PRZEGLĄD STATYSTYCZNY

STATISTICAL REVIEW

Strona WWW „Przegląd Statystyczny”:
http://www.przegladstatystyczny.pan.pl

SPIS TREŚCI

CONTENTS

ZBIGNIEW ŚWITALSKI[1]

STABILITY AND PRICE EQUILIBRIA
IN A MANY-TO-MANY GALE-SHAPLEY MARKET MODEL[2]

1. INTRODUCTION

Much research on markets with indivisible goods have their sources in the classical models of Gale (model of buying *n* houses by *n* buyers, 1960, Ch. V, § 6), Gale, Shapley (college admissions model, 1962) and Shapley, Shubik ("assignment game" with quasi-linear utility, 1971/72). Different variants of these models have many applications in the theory of recruitment systems, auctions markets, labor markets and so on (see, e.g., Roth, Sotomayor, 1992; Crawford, Knoer, 1981; Andersson, Erlanson, 2013; Biro, Kiselgof, 2013, many interesting applications are presented in the survey of Sönmez, Ünver, 2011, the role of such models in modern economics is explained in Roth, 2002).

One of the main topics, considered in the literature related to such models and their generalizations, is the problem of relationships between the concept of stability (or the core outcome) and the concept of competitive equilibrium. Stable outcomes are often defined as allocations of goods among agents such that no coalition of agents can reallocate the goods in such a way that the situation of all members of the coalition will be improved. Competitive equilibrium is in most cases defined as an allocation of goods and a price vector such that each agent obtains the most preferred goods from the set of all feasible goods (for this agent). The two notions are defined differently, but for many models, for which prices of the goods can be defined, it can be proved that equilibria allocations are stable and stable outcomes are equilibria allocations, associated with some price vectors (see, e.g., Shapley, Shubik, 1971/72; Camina, 2006; Sotomayor, 2007; Hatfield et al., 2013; Herings, 2015).

Studying relationships between stability and competitive equilibria is very important. First, we can find in this way some new characterization of stable outcomes and this may have much importance for stability theory, and second, we can characterize competitive equilibria allocations as stable outcomes, which gives possibility of proving results on existence of equilibria (because existence of stable outcomes can be, in many cases, proved with relatively small effort).

[1] University of Zielona Góra, Faculty of Mathematics, Computer Science and Econometrics, 4a Prof. Z. Szafrana St., 65-516 Zielona Góra, Poland, e-mail: z.switalski@wmie.uz.zgora.pl.

Most of the research in the theory of "matching markets" concerning relationships between stability and competitive equilibria are based on the Shapley-Shubik type models and relatively few are devoted to such relationships for the Gale-Shapley model (see Świtalski, 2008, 2010, 2015, 2016; Azevedo, Leshno, 2011).

In the paper of Świtalski (2016) a variant of many-to-many Gale-Shapley market model was presented, for which exact relationships between stable matchings and a kind of generalized equilibria (called there order equilibria) were proved (see theorems 1 and 2 below). Stability in the paper of Świtalski (2016) is understood as pairwise stability in the sense of Echenique, Oviedo (2006) and is equivalent to the concept of stability used by Alkan, Gale (2003). Order equilibria generalize classical price equilibria, and so the results of the paper of Świtalski (2016) can be used to study relationships between stability and price equilibria for the GS-models.

The presented paper may be treated as a continuation of the paper of Świtalski (2016). Using the results of our previous paper, we study in detail relationships between (pairwise) stability and price equilibria for a generalized many-to-many Gale-Shapley market model with choice functions representing preferences of the buyers, reservation prices of the buyers and weak orders representing preferences of the sellers. In our model we assume that preferences of the sellers are determined by (or at least are closely related to) reservation prices of the buyers (see definition 13). The model is a generalization of the standard one-to-one GS-model, but can also be treated as a many-to-many generalization of the Chen, Deng and Ghosh's model (Chen et al., 2014) of matching markets with budgets (we use the term "reservation price" instead of "budget", see the comments at the beginning of section 3).

The main result of our paper is theorem 5 which shows that, under the assumption of path independence of choice functions of the buyers, strongly stable matchings for the generalized GS-model are identical with price equilibria allocations for this model. Using this result and some results of Alkan, Gale (2003) we prove theorems on existence of price equilibria for many-to-many GS-model or for many-to-many version of Chen, Deng and Ghosh's model (Chen et al., 2014) (theorems 7 and 8).

Our results can be treated as a far-reaching generalization of "supply and demand lemma" of Azevedo, Leshno (2011, p. 18; see also introduction in Świtalski, 2016). First, we consider many-to-many model with choice functions at one side of the market and weak orders at the second side (Azevedo and Leshno study many-to-one model with strict linear orders at both sides of the market). Secondly, we consider different conditions of compatibility of reservation prices (= scores in Azevedo, Leshno) with preferences of the buyers (colleges) and two different conditions of stability (see definition 7). We also study in detail how implications relating stable matchings and equilibria allocations (i.e. stable matchings $\Rightarrow$ equilibria allocations and equilibria allocations $\Rightarrow$ stable matchings) depend on particular properties of choice functions (outcast and heritage properties, see theorem 3 and 4). We consequently use the terminology of equilibrium theory (e.g. prices and equilibria allocations) and

reformulate Azevedo and Leshno's result (which can be treated as a special case of our results) using this terminology.

It is worth noting that the many-to-many model we use cannot be treated as a special case of typical many-to-many models (e.g. presented by Echenique, Oviedo, 2006 or used in the contract theory, see e.g. Klaus, Walz, 2009; Kominers, 2012; Hatfield, Kominers, 2016). The reason is that we use choice functions which are not necessarily generated by strict linear orders on the families of subsets of feasible contracts (such assumption is commonly used in "many-to-many" papers). A similar approach with general choice functions for many-to-many models was used by Alkan, Gale (2003).

Alkan, Gale (2003, pp. 290, 291) argue that the model with choice functions may be more suitable for markets in which sellers (or buyers), e.g. colleges, want to satisfy requirements concerned with diversity (e.g. racial or ethnical) of chosen groups (e.g. groups of students). What's more, assumption about linear ordering of subsets of agents, in many cases, is not necessary because all relevant information about preferences of the agent is included in the choice function alone. In our paper we follow the Alkan and Gale's approach and show that there can be proved interesting results about relationships between stability and equilibria without the assumption about linear ordering of subsets of agents (or contracts).

In our model we use standard pairwise stability condition which can also be defined without the assumption about preference ordering of subsets of contracts (our definition agree with the definition of Echenique, Oviedo, 2006 and the one used by Alkan, Gale, 2003) and do not refer to other stability concepts (for example setwise stability or other stabilities defined by Echenique, Oviedo, 2006). Pairwise stability gives possibility of using the result of Alkan, Gale (2003) for proving existence of price equilibria in our model (see theorem 7).

The paper is organized as follows. In section 2 we present the model of market which is taken from Świtalski (2016) and define generalized (order) equilibria for such a model. In section 3 we study the problem of relationships between price equilibria and order equilibria. In section 4 we formulate and prove results on relationships between price equilibria and stable matchings and on existence of price equilibria.

2. THE MODEL

We present here a generalized many-to-many Gale-Shapley market model described in the paper of Świtalski (2016), where the reader can also find the motivation for studying such a model. The model is based on a many-to-many version of the Gale-Shapley (1962) college admissions model (many-to-many models of the GS type are studied, e.g., in the paper of Alkan, Gale, 2003 or Echenique, Oviedo, 2006, in both papers preferences of agents are represented by choice functions, but in the last paper with the additional assumption that the choice functions are generated by strict linear orders on the families of subsets of contracts).

In some cases (for preferences satisfying additional requirements) many-to-many model can be easily reduced to many-to-one or one-to-many model by assuming that, say, buyers, matched to multiple sellers, are represented by sets of buyers with identical preferences and matched to only one seller. We do not follow this way of reasoning. We think that studying many-to-many model in full generality is more elegant and with obvious real-life interpretations. Moreover, for the presented below many-to-many model we can directly use the results of our previous paper (Świtalski, 2016).

In our model we have a finite set of buyers – B, a finite set of sellers – S, preferences of buyers over sellers and preferences of sellers over buyers. As an example (see Echenique, Oviedo, 2006), let B be a set of firms and S – a set of consultants. Each firm wants to hire a set of consultants and each consultant wants to work for a set of firms. Firms rank consultants according to their competences and consultants rank firms according to their subjective preferences.

There are many other examples of many-to-many markets, mainly related to labor markets. For example, Echenique, Oviedo (2006) mention markets for medical interns in the U.K. or teacher (university professor) markets in some countries where teachers (or professors) can work in more than one school (university). Another example from Echenique, Oviedo (2006) is a model of contracting between down-stream firms and up-stream providers.

In what follows we use basic notation and definitions from the paper of Świtalski (2016).

We consider Cartesian product $B \times S$ of the sets B and S. For any relation $u \subset B \times S$ and for any $b \in B$, $s \in S$, we define the sets:

$$u(b) = \{s \in S : (b, s) \in u\}, \tag{1}$$

$$u(s) = \{b \in B : (b, s) \in u\}. \tag{2}$$

We assume that a non-empty set of acceptable (feasible) pairs $F \subset B \times S$ is defined. We interpret an acceptable pair $(b, s) \in F$ as a possible transaction which can be realized in the market or (from the point of view of contract theory of Hatfield et al., 2013), as a possible contract which can be signed by b and s (for example in the "consultants" market, $(b, s) \in F$ means that it is possible for a firm b to hire a consultant s and it is possible for s to work for b). According to (1) and (2), the sets $F(b)$ and $F(s)$ for any $b \in B$ and $s \in S$ can be defined. The set $F(b)$ can be interpreted as the set of sellers which can sign a contract with b, and $F(s)$ – as the set of buyers which can sign a contract with s.

We assume that b can sign many contracts with different sellers, but only one contract with a given seller s, and s can sign many contracts with different buyers, but only one contract with a given buyer b (hence we assume that the "unitarity" assumption is satisfied (see, e.g., Kominers, 2012).

In our model we introduce quotas for buyers and sellers. Let $q(b) \geq 1$ be the quota for b, which is interpreted as maximal number of contracts which b can sign with different sellers and $q(s) \geq 1$ – the quota for s, which is interpreted as maximal number of contracts which s can sign with different buyers. We assume that $\# F(b) \geq q(b)$ and $\# F(s) \geq q(s)$ ($\# A$ denotes cardinality of A).

Preferences of sellers are represented by weak orders. Namely, we assume that on every set $F(s)$, a weak order (transitive and complete relation) $\geq_s$ is defined (i.e. some buyers may be indifferent for the seller s). The symbols $>_s$ and $\approx_s$ denote the respective strict order and indifference relation. Hence the notation $b >_s c$ means that the buyer b is better than the buyer c for the seller s, and $b \approx_s c$ means that b and c are indifferent for s.

Preferences of buyers are represented by choice functions (standard theory of choice functions is described, e.g., by Aizerman, Aleskerov, 1995 or by Aleskerov, Monjardet, 2002, applications for matching markets can be found, e.g., in Echenique, 2007; Klaus, Walzl, 2009; Hatfield et al., 2013).

We assume that a choice function is defined for every feasible set $F(b)$ (for a given buyer b). This means that for every buyer b and every set of feasible sellers $X \subset F(b)$, a set $C(b, X) \subset X$ is defined. The set $C(b, X)$ is interpreted in the following way. Assume that b considers some set of feasible sellers X. Then his decision will be to choose the set $C(b, X)$ as the set of sellers, with which he will sign a contract. We consider only the so-called quota-filling choice functions (Alkan, Gale, 2003), i.e. we assume that:

(i) $\quad C(b, X) = X, \quad$ if $\quad \# X < q(b)$,

(ii) $\quad \# C(b, X) = q(b), \quad$ if $\quad \# X \geq q(b)$.

We do not assume that the choice function of a buyer b is generated by a linear order over the subsets of the set $F(b)$ (hence we follow the model of Alkan and Gale and do not follow standard approaches on many-to-many markets or contract theory as in, e.g., Echenique, Oviedo, 2006; Klaus,Walz, 2009; Kominers, 2012).

It is worth noting that any linear order $>_s$ on $F(s)$ (when there are no indifferences between different buyers) generates, in an obvious way, a quota-filling choice function on $F(s)$ (we take, for any $X \subset F(s)$, the set of $q(s)$ best buyers in X, or the set X, if $\# X < q(s)$).

We need the following properties of the function C (interpretation and comments on these properties can be found in the paper of Świtalski, 2016):

Definition 1. A choice function C satisfies the outcast property if for every $b \in B$ and $X, Y \subset F(b)$ we have

$$Y \subset X \setminus C(b, X) \quad \Rightarrow \quad C(b, X \setminus Y) = C(b, X). \tag{3}$$

Definition 2. A choice function C satisfies the heritage property if for every $b \in B$ and $X, Y \subset F(b)$ we have

$$Y \subset X \quad \Rightarrow \quad Y \cap C(b, X) \subset C(b, Y). \tag{4}$$

We use here the classical terms (outcast and heritage) taken from the literature on choice theory (Aizerman, Aleskerov, 1995; Aleskerov, Monjardet, 2002). In the matching literature outcast property is known also under the name independence (see, e.g., Echenique, 2007) or consistency (Alkan, Gale, 2003) and heritage property under the name substitutability (Echenique, Oviedo, 2006) or persistency (Alkan, Gale, 2003).

Choice functions satisfying the outcast and heritage properties are called path independent (or Plott) choice functions (see Danilov, Koshevoy, 2005). Example of Plott choice function is the choice determined by some linear order (then $C(b, X)$ is the set of $q(b)$ best sellers in X, if $\# X \geq q(b)$, and $C(b, X) = X$ otherwise).

The next examples show that there exist quota-filling Plott choice functions with real-life interpretations which are not generated by any linear order.

Example 1. We fix a buyer $b \in B$. Let $q(b) = 2q$ be a fixed even number ($q \geq 1$). Assume that the set of sellers S is divided into two disjoint subsets Y and Z (i.e. $S = Y \cup Z$ and $Y \cap Z = \varnothing$). For example Y may be a set of men and Y a set of women, or Y – a set of statisticians and Z – a set of computer scientists (in the set of consultants S). Assume that there is a (strict) linear order M on S and for any $X \subset F(b)$ and any $n \geq 1$ define

$$M(X, n) = \begin{cases} \text{the set of } n \text{ best elements in } X \text{ with respect to } M, & \text{if} \quad \#X \geq n, \\ X, & \text{if} \quad \#X < n. \end{cases}$$

We want to take into account, when choosing the agents from X, the gender quotas (or quotas for consultants of specific professions). To this end we can define the following choice function:

$$C(b, X) = \begin{cases} M(X \cap Y, q) \cup M(X \cap Z, q), & \text{if } \# X \cap Y \geq q, \quad \# X \cap Z \geq q, \\ (X \cap Y) \cup M(X \cap Z, 2q - m), & \text{if } \# X \cap Y = m < q, \# X \cap Z \geq q, \\ M(X \cap Y, 2q - m) \cup (X \cap Z), & \text{if } \# X \cap Y \geq q, \quad \# X \cap Z = m < q, \end{cases}$$

and $C(b, X) = X$ otherwise.

Hence we choose the q best consultants from the set $X \cap Y$ and the q best consultants from the set $X \cap Z$, or m consultants from one of the sets and $2q - m$ (or less) best consultants from the second set (if $m < q$ and the first set contains only m consultants)

or the whole set of consultants if both sets ($X \cap Y$ and $X \cap Z$) contain less than q consultants.

It is not difficult to prove that the defined above choice function satisfies both the outcast and heritage properties and hence is a Plott choice function.

Example 2. We assume now that there are two different linear orders K and L defined on the set X (e.g. we order the set of consultants according to two criteria: creativity (K) and experience (L)). We want to have, when choosing the specialists from the set X, a balance between the number of creative members of the group and the number of experienced members. We define the following choice function (the symbols $K(X, q)$ and $L(X, q)$ are defined similarly as $M(X, n)$ in the previous example):

$$C(b, X) \ = \ \left| \begin{array}{l} X, \quad \text{if} \quad \#X < 2q, \\ K(X, q) \cup L(X \setminus K(X, q), q), \quad \text{if} \quad \#X \geq 2q. \end{array} \right.$$

Hence, if we have at our disposal at least $2q$ consultants, we choose first the q most creative ones and then the q most experienced from the rest of the group. It is also easy to show that such choice function satisfies the Plott conditions (outcast and heritage).

We define a generalized GS-model as a 6-tuple (B, S, F, C, P, q), where F is the set of acceptable pairs, C is the family of choice functions (defined for all $b \in B$), P is the family of weak orders (defined for all $s \in S$), and q is the vector of quotas (defined for all $b \in B$ and all $s \in S$).

In the next definitions we define matchings and (strongly) stable matchings for a generalized GS-model (B, S, F, C, P, q) (see Świtalski, 2016). Our definition of stability is equivalent to the definition of Alkan, Gale (2003) and is often called pairwise stability in the matching literature (see, e.g., Echenique, Oviedo, 2006).

Definition 3. A relation $u \subset B \times S$ is a matching if

(i) $u \ \subset \ F$,

(ii) $\# \ u(b) \ \leq \ q(b)$, $\forall b \in B$,

(iii) $\# \ u(s) \ \leq \ q(s)$, $\forall s \in S$.

A matching u can be interpreted as a set of actual transactions realized in the market (contrary to F which can be treated as the set of potential (possible) transactions in the market).

Definition 4. Let $u \subset B \times S$ be a matching. We say that a seller $s \in F(b)$ improves the situation of a buyer $b \in F(s)$ (we write $s >_b u(b)$) if $s \in C (b, u(b) \cup \{s\})$.

Definition 5. Let $u \subset B \times S$ be a matching. We say that a buyer $b \in F(s)$ improves the situation (weakly improves the situation) of a seller $s \in F(b)$ (we write $b >_s u(s)$ or $b \geq_s u(s)$ respectively) if at least one of the following conditions hold:

(i) $\# u(s) < q(s)$,

(ii) $\exists c \in u(s), \ b >_s c \quad (b \geq_s c)$.

Definition 6. A pair $(b, s) \in B \times S$ is a blocking pair (weakly blocking pair) for a matching $u \subset B \times S$ if

(i) $(b, s) \in F \setminus u$,

(ii) $s >_b u(b)$,

(iii) $b >_s u(s) \qquad (b \geq_s u(s))$.

Definition 7. A matching $u \subset B \times S$ is stable (strongly stable) if there are no blocking pairs (weakly blocking pairs) for u.

Now we start describing the notion of generalized equilibrium in our model. Generalized equilibria for the model (B, S, F, C, P, q) are defined with the help of families of sets $W(s) \subset F(s)$ where the sets $W(s)$ are such that being in the set $W(s)$ is for $b \in F(s)$ a necessary condition to sign a contract with s (for example a consultant s determine some minimal conditions under which she can work for a firm b, the set of all firms satisfying such conditions will be denoted by $W(s)$). The feasible sets $F(s)$ are fixed, but the sets $W(s)$ can vary for the given model (B, S, F, C, P, q). Special case of conditions $W(s)$ are the price conditions of the form $W(s) = \{b \in F(s): r(b, s) \geq p(s)\}$, where $p(s)$ is a price of a good offered by s and $r(b, s)$ – maximal price at which b can buy the good offered by s (see Świtalski, 2016).

A family $W = \{W(s)\}$ ($s \in S$) will be called a system of conditions. The set of feasible sellers for a buyer b under the system $W = \{W(s)\}$ is defined as:

$$F(W, b) = \{s \in S: \ b \in W(s)\}.$$

$F(W, b)$ is the set of sellers s, such that b can sign a contract with s (b satisfies the conditions $W(s)$ stated by s). Obviously, $F(W, b) \subset F(b)$, i.e., each seller feasible for b under W is acceptable for b.

We define also the set of "best" sellers (contracts) for b under the system W as

$$M(W, b) \ = \ C(b, F(W, b)),$$

and the demand set for the seller s (under the conditions W) as

$$D(W, s) = \{b \in F(s): \ s \in M(W, b)\}.$$

Demand set $D(W, s)$ is the set of all buyers for which s is among the "best" sellers.

Now we define generalized equilibrium for the model (B, S, F, C, P, q) (see Świtalski, 2016).

Definition 8. A system of conditions $W = \{W(s)\}$ is an *equilibrium system* for the model (B, S, F, C, P, q) if

  (i) $\# D(W, s) \leq q(s)$,   for all $s \in S$.
  (ii) $W(s) = F(s)$,   for all $s \in S$ such that $\# D(W, s) < q(s)$.

For an equilibrium system $W = \{W(s)\}$ we define a matching associated with W as:

$$u(W) = \{(b, s) \in F: \ b \in D(W, s)\}.$$

Definition 9. An *equilibrium* (for a generalized GS-model (B, S, F, C, P, q)) is a pair (u, W) such that W is an equilibrium system and $u = u(W)$.

If (u, W) is an equilibrium we will say that $u = u(W)$ is an equilibrium allocation associated with W.

In the paper of Świtalski (2016) some special kind of generalized equilibria have been defined, namely the so-called order equilibria. These are equilibria for which the conditions $W(s)$ are "order" conditions in the sense that if a buyer b satisfies $W(s)$, then all the buyers better (not worse) than b also satisfy $W(s)$ (so there is some kind of "compatibility" of $W(s)$ with the preferences P).

Formal definitions are the following:

Definition 10. A system of conditions $W = \{W(s)$ is compatible (strongly compatible) with the sellers' preferences if

$$b \in W(s) \quad \wedge \quad c >_s b \quad \Rightarrow \quad c \in W(s), \qquad \text{for all } \ s \in S,$$

$$(b \in W(s) \quad \wedge \quad c \geq_s b \quad \Rightarrow \quad c \in W(s), \qquad \text{for all } \ s \in S).$$

Definition 11. An equilibrium (u, W) for a generalized GS-model (B, S, F, C, P, q) is an *order equilibrium* (strongly order equilibrium) if W is compatible (strongly compatible) with the sellers' preferences.

Let (B, S, F, C, P, q) be a generalized GS-model. We will say that the family C satisfies the outcast (heritage) property if all the choice functions in the family C satisfy this property.

In Świtalski (2016) some relationships between stability and order equilibria were proved. Namely, it was proved that under the outcast property of C, order (strongly order) equilibria allocations are stable (strongly stable) and that under the heritage

property, stable (strongly stable) matchings are order (strongly order) equilibria allocations. The exact formulations (see lemmas 1 and 2 in Świtalski, 2016) are the following:

Theorem 1. Let $M = (B, S, F, C, P, q)$ be a generalized GS-model such that C satisfies the outcast property. Let (u, W) be an order (strongly order) equilibrium for the model M. Then the matching u is stable (strongly stable).

Theorem 2. Let u be a stable (strongly stable) matching in a generalized GS-model $M = (B, S, F, C, P, q)$ such that C satisfies the heritage property. Then there exists a system of conditions W compatible (strongly compatible) with P such that (u, W) is an order (strongly order) equilibrium.

In section 4 we use theorems 1 and 2 to characterize price equilibria for the generalized GS-models with reservation prices (= maximal prices, at which buyers are ready to sign contracts with the sellers).

Models with reservation prices, price equilibria and relationships between price equilibria and order equilibria are described in the next section.

3. PRICE EQUILIBRIA AND ORDER EQUILIBRIA

In most market models competitive equilibria are defined as price equilibria. In Świtalski (2016, section 2) price equilibria were defined for the simplest (one-to-one) version of the Gale-Shapley model. To define such equilibria we have introduced reservation prices $r(b, s)$, where the number $r(b, s)$ is interpreted as maximal price at which buyer b is willing to enter into the transaction with the seller s (to sign a contract with s). Such equilibrium model is analogous to the model of equilibrium for matching markets with budgets described by Chen et al. (2014) (authors interpret the number $r(b, s)$ as budget which is at the disposal of b when signing the contract with s).

If we assume that in the one-to-one Gale-Shapley model sellers' preferences are represented by linear orders and are determined by reservation prices, i.e., if the following condition is satisfied (for any $b, c \in B$ and $s \in S$):

$$b >_s c \quad \Leftrightarrow \quad r(b, s) > r(c, s), \tag{5}$$

then it can be proved that the respective price equilibria allocations are stable (in the sense of Gale, Shapley, 1962) and vice versa (see, Świtalski, 2016, theorem 1). Similar result, with the college admissions interpretation, was proved by Azevedo, Leshno (2011, "supply and demand lemma", p. 18).

We generalize (see theorem 6) this result to many-to-many GS-models described in section 2. We start now with describing many-to-many models with reservation prices and with defining price equilibria for such models.

Consider a generalized GS-model (B, S, F, C, P, q). Assume that for every buyer b and every seller s such that $(b, s) \in F$, similarly as in the one-to-one model, a reservation price $r(b, s)$ is defined such that b is ready to pay no more than $r(b, s)$, when signing the contract (b, s).

A generalized GS-model with prices is defined as a 7-tuple (B, S, F, C, P, q, r), where r denotes the vector of reservation prices (for all $(b, s) \in F$).

Assume that every seller $s \in S$ announces some price $p(s) \geq 0$ interpreted as minimal price at which she is ready to sign contracts with the buyers. We define a price vector p as a sequence $p = (p(s))$ $(s \in S)$ of prices announced by sellers.

We can define price conditions $W(p)(s)$ in the following way:

$$W(p)(s) = \{b \in F(s): \quad r(b, s) \geq p(s)\}.$$

Inequality $r(b, s) \geq p(s)$ can be interpreted as some kind of "budget constraint" (similarly as in the neoclassical model of consumer choice). The set $W(p)(s)$ can be interpreted as the set of buyers, with which s can sign a contract when prices in the market are p. Having defined the conditions $W(p)(s)$ we can easily define price equilibria for the model (B, S, F, C, P, q, r) namely:

Definition 12. Let $M = (B, S, F, C, P, q, r)$ be a generalized GS-model with prices and (u, W) an equilibrium for the model (B, S, F, C, P, q) (according to definition 9). We say that (u, W) is a price equilibrium for the model M if there exists a price vector p such that $W = W(p)$.

If (u, W) is a price equilibrium and $W = W(p)$, we also use the notation $(u, W) = (u, W(p)) = (u, p)$.

To state the results on relationships between stability and price equilibria in many-to-many case (see theorems 3 and 4 below) we use three different conditions of compatibility of reservation prices r with preferences P, one of which is equivalent to (5) and two others are weaker versions of (5).

A triple (b, c, s) is called acceptable if $(b, s) \in F$ and $(c, s) \in F$.

