birth-death process.

2 General birth-death process: For the general birth-death process, e.g. in [1], we assume that rate of changes between states depends on the state of system. The states of system are usually meant as a number of requirement in the system and probabilities of those transitions during time dt we define then as:

from k to $k+1$: $l_k dt$, $k = 0, 1, 2, ..., N_0$, from k to k-1: $m_k\, dt$, $k=1, 2, 3, ..., N_0$

from k to k: $1 - (l_k + m_k)dt$, when $l_{N_0} = m_0 = 0$, $k = 0, 1, 2, ..., N_0$. (2.1.)

We can express a state probability $p_k = P\{S = k\}$ according to [1] in the form

$$p_k = \frac{l_0 l_1 ... l_{k-1}}{m_1 m_2 .. m_k} p_0 \text{ and } p_0 = \left[1 + \sum_{k=1}^{N_0} \frac{l_0 l_1 ... l_{k-1}}{m_1 m_2 .. m_k}\right]^{-1}, \; k = 0, 1, 2, ..., N_0. \qquad (2.2.)$$

3 Rates of service: Let us have a queue with n service points and a number of places in the queue is a positive integer N_0 which can take a finite or infinite countable value. The number of requirements for service has the Poisson distribution with finite rate *l > 0* and a ser-vice in each service point of *n* has an exponential density function $f_i(t) = m_i e^{-m_i t}$ with rates *m*$_i$ *> 0, i = 1,2, ...,n*. Thus, if the state *S* is equal to *k, n £ k £ N*$_0$, then a total rate of service m'_k in that state is not dependent on the understood state and is defined by the sum of service rates over each occupied service point, so we have $m'_n = m_1 + m_2 + ... + m_n$. It is clear to see that for states *S* equal to *k*, *k < n*, the total rate of service depends on the number of occupied service points and also on their own service rates m_k. Otherwise said, the rate of service for the state *S = k* is given by a combination of *k* service points from *n* without their repeating. Those combinations for every state *k Î { 1, 2, ... , n }* is $\binom{n}{k}$ and each rate m_k belongs to each other exactly $\binom{n-1}{k-1}$ times.If we are assuming that entered requirement is served with equal probability $p_i^{(k)} = \quad = p^{(k)} = \frac{1}{\binom{n}{k}}$ for each combination $C_i^{(k)}(m_1,...,m_n)$ of service rates in state *S = k* , we shall define the rate of the service m'_k in the state *S = k* as

$$m'_k = \sum_{i=1}^{\binom{n}{k}} p_i^{(k)} C_i^{(k)}(m_1,...,m_n) = p^{(k)} \sum_{i=1}^{\binom{n}{k}} C_i^{(k)}(m_1,...,m_n) = \frac{k}{n}(m_1 + m_2 + ... + m_n) = k\hat{m}, \qquad (3.1.)$$

when $\hat{m} = \frac{m_1 + m_2 + ... + m_n}{n}$.

Consider now, every combination of service rates $C_i^{(k)}(m_1,...,m_n)$ in the state *S = k* will occur with a probability $p_i^{(k)}$ so that $p_i^{(k)} \neq p_j^{(k)}$, for *i ¹ j, i, j Î { 1, 2, ... ,* $\binom{n}{k}$*}*. Set $s_i^{(k)} = \sum_j p_{j_i}^{(k)}$ when we are summing over all indexes of probabilities $p_j^{(k)}$ corresponding to combinations ob-taining a service rate m_i. The probability $s_i^{(k)}$ indicates a total probability for an appearance of the rate m_i in the given combinations of *k* rates. For a total service rate in the state *S = k* then we have

$$m'_k = \sum_{i=1}^{\binom{n}{k}} p_i^{(k)} C_i^{(k)}(m_1,...,m_n) = \sum_{i=1}^{\binom{n}{k}} \left(\sum_j p_{j_i}^{(k)} \right) m_i = \sum_{i=1}^{\binom{n}{k}} s_i^{(k)} m_i \quad . \qquad (3.2.)$$

4 Closed model M/M/n/m: A closed model of queue with waiting line is called a mo-del with a finite source of requirements for service. It reflects real systems which get an essential weight within applications. The first of all is a problem of the service of the several machine. A repairman (or team of repairmen *n*) tends *m (n < m)* of machines. The machines are running until they drop out. A role of repairman (repairmen) is to eliminate those accidents.

We assume that number of requirements follows the Poisson distribution with parameter $l > 0$ and each service point has an exponencial density function $f_i(t) = m_i e^{-m_i t}$ of service time with $m_i > 0$, for $i = 1,2, \dots, n$. States of system $k\, \hat{I} \{0, 1, \dots , m\}$ mean that exactly k machines underlie in service or they are in waiting line. Always $m - k$ machines occur out of the system and they represent active machines. Intensity of the requirements for service is comparable to the number of machines out of the system and it can express as $(m - k)l$. The total rates of service in the state $S = k$ represents exactly k occupied service points and we can denote it as m'_k, for k £ n. If is $k > n$ then the rate of being service is denoted m'_n.

According to (2.1.) and a previous analysis transition probabilities during time dt will be:

from k to $k+1$: $l_k dt = (m-k) l\, dt$, 0 £ k £ $m - 1$, from k to k-1: $m'_k\, dt$, 1 £ k £ n,

from k to k-1: $m'_n\, dt$, $n+1$£ k £ m, from k to k: 1- $(l_k + m'_k)dt$, when $l_m = m'_0 = 0$, 0 £ k £ m.

In general there is known a solution of the previous problem e.g. in [1], [2], if all service rates m_i are equal to$m$. We shall call that model as a basic model and if we denote $y = \dfrac{l}{m}$ then its probabilities of states hold: $p_k = \binom{m}{k} y^k p_0$, $1 \le k < n$, $p_k = \binom{m}{k} \dfrac{k!}{n^{k-n} .n!} y^k p_0$, $n \le k \le m$,

$$p_0 = \left[\sum_{k=0}^{n-1} \binom{m}{k} y^k + \sum_{k=n}^{m} \binom{m}{k} \frac{k!\, y^k}{n^{k-n} .n!} \right]^{-1} . \qquad (4.1.)$$

It is clear to see that the above model is the specific birth-death process. By means of that mo-del we can express probability of state $S = k$ for a closed queue model with different rates according to (2.2.) as $p_k = \dfrac{m!\, l^k}{(m-k)!\, m'_1 m'_2 .. m'_k} p_0$, 1£ k £ n, $p_k = \dfrac{m!\, l^k}{(m-k)!\, m'_1 m'_2 \dots m'_n (m'_n)^{k-n}} p_0$, $n+1$ £ k £ m, so that $\sum_{k=0}^{m} p_k = 1$.

Probabilities of states p_k will be next arranged if we depose $\hat{m}$ from (3.1.) and $\hat{y} = \dfrac{l}{\hat{m}}$. We shall get $p_k^{(c)} = \dfrac{m!\, l^k}{(m-k)!\, \dfrac{k!}{n^k} (m_1 + \dots + m_n)^k} p_0^{(c)} = \binom{m}{k} \dfrac{l^k}{\hat{m}^k} p_k^{(c)} = \binom{m}{k} \hat{y}^k p_0^{(c)}$, 1£ k £ n,

$$p_k^{(c)} = \frac{m!\, l^k}{(m-k)!\, \dfrac{n!}{n^n} (m_1 + \dots + m_n)^n (m_1 + \dots + m_n)^{k-n}} p_0^{(c)} = \binom{m}{k} \frac{k!\, l^k}{n!\, n^{k-n} \hat{m}^k} p_0^{(c)} = \binom{m}{k} \frac{k!}{n!\, n^{k-n}} \hat{y}^k p_0^{(c)},$$

for $n +1$ £ k £ m, so as $p_0^{(c)}$ express with condition $\sum_{k=0}^{m} p_k^{(c)} = 1$. (4.2.)

5 Open model M/M/n/N: On the other hand if we shall assume an infinite source of requirement to the queues they shall be called open models then. Consider now so-called open system M/M/n/N with n parallel service points and waiting line in which our previous assumptions stand. Thus we assume different service rates with an exponential density function accor-ding to 2nd and 3rd chapters. Let $l > 0$ be the intensity of incoming stream of requirement to the queue. A number of places in queue takes a given value and it is denoted as N. If we consider the state of system as a number of requirements in the system we can express transition probabilities during dt as:

from k to $k+1$: $l_k dt = l dt$, 0 £ k £ N, from k to $k-1$: m_k' dt, 1 £ k £ n,

from k to $k-1$: m_n' dt, $n + 1$ £ k £ N, from k to k: 1- $(l_k + m_k')dt$, 0 £ k £ N, $l_N = m_0' = 0$.

Substituting corresponding rates to the (2.2.) and next modifying we shall get proba-bilities for a queue with a finite waiting line and different service rates in the form:

$$p_k^{(N)} = \frac{\hat{y}^k}{k!} p_0^{(N)}, \text{ for } 1 \le k < n, \quad p_k^{(N)} = \frac{\hat{y}^k}{n!\, n^{k-n}} p_0^{(N)}, \text{ for } n \le k \le N, \qquad (5.1.)$$

subject to $p_0^{(N)} = \left[\sum_{k=0}^{n-1} \frac{\hat{y}^k}{k!} + \sum_{k=n}^{N} \frac{\hat{y}^k}{n!\, n^{k-n}} \right]^{-1}$, where $\hat{y} = \frac{l}{\hat{m}}$ and $\hat{m} = \frac{m_1 + \ldots + m_n}{n}$.

Using a well-known technique under previous assumptions for different rates of service and equal probability of the service we can derive for $N \to \infty$ state probabilities of the system M/M/n/∞. So we have

$$p_k^{(\infty)} = \lim_{N\to\infty} p_k^{(N)} = \frac{\hat{y}^k}{k!} p_0^{(\infty)}, \; 1 \le k < n, \; p_k^{(\infty)} = \lim_{N\to\infty} p_k^{(N)} = \frac{\hat{y}^k}{n!\, n^{k-n}} p_0^{(\infty)}, \quad n \le k, \qquad (5.2.)$$

and $p_0^{(\infty)}$ is determined by a condition $p_0^{(\infty)} = \left[\sum_{k=0}^{n-1} \frac{\hat{y}^k}{k!} + \sum_{k=n}^{\infty} \frac{\hat{y}^k}{n!\, n^{k-n}} \right]^{-1} = \left[\sum_{k=0}^{n-1} \frac{\hat{y}^k}{k!} + \frac{\hat{y}^n}{(n-1)!\,(n-\hat{y})} \right]^{-1}$.

If we replace $\hat{y} = \frac{l}{\hat{m}}$ with $\hat{y} = \frac{l}{m}$ in (5.1.) and (5.2.), for $m = m_1 = m_2 = \ldots = m_n$, we shall get the models called again basic models.

For a queue M/M/n/∞ we need to remind also a condition for its stability treatment. In the basic model that condition has a form $l < nm$. It indicates that a total capacity of service nm has exceed an incoming rate l to the queue. In the case we reason different service rates we shall get an analogous request for an incoming rate and service rates $l < m_1 + m_2 + \ldots + m_n$.

6 Unreliable service points: We can take advantage of a previous technique from 3rd chapter used for different services rates for solving queues with unreliable service points. Ven-cetel introduces in [3] a very simple approach to the queue model with unreliable service points which consists in a correction of the model parameters for a single service point. We mean a queue with a given probability of the successful service p which indicates a probability to terminate successfully of service for attendant requirement in the queue. Moreover, a probability $1 - p$ indicates a probability of failure service for attendant requirement in the queue. Let m denote a rate of the service for the basic model in each service point. Let next r_k express a probability of the successful service in the kth service point and 1 - r_k is the probability of the failure of the service. Then $m_k = r_k m$ yields the rate of the successful service for kth service point. Assuming equal probabilities of access to an arbitrary service point $p^{(k)} = \frac{k}{n}$, k Î{1, 2, ... , n} we shall get models with different service rates from 2nd and 3rd chapters. We derived probabilities of states in formulas (4.2.), (5.1.), (5.2.) for the meant queues where is $\hat{m} = \frac{m_1 + ... + m_n}{n} = \frac{r_1 m + ... + r_n m}{n}$. Thus the queue model with different rates of service seems to be a special case of the basic model with unreliable service points. We can consider that those different rates of service represent a fundamental conditions for functioning of the queue. Consider that an efficiency of every point and their reliability are different. That point of view leads to a very general model with the different rates for the unreliable service points. Its solving method is assembling both previous methods. Let an every service point operate with the service rate u_k and corresponding probability of the successful service will be r_k. Then $m_k = r_k u_k$ expresses the rate of the successful service for kth service point. Under condition of the equal probability of the access to service (3.1.), we get again a model with the different service rates with probabilities of the states (4.2.), (5.1.), (5.2.), setting $\hat{m} = \frac{r_1 u_1 + r_2 u_2 + ... + r_n u_n}{n}$.

7 Optimal rate of service: Let us look at a problem of optimising rate of service for the queue M/M/n/∞. Then we can use a submitted technique for the other understood models with the different service rates after minimal modifications.

Let us have a queue with n service points with an exponential density function $f_i(t) = m_i e^{-m_i t}$ of the service time, with $m_i > 0, 1 \le i \le n$. Let c_1 be average service costs per time unit and let c_2 be average store costs per time unit both reduced per one requirement. Next let $d_S(n,\hat{m})$ be the mean-value of requirements in the queue depended on the number of the service points *n* and the average rate of service $\hat{m}$ from (3.1.). By course of that we have done

$$d_S(n,\hat{m}) = \sum_{k=0}^{\infty} k p_k^{(\infty)} = \hat{y} + \frac{\hat{y}^{\,n+1}}{(n-1)!\,(n-\hat{y})^2}\, p_0^{(\infty)}, \text{ when } p_0^{(\infty)} = \left[\sum_{k=0}^{n-1} \frac{\hat{y}^{\,k}}{k!} + \frac{\hat{y}^{\,n}}{(n-1)!\,(n-\hat{y})}\right]^{-1}. \quad (7.1.)$$

A total costs function $C(n,\hat{m}) = c_1 n\hat{m} + c_2 d_S(n,\hat{m})$ includes the service costs with the costs for waiting in line. Moreover we assume that a number of service points n is given and the rate $l > 0$ of the Poisson incoming stream of requirement we shall take as a constant value to the optimising variable $\hat{m}$. Valuations of the criteria function $C(n,\hat{m})$ and $d_S(n,\hat{m})$ will be dependent on the continuous variable $\hat{m}$ and it will be denote $C(\hat{m}) = c_1 n\hat{m} + c_2 d_S(\hat{m})$. (7.2.)

Then we shall specify an optimal average value of rate $\hat{m}$ by derivative of the criteria function $C(\hat{m})$ with respect $\hat{m}$. We shall compute a necessary condition of the existence of the minimum value setting the first derivative (7.2.) equal to zero. Thus we have

$$\frac{dC(\hat{m})}{d\hat{m}} = \frac{d\left[c_1 n\hat{m} + c_2 d_S(\hat{m})\right]}{d\hat{m}} = 0. \quad (7.3.)$$

Employment of general form of derivative (7.3.) leads to equations of high orders whose solution is possible only with computational approach and it does not let us analyse the solution in consideration of the costs. Let $\hat{m}^*$ denote an optimal average rate of service from (7.3.). We have another problem how to optimise an optimal rate $\hat{m}_i^*$ for every service point. Denote Δh as a difference of the optimal average rate $\hat{m}^*$ and the average rate $\hat{m}$. If it is $\hat{m}^* - \hat{m} > 0$ then $\Delta h = +h$ and if it is $\hat{m}^* - \hat{m} < 0$ then $\Delta h = -h$. Next denote $j_i = \frac{m_i}{m_1 + \ldots + m_n}$, *i =1,...,n*. The optimal average service rate we can express as $\hat{m}^* = \hat{m} + \Delta h$. After reform we have done

$$\hat{m}^* = \frac{m_1 + \ldots + m_n + n\Delta h}{n} = \frac{\left(m_1 + \frac{m_1}{m_1 + \ldots + m_n} n\Delta h\right) + \ldots + \left(m_n + \frac{m_n}{m_1 + \ldots + m_n} n\Delta h\right)}{n} =$$

$$= \frac{(m_1 + j_1 n\Delta h) + \ldots + (m_n + j_n n\Delta h)}{n} = \frac{m_1^* + \ldots + m_n^*}{n}.$$

So, to obtain an optimal average service rate $\hat{m}_i^*$ for the *ith* service point it is needed to revise initial rates m_i according to $m_i^* = m_i + \dfrac{m_i}{m_1 + \ldots + m_n} n\Delta h = m_i + j_i n\Delta h = m_i + j_i n\left(\hat{m}^* - \hat{m}\right)$.

8 Conclusions: From the previous results it follows that basic models of the queue with equal service rates and the queues with the different rates have a more simple relationship than we would expect. Establishing the different rates of the service enables a generalisation of the solution also for the queues with the unreliable service points which frame a temporal mo-del between the ba-sic model and the model with the different service rates. The above descri-bed method optimising an average rate of service also allows us to use it for optimising all of the previous models of the queue with a respect to their a costs. So we have a solid common technique for solving and optimising the whole class of the queue models.

Acknowledgement

The solved problem is a part of the research project supported by the Scientific grant Agency of Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences under grant No. 1/7211/20.