Definition 13. We say that the prices r are
(i) compatible with the preferences P (shortly – COMP) if for all acceptable triples (b, c, s) we have

$$b >_s c \quad \Rightarrow \quad r(b, s) \geq r(c, s), \tag{6}$$

   (ii) strongly compatible with the preferences P (shortly – SCOMP) if for all accep-
table triples (b, c, s) we have

$$b \geq_s c \quad \Rightarrow \quad r(b, s) \geq r(c, s), \tag{7}$$

   (iii) very strongly compatible with the preferences P (shortly – VSCOMP) if for
all acceptable triples (b, c, s) we have

$$b \geq_s c \quad \Leftrightarrow \quad r(b, s) \geq r(c, s). \tag{8}$$

It is easy to see that (8) is the strongest condition and (7) is stronger than (6). In other
words, implications (8) $\Rightarrow$ (7) $\Rightarrow$ (6) are valid, although none of these implications
can be reversed. It is also easy to check that (8) is equivalent to (5).

    Conditions (6)–(8) mean that there is some relationship between reservation prices
and the sellers' preferences. In the case of the strongest condition VSCOMP = (8) it
means that preferences are determined by reservation prices similarly as in the one-
to-one case (see (5)). Observe that the conditions (6)–(8) do not restrict the domain
of possible preference orderings of the sellers, because for any weak order $\geq_s$ we
can obviously find reservation prices r satisfying VSCOMP (and hence COMP and
SCOMP).

    Using the conditions (6)–(8) we study now relationships between price equilibria
and order (or strongly order) equilibria defined in section 2 (such relationships are
necessary for transforming theorems 1 and 2 into theorems about price equilibria).

    The following proposition shows that under the conditions of (strong) compatibility
price equilibria are order (or strongly order) equilibria

    Proposition 1. If prices r in a model $M = (B, S, F, C, P, q, r)$ are (strongly)
compatible with the preferences P, then any price equilibrium for M is a (strongly)
order equilibrium for (B, S, F, C, P, q).

Proof. Assume that the prices r are compatible with P. Then, for any acceptable triple
(b, c, s) we have:

$$b >_s c \quad \Rightarrow \quad r(b, s) \geq r(c, s). \tag{9}$$

We want to prove (by definition 10) that for any price vector $p = (p(s))$ and any
acceptable triple (b, c, s) we have:

$$r(c, s) \geq p(s) \quad \wedge \quad b >_s c \quad \Rightarrow \quad r(b, s) \geq p(s). \tag{10}$$

It is easy to see that (9) implies (10). The proof for strong compatibility and strong
equilibrium is quite analogous (we change $b >_s c$ by $b \geq_s c$). $\qquad\qquad\square$

Of course we can also ask the reverse question: whether every (strongly) order equilibrium is a price equilibrium. The following proposition shows that this is true under the condition of very strong compatibility (taking into account the previous proposition we can prove even "if and only if" statement in this case).

Proposition 2. If prices r in a model $M = (B, S, F, C, P, q, r)$ are very strongly compatible with the preferences P, then an equilibrium (u, W) is a price equilibrium for M if and only if it is a strongly order equilibrium for (B, S, F, C, P, q).

Proof. ($\Rightarrow$) Let (u, W) be a price equilibrium for M. Very strong compatibility of r implies strong compatibility of r, hence, by proposition 1, (u, W) is a strongly order equilibrium for (B, S, F, C, P, q).
($\Leftarrow$) Let (u, W) be a strongly order equilibrium for (B, S, F, C, P, q). By the definition of equilibrium system, all the sets $W(s)$ are non-empty. Let $c(s)$ be the worst buyer (one of the worst buyers) in the set $W(s)$. We define $p(s) = r(c(s), s)$. We want to prove that (u, W) is a price equilibrium for M. It suffices to show that $W = W(p)$, where $p = (p(s))$ $(s \in S)$ or, equivalently, to show that $W(s) = W(p)(s)$ for all $s \in S$. To prove that $W(s) \subset W(p)(s)$, let $b \in W(s)$. Then $b \geq_s c(s)$ (because $c(s)$ is worst in $W(s)$). By (8), $r(b, s) \geq r(c(s), s) = p(s)$, hence $b \in W(p)(s)$.
To prove that $W(p)(s) \subset W(s)$, let $b \in W(p)(s)$. Then $b \in F(s)$ and $r(b, s) \geq p(s) = r(c(s), s)$. By (8), $b \geq_s c(s)$. We have $c(s) \in W(s)$ and (u, W) is a strongly order equilibrium, hence, by definition 10, $b \in W(s)$.   □

Propositions 1 and 2 will be used to prove the characterization results (theorems 3 and 4) in section 4.
Denote by PE – the condition of being a price equilibrium, by OE – the condition of being an order equilibrium, by SOE – the condition of being a strongly order equilibrium.
Taking into account that SOE $\Rightarrow$ OE, propositions 1 and 2 imply the following statements:

$$\text{COMP} \quad \Rightarrow \quad (\text{PE} \Rightarrow \text{OE}), \tag{11}$$

$$\text{SCOMP} \quad \Rightarrow \quad (\text{PE} \Rightarrow \text{SOE} \Rightarrow \text{OE}), \tag{12}$$

$$\text{VSCOMP} \quad \Rightarrow \quad (\text{PE} \Leftrightarrow \text{SOE} \Rightarrow \text{OE}). \tag{13}$$

The next examples show that the implications at the right side of the implications (11)–(13) cannot be reversed or that they cannot be made stronger (example 1 shows that under the assumptions COMP, SCOMP or VSCOMP, OE $\Rightarrow$ PE may be not true, example 2 shows that under COMP or SCOMP, SOE $\Rightarrow$ PE may be not true and example 3 shows that under COMP, PE $\Rightarrow$ SOE may be not true).

Example 1. Let $B = \{b, c\}$, $S = \{s\}$, $F = \{(b, s), (c, s)\}$, $C(b, \{s\}) = \{s\}$, $C(c, \{s\}) = \{s\}$, b and c are indifferent for s, all quotas are equal to 1, $r(b, s) = r(c, s) = 1$, $W(s) = \{b\}$. Then VSCOMP = (8) is satisfied, and hence COMP = (6) and also SCOMP = (7) (because VSCOMP $\Rightarrow$ SCOMP $\Rightarrow$ COMP). It is easy to see that (u, W) is an order (but not strongly order) equilibrium with allocation $u = \{(b, s)\}$ and (u, W) is not a price equilibrium (for any equilibrium prices p we should have $W(p)(s) \neq \varnothing$, hence $W(p)(s) = \{b, c\} \neq W(s)$, because $r(b, s) = r(c, s) = 1$). Hence (u, W) satisfies OE but not PE and so the implication OE $\Rightarrow$ PE is not true.

Example 2. Let $B = \{a, b, c\}$, $S = \{s, t, v\}$, $F = B \times S$ and all quotas are equal to 1. Preferences of the buyers and sellers are the following:

a: s t v s: $[a\ b]$ c
b: t s v t: a b c
c: s v t v: c a b

All preference orderings are strict (linear orders) except preferences for the seller s, for which the buyers a and b are indifferent (it is denoted by $[a\ b]$). The choice functions of the buyers are determined by the preference orderings in an obvious way (we choose the best seller from any non-empty set of sellers).

We consider the matching $u = \{(a, s), (b, t), (c, v)\}$. It is easy to see that u is strongly stable (and hence stable). The only weakly blocking pair could be (c, s), but this is impossible, because $a >_s c$. Hence, by theorem 2, there exists a strongly order equilibrium (u, W) and, by the proof of theorem 2 (see Świtalski, 2016, proof of lemma 2), we can take $W(s) = \{a, b\}$, $W(t) = \{a, b\}$, $W(v) = \{c\}$. Assume that all reservation prices for s are equal to 1. Hence the conditions COMP and SCOMP are satisfied. If (u, p) would be a price equilibrium, then $W(p)(s) = \{a, b, c\}$ (because $W(p)(s) \neq \varnothing$). Hence $a, c \in D(W(p), s)$, and so $\# D(W(p), s) \geq 2$. Thus the system of conditions $W(p)$ would not be an equilibrium system (because $q(s) = 1$). Hence (u, W) satisfies SOE, but not PE (there is no price equilibrium at all in this case) and the implication SOE $\Rightarrow$ PE is not valid.

Example 3. Consider once more example 1 but with reservation prices $r(b, s) = 2$, $r(c, s) = 1$. Hence the condition COMP is satisfied. Take $p(s) = 2$. Then $W(p)(s) = \{b\}$ and it is easy to see that $(u, W(p))$ is a price equilibrium with $u = \{(b, s)\}$. On the other hand it does not satisfy SOE (because $b \in W(p)(s)$, $c \notin W(p)(s)$ and $c \geq_s b$, by indifference of b and c). Hence PE $\Rightarrow$ SOE is not true in this case.

4. PRICE EQUILIBRIA AND STABLE MATCHINGS

Now we state the main results of our paper, namely the results on relationships between stable matchings and price equilibria allocations for the models (B, S, F, C, P, q, r). Proofs of these results will be based on theorems 1 and 2 and propositions 1

and 2. Firstly, using theorem 1, we prove that, under the assumptions of compatibility COMP (SCOMP), any price equilibrium allocation is stable (strongly stable).

Theorem 3. Let $M = (B, S, F, C, P, q, r)$ be a generalized GS-model with prices in which choice functions C satisfy the outcast property and prices r are compatible (strongly compatible) with the preferences P. Let (u, p) be a price equilibrium for the model M. Then u is stable (strongly stable).

Proof. If (u, p) is a price equilibrium, then $(u, p) = (u, W(p))$. Prices are compatible (strongly compatible) with the preferences P, hence, by proposition 1, equilibrium $(u, W(p))$ is an order (strongly order) equilibrium. Hence, by theorem 1, u is stable (strongly stable). □

Theorem 3 implies that under any compatibility condition (COMP, SCOMP, VSCOMP), in the generalized models with outcast choice functions, price equilibria allocations are stable or even (under SCOMP or VSCOMP) strongly stable.

Now we want to reverse theorem 3. Namely we want to state conditions under which any stable (or strongly stable) matching in a model (B, S, F, C, P, q, r) is a price equilibrium allocation.

To this end, we could use theorem 2 which says that if C satisfies the heritage property, then any stable (strongly stable) matching is an order (strongly order) equilibrium allocation. Unfortunately, neither COMP nor SCOMP nor VSCOMP conditions do not guarantee that the obtained order equilibrium allocation would be a price equilibrium allocation (as it is shown by example 1) and neither COMP nor SCOMP conditions do not guarantee that the obtained strongly order equilibrium allocation would be a price equilibrium allocation (example 2).

The only result we can obtain with the help of theorem 2, which could be interpreted as reversion of theorem 3 is the following.

Theorem 4. Let $M = (B, S, F, C, P, q, r)$ be a generalized GS-model with prices in which choice functions C satisfy the heritage property and prices r are very strongly compatible with the preferences P. Let u be a strongly stable matching. Then there exist prices p, such that (u, p) is a price equilibrium.

Proof. By theorem 2, there exist conditions W such that (u, W) is a strongly order equilibrium. By proposition 2, (u, W) is a price equilibrium for M, hence there exist prices p such that $(u, W) = (u, W(p)) = (u, p)$ and (u, p) is obviously a price equilibrium. □

Example 2 shows that the assumption about very strong compatibility of r is essential in theorem 4 (in this example u is strongly stable and prices r are strongly compatible with P, yet there is no price equilibrium of the form (u, p)).

Also the assumption about strong stability of u is essential in theorem 4 (in the example 1 we have stable matching u, prices r are very strongly compatible with P, yet there are no prices p such that (u, p) is a price equilibrium).

Combining theorems 3 and 4 we obtain the following result on equivalence of strongly stable matchings with price equilibria allocations in the generalized GS-models with prices.

Theorem 5. Let $M = (B, S, F, C, P, q, r)$ be a generalized GS-model with prices in which choice functions C satisfy Plott condition and reservation prices r are very strongly compatible with the preferences P. Then a matching u is strongly stable if and only if it is a price equilibrium allocation.

Unfortunately, as the example 2 shows, there is no, in general, similar equivalence between stable matchings and price equilibria allocations. Yet, if we take a model with linear preferences (i.e. one in which preferences of the sellers are linear orders), then such equivalence is obvious, because stability = strong stability in this case. So we have the following result (a many-to-many generalization of theorem 1 in Świtalski, 2016):

Theorem 6. Let $M = (B, S, F, C, P, q, r)$ be a generalized GS-model with prices, with choice functions C satisfying Plott condition, with linear preferences P and with the reservation prices r satisfying VSCOMP = (8) (equivalently (5)) condition.

Then a matching u is stable if and only if it is a price equilibrium allocation.

Now we consider the problem of existence of equilibrium prices for the generalized GS-models with prices. It is easy to see that in general there can be models for which there are no equilibrium prices at all (as in the example 1). Yet, theorem 6 combined with some results of Alkan, Gale (2003) implies that for the models with linear preferences of the sellers we can prove the existence theorem, namely:

Theorem 7. Let $M = (B, S, F, C, P, q, r)$ be a generalized GS-model with prices, with choice functions C satisfying Plott condition, with linear preferences P and with the reservation prices r satisfying VSCOMP condition. Then there exists a price equilibrium (u, p) for the model M.

Proof. Using the results of Alkan, Gale (2003) we can prove, similarly as in the proof of theorem 3 in the paper of Świtalski (2016), that there exists a stable matching u for the model (B, S, F, C, P, q). Obviously u is also stable for the model $M = (B, S, F, C, P, q, r)$ and, by theorem 6, it is a price equilibrium allocation. Hence, there exists a price equilibrium (u, p) for the model M.   □

Using theorem 7 we can also prove some existence result for a many-to-many variant of the model of Chen, Deng and Ghosh (CDG-model, 2014). Define generalized

CDG-model as a 6-tuple (B, S, F, C, q, r) (C – choice functions, r – reservation prices = budgets in the terminology of Chen, Deng and Ghosh). Having reservation prices r we can define preference relation $P(r, s)$ for a seller s as:

$$b \; P(r, s) \;\; c \quad \Leftrightarrow \quad r(b, s) \;\geq\; r(c, s).$$

Definition 14. Let $M = (B, S, F, C, q, r)$ be a generalized CDG-model. We say that r are differentiated reservation prices if for any $b, c \in B$ and any $s \in S$ we have

$$b \neq c \quad \Rightarrow \quad r(b, s) \;\neq\; r(c, s)$$

(i.e. any two different buyers have different reservation prices for the contract with the same seller).

Obviously reservation prices are differentiated if and only if all preferences $P(r, s)$ are linear orders. Hence if we have a CDG-model (B, S, F, C, q, r) with differentiated r, then we can construct a generalized GS-model $(B, S, F, C, P(r), q, r)$ with $P(r)$ – family of linear orders $P(r, s)$. By theorem 7 we obtain the following result:

Theorem 8. Let $M = (B, S, F, C, q, r)$ be a generalized CDG-model with prices, with Plott choice functions C and differentiated reservation prices r. Then there exists a price equilibrium (u, p) for the model M.

5. CONCLUDING REMARKS

In our paper we have studied relationships between stable (strongly stable) matchings and price equilibria for generalized many-to-many Gale-Shapley market models with choice functions representing preferences of the buyers, weak orders representing preferences of the sellers and reservation prices of the buyers. We have shown that strongly stable matchings, under the assumptions of path independency of choice functions and very strong compatibility of reservation prices with the preferences of the sellers, are identical with price equilibria allocations. Unfortunately, in general, there is no similar characterization for stable matchings (in Świtalski, 2016 it is shown that stable matchings can be characterized by order equilibria). A special case in which such characterization for stable matchings is possible is the one with linear preferences of the sellers, because stability = strong stability in this case.

We have also shown how to use characterization results to prove existence of price equilibria for many-to-many GS-models with Plott (= path independent) choice functions and linear preferences of the sellers or for many-to-many Chen-Deng-Ghosh-models with Plott choice functions and differentiated reservation prices. Simple examples show that in the cases of non-linear preferences (for the GS-models) or non-differentiated prices (for the CDG-models) price equilibria may not exist.

We have used standard pairwise stability condition and this helped us to prove the existence result by using Alkan and Gale theory (2003). There are other stability conditions for many-to-many models (see, e.g. Echenique, Oviedo, 2006), but they are defined under the assumption that choice functions are generated by strict linear orders on the families of subsets of feasible contracts. It would be interesting to study relationships between stability and price equilibria for other stability concepts (different from pairwise stability) and for the model with general choice functions. But this could be the problem for further research.

A method of studying many-to-many model is the transforming of such model into equivalent many-to-one or one-to-many model by cloning the agents at one side of the market (under suitable assumptions). We did not follow this way, because we could directly use the results from our previous paper (Świtalski, 2016), but of course it would be interesting to study possibility of such transformation.

It would be very important for matching theory to have a model of matching market, in which we could prove relationships between stability and equilibria as general as possible. Our model cannot be embedded directly into contract theory, hence an interesting question would be the possibility of building a general model including our model and contract models (for example the model of Hatfield et al., 2013) for which similar results on stability and equilibria will hold.

REFERENCES

Aizerman M., Aleskerov F., (1995), *Theory of Choice*, North-Holland, Amsterdam.
Aleskerov F., Monjardet B., (2002), *Utility Maximization, Choice and Preference*, Springer, Berlin.
Alkan A., Gale D., (2003), Stable Schedule Matching under Revealed Preference, *Journal of Economic Theory*, 112, 289–306.
Andersson T., Erlanson A., (2013), Multi-item Vickrey-English-Dutch Auctions, *Games and Economic Behavior*, 81, 116–129.
Azevedo E. M., Leshno J. D., (2011), A Supply and Demand Framework for Two-sided Matching Markets, *working paper*, www.aeaweb.org (last access: 28.07.2016).
Biro P., Kiselgof S., (2013), College Admissions with Stable Score-Limits, *discussion paper*, Institute of Economics, Research Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budapest.
Camina E., (2006), A Generalized Assignment Game. *Mathematical Social Sciences*, 52, 152–161.
Chen N., Deng X., Ghosh A., (2014), Competitive Equilibria in Matching Markets with Budgets, *www.arxiv.org* (last access: 28.07.2016).
Crawford V. P., Knoer E. M., (1981), Job Matching with Heterogeneous Firms and Workers, *Econometrica*, 49, 437–450.
Danilov V., Koshevoy G., (2005), Mathematics of Plott Choice Functions, *Mathematical Social Sciences*, 49, 245–272.
Echenique F., (2007), Counting Combinatorial Choice Rules, *Games and Economic Behavior*, 58, 231–245.
Echenique F., Oviedo J., (2006), A Theory of Stability in Many-to-Many Matching Markets, *Theoretical Economics*, 1, 233–273.
Gale D., (1960), *The Theory of Linear Economic Models*, The University of Chicago Press, Chicago.

Gale D., Shapley L. S., (1962), College Admissions and the Stability of Marriage, *American Mathematical Monthly*, 69, 9–15.

Hatfield J. W., Kominers S. D., (2016) Contract Design and Stability in Many-to-Many Matching, *Games and Economic Behavior* (in press).

Hatfield J. W., Kominers S. D., Nichifor A., Ostrovsky M., Westkamp A., (2013), Stability and Competitive Equilibrium in Trading Networks, *Journal of Political Economy*, 121 (5), 966–1005.

Herings P. J., (2015), Equilibrium and Matching under Price Control, *working paper* RM/15/001, Maastricht University, Graduate School of Business and Economics.

Klaus B., Walzl M., (2009), Stable Many-to-Many Matchings with Contracts, *Journal of Mathematical Economics*, 45, 422–434.

Kominers S. D., (2012), On the Correspondence of Contracts to Salaries in (Many-to-Many) Matching, *Games and Economic Behavior*, 75, 984–989.

Roth A. E., (2002), The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics, *Econometrica*, 70, 1341–1378.

Roth A. E., Sotomayor M. A., (1992), *Two-Sided Matching. A Study in Game-Theoretic Modeling and Analysis*, Cambridge University Press.

Shapley L. S., Shubik M., (1971/72), The Assignment Game I: The Core, *International Journal of Game Theory*, 1, 111–130.

Sönmez T., Ünver M. U., (2011), Matching, Allocation and Exchange of Discrete Resources, in: Benhabib J., Bisin A., Jackson M. O., (eds.), *Handbook of Social Economics*, 781–852.

Sotomayor M., (2007), Connecting the Cooperative and Competitive Structures of the Multiple Partners Assignment Game, *Journal of Economic Theory*, 134, 155–174.

Świtalski Z., (2008), Stability and Equilibria in the Matching Models, *Scientific Research of the Institute of Mathematics and Computer Science, Częstochowa University of Technology*, 2 (7), 77–85.

Świtalski Z., (2010), Equilibria in Markets with Two-Sided Preferences (in Polish)*, Przegląd Statystyczny*, 4 (57), 54–69.

Świtalski Z., (2015), Some Properties of Competitive Equilibria and Stable Matchings in a Gale-Shapley Market Model, *Studia Ekonomiczne, Scientific Papers of the University of Economics, Katowice*, 248, 222–232.

Świtalski Z., (2016), Stability and Generalized Competitive Equilibria in a Many-to-Many Gale-Shapley Market Model, *Przegląd Statystyczny*, 3 (63), 237–253.

STABILNOŚĆ I RÓWNOWAGI CENOWE W MODELU RYNKU GALE'A-SHAPLEYA TYPU „MANY-TO-MANY"

S t r e s z c z e n i e

W artykule zbadano zależności między uogólnionymi równowagami konkurencyjnymi zdefiniowanymi w pracy Świtalskiego (2016), a równowagami cenowymi dla pewnego wariantu modelu rynku Gale'a-Shapleya (typu „many-to-many"), a także między równowagami cenowymi a skojarzeniami stabilnymi dla tego modelu. Uzyskane wyniki wykorzystano do udowodnienia twierdzeń o istnieniu równowag cenowych w modelu GS typu many-to-many oraz w pewnym modelu typu many-to-many uogólniającym model zawarty w pracy Chen i inni (2014).

Słowa kluczowe: skojarzenie stabilne, teoria Gale'a-Shapleya, model „many-to-many", równowaga cenowa, funkcje wyboru, dyskretny model rynku

STABILITY AND PRICE EQUILIBRIA
IN A MANY-TO-MANY GALE-SHAPLEY MARKET MODEL

Abstract

In the paper we study relationships between generalized competitive equilibria defined in the paper of Świtalski (2016) and price equilibria for some variant of many-to-many market model of Gale-Shapley type and between price equilibria and stable matchings for such a model. Obtained results are used for proving theorems on existence of price equilibria in the many-to-many GS-model and in the many-to-many model generalizing the model of Chen, Deng and Ghosh (Chen et al., 2014).

Keywords: stable matching, Gale-Shapley theory, many-to-many model, price equilibrium, choice function, discrete market model

AGNIESZKA LIPIETA[1]

ADJUSTMENT PROCESSES ON THE MARKET
WITH COUNTABLE NUMBER OF AGENTS AND COMMODITIES[2]

1. INTRODUCTION

Innovations are the driving force for the economic development (see Schumpeter, 1912), hence the modelling the structures convenient to analyzing innovative processes remains at the core of interest of the economic theory.

The introducing of new products, new technologies, new ways of production etc., can be easily noticeable during the analysis of commodity bundles and producers' plans of action in different points of time. Economic agents, operating on the market, can observe and retrieve the diversity of feasible goods as well as the structure of the supply and the demand. If we want to focus on producers' and consumers' characteristics, then it is convenient to use the Arrow and Debreu apparatus (see Arrow, Debreu, 1954; Debreu, 1959) to model the economic dependencies on such a period of time on which the activities of economic agents are not changed. Such set-up is also useful in formulating and proving the sufficient conditions for existence equilibrium in the private ownership economy (see Arrow, Debreu, 1954; Mas-Colell et al., 1995). However, modeling economic processes resulting in equilibrium needs to involve time.

Since years, many researches have been done to explain how an economy evolves over time. Evolution of economic structures can be caused, among others, by modification activities of economic agents, revealing in introducing new commodities, in increasing or decreasing in amounts of existed commodities, or in eliminating some goods from the market. The set of economic agents may be changed on the observable period of time as some of economic agents might enter or exit the market now or in the future. The above are taken into consideration in the model presented in the current paper.

Generally, the models of evolution of an economy can be divided into two groups: the models where time is the discrete valuable and the models with continuous time. To the first group belong the two-periods and the multi-periods economies under and without risk, as well as the models in which economic processes are modeled

by difference equations. Some results the reader can find, for example in Radner (1972), Magill, Quinzii (2002), Mas-Colell et al. (1995), Acemoglu (2009), Arrow, Intriligator (1987), Chiang (1992). The second group consists of the models in which the economic processes are examined by the use of differential equations. They are, among others, the Domar model, the classical Solow model, the Romer model as well as their modifications (see for example Romer, 2012; Acemoglu, 2009; Chiang, 1992; Malawski, 1999). Such approach is typical for the models studied in the growth theory.

Some results on the analysis of transitions of economic systems the reader can also find in Lipieta, Malawski (2016) and in Lipieta (2013). The examples of using difference equations in modelling some economic processes are presented, for instance, in Lipieta (2015, 2016).

The specific mathematical properties of the topological apparatus used by Kenneth Arrow and Gérard Debreu and separately by Lionel W. McKenzie (see also Panek, 1993) encourage to consider time in an economy defined with similar tools, especially that the analysis of Schumpeter's conceptions of the economic evolution (see for instance Schumpeter, 1912), leads to the Walras's approach in modelling innovative mechanisms (see also Shionoya, 2015; Lipieta, Malawski, 2016). The implementation of the Arrow and Debreu stationary economy into dynamic processes is not new as there are lots of papers devoted to that problem as well as lots of its solutions (for example Arrow, Intriligator, 1987; Ciałowicz, Malawski, 2011, 2017; Panek, 1997). However, there is no a coherent and unified model of economic evolution in the scientific literature, in which the innovative changes in an economy, could be model by the use of the Arrow and Debreu topological apparatus.

In this context, the aim of this paper is to determine a system of difference equations (see for instance Chiang, 1984) defined in the environment of an economy with countable number of agents and commodities. The above could be useful in modeling some aspects of economic life, especially innovative changes as well as so called adapting processes (see Andersen, 2009), which moves an economic system to equilibrium. In difference to the multi-periods economies, we model the situation in which the sets of commodities, consumers and producers (firms) can be changed on the analyzed time interval. In difference to the models in the growth theory, where the strong mathematical properties of the economic objects under study are required, the model presented in the current paper does not require additional mathematical assumptions. Therefore, the model of economic evolution presented in the paper can be used for exploration of many discontinuous processes such as innovative processes or the processes of bankruptcy of firms.

The paper consists of five parts: in the second part, the private ownership economy with countable number of agents and commodities is defined, the third part is devoted to modelling transformations of the above economy on a given time interval, defined by the use of the specific kind of dynamic system with discrete time. In the fourth part some qualitative properties of adjustment processes are specified while the five part contains conclusions.

<div align="center">

2. THE PRIVATE OWNERSHIP ECONOMY
WITH COUNTABLE NUMBER OF AGENTS AND COMMODITIES

</div>

The activities of two kinds of economic agents: producers and consumers are under our consideration. To emphasize the fact that the number of economic agents as well as that the number of commodities can be changed on the analyzed time interval, we construct the modification of the private ownership economy (see Debreu, 1959; Mas-Colell et al., 1995), to the economy with countable number of agents and commodities, defined in the form of the multi-range relational system (see Adamowicz, Zbierski, 1997; Malawski, 1999), analogously to the definition of the economy presented in Lipieta (2010). Let

  – $A = (a_i)_{i \in \mathbb{N}}$ – be a countable set of consumers,
  – $B = (b_j)_{j \in \mathbb{N}}$ – be a countable set of producers.

Hence $K \stackrel{\text{def}}{=} A \cup B$ is the countable set of economic agents. Let $\ell \in \{1, 2, \dots\}$,

$$\mathcal{R}^\ell \stackrel{\text{def}}{=} \mathbb{R}^\ell \times \{0\} \times \{0\} \times \dots$$

as well as

$$\mathcal{R} = \{(x_n)_{n \in \mathbb{N}} : \exists n_0 \in \mathbb{N} \; \forall n > n_0 \; x_n = 0\}. \tag{1}$$

For every $\ell \in \{1, 2, \dots\}$, $\mathcal{R}^\ell \subset \mathcal{R}$. Suppose that $\ell \in \{1, 2, \dots\}$ commodities are on the market. Producers' activities in space $\mathcal{R}^\ell$ with respect to achievable technologies are demonstrated by correspondence of production sets

$$y : B \ni b \to Y^b \subset \mathcal{R}^\ell,$$

which to every producer assigns his feasible plans of action. Moreover, it is assumed that

$$\exists n \in \{1, 2, \dots\} \; \forall j > n \; y(b_j) \stackrel{\text{def}}{=} \{0\},$$

what illustrates the assumption that a finite number of producers operate on the market while every producer b_j, for $j > n$, is an inactive producer at the given moment. Due to such set-up, it is underlined that an unknown number of producers might enter or exit the market in the future.

Assume that a price vector $p \in \mathcal{R}^\ell$ is given.

Definition 1. A two-range relational system

$$P_q^\ell = (B, \mathcal{R}; y, p)$$

is called the ℓ-dimensional quasi-production system.

In the quasi-production systems, the aim of producers is not specified, hence quasi-production systems could be regarded as the area for modeling the producers' activities under the perfect or the bounded rationality assumption (see Simon, 1955; Lipieta, Malawski, 2016).

Definition 2. The three-range relational system

$$C_q^\ell = (A, \mathcal{R}, \Xi; \chi, \epsilon, \varepsilon, p),$$

where:
- $\mathcal{R}$ is of the form (1),
- $\Xi \subset \mathcal{R}^\ell \times \mathcal{R}^\ell$ is the family of all preference relations in $\mathcal{R}^\ell$,
- $\chi: A \ni a \to \chi(a) = X^a \subset \mathcal{R}^\ell$ is the correspondence of consumptions sets, where

$$\exists\, m \in \{1, 2, \dots\}\ \forall i > m\ \chi(a_i) \overset{\text{def}}{=} \{0\},$$

which analogously means that every consumer a_i, for $i > m$, is an inactive consumer,
- $\epsilon: A \ni a \to \omega^a \in \mathcal{R}^\ell$ is the initial endowment mapping,
- $\varepsilon \subset A \times (\mathcal{R}^\ell \times \mathcal{R}^\ell)$ is the correspondence, which to every consumer $a \in A$ assigns a preference relation $\preccurlyeq^a$ from set Ξ restricted to set $\chi(a) \times \chi(a)$,

is called the ℓ-dimensional quasi-consumption system.

Definition 3. The structure

$$\mathcal{E}_q^\ell = \left(\mathcal{R}, P_q^\ell, C_q^\ell, \theta, \omega\right),$$

where:
- $\mathcal{R}$ is of the form (1),
- P_q^ℓ is the ℓ-dimensional quasi-production system,
- C_q^ℓ is the ℓ-dimensional quasi-consumption system,
- $\omega = \sum_{a \in A} \omega^a \in \mathcal{R}^\ell$,
- for $a \in A$ and $b \in B$, number $\theta(a, b)$ is the share of consumer a in the profit of producer b as well as mapping $\theta: A \times B \to [0,1]$ satisfies $\forall b \in B\ \sum_{a \in A} \theta(a, b) = 1$,

is called the ℓ-dimensional private ownership economy.

The number ℓ is called the dimension of the private ownership economy, the number of economic agents active on the market is not greater than $m + n$.

Definition 4. If $P_q^\ell = (B, \mathcal{R}; y, p)$ is the ℓ-dimensional quasi-production system, where

$$\forall\, b \in B\ \ \eta^b(p) \overset{\text{def}}{=} \{y^{b*} \in y(b): p \circ y^{b*} = \max\{p \circ y^b: y^b \in y(b)\}\} \neq \emptyset,$$

then

- $\eta: B \ni b \to \eta^b(p) \subset Y^b$ is called the correspondence of supply at price system p,
- $\pi: B \ni b \to \pi(b) = p \circ y^{b*} \in \mathbb{R}$ is called the maximal profit function at price system p,
- the quasi-production system P_q^{ℓ} is called the ℓ-dimensional production system and denoted by

$$P^{\ell} = (B, \mathcal{R}; y, p, \eta, \pi).$$

Let us notice that in contrast to quasi-production systems, in production systems the aim of producers is the profit maximization.

Definition 5. If, for every $a \in A$, at the given price vector $p \in \mathcal{R}^{\ell}$

$$\beta^a(p) = \{x \in \chi(a): p \circ x \leq p \circ \omega^a + \sum_{b \in B} \theta(a, b) \cdot \pi^b(p)\} \neq \emptyset$$

and

$$\varphi^a(p) = \{x^{a*} \in \beta^a(p): \forall x^a \in \beta^a(p) \; x^a \preccurlyeq^a x^{a*}, \preccurlyeq^a \in \Xi\} \neq \emptyset,$$

then

- $\beta: A \ni a \to \beta^a(p) \subset \mathcal{R}^{\ell}$ is the correspondence of budget sets at price system p, which to every consumer $a \in A$ assigns his set of budget constrains $\beta^a(p) \subset \chi(a)$ at price system p and initial endowment ω^a; number

$$w^a = p \circ \omega^a + \sum_{b \in B} \theta(a, b) \cdot \pi^b(p) \tag{2}$$

is called the wealth of consumer a,
- $\varphi: A \ni a \to \varphi^a(p) \subset \mathcal{R}^{\ell}$ is the demand correspondence at price system p, which to every consumer $a \in A$ assigns the consumption plans maximizing his preference on the budget set $\beta^a(p)$,
- the ℓ-dimensional quasi-consumption system C_q^{ℓ} is the ℓ-dimensional consumption system and is denoted by

$$C_q^{\ell} = C^{\ell} = (A, \mathcal{R}, \Xi; \chi, \epsilon, \varepsilon, p, \beta, \varphi).$$

We assume that consumers aim in the maximization of preferences on budget sets, however in quasi-consumption systems there may be no upper bound for a consumer's preference relation on the adequate budget set.

Definition 6 (see also in Lipieta, 2010). If P_q^{ℓ} is the ℓ-dimensional production system ($P_q^{\ell} = P^{\ell}$) and C_q^{ℓ} is the ℓ-dimensional consumption system ($C_q^{\ell} = C^{\ell}$), then the ℓ-dimensional private ownership economy $\mathcal{E}_q^{\ell}$ is called the ℓ-dimensional Debreu economy.

If economy $\mathcal{E}_q^\ell$ is the ℓ-dimensional Debreu economy, then we will write $\mathcal{E}_p^\ell = (\mathcal{R}, P^\ell, C^\ell, \theta, \omega)$ instead of $\mathcal{E}_q^\ell = (\mathcal{R}, P_q^\ell, C_q^\ell, \theta, \omega)$ or $\mathcal{E}_q^\ell = \mathcal{E}_p^\ell$. The commodity space of every ℓ-dimensional private ownership economy is the subset of the space of real sequences. In this meaning, every economy $\mathcal{E}_q^\ell$ can be viewed as the economy with the countable number of commodities. If $x^a \in X^a$ for every $a \in A$, $y^b \in Y^b$ for every $b \in B$ as well as

$$\sum_{a \in A} x^a - \sum_{b \in B} y^b = \omega,$$

then the sequence (x, y), where $x = (x^{a_1}, x^{a_2}, \dots)$, $y = (y^{b_1}, y^{b_2}, \dots)$, is called the feasible allocation. The sequence

$$(x^*, y^*, p), \tag{3}$$

where $x^* = (x^{a_1*}, x^{a_2*}, \dots)$, $y^* = (y^{b_1*}, y^{b_2*}, \dots)$, for which

- $\forall a \in A \; x^{a*} \in \varphi^a(p)$,
- $\forall b \in B \; y^{b*} \in \eta^b(p)$,
- $\sum_{a \in A} x^{a*} - \sum_{b \in B} y^{b*} = \omega$,

is called the state of equilibrium in economy $\mathcal{E}_p^\ell$. If there exists a state of equilibrium in economy $\mathcal{E}_p^\ell$, then we say that the economy $\mathcal{E}_p^\ell$ is in equilibrium as well as the price vector p is called the equilibrium price vector and is denoted by p^*.

3. ADJUSTMENT PROCESSES
IN THE ℓ-DIMENSIONAL PRIVATE OWNERSHIP ECONOMY

The definitions presented below are borrowed from Arrow, Intriligator (1987) and are adapted to the private ownership economy with the countable number of commodities. Let $\tau \in \{1, 2, \dots\}$ be the number of points of time indexed by t, $t = 0, 1, 2, \dots, \tau$. As in Lipieta (2015) and (2016), we say that the economic process is the sequence of actions of economic agents on time interval $[0, \tau]$, resulting in offered goods and services. The set of possible resource allocations will be denoted by Z.

The sequence of characteristics, determining an individual as agent $k \in K$ in the given economic process, is called the environment of that agent. The environment of agent k is denoted by e^k, whereas symbol E^k stands for the set of all his feasible environments. The set

$$E \stackrel{\text{def}}{=} E^{k_1} \times E^{k_2} \times \dots$$

is called the set of environments.

From now, if $\tau > 1$, then every natural number t such that $0 < t < \tau$, is identified with time interval $[t - 1, t)$ on which the activities of producers and consumers are constant. The lengths (ranges) of times intervals do not have to be equal. Saying "at time t", we mean "at time interval $[t - 1, t)$" for $0 < t < \tau$, or at the moment of time $t = 0$, or at the moment of time $t = \tau$.

By the fact that activities of producers and consumers are constant on the considered time intervals, we assume that the environment of every agent k is also constant on every time interval t. The environment of agent k at time interval t is denoted by $e^k(t)$. By the above

$$e^k(t) \in E^k, \ e^k(t) = const \text{ for } k \in K \text{ and } t \in \{1, 2, \dots, \tau\}.$$

The set of messages (information) to be used on the market by agent k is denoted by M^k. The messages of agent k are denoted by m^k ($m^k \in M^k$). Moreover, $m^k(t)$ means the message of agent k at time t. As in case of environments,

$$m^k(t) \in M^k, \ m^k(t) = const \text{ for } k \in K \text{ and } t \in \{1, 2, \dots, \tau\}.$$

The vector

$$m = (m^{k_1}, m^{k_2}, \dots)$$

is called the message, if $m^k \in M^k$ for every $k \in K$. Now, we can put the following definition:

Definition 7. The structure

$$(M, f, h), \tag{4}$$

where:
- $M \subset M^{k_1} \times M^{k_2} \times \dots$ is the set of messages,
- $f = (f^{k_1}, f^{k_2}, \dots): M \times E \to M$ is the response function, while $f^k: M \times E \to M^k$ is the response function of agent k,
- $m^k(t + 1) = f^k(m(t), e(t))$, $t = 1, \dots, \tau - 1$; $k \in K$ is the process of exchanging messages, where $m(t) = (m^{k_1}(t), m^{k_2}(t), \dots)$,
- $h: M \to Z$ is the outcome function, which to every message m assigns the allocation which are the result of analysis of the message m by economic agents,

is called the adjustment process on time $[0, \tau]$.

Definition 8. Let the structure (M, f, h) be an adjustment process. A message

$$\bar{m} = (\bar{m}^{k_1}, \bar{m}^{k_2}, \dots) \in M$$

is said to be stationary if, for every $k \in K$, it satisfies the equation $\bar{m}^k = f^k(\bar{m}, e)$.

Let (M, f, h) be an adjustment process of the form (4). If the components of the environment at time $t = 0$

$$e(0) = (e^{k_1}(0), e^{k_2}(0), \dots) \in E$$

form the Debreu economy, then the adjustment process (4) is called the adjustment process in the Debreu economy (see also Lipieta, 2016).

We aim in modeling adjustment processes of the economic evolution using the Arrow-Debreu apparatus (see for example Arrow, Debreu, 1954), however we admit that the number of commodities, the number of active economic agents as well as the plans of action of economic agents can be changed in time.

Let $\tau \in \{1, 2, \dots\}$. Number $\ell_t \in \{1, 2, \dots\}$ means the dimension of the commodity-price space at time $t \in \{0, 1, \dots, \tau\}$. We admit that if a commodity $l \in \{1, \dots, \ell_t\}$ is not used in production or consumption at time $t + 1$, then in every producers' and consumers' plans of action at time $t + 1$, l-th coordinate is equal to zero. Under such assumption, we can assume that $\ell = \ell_0 \leq \ell_1 \leq \cdots \leq \ell_\tau$.

Let $\mathcal{E}_q^{\ell_t} = (\mathcal{R}, P_q^{\ell_t}, C_q^{\ell_t}, \theta, \omega)$ be the ℓ_t-dimensional private ownership economy. If an economic agent active at time t disappears from the market at time $t + 1$, then he becomes the inactive agent (a producer or a consumer) with zero plans of action at time $t + 1$. If an economic agent with zero plan of action at time t (so the inactive agent at time t) enters the market at time $t + 1$, then he becomes the active agent with non-zero plans of action. For every t and ℓ_t:

– $Y^b(t) \subset \mathcal{R}^{\ell_t} \subset \mathcal{R}^{\ell_\tau} \subset \mathcal{R}$ means the set of plans of action of producer $b \in B$, feasible to realization at time t,

– $y^b(t)$ – the plan of producer b realized at time t, $y^b(t) \in Y^b(t)$.

In the same way, the characteristics of consumers: $X^a(t) \subset \mathcal{R}^{\ell_\tau}$ and $\epsilon_t(a) \in \mathcal{R}^{\ell_t} \subset \mathcal{R}^{\ell_\tau}$ at time t, for $a \in A$, are defined. The correspondence of preference relations at time t is denoted by $\varepsilon_t(a) = \preccurlyeq_{\ell_t}^a$, where $\preccurlyeq_{\ell_t}^a \subset X^a(t) \times X^a(t) \subset \mathcal{R}^{\ell_t} \times \mathcal{R}^{\ell_t} \subset \mathcal{R}^{\ell_\tau} \times \mathcal{R}^{\ell_\tau} \subset \mathcal{R} \times \mathcal{R}$ means the preference relation of consumer a at time t.

On the basis of the above notation, the environment $e^k(t)$ of every economic agent $k \in K = A \cup B$ at time t is defined. Namely

$$e^k(t) = \left(\tilde{y}_t(k), \tilde{\chi}_t(k), \tilde{\epsilon}_t(k), \tilde{\varepsilon}_t(k), \tilde{\theta}_t(k, \cdot) \right), \tag{5}$$

where:

– $\tilde{y}_t(k) = Y^k(t)$ for $k \in B$, $\tilde{y}_t(k) = \{0\}$ for $k \notin B$,

– $\tilde{\chi}_t(k) = X^k(t)$ for $k \in A$, $\tilde{\chi}_t(k) = \{0\}$ for $k \notin A$,

– $\tilde{\epsilon}_t(k) = \omega^k$ for $k \in A$, $\tilde{\epsilon}(k) = 0$ for $k \notin A$,

– $\tilde{\varepsilon}_t(k) = \preccurlyeq_{\ell_t}^a$ for $k \in A$, $\tilde{\varepsilon}(k) = \{\emptyset\}$ for $k \notin A$,

– the mapping $\tilde{\theta} : K \times K \to [0,1]$ satisfies:
  $\tilde{\theta}(k, \cdot) \equiv 0$ for $k \notin A$, $\tilde{\theta}(\cdot, k) \equiv 0$ for $k \notin B$ and $\forall b \in B \ \sum_{a \in A} \tilde{\theta}(a, b) = 1$;

moreover, for $a \in A$ and $b \in B$, number $\tilde{\theta}(a, b)$ is the share of consumer a in the profit of producer b.

By the above, the set of environments E^k of every agent $k \in K$ is of the form

$$E^k = P(\mathcal{R}) \times P(\mathcal{R}) \times \mathcal{R} \times P(\mathcal{R} \times \mathcal{R}) \times \mathcal{F}(K, [0,1]),$$

with $\mathcal{F}(K, [0,1]) \overset{\text{def}}{=} \{f \mid f: K \to [0,1]\}$. The set of environment is given by

$$E \overset{\text{def}}{=} E^{k_1} \times E^{k_2} \times \dots \ .$$

The rest of components of the adjustment process in the meaning of Definition 7 is defined in the standard way (compare to Arrow, Intriligator, 1987). Namely, the message of every agent $k \in K$ at time $t = 0,1, \dots \tau$, is understood as:

$$m^k(t) \overset{\text{def}}{=} (p(t), y^k(t), x^k(t)), \tag{6}$$

where $x^k(t) = 0 \in \mathcal{R}$ for $k \notin A$ and $y^k(t) = 0 \in \mathcal{R}$ for $k \notin B$.

Consequently, $M^k = \mathcal{R} \times \mathcal{R} \times \mathcal{R}$. Define $M \subset M^{k_1} \times M^{k_2} \times \dots$ by the formula

$$M \overset{\text{def}}{=} \left\{ \begin{array}{c} (m^{k_1}(t), m^k(t), \dots): \sum_{k \in K} x^k(t) - \sum_{k \in K} y^k(t) = \omega(t) \\ \wedge \, \forall t = 0, 1, \dots, \tau \, \exists \, p(t) \in \mathcal{R} \, \forall k \in K: m^k(t) = (p(t), x^k(t), y^k(t)) \end{array} \right\}.$$

Suppose that $M \neq \emptyset$. Every message $m \in M$ has to be a feasible message at any time $t \in \{0,1, \dots, \tau\}$. Hence $m = m(t)$ for $t \in \{0,1, \dots, \tau\}$. The response function of every agent k to the message $m(t) \in M$, for $t = 0, 1, \dots, \tau - 1$, is of the form:

$$f^k(m(t), e(t)) = (p(t+1), y^k(t+1), x^k(t+1)). \tag{7}$$

As a reply to prices $p(t)$ at time t, every agent k chooses his plan of action at time $t = 0, 1, \dots, \tau$. If $x(t) = (x^{a_1}(t), x^{a_2}(t), \dots)$ and $y(t) = (y^{b_1}(t), y^{b_2}(t), \dots)$, then

$$Z \overset{\text{def}}{=} \left\{ \begin{array}{c} (x, y): \exists \, t = 0, 1, \dots, \tau \, (x = x(t) \, \wedge \, y = y(t)) \\ \sum_{a \in A} x^a(t) - \sum_{b \in B} y^b(t) = \omega \end{array} \right\}. \tag{8}$$

In this situation, the outcome function $h: M \to Z$ is of the form:

$$h(m) = h(m^{k_1}(t), m^k(t), \dots) =$$

$$= h\left(\left(p(t), y^{k_1}(t), x^{k_1}(t) \right), \left(p(t), y^{k_2}(t), x^{k_2}(t) \right), \dots \right) \overset{\text{def}}{=} \tag{9}$$

$$= \left((x^{k_1}(t), x^{k_2}(t), \dots), (y^{k_1}(t), y^{k_2}(t), \dots) \right).$$

Precisely, function h assigns to a message at time t the sequence of feasible allocation of economic agents transferred by this messages.

If a message $m(t)$, for a $t \in \{0, 1, \ldots, \tau - 1\}$, is the stationary one, then $p(t) = p(t+1)$, and the state

$$(x(t), y(t), p(t))$$

has to be the state of equilibrium in economy $\mathcal{E}_q^{\ell_t}$. Consequently, the economy $\mathcal{E}_q^{\ell_t}$ is the private ownership Debreu economy ($\mathcal{E}_q^{\ell_t} = \mathcal{E}_p^{\ell_t}$). If there is $t_0 \in \{0, 1, \ldots, \tau - 1\}$ such that message $m(t_0)$ defined in (6) is stationary, then for every $t \in \{t_0, \ldots, \tau\}$ messages $m(t), \ldots, m(\tau - 1)$ are stationary as well as $m(t) = \cdots = m(\tau)$. Consequently, the economies $\mathcal{E}_q^{\ell_t}$, for $t \in \{t_0, \ldots, \tau\}$, are the Debreu economies in equilibrium ($\mathcal{E}_q^{\ell_t} = \mathcal{E}_p^{\ell_t}$).

Definition 9. An adjustment process (4) with the environments (5), the messages of the form (6), the response functions defined in (7) and the outcome function (8), is called the transformation process of economy $\mathcal{E}_q^{\ell_0}$.

If ℓ-dimensional private ownership economy $\mathcal{E}_q$, is built by the components of an environment $e(\tau)$ of the transformation process (4) of economy $\mathcal{E}_q^{\ell_0}$, then $\ell = \ell_\tau$, $\mathcal{E}_q = \mathcal{E}_q^{\ell_\tau}$ as well as $\mathcal{E}_p$ is said to be the transformation (or the evolution) of economy $\mathcal{E}_q^{\ell_0}$. This relationship will be noted by $\mathcal{E}_q^{\ell_0} \subset \mathcal{E}_q^{\ell_\tau} = \mathcal{E}_q$.

The transformation process of private ownership economy $\mathcal{E}_q^{\ell_0}$ can be used for modelling the Schumpeterian vision of economic development. Namely, if

$$\exists b_0 \in B \ \exists y^{b_0}(\tau) \ \forall b \in B \ y^{b_0}(\tau) \notin \bigcup_{b \in B} Y^b(0), \tag{10}$$

then the innovative changes are noticeable during the transformation process. If $\ell_0 = \ell_\tau$, then the above condition means that at least one new technology reveals in producers' activities in the framework of the economy $\mathcal{E}_q^{\ell_\tau}$ in comparison to economy $\mathcal{E}_q^{\ell_0}$. If $\ell_0 < \ell_\tau$, then at least one new product or new technology appear in the final economy $\mathcal{E}_q^{\ell_\tau}$ in comparison to initial economy $\mathcal{E}_q^{\ell_0}$. The producer b_0 satisfying condition (10) is called the innovator. If the profit of innovator b_0 realized in the economy $\mathcal{E}_q^{\ell_\tau}$ is greater than in initial economy $\mathcal{E}_q^{\ell_0}$, then it is said that innovator b_0 is the successful innovator (see also Lipieta, 2013). More about innovations and innovative changes modeled in the Arrow-Debreu apparatus, the reader can find, for example in Malawski (2013) or in Lipieta, Malawski (2016).

The above defined transformation process in the private ownership economy can also be used to model the procedure of adjustment producers' or consumers' plans of action as well as prices to equilibrium, without changing the set of commodities. Such an adjustment process can be viewed as the adapting process (see Andersen, 2009), during which economic agents adapt innovations and which results in a new

state of equilibrium in the final transformation of the economy under study. Formally, the transformation process (4) of economy $\mathcal{E}_q^{\ell_0}$ is called the adapting process, if $\mathcal{E}_q^{\ell_\tau}$ is the Debreu economy and $\ell_0 = \ell_\tau$.

4. COMPARATIVE ANALYSIS OF TRANSFORMATIONS OF A DEBREU ECONOMY

Now we face the challenge of formulating criteria for comparing the transformation processes of a given initial private ownership economy $\mathcal{E}_q^{\ell_0}$. The transformation processes on the same time interval can be compared on the basis of qualitative properties of the final private ownership economies – built by components of environments at time $t = \tau$.

As earlier, the moment of time $t = 0$ is the starting point, number $\tau \in \{1,2,\dots\}$ – the ending point of two transformation processes (M, f, h) and $(\widetilde{M}, \tilde{f}, \tilde{h})$ of given economy $\mathcal{E}_q^{\ell_0}$. Assume that the components of environments $e(\tau)$ and $\tilde{e}(\tau)$ of transformation processes (M, f, h) and $(\widetilde{M}, \tilde{f}, \tilde{h})$ form the private ownership economies $\mathcal{E}_q^{\ell_\tau}$ and $\mathcal{E}_q^{\tilde{\ell}_\tau}$. Suppose firstly that economies $\mathcal{E}_q^{\ell_\tau}$ and $\mathcal{E}_q^{\tilde{\ell}_\tau}$ are the Debreu economies with states of equilibrium (see (3)) denoted by (x^*, y^*, p^*) and $(\tilde{x}^*, \tilde{y}^*, \tilde{p}^*)$, adequately.

Similarly as in Lipieta (2013), we say that a producer $b \in B$ is better off in Debreu economy $\mathcal{E}_p^{\tilde{\ell}_\tau}$ than in Debreu economy $\mathcal{E}_p^{\ell_\tau}$ if and only if,

$$p^*(\tau) \circ y^{b*}(\tau) < \tilde{p}^*(\tau) \circ \tilde{y}^{b*}(\tau). \tag{11}$$

Condition (11) means that the maximal profit of producer b in economy $\mathcal{E}_p^{\tilde{\ell}_\tau}$ is greater than in economy $\mathcal{E}_p^{\ell_\tau}$. In contrast to Lipieta (2013) it is said that a consumer $a \in A$ is better off in Debreu economy $\mathcal{E}_p^{\tilde{\ell}_\tau}$ than in Debreu economy $\mathcal{E}_p^{\ell_\tau}$ if and only if,

$$p^*(\tau) \circ x^{a*}(\tau) < \tilde{p}^*(\tau) \circ \tilde{x}^{a*}(\tau). \tag{12}$$

If there are the states of equilibrium in Debreu economies (see (3)) $\mathcal{E}_p^{\ell_\tau}$ and $\mathcal{E}_p^{\tilde{\ell}_\tau}$, then

$$p^*(\tau) \circ x^{a*}(\tau) = w^a(\tau) \text{ as well as } \tilde{p}^*(\tau) \circ \tilde{x}^{a*}(\tau) = \widetilde{w}^a(\tau)$$

(see (2)). Hence, condition (12) means that the wealth of consumer a in economy $\mathcal{E}_p^{\tilde{\ell}_\tau}$ is greater than in economy $\mathcal{E}_p^{\ell_\tau}$. The wealth of the Debreu economy $\mathcal{E}_p^{\ell_\tau}$, namely number

$$w(\tau) = \sum_{a \in A} w^a(\tau)$$

can be viewed as the result of consumers' wealths. Hence, we say that economy $\mathcal{E}_p^{\tilde{\ell}_\tau}$ is better off than economy $\mathcal{E}_p^{\ell_\tau}$ if

$$\sum_{a \in A} w^a(\tau) < \sum_{a \in A} \widetilde{w}^a(\tau). \tag{13}$$

Let (M, f, h) and $(\widetilde{M}, \tilde{f}, \tilde{h})$ be two adjustment processes of given Debreu economy $\mathcal{E}_p^{\ell_0}$ on time interval $[0, \tau]$. It is said that the adjustment process $(\widetilde{M}, \tilde{f}, \tilde{h})$ is more effective than adjustment process (M, f, h), if economy $\mathcal{E}_p^{\tilde{\ell}_\tau}$ is better off than economy $\mathcal{E}_p^{\ell_\tau}$.