References

[1] Gross, D., Harris, C., M.: *Fundamentals of Queueing Theory,* Wiley, New York 1985, 1998

[2] Saaty, T., L.: *Elements Of Queueing Theory and Its Applications,* Wiley, New York, 1963

[3] Vencetel, E., S.: *Issledovanie operacij*, Sovetskoe radio, Moskva 1972

RNDr. REBO JÚLIUS,

ŽILINSKÁ UNIVERZITA V ŽILINE,

DP FAKULTY RIADENIA A INFORMATIKY,

BAKALÁRSKA 2, 971 01 PRIEVIDZA

e–mail : rebo@utcpd.sk

INTERVENTION MODELS IN TIME SERIES OF UNEMPLOYMENT

Eva Rublíková
Department of Statistics of Economic University

The economic time series of unemployment are frequently affected by policy changes and others events that are known to have occurred at a particular point of time. Events of this type, whose timing are known, have been termed interventions by Box and Tiao (1975). Interventions can affect a time series in a several ways. They can change the level, either abruptly or after some delay, change trend, or lead to other, more complicated, response pattern. Ignoring interventions can lead to an inadequate ARIMA model being fitted and a poor forecast being made.

Interventions can be incorporated into univariate ARIMA model by extending it to include deterministic (or dummy) input variables I_t. I_t is the dummy or indicator sequence taking values 1 and 0 to denote the occurrence or nonoccurrence of the exogenous intervention. The following dummy variables have been found to be useful for representing various forms of interventions:

l) A *pulse* variable, which models an intervention lasting only for the observation T,

$$I_t = e_t^{(T)}, \text{ where } e_t^{(T)} = \begin{cases} 1, t = T \\ 0, t \neq T. \end{cases}$$

2) A *step* variable, which models a step change in y_t beginning at observation T,

$$I_t = x_t^{(T)}, \text{ where } x_t^{(T)} = \begin{cases} 1, t \rangle T \\ 0, t \geq T. \end{cases}$$

3) An *extended pulse* variable, useful for modeling ´policy on-policy off´ interventions,

$$I_t = h_t^{(T_1,T_2)}, \text{ where } h_t^{(T_1,T_2)} = \begin{cases} 1, \ \mathrm{T}_1 \leq \ t \ \leq \ T_2 \\ 0, \ \text{otherwise,} \end{cases}$$

noting that $h_t^{(T_1,T_2)} = \sum_{j=0}^{T_2-T_1} e(T_1+j)_t = (1+B+\ldots+B^{T_2-T_1})s_t^{(T_2)}$.

If y_t is generated by an ARIMA(p,q) process, then an intervention model may be postulated as

$$y_t = n(B)I_t + N_t \qquad (1)$$

where

$$N_t = \frac{q(B)}{f(B)}a_t \qquad (2)$$

is the ´noise ´ model, $n(B)$ is a (possibly infinite) polynomial which may admit a rational form such as

$$n(B) = \frac{w(B)}{d(B)}B^b, \qquad (3)$$

where

$$w(B) = w_0 - w_1 B - w_2 B^2 - \ldots - w_m B^m,$$

$$d(B) = 1 - d_1 B - \ldots - dl_r B^r,$$

where b measures the delay in effect (or dead time).

In general, if polynomial $d(B) = 1$ ($r = 0$), finite responses of length m are obtained, whereas if $r > 0$, responses of infinite length are obtained.

For various specifications of $n(B)$ we have various responses to pulse and a step change ´input´ that are of practical interest.

1) If time series y_t is affected by a pulse input $I_t = e_t^{(T)}$, the polynomial $n(B)$ has several forms:

a) $n(B) = \dfrac{w}{1-dB}$,

what means, that I_t has only a transient effect on y_t with w measuring the initial increase and d the rate of decline. If $d = 0$, then only an instantaneous effect is felt, whereas if $d = 1$, the pulse input is really a step change, and the effect is permanent;

b) $$n(B) = \frac{w_0}{1-dB} + \frac{w_1}{1-B},$$

represents the situation where, apart from the transient effect w_0, the possibility is entertained that a permanent gain (or loss), w_1 in y_t is obtained;

c) $$n(B) = w_0 + \frac{w_1}{1-dB} + \frac{w_2}{1-B},$$

shows the case of an immediate positive response followed by a decay and, possibly, a permanent residual effect, and this might well represent the dynamic response of unemployment to a policy decision.

2) If time series y_t is affected by a step input $I_t = x_t^{(T)}$, the polynomial $n(B)$ has also several forms:

a) $$n(B) = w,$$

shows an immediate step response of w;

b) $$n(B) = \frac{w}{1-dB},$$

shows the situation of first-order dynamic response and an eventual, or long-run response, of $\frac{w}{1-d}$;

c) $$n(B) = \frac{w}{1-B},$$

represents the case when $d = 1$, in which the step change produces a ramp or trend in y_t.

Obviously, these models can be readily extended to represent many situations of potential interest.

If the noise model is of multiplicative form

$$N_t = \frac{q(B)\Theta(B^s)}{(1-B)^d(1-B^s)^D f(B)\Phi(B^s)} a_t, \tag{4}$$

and if there are J interventions, the model given by equations (1)-(3) can be extended to

$$(1-B)^d(1-B^s)^D y_t = \sum_{j=1}^{J} \frac{w_j(B)}{d_j(B)} B^{b_j}(1-B)^d(1-B^s)^D I_{jt} + \frac{q(B)\Theta(B^s)}{f(B)\Phi(B^s)} a_t . \qquad (5)$$

In building models of the type (5), parsimonious forms are initially postulated to represent the expected effects of the interventions, with more complex form only being considered if knowledge of the dynamics of the intervention, or subsequent empirical evidence, suggests so. The identification procedures may be applied to data prior to the occurrence of the interventions if a sufficiently large number of such observations are available. If the effects of interventions are expected to be transient, then the identification procedures may be applied to the entire data set. Alternatively, the response polynomials $n_j(B) \cong d_j^{-1}(B) w_j(B) B^{b_j}$ may be estimated by least squares method, for suitably large maximum orders, and the identification procedures applied to the residuals $y_t - \sum \hat{n}_j(B) I_{jt}$.

Once a model (5) has been specified, the intervention and noise parameters can be estimated simultaneously by maximum likelihood or nonlinear least squares.

The mentioned theory of intervention models will be applied to the variables of unemployment of shool-leavers to express the effects of policy decisions of the gowerment.

References

1. Cromwell, J.B., Labys, W.C. and Terraza, M. (1999). Univariate Tests for Time Series Models, Sage University Paper.
2. Enders, W. (1995). Applied Econometric Time Series, John Wiley and Sons.
3. Mills, T. C. (1990). Time Series for Economists, Cambridge University Press.
4. Pacáková, V. (1999). Identification of Problem Groups and Problem regions with Regard to the Unemployment in the Slovak Republic. 21th Polish Symposium on Spatial Econometrics, Zakopane, Poland.

5. Pacáková, V. (1998). Probability -statistical Models in Labour Market Analysis. 20th Polish Symposium on Spatial Econometrics, Zakopane, Poland.
6. Rublíková, E., Pacáková, V. (2000). Štatistické modely v analýzach trhu práce, Ekonóm, Bratislava.
7. Hill M.M.: Gender Differences in Pay in Lisbon Hotels: A pilot study in:Anniversary volume from the Sixth international conferenve "Quantitative Methods in Buseness and Management", Bratislava, Nov 4-5, 1999, FHI EU Bratislava, p. 98-104
8. Figueira, J.- M. A.: Some Questions About Modeling Pensions Funds. in: Anniversary volume from the Sixth international conference :Quantitative Methods in Business and Management", Bratislava, Nov 4-5, 1999. FHI EU Bratislava, p. 81-87
9. Linda, B.- Kahoutová, J.: Stochastic model of running of means of transport on the rote. in.: Mathematical methods in economy and industry, 11-th joint Czech-German-Slovak Conference, Liberec 1998

Doc. RNDr. Eva Rublíková, PhD.
Department of Statistics
FEI Economic University
Dolnozemská cesta 1/b
85235 Bratislava

A Note on Stochastic Dynamic Programming for the Mean Variance Model

Milan Sitar[1]

Department of Probability and Statistics
Faculty of Mathematics and Physics, Charles University

Let us consider a Markov reward process, where two important aspects of the model are changed:

1. the rewards are random variables instead of known constants and
2. we allow for any decision rule over the moment set of the reward distribution, rather than assuming maximization of the expected value of the reward outcome.

These modifications provide a natural setting for the rewards to be normally distributed, and thus, applying the mean variance models becomes possible. Mean variance models were originally proposed for the portfolio selection problem by Markowitz [5] and Tobin [4].

This problem was originally formulated in Goldwerger [1] and studied for the maximizing of the mean-variance $\left(\frac{m}{s}\right)$ of the reward outcome. Unfortunately, the key formula presented in this paper is incorrect, see [2] for a counterexample. Moreover, the principle of optimality cannot be applied easily for this model. The aim of the note is to present correct expression for the mean-variance $\left(\frac{m}{s}\right)$ (cf. [1], equation (13)). Notations used in [1] will be followed as close as possible.

1 A GENERAL DYNAMIC PROGRAMMING ALGORITHM

First of all we recall the dynamic programming algorithm. We suppose that we have *N* possible states, which are presented by rows of the stochastic matrix $\{p_{ij}\}$. This matrix describes the probabilistic structure of the Markov process. In each state we will have a finite number of alternative choices. If the process is now in state *i* and action k has been chosen from the alternatives, then p_{ij}^k is the probability of transition to state *j*, $R_{ij}^k(x)$ is the probability distribution of rewards. Thus

$$p_{ij} = P \text{ (transition from } i \text{ to } j), \tag{1}$$

$$R_{ij}(x) = P(X = x \,|\, \text{transition from } i \text{ to } j), \tag{2}$$

Let be $F_{ij}(x)$ the joint probability function of the rewards, then we can write

$$F_{ij}^k(x) = p_{ij}^k R_{ij}^k \tag{3}$$

We define the total probability function associated with given reward x for the next transition, when the system is now in state *i* and alternative k is chosen, as

$$F_i^k(x) = \sum_{j=1}^{N} F_{ij}^k(x) \tag{4}$$

[1] This research was supported by the Grant Agency of the Czech republic under Grant 402/99/11336

We assume a finite planning horizon of T periods of time. We will seek the total probability function of the stochastic present value of the system denoted by $H_i^k(T,y)$. This probability distribution over y (the stochastic discounted present value of all future rewards) depends on the present state i, on the planning horizon T and on the chosen alternative k at each point of time ($y=\sum_{t=1}^{T}(1+r)^{-1}x_t$, where r is the discount rate and x_t the reward at time t – we suppose that x_t, t = 1,..., T, are independently distributed).

When we seek an optimal solution of the dynamic programming problem for a stochastic Markovian process, we use Bellman ‘s principle of optimality (see [3]). An optimal policy has the property that whatever the initial state and initial decision are, the remaining decision must constitute an optimal policy with regard to the state resulting from the first decision. Therefore we first find the optimal decision $k^{\bullet}(1)$ for the period before the last (we denote z = 1, $t = T - z$), then knowing this decision and resulting probability distribution function $H_i^{k^{\bullet}}(1,y)$ we go one period backward and find the optimal decision $k^{\bullet}(2)$ (the decision horizon is two periods (z = 2)), and so on.

For any $z\geq 1$ we denote the decision function by G and the chosen alternative by $k^{\bullet}$, so that $k^{\bullet}(z)=G\{H_i^k(z,y)\}$. We denote by $f_i(z,y)$ the probability distribution function of starting in state i at stage z of the stochastic present value y, which results from the above decision rule, that is $f_i(z,y)=H_i^{k^*}(z,y)$. We can use a notation D to designate the above two-stage decision process:

$$f_i(z,y)=D\{H_i^k(z,y)\} \tag{5}$$

Therefore for z = 1 we obtain

$$f_i(1,y)=D\{H_i^k(1,y)\}=D\{F_i^k[(1+r)y]\} \tag{6}$$

When the system is in period T–2 in state i, then for any chosen alternative k we get

$$R_{ij}^k(x)*f_j\left[1,(1+r)^{-1}y\right]=P(Y=x+(1+r)^{-1}y \mid \text{the next transition from } i \text{ to } j), \tag{7}$$

where Y denotes a stochastic present value and * denotes convolution. Then we can obtain joint probability function of the stochastic present value at T–2 and the next transitions associated with alternative k denoted $H_{ij}^k(2,y)$ as

$$H_{ij}^k(2,y)=p_{ij}^k\left[R_{ij}^k(x)*f_j\left(1,(1+r)^{-1}y\right)\right] \tag{8}$$

The total probability function associated with a given y, when the system is in state i at T–2 and alternative k is chosen, is

$$H_i^k(2,y)=\sum_{j=1}^{N}H_{ij}^k(2,y)=\sum_{j=1}^{N}p_{ij}^k\left[R_{ij}^k(x)*f_j\left(1,(1+r)^{-1}y\right)\right] \tag{9}$$

Then

$$f_i(2,y)=D\{H_i^k(2,y)\}=D\left\{\sum_{j=1}^{N}p_{ij}^k\left[R_{ij}^k(x)*f_j\left(1,(1+r)^{-1}y\right)\right]\right\} \tag{10}$$

We can generalize equation (10) as

$$f_i(z,y)=D\{H_i^k(z,y)\}=D\left\{\sum_{j=1}^{N}p_{ij}^k\left[R_{ij}^k(x)*f_j\left(z-1,(1+r)^{-1}y\right)\right]\right\} \tag{11}$$

since the convolution operation is commutative and associative.

2 DYNAMIC MEAN – VARIABILITY APPROACH ALGORITHM

Under the assumptions of the mean variability approach we maximize the ratio of the mean over the standard deviation of the portfolio's stochastic present value distribution. The definition of the decision process is therefore:

a) All distribution functions are from the normal family of distributions.

b) For every state of the world (th i-th row of the stochastic matrix) we define the set of alternative K, by the superscript k: $K = \{k \mid k$ is superscript denoting the alternatives for every state $i\}$.

c) With every alternative there is an associated $\left(\frac{m}{s}\right)$.

d) The distribution function thet results from the choice of an alternative is denoted by $f_i^t(t, y)$, where i denote the state and t is a time index and y is the stochastic present value.

Therefore

$$X \cong R_{ij}^k(x) = N\left(m_{ij}^k, s_{ij}^{2k}\right)$$

and when $Y(z)$ is a stochastic variable with distribution function $f_i\ (z, y)$ the denoting

$$E\left[Y(z)\right] = m_j(z) \text{ and } var\left[Y(z)\right] = s_j^2(z)$$

we have for $z = 1$:

$$f_i(1, y) = H_i^{k*}(1, y) = D\left\{H_i^k(1, y)\right\} = D\left\{\sum_{j=1}^{N} F_{ij}^k(x)\right\} = D\left\{\sum_{j=1}^{N} p_{ij}^k R_{ij}^k(x)\right\}$$

where $y = (1+r)^{-1} x$ and $k*$ is an optimal at time T–1 and

$$E\left[Y(1)\right] = m_j(1) = \sum_{j=1}^{N} p_{ij}^{k*} m_{ij}^{k*}$$

$$var\left[Y(1)\right] = s_j^2(1) = \sum_{j=1}^{N} p_{ij}^{k*} m_{ij}^{(2)k*} - \left(\sum_{j=1}^{N} p_{ij}^{k*} m_{ij}^{k*}\right)^2$$

where $m_{ij}^{(2)k*}$ is the second moment of R_{ij}^{k*}. The distribution of $Y(z)$ may not be normal as in [1]. For $z > 1$ we can write

$$f_i(z, y) = H_i^{k*}(z, y) = D\left\{H_i^k(z, y)\right\} \tag{12}$$

$$= D\left\{\sum_{j=1}^{N} p_{ij}^k\left[R_{ij}^k(x) * f_j\left(z-1, (1+r)^{-1} y\right)\right]\right\}. \tag{13}$$

The convolution $R_{ij}^k(x) * f_j\left(z-1, (1+r)^{-1} y\right)$ has the mean $m_{ij}^k + (1+r)^{-1} m_j(z-1)$ and the variance $s_{ij}^{2k} + (1+r)^{-2} m_j^2(z-1)$. Therefore the total probability function associated with kth alternative reads

$$H_i^k(z, y) = \sum_{j=1}^{N} p_{ij}^k\left[R_{ij}^k(x) * f_j\left(z-1, (1+r)^{-1} y\right)\right] \tag{14}$$

with the mean

$$m_i^k(z) = \sum_{j=1}^{N} p_{ij}^k m_{ij}^k + (1+r)^{-1} \sum_{j=1}^{N} p_{ij}^k m_j(z-1)$$

and the variance

$$\begin{aligned} s_i^{2k}(z) = {} & \sum_{j=1}^{N} p_{ij}^k s_{ij}^{2k} + \sum_{j=1}^{N} p_{ij}^k \left(s_{ij}^k\right)^2 + 2(1+r)^{-1} \sum_{j=1}^{N} p_{ij}^k m_{ij}^k m_j(z-1) + (1+r)^{-2} \sum_{j=1}^{N} p_{ij}^k s_j(z-1) + \\ & + (1+r)^{-2} \sum_{j=1}^{N} p_{ij}^k \left[m_j(z-1)\right]^2 - \left(\sum_{j=1}^{N} p_{ij}^k m_{ij}^k\right)^2 - 2(1+r)^{-1} \sum_{j=1}^{N} p_{ij}^k m_{ij}^k \sum_{j=1}^{N} p_{ij}^k m_j(z-1) - \\ & - (1+r)^{-2} \left[\sum_{j=1}^{N} p_{ij}^k m_j(z-1)\right]^2 \end{aligned}$$

This variance formula is different from that in [1]. The present formula is in accordance with the counter example in [2].

REFERENCES

[1] GOLDWERGER, J., Dynamic Programming for a Stochastic Markovian Process with an Application to the Mean Variance Models, *Management Science* 23 (1977), 612-620.