By (2), condition (13) is equivalent to the following:

$$p^* \circ \omega + \sum_{b \in B} p^* \circ y^{b*} < \tilde{p}^* \circ \widetilde{\omega} + \sum_{b \in B} \tilde{p}^* \circ \tilde{y}^{b*}.$$

The above inequality means that the consumers' wealth depends on the size of producers' profits and the wealth of total endowment in the final economy.

If $\mathcal{E}_q^{\ell_\tau}$ is not the Debreu economy, then we put in criterion (12) the realized allocation $x^a(\tau)$ instead of equilibrium consumption plan $x^{a*}(\tau)$. Similarly in (11), equilibrium production plan $y^{b*}(\tau)$ is replaced by realized production plan $y^b(\tau)$. If $\mathcal{E}_q^{\tilde{\ell}_\tau}$ is not the Debreu economy, then it is done in the same way.

On the basis of the above, it is said that
- a producer $b \in B$ is better off in an economy $\mathcal{E}_q^{\tilde{\ell}_\tau}$ than in an economy $\mathcal{E}_q^{\ell_\tau}$, if and only if,

$$p(\tau) \circ y^b(\tau) < \tilde{p}(\tau) \circ \tilde{y}^b(\tau),$$

- a consumer $a \in A$ is better off in an economy $\mathcal{E}_q^{\tilde{\ell}_\tau}$ than in an economy $\mathcal{E}_q^{\ell_\tau}$, if and only if,

$$p(\tau) \circ x^a(\tau) < \tilde{p}(\tau) \circ \tilde{x}^a(\tau).$$

Now

$$p(\tau) \circ x^a(\tau) = w^a(\tau) \text{ as well as } \tilde{p}(\tau) \circ \tilde{x}^a(\tau) = \tilde{w}^a(\tau),$$

and, as above, we say that economy $\mathcal{E}_q^{\tilde{\ell}_\tau}$ is better off than economy $\mathcal{E}_q^{\ell_\tau}$, if

$$\sum_{a \in A} w^a(\tau) < \sum_{a \in A} \tilde{w}^a(\tau).$$

At the end let us notice that the adjustment process and consequently the transformation process of the economy $\mathcal{E}_p^{\ell_0}$ are also the economic mechanism in the sense of Hurwicz (see also Arrow, Intriligator, 1987; Hurwicz, Reiter, 2006). Hence in the same way as for economic mechanisms, we can say about qualitative properties for adjustment processes (see Lipieta, 2013; Lipieta, Malawski, 2016). Namely:

Definition 10. An adjustment process, in which prices of commodities are elements of the message space is called the price adjustment process. If in the given adjustment process at least one agent from the given set will be better off due to a given criterion,

without making the rest of agents (from this set) worse off, then this adjustment process will be called the qualitative one with respect to the given set.

On the basis of the above, we can say that the transformation process of a private ownership economy (see Definition 9) is the price adjustment process. Moreover, if the final economy is the innovative extension of the initial economy (see Lipieta, 2013), then the transformation process of the initial economy is the qualitative adjustment process with respect to the set of successful innovators.

5. CONCLUSIONS

The modifications of the definitions of production and consumption systems – the components of the Debreu economy presented in part 2 can simplify comparing of two Debreu economies with the same set of economic agents. Such two structures can be interpreted as the mathematical models of a real economy in two points of time where "the subsequent" economy can be understood as the transformation of "the earlier" economy.

On the other hand, the transformation process of a private ownership economy defined in part 3 is an attempt to put the initial stationary model "in motion" to make it possible to study changes in the economy modeled in the Arrow-Debreu apparatus.

The part three seems to be the basis for further studying of properties of transformation processes as it contains the criteria for the choice of the best or at least "good enough" process from the point of view of producers or consumers.

REFERENCES

Acemoglu D., (2009), *Introduction to Modern Economic Growth*, Princeton University Press, Princeton and Oxford.

Adamowicz Z., Zbierski P., (1997), *Logic of Mathematics: A Modern Course of Classical Logic*, Springer.

Andersen E. S., (2009), *Schumpeter's Evolutionary Economics*, Anthem Press, London.

Arrow K. J., Debreu G., (1954), Existence of an Equilibrium for a Competitive Economy, *Econometrica*, 22, 265–290.

Arrow K. J., Intriligator M. D., (eds.), (1987), *Handbook of Mathematical Economics*, 3, Amsterdam, North-Holland.

Ciałowicz B., Malawski A., (2011), The Role of Banks in the Schumpeterian Innovative Evolution an Axiomatic Set-up, in: Pyka A., Derengowski F., da Graca M., (eds.), *Catching Up, Spillovers and Innovations Networks in a Schumpeterian Perspective*, Springer, Heidelberg, Dordrecht, London, New York.

Ciałowicz B., Malawski A., (2017), Innovativeness of Banks as a Driver of Social Welfare, *Central European Journal of Economic Modelling and Econometrics*, 9 (2), 97–113.

Chiang A. C., (1984), *Fundamental methods of mathematical economic*, McGraw-Hill, Inc.

Chiang A. C., (1992), *Elements of Dynamic Optimization*, New York: McGraw-Hill, Inc.

Debreu G., (1959), *Theory of Value*, New York, Viley.

Hurwicz L., Reiter S., (2006), *Designing Economic Mechanism*, Cambridge University Press, New York.

Lipieta A., (2010), The Debreu Private Ownership Economy with Complementary Commodities and Prices, *Economic Modelling*, 27, 22–27.

Lipieta A., (2013), Mechanisms of Schumpeterian Evolution in: Malawski A., (ed.), *Innovative Economy as the Object Investigation in Theoretical Economics*, Cracow University of Economics Press, 94–119.

Lipieta A., (2015), Existence and Uniqueness of the Producers' Optimal Adjustment Trajectory in the Debreu-Type Economy, *Mathematical Economics*, 11 (18), 55–68.

Lipieta A., (2016), Adjustment Processes in the Debreu-Type Economy, in: Jurek W., (ed.), *Matematyka i informatyka na usługach ekonomii. Wybrane współczesne problemy wzrostu gospodarczego informatyki ekonomicznej*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, 75–86.

Lipieta A., Malawski A., (2016), Price Versus Quality Competition: In Search for Schumpeterian Evolution Mechanisms, *Journal of Evolutionary Economics*, 26 (5), 1137–1171.

Magill M., Quinzii M., (2002), *Theory of Incomplete Markets*, MIT Press, Cambridge.

Malawski A., (1999), *Metoda Aksjomatyczna w Ekonomii*, Ossolineum, Wrocław.

Malawski A., (ed.), (2013), *Innovative Economy as the Object Investigation in Theoretical Economics*, Cracow University of Economics Press.

Mas-Colell A., Whinston M. D., Green J. R., (1995), *Microeconomic Theory*, Oxford University Press, New York.

Panek E., (1993), *Elementy ekonomii matematycznej. Statyka*, Wydawnictwo Naukowe PWN, Warszawa.

Panek E., (1997), *Elementy ekonomii matematycznej. Równowaga i wzrost*, Wydawnictwo Naukowe PWN, Warszawa.

Radner R., (1972), Existence of Equilibrium of Plans, Prices and Price Expectations in a Sequence of Markets, *Econometrica*, 40 (2), 289–303.

Romer D., (2012), *Advanced Macroeconomics*, McGraw-Hill (4th edition).

Schumpeter J. A., (1912), *Die Theorie der wirtschaftlichen Entwicklung*, Leipzig: Duncker & Humblot, English translations: *The Theory of Economic Development*, Cambridge, MA: Harvard University Press 1934 and *A Galaxy Book*, New York, Oxford University Press 1961.

Shionoya Y., (2015), Schumpeter and Evolution: an Ontological Exploration. http://www.lib.hit-u.ac.jp/service/tenji/amjas/Shionoya18.pdf. (accessed 29.12.2015).

Simon H. A., (1955), A Behavioural Model of Rational Choice, *Quarterly Journal of Economics*, 69, 99–118.

PROCESY DOSTOSOWAWCZE NA RYNKU
Z PRZELICZALNĄ LICZBĄ AGENTÓW I TOWARÓW

S t r e s z c z e n i e

Większość składowych ekonomii z własnością prywatną to odwzorowania niezależne od czasu, choć opisują działania podmiotów gospodarczych rozgrywające się w czasie. Dlatego struktura ta jest interpretowana jako stacjonarny model gospodarki, w której działalność podmiotów ekonomicznych na rynkach jest stała w analizowanym przedziale czasu. Matematyczne własności przestrzeni towarów i cen ekonomii z własnością prywatną mogłyby być przydatne w analizie zmian działalności agentów ekonomicznych. Stąd potrzeba określenia w jaki sposób ekonomia z własnością prywatną mogłaby ewoluować w czasie.

W tym kontekście celem artykułu jest modelowanie ewolucji gospodarki zdefiniowanej w aparacie pojęciowym Arrowa i Debreu, z wykorzystaniem równań różnicowych. W rezultacie otrzymujemy spójny i jednolity opis tej ewolucji, który może być zastosowany, m.in., do analizy mechanizmów schumpeterowskiego rozwoju gospodarczego, odmiennie od metod używanych zwykle w teorii wzrostu.

Słowa kluczowe: ekonomia z własnością prywatną, procesy dostosowawcze

ADJUSTMENT PROCESSES ON THE MARKET
WITH COUNTABLE NUMBER OF AGENTS AND COMMODITIES

A b s t r a c t

Most components of the private ownership economy are the mappings independent on time, although they model activities of economic agents which take place in time. Therefore this structure is interpreted as the stationary model in which actions of economic agents on the market are constant on the analyzed time interval. The mathematical properties of the commodity-price space of the private ownership economy could be convenient in analyzing changes in the activities of economic agents. Hence, there is a need to determine how a private ownership economy could evolve over time.

In this context, the aim of the paper is to model evolution of the economy defined in the Arrow and Debreu apparatus by the use of difference equations. As a result, we get a coherent and unified description of the evolution of an economy that can be used, among others, in the analysis of the mechanisms of Schumpeter's economic development, differently from the methods usually used in the growth theory.

Keywords: private ownership economy, adjustment process

MACIEJ NOWAK[1], TADEUSZ TRZASKALIK[2]

QUASI-HIERARCHICAL APPROACH TO DISCRETE MULTIOBJECTIVE STOCHASTIC DYNAMIC PROGRAMMING[3]

1. INTRODUCTION

Many decision problems are dynamic by their very nature. In such cases the decision is not made once, but many times. Partial choices are mutually related, since earlier decisions influence which decisions can be considered in the consecutive stages of the process.

The consequences of decisions become apparent in the near or remote future, which is uncertain by its very nature. Precise assessment of the results of the choices made is usually not possible. The information which is at the disposal of the decision maker is much more often incomplete and fragmentary. In such a situation he or she should, as far as possible, expand his/her knowledge of the problem under investigation. Although it is usually not possible to obtain data allowing to apply a deterministic model, these efforts can result in a partial knowledge thanks to which it is possible to estimate probability distributions describing values of the criteria obtained for the decision alternatives under consideration. In such cases we deal with what is called in the literature the problem of decision making under risk.

In such situations we can apply methods using discrete stochastic dynamic programming approach based on Bellman's optimality principle (Bellman, 1957). For these processes it is characteristic that at the beginning of each stage, the decision process is in a certain state. In each state, a set of feasible decisions is available. The process is discrete when all sets of states and decisions are finite. These processes are stochastic which means that the probability of achieving the final state for the given stage is known when at the beginning of this stage the process was in one of the admissible states and when a feasible decision has been made.

---

[1] University of Economics in Katowice, Faculty of Informatics and Communication, Institute of Analysis and Managerial Decision Support, Department of Operations Research, St. 1 Maja 50, 40-287 Katowice, Poland, corresponding author – e-mail: maciej.nowak@ue.katowice.pl.

[2] University of Economics in Katowice, Faculty of Informatics and Communication, Institute of Analysis and Managerial Decision Support, Department of Operations Research, St. 1 Maja 50, 40-287 Katowice, Poland.

We will consider additive multi-criteria processes. At each stage, we estimate the realisation of the process using stage criteria. The sum of the stage criteria gives the value of the multi-stage criterion. In the classical approach, the task consists in obtaining a strategy for which the expected value of the given criterion is optimal. Multi-criteria problems can be regarded as hierarchical problems. This means that the decision maker is able to formulate a hierarchy of criteria so that the most important criterion is assigned the number 1; the number 2 is reserved for the second-most important criterion, and so on. We assume that all criteria considered in the problem can be numbered in this way.

Usually we solve the hierarchical problem sequentially. First we find the set of solutions which are optimal with respect to the most important criterion. Out of this set, we select the subset of solutions which are optimal with respect to the criterion number 2. We continue this procedure until we determine the subset of solutions which are optimal with respect to the least important criterion.

The hierarchical approach has a certain essential shortcoming. It turns out that very often the subset of solutions, obtained when an important criterion in the hierarchy is considered, has only one element. As a result, the selection of the solution with respect to less important criteria is determined and these criteria do not play an essential role in the process of determining the final solution. It is why a quasi-hierarchy approach is proposed for solving hierarchical problems.

The quasi-hierarchical approach to hierarchical multi-objective stochastic programming seems quite new. Below we list some related theoretical and application papers.

Elmaghraby (1970) discusses some models most often encountered in Management Science applications: the shortest path problem between two specified nodes; the shortest distance matrix; as well as the special case of directed acyclic networks. One of related topics is finding the k-th shortest path.

The extension of the approach proposed above to discrete (deterministic) dynamic programming problem can be found in Trzaskalik (1990). The algorithm described there is applied to solve hierarchical deterministic dynamic programming problem.

Tempelmeier, Hilger (2015) consider the stochastic dynamic lot sizing problem with multiple items and limited capacity under two types of fill rate constraints. It is assumed that according to the static-uncertainty strategy, the production periods as well as the lot sizes are fixed in advance for the entire planning horizon and are executed regardless of the realisation of the demands.

Woerner et al. (2015) analyse Markov Decision Processes over compact state and action spaces. They investigate the special case of linear dynamics and piecewise-linear and convex immediate costs for the average cost criterion. This model is very general and covers many interesting examples, for instance in inventory management.

Shapiro (2012) analyse relations between the minimax, risk averse and nested formulations of multi-stage stochastic programming problems. In particular, it discusses conditions for time consistency of such formulations of stochastic problems.

Topaloglou et al. (2008) develop a multi-stage stochastic programming model for international portfolio management in a dynamic setting. They consider portfolio rebalancing decisions over multiple periods in accordance with the contingencies of the scenario tree. The solution jointly determines capital allocations to international markets, the selection of assets within each market, and appropriate currency hedging levels.

Hatzakis, Wallace (2006) describe a forward-looking approach for the solution of dynamic (time-changing) problems using evolutionary algorithms. The main idea of the proposed method is to combine a forecasting technique with an evolutionary algorithm. The location, in variable space, of the optimal solution (or of the Pareto optimal set in multi-objective problems) is estimated using a forecasting.

Dempster (2006) gives a comprehensive treatment of EVPI-based sequential importance sampling algorithms for multi-stage, dynamic stochastic programming problems. Both theory and computational algorithms are discussed.

Bakker et al. (2005) analyse the problem of robot planning (e.g. for navigation) with hierarchical maps. The authors present an algorithm for hierarchical path planning for stochastic tasks, based on Markov decision processes and dynamic programming.

Sethi et al. (2002) review the research devoted to proving that a hierarchy based on the frequencies of occurrence of different types of events in the systems results in decisions that are asymptotically optimal as the rates of some events become large compared to those of others. The paper also reviews the research on stochastic optimal control problems associated with manufacturing systems, their dynamic programming equations, existence of solutions of these equations, and verification theorems of optimality for the systems.

Sethi, Zhang (1994) present an asymptotic analysis of hierarchical manufacturing systems with stochastic demand and machines subject to breakdown and repair as the rate of change in machine states approaches infinity.

Daellenbach, De Kluyver (1980) present and illustrate a technique for finding MINSUM and MINMAX solutions to multi-criteria decision problems, called Multi Objective Dynamic Programming, capable of handling a wide range of linear, nonlinear, deterministic and stochastic multi-criteria decision problems. Multiple objectives are considered by defining an adjoin state space and solving an $(N + 1)$ terminal optimisation problem.

Two monographs Trzaskalik (1991, 1998) are also worth mentioning here. They present proposals for formulating and solving hierarchical problems of multiobjective dynamic programming approached deterministically.

In our paper we present a method based on a quasi-hierarchical approach. We assume that the decision maker is able to define a hierarchy of criteria and to determine the extent to which the optimal value of a higher-priority criterion can be made worse in order to improve the value of lower-priority criteria. To find the final solution of the problem, we start with determining the solutions for which the criteria take values no lower than the thresholds determined by the decision maker. Next, we use the criteria hierarchy to determine the optimal solution of the problem.

The main algorithm presented in this paper is based on the observation used previously in Trzaskalik (1991). Let us note that, except for the case of alternative solutions, when the optimal strategy is modified by changing the decision in any feasible process state, the expected value of the given criterion deteriorates. Therefore, it is necessary to consider all the strategies that differ from the optimal strategy in one of the feasible states and to select those which are within a determined tolerance interval. One should then analyse again the strategies found, changing the value in one of the feasible states. This process should be continued as long as it is possible to change the strategy for one of the feasible states, which provides a new strategy with the expected value of the realisation within the given tolerance interval.

The main idea of the approach discussed in the paper was previously presented on International Symposium of Management Engineering ISME 2015 Kitakyushu, Japan and International Conference of German, Austrian and Swiss Operations Research Societies (GOR, OGOR, SVOR/ASRO), University of Vienna, Austria, 2015 (Nowak, Trzaskalik, 2017). The final version of the paper, presented below includes full literature review, revised algorithms and detailed description of illustrative examples, not published before.

2. SINGLE-CRITERION STOCHASTIC DYNAMIC PROGRAMMING

We will use the following notation (Trzaskalik, 1991, 1998):

T – number of stages of the decision process under consideration,

y_t – state of the process at the beginning of stage t ($t+1,…,T$),

Y_t – finite set of process states at stage t,

x_t – feasible decision at stage t,

$X_t(y_t)$ – finite set of decisions feasible at stage t, when the process was in state $y_t \in Y_t$ at the beginning of this stage,

$F_t(y_{t+1} \mid y_t, x_t)$ – value of stage criterion at stage t for the transition from state y_t to state y_{t+1}, when the decision taken was $x_t \in X_t(y_t)$,

$P_t(y_{t+1} \mid y_t, x_t)$ – probability of the transition at stage t from state y_t to state y_{t+1}, when the decision taken was $x_t \in X_t(y_t)$.

$P(y_1)$ – probability of distribution in the set of initial stages $y_1 \in Y_1$.

The following holds:

$$\forall_{t \in \overline{1,T}} \forall_{y_t \in \mathbf{Y}_t} \forall_{x_t \in \mathbf{X}_t(y_y)} \sum_{y_{t+1} \in \mathbf{Y}_{t+1}} P_t(y_{t+1} \mid y_t, x_t) = 1 \qquad (1)$$

$\{x\}$ – strategy – a function assigning to each state $y_t \in Y_t$ exactly one decision $x_t \in X_t(y_t)$,

$\{X\}$ – the set of all strategies of the process under consideration,

$\{x_{\overline{t,T}}\}$ – shortened strategy, encompassing stages from t to T.

Let us assume that we have selected a certain strategy $\{\bar{x}\} \in \{X\}$. The expected value of this strategy is calculated as below.

Algorithm 1
1. For each state $y_T \in Y_T$ we calculate

$$G_T(y_T, \{\bar{x}_{\overline{T,T}}\}) = \sum_{y_{T+1} \in Y_{T+1}} F_T(y_{T+1} \mid y_T, \bar{x}_T) P_T(y_{T+1} \mid y_T, \bar{x}_T). \qquad (2)$$

2. For each stage t, $t \in \overline{T-1,1}$ we calculate the expected value

$$G_t(y_t, \{\bar{x}_{\overline{t,T}}\}) = \sum_{y_{t+1} \in Y_{t+1}} (F_t(y_{t+1} \mid y_t, \bar{x}_t) + G_{t+1}(y_{t+1}, \{\bar{x}_{\overline{t+1,T}}\})) P_t(y_{t+1} \mid y_t, \bar{x}_t). \qquad (3)$$

3. The expected value of the strategy $\{\bar{x}\} \in \{X\}$ is calculated from the formula:

$$G\{\bar{x}\} = \sum_{y_1 \in Y_1} G_1(y_1, \{\bar{x}\}) P_1(y_1). \qquad (4)$$

Using Bellman's optimality principle (Bellman, 1957), we determine the optimal expected value for the process and optimal strategy.

Algorithm 2
1. For each state $y_T \in Y_T$ we calculate the optimal expected values

$$G_T^*(y_T) = \max_{x_T \in X_t(y_t)} \sum_{y_{T+1} \in Y_{T+1}} F_T(y_{T+1} \mid y_T, x_T) P_T(y_{T+1} \mid y_T, x_T) \qquad (5)$$

   and find the decision $x_t^*(y_t)$, for which this maximum is attained. This decision forms a part of the optimal strategy being constructed.
2. For stage t, $t \in \overline{T-1,1}$ and each state $y_t \in Y_t$ we calculate the optimal expected values

$$G_t^*(y) = \max_{x_t \in X_t(y_t)} \sum_{y_{t+1} \in Y_{t+1}} (F_t(y_{t+1} \mid y_t, x_t) + G_{t+1}^*(y_{t+1})) P_t(y_{t+1} \mid y_t, x_t) \qquad (6)$$

   and find the decision $x_t^*(y_t)$, for which this maximum is attained. This decision forms a part of the optimal strategy being constructed.
3. The optimal expected value of the process realisation is calculated from the formula:

$$G\{x^*\} = \sum_{y_1 \in Y_1} G_1^*(y_1, \{x^*\}) P_1(y_1). \qquad (7)$$

3. DETERMINATION OF NEAR OPTIMAL STRATEGIES

The strategy $\{x^m\}$ is called near optimal if the expected value of its realisation differs from the expected value of the realisation of the optimal strategy $\{x^*\}$ by at most the given value z, that is

$$G\{x^*\} - G\{x^m\} \le z \tag{8}$$

where $z > 0$.

We will use the following notation:

LS – the list of optimal and near optimal strategies,

LSB – the list of strategies to be investigated, that is of strategies which can be modi-fied to determine further near optimal strategies,

LSC – the list of strategies considered in the algorithm,

M$\{x\}$ – the set of modified strategies which differ from the strategy $\{x\}$ by a decision in one state.

Algorithm 3

1. Set: LS$:= \varnothing$, LSB$:= \varnothing$, LSC $= \varnothing$.
2. Using **Algorithm 1** determine the set of strategies $\{X^*\}$, for which the given criterion attains the optimal value.
3. Add the strategies from the set $\{X^*\}$ to the sets LS and LSB:

   LS$:=$ LS $\cup \{X^*\}$,

   LSB $:=$ LBS $\cup \{X^*\}$.

4. If LSB $= \varnothing$, go to step 11.
5. Select the next strategy $\{x\}$ from the set LSB; delete it from this set:

   LSB$:=$ LSB $\setminus \{x\}$.

6. Determine all the modified strategies, which differ from the strategy $\{x\}$ by a deci-sion taken in one state and add them to the set M$\{x\}$.
7. Check if the set M$\{x\}$ contains the strategies which are also in the sets LS, LSB and LSC. Delete the duplicate strategies from the set M$\{x\}$:

   M$\{x\} =$ M$\{x\} \setminus$ (M$\{x\} \cap$ LS) $\setminus$ (M$\{x\} \cap$ LSB) $\setminus$ M$\{x\} \cap$ LSC.

8. Check if M$\{x\} \ne \varnothing$. If not, go to step 4.
9. For the consecutive strategies $\{x^m\} \in$ M$\{x\}$:
   a) using formulas (2) i (3) calculate the expected value of the given criterion obtained by applying the strategy $\{x^m\}$,
   b) add the strategy $\{x^m\}$ to the set LSC:

      LSC$:=$ LSC $\cup \{x^m\}$,

c) if the expected value of the given criterion is no lower than *Z*, add the strategy $\{x^m\}$ to the sets LS and LSB:

LS:= LS $\cup$ { x^m },

LSB:= LSB $\cup$ { x^m}.

10. Go to step 4.
11. End of procedure.

This algorithm modifies the strategies from the set LSB by changing a decision in one state only.

For each new strategy we check if it generates a solution different from the ones determined previously. If so, we calculate the expected value of the given criterion and check if it satisfies the condition formulated by the decision maker. If this is not the case, such a strategy does not have to be further analysed, since its further modification cannot lead to an improvement of the criterion value. The procedure ends when the set LSB is empty.

4. APPLICATION OF THE QUASI-HIERARCHICAL APPROACH TO THE SOLUTION OF THE MULTIOBJECTIVE PROBLEM

Let us assume that the solution of the dynamic problem is evaluated with respect to *K* multi-stage criteria, each of which is the sum of *T* stage criteria. The evaluation of each strategy with respect to each criterion is based on the expected value. We assume that the decision maker ordered the criteria starting with the one he or she regards as the most important. We assume therefore that he is first of all interested in the optimisation of the criterion number 1, then of the criterion number 2, etc. The determination of the solution by means of the quasi-hierarchical approach is performed as follows:

Algorithm 4
1. Determine the optimal solutions of the problem with respect to each criterion.
2. Present the optimal values of each criterion to the decision maker.
3. Ask the decision maker to determine the aspiration thresholds Z_k, that is the values which should be attained by each criterion in the final solution.
4. For each criterion determine the set LS_k of strategies satisfying the requirements determined by the decision maker.
5. Set $J = K$.
6. Determine the set LS which is the intersection of the sets LS_k:

$$LS := \bigcap_{k \in \overline{1,J}} LS_k.$$

7. If LS $\neq \varnothing$, go to step 9.
8. Set $J := J - 1$. Go to step 6.

9. From among the solutions in the set LS select those for which the first criterion attains the highest value. If there are more than one such solutions, then in your selection take into account the values of the next criteria in the order determined by the hierarchy formulated by the decision maker.

10. If $\mathrm{LS} := \bigcap_{k \in \overline{1,J}} \mathrm{LS}_k = \varnothing$, check if the strategy obtained by this procedure satisfies the decision maker. If he is not satisfied, choose the final solution from the set LS not applying the proposition, formulated in 9. If he is still not satisfied, return to step 3.