[2] MILLER, B. L., On “Dynamic Programming for a Stochastic Markovian Process with an Application to the Mean Variance Models” bz J. Goldwerger, *Management Science* 24 (1978), 1779.

[3] BELLMAN, R., Dynamic Programming, *Princeton University Press*, Princeton, New Jersey, 1957.

[4] TOBIN, J., Liquidity Preference as Behavior toward Risk, *Review of Economic Studies*, 1958.

[5] MARKOWITZ, H., Portfolio Selection – Efficient Diversification of Investments, *Wiley*, New York, 1959.

Mgr. Milan Sitar

Department of Probability and Statistics

Faculty of Mathematics and Physics, Charles University

Sokolovská 83, 186 00 Prague 8

Suboptimal and Pareto Optimal Solutions for Variance Penalized Markov Decision Chains [1]

Karel Sladký

Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic

Milan Sitar

Department of Probability and Statistics, Faculty of Mathematics and Physics, Charles University

1. Introduction

Mean variance selection rules were proposed for the portfolio selection problem in [5] and [11] (see also [6]). Following the mean variance selection rule, the investor selects from among a given set of investment alternatives only investments with a higher mean and lower variance than a member of the given set.

In this note we investigate how the mean variance selection rule can work in the Markovian decision models. In particular, we adapt notions and notations used in Markov decision processes (cf. e.g. [8], [9]) and in contrast to the classical models we assume that instead of maximizing the long run average expected return (i.e. the mean reward per transition) we consider more sophisticated criteria taking into account also the variance of the total (long run) reward (details can be found in Section 3). In the present article we focus our attention on finding suboptimal and Pareto optimal solution to the variance penalized Markovian decision processes.

Some research in the direction has been reported in the literature. In particular, Goldwerger [1] suggests a dynamic programming method for maximizing the ratio of the mean reward to the standard deviation (i.e. the square root of average variance) in undiscounted Markov decision chains. Unfortunately, this reseach was not successful, see e.g. [7] presenting a counterexample to the proposed algorithm. In Sobel [12] the problem of maximization of the mean to the standard deviation ratio is analyzed for undiscounted unichained model using the methods of non-linear and parametric linear programming. Detailed analysis of variance penalized Markov decision processes

[1] This research was supported by the Grant Agency of the Czech Republic under Grants 402/99/1136 and 402/98/0742.

(including also undiscounted multichain and discounted cases) by methods of mathematical programming was performed in [2] and [3].

2. Markov decision chains with first and second moment average optimality

Let us consider a nonhomogeneous Markov chain $X = \{X_n, n = 0,1,\ldots\}$ with finite state space $S = \{1,2,\ldots,N\}$. If after the n-th transition the chain is observed to be in state $i \in S$ then a decision k in a finite set D_i must be selected. Selecting decision k in state i state j is reached in the next transition with probability p_{ij}^k and an immediate reward r_{ij}^k is earned. Moreover, we introduce the expected reward (i.e. the first moment of the random reward) earned in state $i \in S$ if decision $k \in D_i$ is selected $r_i^k = \sum_{j=1}^N p_{ij}^k r_{ij}^k$ along with the corresponding second moment $s_i^k = \sum_{j=1}^N p_{ij}^k \left(r_{ij}^k\right)^2$ implying that the expected *one-stage myopic* reward variance of the random reward earned in state $i \in S$ if decision $k \in D_i$ is selected is equal to

$$\left(s_{\,i}^{\,k}\right)^2 = s_i^k - \left(r_i^k\right)^2 = \sum_{j=1}^N p_{ij}^k \left(r_{ij}^k\right)^2 - \left(\sum_{j=1}^N p_{ij}^k r_{ij}^k\right)^2. \qquad (2.1)$$

Such a Markov chain is called a *controlled Markov chain* or *Markov decision chain*. The data p_{ij}^k, r_{ij}^k are assumed to be known to the decision maker. Let $D = D_1 \times D_2 \times \cdots \times D_N$.

Policy Π controlling the chain is a rule prescribing the decision to be taken after each transition in any state of the chain. We restrict on Markov (memoryless) policies, i.e. decision rules taking into account only the number of transitions n and the current state $X_n \in S$ of the chain. A policy which takes at all times the same decision rule is called stationary. We write $\Pi = \left(p^0, p^1, p^2, \ldots\right)$ where $p^n \in D$ for every $n = 0,1,2,\ldots$ and $p_i^n \in D_i$ is the decision at the n-th transition when the chain is in state i. Stationary policy Π is identified by $\Pi \sim (p)$.

Let $P\left(p^n\right)$ be the $N \times N$ matrix whose ij-th element equals $p_{ij}^{p_i^n}$ and let $P^m(\Pi) = \prod_{n=0}^{m-1} P\left(p^n\right)$ (for convenience we set $P^0(\Pi) = I$, the identity matrix). Obviously, $P^{n+1}(\Pi) = P^n(\Pi) P\left(p^n\right)$. Similarly, $r\left(p^n\right)$, resp. $s\left(p^n\right)$, denotes $N \times 1$ vector whose i-th element equals $r_i^{p_i^n}$, resp. $s_i^{p_i^n}$. If $\Pi \sim (p)$ (i.e. if Π is stationary) then $P^m(\Pi) = \left(P(p)\right)^m$. Recall that the limit $P^*(p) = \lim_{m\to\infty} m^{-1}\sum_{n=0}^{m-1}(P(p))^n$ exists; if $P(p)$ is unichained (i.e. $P(p)$ contains single class of recurrent states) the rows of $P^*(p)$ are identical.

In what follows we denote by $v^m(\Pi)$ the vector of total expected rewards (i.e. the sum of expected rewards) earned in the m next transitions. In particular, $v^m(\Pi)$ is the $N\times 1$ vector whose i-th element, denoted $v_i^m(\Pi)$, is the total expected reward earned provided the chain starts in state $i\in S$ and policy Π is followed. Obviously

$$v^m(\Pi) = \sum_{n=0}^{m-1} P^n(\Pi)\, r(p^n) \quad \text{with} \quad \lim_{m\to\infty} m^{-1}v^m(\Pi) = g(\Pi) \tag{2.2}$$

provided the limit exists (obviously, the i-th element of $g(\Pi)$, denoted $g_i(\Pi)$, is the long run average expected reward, or the mean reward, if the Markov chain starts in state i).

Observe that for stationary policy $\Pi \sim (p)$ we get $v^m(\Pi) = \sum_{n=0}^{m-1}(P(p))^n r(p)$ and for m tending to infinity we have

$$g(\Pi) = \lim_{m\to\infty} m^{-1}\sum_{n=0}^{m-1}(P(p))^n\, r(p) = P^*(p)r(p). \tag{2.3}$$

Moreover, if $P(p)$ is unichained, the rows of $P^*(p)$ are identical and $g(p)$ is a constant vector (i.e., the i-th element of $g(\Pi)$, denoted $g_i(\Pi)$, equals some constant number).

Policy $\hat{\Pi}^{(1)}$ is called (first moment) average optimal if

$$\liminf_{m\to\infty}\; m^{-1}v^m\left(\hat{\Pi}^{(1)}\right) \ge \liminf_{m\to\infty}\; m^{-1}v^m(\Pi) \quad \text{for every policy } \Pi. \tag{2.4}$$

Similarly, let $u^m(\Pi)$ be the $N\times 1$ vector whose i-th element is the sum of expected second moments of s_i^k obtained in each of the m next transitions, provided the chain starts in state $i\in S$ and policy Π is followed. Obviously

$$u^m(\Pi) = \sum_{n=0}^{m-1} P^n(\Pi) s(p^n) \quad \text{with} \quad \lim_{m\to\infty} m^{-1}u^m(\Pi) = g^{(2)}(\Pi) \tag{2.5}$$

provided the limit exists (again, the i-th element of $g^{(2)}(\Pi)$ is denoted by $g_i^{(2)}(\Pi)$).

We say that policy $\hat{\Pi}^{(2)}$ is second moment average minimal if

$$\liminf_{m\to\infty} m^{-1}u^m\left(\hat{\Pi}^{(2)}\right) \le \liminf_{m\to\infty} m^{-1}u^m(\Pi) \quad \text{for every policy } \Pi. \tag{2.6}$$

3. Markov decision chains with mean and square mean variance optimality

In what follows we make

Assumption GA. $P(p)$ has single class of recurrent states and $r(p) > 0$ for any $p \in D$.

Supposing that $g(\Pi) = \lim_{m\to\infty} m^{-1} v^m(\Pi) > 0$ exists, instead of one-stage myopic reward variance in state i given by (3.3), we shall consider the expected one-stage reward variance in state i (if decision k according to policy $\Pi = (p^n)$ is selected in state i) as

$$\sum_{j\in S} p_{ij}^k \left[r_{ij}^k - g_i(\Pi)\right]^2 = \sum_{j\in S} p_{ij}^k \left(r_{ij}^k\right)^2 - 2g_i(\Pi)\sum_{j\in S} p_{ij}^k r_{ij}^k + \left[g_i(\Pi)\right]^2 \tag{3.1}$$

implying that the vector of one-stage expected reward variances at the n-th stage is equal to $s(p^n) - 2g_i(\Pi)r(p^n) + g_i(\Pi)g(\Pi)$ (observe that $g_i(\Pi)g(\Pi)$ is an $N\times 1$ constant vector). Then (following policy $\Pi = (p^n)$) the vector of total expected variances of rewards earned in the m next transition is given by

$$d^m(\Pi) = u^m(\Pi) - 2g_i(\Pi)v^m(\Pi) + mg_i(\Pi)g(\Pi). \tag{3.2}$$

Letting $m \to \infty$ for the average expected reward variance we get

$$d(\Pi) = \lim_{m\to\infty} m^{-1}u^m(\Pi) - g_i(\Pi)g(\Pi) = g^{(2)}(\Pi) - g_i(\Pi)g(\Pi). \tag{3.3}$$

Observe that for the i-th element of (3.3) we have

$$d_i(\Pi) = \lim_{m\to\infty} m^{-1}u_i^m(\Pi) - \left[g_i(\Pi)\right]^2 = g_i^{(2)}(\Pi) - \left[g_i(\Pi)\right]^2 \tag{3.4}$$

(recall that under Assumption GA $g^{(2)}(\Pi), g(\Pi)$ are constant vectors).

In what follows instead of average expected reward or average expected reward variance we shall consider either

i) The ratio of the long run average expected reward variance to the long run average expected reward (the mean reward), called the *mean variance optimality*, or

ii) The ratio of the long run average expected reward variance to the square of the mean reward, called the *square mean variance optimality*.

i) *Mean variance optimality*

The goal is to minimize the ratio of the long run average expected reward variance to the mean reward. In particular, we are interested in

$$\liminf_{m\to\infty} \frac{d_i(\Pi)}{g_i(\Pi)} = \liminf_{m\to\infty} \frac{m^{-1}u_i^m(\Pi) - [g_i(\Pi)]^2}{m^{-1}v_i^m(\Pi)} = \liminf_{m\to\infty} \frac{u_i^m(\Pi) - m[g_i(\Pi)]^2}{v_i^m(\Pi)}. \tag{3.5}$$

Policy $\hat{\Pi}$ is called *mean variance optimal* if for every policy $\Pi = (p^n)$ and every $i \in S$

$$\liminf_{m\to\infty} \frac{u_i^m(\Pi) - m[g_i(\Pi)]^2}{v_i^m(\Pi)} \geq \liminf_{m\to\infty} \frac{u_i^m(\hat{\Pi}) - m[g_i(\hat{\Pi})]^2}{v_i^m(\hat{\Pi})}. \tag{3.6}$$

Supposing that the following limits exist $\lim_{m\to\infty} m^{-1} v_i^m(\Pi) = g_i$, $\lim_{m\to\infty} m^{-1} u_i^m(\Pi) = g_i^{(2)}(\Pi)$, condition (3.6) is fulfilled if and only if for every policy $\Pi = (p^n)$ and every $i \in S$

$$\liminf_{m\to\infty} \left\{\frac{u_i^m(\Pi)}{v_i^m(\Pi)} - g_i(\Pi)\right\} \geq \liminf_{m\to\infty} \left\{\frac{u_i^m(\hat{\Pi})}{v_i^m(\hat{\Pi})} - g_i(\hat{\Pi})\right\} \tag{3.7}$$

or equivalently if

$$\frac{g_i^{(2)}(\Pi)}{g_i(\Pi)} - g_i(\Pi) \geq \frac{g_i^{(2)}(\hat{\Pi})}{g_i(\hat{\Pi})} - g_i(\hat{\Pi}). \tag{3.8}$$

ii) *Square mean variance optimality*

The goal is to minimize the ratio of the long run average expected reward variance to the square of the mean reward. In particular, we are interested in the ratio

$$\lim_{m\to\infty} \frac{d_i(\Pi)}{[g_i(\Pi)]^2} = \lim_{m\to\infty} \frac{m^{-1}u_i^m(\Pi) - [g_i(\Pi)]^2}{[m^{-1}v_i^m(\Pi)]^2} = \frac{g_i^{(2)}(\Pi)}{[g_i(\Pi)]^2} - 1. \tag{3.9}$$

Policy Π^* is called *square mean variance optimal* if for every policy $\Pi = (p^n)$ and every $i \in S$

$$\liminf_{m\to\infty} \frac{u_i^m(\Pi) - m[g_i(\Pi)]^2}{m^{-1}[v_i^m(\Pi)]^2} \geq \liminf_{m\to\infty} \frac{u_i^m(\Pi^*) - m[g_i(\Pi^*)]^2}{m^{-1}[v_i^m(\Pi^*)]^2} \tag{3.10}$$

or equivalently if

$$\frac{g_i^{(2)}(\Pi)}{[g_i(\Pi)]^2} \geq \frac{g_i^{(2)}(\Pi^*)}{[g_i(\Pi^*)]^2}. \tag{3.11}$$

4. Suboptimal and Pareto Optimal Solutions for Mean Variance and Square Mean Variance Models

The results summarized in the following two propositions are well known to the workers in discrete dynamic programming. The proofs and details concerning the remaining results of this section can be found in [10].

Proposition 4.1. There exists $\hat{p}^{(1)} \in D$, resp. $\hat{p}^{(2)} \in D$, unique $N \times 1$ constant vector, $\hat{g}^{(1)} = g(\hat{p}^{(1)})$, resp. $\hat{g}^{(2)} = g^{(2)}(\hat{p}^{(2)})$ and an $N \times 1$ vector $\hat{w}^{(1)}$, resp. $\hat{w}^{(2)}$, (unique up to an additive constant) such that for every $p \in D$

$$\hat{g}^{(1)} + \hat{w}^{(1)} = r(\hat{p}^{(1)}) + P(\hat{p}^{(1)})\hat{w}^{(1)} \geq r(p) + P(p)\hat{w}^{(1)}, \tag{4.1}$$

$$\hat{g}^{(2)} + \hat{w}^{(2)} = s(\hat{p}^{(2)}) + P(\hat{p}^{(2)})\hat{w}^{(2)} \leq s(p) + P(p)\hat{w}^{(2)}. \tag{4.2}$$

Stationary policy $\hat{\Pi}^{(1)} \sim (\hat{p}^{(1)})$ is (first moment) average optimal (cf. (2.4)), i.e. $g_i(\hat{p}^{(1)}) = \max_{p \in D} g_i(p)$; stationary policy $\hat{\Pi}^{(2)} \sim (\hat{p}^{(2)})$ is second moment average minimal, i.e. $g_i^{(2)}(\hat{p}^{(2)}) = \min_{p \in D} g_i^{(2)}(p)$, cf. (2.6).

Proposition 4.2. There exists $\hat{p}^{(0)} \in D$, unique $N \times 1$ constant vector $\hat{g}$ and an $N \times 1$ vector $\hat{w}$ (unique up to an additive constant) such that for every $p \in D$

$$\hat{w} = s(\hat{p}^{(0)}) - r(\hat{p}^{(0)})\hat{g} + P(\hat{p}^{(0)})\hat{w} \leq s(p) - r(p)\hat{g} + P(p)\hat{w}. \tag{4.3}$$

Stationary policy $\hat{\Pi}^{(0)} \sim (\hat{p}^{(0)})$ minimizes the ratio of the second to the first moments of average rewards, i.e. $\dfrac{g_i^{(2)}(\hat{p}^{(0)})}{g_i(\hat{p}^{(0)})} = \min\limits_{p \in D} \dfrac{g_i^{(2)}(p)}{g_i(p)}$.

From Proposition 4.1 and 4.2 we immediately get

Proposition 4.3. Mean variance, resp. square mean variance, optimal policy (cf. (3.8), resp. (3.11)) is bounded from below by

$$\frac{g_i^{(2)}(\hat{p}^{(0)})}{g_i(\hat{p}^{(0)})} - g_i(p^{(1)}), \quad \text{resp.} \quad \frac{g_i^{(2)}(\hat{p}^{(0)})}{g_i(\hat{p}^{(0)})g_i(p^{(1)})}. \tag{4.4}$$

To analyze properties of the variance penalized optimality conditions we present the following useful result.