11. End of procedure.

In the procedure we determine the set of alternatives which satisfy all the requirements determined by the decision maker. In many cases it may turn out that such solutions do not exist. In such cases we try to determine solutions satisfying the requirements formulated for those criteria, which the decision maker regards as the most important ones. Gradually, we therefore omit the requirements formulated for the least important criteria, until the set LS containing the solutions satisfying the requirements of the decision maker ceases to be empty. From among the solutions contained in this set we select the one for which the first criterion attains the highest value. If there are more than one such solutions, then in our selection we take into account the values of the next criteria according to the hierarchy defined by the decision maker.

5. ILLUSTRATIVE EXAMPLES

We will illustrate proposed method by means of illustrative examples. The description of the illustrative process under consideration can be found below.

We consider a three-stage decision process. The sets of states for the consecutive stages are as follows:

$$Y_1 = \{1,2\}, \quad Y_2 = \{3,4\}, \quad Y_3 = \{5,6\}.$$

We have the following set of final states of the process:

$$Y_4 = \{7,8\}.$$

The sets of feasible decisions are as follows:

$$X_1(1) = \{A, B\}, \qquad X_1(2) = \{C, D\},$$
$$X_2(3) = \{E, F\}, \qquad X_2(4) = \{G, H\},$$
$$X_3(5) = \{I, J\}, \qquad X_3(6) = \{K, L\}.$$

The graph of the process is given in figure 1. Rectangles denote states of the process in the consecutive stages, circles – random nodes.
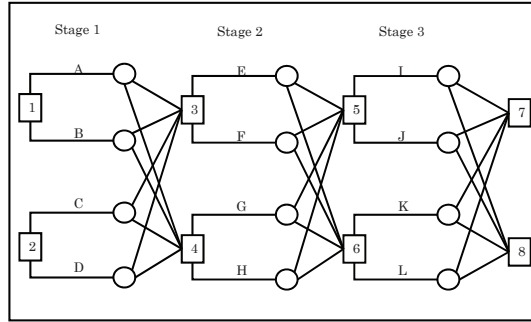
Figure 1. Graph of the process
Source: own elaboration.

The possible stage realisations of the process, probabilities of their occurrence, as well as the values of the stage criteria functions are shown in table 1.

Table 1.

Numeric values

Stage	$(y_{t+1}|y_t, x_t)$	$P(\cdot)$	$F^1(\cdot)$	$F^2(\cdot)$	$F^3(\cdot)$	Stage	$(y_{t+1}|y_t, x_t)$	$P(\cdot)$	$F^1(\cdot)$	$F^2(\cdot)$	$F^3(\cdot)$
1	(3|1,A)	0.4	6	15	22	2	(5|4,G)	0.6	5	15	20
1	(4|1,A)	0.6	8	17	14	2	(6|4,G)	0.4	6	18	13
1	(3|1,B)	0.7	6	15	22	2	(5|4,H)	0.8	5	15	20
1	(4|1,B)	0.3	8	17	14	2	(6|4,H)	0.2	6	18	13
1	(3|2,C)	0.5	6	15	22	3	(7|5,I)	0.8	5	30	12
1	(4|2,C)	0.5	8	17	14	3	(8|5,I)	0.2	1	12	15
1	(3|2,D)	0.8	6	15	22	3	(7|5,J)	0.3	5	30	12
1	(4|2,D)	0.2	8	17	14	3	(8|5,J)	0.7	1	12	15
2	(5|3,E)	0.5	5	15	20	3	(7|6,K)	0.2	5	30	12
2	(6|3,E)	0.5	6	18	13	3	(8|6,K)	0.8	1	12	15
2	(5|3,F)	0.3	5	15	20	3	(7|6,L)	0.9	5	30	12
2	(6|3,F)	0.7	6	18	13	3	(8|6,L)	0.1	1	12	15

Source: own data.

The probability distribution in the set of initial states is as follows:

$P(1) = 0.4,$ $P(2) = 0.6.$

For clarity and due to small size of this illustrative problem, the existing strategies can be written down and numbered from 1 to 64. This numbering is presented in table 2.

List of strategies

No	Decision	No	Decision	No	Decision	No	Decision
1	(A,C,E,G,I,K)	17	(A,D,E,G,I,K)	33	(B,C,E,G,I,K)	49	(B,D,E,G,I,K)
2	(A,C,E,G,I,L)	18	(A,D,E,G,I,L)	34	(B,C,E,G,I,K)	50	(B,D,E,G,I,L)
3	(A,C,E,G,J,K)	19	(A,D,E,G,J,K)	35	(B,C,E,G,J,K)	51	(B,D,E,G,J,K)
4	(A,C,E,G,J,L)	20	(A,D,E,G,J,L)	36	(B,C,E,G,J,L)	52	(B,D,E,G,J,L)
5	(A,C,E,H,I,K)	21	(A,D,E,H,I,K)	37	(B,C,E,H,I,K)	53	(B,D,E,H,I,K)
6	(A,C,E,H,I,L)	22	(A,D,E,H,I,L)	38	(B,C,E,H,I,L)	54	(B,D,E,H,I,L)
7	(A,C,E,H,J,K)	23	(A,D,E,H,J,K)	39	(B,C,E,H,J,K)	55	(B,D,E,H,J,K)
8	(A,C,E,H,J,L)	24	(A,D,E,H,J,L)	40	(B,C,E,H,J,L)	56	(B,D,E,H,J,L)
9	(A,C,F,G,I,K)	25	(A,D,F,G,I,K)	41	(B,C,F,G,I,K)	57	(B,D,F,G,I,K)
10	(A,C,F,G,I,L)	26	(A,D,F,G,I,L)	42	(B,C,F,G,I,L)	58	(B,D,F,G,I,L)
11	(A,C,F,G,J,K)	27	(A,D,F,G,J,K)	43	(B,C,F,G,J,K)	59	(B,D,F,G,J,K)
12	(A,C,F,G,J,L)	28	(A,D,F,G,J,L)	44	(B,C,F,G,J,L)	60	(B,D,F,G,J,L)
13	(A,C,F,H,I,K)	29	(A,D,F,H,I,K)	45	(B,C,F,H,I,K)	61	(B,D,F,H,I,K)
14	(A,C,F,H,I,L)	30	(A,D,F,H,I,L)	46	(B,C,F,H,I,L)	62	(B,D,F,H,I,L)
15	(A,C,F,H,J,K)	31	(A,D,F,H,J,K)	47	(B,C,F,H,J,K)	63	(B,D,F,H,J,K)
16	(A,C,F,H,J,L)	32	(A,D,F,H,J,L)	48	(B,C,F,H,J,L)	64	(B,D,F,H,J,L)

Source: own elaboration.

Example 1

Applying Algorithm 1 find expected value for criterion F^1 and the strategy $\{x^{28}\}$. For $t = 3$ we have (formula (2)):

$$G_3(6,\{x^{28}\}) = F_3^1(7|6,L)\, P_3(7|6,L) + F_3^1(8|6,L)\, P_3(8|6,L) = 5\times0.9 + 1\times0.1 = 4.6,$$

$$G_3(5,\,\{x^{28}\}) = F_3^1(7|5,J)\times P_3(7|5,J) + F_3^1(8|5,J)\times P_3(8|5,J) = 5\times0.3 + 1\times0.7 = 2.2.$$

For $t = 2$ we have (formula (3)):

$$G_2(4,\{x^{28}\}) =$$
$$= [F_2^1(5|4,G) + G_3(5,\{x^{28}\})]\times P_2(5|4,G) + [F_2^1(6|4,G) + G_3(6,\{x^{28}\})]\times P_2(6|4,G) =$$
$$= (5+2.2)\times0.6 + (6+4.6)\times0.4 = 8.56,$$

$$G_2(3,\{x^{28}\}) =$$
$$= [F_2^1(5|3,F) + G_3(5,\{x^{28}\})]\times P_2(5|3,F) + [F_2^1(6|3,F) + G_3(6,\{x^{28}\})]\times P_2(6|3,F) =$$
$$= (5+2.2)\times0.3 + (6+4.6)\times0.7 = 9.58.$$

For $t = 1$ we have (formula (3)):

$G_1(2, \{x^{28}\}) =$
$= [F_1^1(3|2,D) + G_2(3, \{x^{28}\})] \times P_1(3|2,D) + [F_1^1(4|2,D) + G_2(4, \{x^{28}\})] \times P_1(4|2,D) =$
$= (6+9.58) \times 0.8 + (8+8.56) \times 0.2 = 15.776,$

$G_1(1, \{x^{28}\}) =$
$= [F_1^1(3|1,A) + G_2(3, \{x^{28}\})] \times P_1(3|1,A) + [F_1^1(4|1,A) + G_2(4, \{x^{28}\})] \times P_1(4|1,A) =$
$= (6+9.58) \times 0.4 + (8+8.56) \times 0.6 = 16.168.$

The expected value for the strategy $\{x^{28}\}$ is calculated from the formula (4):

$G(\{x^{28}\}) =$
$= G_1(1, \{x^{28}\}) \times P(1) + G_1(2, \{x^{28}\}) \times P(2) = 15.772 \times 0.4 + 14.624 \times 0.6 = 15.0832.$

Example 2
Applying Algorithm 2 find optimal expected values for criterion F^1 and optimal strategy for this criterion.
For $t = 3$ we have (formula (5)):

$G_3^*(6) = \max \{F_3^1(7|6,K) \times P_3(7|6,K) + F_3^1(8|6,K) \times P_3(8|6,K),$
$\qquad\qquad F_3^1(7|6,L) \times P_3(7|6,L) + F_3^1(8|6,L) \times P_3(8|6,L)\} =$
$\qquad = \max \{(5 \times 0.2 + 1 \times 0.8), (5 \times 0,9 + 1 \times 0.1) = 4.6,$

$x_3^*(6) = L,$

$G_3^*(5) = \max \{F_3^1(7|5,I) \times P_3(7|5,I) + F_3^1(8|5,I) \times P_3(8|5,I),$
$\qquad\qquad F_3^1(7|5,J) \times P_3(7|5,J) + F_3^1(8|5,J) \times P_3(8|5,J)\} =$
$\qquad = \max \{(5 \times 0.8 + 1 \times 0.2), 5 \times 0,3 + 1 \times 0.7) = 4.2,$

$x_3^*(5) = I.$

For $t = 2$ we have (formula (6)):

$G_2^*(4) = \max \{[F_2^1(5|4,G) + G_3^*(5)] \times P_2(5|4,G) + [F_2^1(6|4,G) + G_3^*(6)] \times P_2(|4,G),$
$\qquad\qquad [F_2^1(5|4,H) + G_3^*(5)] \times P_2(6|4,H) + [F_2(6|4,H) + G_3^*(6)] \times P_2(6|4,H)\} =$
$\qquad = \max \{[(5+4.2) \times 0.6 + (6+4.6) \times 0.4], [(5+4.2) \times 0.8 + (6+4.6) \times 0.2]\} = 9.76,$

$x_2^*(4) = G,$

$G_2^*(3) = \max \{[F_2^1(5|3,E) + G_3^*(5)] \times P_2(5|3,E + [F_2^1(6|3,E) + G_3(6)] \times P_2(6|3,E),$
$\qquad\qquad [F_2^1(5|3,E) + G_3^*(5)] \times P_2(5|3,E) + [F_2^1(6|3,E) + G_3^*(6)] \times P_2(6|3,E)\} =$
$\qquad = \max \{[(5+4.2) \times 0.5 + (6+4.6) \times 0.5], [(5+4.2) \times 0.3 + (6+4.6) \times 0.7]\} = 9.76,$

$x_2^*(3) = F.$

For $t = 1$ we have (formula (6)):

$G_1^*(2) = \max \{[F_1^1(3|2,C) + G_2^*(3)] \times P_1(3|2,C) + [F_1^1(4|2,C) + G_2^*(4)] \times P_1(4|2,C),$
$\qquad\qquad [F_1^1(3|2,D) + G_2^*(3)] \times P_1(3|2,D) + [F_1^1(4|2,D) + G_2^*(4)] \times P_1(4|2,D)\} =$
$\qquad = \max \{[(6+10.18) \times 0.5 + (8+9.76) \times 0.5], [(6+10.18) \times 0.8 + (8+9.76) \times 0.2]\} =$
$\qquad = 16.97,$

$x_1^*(2) = C,$

$G_1^*(1) = \max \{F_1^1(3|1,A) + G_2^*(3)] \times P_1(3|1,A) + [F_1^1(4|1,A) + G_2^*(4)] \times P_1(4|1,A),$
$\qquad\qquad F_1^1(3|1,B) + G_2^*(3)] \times P_1(3|1,B) + [F_1^1(4|1,B) + G_2^*(4)] \times P_1(4|1,B)\} =$
$\qquad = \max \{[(6+10.18) \times 0.4 + (8+9.76) \times 0.6], [(6+10.18) \times 0.7 + (8+9.76) \times 0.3]\} =$
$\qquad = 17.128,$

$x_1^*(1) = A.$

The optimal expected value is calculated from the formula (7):

$G\{x^*\} = G_1^*(1) \times P(1) + G_1^*(2) \times P(2) = 17.128 \times 0.4 + 16.97 \times 0.6 = 17.0332.$

Example 3

We have to determine all the strategies for which the expected value of the criterion number 1 differs from the optimal value by at most 2%.

The determination of near optimal strategies, described in **Algorithm 3** proceeds as follows:

1. We set

   LS:= $\varnothing$, LSB:= $\varnothing$, LSC = $\varnothing$.

2. Using Algorithm 1 we find the set optimal strategy:

   $\{X^*\} = \{\{x^{10}\}\}$

   for which the given criterion attains the optimal value equal to 17.0332. We have (see table 2):

   $\{x^{10}\} = (A,C,F,G.I,L).$

3. We add the strategy found to the sets LS and LSB. We have:

   LS = LS $\cup \{X^*\} = \{\{x^{10}\}\}$,

   LSB = LSB $\cup \{X^*\} = \{\{x^{10}\}\}$.

4. Since LSB $\neq \varnothing$, we go to step 5.
5. We select the strategy $\{x^{10}\}$ from the set LSB and delete it from this set:

   LSB:= LSB $\setminus \{x^{10}\} = \varnothing$.

6. We determine all the modified strategies which differ from the strategy $\{x^{10}\}$ by a decision taken in one state:

    $\{x^9\}$ = { A,C,F,G,I,K}, $\{x^{12}\}$ = { A,C,F,G,J,L }, $\{x^{14}\}$ = { A,C,F,H,I,L },

    $\{x^2\}$ = { A,C,E,G,I,L }, $\{x^{26}\}$ = { A,D,F,G,I,L }, $\{x^{42}\}$ = { B,C,F,G,I,L },

    and add them to the set $M\{x^{10}\}$:

    $M\{x^{10}\}$ = {$\{x^9\}$, $\{x^{12}\}$, $\{x^{14}\}$, $\{x^2\}$, $\{x^{26}\}$, $\{x^{42}\}$}.

7. We check if the set $M\{x^{10}\}$ contains strategies which are also contained in the sets LS, LSB and LSC. We obtain:

    $M\{x^{10}\} \cap LS = \varnothing$,

    $M\{x^{10}\} \cap LSB = \varnothing$,

    $M\{x^{10}\} \cap LSC = \varnothing$.

8. We have $M\{x^{10}\} \neq \varnothing$.
9. We consider further strategies $\{x^m\} \in M\{x^{10}\}$.

    Strategy $\{x^9\}$
     a) we calculate the expected value

        $G\{x^9\}$ = 15.5268,

     b) we add the strategy $\{x^9\}$ to the set LSC:

        LSC:= LSC $\cup$ $\{x^9\}$ = {$\{x^9\}$},

     c) since

        $G\{x^9\}$ = 15.5268 < 16.6925,

        we do not add the strategy $\{x^9\}$ to the sets LS and LSB.

    Strategy $\{x^{12}\}$
     a) we calculate the expected value

        $G\{x^{12}\}$ = 16.192,

     b) we add the strategy $\{x^{12}\}$ to the set LSC:
        LSC:= LSC $\cup$ $\{x^{12}\}$ = {$\{x^9\}$, $\{x^{12}\}$},

     c) since

        $G\{x^{12}\}$ = 16.192 < 16.6925,

        we do not add the strategy $\{x^{12}\}$ to the set LS or to the set LSB.

    Strategy $\{x^{14}\}$
     a)   we calculate the expected value

        $G\{x^{14}\}$ = 16.882,

b) we add the strategy $\{x^{14}\}$ to the set LSC:

   LSC:= LSC $\cup$ $\{x^{14}\}$ = $\{\{x^9\}, \{x^{12}\}, \{x^{14}\}\}$,

 c) since

   $G\{x^{14}\}$ = 16.882 > 16.6925,

   we add the strategy $\{x^{14}\}$ to the sets LS and LSB:

   LS:= LS $\cup$ $\{x^{14}\}$ = $\{\{x^{10}\}, \{x^{14}\}\}$,

   LSB:= LSB $\cup$ $\{x^{14}\}$ = $\{\{x^{14}\}\}$.

Strategy $\{x^2\}$

a) we calculate the expected value

   $G\{x^2\}$ = 16.9044,

b) we add the strategy $\{x^2\}$ to the set LSC:

   LSC:= LSC $\cup$ $\{x^2\}$ = $\{\{x^9\}, \{x^{12}\}, \{x^{14}\}, \{x^2\}\}$,

c) since

   $G\{x^{63}\}$= 16.9044 > 16.6925,

   we add the strategy $\{x^2\}$ to the set LS and to the set LSB:

   LS:= LS $\cup$ $\{x^2\}$ = $\{\{x^{10}\}, \{x^{14}\}, \{x^2\}\}$,

   LSB:= LSB $\cup$ $\{x^{14}\}$ = $\{\{x^{14}\}, \{x^2\}\}$.

Strategy $\{x^{26}\}$

a) we calculate the expected value

   $G\{x^{26}\}$ = 16.7488,

b) we add the strategy $\{x^2\}$ to the set LSC:

   LSC:= LSC $\cup$ $\{x^{26}\}$ = $\{\{x^9\}, \{x^{12}\}, \{x^{14}\}, \{x^2\}, \{x^{26}\}\}$,

c) since

   $G\{x^{26}\}$ = 16.7488 > 16.6925,

   we add the strategy $\{x^2\}$ to the set LS and to the set LSB:

   LS:= LS $\cup$ $\{x^2\}$ = $\{\{x^{10}\}, \{x^{14}\}, \{x^2\}, \{x^{26}\}\}$,

   LSB:= LSB $\cup$ $\{x^{14}\}$ = $\{\{x^{14}\}, \{x^2\}, \{x^{26}\}\}$.

Strategy $\{x^{42}\}$

a) we calculate the expected value

   $G\{x^2\}$ = 16.8436,

b)   we add the strategy $\{x^{42}\}$ to the set LSC:

LSC:= LSC $\cup$ $\{x^2\}$ = $\{\{x^9\}, \{x^{12}\}, \{x^{14}\}, \{x^2\}, \{x^{42}\}\}$,

c)   since

$G\{x^{63}\}$ = 16.8436 > 16.6952,

we add the strategy $\{x^2\}$ to the set LS and to the set LSB:

LS:= LS $\cup$ $\{x^2\}$ = $\{\{x^{10}\}, \{x^{14}\}, \{x^2\}, \{x^{42}\}\}$,

LSB:= LSB $\cup$ $\{x^{14}\}$ = $\{\{x^{14}\}, \{x^2\}, \{x^{42}\}\}$.

10. We go to step 4.

 4. Since LSB $\neq \varnothing$, we go to step 5.

Continuing this procedure, in the consecutive steps we determine the next strategies which are near optimal and satisfy the condition:

$G\{x\}$ > 16.6952.

The set of optimal and near optimal strategies, for which the expected value of the given criterion differs from the optimal value by no more than 2%, that is for which $G\{x\} \geq 16.6952$, contains the following strategies:

LS = $\{\{x^{10}\}, \{x^{14}\}, \{x^2\}, \{x^{42}\}, \{x^6\}, \{x^{26}\}, \{x^{46}\}\}$.

Example 4

Now we regard the considered process as a three-criteria hierarchical process, in which the most important is the first criterion, the second-most important is the second criterion, and the least important is the third criterion.

The determination of the final strategy using the quasi-hierarchical procedure described in **Algorithm 4** is performed as follows:

1. We determine the optimal solution of the problem with respect to each criterion. The optimal strategy with respect to the first criterion is $\{x^{10}\}$. The expected value for this strategy with respect to the first criterion is 17.0332.
    The optimal strategy with respect to the second criterion is $\{x^{10}\}$. The expected value for this strategy with respect to the second criterion is 60.0624.
    The optimal strategy with respect to the third criterion is $\{x^{55}\}$. The expected value for this strategy with respect to the third criterion is 51.3124.
2. We present the optimal values of each criterion to the decision maker.
3. Based on the information obtained, the decision maker decided, that the expected values of all the criteria can differ from the optimal value at most 2%. It means, that the aspiration levels are as follows:
    – Z_1 = 16.6925 for criterion 1,
    – Z_2 = 58.8612 for criterion 2,
    – Z_3 = 50.2862 for criterion 3.

4. For each criterion we determine the set of strategies satisfying the requirements determined by the decision maker. We obtain:

$LS_1 = \{\{x^{10}\}, \{x^{14}\}, \{x^2\}, \{x^{42}\}, \{x^6\}, \{x^{26}\}, \{x^{46}\}\}$,

$LS_2 = \{\{x^{10}\}, \{x^{42}\}, \{x^{26}\}, \{x^{58}\}, \{x^{62}\}, \{x^2\}, \{x^{30}\},$
$\quad \{x^{46}\}, \{x^{14}\}, \{x^{34}\}, \{x^{18}\}, \{x^{50}\}, x^6\}, \{x^{38}\}, \{x^{22}\}, \{x^{54}\}\}$,

$LS_3 = \{\{x^{55}\}, \{x^{51}\}, \{x^{23}\}, \{x^{53}\}, \{x^{56}\}\}$.

5. We set $J = 3$.
6. We determine

$LS_1 \cap LS_2 \cap LS_3 = \varnothing$.

7. Since $LS = \varnothing$, we go to step 8.
8. We set $J := J - 1 = 2$. We go to step 6.
6. We determine

$LS = LS_1 \cap LS_2 = \{\{x^{10}\}, \{x^2\}, \{x^{14}\}, \{x^{42}\}, \{\{x^{26}\}, \{x^{46}\}\}$.

7. Since $LS \neq \varnothing$, we go to step 9.
9. From among the solutions from the set LS we select the strategy $\{x^{10}\}$, which is optimal for both the first and the second criteria.
10. Since for $J = 3$ we have $LS := \bigcap_{k \in \overline{1,J}} LS_k = \varnothing$ we check if the strategy obtained in our procedure satisfies the decision maker. The expected value of the strategy $\{x^{10}\}$ for criterion 3 is 46.3526. It is easy to check that this is the worst strategy with respect to the expected value for criterion 3, hence the selection of this strategy as the final strategy is very unsatisfactory for the decision maker. We return to step 3.

 3 The decision maker decided to slightly lower the aspiration level with respect to the first criterion. The new aspiration level with respect to criterion 1 is
$Z_1 = 16.68$.

 4. We determine the sets:

$LS_1 = \{\{x^{10}\}, \{x^{14}\}, \{x^2\}, \{x^{42}\}, \{x^6\}, \{x^{26}\}, \{x^{46}\}, \{x^{34}\}\}$,

$LS_2 = \{\{x^{10}\}, \{x^{42}\}, \{x^{26}\}, \{x^{58}\}, \{x^{62}\}, \{x^2\}, \{x^{30}\}, \{x^{46}\}, \{x^{14}\}, \{x^{34}\}, \{x^{18}\}, \{x^{50}\},$
$\quad \{x^6\}, \{x^{38}\}, \{x^{22}\}, \{x^{54}\}\}$,

$LS_3 = \{\{x^{55}\}, \{x^{51}\}, \{x^{23}\}, \{x^{53}\}, \{x^{56}\}\}$.

 5. We set $J = 3$.
 6. We have

$LS_1 \cap LS_2 \cap LS_3 = \varnothing$.

7. Since $LS = \varnothing$, we go to step 8.
8. We set $J := J - 1 = 2$. We go to step 6.

6. We determine

   LS = LS$_1$ ∩ LS$_2$.= {{x^{10}},{x^{14}},{x^2},{x^{42}}, {x^{26}},{x^{46}}, {x^{34}}}.

7. Since LS ≠ ∅, we go to step 9.
9. From among the solutions belonging to the set LS the best strategy is again {x^{10}}, disqualified previously by the decision maker. That is why we consider the next strategies from the set LS$_3$.
10. We present the strategy {x^{34}}, for which G^3{x^{34}}= 47.8966, to the decision maker. This value, although still far from the aspiration level accepted by the decision maker for the third criterion, has been accepted by him.
11. End of procedure.

6. SUMMARY

The quasi-hierarchical approach is a frequently used method of solving multi-criteria problems. It requires that the decision maker order the criteria of the evaluation of the decision process. In our paper we have presented a way of applying this approach to solving the multi-criteria problem. It has been assumed that the decision maker is able to order the criteria from the most important one to the least important one, and that based on the information about optimal solutions with respect to each criterion he or she can formulate the conditions to be satisfied by the strategies to be taken into account in the determination of the final solution of the problem.

In our paper we assumed that the evaluation of the quality of the individual solutions with respect to each criterion was based on the expected value. This is not, however, the only possible way of analysing the problem. For the evaluations of the solutions one can use also measures based on the probability of the occurrence of a given event, as well as the conditional expected value. In future papers we intend to propose a quasi-hierarchical method taking into account criteria of this type.

The approach proposed in the paper can be applied to solve a variety of problems. In future research, we are going to show how the quasi-hierarchical dynamic approach can be used to solve the project portfolio selection problem and production capacity planning problem.

REFERENCES

Bakker B., Zivkovic Z., Krose B., (2005), Hierarchical Dynamic Programming for Robot Path Planning, in: *Intelligent Robots and Systems* 2005 (IROS 2005), 2756–2761.

Bellman R., (1957), *Dynamic Programming*, Princenton University Press.

Daellenbach H. G., De Kluyver C. A., (1980), Note on Multiple Objective Dynamic Programming, *Journal of the Operational Research Society*, 31 (7), 591–594.

Dempster M. A. H., (2006), Sequential Importance Sampling Algorithms for Dynamic Stochastic Programming, *Journal of Mathematical Sciences*, 133 (4), 1422–1444.

Elmaghraby S. E., (1970), The Theory of Networks and Management Science, part I, *Management Science*, 17 (1), 1–34.

Hatzakis I., Wallace D., (2006), Dynamic Multi-Objective Optimization with Evolutionary Algorithms: a Forward-looking Approach, in: *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation*, 1201–1208.

Nowak M., (2014), Quasi-Hierarchical Approach in Multiobjective Decision Tree, *Studia Ekonomiczne Uniwersytetu Ekonomicznego w Katowicach*, 208, 59–73 (in Polish).

Nowak M., Trzaskalik T., (2017), Optimal and Near Optiomal Strategies in Discrete Stochastic Multiobjective Quasi-Hierarcical Dynamic Poblems, in: Doerner K. F., Ljubic I., Pflug G., Tragler G., (eds.), *Operations Research Proceedings*, Springer, 295–300.

Sethi S. P., Yan H., Zhang H., Zhang Q., (2002), Optimal and Hierarchical Controls in Dynamic Stochastic Manufacturing Systems: A Survey. *Manufacturing & Service Operations Management*, 4 (2), 133–170.

Sethi S., Zhang Q., (1994), Hierarchical Production Planning in Dynamic Stochastic Manufacturing Systems: Asymptotic Optimality and Error Bounds, *Journal of Mathematical Analysis and Applications*, 181 (2), 285–319.

Shapiro A., (2012), Minimax and Risk Averse Multistage Stochastic Programming, *European Journal of Operational Research*, 219 (3), 719–726.

Tempelmeier H., Hilger T., (2015), Linear Programming Models for a Stochastic Dynamic Capacitated Lot Sizing Problem, *Computers & Operations Research*, 59, 119–125.

Topaloglou N., Vladimirou H., Zenios S. A., (2008), A Dynamic Stochastic Programming Model for International Portfolio Management, *European Journal of Operational Research*, 185 (3), 1501–1524.

Trzaskalik T., (1990), *Multiobjective Discrete Dynamic Programming. Theory and Applications in Economics*, The Karol Adamiecki University of Economics in Katowice Press, Katowice (in Polish).

Trzaskalik T., (1998), *Multiobjective Analysis in Dynamic Environmnent*, The Karol Adamiecki University of Economics in Katowice Press, Katowice, 1998.

Woerner S., Laumanns M., Zenklusen R., Fertis A., (2015), Approximate Dynamic Programming for Stochastic Linear Control Problems on Compact State Spaces, *European Journal of Operational Research*, 241 (1), 85–98.

PODEJŚCIE QUASI-HIERARCHICZNE W DYSKRETNYM WIELOKRYTERIALNYM STOCHASTYCZNYM PROGRAMOWANIU DYNAMICZNYM

S t r e s z c z e n i e

W pracy rozważany jest wieloetapowy wielokryterialny proces podejmowania decyzji w warunkach ryzyka. W celu jego rozwiązania wykorzystano dyskretne stochastyczne programowanie dynamiczne oparte na zasadzie optymalności Bellmana. Zakłada się, że decydent jest w stanie zdefiniować quasi-hierarchię rozważanych kryteriów, co oznacza, że jest on w stanie określić w jakim zakresie optymalna wartość oczekiwana dla kryteriów o wyższym priorytecie może być pogorszona w celu poprawy wartości oczekiwanej kryterium o priorytecie niższym. Proces uzyskania rozwiązania końcowego może być realizowany interaktywnie. Obserwując kolejno proponowane rozwiązania, decydent może modyfikować poziomy aspiracji dla rozważanych kryteriów, otrzymując ostatecznie rozwiązanie satysfakcjonujące. Metoda została zilustrowana przykładem opartym na danych umownych.

Słowa kluczowe: programowanie dynamiczne, podejmowanie decyzji w warunkach ryzyka, podejście interaktywne, metoda quasi-hierarchiczna

QUASI-HIERARCHICAL APPROACH TO DISCRETE MULTIOBJECTIVE STOCHASTIC DYNAMIC PROGRAMMING

Abstract

In this paper we consider a multi-stage, multi-criteria discrete decision process under risk. We use a discrete, stochastic dynamic programming approach based on Bellman's principle of optimality. We assume that the decision maker determines a quasi-hierarchy of the criteria considered; in other words, he or she is able to determine to what extent the optimal expected value of a higher-priority criterion can be made worse to improve the expected value of a lower-priority criterion. The process of obtaining the final solution can be interactive. Based on the observations of the consecutive solutions, the decision maker can modify the aspiration levels with respect to the criteria under consideration, finally achieving a solution which satisfies him/her best. The method is illustrated on an example based on fictitious data.

Keywords: dynamic programming, decision making under risk, interactive approach, quasi-hierarchical method

LECH KRUŚ[1], IRENA WORONIECKA-LECIEJEWICZ[2]

MONETARY-FISCAL GAME ANALYZED USING A MACROECONOMIC MODEL FOR POLAND

1. INTRODUCTION

The paper deals with the choice of policy mix in the context of mutual decision conditioning between the fiscal authority (the government) and the monetary authority (the central bank). Mathematical modeling, game theory and multicriteria optimization methods are applied. The policy mix means a combination of a monetary and a fiscal policy with a given restrictiveness/expansiveness level of each of them.

There exists a relatively rich bibliography dealing with interactions of fiscal and monetary policies. Blinder (1983) and after him Bennett, Loayza (2001) considered a simplified monetary-fiscal game with the fiscal and monetary authorities as players having respectively two fiscal and two monetary strategies: restrictive and expansive ones. The authors shown that an independent actions of the authorities may lead to the Nash equilibrium which is not Pareto optimal. They presented a similar interpretation relating to the prisoner's-dilemma problem and similar arguments for coordination of the policies. Nordhaus (1994) analyzed the problem of independence versus coordination of fiscal and monetary policies using a monetary-fiscal game. The game is based on a simple hypothetical macroeconomic model with utility functions of the government and of the central bank, dependent on their policy instruments. He presented an extended discussion relating to the Nash equilibria, Pareto optimality of payoffs, possible conflicts of interests of the authorities and suggestions for policy coordination. The Nordhaus game model is a starting point for further research. In the monograph (Marszałek, 2009, p. 131–132) a list of selected game models describing relations between the government and the central bank is presented and the models are characterized. Dixit, Lambertini (2001) considering a monetary-fiscal game underlined the importance of players credibility and fiscal discipline for results of the game. Lambertini, Rovelli (2003) continued the above research comparing the Nash and Stackelberg games. Many authors discuss and explain the facts that solutions in the above models of noncooperative monetary-fiscal games are not optimal and lead to a suboptimal policy mix, see for example the papers (Darnault, Kutos, 2005;

[1] Polish Academy of Sciences, Systems Research Institute, 6 Newelska St., 01-447 Warsaw, Poland, corresponding author – e-mail: krus@ibspan.waw.pl.

[2] Polish Academy of Sciences, Systems Research Institute, 6 Newelska St., 01-447 Warsaw, Poland.

Wojtyna, 1996; Marszałek, 2009). There are also some papers discussing policy mix problems, formulating arguments for policy coordination and presenting interesting results with the use of statistical data for Poland (Darnault, Kutos, 2005; Stawska, 2014). Libich et al. (2015) presented analysis and comparison of selected countries in so called *monetary vs fiscal leadership space*. Poland is located in the central part of the space. It means that in the case of Poland the fiscal authority does not dominate the monetary authority and nor vice versa. Cevik et al. (2015) examined the interactions between fiscal and monetary policy for some former transition, emerging European economies, also for Poland, over the 1995–2010 period by using a Markov regime-switching model. Empirical results suggest that monetary and fiscal policy rules exhibit switching properties between active and passive regimes. Libich, Nguyen (2015) analyzed strategic interaction between the central bank and government in the post global financial crisis period of 2010–2014. They concentrate on inflation targeting and its possible effect on both monetary and fiscal outcomes. Analysis of monetary and fiscal policies in Poland was also presented in OECD Economic Surveys (Monetary and fiscal policies to head off overheating, Poland 2008).

There are no publications dealing with interactions of the fiscal and monetary policies analyzed with the use of computational game models for Poland. The research presented in this paper tries to cover this gap. It is the first presentation of results within the monetary-fiscal games based on the macroeconomic model for Poland.

2. SUBJECT OF THE PAPER

This paper presents current results of the research carried on within the game theory, macroeconomic modeling and optimization methods applied for analysis of the policy mix problem. We try to analyze an efficiency of decisions made by the authorities, considering the Nash equilibria and Pareto optimality of their decisions. We try also answer the questions: how priorities of the monetary and fiscal authorities relate to the choice of the authorities' strategies; when and under what conditions the independent choice of strategies by the monetary and fiscal policies leads to the decisions which are economically effective and when a coordination of the decisions is required.

A noncooperative game called the monetary-fiscal game is formulated and analyzed in which the fiscal and monetary authorities play roles of players. Strategies of the monetary authority relate to the monetary policies with different restrictiveness/ expansiveness level and are characterized by the real interest rate. Similarly, strategies of the fiscal authority mean the budget policies with different restrictiveness/expansiveness level. They are characterized by the budget deficit in relation to the GDP. The level of restrictiveness of each policy is defined by a value of the respective policy instrument. Each authority tries to obtain his respective economic target: a desired value of the GDP dynamics in the case of the fiscal authority, and a desired value of inflation in

the case of the monetary authority. It is assumed that the authorities make decisions independently.

A macroeconomic model for the Polish economy has been formulated on the basis of the New Neoclassical Synthesis concept. It includes four fundamental equations referring to the output gap, inflation, expected inflation and the Taylor rule of the interest rate. It allows analyzing of the economic situation in time. It takes into account the influence of interest rate on economy. The classical form of the model has been extended to include the influence of the fiscal policy. The model has been estimated using quarterly time series of data for Poland from the period 2000–2014.

A computer-based system calculating results of the game has been constructed using the above model. A sequence of simulations have been made in which payoffs of the game were derived for alternative monetary and fiscal policies. This paper presents continuation of the research described in the previous papers of the authors (Kruś, Woroniecka-Leciejewicz, 2015; Woroniecka-Leciejewicz, 2015a, b, 2010, 2008, 2007).

This paper is organized as follows. The next section 3 presents mathematical formulation of the game. The proposed macroeconomic NNS-MFG model is described in section 4. Section 5 presents results of the model estimation and examples of simulation runs. Analysis of the proposed monetary-fiscal game are shown and discussed in section 6. Conclusions are in section 7.

3. MATHEMATICAL FORMULATION OF THE GAME

Relations between the fiscal authority and the monetary authority can be described by a noncooperative, static, deterministic game. It is a single stage, deterministic, non-zero sum, perfect information game played by the central bank and the government. Each player takes decision independently taking into account possible reaction of the counter player. The game is defined in the strategic form as follows:

(i)   There are two players $i = 1, 2$: the fiscal authority (the government) and the monetary authority (the central bank).

(ii) For each player a set Ω^i of pure strategies is defined. The strategies of the fiscal authority are those of the budgetary policy – from the extremely restrictive to the extremely expansive. The measure, denoted by b, of the degree of restrictiveness/expansiveness of the fiscal policy is constituted here by the level of budget deficit in relation to GDP. The strategies of the monetary authority range from the extremely restrictive one to the extremely expansive. The degree of restrictiveness/expansiveness is equivalent simply to the value of the real interest rate and denoted by r. Let Ω denote the Cartesian product of the sets of the strategies $\Omega = \Omega^1 \times \Omega^2$.

(iii) For each player $i = 1, 2$, a function $h^i: \Omega \rightarrow \boldsymbol{R}$ is given defining outcome of the player i for given strategies undertaken by the both players. The outcome of the fiscal authority is measured by the GDP growth rate, denoted by y, where

$y = h^1(b, r)$. In the case of the monetary authority it is the inflation value, denoted by p, where $p = h^2(b, r)$. The functions h^i, $i = 1, 2$, are defined by the model relations.

(iv) For each player $i = 1, 2$, a preference relation is given in the set of the attainable outcomes. It is assumed here that each authority tries to achieve a given goal: the fiscal authority – a desired value of GDP growth, the monetary authority – a desired value of inflation.

Outcomes of the game in the discrete form are presented in table 1. Payoffs in the table are denoted in the following manner: y_{ij} – payoff of the fiscal authorities (GDP growth rate) in the case where the government applies the fiscal strategy F_i and the central bank applies the monetary strategy M_j; p_{ij} – cost to the monetary authorities (inflation) for the same pair of policies. The symbol b_i denotes the budgetary deficit in relation to GDP, corresponding to the i-th fiscal strategy, while r_j denotes the real interest rate, ascribed to the j-th monetary strategy.

It is assumed that the fiscal and monetary authorities take decisions independently, and the Nash equilibrium state in such a game is identified with the choice of a given combination of the budgetary and monetary policies.

Table 1.

The monetary-fiscal game – table of payoffs

Strategies		Central bank – the monetary policy			
		← restrictive			expansive →
		Monetary strategy M_1 (interest rate r_1)	Monetary strategy M_2 (interest rate r_2)	…	Monetary strategy M_n (interest rate r_n)
Government – fiscal policy	restrictive ↑	Fiscal strategy F_1 (budgetary deficit b_1)	p_{11} y_{11}	p_{12} y_{12}	… …	p_{1n} y_{1n}
		Fiscal strategy F_2 (budgetary deficit b_2)	p_{21} y_{21}	p_{22} y_{22}	… …	p_{2n} y_{2n}
	expansive ↓	…			…	
		Fiscal strategy F_m (budgetary deficit b_m)	p_{m1} y_{m1}	p_{m2} y_{m2}	… …	p_{mn} y_{mn}

Source: Woroniecka-Leciejewicz (2010, p. 191).

4. MACROECONOMIC NNS-MFG MODEL

This section describes a recursive macroeconomic model based on the New Neoclassical Synthesis (NNS) concept. The model – called NNS-MFG has been constructed to analyze the discussed monetary-fiscal game (MFG).

    Development of macroeconomic modeling based on dynamic stochastic general equilibrium concepts is observed in the last years, trying to find a consensus among alternative theoretical views on the key macroeconomic phenomena and modeling problems. The models proposed describe among others a temporary influence of the monetary policy on economic activity. The New Neoclassical Synthesis theory tries to combine positives of concurrent modern theories. It adopts the concepts of inter-temporal optimal behavior of households and firms, rational expectations and permanently balanced markets from the New Classical Economics and Real Business Cycle schools. On the other hand it accepts the assumption of monopolistic competition taken from the New Keynesian Economics school. This theoretical concept is called in literature as the New Neoclassical Synthesis (Goodfriend, King, 1997), the Neo-Wicksellian Model (Woodford, 2003), the New Keynesian Model (Blanchard, 2009), the new consensus in macroeconomics (Arestis, 2009), the New Keynesian macroeconomics (Spahn, 2009). Theoretical backgrounds of the NNS concept, discussion of doubts and proposals of possible extensions are presented by Bludnik (2010). The NNS models are constructed around three relations having deep roots in the economic theory and treated as essential in description of transmission of the monetary policy impulses. The three relations refer to the IS curve, the New Keynesian Philips curve and the Taylor rule. The basic NNS model is formally presented in the papers (Goodhart, 2007, p. 4; Galí, 2009, p. 2−3).

    The proposed macroeconomic model NNS-GMF has been constructed to derive payoffs of the monetary-fiscal game in simulation experiments. The model has to fulfill the following prerequisites – it should enable analysis of the impact of the monetary and fiscal policies and their instruments: the real interest rate and the budget deficit in relation to GDP on the state of economy, i.e. on the GDP growth and inflation. It should be a dynamic model, enabling observation of the economic activity in time.

    According to the ideas of the control theory the players strategies are input variables for which the state of economy and payoffs treated as outputs of the model are derived in recursive calculations. Therefore the real interest rate and the budget deficit in relation to GDP are exogenous variables. On the other hand the model is constructed on the basis of the NNS model (New Neoclassical Synthesis, Goodhart, 2007; Galí, 2009), which includes three key relations describing mechanisms of transmission of impulses of the monetary policy, i.e. the IS curve (equation of the demand gap), the Philips curve (the inflation equation) and the Taylor rule. The NNS model enables observation of the economic activity in time and takes into account the influence of the interest rate on the economy. The proposed NNS-MFG differs from the classic NNS model. Using the model one can observe not only effects of the monetary policy instruments but also instruments of the fiscal policy. For this reason it takes into account the budget expenditures also. The production gap equation has an analogic form as the equation in the basic NNS model, however the explanatory variables have no anticipative character. This simplification is typically applied in empirical studies, see for example Batini, Haldane (1999), Muinhos (2001), Freitas, Muinhos (2001), Kokoszczyński

et al. (2002). Because of the recursive calculation requirement a delayed variable is used in the place of the variable expressing the expected output gap.

The papers Budnik et al. (2009), Greszta et al. (2012) refer to the model NECMOD applied in NBP. It is relatively large and advance model constructed to prepare forecasts of main macroeconomic categories – first of all inflation, but also GDP and its components as well as other quantities important to pursue an effective and responsible monetary policy. The NNS-MFG is relatively simple and has been constructed not to prepare so accurate macroeconomic forecasts but to analyze the discussed monetary-fiscal game, especially interactions of monetary and fiscal strategies in the game as well as impact of targets assumed by the authorities on the game solution concepts.

Equations of the recursive model are presented below. Notation in the equations is assumed according to the basic NNS models.

Equation of the output gap

The equation, referring to the dynamic, inter-period version of the IS curve, describes an aggregated demand as the result of the optimal decisions made by a representative consumer. It has the following form:

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 \left(r_t - \pi_t^e - r_t^n \right) + \alpha_3 g_t, \tag{1}$$

where $x_t = y_t - y_t^n$, $g_t = G_t - G_t^n$.

The output gap x_t is defined as the difference of the current real production y_t and its natural level y_t^n in the equilibrium state with the perfectly elastic prices. The production is measured by the real Gross Domestic Product (GDP). A current value of the production gap depends on its delayed value and on the interest rate gap, where the interest rate gap is defined as the difference of the real interest rate and its natural level r_t^n. The real interest rate is calculated as the difference: the nominal interest rate r_t (WIBOR 1M) minus the expected inflation π_t^e.

The proposed model takes additionally into account in the first equation effects of the fiscal policy – an influence of the real budget expenditure G_t in the gap category, i.e. as the deviation from its natural value G_t^n. The natural levels of the product y_t^n, of the interest rate r_t^n and the budget expenditure G_t^n have been calculated using the Hodrick–Prescott filter. The production gap and the budget expenditure gap are defined in two versions as the absolute deviation from the natural value.

Inflation equation

The equation is known as the New Keynesian version of the Phillips curve. It presents a function of the aggregated supply based on price decisions of firms in the conditions of imperfect competition (Calvo, 1983). Inflation depends on the expected inflation π_t^e and on the output gap x_t. The equation has the form:

$$\pi_t = \beta_0 + \beta_1 \pi_{t-1}^e + \beta_2 x_t. \tag{2}$$

Equation of the expected inflation

The expected inflation is explained by its delayed value and by the current inflation. The equation has the form:

$$\pi_t^e = \delta_0 + \delta_1 \pi_{t-1}^e + \delta_2 \pi_t. \tag{3}$$

Equation of the interest rate (Taylor rule)

The equation describes a rule deriving the nominal interest rate by the central bank. The central bank derives the nominal interest rate in reaction on the deviation of inflation from the target π_t^* and on the current economic situation measured by the production gap. In this model it is the inflation target assumed by the National Bank of Poland in the Monetary Policy Guidelines. The equation describes reaction of the central bank according to the Taylor rule and has the form:

$$r_t = \varphi_0 + \varphi_1 r_{t-1} + \varphi_2 (\pi_{t-1} - \pi_{t-1}^*) + \varphi_3 x_{t-1}. \tag{4}$$

The delays are introduced in the equation similarly as in the previous equations. The delays relate to the difference of the current inflation and the target and to the output gap. They are introduced because of the recursive use of the model.

5. MODEL ESTIMATION

The above NNS-MFG model including the equations (1–4) has been estimated as a system of simultaneous equations using the Three-Stage Least Squares Method (3SLS) in the econometric GRETL package. Time series for the Polish economy from the period 2000–2014 (quarterly data) have been used in estimation. Time series used for estimation include relatively long time with the phases of strong and weak policy mix.

The statistical data have been collected from the following sources: the Central Statistical Office of Poland, the National Bank of Poland (NBP), the Ipsos group. Names and description of the exogenous and endogenous variables used for estimation are presented in table 2.

The interest rates WIBOR 1M and WIBOR 3M were considered in the model construction. The interest rate WIBOR 1M was finally assumed according to the research results indicating a stronger reaction of the WIBOR rates on the basic NBP interest rate for shorter maturity times and a lower reaction for the longer times (Janecki, 2012). On the other hand we obtained better estimation results for WIBOR 1M than for WIBOR 3M.

Results of estimation (model variant 1) are as follows (GRETL outputs; standard deviations for the estimated coefficients are given in brackets):

The variables used in the model estimation

Variable	Description
output_gap (endogenous)	The output gap (denoted by x in equations (1–4)) is defined as the difference between the real GDP and the natural level of output presented by the Central Statistical Office (GUS) in time series according to the principles of the "European System of National and Regional Accounts" (ESA); GDP in constant prices. The natural level is calculated as a long term trend of the of GDP using the Hodrick–Prescott filter.
output_gap_1	The output gap, one period delayed
Inflation (endogenous)	Inflation (π) is calculated on the basis of the consumer price index, analogic period of the previous year = 100 (GUS data).
expected_infl (endogenous)	Expected inflation is measured as the average inflation level expected in the next year (NBP, Ipsos data)
expected_infl_1	Expected inflation, one period delayed
WIBOR (exogenous)	The interest rate WIBOR 1M, nominal, at the beginning of each period (data from Money.pl (http://www.money.pl/))
WIBOR_1	The interest rate WIBOR 1M (nominal), one period delayed
WIBOR_gap (exogenous)	The interest rate gap (r^n) is the difference of the real interest rate (WIBOR 1M) and the natural rate, while the real interest WIBOR 1M is derived as the difference of the nominal rate WIBOR (r) and expected inflation (π^ε). The natural (real) interest rate is calculated as a long term trend of the real interest rate using the Hodrick–Prescott filter.
expend_gap (exogenous)	The gap of the expenditure (g) of the public sector means deviation of the real public expenditure (G) from its natural level (G^n). The natural expenditure is calculated as a long term trend of the real public expenditure using the Hodrick–Prescott filter.
infl_target_dif (exogenous)	The difference between inflation (π) and the inflation target (π^*) indicated by NBP (Monetary Policy Guidelines data)
infl_target_dif_1	The difference between inflation and the inflation target, one period delayed

Source: own elaboration.

Equation 1

$$\text{output\_gap} = 0.0273 + 0.6938 \text{ output\_gap\_1} - 0.4243 \text{ WIBOR\_gap} + 0.1376 \text{ expend\_gap}$$
$$\quad\quad (0.1195)\ (0.0840)\quad\quad\quad (0.1247)\quad\quad\quad\quad (0.0621)$$

Equation 2

$$\text{inflation} = 0.5550 + 0.7532 \text{ expected\_infl\_1} + 0.3384 \text{ output\_gap}$$
$$\quad\quad (0.1893)\ (0.0578)\quad\quad\quad\quad (0.0911)$$

Equation 3

expected_infl = – 0.1963 + 0.1431 expected_infl_1 + 0.9152 inflation
                   (0.1028) (0.0587)   (0.0742)

Equation 4

WIBOR = 0.2653 + 0.9227 WIBOR_1 + 0.1967 infl_target_dif_1 + 0.2297 output_gap_1
            (0.0946) (0.0132)   (0.0290)   (0.0346)

Table 3.

Model, equation system, GRETL outputs

Equation 1: Estimation 3SLS, observations 2001:1–2014:4 (N = 56)
Dependent variable (Y): output_gap
Instruments: const output_gap_1 WIBOR_gap expend_gap expected_infl_1 WIBOR_1 infl_target_dif_1

Variable	Coefficient	Std. error	z	p-value
const	0.0272869	0.119475	0.2284	0.8193
output_gap_1	0.693819	0.0840444	8.255	1.51e-016
WIBOR_gap	-0.424272	0.124742	-3.401	0.0007
expend_gap	0.137646	0.0620940	2.217	0.0266

Mean dependent var	-0.113802	S.D. dependent var	1.541290	
Sum squared resid	46.33381	S.E. of regression	0.909610	
R-squared	0.647118	Adjusted R-squared	0.626760	

Equation 2: Estimation 3SLS, observations 2001:1–2014:4 (N = 56)
Dependent variable (Y): inflation
Instruments: const output_gap_1 WIBOR_gap expend_gap expected_infl_1 WIBOR_1 infl_target_dif_1

Variable	Coefficient	Std. error	z	p-value
const	0.554993	0.189279	2.932	0.0034
expected_infl_1	0.753164	0.0577744	13.04	7.61e-039
output_gap	0.338427	0.0911186	3.714	0.0002

Mean dependent var	2.606757	S.D. dependent var	1.678949	
Sum squared resid	36.97738	S.E. of regression	0.812595	
R-squared	0.765484	Adjusted R-squared	0.756635	

Equation 3: Estimation 3SLS, observations 2001:1–2014:4 (N = 56)
Dependent variable (Y): expected_infl
Instruments: const output_gap_1 WIBOR_gap expend_gap expected_infl_1 WIBOR_1 infl_target_dif_1

Variable	Coefficient	Std. error	z	p-value
const	-0.196344	0.102801	-1.910	0.0561
expected_infl_1	0.143050	0.0587171	2.436	0.0148
inflation	0.915195	0.0741551	12.34	5.41e-035

Mean dependent var	2.586357	S.D. dependent var	1.736927	
Sum squared resid	11.16849	S.E. of regression	0.446584	
R-squared	0.936581	Adjusted R-squared	0.934188	

Equation 4: Estimation 3SLS, observations 2001:1–2014:4 (N = 56)
Dependent variable (Y): <u>WIBOR</u>
Instruments: const output_gap_1 WIBOR_gap expend_gap expected_infl_1 WIBOR_1 infl_target_dif

Variable	Coefficient	Std. error	z	p-value
const	0.265344	0.0946242	2.804	0.0050
WIBOR_1	0.922672	0.0131957	69.92	0.0000
infl_target_dif	0.196669	0.0290075	6.780	1.20e-011
output_gap_1	0.229699	0.0345724	6.644	3.05e-011

Mean dependent var	5.858571	S.D. dependent var	3.643124	
Sum squared resid	8.191705	S.E. of regression	0.382466	
R-squared	0.988783	Adjusted R-squared	0.988136	

Source: own elaboration.