Theorem 4.4. Let (stationary) policy $\hat{\Pi}^{(2)} \sim (\hat{p}^{(2)})$ be second moment average minimal, see (2.6), and (stationary) policy $\hat{\Pi}^{(1)} \sim (\hat{p}^{(1)})$ be (first moment) average optimal, see (2.4). Then there exist stationary policies

$$\hat{\Pi}^{(1)} = \Pi^{(0)} \sim (p^{(0)}), \quad \Pi^{(1)} \sim (p^{(1)}), \quad \Pi^{(2)} \sim (p^{(2)}), \dots, \hat{\Pi}^{(2)} = \Pi^{(r)} \sim (p^{(r)}) \tag{4.5}$$

such that $k = 0,1,\dots,r-1$

$$p_i^{(k)} \neq p_i^{(k+1)} \text{ for one } i_k = i \in S, \quad p_i^{(k)} = p_i^{(k+1)} \text{ for every } i \neq i_k, \tag{4.6}$$

that are Pareto optimal, i.e. there exists no policy $\Pi \sim (p)$ such that

$$\lim_{m\to\infty} m^{-1} u_i^m(\Pi) = g_i^{(2)}(\Pi) < \lim_{m\to\infty} m^{-1} u_i^m(\Pi^{(k)}) = g_i^{(2)}(\Pi^{(k)})$$

$$\text{and simultaneously } \lim_{m\to\infty} m^{-1} v_i^m(\Pi) = g_i(\Pi) > m^{-1} v_i^m(\Pi^{(k)}) = g_i(\Pi^{(k)}) \tag{4.7}$$

for at least one $k = 0,1,\dots,r-1$ and some $i \in S$.

Moreover, every randomized policy arising by randomization of the two subsequent policies $\Pi^{(k)}$ and $\Pi^{(k+1)}$ (for $k = 0,1,\dots,r-1$) is also Pareto optimal.

Furthermore, we can obtain similar results closely connected with mean variance optimal policies. Recall that by Propositions 4.2 and 4.1 (stationary) policy $\hat{\Pi}^{(0)} \sim (\hat{p}^{(0)})$ minimizes $\lim_{m\to\infty} \frac{u_i^m(\Pi)}{n_i^m(\Pi)} = \frac{g_i^{(2)}(\Pi)}{g_i(\Pi)}$ (over all policies $\Pi \sim (p)$) and (stationary) policy $\hat{\Pi}^{(1)} \sim (\hat{p}^{(1)})$ maximizes average reward $g(\Pi)$. In particular, Theorem 4.4 still holds if we replace (4.5) by

$$\hat{\Pi}^{(0)} = \Pi^{(0)} \sim (p^{(0)}), \quad \Pi^{(1)} \sim (p^{(1)}), \quad \Pi^{(2)} \sim (p^{(2)}), \dots, \hat{\Pi}^{(1)} = \Pi^{(r)} \sim (p^{(r)}) \tag{4.8}$$

and (4.7) by ... there exists no policy $\Pi \sim (p)$ such that

$$\lim_{m\to\infty} \frac{u_i^m(\Pi)}{v_i^m(\Pi)} = \frac{g_i^{(2)}(\Pi)}{g_i(\Pi)} > \lim_{m\to\infty} \frac{u_i^m(\Pi^{(k)})}{v_i^m(\Pi^{(k)})} = \frac{g_i^{(2)}(\Pi^{(k)})}{g_i(\Pi^{(k)})} \text{ and simultaneously } g_i(\Pi) > g_i(\Pi^{(k)}).$$

References

[1] J. Goldwerger: Dynamic programming for a stochastic Markovian process with an application to the mean variance models. Manag. Sci. *23* (1977), 612-620.

[2] J. Filar, L.C.M. Kallenberg and H.-M. Lee: Variance penalized Markov decision processes. Mathem. Oper. Research *14* (1989), 147-161.

[3] Ying Huang and L.C.M. Kallenberg: On finding optimal policies for Markov decision chains: a unifying framework for mean-variance-tradeoffs. Mathem. Oper. Research *19* (1994), 434-448.

[4] H. Kawai: A variance minimization problem for a Markov decision process. European J. Oper. Research *31* (1987), 140-145.

[5] H. Markowitz: Portfolio selection. J. Finance *7* (1952), 77-91.

[6] H. Markowitz: Portfolio Selection – Efficient Diversification of Investments. Wiley, New York 1975.

[7] B.L. Miller: On „Dynamic programming for a stochastic Markovian process with an application to the mean variance models“ by J. Goldwerger. Manag. Sci. *24* (1978), 1779.

[8] M.L. Puterman: Markov Decision Processes - Discrete Stochastic Dynamic Programming. Wiley, New York 1994.

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

[10] K. Sladký and M. Sitar: Mean Variance Models in Markovian Decision Processes: Optimality Conditions and Algorithms. Research Report No. 1982, Institute of Information Theory and Automation AS CR, Prague 2000.

[11] J. Tobin: Liquidity preferences as behavior toward risk. Review of Economic Studies *25* (1958), 65-86.

[12] M.J. Sobel: Maximal mean/standard deviation ratio in an undiscounted MDP. Oper. Research Lett. *4* (1985), 157-159.

[13] D.J. White: Mean, variance and probability criteria in finite Markov decision processes: A review. J. Optim. Theory Appl. *56* (1988), 1-29.

Ing. Karel Sladký, CSc.

Department of Econometrics

Institute of Information Theory and Automation

Academy of Sciences of the Czech Republic

Pod vodárenskou veží 4, 182 08 Praha 8

Mgr. Milan Sitar

Department of Probability and Statistics

Faculty of Mathematics and Physics, Charles University

Sokolovská 83, 186 00 Praha 8

Quantitative Methods in Competition Analysis

Ing. Dagmar Straková, Protimonopolný úrad SR

From economic standpoint competition concerns market structure and subsequent behaviour of buyers and sellers in a market. It is a way of production organisation and setting of prices. Economic role of competition is to encourage producers and owners of resources to offer their goods and services in a cheap way and skilfully done with the aim of consumer welfare maximalisation that means maximalisation of economy effectiveness.

However, the aim of firms themselves is to gain market power, i.e. intentional control over prices and other relevant factors e.g. quantity, discounts, that are decisive for business transactions. This market power can be obtained through competition restriction and that is a market imperfection which leads to ineffective allocation of sources and subsequent impair of industry output and economic welfare. Market imperfections allow sellers to decrease output of production deliberately with increasing prices to the detriment of consumers and society as a whole.

Quantitative methods are the instrument which helps to define the structure and behaviour of industry. Empirical analysis of economic issues becomes increasingly a basic characteristic of antitrust proceedings abroad especially when dealing with concentrations, abuse of a dominant position and agreements restricting competition.

The USA has got a long tradition of economic analysis stemming from the seventies when economists at the Department of Justice have got a greater influence as well as from the nature of the USA antitrust policy which requires lots of backing and facts finding evidence.

European antitrust law has started to pay attention to empirical evidence of economists only recently. European Commission Notice the Definition of the Relevant Market for the Purposes of Community Competition Law, 1997, defences quantitative techniques explicitly: "There are a number of quantitative tests that have specifically been designed for the purpose of delineating markets. These tests consist of various econometric and statistical approaches: estimates of elasticity and cross-price elasticity for the demand of a product, tests based on similarity of price movements over time, the analysis of causality between price series and similarity of price levels and/or their convergence. The Commission takes into account the available quantitative evidence capable of withstanding rigorous scrutiny for the purposes of establishing patterns of substitution in the past"[1].

Econometric methods are suitable for rich data antitrust cases. An analyst has to know industry and behaviour of main actors in the market. He must appraise credibility and suitability of data before implementing empirical test [2]. Suitable method depends on quality of data, time and sources which are attainable. Results of econometric analysis should be compared with sources of evidence like documents, consumer reviews and marketing studies. If economic issues and a hypothesis are formulated clearly, if a model is well-defined and collected data are credible, then econometric methods are able to give the right answer and they represent a backing instrument when assessing the competition issue.

Analysis of concentrations, agreements restricting competition and abuse of a dominant position have the similar analytical proceeding consisting of the following steps:

1. identification of participants of proceedings, preliminary analysis of their activities – the subject of entrepreneurial activity, property and personnel linkage, definition of relevant jurisdiction,
2. definition of influenced markets from the product, geography and time point of view, assessment of position of firms in the relevant markets,
3. assessment of impacts on competition in the relevant markets in connection with behaviour restricting competition or proposed concentration
4. assessment of possible overall economic advantages.

Quantification can be involved in each of these steps whereby in every country quantitative techniques can differ. However, they are the only instrument for assessment of structures and behaviour of industry. Quantitative techniques are used especially for

- Market definition
- Market structure analysis
- Empirical analysis of competitive behaviour of firms, e.g. predatory pricing
- Assessment of effects of a concentration.

Market definition is the most important task when analysing economic competition. Statement on competition implications depends on size and form of relevant market. To define a relevant market means to characterise circumstances for market power performance. Key element is extent of supply and demand substitution. Primary means for a product market delineation is the evaluation of demand substitution – the extent to which buyers would easy substitute among alternative products or sources of supplies is characterised. When evaluating reactions of consumers one can stem from price and cross-price elasticity of demand. Interchangeability of supply, i.e. reactions of customers represent the secondary mean for definition of market boundaries. Level of substitutability of products, necessary for location of products in the same market, is high. Market has to be well-defined because a narrow definition of market takes into account insubstantial issues and too wide definition can conceal the very competition problems.

Very well known test is the test of hypothetical monopolist or cartel, called the SSNIP test in the USA (Small but Significant, Non-transitory Increase in Price). The test identifies the smallest number of products or producers, where the hypothetical monopolist or cartel which controls supply of all the relevant products, could increase profits by instituting a small but appreciable permanent increase in price over the competitive level. Also European Commission notice on market definition [1] makes reference to this approach as well as the U.K. new Competition Law [3].

Price trends analysis, demand test and system of demand equations are the quantitative tests used for market definition. In general, tests based on price trends should be used carefully as they do not enable to evaluate whether prices could be increased by market participants in profiting way. However, for the reason of not having data other models are not used and therefore markets are defined on the grounds of price tests [4].

<u>Degree of market concentration</u> was considered one of the main structural characteristics for a long time and market structure analysis has become the key indicator of actual and potential market power.

Traditional competition analysis stems from Bain's paradigm on relationship among market structure, behaviour and market performance. Under this theory a market structure (number of sellers, feasibility of entry) predicts to the great extent the degree of market participants behaviour (pricing policy, advertisement) and a market performance in the form of effectiveness, technical progress is the result of the market behaviour .

Nowadays, there does not exist unambiguous opinion on the paradigm. Under some economists there does not exist evidence proving hypothesis that structure predetermines performance. Industries are very different and at the same time so complex that a simple generalisation is not possible. Therefore, the most preferred approach nowadays is case-by-case approach, i.e. individual assessment of the every case [5].

Concentration indices

Index used within econometric analysis of competition is called *a concentration index*. This index sets up a rate of individual firm to the total production of industry or within defined group of firms.

$$a_{ii} = \frac{q_i}{Q} \qquad \text{where } i = 1, ..., n \text{ and } \sum_{i=1}^{n} a_i = 1 \qquad [6]$$

Concentration rate CR_m points on the position of the m highest shares of firms

$$CR_m = \sum_{i=1}^{m} a_i \qquad \text{ordering the firms so that } a_1 \geq .... \geq a_m \geq ... \geq a_n \qquad [6]$$

CR_m is defined as a percentage rate of total industry sales /or capacity, employment, physical output/ of the largest companies set up under market shares. In the USA CR_m is measured for the first four firms CR4 whereby production industries are published for 8, 20 50 firms [7].

The most popular instrument is the *Herfindhal- Hirschman Index* (HHI) which is equal to the sum of the squares of the market shares

$$R_m = \sum_{i=1}^{m} a_i^{\ 2} \qquad [7]$$

Under the USA guidelines unconcentrated market has the HHI less than 1000, a moderately concentrated market has the HHI between 1000 and 1800 and the market is concentrated when HHI is higher than 10 000. Concentration which leads to the increase of HHI less than 100 points and HHI is between 1000 and 1800 is not usually investigated. If HHI is above 1800 and an increase in the HHI is more than 50 points then a concentration could invoke serious competition harm.

Even if HHI seems to be a rational way of concentration measurement none of theories and econometric evidences have proven that HHI is a sufficient instrument for definition of concentration effects [2]. Every serious case of concentration has to point not only on concentration degree under HHI but especially on concentration effects on behaviour of market participants, e.g. probability of a collusion.

The question which of the mentioned indexes is the best one has not been answered yet. In every case it is necessary to proceed in empirical way and characterise the most suitable index for the particular behavioural relations.

When assessing impacts on competition stemming from behaviour of a firm or proposed concentration prices are analysed. Even a simple price analysis can give enough information as under Hayek‘s statement all relevant economic information are involved in data on price of product [9]. Matters are analysed in many ways:

- Time series of prices are evaluated in conditions of structural changes
- Actual price trends can be compared with that what a competition model in the absence of an event would predict (for example bidding models) [4].

Assessment of overall economic advantages in connection with behaviour restricting competition is also called defence of efficiency. Economists define three types of efficiency which are problematic when doing empirical verification:

Allocative efficiency means that prices reflect costs in such a way that firms produce relatively more of what people want and are willing to pay for. The result is that sources within economy are allocated in the way that the output which is the most appreciated by customers is produced. Co-operation among independent firms in the market or concentrations which decrease number of firms in the market, do not usually support allocative efficiency.

Technical efficiency means that under given output production uses the most effective combination of inputs without breaks. Healthy competition supports technical efficiency. If concentration increases beyond a certain level, technical efficiency is reduced, there is a fall in productivity.

Dynamic efficiency means that there is an optimal trade-off between existing consumption and investment in innovation and technological progress. Competitive pressures in the market support innovation. Firm in the front of technological progress has a chance to acquire market power.

The role of economic competition is to improve allocative efficiency without restricting the remaining two efficiencies.

Antimonopoly Office of the Slovak Republic is responsible for the supervision of competition rules in our country. Activity of the Office is focused primarily on observation of the Act on Protection of the Economic Competition as amended [9]. Quantitative analysis in competition area has not become yet the instrument regularly used when assessing market structures. In sporadic way Herfindhal index HHI was used as a backing argument when assessing degree of concentration in the relevant market under a concentration. During 1996 a methodology for industry quantitative analysis was tested in co-operation with University of Economics and it was confirmed that it is necessary to provide single data collection and to work out a methodological apparatus in the next future. After the validity of the just prepared new competition law harmonised with the EU legislation the Office is to pay increased attention to creation of methodology for quantitative analysis of competitive environment, degree of industry concentration, market structure analysis and utilisation of econometric and optimalisation progress.

References

[1]COMMISSION NOTICE on the definition of the relevant market for the purposes of Community competition law, OJ C 372 on 9/12/1997

[2] FISHER F.M. *Industrial Organization, Economics and the Law*. Massachusetts Institute of Technology 1991, p. 446 - 469

[3] UK Competition Act 1998

[4] LECG LTD *Quantitative techniques in competition analysis*. London, Office of Fair Trading Research Paper 17, 1999,.

[5] W.Kip Viscusi, John M. Vernon, Joseph E. Harrington, Jr. - Economics of Regulation and Antitrust, 1998

[6] TIROLE J. *The Theory of Industrial Organization*. Massachusetts Institute of Technolo gy, 9th printing 1997,

[7] SCHERER F.M., ROSS D. *Industrial Market Structure and Economic Perforrmance*. Boston, 3rd ed. Houghton Mifflin Company 1990,

[8] HAYEK F. *The Used of Knowledge in Society,* American Economic Review 35, 1945

[9] Act No. 188/1994 Coll. of Laws on Protection of the Economic Competition as amended

Ing. Dagmar Straková

Protimonopolný úrad SR

Drienová 24

826 03 Bratislava

GFDS Knowledge Base Operations and Their Applications in Economics

Michal Holcapek, Matej Turcan
On behalf of Institutional Research of Department of Mathematical Methods in Economics,
Faculty of Economics, VŠB-TU Ostrava

Abstract: In this paper we present operations defined on Knowledge Bases of General Fuzzy Decision Systems (GFDS). It is shown how it works on two Knowledge Bases of GFDS for a bank, that is checking the credit solvency of an applicant for the credit.

Keywords: General Fuzzy Decision System (GFDS), Operations on Knowledge Base

1 Introduction

In practice we usually need to express an expert knowledge of more than one expert. Expert rules are usually formulated by linguistic terms which is why it is difficult to transform them into classical mathematical terms or to apply them to computer-aid processing. For example, a bank want to work with an expert system consisted of expert rules (with a knowledge base) which would make a decision about the credits according to the conditions of this bank. It's clear, that the expert decision is a subjective. Therefore the bank takes various suggestions of knowledge base and these one combine by some operation to only one knowledge base of an expert system of this bank. Various approaches can be seen in [2] and [4]. For our considerations we'll use the fuzzy decision systems described in [3] on the base of the model of GFDS, created by Jirí. Mockor (see [1]). We would like to describe the knowledge base, which is one of the elements of decision system, that works better than original ones and define a general operation in the set of all knowledge bases. It's well known, that the fuzzy set theory offers adequate instruments for dealing with uncertainty.
By modelling linguistic variables in form of fuzzy sets, it is possible to transform expert rules into mathematical terms. Moreover the fuzzy sets offers a great collection of operators which are able to aggregate and combine these rules as well as the knowledge bases..
Throughout this paper **U** denotes a universe (a discrete set of variants), $\boldsymbol{F}(\mathbf{U})$ the fuzzy power set on **U**, **C** a set of criterions, $(\mathbf{U},\mathbf{C})$ the decision space. For our purpose this space will be constant and we will be working only inside this space. Further $\boldsymbol{P}_{\mathbf{U}} \subseteq \bigcup \boldsymbol{F}(\mathbf{U})^k$, where $1 \le k \le n$, is an abstract space over the universe **U** and $\boldsymbol{P}_{\mathbf{U}}^{n} \subseteq \boldsymbol{F}(\mathbf{U})^{n}$ is n-dimensional abstract space over the universe **U** created by n-dimensional vectors with fuzzy sets as their components. $F_{\mathbf{U}}^{n}$ is the set of all abstract spaces $\boldsymbol{P}_{\mathbf{U}}^{n}$ and $F_{\mathbf{U}}$ is the set of all possible abstract spaces $\boldsymbol{P}_{\mathbf{U}}$. Vectors of fuzzy sets, that have for all elements of the universe the same value of membership function equal 1 or 0 are special elements of an abstract space $\vec{1} = (1,...,1)$ or $\vec{0} = (0,...,0)$. For the universe [0,1] we use the above symbols without the lower index, i.e. $\boldsymbol{P}^{n}$, F^{n}, F.

2 General Fuzzy Decision System (GFDS)

We begin with definition of GFDS, what introduces a general notion.

Definition 2.1
Let ($\boldsymbol{U}$,**C**) *be a decision space. Then a general fuzzy decision system over* (**U**,**C**) *is a set* $R_g = (\{G_g \mid g\in \mathbf{K}\},\{w_c \mid c\in \mathbf{C}\},\{\boldsymbol{V}_{g,\mathrm{C}} \mid g\in \mathbf{K}, \mathrm{C}\in \mathbf{C} \}, s, f)$,
where
1. **K** *is a finite set of goals,*
2. $G_g \,\tilde{\subset}\, [0,1]$ *is a fuzzy set defined for all* $g\in \mathbf{K}$,
3. for every $c\in\mathbf{C}$ *there is defined a fuzzy set* $w_c : \mathbf{U} \rightarrow \boldsymbol{F}([0,1])$,
4. $\boldsymbol{V}_{g,\mathrm{C}}$ *is a knowledge base over the set of criterions* $\mathrm{C} \subset \mathbf{C}$, *determined by the goal* $g\in\mathbf{K}$,
5. s is a fuzzy inclusion relation,
6. $f : [0,1]^{|\mathrm{C}|} \rightarrow [0,1]$ *is an aggregation function.*

Below we state short intuitive interpretation of this abstract definition. For more information see [3].

1. A finite set **K** is a set of goals we want to deal with.
2. A level of satisfaction of any goal $g\in\mathbf{K}$ could be prescribed and is dependent on the goal's importance. A fuzzy set G_g then represents a required level of satisfaction of a goal $g\in\mathbf{K}$.
3. A function $w_c : \mathbf{U} \rightarrow \boldsymbol{F}([0,1])$, then describes a level of satisfaction of a criterion $c\in\mathbf{C}$ by a variant u from **U**. The more fuzzy set w_c is similar to the crisp set $\{1\} \subset [0,1]$, the better a variant $u\in\mathbf{U}$ satisfies a criterion $c\in\mathbf{C}$. For the goal $g\in \mathbf{K}$ and the set of criterions $\mathrm{C} = \{ c_1, c_2,\ldots, c_n \} \subset \mathbf{C}$ we can define a set of vectors $\boldsymbol{W}_{g,\mathrm{C}} = \{(w_{c_1}(u), w_{c_2}(u),\ldots, w_{c_n}(u)) \mid u\in\mathbf{U} \}$.
4. A set $\boldsymbol{V}_{g,\mathrm{C}} = \{ \vec{V}_i \mid i = 1, 2,\ldots, m \}$, where $\vec{V}_i = (V_{i,1}, V_{i,2},\ldots, V_{i,n}, V_{i,n+1})\in\boldsymbol{P}^{n+1}$, is the base of expert knowledge. Any element $\vec{V}_i$ represents, in fact, some expert knowledge about the influences of levels of satisfaction of criterions c_j onto level of satisfaction of a goal $g\in\mathbf{K}$. In this notation, an expert (for example) states that if a level of a satisfaction of a criterion c_j can be expressed as a fuzzy set , $V_{i,j} \,\tilde{\subset}\, [0,1]$, then a level of satisfaction of a goal g can be expressed by a fuzzy set $V_{i,n+1} \,\tilde{\subset}\, [0,1]$. The more fuzzy sets $V_{i,j}$ and $V_{i,n+1}$ are similar to the crisp set $\{1\} \subset [0,1]$, the higher levels of satisfaction of criterions or goal are required.
5. The fuzzy inclusion relation s could be defined as follows:
$$s(A,B) = 1 - \frac{1}{|A|} \sum_{\substack{u\in\mathbf{U}\\ A(u)>B(u)}} (A(u) - B(u)), \text{ where } |A| = \sum_{u\in Supp(A)} A(u)$$
6. An aggregation function could be for example defined as: $f(a_1, \ldots , a_n) = min(a_1, \ldots , a_n)$.

The relation s and function f enable us to calculate a total utility function, for example in the following form

$$h_{R,g}(u) = \sup_{\substack{\vec{V}\in\boldsymbol{V}_{g,C}\\ \vec{V}=(V_1,\ldots,V_n,V_{n+1})}} \left[f(s(w_{c_1}(u),V_1),\ldots,s(w_{c_n}(u),V_n)) \circ s(V_{n+1},G_g) \right],$$

where "$\circ$" represents operation "$\wedge$" or "$\cdot$", for any variant $u\in\mathbf{U}$. The higher the value of $h_{R,g}(u)$ is, the better the variant u is from the point of view of R_g.

3 Knowledge Base and Knowledge Bases Operations

The base of expert knowledge is the heart of our general fuzzy decision system Besides the description of this knowledge base in words in form of rules IF-THEN, we need its mathematical description, that is equivalent and enables us to handle with it (evaluate, aggregate, connect and so on). For this reason we use an abstract space described in [3]. We begin with definition of a knowledge base that introduces a general notion.

Definition 3.1

Let $\mathrm{C} = \{ c_1, c_2, \ldots, c_n \}$ *be a set of criterions,* $g \in \mathbf{K}$ *be the goal and* F *be the set of all abstract spaces over* $[0,1]$. *Then the triple* $\boldsymbol{V}_{g,\mathrm{C}} = (Ant, Suc, \rho)$, *where*

1. $Ant = \boldsymbol{P}^n \in \mathrm{F}^n \subset \mathrm{F}$
2. $Suc = \boldsymbol{P}^1 \in \mathrm{F}^1 \subset \mathrm{F}$
3. $\rho \subseteq Ant \times Suc$ *is a relation,*

is called the knowledge base over the couple (C, g).

It's clear, that the knowledge base could be defined as an abstract space $\boldsymbol{P}^{n+1}$ from the set F^{n+1}, where first n components of vectors from $\boldsymbol{P}^{n+1}$ create an antecedent and the last one creates a succedent.

Definition 3.2

Let $\vec{A}, \vec{B} \in \boldsymbol{P}_U$, *where* $\vec{A} = (A_1, \ldots, A_k)$ *and* $\vec{B} = (B_1, \ldots, B_l)$. *Then the binary operation "•" such that* $\vec{A} \bullet \vec{B} \in (A_1, \ldots, A_k, B_1, \ldots, B_l)$ *is called the connection of the vectors of the abstract space.*

According to the definition 3.2 we can write $\boldsymbol{P}^{n+1} = \{ \vec{A} \bullet \vec{S} \mid \forall \vec{A} \in Ant \ \forall \vec{S} \in Suc : (\vec{A}, \vec{S}) \in \rho \}$. To define operations between two knowledge bases we have to require bases from abstract spaces of the same dimension. We consider two knowledge bases $\boldsymbol{V}_{g,C_1} = (Ant_1, Suc_1, \rho_1)$, and $\boldsymbol{V}_{g,C_2} = (Ant_2, Suc_2, \rho_2)$, where C_1 and C_2 are two corresponding sets of criterions. We denote $C = C_S = C_1 \cup C_2$ and $C_P = C_1 \cap C_2$. We´ll solve two cases:

1. C_P is an empty set, then we can define for all $\vec{A} \in Ant_1 : \vec{A}^* = \vec{A} \bullet \vec{1}$, where $\vec{1}$ has $|C_2|$ components. We denote the set of all vectors $\vec{A}^*$ by Ant_1^*, and $\boldsymbol{V}_{g,C}^{*1} = (Ant_1^*, Suc_1, \rho_1^*)$, where $\rho_1^*(\vec{A}^*) = \rho_1(\vec{A})$. By analogy for all $\vec{B} \in Ant_2 : \vec{B}^* = \vec{1} \bullet \vec{B}$, where $\vec{1}$ has $|C_1|$ components. We denote the set of all vectors $\vec{B}^*$ by Ant_2^*, and $\boldsymbol{V}_{g,C}^{*2} = (Ant_2^*, Suc_2, \rho_2^*)$, where $\rho_2^*(\vec{B}^*) = \rho_2(\vec{B})$. Both bases $\boldsymbol{V}_{g,C}^{*1}$ and $\boldsymbol{V}_{g,C}^{*2}$ are evidently from F^{n+1}, where $n = |C| = |C_1| + |C_2|$
2. C_P is a nonempty set and $|C_P| = m$, then we can suppose that all levels of satisfaction of all criterions $c \in C_P$ are described by the first m components of all vectors from sets Ant_1 and Ant_2. Each vector of these sets could be described as a connection of two vectors. First vector contains the common components of both antecedents and the second one contains the rest of components. It means that for all $\vec{A} \in Ant_1 : \vec{A} = \vec{A}_P \bullet \vec{A}_Z$ and for all $\vec{B} \in Ant_2 : \vec{B} = \vec{B}_P \bullet \vec{B}_Z$. Now we can define for all $\vec{A} \in Ant_1 : \vec{A}^* = \vec{A}_P \bullet \vec{A}_Z \bullet \vec{1}$, where

$\vec{1}$ has $|C_2|$ - m components. By analogy for all $\vec{B} \in Ant_2 : \vec{B}^* = \vec{B}_P \bullet \vec{1} \bullet \vec{B}_Z$, where $\vec{1}$ has $|C_1|$ - m components.

If we denote the set of all vectors $\vec{A}^*$ as Ant_1^* and the set of all vectors $\vec{B}^*$ as Ant_2^*, we can create two knowledge bases $\boldsymbol{V}^{*\,1}_{g,C}$ =(Ant_1^*, Suc_1, ρ_1^*), where $\rho_1^*(\vec{A}^*) = \rho_1(\vec{A})$, and $\boldsymbol{V}^{*\,2}_{g,C}$ = (Ant_2^*, Suc_2, ρ_2^*), where $\rho_2^*(\vec{B}^*) = \rho_2(\vec{B})$, from $\mathrm{F}^{\,n+1}$, where n = $|C_1|+|C_2|$-m.

As a knowledge base is a subset of an abstract space, we'll define operations between two knowledge bases generally as operations between two abstract spaces of the same dimension. We consider the relation E $\subseteq \bigcup_{\boldsymbol{P}_U \in \Phi_U} \boldsymbol{P}_U \times \bigcup_{\boldsymbol{P}_U \in \Phi_U} \boldsymbol{P}_U$. For each ordered couple of abstract spaces $(\boldsymbol{P}_U, \boldsymbol{Q}_U) \in \mathbf{F}_U \times \mathbf{F}_U$ we define a relation that is the restriction of relation E on the Cartesian product $\boldsymbol{P}_U \times \boldsymbol{Q}_U$, i.e. E($\boldsymbol{P}_U$, $\boldsymbol{Q}_U$)=En ($\boldsymbol{P}_U \times \boldsymbol{Q}_U$). Let´s remark, that this relation could be empty.

We'll understand operations between abstract spaces as operations between vectors of these spaces, i.e. operations between fuzzy sets, that form components of vectors. In many cases there is profitable to make operation only between some of vectors. This fact can be described by relation E. Let's now define an operation on the set of abstract spaces of the same dimension.

Definition 3.3

Let $\mathrm{F}^{\,n}_{U}$ *be a set of abstract spaces of the same dimension* n *over* **U** *and* E *be the above mentioned nonempty relation between abstract spaces* $\boldsymbol{P}^{n}_{U}$, $\boldsymbol{Q}^{n}_{U} \in \mathrm{F}^{\,n}_{U}$.

A fuzzy operation "$\circ_E$" *is a mapping* $\circ_E : \mathrm{F}^{\,n}_{U} \times \mathrm{F}^{\,n}_{U} \to \mathrm{F}^{\,n}_{U}$, *such that*

$$\boldsymbol{P}^{n}_{U} \circ_E \boldsymbol{Q}^{n}_{U} = \{ \vec{A} \circ_{E(\boldsymbol{P}^{n}_{U}, \boldsymbol{Q}^{n}_{U})} \vec{B} \mid \forall \vec{A} \in \boldsymbol{P}^{n}_{U} \ \forall \vec{B} \in \boldsymbol{Q}^{n}_{U} \},$$

where
$$\vec{A} \circ_{E(\boldsymbol{P}^{n}_{U}, \boldsymbol{Q}^{n}_{U})} \vec{B} \begin{cases} \vec{0} & \Leftrightarrow \ (\vec{A}, \vec{B}) \notin E(\boldsymbol{P}^{n}_{U}, \boldsymbol{Q}^{n}_{U}), \\ (A_1 \circ B_1, \ldots, A_n \circ B_n) & \Leftrightarrow \ (\vec{A}, \vec{B}) \in E(\boldsymbol{P}^{n}_{U}, \boldsymbol{Q}^{n}_{U}), \end{cases}$$

for $\boldsymbol{P}^{n}_{U}$, $\boldsymbol{Q}^{n}_{U} \in \mathrm{F}^{\,n}_{U}$.

The operation "$\circ$" *was specified by the index* E, *as this operation depends on the relation* E, *that determines the couples of abstract spaces and also the couples of vectors, between which the operation is defined.*

Now we can define the operation between two knowledge bases in the following way.

$$\boldsymbol{V}_{g,C_1} \circ_E \boldsymbol{V}_{g,C_2} = \boldsymbol{V}^{*\,1}_{g,C} \circ_E \boldsymbol{V}^{*\,2}_{g,C},$$

where the type of an operation depends on the experience of an expert. It could be for example $\cap, \cup, \ldots$

It is clear, that operations between knowledge bases could be extended onto operations between two decision systems. But it is not the topic of this article. For more information see [3].

4 Example

Imagine a bank wanting to find a decision system, especially its knowledge base, that is able to decide if the financial standing of credit applicants (firms) is good enough to give them a credit The system should estimate whether each firm is enable to repay it in time. Suppose that the bank wants to decide using an information about at least two of the following most important criterions:

- c_1 – Optimal Debt Ratio (indicates the covering of client's assets by external financial sources),
- c_2 – Optimal Profit to Owner's Net Worth Ratio (indicates the return of own sources),
- c_3 – Optimal Total Liquidity Ratio (indicates the client's handy solvency).

Assume that the bank disposes of two experts and each of them offers a proposal of his own base of rules for the same goal g - to give a credit that will be repaid in time. Bank wants to find the new knowledge base $\boldsymbol{V}_{g,C}$, that represents (joins together) an expert knowledge of both experts.

Denote their knowledge bases as $\boldsymbol{V}_{g,C_1}$ and $\boldsymbol{V}_{g,C_2}$, where $C_1 = \{ c_1, c_2 \}$ and $C_2 = \{ c_1, c_2, c_3 \}$.

The following tables show expert rules of both knowledge bases:

$\boldsymbol{V}_{g,C_1}$

c_1	c_2	g
S_1	S_1	S_1
S_1	M_1	S_1
S_1	B_1	M_1
M_1	S_1	M_1
M_1	M_1	M_1
M_1	B_1	B_1
B_1	S_1	M_1
B_1	M_1	B_1
B_1	B_1	B_1

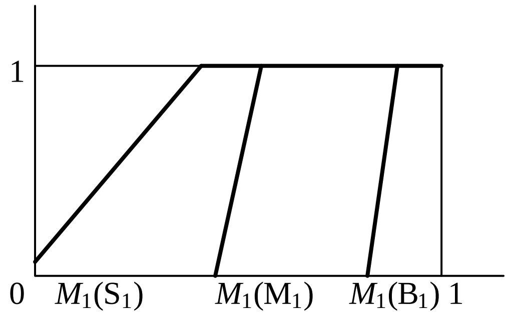
$\boldsymbol{V}_{g,C_2}$

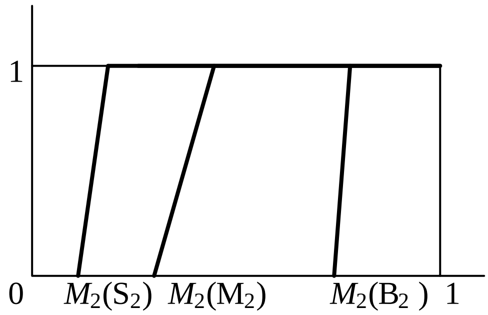
c_1	c_2	c_3	g
S_2	S_2	S_2	S_2
S_2	S_2	M_2	S_2
S_2	S_2	B_2	S_2
S_2	M_2	S_2	S_2
S_2	M_2	M_2	M_2
S_2	M_2	B_2	M_2
S_2	B_2	S_2	M_2
S_2	B_2	M_2	M_2
S_2	B_2	B_2	M_2
M_2	S_2	S_2	S_2
M_2	S_2	M_2	S_2
M_2	S_2	B_2	M_2
M_2	M_2	S_2	M_2
M_2	M_2	M_2	M_2
M_2	M_2	B_2	M_2
M_2	B_2	S_2	M_2
M_2	B_2	M_2	B_2
M_2	B_2	B_2	B_2
B_2	S_2	S_2	M_2
B_2	S_2	M_2	M_2
B_2	S_2	B_2	M_2
B_2	M_2	S_2	M_2
B_2	M_2	M_2	B_2
B_2	M_2	B_2	B_2
B_2	B_2	S_2	M_2
B_2	B_2	M_2	B_2
B_2	B_2	B_2	B_2

Expert rules are described by fuzzy linguistic variables: $L_1=\{X_1, T(X_1), [0,1], G, M_1\}$ for the first base and $L_2=\{X_2, T(X_2), [0,1], G, M_2\}$ for the second one, where the variable X_i, $i = 1,2$, is always the same for all criterions and the goal g. The corresponding sets of terms are $T(X_1)=\{S_1,M_1,B_1\}$ and $T(X_2)=\{S_2,M_2,B_2\}$, where:

S_i means *small,*
M_i means *medium,*
B_i means *big*

level of satisfaction of the criterion or of the goal. Evaluations of these terms $M_i(S_i)$, $M_i(M_i)$, $M_i(B_i)$, where i = 1,2, are represented by the fuzzy sets on a universe [0,1]. See the following two pictures.