Detailed estimation results for the NNS-MFG model are presented in the table 3 (GRETL package, system of equations, 3SLS). The estimation results show an acceptable goodness of fit. All the variables are statistically significant. The R-squared values are greater than 90% in the case of equations 3 and 4. The worse estimation has been obtained in the case of equation 1 and 2 with R-squared values: 63% and 76% respectively. However also in this case all the variables are statistically significant. The estimation inaccuracy is caused by model simplifications. The model describes influences of the economic policies only, it does not describe any influence of exogenous factors in the explicit form. Figure 1 presents matching of the endogenous variables: the theoretical values calculated by the estimated model compared to the empirical values.
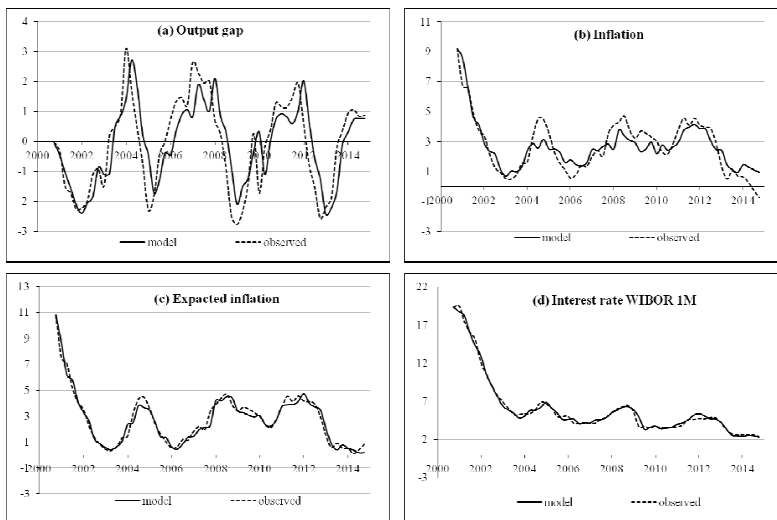

Figure 1. Estimated (3SLS) and observed values of the variables of the model:
(a) output gap, (b) inflation, (c) expected inflation, (d) interest rate WIBOR 1M
Source: own elaboration.

6. ANALYSIS OF THE MONETARY-FISCAL GAME

Strategies, outcomes and payoffs of the game were analyzed using the NNS-MFG model presented in section 4 for the estimation results shown in section 5. Computer simulations have been made for different variants of the model parameters and different initial values of the model variables. Selected simulation results are presented.

Simulation assumptions

The initial state of the economy is represented by the model variables on the basis of the empirical data in the last quarter of 2000. The model variables are calculated using the NNS-MSG model since the first quarter of 2001, while the nominal interest rate is calculated by the Taylor rule and the public expenditure gap according to the statistical data on the real public expenditure. In a selected period (in the presented results: 8 quarters since 1-st quarter of 2008) an impulse changing the policy mix is introduced. The instruments of the policy mix, i.e. real interest rate and the budget deficit in relation to GDP are assumed on a given constant level in this period of time. The nominal interest rate is calculated as the real rate plus expected inflation. The real public expenditure and the budget expenditure gap are calculated on the basis of the budget deficit and the tax rate. After the time the real interest rate is derived according to the Taylor rule and the budget deficit to GDP ratio – according to the ex post data.
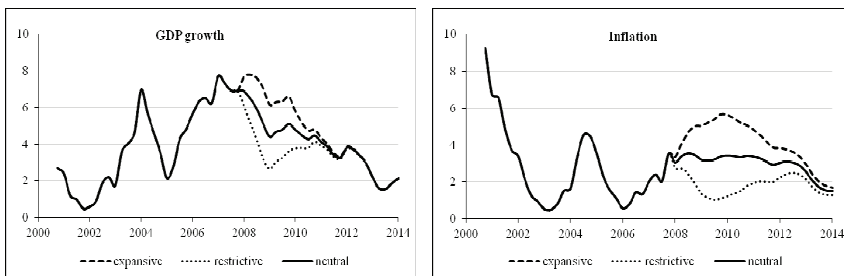

Figure 2. GDP growth rate and inflation in 2000–2014 simulated
for the three variants of policy mix: expansive, restrictive and neutral
Source: own elaboration.

Figure 2 illustrates dynamic effects of the policy mix changed in the considered period of time. Three variants of the policy mix in this period are compared: a policy more expansive than the policy historically implemented (real interest rate was assumed 1 percent points lower and the budget deficit in relation to GDP – 1 percent point greater than the historical values), a policy more restrictive than the historical policy (real interest rate was assumed 1 percent points greater and the budget deficit in relation to GDP – 1 percent point lower than the historical values), and a neutral policy, when the instruments were assumed on the historical level. It can be observed that effects of the introduced changes of the policies are temporary, shorter in the case

of the GDP growth and longer in the case of inflation. The more expansive policy mix results in a greater GDP growth and in a greater inflation in comparison to the neutral path. The effects of the more restrictive policy are reverse.

Payoffs of the monetary-fiscal game being effects of the changed policy mix are measured by the average annual production growth (denoted as y in the game formulation in section 3) and by average annual inflation (denoted by p) in the period of 8 quarters since the changes of the policy mix have been introduced. The payoffs were calculated using the NNS-MFG model relations for the interest rate (r) and budget deficit in relation to GDP (b) treated as exogenous variables in the considered period of time. The functions $h^1(b, r)$ and $h^2(b, r)$ introduced in the game formulation denote the dependence of the payoffs on the policy instruments due to the model relations.

Admissible values of the policies' instruments have been assumed in a form of intervals. The interest rate has been changed in the interval [6%, -1%] and the budget deficit in relation to GDP – in the interval [-1, 6%]. The computer-based system simulating the game and all calculations were made in the MSExcel environment using the VBScript language and the embedded optimization solver. Selected results are presented and discussed below.
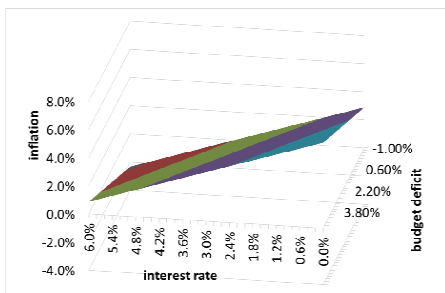

Figure 3. Outcomes of the monetary authority
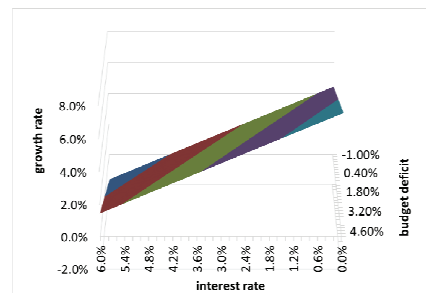Source: own elaboration.

Figure 4. Outcomes of the fiscal authority
Source: own elaboration.

Figures 3 and 4 present the outcomes of the authorities, as dependent on assumed strategies. Inflation (figure 3) can be obtained on a low level when a restrictive monetary policy and a restrictive fiscal policy are applied. More expansive monetary and fiscal policies lead to an increase of inflation and of the economic growth. On the other hand more restrictive monetary and restrictive fiscal policies lead to a decrease of the economic growth (figure 4).

Let us assume that the monetary and fiscal authorities try to achieve given targets of their policies. Let the monetary authority assume the inflation goal on the level p^g, and let the fiscal authority try to achieve the GDP growth rate on the level y^g. Let Ω denotes the set of admissible pairs (b, r) of strategies. The best response strategies of the authorities can be obtained as solutions of the optimization problems:

Min $|h^1(b, r) - y^g|$ with respect to $b \in \Omega^1$ solved for all $r \in \Omega^2$, in the case of the fiscal authority and

Min $|h^2(b,\ r) - p^g|$ with respect to $r \in \Omega^2$, solved for all $b \in \Omega^1$, in the case of the monetary authority, respectively.

Examples of the best response strategies derived for different targets of the authorities are presented in figure 5. Figure 5, part (a) presents the best response strategies of the monetary authority for the three different targets: inflation = 2%, 2.5%, 3%, and the best response strategies of the fiscal authority for the target: GDP growth = 3.5%. Figure 5, part (b) presents the best response strategies of the fiscal authority for the three different targets: GDP growth = 3%, 3.5%, 4%. and the best response strategies of the monetary authority for the target: inflation = 2.5%. The Nash equilibria which are Pareto optimal in the assumed interval of the policies' instruments are shown.
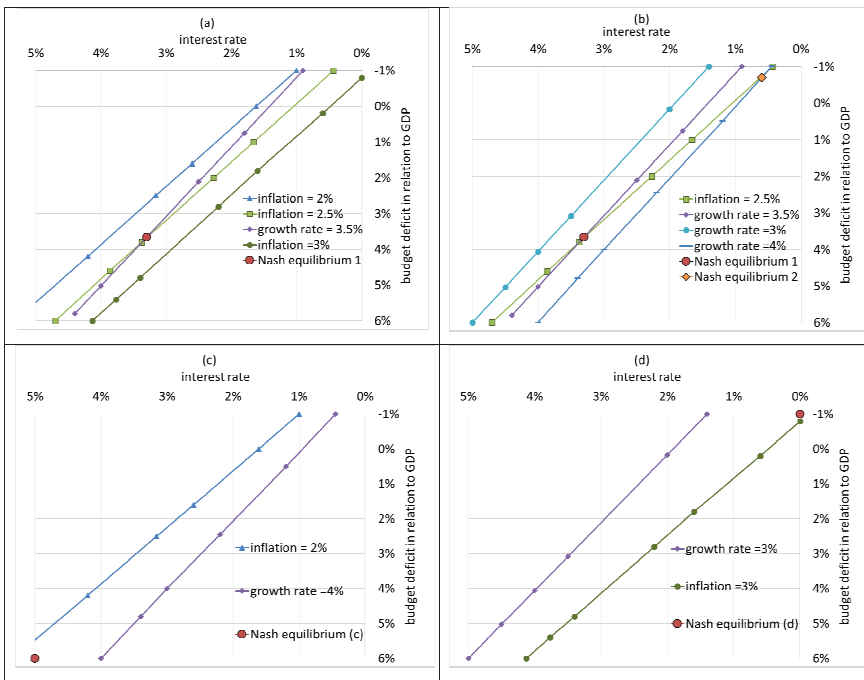
Figure 5. The best response strategies:
(a) for different monetary targets and the fiscal target – growth rate = 3.5%;
(b) for different fiscal targets and the monetary target – inflation = 2.5%;
(c) for the targets: growth rate = 4% and inflation = 2%;
(d) for the targets: growth rate = 3% and inflation = 3%
Source: own elaboration.

It can be observed how the level of restrictiveness/expansiveness of the monetary policy depends on the level of restrictiveness/expansiveness of the fiscal policy. A more expansive fiscal policy leads to a more restrictive monetary policy taken by the central

bank trying to limit inflation excessing the inflation target. If the budget deficit is higher, then the required inflation is obtained for respectively higher interest rates. Analogously, if the government carries out a more restrictive budget policy, then the central bank will apply a less restrictive (more expansive) monetary policy with relatively lower interest rates.

On the other hand a more restrictive monetary policy causes in reaction a more expansive budget policy. If the interest rate is higher, then the required growth rate can be achieved by applying a more expansive fiscal policy supporting a higher growth rate. That means the government should assume a relatively greater budget deficit. Inversely the government can implement more restrictive fiscal policy limiting the budget deficit in reaction on a more expansive monetary policy.

The simulation results show how changes of the targets of fiscal and monetary policies influence on the best response strategies and on the Nash equilibrium state, i.e. on the choice of the respective policy mix. More ambitious target of fiscal policy with a high required economic growth causes that the best response budget strategy moves into more expansive one and vice versa in the opposite case. The higher inflation targets assumed by the monetary authority cause that the best response strategies of the central bank move into more expansive monetary policies. In the opposite case, of the lower inflation target, the best response monetary policy moves into a more restrictive one. Changes of the targets assumed by the fiscal and monetary authorities result in respective positioning of the Nash equilibrium.

There are two cases of the best response strategies in the considered game. The first when the best response strategies cross in the set of admissible strategies, and the second when they do not cross in this set. The cross point in the first case defines the Nash equilibrium. One can easily see that a deviation of any strategy from the point leads to a worse payoff of the respected player (Nash, 1951). In the second case the real Nash equilibrium is out of the set of admissible strategies. The theoretical Nash equilibrium in the set also exists but on the boundary of the set of admissible strategies. One can find also that a deviation from the point leads to worse payoffs of at least one of the players.

Figure 5 (c) illustrates the case when the fiscal and monetary authorities assume too ambitious targets in the given economic state. The monetary authority assumes the restrictive inflation target (2%) and the fiscal authority would like to achieve the high growth (4%). The best response strategies do not cross in the assumed intervals of the policies' instruments. The theoretical Nash equilibrium is at the most restrictive monetary and at the most expansive fiscal policy (the point: $r = 5\%$, $b = 6\%$ in the figure). It is not Pareto optimal. Another example of not ambitious targets is presented in figure 5 (d) for the soft inflation target = 3% and the not demanding fiscal target: GDP growth = 3%. The theoretical Nash equilibrium is in this case at the most expansive monetary and the restrictive fiscal policy (the point: $r = 0\%$ and $b = -0.9\%$ in the figure).

The results show that efficiency of the policy mix depends on the targets assumed by the monetary and fiscal authorities, therefore the respective coordination of the targets will be beneficial for both authorities.

The Nash equilibrium (Nash 1) in figure 5 describes the state which can be compared to the real state of the economy in Poland in the analyzed period of 8 quarters 2008:1–2009:4. Comparing the obtained results to the historical data (average annual growth rate = 3.3% and average annual inflation = 3.8%) in this period one can state that the calculated equilibrium state indicates possibility of policies giving better economic effects: a greater growth rate = 3.5% and lower annual inflation = 2.5%. More expansive policy mix, especially the fiscal policies could be applied with the real interest rate = 3.3%, close to the historical one equal to 3.4% on the average in the period 2008–2009, and with the budget deficit to GDP ratio = 3.8% greater than the historical one equal to 2.7% on the average in this period. It is an open question why the authorities did not carry out these possible better policies. The decision making process was in this time rather difficult, when negative effects of economic recession at the end of 2008 and in 2009 were observed. More general analysis of alternative policies mix in comparison to the historical ones carried out in Poland is included in a separate paper Kruś, Woroniecka-Leciejewicz (2016).

The obtained simulation results indicate a possibility of a case when the Nash equilibrium is Pareto optimal, but also a case when the Nash equilibrium leads to the solution not beneficial for one or both players. The last case is known in the literature as the prisoners' dilemma, when the policy coordination is desired. It is shown in the paper that the respective coordination of monetary and fiscal targets may lead to Pareto optimal Nash equilibria and effective policy mix.

7. CONCLUSIONS

This paper presents selected results of the research dealing with mutual interactions of the monetary and fiscal policies. The results have been obtained using the game theory and optimization methods. A dynamic macroeconomic model, called NNS-MFG model, has been formulated and estimated using the statistical data for Poland. A noncooperative monetary-fiscal game has been formulated in which payoffs of players – namely monetary and fiscal authorities are calculated using the model equations. The results for the monetary-fiscal game presented in this paper are the first obtained for Poland with the use of a macroeconomic model.

The macroeconomic NNS-MFG model describes influences of the instruments of the monetary and fiscal policies on the state of the economy, i.e. influences of the real interest rate and of the budget deficit in relation to GDP on the growth rate and inflation. It is based on concept of the New Neoclassical Synthesis model. It includes four equations describing the production gap, inflation, expected inflation and the Taylor rule.

The basic NNS model describes a transmission of the monetary policy impulses. In comparison to the basic NNS model, this model has been extended to describe influences of the fiscal policy. It takes into account the budget expenditure gap. The model parameters have been estimated using time series 2000–2014. The Three-Stage Least Squares method in the GRETL package has been used for the model treated as a system of simultaneous linear equations. The estimated model has been implemented in the form of a recursive algorithm in a computer-based system. The system calculates the payoffs of players and other variables of the model dependently on strategies implemented by the players. The system derives also the best response strategies dependently on the targets assumed by the monetary and fiscal authorities, as well as the Nash equilibria and the Pareto optimal outcomes.

A number of simulations has been made and the obtained results have been analyzed. The system derives detailed quantitative results, but also some qualitative conclusions can be formulated. There are some values of the targets assumed by the monetary and fiscal authority, for which the best response strategies cross in the interval of assumed admissible values of the instruments. The cross point relates to the Nash equilibrium and the equilibrium is Pareto optimal. However for some values of the targets the Nash equilibrium can be non Pareto optimal. For example in the case of too ambitious targets of the authorities the Nash equilibrium shifts into the most restrictive monetary policy and the most expansive fiscal policy what leads to non-effective non Pareto optimal solutions. The results show that in such cases a coordination of the monetary and fiscal policies is required.

Summarizing, the main result of the presented research consists in construction of a computer-based tool supporting analysis of the monetary fiscal game with the use of a respective macromodel estimated for Poland. More detailed results include the proposed and constructed NNS-MFG model, estimation of the model parameters using the statistical data for Poland, procedures calculating payoffs of the proposed monetary-fiscal game, formulation of respective optimization problems and procedures deriving the best response strategies of players, numerical results discussed above. The presented numerical results illustrate only selected features of the system.

The system may support looking for the Pareto-optimal consensus of the authorities in the policy mix problem. It can be checked when the targets assumed by the fiscal and monetary policies lead to the Pareto-optimal Nash equilibrium and the equilibrium can be derived. On the other hand one can check when the priorities of the monetary and fiscal authorities lead to non-effective Nash equilibria and when coordination of monetary and fiscal policies is required.

The constructed macroeconomic model is relatively simple, but the proposed approach can be applied also for more extended versions of the model. Such an extended nonlinear model describing the influence of the policy mix instruments on the economy in a more adequate way is planned. Further research include also analysis of the problem using dynamic game concepts when a sequence of decisions made by the authorities is considered. Another direction deals with development

of multicriteria optimization tools supporting analysis and consensus seeking. The methods of multicriteria bargaining support proposed in Kruś (2011, 2014) can be applied to construct such optimization tools.

REFERENCES

Arestis P., (2009), *New Consensus Macroeconomics: A Critical Appraisal*, Working Paper, 05–09, Cambridge Centre for Economic and Public Policy.

Batini N., Haldane A., (1999), Forward-Looking Rules for Monetary Policy, *Bank of England Working Paper* No. 91.

Bennett H., Loayza N., (2001), Policy Biases when the Monetary and Fiscal Authorities have Different Objectives, *Central Bank of Chile Working Papers*, no. 66.

Blackburn K., Christensen M., (1989), Monetary Policy and Policy Credibility: Theories and Evidence, *Journal of Economic Literature*, 27, 1–45.

Blanchard O., (2009), The State of Macro, *Annual Review of Economics*, 1, September, 209−228.

Blinder A. S., (1983), Issues in the Coordination of Monetary and Fiscal Policy, in: *Monetary Policy in the 1980s*, Federal Reserve Bank of Kansas City, 3–34.

Blinder A. S., (2000), Central Bank Credibility, Why Do We Care? How Do We Build It?, *American Economic Review*, December 2000, 1421–1431.

Bludnik I., (2010), Nowa synteza neoklasyczna w makroekonomii (The New Neoclassical Synthesis in the Macroeconomics), Bank i Kredyt, 41 (2), 43–70.

Budnik K., Greszta M., Hulej M., Kolasa M., Murawski K., Rot M., Rybaczyk B., Tarnicka M., (2009), The New Macroeconometric Model of the Polish Economy, National Bank of Poland Working Paper, No. 64, Warszawa.

Calvo G. A., (1983), Staggered Prices in a Utility-Maximizing Framework, *Journal of Monetary Economics*, 12 (3), 383−398.

Cevik E. I., Dibooglu S., Kutan A. M., (2014), Monetary and Fiscal Policy Interactions: Evidence from Emerging European Economies, *Journal of Comparative Economics*, 42 (4), 1079–1091.

Darnault N., Kutos P., (2005), Poland's Policy Mix: Fiscal or Monetary Leadership?, *ECFIN Country Focus*, 2 (1), 1–6.

Demertzis M., Hughes Hallet A., Viegi N., (2002), An Indepeded Central Bank Faced with Elected Governments: A Political Economy Conflict, Research Memorandum WO no. 686, De Nederlandsche Bank, May 2002.

Dixit A., Lambertini L., (2001), Monetary-Fiscal Policy Interactions and Commitment versus Discretion in a Monetary Union, *European Economic Review*, 45, 977–987.

Eijffinger S. W., DeHaan J., (1996), *The Political Economy of Central Bank Independence*, Princeton University, Princeton.

Freitas P. S., Muinhos M. K., (2001), A Simple Model for Inflation Targeting in Brazil, Working Paper Series, no. 18, Banco Central do Brasil.

Galí J., (2009), *The New Keynesian Approach to Monetary Policy Analysis: Lessons and New Directions*, http://www.crei.cat/people/gali/jg09cfs.pdf.

Gjedrem S., (2001), Monetary Policy - the Importance of Credibility and Confidence, *BIS Review*, 7, 1–13.

Goodfriend M., King R. G., (1997), The New Neoclassical Synthesis and the Role of Monetary Policy, in: Bernanke B., Rotemberg J., (eds.), *NBER Macroeconomics Annual*, MIT Press, Cambridge.

Goodhart C., (2007), *Whatever Became of the Monetary Aggregates?*, Peston Lecture in Honour of Maurice, Lord Peston, delivered at Queen Mary College, London.

Greszta M., Hulej M., Lewińska R., Michałek A., Pońsko B., Schultz B., (2012), Reestymacja kwartalnego modelu gospodarki polskiej NECMOD 2012 (Reestimation of the Quarterly Model of the Polish Economy NECMOD 2012), Narodowy Bank Polski, Warszawa.

http://www.nbp.pl/home.aspx?f=/polityka_pieniezna/dokumenty/necmod.html

Janecki J., (2012), Reakcja rynkowych stóp procentowych na zmiany stopy procentowej banku centralnego w Polsce w latach 2001–2011 (Reaction of Market Interest Rates on the NBP Interest Rates in Poland, 2001–2011), *Materiały i Studia NBP*, 272, Warszawa.

Kokoszczyński R., Łyziak T., Pawłowska M., Przystupa J., Wróbel E., (2002), Mechanizm transmisji polityki pieniężnej – współczesne ramy teoretyczne, nowe wyniki empiryczne dla Polski (The Transmission Mechanism of Monetary Policy – A Contemporary Theoretical Framework, New Empirical Results for Poland), *Materiały i Studia NBP*, 51, Warsaw.

Kruś L., (2011), Wielokryterialne decyzje kooperacyjne, metody wspomagania komputerowego (Multicriteria Cooperative Decisions, Computer Based Support), Instytut Badań Systemowych PAN, *Seria: Badania systemowe*, 70, Warsaw.

Kruś L., (2014), Computer Based Support in Multicriteria Bargaining with Use of the Generalized Nash Solution Concepts, in: *Modern Approaches in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Networks and Related Topic*s, Vol. II: Applications, K. Atanassov et al., (eds.), SRI PAS, Warsaw, 43–60.

Kruś L., Woroniecka-Leciejewicz I., (2015), Fiscal-Monetary Game Analyzed with Use of a Dynamic Macroeconomic Model, in: Kamiński B., Kersten G., Szapiro T., (eds.), GDN 2015, Springer LNBIP 218, 199–208.

Lambertini L., Rovelli R., (2003), Monetary and Fiscal Policy Coordination and Macroeconomic Stabilization. A Theoretic Analysis, Bolonia.

Libich J., Nguyen D., (2015), Strategic Monetary-Fiscal Interactions in a Downturn, *Economic Record*, 91 (293), 172–190.

Libich J., Nguyen D., Stehlik P., (2015), Monetary Exit and Fiscal Spillovers, *European Journal of Political Economy*, 40, 184–206.

Marszałek P., (2005), Zastosowanie teorii gier do badania koordynacji polityki pieniężnej i polityki fiskalnej (Game Theory Applied in the Study of Coordination of the Monetary and Fiscal Policy), in: Studia z bankowości centralnej, Przybylska-Kapuścińska W., (ed.), *Zeszyty Naukowe AE w Poznaniu*, 56, 224–247, Poznań, 2005.

Marszałek P., (2009), *Koordynacja polityki pieniężnej i fiskalnej jako przesłanka stabilności poziomu cen* (Coordination of the Monetary and Fiscal Policies as the Prerequisite for the Price Level Stability), PWN, Warsaw.

Monetary and Fiscal Policies to Head off Overheating, in: *OECD Economic Surveys: Poland 2008*, OECD Publishing, 47–71.

Muinhos M. K., (2001), Inflation Targeting in an Open Financially Integrated Emerging Economy: the Case of Brazil, Working Paper Series, no. 26, Banco Central do Brasil.

Nash J. F., (1951), Non-cooperative Games, *Annals of Mathematics*, 54 (2), 286–295.

Nordhaus W. D., (1994), Policy Games: Coordination and Independence in Monetary and Fiscal Policies. *Brookings Papers on Economic Activity*, 2, 139–215.

Sargent T., Wallace N., (1981), Some Unpleasant Monetarist Arithmetic, *Federal Reserve Bank of Minneapolis Quarterly Review*, 5, 1–17.

Spahn H.-P. (2009), *The New Keynesian Microfoundation of Macroeconomics*, Diskussionspapiere, 317, Institut für Volkswirtschaftslehre der Universität Hohenheim.

Stawska J., (2014), Wpływ policy-mix na wzrost gospodarczy i poziom bezrobocia w Polsce (An Influence of Policy-Mix on the Economic Growth and Unemployment in Poland), Finanse, Zeszyty Naukowe Uniwersytetu Szczecińskiego, 804, Rynki Finansowe, Ubezpieczenia, 67, 667–677.

Taylor J. B., (1993), Discretion versus Policy Rules in Practice, Carnegie-Rochester [1] *Conference Series on Public Policy*, 39, 195–214.

Walsh C., (2001), Transparency in Monetary Policy, FRBSF Economic Letter 2001, 26 (September 7).

Wierzbicki A. P., (1986), On the Completeness and Constructiveness of Parametric Characterizations to Vector Optimization Problems, *OR Spectrum*, Springer Verlag, 8, 73–87.

Wierzbicki A. P., Kruś L., Makowski M., (1993), The Role of Multi-Objective Optimization in Negotiation and Mediation Support, *Theory and Decision*, 34 (2), 201–214.

Wierzbicki A. P., Makowski M., Wessels J., (2000), *Model-based Decision Support Methodology with Environmental Applications,* Kluwer Academic Press, Dordrecht, Boston.

Wojtyna A., (1996), Niezależność banku centralnego a teoretyczne i praktyczne aspekty koordynacji polityki pieniężnej i fiskalnej (Independence of the Central Bank and the Theoretical and Practical Aspects of Coordination of the Monetary and Fiscal Policy), Bank i Kredyt, 6, 87–101.

Wojtyna A., (1998), Szkice o niezależności banku centralnego (Essays on the Independence of the Central Bank), PWN, Warsaw.

Woodford M., (2003), *Interest and Prices*, Princeton University Press, Princeton.

Woroniecka I., (2007), Analiza priorytetów banku centralnego w polityce stóp procentowych (Analysis of Priorities of the Central Bank in the Interest Rate Policy), *Ekonomista*, 4, 559–580.

Woroniecka-Leciejewicz I., (2008), Dylemat więźnia i inne przypadki grze monetarno-fiskalnej fiskalnej (Prisoner's Dilemma and Other Cases in the Monetary-Fiscal Game), in: Owsiński J. W., Nahorski Z., Szapiro T., (eds.), *Badania operacyjne i systemowe: decyzje, gospodarka, kapitał ludzki i jakość*, Seria Badania Systemowe, 64, IBS PAN, Warsaw, 161–172.