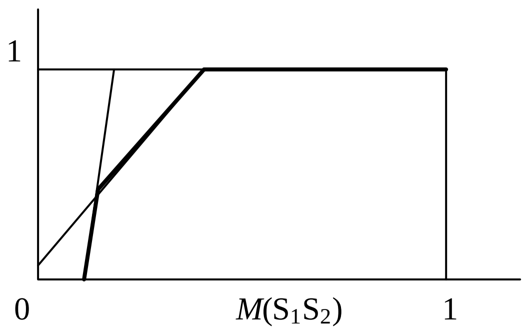




The task is to find a new expert knowledge base $\boldsymbol{V}_{g,C}$. It can be derived from primary bases $\boldsymbol{V}_{g,C_1}$ and $\boldsymbol{V}_{g,C_2}$ as, for example, $\boldsymbol{V}_{g,C_1} \cap_E \boldsymbol{V}_{g,C_2} = \boldsymbol{V}^{*\,1}_{g,C} \cap_E \boldsymbol{V}^{*\,2}_{g,C}$, where operation $\cap$ is taken as fuzzy intersection of two fuzzy sets, i.e. $(A \cap B)(u) = \min(A(u), B(u))$, where $A, B \in \boldsymbol{F}(\mathbf{U})$. We'll get a new linguistic variable $L=\{Y, T(Y), [0,1], G, M\}$, where Y is the same for criterions and the goal g. All possible values of $T(Y)$ are shown in the next table, where the intersection of two terms could be described by: term_1 & term_2, for example small_1 & small_2, where "&" could be taken as "and" in the natural language.

&	S_2	M_2	B_2
S_1	S_1S_2	S_1M_2	S_1B_2
M_1	M_1S_2	M_1M_2	M_1B_2
B_1	B_1S_2	B_1M_2	B_1B_2

Evaluations of these terms could be represented by the fuzzy sets, derived as intersections of fuzzy sets, that evaluate terms of L_1 and L_2. The next picture shows the evaluation of S_1S_2 as $M(S_1S_2) = M_1(S_1) \cap M_2(S_2)$.



Operation $\cap_E$ means, that the operation intersection of knowledge bases is restricted by relation E. In our case, we'll make operations only between such expert rules, that are similar, i.e. the terms for corresponding criterions are adjoining in an ordered sequence S_i, M_i, B_i. To simplify the table of relation E, we use symbol $\vec{V}_i$ for the expert rule from the base $\boldsymbol{V}_{g,C_1}$ and symbol $\vec{V}_{ij} = \{\vec{V}_{i1}, \vec{V}_{i2}, \vec{V}_{i3}\}$, that represents three expert rules from the base $\boldsymbol{V}_{g,C_2}$. These rules have the same terms for the first two criterions as the rule $\vec{V}_i$ and the term for the third criterion is then S_2 or M_2 or B_2. The relation E is described in the next table, where 1 or 0 mean that the corresponding couple is or is not in E, respectively.

E	$\vec{V}_{1j}$	$\vec{V}_{2j}$	$\vec{V}_{3j}$	$\vec{V}_{4j}$	$\vec{V}_{5j}$	$\vec{V}_{6j}$	$\vec{V}_{7j}$	$\vec{V}_{8j}$	$\vec{V}_{9j}$
$\vec{V}_1$	1	1	0	1	0	0	0	0	0
$\vec{V}_2$	1	1	1	0	1	0	0	0	0
$\vec{V}_3$	0	1	1	0	0	1	0	0	0
$\vec{V}_4$	1	0	0	1	1	0	1	0	0
$\vec{V}_5$	0	1	0	1	1	1	0	1	0
$\vec{V}_6$	0	0	1	0	1	1	0	0	1
$\vec{V}_7$	0	0	0	1	0	0	1	1	0
$\vec{V}_8$	0	0	0	0	1	0	1	1	1
$\vec{V}_9$	0	0	0	0	0	1	0	1	1

Now could be created the knowledge base $\boldsymbol{V}_{g,C}$. This base will have at most 31(i.e. number of 1 in the above table) $\times$ 3 = 93 expert rules. In fact some of them should be reduced. As there is not place enough here, just a small part of the table of this knowledge base could be shown.

c_1	c_2	c_3	g
S_1S_2	S_1S_2	S_2	S_1S_2
S_1S_2	S_1S_2	M_2	S_1S_2
S_1S_2	S_1S_2	B_2	S_1S_2
S_1S_2	S_1M_2	S_2	S_1S_2
S_1S_2	S_1M_2	M_2	S_1M_2
S_1S_2	S_1M_2	B_2	S_1M_2
S_1S_2	M_1S_2	S_2	S_1S_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
B_1B_2	B_1B_2	B_2	B_1B_2

5 Conclusion

In this article we defined GFDS as a mathematical model of decision system. We have introduced a knowledge base as a set of ordered n-tuples of fuzzy sets (or as a subset of an abstract space). We also showed very simple principle, which enables us to do some operations between knowledge bases of two different decision systems. If we deal with operations, connected with relation E, that are associative (see [3]), we could apply this principle to more than two knowledge bases. If an operation, connected with relation E, is not associative, then we could obtain different knowledge bases for different sequences of decision systems.

There is a possibility to generalize an operation between abstract spaces, connected with 2-dimensional relation E, onto an operation among abstract spaces, connected with n-dimensional relation E. But this problem we haven't solved yet.

References

[1] Mockor, J.: *General Fuzzy Decision Systems*, Acta Mathematica et Informatica Universitatis Ostraviensis 5, 1997

[2] Rommelfanger, H.: *Fuzzy Decision Support Systems*, Springer Verlag, Berlin, Heidelberg, 1994

[3] Turcan.M.: *Teorie fuzzy rozhodovacích systému a její aplikace v rozhodování banky.* Disertacní práce, Ekf VSB-TU Ostrava, 2000

[4] Zimmermann, H.J.: *Fuzzy Sets, Decision-Making and Expert Systems*, Kluver Academic. Publ., Dordrecht, 1987

Michal Holcapek, Matej Turcan

On behalf of Institutional Research of Department of Mathematical Methods in Economics, Faculty of Economics, VŠB-TU Ostrava

E-mail: michal.holcapek@vsb.cz, matej.turcan@vsb.cz

Banks Concentration and Efficiency

Ladislav Uncovský
University of Economics, Bratislava.

The problem of concentration and its relation to efficiency is often studied in industrial economy and organisation. Such investigation, concerning manufacturing industry in Slovakia , is published in Uncovský L.-Brezina I (12).

The problem of concentration and efficiency of the commercial banking sector in Slovakia. is issue of the paper (13, 14). The motivation of such an inquiry is connected with the fact, that this sector consists of relatively small number of firms and in (8,9,10) there are published data showing their results for 1996, 1997 and 1998.

In case of commercial banks there is a problem of indices, used for expressing concentration. For manufacturing industries the volume of sales is used. For banks three volumes are relevant for measuring performance, the volume of loans, the volume of assets and the volume of deposits. As single indicator, the position of banks (deposits - loans) may be used, but there are arguments against use of this value. .\

The concentration of Slovak commercial banks studied J. Makúch and R. Preisinger in (5). Their results are interesting, as they cover the period of 1993 - 1996. They used rate of concentration of 5 largest banks (CR5) and rate of the one largest bank (CR1). The results are following: (Table 1)

Table 1.

		1993	1994	1995	1996
CR1	Loans	40,83	36,89	35,97	31,97
	Deposits	53,72	48,13	44,99	39,26
CR5	Loans	92,87	84,83	79,05	68,7
	Deposits	92,17	86,21	80,61	72,49

Own computations show these values of concentration rates:

Table 2.

		1996	1997	1998
CR1	Deposits	0,363581	0,337597	0,3075
	Loans	0,3466	0,319136	0,3143
	Assets	0,26450	0,24036	0,21832
CR5	Deposits	0,75338	0,724699	0,70159
	Loans	0,7664358	0,723375	0,68352
	Assets	0,70244	0,65701	0,60280
HHI	Deposits	0,20946	0,18941	0,16194
	Loans	0,183905	0,15323	0,14992
	Assets	0,15545	0,13091	0,11021

Sources: The journal TREND and data of the National Bank of Slovakia.

From data, published in the Central European Economic Review indices, concerning concentration of commercial banks in Slovakia, Czechia, Hungaria, Poland, Romania and Russia, CR5 and HHI for 1996 and 1997 were calculated.

Table 3. CR5 for 1996 - 1998.

	1996			1997			1998		
	Deposits	Loans	Assets	Deposits	Loans	Assets	Deposits	Loans	Assets
SK	0,75240	0,76644	0,72044	0,73442	0,71720	0,65701	0,70159	0,68352	0,60280
CZ	0,82907	0,83409	0,78826	n.a.	0,80447	0,74329	n.a.	0,86474	0,82000
HU	n.a.	n.a.	0,56615	n.a.	n.a.	0,56986	n.a.	0,57851	0,60013
PL	0,58689	0,46870	0,52784	n.a.	0,44288	0,49483	n.a.	0,50157	0,54514
RUS	0,34082	0,47073	0,45327	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
RO	0,78961	0,93332	0,77797	n.a.	n.a.	0,77642	n.a.	0,88734	0,85943

Source: Central European Economic Review,October 1997,1998,1999

Table 4. CR1 for 1996, 1997, 1998

	1996			1997			1998		
	Deposits	Loans	Assets	Deposits	Loans	Assets	Deposits	Loans	Assets
SK	0,36291	0,34660	0,26450	0,34322	0,31641	0,24036	0,30752	0,31429	0,21832
CZ	0,29318	0,30842	0,26817	n.a.	0,27766	0,25456	n.a.	0,28951	0,28472
HU	n.a.	n.a.	0,27419	n.a.	n.a.	0,26821	n.a.	0,19589	0,27882
PL	0,27243	0,15549	0,19619	n.a.	0,14809	0,19619	n.a.	0,18218	0,19452
RUS	n.a.	n.a.	n.a.	n.a,	n.a.	n.a.	n,a.	n.a,	n.a.
RO	0,25533.	0,34725	0,27408	n.a.	n,a,	0,25490	n.a.	0,31711	0,33838

Source: Central European Economic Review,October 1997,1998,1999

Table 5. HHI for 1996, 1997, 1998

	1996			1997			1998		
	Deposits	Loans	Assets	Deposits	Loans	Assets	Deposits	Loans	Assets
SK	0,20871	0,17982	0,15545	0,18941	0,15323	0,13091	0,16199	0,14992	0,11021
CZ	0,18753	0,17909	0,15725	n.a.	0,15751	0,14527	n.a.	0,18748	0,17619
HU	n,a.	n.a.	0,10741	n.a.	n.a.	0,10682	n.a.	0,09138	0,11748
PL	0,11646	0,06442	0,08076	n.a.	0,05943	0,07562	n.a.	0,07883	0,09407
RUS	0,03754	0,0746	0,08083	n.a.	0,0782	0,02935	n.a.	n.a.	n.a.
RO	0,14744	0,24157	0,15573	n.a.	n.a.	0,14945	n.a.	0,21504	0,21221

Source: Central European Economic Review,October 1997,1998,1999

In (4) the European Central Bank investigated concentration of commercial banks in EMU countries. The following three tables show the results:

Table 6. Deposits of 5 largest banks in %.

	1980	1985	1990	1995	1996	1997
SE	n.a.	57.94	61.36	84.31	81.77	86.9
NL	n.a.	85.00	79.50	81.90	81.30	84.20
GR	89.93	89.24	87.67	82.95	82.08	79.57
PT	62.00	64.00	62.00	76.00	81.00	79.00
DK	72.00	70.00	82.00	76.00	74.00	72.00
FR	n.a.	46.00	58.70	68.10	68.80	68.60
BE	n.a.	62.00	67.00	62.00	61.00	64.00
FI	52.80	54.20	46.08	64.17	62.69	63.12
AT	n.a.	32.01	31.95	36.37	35.77	39.06
ES	37.20	35.10	31.40	39.20	39.78	38.16

IT	n.a.	19.90	18.60	42.10	40.40	36.70
LU	n.a.	n.a.	n.a.	22.48	27.76	28.02
UK	n.a.	n.a.	n.a.	25.00	27.00	26.00
DE	n.a.	n.a.	11.57	12.55	14.02	14.19

Source: Ref. (3).

Table 7. Loans of the 5 largest banks in %

	1980	1985	1990	1995	1996	1997
SE	n.a.	62.65	64.89	90.06	86.45	87.84
NL	n.a.	67.10	76.70	78.50	78.10	80.60
GR	87.79	93.16	89.67	80.75	78.65	79.90
DK	73.00	71.00	82.00	79.00	85.00	75.00
PT	60.00	60.00	57.00	73.00	76.00	75.00
BE	55.00	54.00	58.00	61.00	63.00	66.00
FI	49.85	49.68	49.65	59.93	57.98	56.23
FR	n.a.	48.70	44.70	46.80	48.60	48.30
IE	44.40	47.70	42.90	47.50	46.40	46.80
ES	36.70	35.10	33.40	43.12	42.54	42.13
AT	n.a.	28.87	30.07	34.01	33.38	39.31
LU	n.a.	n.a.	n.a.	15.13	30.06	28.63
UK	n.a.	n.a.	n.a.	25.00	26.00	26.00
IT	n.a.	16.60	15.10	26.30	26.60	25.90
DE	n.a.	n.a.	13.48	13.83	13.26	13.71

Source Ref.(3)

Table 8. Assets of 5 largest banks in %.

	1980.	1985.	1990	1995	1996	1997
SE	n.a.	60.22	70.02	85.85	86.21	89.71
NL	n.a.	69.30	73.40	76.10	75.40	79.40
FI	51.43	51.72	53.48	68.60	73.56	77.77
PT	60.00	61.00	58.00	74.00	80.00	76.00
DK	62.00	61.00	76.00	74.00	78.00	73.00
GR	85.44	82.06	83.32	75.66	71.72	71.05
BE	54.00	48.00	48.00	54.00	55.00	57.00
AT	n.a.	35.88	34.64	39.19	38.96	48.26
ES	40.10	38.10	34.90	45.55	44.35	43.60
IE	59.10	47.50	44.20	44.40	42.20	40.70
FR	n.a.	46.00	42.50	41.30	41.20	40.30
UK	n.a.	n.a.	n.a.	27.00	28.00	28.00
IT	n.a.	20.90	19.10	26.10	25.40	24.60
LU	31.06	26.83	n.a.	21.23	21.81	22.43
DE			13.91	16.67	16.08	16.68

Source: Ref. (3)

As concentration itself is concerned the Slovak banks are highly concentrated. The 5 largest banks share of deposits is, comparing with EMU banks on 7th place, behind Portugal and before Denmark. (in 1997). Similar ist the place as far as loans are concerned. The share of assets is relatively lower, but the 7th place of Slovak banks is due in this case too.

Among the other postcommunist countries, the most concentrated commercial banks are the banks of the Czech republic. They are on the level of EMU banks third place. The banks of Hungary and Poland are less concentrated and their values are on the level of EMU average. Beyond this average is the concentration of Rumanian banks. On the other side the concentration of Russian banks is high, nearly on the level of Czech banks.

In (4) three groups of countries is identified with regard to the status of concentration“

1. countries with high concentration above 70% (SE, NL, FI, PT, DK , GR)
2. countries with medium concentration between 40% and 60% (AT, BE, ES, IE, FR)
3. countries with relatively low concentration bellow 30% (DE, LU, UK, IT0.

From the mentioned CEE countries, SK, CZ and RO pertains to the first group, the other (HU, PL nad RUS) to the second group.

The problem of concentration is investigated first of all as a fact, determining economic results the given branch. For commercial banks not only the overall concentration of the branch is important, but inside of it the size distribution of banks . The issue is to find correlation between performance indicators and profit . Such type of analysis is in Uncovský (13,14).

For analysing relations of concentration and profitability for particular countries, one possible way is to compare ranking of given countries banks according their deposits, loans and assets with ranking according their ROA and ROE indices.This is done in mentioned papers.

The scope of concentrations analysis is to find relationship between concentration and efficiency. As shown in Uncovský,L.- Brezina, I.(12) and Uncovský,L (13) a (14) it is not easy ro find significant correlation between these phenomena. In (4) this problem is not mentioned. There is a notice on p. 20., concerning relationship between profitability and size (in terms of total assets). A positive relationship between them is considered to be„not a general rule“.

Results show very low correlation between ROA and values of concentration. Higher are the values of correlation between ROE and values of concentration. The highest value is the that of ROE and volume of 5 largest banks loans. These results do not allow to state a significant correlation between concentration and efficiency of banks in EMU countries. Similar is the result of Slovak banks results comparison.

References:

1. Central European Economic Review, 8. october 1997
2. Central European Economic Review, 16. october 1998
3. Central European Economic Review, october 1999
4. European Central Bank: Possible Effects of EMU on the EU Banking Systems in the medi um to Long Term, Frankfurt a/M, ;1999
5. Makúch,J.-Preisinger,R.: Bankový sektor v SR, paper, presented on the seminar :Perspectives of banking systems in Central Europe, 1997 (in Slovak)
6. Prádelová, L.: Analýza koncentrácie v bankovníctve, Diploma thesis, Bratislava, 1999
7. Ryšavá S.: Koncentrácia bankovníctva v Slovenskej republike, Diploma thesis, Bratislava, 2000
8. TREND Top 96, Bratislava, 1996
9. TREND Top 97 , Bratislava, 1997
10.TREND, 17.6.1998, Banky
11.TREND Top 99 vo financníctve, Bratislava, 1999
12 Uncovský,L.- Brezina,I.: Concentration of Industry in the Slovak Republic, Ekonomický casopis, 45, 1997, No 19, 758 - 778.
13..Uncovský, L.:Koncentrácia v slovenskom bankovníctve, in Ekonomická transformácia, jej sociálne a ekologické dôsledky III, Zborník, november 1998, p. 242- 248.
14. Uncovský, L.: Concentration and Efficiency of Slovak Banking, in Kvantitatívne metódy v ekonomike, IX., Zborník, 1999.
15. Uncovský, L.:Kapacita a výkony bankovníctva EMU a Slovenska, Biatec, júl 1999, p. 2 - 5.

Ladislav Uncovský, Department of Operations Research and Econometrics,

University of Economics, Dolnozemská 1/b Bratislava, Slovak Republic

Non-Accelerating Inflation Rate of Unemployment:
Estimation for the Czech Republic

Osvald Vašícek, Martin Fukac*

Abstract

Quantification of the Non-Accelerating Inflation Rate of Unemployment (NAIRU) is often discussed topic, mainly in the sense of its implementation into the monetary authority decision making. We believe that the NAIRU estimation for the Czech Republic has not been published yet. Identified model for the short-run NAIRU is based on the new classics theoretical approach. The model is treated within state-space framework allowing both time-varying parameters and unobserved variables to be identified using the Kalman filter (with backward smoothing, alternatively). Adaptive identification method provides more robust results in comparison with Kalman filter employed by Estrella, Mishkin (1998), Gordon (1996) or Staiger, Stock, Watson (1996). The KFS application proved its usefulness in modelling economies in transition. Even if the NAIRU does not represent an operative criterion for the monetary policy, it can be a useful information source for its formation. The Czech NAIRU model is estimated on quarterly data.

Keywords

short-run NAIRU, long-run NAIRU, Phillips curve, unemployment gap, extended Kalman filter with backward smoothing, unobserved states, money illusion

1 Short Run NAIRU Conceptual Model

Prior to the conceptual model description, it is necessary to distinguish short-run NAIRU, long-run NAIRU and natural rate of unemployment. The definition of exogenous shocks affecting the NAIRU and the price level is critical issue for this distinction. The short-run NAIRU is defined as a rate of unemployment consistent with inflation rate stabilised at its current level for the next period (month, quarter, semester, year). Its path is affected by shocks with both short run and long lasting impact. The long-run NAIRU is defined as the equilibrium rate towards which unemployment converges in the absence of temporary supply

* Masaryk University, Faculty of Economics and Administration, Department of Applied Mathematics and Informatics, Lipová 41a, 659 79 Brno, Czech republic. E-mail: osvald@econ.muni.cz, fukac@econ.muni.cz

influences, ones the dynamic adjustment of inflation is completed. We refer to the short-run NAIRU thereinafter.

The conceptual model is based on the new classical economics in what explains fluctuations of unemployment rate around its natural rate, short run NAIRU respectively. A model of money illusion address to explain the fluctuations in terms of agents' surprises (unanticipated innovations). The reason for the mistaken expectation is that they cannot distinguish relative from general price movements (recall Lucas's misperception model). We proceed from following assumption in deriving the theoretical model[1]:

- agents' decision depends on relative prices,
- agents' behavior is rational regarding to a known information set,
- labour suppliers are located in a large number of physically separated competitive markets,
- labour demand is distributed unevenly across markets, so the labour price (wage) varies across the markets,
- labour suppliers are price takers and
- all markets are identical in their behavioral parameters.

We leave the Phillips curve derivation from individual labour demands for a future publication. In this article we only suggest the Phillips curve[2] to be established within the microeconomic framework.

$$p_t - p_t^e = a(u_t - u_t^*) + \text{ß}\mathbf{z}_t + e_t \qquad a < 0 \qquad t = 1,2,...,T \;, \tag{1}$$

where p is an inflation rate, p^e is an expected inflation rate, u is an unemployment rate, u^* is a short-run or long-run NAIRU, $\mathbf{z}$ is a term capturing exogenous shocks affecting aggregate price level (in the sense of Gordon(1996), Estrella and Mishkin(1998)) and e_t is i.i.d. with

$$\mathrm{E}(e_t) = 0 \text{ a } \mathrm{E}(e_t\ e_t^{'}) = \mathbf{I}s^2 \;. \tag{2}$$

[1] see Sargent (1987)
[2] see Gordon (1996)

If $p^e_t < p_t$ or $p^e_t > p_t$ (price level growth is not anticipated in agents' expectations), then the agents see this disproportion as a change of relative prices, as a change in real purchase power respectively. This leads to a change in labour supply and in unemployment. If $p^e_t = p_t$ then the price level change is understood as a change in absolute prices and agents do not react.

NAIRU equation closes the model concept. The NAIRU is determined by lagged unemployment gap and the same exogenous shocks $\mathbf{z}_t$ as the price level, or

$$\tilde{u}_t = f\tilde{u}_{t-1} + d\Delta r_t + ?\mathbf{z}_t + w_t \quad , \qquad (3)$$
$$\tilde{u}_t = u_t - u_t^* \quad ,$$

where $\tilde{u}$ is the unemployment gap, Δr is a change in interest rate (monetary policy tool), $\mathbf{z}$ is the exogenous shocks vector and w is stochastic component with the same properties as e.

Equations (1) and (3) are added assumptions on expectation formation (autoregressive process) and on composition of exogenous effects hitting domestic price level. Then the model becomes

$$\begin{aligned} \Delta p_t &= \mathrm{a}_1\tilde{u}_{t-1} + \mathrm{a}_2\tilde{u}_{t-2} + \mathrm{a}_3\Delta s_{t-1} + \mathrm{a}_4\Delta s_{t-2} + \mathrm{a}_5\Delta o_{t-1} + \mathrm{a}_6\Delta o_{t-2} + e_t \quad , \\ \tilde{u}_t &= \mathrm{b}_1\tilde{u}_{t-1} + \mathrm{b}_2\tilde{u}_{t-2} + \mathrm{b}_3\Delta r_{t-1} + \mathrm{b}_4\Delta r_{t-2} + \mathrm{b}_5\Delta s_{t-1} + \mathrm{b}_6\Delta s_{t-2} + \mathrm{b}_7\Delta o_{t-1} + \mathrm{b}_8\Delta o_{t-2} + w_t \quad , \end{aligned} \qquad (4)$$

where Δp is the first difference of net inflation rate[3], $\tilde{u}$ is the unemployment gap, Δs is the first difference of nominal effective exchange rate logarithm[4], Δo is the first difference of oil price index logarithm, Δr is a change of 3-month interest rate, and e and w are stochastic parts defined above.

As the identification method we use the Kalman filter with backward smoothing arranged to allow simultaneous estimation of both time-varying parameters and unobserved states. Here, state is an unobserved economic variable identified by means of structural assumptions and a set of observed variables (signal extraction problem). We need to transform the model (4) into the state-space representation in order to estimate time varying parameters and the

[3] Price deregulation (under government control) effects are excluded from net inflation.
[4] Nominal effective exchange rate is defined as an 65% CZK/DEM -- 35% CZK/USD basket.

unemployment gap simultaneously. We let parameters to be time varying to detect and capture possible structural breaks.

2 Identification Results

Quarterly time series are employed in the model, see Fig.1 for the data. The results of model (4) identification with time-varying parameters are found in Fig.2 and Fig.4. Time-varying parameters trajectories and their confidence intervals are depicted in Fig.3. Time invariant parameter estimates and their standard errors are reported in Tab.1.

Figure 1: Model inputs

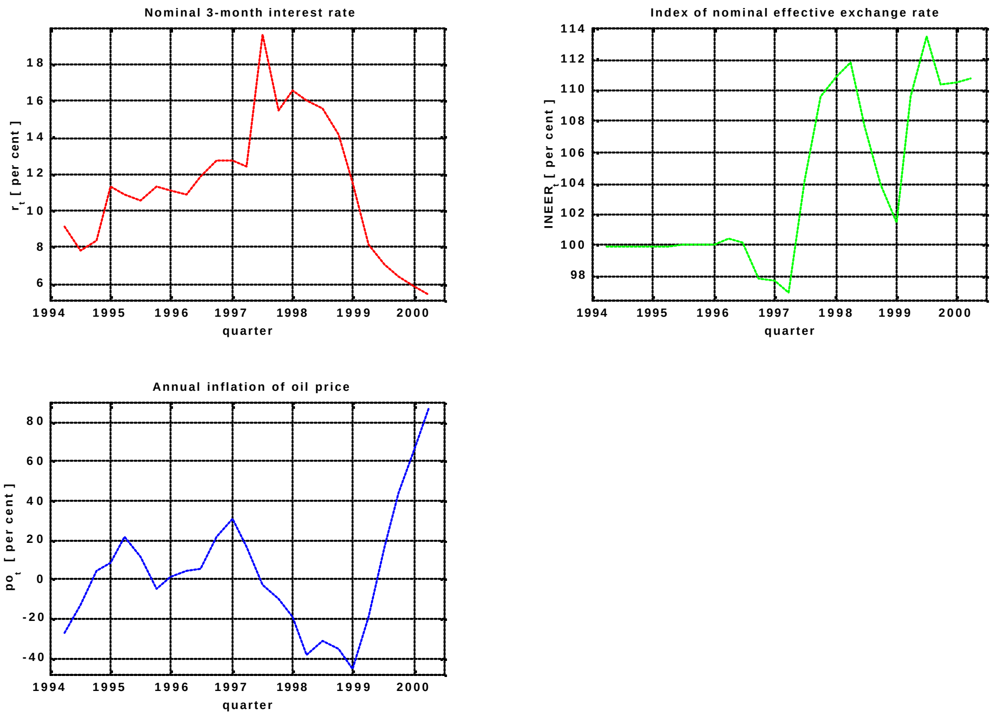


Table 1: Time invariant parameter estimates and their standard deviation

Model equation		Parameter	Parameter estimate (standard error in parenthesis)
Δp_t	$\tilde{u}_{t-1}$	a_1	$-2{,}335\ (0{,}216)$
	$\tilde{u}_{t-2}$	a_2	$0{,}674\ (0{,}169)$
	Δe_{t-1}	a_3	$0{,}040\ (0{,}014)$
	Δe_{t-2}	a_4	$-0{,}085\ (0{,}017)$
	Δo_{t-1}	a_5	$0{,}003\ (0{,}002)$
	Δo_{t-2}	a_6	$0{,}010\ (0{,}003)$
$\tilde{u}_t$	$\tilde{u}_{t-1}$	b_1	$1{,}673\ (0{,}049)$
	$\tilde{u}_{t-2}$	b_2	$-0{,}994\ (0{,}044)$
	Δr_{t-1}	b_3	$-0{,}023\ (0{,}016)$
	Δr_{t-2}	b_4	$-0{,}040\ (0{,}018)$
	Δe_{t-1}	b_5	$-0{,}052\ (0{,}013)$
	Δe_{t-2}	b_6	$0{,}035\ (0{,}010)$
	Δo_{t-1}	b_7	$0{,}006\ (0{,}002)$
	Δo_{t-2}	b_8	$-0{,}008\ (0{,}002)$
Innovation variance			
var(e_t)			0,0234
var(w_t)			0,2514

Figure 2: Estimation of NAIRU model (4)

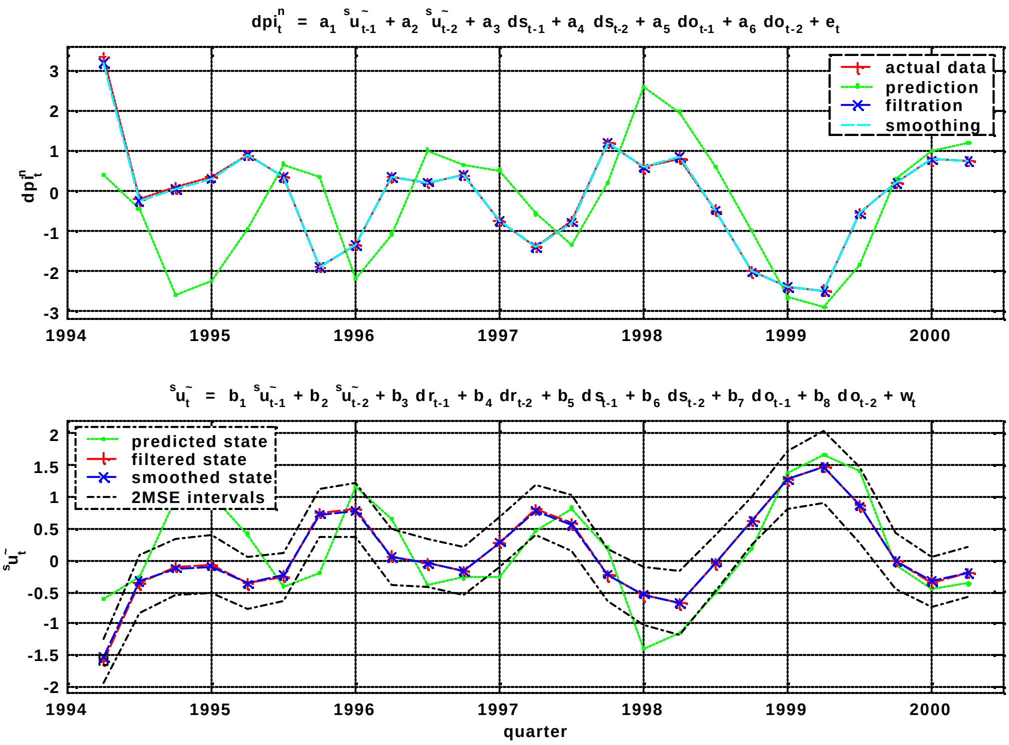


Figure 3: Estimated time varying parameters paths and their confidence intervals

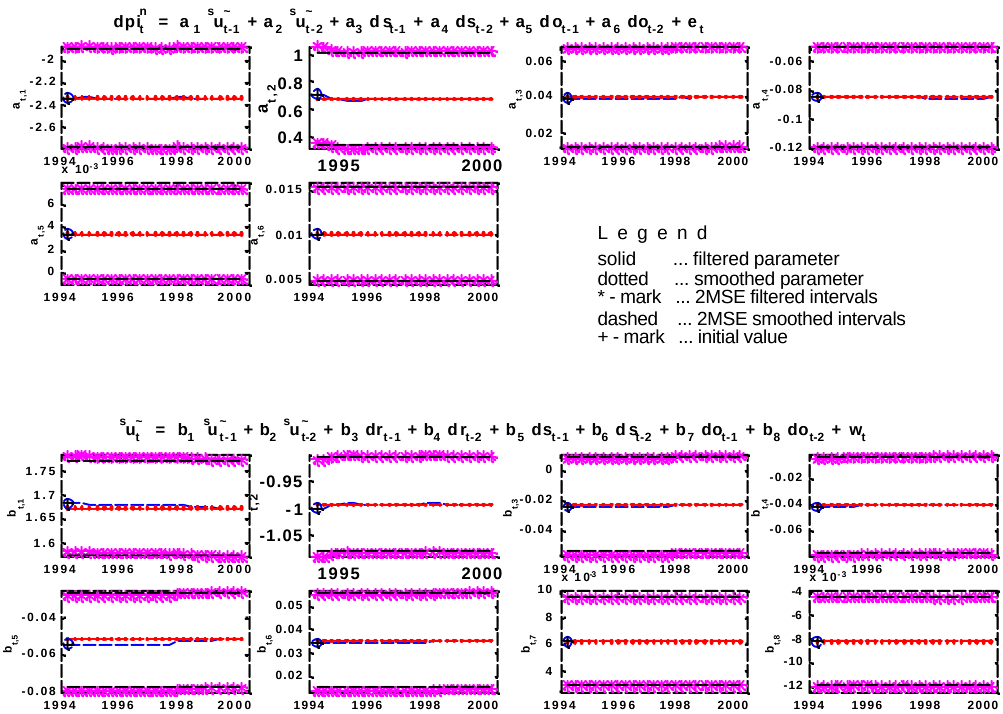
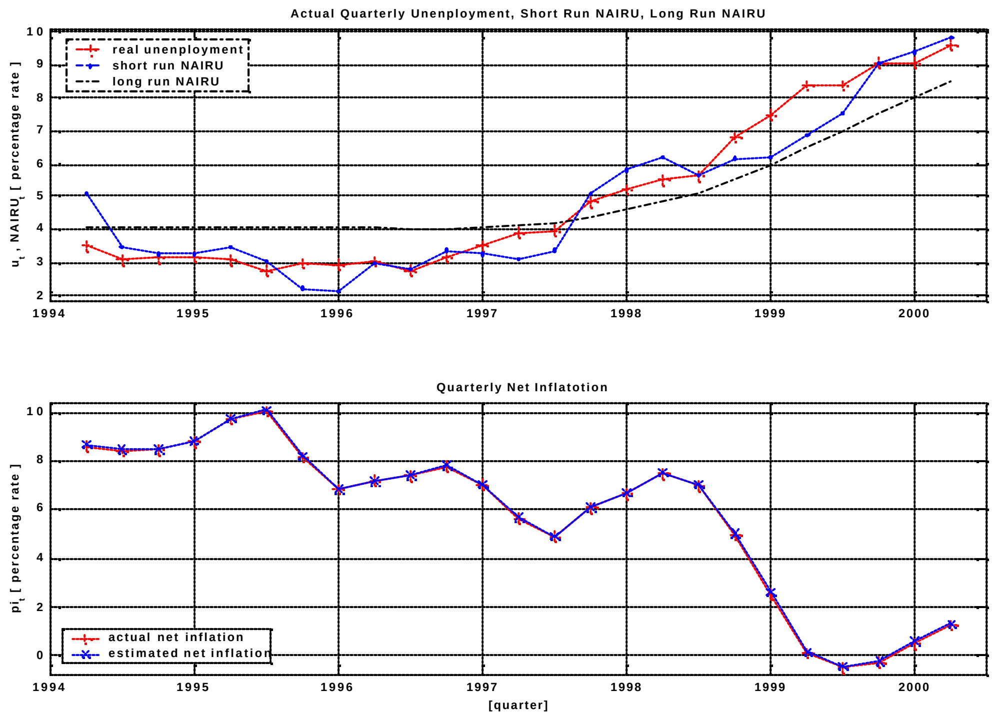


Figure 4: Short-run NAIRU, long-run NAIRU and net inflation estimates



Path of time varying parameters seems to be stable (see Fig.3). Their filtered estimates vary only marginally around the smoothed constant path. In this sense we can describe the model with time invariant parameters without any loss in informational content (see Tab.1).

We can say about the development of unemployment rate and NAIRU in 1994:I.Q-1996: IV.Q (Fig.4) that it is in part approximately steady because the net inflation is "stable" at that time period. First well-marked break in this trend is in 1997 when the unemployment growth begins overtaking the NAIRU growth and this tendency continues till 1999. The higher level of estimated NAIRU in 1997 is caused by growth of 3-month interest rate and by a significant effective exchange rate depreciation. The end of 1999 and beginning of 2000 stands for follow-up in being NAIRU and unemployment rate at the same level. However, we can hardly assess its stability, because even if the interest rate path and effective exchange rate path seems to be stabilized, the oil price shock occurs.

Averaging of unemployment rate on annual intervals approximates long-run NAIRU in Fig.4. Even if we cannot educe strict conclusions about real long-run NAIRU from this, we can use it for hypothesis (1) testing. It is evident from Fig.4 that the relationship between p, u and u^* is in accordance with (1). If $u < u^*$, the inflation rate accelerates (1994-1996). If $u > u^*$, the inflation rate decelerates (1997-2000).

3 Conclusion

This early attempt to estimate the short-run NAIRU for the Czech economy is performed on the basis of a simple two-equation model. The Kalman filter with backward smoothing is applied as identification method. The NAIRU model was estimated with time varying parameters. This is convenient for transitive economy conditions. The NAIRU concept is unusable as an operative criterion for monetary authority decision making, but we believe that it can be an useful information source after augmentation of the basic model and its set in wider macroeconomic framework

References

1. Espinosa Vega, M.A.; Russell, S. (1997) : *History and Theory of the NAIRU: A Critical Review*. Federal Economic Review, Reserve Bank of Atlanta , Second Quater, pp.4-25

2. Estrella, A.; Mishkin, F.S. (1998) : *Rethinking the Role of NAIRU in Monetary Policy: Implications of Model Formulation and Uncertainty.* NBER Working Paper Series, Working Paper No. 6518, Cambridge

3. Gordon, R.J. (1996) : *The Time-Varying NAIRU and Its Implications for Economic Policy*. NBER Working Paper Series, Working Paper No. 5735, Cambridge

4. Hamilton, J.D. (1994) : *Time Series Analysis*. Princeton University Press, Princeton, ISBN 0-691-04289-6

5. Sargent, T. J.(1987) : *Macroeconomic Theory – Second Edition.* Academic Press, London, ISBN 0-12-619751-2

6. Staiger, D.; Stock, J.H.; Watson, M.W. (1996): *How Precise Are Estimates of the Natural Rate of Unemployment*. . NBER Working Paper Series, Working Paper No. 5477, Cambridge

7. Šimandl, M. (1995) : *Identifikace systému a filtrace*. Skripta, Západoceská univerzita fakulta aplikovaných ved, Plzen. ISBN 80-7082-170-1

8. Vašícek, O. (1999): *Adaptive identification, quantitative analysis and forecast of macroeconomic effects of desinflation strategy by dynamic model*. Sborník Quantitative Methods for Economics, Bratislava, pp.245-251

An Uncertainty Principle in Economics[*/]

Miloslav Vošvrda

Institute of Information Theory and Automation

Academy of Sciences of the Czech Republic

Abstract**:** *One of the central tenets of modern financial economics is the necessity of some trade-off between* ***a risk and an expected return****. If a security's expected price change is positive, it is needed* ***a reward*** *to attract investors to hold the capital asset and bear* ***the corresponding risks****. If an investor is* ***sufficiently risk averse****, he might gladly pay to avoid holding* ***a security*** *which has* ***unforecastable returns****. Correct estimation and prediction of the volatility is thus most important for major financial institutes, because volatility is directly related to usual risk measures.*

Keywords: Risk, Expected return, Stochastic Volatility, Uncertainty Principle**,** Cramer-Rao Bounds, Informational Content

JEL Classifications: C1, G14

Introduction

Modern financial economics theory is concentrated on the informational efficiency of the options markets. Capital assets prices, conveying information in the option markets, are depended on a volatility. Long time it is recognized

[*/] The research was supported by GACR 402/00/0439, 402/98/0742 and GAAS K 1075601

on capital markets that capital asset returns exhibit volatility clustering. Correct estimation and prediction of volatility are most important for major financial institutes, because the volatility is directly related to usual risk measures. Since all modern theories of capital asset pricing relate first moment(risk premium) to second moments(measures of risk). Today it is clear that standard ARCH models can only be the starting point of model. Therefore modern theory is concentrated on time series techniques that have been developed for modeling σ_t^2 . The integration of time series techniques for conditional mean into structural econometric model building has led to a much deeper and richer understanding of the underlying dynamics. **The instantaneous variance of return implicit in the price of the call option can be interpreted as an ex ante forecast of the average volatility of the underlying capital asset over the life of the option** (see Merton(1973)). The ability of implied volatility to predict the future volatility of an underlying capital asset is considered as a measure of information content of call prices. The prediction of future market volatility is of interest due to the theoretical relation between the expected market risk premium and the ex ante volatility of the market(Merton(1980)). Recent studies using the GARCH framework find that volatility is persistent over time (Bollerslev, Engle, and Wooldridge (1988)). *This suggests that the GARCH model may be useful in predicting future volatility.* A related model for conditional volatility is based on the empirical evidence which showed that **stock returns are negatively associated with unexpected increases in volatility**. This negative association implies **that conditional volatility has an asymmetric relation to past data**. To capture this asymmetry, the GARCH model is modified by making a logarithmic transformation of the conditional variance equation (EGARCH). The GARCH model for volatility provides a more general framework for evaluating the incremental information content of implied volatilities than has been previously available. The GARCH framework explicitly accounts for the time series behavior of the volatility and the relation between expected returns and the conditional market volatility as well. A predictive content of the implied volatility was added to the GARCH model as an exogenous variable and was assessed that the implied volatility is an important determinant of conditional volatility (Day and Lewis(19992)).

This paper is focused on adding a new estimator of volatility constructed on Rao-Cramer-Wolfowitz information bound. It is showed that this estimator is efficient and prices constructed on this estimator convey all information and therefore markets are cleared. An information content of the capital asset market is defined on probability distributions capital assets prices. It is assumed that a family of probability distributions of capital assets prices is an exponential family of probability distributions. The mean-reverting Ornstein-Uhlenbeck process is used for modeling of a stochastic volatility (Oksendal B. (1998)). A byproduct of this analysis is formulated a new notion **of the economic uncertainty principle** saying that a product of the variance volatility estimator and the information content of capital asset market is bounded from below by 1. The economic uncertainty principle says that for obtaining more accurate estimation of the volatility is necessary to increase information content of capital asset market as well. But the information content of the capital asset market is doubtless an exogenous variable. Therefore is possible to obtain only such an accuracy of the estimation for the volatility which is prescribed by an information content of the family (exponential form) of probability distributions capital assets prices.

Stochastic Volatility

Each market participant has to calculate forecasts for a capital asset price and for a volatility. Let us introduce equations for these items:Consider a stochastic differential equation(SDE) for a capital asset price $\boldsymbol{S_t}$

$$dS_t = S_t \cdot (m \cdot dt + s_t \cdot dW_{1t}) \tag{1}$$

where m - the drift parameter is constant, s_t - the diffusion parameter is assumed to change over time by another SDE , W_t - the Wiener process. A solution of the SDE has the following form

$$S_t = S_0 \cdot \exp\left(-\frac{1}{2} t \cdot s_t^2 + m \cdot t + W_{1t} \cdot s_t \right) \tag{2}$$

A model for the stochastic volatility can be expressed by the another SDE formed as follows

$$ds_t = l \cdot \left(s^0 - s_t\right) dt + a\, s_t dW_{2t} \tag{3}$$

where s^0 - long-run mean of the volatility, s_t - the actual volatility

l - the adjustment parameter, $a > 0$ -is a parameter.

By assumption, a volatility of the capital asset price has a long-run mean of σ^0 . But at any time t, the actual volatility may deviate from this long-run mean and the adjustment parameter is λ. The increments dW_{2t} are unpredictable shocks to volatility that are independent of the shocks to capital asset prices S_t . A solution of this SDE by the method from Oksendal has the following form

$$s_t = s_0 \cdot \exp\left(l \cdot \int_0^t \frac{s^0 - s_s}{s_s} ds + a \cdot (W_{2t} - W_{20}) \right) \tag{4}$$

Let us now assume that σ^0 is unknown parameter. We will estimate needed parameters from the realizations of $\{\boldsymbol{S_t},\ \boldsymbol{t} > \boldsymbol{0}\}$ where

$$\begin{aligned} dS_t &= a(S_t, m)dt + b(S_t, s_t)dW_t \\ a(S_t, m) &= m \cdot S_t \\ b(S_t, s_t) &= s_t \cdot S_t \end{aligned} \tag{5}$$

where σ_t represents the expression (4).

A solution of this SDE has the following form

$$S_t = S_0 \cdot \exp\left(-\frac{1}{2} a^2 t \cdot s_t^2 + l \cdot t \cdot (s^0 - s_t) + a \cdot W_t \cdot s_t\right) \tag{6}$$

Thus each market participant has to forecast both a capital asset price and a level of the stochastic volatility. We realize an estimation of long-run volatility σ^0 on an exponential family of probability distributions.

An estimator for $s^{\mathbf{0}}$

Let $\tau=\tau(s)$ be increasing Markov stopping times adapted to the information set $\Phi_t, t \geq 0$, generated mainly by historical assets price time series and $d = d(t(s), s)$ be an estimator for the parameter σ^0 . We will work with an exponential family of probability distributions only. Let us introduce an exponential family of probability distributions:

For random processes the $\{S_t,\ t > 0\}$ and the Wiener process **W** we can derive a density of probability distribution in the following form

$$j\ (s,W) = \exp(\int_0^{t(W)} a(s,m)dW_t + \int_0^{t(W} b(s,s_t)dW_t - \\ -\frac{1}{2}\int_0^{t(W)} a^2(s,W)dt - \frac{1}{2}\int_0^{t(W)} b^2(s,s_t)dt) \tag{7}$$

Let us introduce several technical assumptions (see an appendix)

for both processes and probability distributions as well. Let (**C**,***B***) be a space of continuous functions on $< 0, \infty)$ $S_t, t \geq 0, S_0 = 0$ and ***B*** is a minimal σ-algebra $\{S: S_v, v \leq t\}$. Let m_W and m_S be measures on (**C**,***B***).

Theorem

Let $d = d(t(s), s)$ be an unbiased estimator of parameter s **0**, i.e.,

$$E[d(t(s), s)] = s^0 \tag{8}$$

then

$$E[d(t,s) - s^0]^2 \geq \frac{1}{E[\int_0^{t(s)} (\frac{¶}{¶s_t} b(S_t, s_t))^2 dt]} \tag{9}$$

Proof: A framework of this proof has used of the Rao-Cramer-Wolfowitz's theorem (see Liptser, R.S., Shiryaev, A.N.).

An Economic Uncertainty Principle

An economic uncertainty principle has the following form

(10)

$$E[d(t,s) - s^0]^2 \cdot E[\int_0^{t(s)} S_t^2 dt] \geq 1$$

An information content of the capital asset price we call the following expression

$$(E[\int_0^{t(s)}(\frac{\P}{\P s_t}b(S_t, s_t))^2 dt])^{-1} = (E[\int_0^{t(s)}(\frac{\P}{\P s_t}s_t \cdot S_t)^2 dt])^{-1} = (E[\int_0^{t(s)} S_t^2 dt])^{-1} \quad (11)$$

Example

Let the capital asset price equation has the following form

$$S_t = S_0 \cdot \exp[l \cdot (s^0 - s_t) \cdot t - \frac{1}{2} \cdot a^2 t \cdot s_t^2 + a \cdot s_t \cdot W_t] \quad (12)$$

where **r** is nominal interest rate, **S0** is initial stock price, and **n** is a discrete time. Parameters values used in the following examples are:

$r = 0.18,\ S_0 = 1,\ a = 0.2,\ s^0 = 0.09,\ n = 20$ A behavior this equation is shown on the fig.1

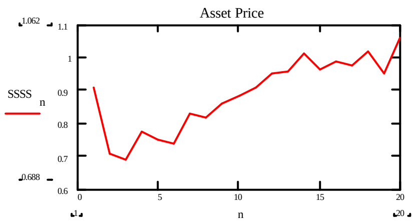


Fig.1

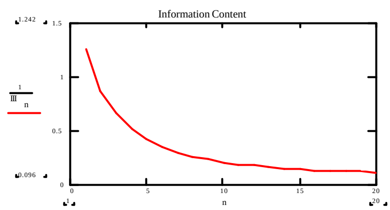


Fig.2

The information content of the capital asset price in this example has the following expression

$$(E[\int_0^{t(s)} S_t^2 dt])^{-1}$$

$$= (E[\int_0^{t(s)}(S_0 \cdot \exp[l \cdot (s^0 - s_t) \cdot t - \frac{1}{2} \cdot a^2 t \cdot s_t^2 + a \cdot s_t \cdot W_t)^2 dt])^{-1} \quad (13)$$

$$= (\frac{S_0^2}{2l(s^0 - s_t) - a^2 \cdot s_t^2} \cdot E[e^{(2l(s^0 - s_t) - a^2 \cdot s_t^2) \cdot t(s)} - 1])^{-1}$$

The behaviour of the information content of the capital asset price in this example, i.e., equation (13), is demonstrated on the fig. 2.

Conclusions

A notion of the information content of the capital asset price is very important for finding of low bound for variance of asset price time series. In this case it is not necessary to use an implied volatility estimator for forecasting capital asset price. Now, we are able to construct an efficient estimator of capital asset price with exponential family distribution for capital asset prices and Wiener processes.

References

[1] Bollerslev, Engle, and Wooldridge (1988): A capital asset pricing model with time varying covariances, Journal of Political Economy 96, 116-131

[2] Canina L., Figlewski S. (1993): The Informational Content of Implied Volatility, The Review of Financial Studies, 6, 3, 659-681

[3] Day T.E., Lewis C. M. (1992): Stock Market Volatility and the Information content of Stock Index Options, Journal of Econometrics, 52, 267-287

[4] Domowitz and Hakkio (1985): Conditional variance and the risk premium in the foreign exchange market, Journal of International Economics 19,47-66

[5] Liptser, R.S., Shiryaev, A.N. (1974): Statistics of Random Processes, Nauka, Moscow, (in Russian)

[6] Merton R.C.(1973): The theory of rational option pricing, Bell Journal of Economics and Management Sciences 4, 141-183

[7] Merton R.C.(1980): On estimating the expected return on the market, Journal of Financial Economics 8, 323-361

[8] Oksendal B. (1998): Stochastic Differential Equations, Springer

[9] Shin F.B., Kil D.H. (1998): Classification Cramer-Rao Bounds on Stock Price Prediction, J. Forecast., 17, 389-399

Appendix

$$\frac{¶}{¶s_t}\int_0^{t(W)} b(S_t, s_t)dW_t = \int_0^{t(W)} \frac{¶}{¶s_t} b(S_t, s_t)dW_t$$

$$\int_0^{\infty} |a(S_t, m)|dt < \infty, \int_0^{\infty} |b(S_t, s_t)|dt < \infty, m_W - \text{and } m_S - a.s.$$

$$0 < E[\int_0^{t(s)} (\frac{¶}{¶s_t} b(S_t, s_t))^2 dt] < \infty \qquad \frac{d}{ds_t} E_j[d(t(s), s)] = E_j[d(t(s), s)\cdot \frac{¶j(s,W)}{¶s_t}]$$

$$\int_0^{\infty} a^2(S_t, m)dt < \infty, \int_0^{\infty} b^2(S_t, s_t)dt < \infty, a.s. \qquad \int_0^{t(W)} (\frac{¶}{¶t} a(S_t, m))^2 dt < \infty, \int_0^{t(W)} (\frac{¶}{¶s_t} b(S_t, s_t))^2 dt <$$