Woroniecka-Leciejewicz I., (2010), Decision Interactions of Monetary and Fiscal Authorities in the Choice of Policy Mix. The special issue *Journal of Organisational Transformation and Social Change*: „Corruption and Good Governance", Intellect - Publishers of Original Thinking, UK, 7 (2), 189–210.

Woroniecka-Leciejewicz I., (2015a), Equilibrium Strategies in a Fiscal-monetary Game. A Simulation Analysis, *Operations Research & Decisions*, 25 (2), 75–100.

Woroniecka-Leciejewicz I., (2015b), Wpływ instrumentów policy mix na gospodarkę – ujęcie modelowe (The Influence of Policy Mix Instruments on the Economy – A Modeling Approach), Zeszyty Naukowe WSISiZ, *Współczesne Problemy Zarządzania*, 7–33.

ANALIZA GRY MONETARNO-FISKALNEJ Z WYKORZYSTANIEM MAKROEKONOMICZNEGO MODELU DLA POLSKI

S t r e s z c z e n i e

W pracy sformułowano i przeanalizowano grę monetarno-fiskalną opisującą interakcje między władzami monetarnymi i fiskalnymi. Każdy z graczy stara się osiągnąć założony cel: władze monetarne – pożądany poziom inflacji, władze fiskalne – pożądany wzrost gospodarczy. Zbudowano model makroekonomiczny oparty na koncepcji nowej syntezy neoklasycznej. Model ten, zawierający cztery podstawowe równania: luki popytowej, inflacji, inflacji oczekiwanej i regułę Taylora, został rozbudowany z uwzględnieniem wydatków budżetowych. Dokonano estymacji modelu na dla gospodarki polskiej na podstawie szeregów czasowych 2000–1014.

Zbudowano system komputerowy wyznaczający wypłaty gry. Przeprowadzono szereg symulacji z wykorzystaniem zbudowanego modelu. Przedstawiono skutki polityk monetarnych i fiskalnych alternatywnych do realizowanych w przeszłości. Wyznaczono i przeanalizowano optymalne strategie odpowiedzi oraz strategie równowagi Nasha. Wskazano na przypadki, gdy równowagi Nasha są Pareto optymalne oraz takie, które prowadzą do wyboru skrajnie restrykcyjnych lub skrajnie ekspansywnych polityk, a uzyskiwana równowaga nie jest Pareto optymalna. Wskazana jest wtedy koordynacja polityk.

Słowa kluczowe: gar monetarno-fiskalna, model makroekonomiczny, policy mix, równowaga Nasha, Pareto optymalność

MONETARY-FISCAL GAME ANALYZED USING A MACROECONOMIC MODEL FOR POLAND

A b s t r a c t

In the paper a monetary-fiscal game is formulated and analyzed. It describes interactions of the monetary and fiscal authorities. Each authority tries to achieve its own goal: the fiscal authority – assumed GDP growth, and the monetary authority – an inflation level. A macroeconomic model for the Polish economy has been formulated on the basis of the New Neoclassical Synthesis concept and respectively extended to describe effects of the fiscal instruments. The model parameters have been estimated using statistical data of the Polish economy from the period 2000–2014.

A computer-based system calculating results of the game has been constructed using the above model. A sequence of simulations have been made in which payoffs of the game were derived for alternative monetary and fiscal policies. Results of the policy mix strategies alternative to the historical policies in Poland were also considered.

In the paper the best response strategies of the authorities and the Nash equilibria are analyzed when the authorities assume independently their goals. The simulation results are presented and discussed for the cases when the Nash equilibria is Pareto optimal but also when it is not Pareto optimal. It is shown also that the best response strategies may lead to the extremely restrictive or expansive policies.

Keywords: monetary-fiscal game, macroeconomic model, policy mix, Nash equilibrium, Pareto optimality

JAN PURCZYŃSKI[1], KAMILA BEDNARZ-OKRZYŃSKA[2]

THE RAYBIT MODEL AND THE ASSESSMENT OF ITS QUALITY IN COMPARISON WITH THE LOGIT AND PROBIT MODELS

1. INTRODUCTION

A prevailing amount of methods of econometric model analysis refers to the situation when variables (both dependent and explanatory) are continuous variables. This is the case of a quantitative model – a quantitative dependent variable. If the variable can take a finite number of values, it is referred to as a discrete or a qualitative variable. Gruszczyński (2012) draws attention to an increasing importance of qualitative models, as they constitute a basic tool for describing microeconometric models used in empirical corporate finance. In the case when a variable takes only two values it is called a dichotomous variable (also binomial or binary variable).

The simplest method of solving an equation with a binary dependent variable is a linear probability model, the solution of which has one vital drawback, namely, a possibility of obtaining the probability which falls outside the interval [0;1] (Maddala, 1992). In order to get rid of this drawback, it is assumed that the probability corresponds to the cumulative distribution function of a random variable. In the case of a logistic distribution, a logit model is obtained, and in the case of a normal distribution – a probit model (Maddala, 1992).

In the literature other types of transformations can be found. Nerlove (1973) provides the following formulas:

$$F(x) = \frac{1}{2}\left(1 + \sin x\right), \quad -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}, \tag{1}$$

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan x, \quad -\infty < x < \infty, \tag{2}$$

$$F(x) = \tanh x, \quad -\infty < x < \infty. \tag{3}$$

---

[1] University of Szczecin, Faculty of Management and Economics of Services, Department of Quantitative Method, 8 Cukrowa St., 71-004 Szczecin, Poland, corresponding author – e-mail: jan.purczynski@wzieu.pl.

[2] University of Szczecin, Faculty of Management and Economics of Services, Department of Quantitative Method, 8 Cukrowa St, 71-004 Szczecin, Poland.

McFadden (1984) lists the following transformations: the cumulative distribution function of the Student's t-distribution, the cumulative distribution function of the Cauchy distribution and the arctan model (equation (2)). Finney (1973) lists four transformations: the arctan model, the rational function, the $\sin^2 x$ and parabolic function.

In this paper the author proposes his own model, in which the probability is expressed by a Rayleigh cumulative distribution function, hence the name of the model – raybit. The Rayleigh distribution is a special case of the Weibull distribution, the cumulative distribution function of which is given by (Rine, 2009):

$$F(x) = 1 - \exp\left[-\left(\frac{x-a}{b}\right)^c\right]. \tag{4}$$

By assuming in equation (4) $a = 0$, $b = 1$ and $c = 2$, the Rayleigh cumulative distribution function is obtained:

$$F(x) = 1 - \exp\left(-x^2\right). \tag{5}$$

While conducting computer simulations described in section 5 of the paper, it was observed that the values of parameters a nad b (equation (4)) do not affect the results of the proposed method. In order to obtain the simplest form of the Rayleigh cumulative distribution function (equation (5)), $a = 0$ and $b = 1$ were assumed.

The continuous random variable with a Weibull distribution (Rayleigh) has been widely applied in modeling physical and economic phenomena (Polakow, Dunne, 1999; Celik, 2003).

The random variable with a Weibull distribution is also applied in binary variable analysis, yet the cumulative distribution function is given by a relation other than equation (4) (Chou, 1983):

$$F(x) = \exp\left[-\exp\left(-x\right)\right]. \tag{6}$$

This misunderstanding is explained by Train (2009), namely, the distribution (6) is also called Gumbel and type I extreme value, and quite often is mistakenly referred to as the Weibull distribution. The Gumbel distribution is often used in modeling extreme values (Koutsoyiannis, 2003).

Hence it can be concluded that the Rayleigh distribution (equation (5)) has not been used in modeling a discrete variable yet.

2. PROBABILITY MODELS FOR A BINARY VARIABLE

It is assumed that a variable Y can take two values: one or zero, corresponding to the fact of making or not making a decision – an occurrence of an event A.

The subject of the analysis are the models of a binary variable for grouped data.

If among n_i of decision-makers, y_i of them made a sensible decision, then a quotient

$$p_i = \frac{y_i}{n_i}, \ (i = 1,2,...,I) \tag{7}$$

represents an empirical frequency of making a decision in an i-th group of decision-
-makers.

The easiest model is a linear model of probability (Judge et al., 1980):

$$\boldsymbol{p} = \boldsymbol{X\alpha} + \boldsymbol{\varepsilon}, \tag{8}$$

where:
$\boldsymbol{p}$ – I-dimensional vector of empirical probabilities,
$\boldsymbol{X}$ – $[I \times (k+1)]$ dimensional matrix including k number of explanatory variables,
$\boldsymbol{\alpha}$ – $(k+1)$ vector of parameters,
$\boldsymbol{\varepsilon}$ – I-dimensional vector of random elements.
Based on equation (8), the following can be observed

$$p_i = P_i + \varepsilon_i, \tag{9}$$

where:
p_i – empirical probability of an occurrence of an event A for an i-th value of a vector
of explanatory variables,
P_i – probability of an occurrence of an event A for an i-th value of a vector of explana-
tory variables,
ε_i – a disturbance: $E(\varepsilon_i) = 0$ and $cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$.
Since a variable y_i (equation (7)) has a binomial distribution, the variance of
a disturbance is given by relation (Judge et al., 1980)

$$V(\varepsilon_i) = \frac{P_i(1 - P_i)}{n_i}, \tag{10}$$

which means that the disturbances appearing in equation (9) are heteroskedastic.

Due to the drawback mentioned in the Introduction of this paper (p. 1), the linear
probability model will not be further discussed.

It is assumed that the probability P_i, with which the decision in question is made
in an i-th group of decision-makers, is a function F of a variable $\boldsymbol{x_i}^T\boldsymbol{\alpha}$

$$P_i = F(\boldsymbol{x_i}^T\boldsymbol{\alpha}), \tag{11}$$

where F is a cumulative distribution function, $\boldsymbol{x_i}^T$ is an i-th row of an explanatory
variable matrix.

The most commonly applied cumulative distribution functions are as follows:
–   a logit model, hereafter referred to as LOG

$$P_i = L(x_i^T \boldsymbol{\alpha}) = \left[1 + e^{-x_i^T \boldsymbol{\alpha}} \right]^{-1}, \tag{12}$$

where L denotes the cumulative distribution function of a logistic distribution
–   a probit model, hereafter referred to as PRO

$$P_i = \Phi(x_i^T \boldsymbol{\alpha}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_i^T \boldsymbol{\alpha}} e^{-\frac{t^2}{2}} dt, \tag{13}$$

where Φ denotes the cumulative distribution function of a standardized normal distribution.

Depending on the model, a vector v is called:

–   an observed logits $\qquad\qquad v_i = \ln\left(\frac{p_i}{1-p_i} \right),$ $\qquad\qquad$ LOG $\quad$ (14)

–   an observed probits $\qquad\qquad v_i = \Phi^{-1}(p_i),$ $\qquad\qquad$ PRO $\quad$ (15)

where $\Phi^{-1}(\ )$ – the inverse function to the cumulative distribution function of a standardized normal distribution.

The following relations can be observed (Amemiya, 1981; Judge et al., 1980):

$$v_i = x_i^T \boldsymbol{\alpha} + u_i , \;\; E(u_i) = 0,$$

–   for the logit model $\qquad\qquad V(u_i) = \frac{1}{n_i P_i(1-P_i)},$ $\qquad\qquad$ (16)

–   for the probit model $\qquad V(u_i) = \frac{P_i(1-P_i)}{n_i \left\{ \varphi \left[\Phi^{-1}(P_i) \right] \right\}^2},$ $\qquad\qquad$ (17)

where φ – a standard normal density.
In this paper, the Rayleigh cumulative distribution function, given by equation (5), is considered as a function F in equation (11),

$$P_i = R(x_i^T \boldsymbol{\alpha}) = 1 - \exp\left[-\left(x_i^T \boldsymbol{\alpha} \right)^2 \right]. \tag{18}$$

A vector v of the observed raybit is given by formula:

$$v_i = \sqrt{-ln(1 - p_i)}.$$
(19)

From equations (18) and (19) it follows that

$$R^{-1}(P_i) = x_i^T \alpha = \sqrt{-\ln(1 - P_i)}.$$
(20)

Starting from equation

$$\sqrt{-ln(1 - p_i)} = \sqrt{-\ln(1 - P_i - \varepsilon_i)},$$

and adopting approximate formulas, applicable for small values δ ($\delta \approx 0$):

$$\ln(1 \pm \delta) \approx \pm \delta,$$

$$\sqrt{1 + \delta} \approx 1 + \frac{1}{2}\delta,$$

the following is derived

$$R^{-1}(p_i) = x_i^T \alpha + \frac{\varepsilon_i}{2 \cdot (1 - P_i) \cdot \sqrt{-\ln(1 - P_i)}} = x_i^T \alpha + \eta_i.$$
(21)

From equations (10) and (21), the following is obtained:

$$V(\eta_i) = \frac{P_i}{4n_i(1 - P_i)\ln\dfrac{1}{1 - P_i}}.$$
(22)

Which means that the random variable η_i is heteroskedastic.

In the analysis of each model the following three steps can be singled out (Judge et al., 1980; Jajuga, 1989):

A. The first step
Estimation of a vector α_0 of parameters α

$$\alpha_0 = \left(X^T W^{-1} X\right)^{-1} X^T W^{-1} v,$$
(23)

where: W is a diagonal covariance matrix (of a size $I \times I$), where the elements on the main diagonal equal:

$$w_i = \left[n_i p_i (1 - p_i)\right]^{-1},$$ LOG   (24)

$$w_i = \frac{p_i(1 - p_i)}{n_i\left\{\varphi\left[\Phi^{-1}(p_i)\right]\right\}^2}, \qquad\qquad \text{PRO} \quad (25)$$

$$w_i = \frac{p_i}{4n_i(1 - p_i)\ln\dfrac{1}{1 - p_i}}. \qquad\qquad \text{RAY} \quad (26)$$

The estimation of theoretical probability:

$$p_{0i} = L\!\left(x_i^T \alpha_0\right), \qquad\qquad \text{LOG} \quad (27)$$

$$p_{0i} = \Phi(x_i^T \alpha_0), \qquad\qquad \text{PRO} \quad (28)$$

$$p_{0i} = R(x_i^T \alpha_0). \qquad\qquad \text{RAY} \quad (29)$$

B. The second step
By applying the ordinary least squares (OLS), the following is obtained:

$$\alpha_1 = \left(X^T X\right)^{-1} X^T v, \qquad\qquad (30)$$

where: v is defined by formulas (14), (15), (19).
The estimation of theoretical probability

$$p_{1i} = L\!\left(x_i^T \alpha_1\right), \qquad\qquad \text{LOG} \quad (31)$$

$$p_{1i} = \Phi(x_i^T \alpha_1), \qquad\qquad \text{PRO} \quad (32)$$

$$p_{1i} = R(x_i^T \alpha_1). \qquad\qquad \text{RAY} \quad (33)$$

C. The third step
Estimation of a vector α_2 of parameters α

$$\alpha_2 = \left(X^T W_1^{-1} X\right)^{-1} X^T W_1^{-1} v, \qquad\qquad (34)$$

where: v is defined by formulas (14), (15), (19).
W_1 is a diagonal covariance matrix, where the elements on the main diagonal equal:

$$w_{1i} = \left[n_i p_{1i}(1 - p_{1i})\right]^{-1}, \qquad\qquad \text{LOG} \quad (35)$$

$$w_{1i} = \frac{p_{1i}(1 - p_{1i})}{n_i\left\{\varphi\left[\Phi^{-1}(p_{1i})\right]\right\}^2}, \qquad\qquad \text{PRO} \quad (36)$$

$$w_{1i} = \frac{p_{1i}}{4n_i\left(1 - p_{1i}\right)\ln\dfrac{1}{1 - p_{1i}}}, \qquad\qquad \text{RAY} \quad (37)$$

where p_{1i} is given by equations (31), (32) and (33).

The estimation of theoretical probability:

$$p_{2i} = L\left(x_i^T \boldsymbol{\alpha_2}\right), \qquad\qquad \text{LOG} \quad (38)$$

$$p_{2i} = \Phi(x_i^T \boldsymbol{\alpha_2}), \qquad\qquad \text{PRO} \quad (39)$$

$$p_{2i} = R(x_i^T \boldsymbol{\alpha_2}). \qquad\qquad \text{RAY} \quad (40)$$

In the literature (Judge et al., 1980; Jajuga, 1989) there are two alternative methods described: the probability p_0 and the probability p_2 (in which case the probability p_1 is used to determine p_2). In this paper the probability p_1 is taken into account in the same way as p_0 and p_2.

The forms of the likelihood function for the logit and probit models can be found in the paper by Chow (1983). In the case of the raybit model, the likelihood function is given by:

$$L(\boldsymbol{\alpha}) = \prod_{i=1}^{I}\left[1 - e^{-\left(x_i^T\boldsymbol{\alpha}\right)^2}\right]^{n_i'}\left[e^{-\left(x_i^T\boldsymbol{\alpha}\right)^2}\right]^{n_i - n_i'},$$

where n_i' is the number of decision-makers for whom a variable $y_i = 1$ (equation (7)).

The log-likelihood function of the model is given by:

$$lnL(\alpha_0, \alpha_1) = \sum_{i=1}^{I} n_i' \cdot ln\left[1 - e^{-(\alpha_0 + \alpha_1 x_{1i} + \ldots, \alpha_k x_{ki})^2}\right] - \sum_{i=1}^{I}\left(n_i - n_i'\right)\left(\alpha_0 + \alpha_1 x_{1i} + \ldots \alpha_k x_{ki}\right)^2. (41)$$

The necessary condition for extremum leads to the set of equations:

$$\sum_{i=1}^{I}\left\{\frac{n_i'}{1 - \exp\left[-\left(\alpha_0 + \alpha_1 x_{1i} + \ldots \alpha_k x_{ki}\right)^2\right]} - n_i\right\}\left(\alpha_0 + \alpha_1 x_{1i} + \ldots \alpha_k x_{ki}\right) = 0, \quad (42a)$$

$$\sum_{i=1}^{I}\left\{\frac{n_i'}{1 - \exp\left[-\left(\alpha_0 + \alpha_1 x_{1i} + \ldots \alpha_k x_{ki}\right)^2\right]} - n_i\right\}\left(\alpha_0 + \alpha_1 x_{1i} + \ldots \alpha_k x_{ki}\right) \cdot x_{1i} = 0, \quad (42b)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \;,$$

$$\sum_{i=1}^{I}\left\{\frac{n_i'}{1 - \exp\left[-\left(\alpha_0 + \alpha_1 x_{1i} + \ldots \alpha_k x_{ki}\right)^2\right]} - n_i\right\}\left(\alpha_0 + \alpha_1 x_{1i} + \ldots \alpha_k x_{ki}\right) \cdot x_{ki} = 0. \quad (42c)$$

3. ESTIMATING THE ERROR OF THE MODEL

The most popular measure of goodness of fit of a model is the mean square error (MSE):

$$MSE = \frac{1}{I} \sum_{i=1}^{I} (p_i - \hat{p}_i)^2,$$

(43)

where:
p_i – empirical probability (equation (7)),
$\hat{p}_i$ – the estimation of theoretical probability.

As $\hat{p}_i$ the results of the following four methods (p_{0i}, p_{1i}, p_{2i}, p_{MLi}) are taken.

Guzik et al. (2005) recommends equation (43) as a criterion of goodness of fit of a theoretical probability model.

Another measure is the mean absolute error (MAE):

$$MAE = \frac{1}{I} \sum_{i=1}^{I} |p_i - \hat{p}_i|.$$

(44)

Due to the heteroskedasticity of the disturbance, many authors (cf. Amemiya, 1981; Jajuga, 1989; Maddala, 2006) propose a criterion called the Weighted Mean Squared Error (WMSE):

$$WMSE = \sum_{i=1}^{I} \frac{n_i (p_i - \hat{p}_i)^2}{p_i (1 - p_i)}.$$

(45)

The main problem lies in the fact that the variance of MSE (equation (43)) and MAE (equation (44)) depends heavily on the value of the empirical probability. Therefore, a recommended measure of goodness of fit is the weighted mean squared error (equation (45)). This issue was discussed in the paper by Purczyński et al. (2015), where computer simulations were carried out using a random number generator with a binominal distribution. As a result of these studies, yet another measure of goodness of fit was proposed, namely the Weighted Mean Absolute Error (WMAE):

$$WMAE = \sum_{i=1}^{I} \frac{n_i |p_i - \hat{p}_i|}{\sqrt{p_i (1 - p_i)}}.$$

(46)

Adopting as a criterion a constant value of a variance for a changing empirical probability, it was shown, in the aforementioned paper, that the least useful measure of goodness of fit of the model is MSE (equation (43)), a slightly better measure is MAE (equation (44)), still better is WMSE (equation (45)), however the best and the mostly recommended one is the weighted mean absolute error (equation (46)).

4. COMPUTATIONAL EXAMPLES

A computational example was conducted based on the data taken from Household Budget Survey in 2012, CSO Warsaw 2013, which refers to the likelihood of possessing the PC by a household. The data presented in table 1 refer to the year 2012. Column 5 includes the number of households n_i' equipped with the PC.

Table 1.

Households equipped with PCs

Lp.	Number of residents in thousand	Surveyed residents in thousand x_{1i}	Available income per person x_{2i}	Households surveyed n_i	Empirical probability p_i	Number of households possessing PC n_i'
	1	2	3	3	4	5
1	less than 20	10	1199.58	4296	0.652	2801
2	20–99	60	1272.82	6447	0.676	4358
3	100–199	150	1320.44	2719	0.707	1922
4	200–499	350	1497.20	3455	0.722	2495
5	500 and more	870	2011.66	4768	0.769	3667
6	rural	0.4	1027.63	15742	0.642	10106

Source: Household Budget Survey in 2012, GUS, Warsaw.

Column 1 of table 1 contains the number of residents of Polish towns in which the people, included in the survey and given in column 3, live. Column 2 (rows 1–4) contains the values which correspond to the center of the interval. In the case of towns of the population 500,000 and more, the mean was calculated for five Polish towns fulfilling this condition. Row 6 represents rural residents. Starting from the number of rural residents and the number of villages, the average number of rural residents was estimated at 360 persons, which was rounded off to 0.4 thousand. The household possessing the PC was chosen as the first model, where an explanatory variable x_{1i} was the number of residents (column 2). Table 2 contains the results of calculations in the form of errors of the following models: logit, probit, raybit.

As far as labeling is concerned, the model errors MAE0, MAE1, MAE2 demonstrate the results of calculations obtained using equation (44) for estimating the probability $p0$ – equations (27), (28) and (29). However MAEML represents the results of the Maximum Likelihood Method for the error given by equation (44).

Taking into account the data included in table 2, the values of errors for particular methods obtained for a given equation were compared. For instance, for the data included in column 1 and rows 1, 2, 3 it can be noticed that in the case of equation

(44) (MAE) and the results obtained for the estimation of the probability p_0, the raybit model yields the smallest error. By conducting further comparisons, it was observed that the raybit method yielded the smallest errors in 13 cases. In the remaining three cases, the logit method yielded the smallest errors.

Table 2.

Errors of the models: logit, probit and raybit for explanatory variable x_{1i}

		MAE0	MAE1	MAE2	MAEML	MSE0	MSE1	MSE2	MSEML
		1	2	3	4	5	6	7	8
1	Logit	0.01347	0.012963	0.01346	0.01347	0.0002569	0.0002120	0.0002532	0.0002550
2	Probit	0.01371	0.01320	0.01370	0.01371	0.0002640	0.0002178	0.0002619	0.0002634
3	Raybit	0.01346	0.012960	0.01345	0.01346	0.0002571	0.0002119	0.0002528	0.0002548
4		WMAE0	WMAE1	WMAE2	WMAEML	WMSE0	WMSE1	WMSE2	WMSEML
5	Logit	955.388	1143.383	962.856	959.318	31.592	42.104	31.559	31.570
6	Probit	973.572	1157.956	978.093	975.187	32.589	42.896	32.571	32.581
7	Raybit	953.510	1143.878	961.957	958.031	31.543	42.159	31.507	31.518

Source: own elaboration.

Another model related to the household possessing the PC assumed an explanatory variable x_{2i} as an available income per person (column 3 in table 1). The results of calculations are shown in table 3 with the labeling identical as in table 2.

Table 3.

Errors of the models: logit, probit and raybit for explanatory variable x_{2i}

	MAE0	MAE1	MAE2	MAEML	MSE0	MSE1	MSE2	MSEML
Logit	0.009538	0.010287	0.009561	0.009557	0.0001530	0.0001492	0.0001523	0.0001526
Probit	0.009732	0.010510	0.009720	0.009723	0.0001580	0.0001535	0.0001586	0.0001578
Raybit	0.009545	0.010289	0.009558	0.009563	0.0001530	0.0001492	0.0001523	0.0001525
	WMAE0	WMAE1	WMAE2	WMAEML	WMSE0	WMSE1	WMSE2	WMSEML
Logit	454.606	583.21	456.855	456.444	14.4814	15.332	14.473	14.4754
Probit	473.849	604.317	472.204	467.372	14.956	15.890	14.960	14.939
Raybit	456.119	582.982	456.191	457.701	14.4815	15.338	14.472	14.4749

Source: own elaboration.

On the basis of the data included in table 3 it can be concluded that the smallest errors are obtained through the raybit method – in 9 cases, and the logit method – in 7 cases. Considering the total values of errors of MAE, MSE, WMAE and WMSE presented in tables 2 and 3, it can be noticed that the smallest values of the afore-mentioned errors were obtained for the following probabilities: p_0 – 6 cases, p_1 – 10 cases, p_2 – 6 cases, p_{ML} – 2 cases. It only validates the application of the probability p_1 in the same way as p_0 and p_2. There was no point in examining the model with two explanatory variables x_{1i} and x_{2i}, since they are strongly correlated – the Pearson's correlation coefficient equaling 0.9985.

5. RESULTS OF COMPUTER SIMULATIONS

In order to verify the applicability of particular models (logit, probit, raybit), computer simulations were conducted.

In accordance with equation (47) a random variable S with a Bernoulli distribution was determined (Devroye, 1986) and takes the value:

$$s_k = \begin{cases} 1 & for \quad r_k \leq P \\ 0 & for \quad r_k > P \end{cases}, \tag{47}$$

where $r_k \in [0;1]$ are uniform random variables,
P – theoretical probability,
$k = 1,2,\ldots, M$.

The observed value of a binomially distributed random variable Z is given by (Devroye, 1986):

$$z = \sum_{k=1}^{M} s_k.$$

The generated empirical value of probability was derived from:

$$p = \frac{z}{M}, \tag{48}$$

where M is the number of random variable in Bernoulli process.

The calculations were conducted for M = 50. The interval [0 ; 1] was divided into ten sub-intervals of the length 0.1 each. For each sub-interval of the form $[A_n; A_{n+1}]$, where $A_n = 0.1 \cdot n$; $n = 0,1,2,...,9$ the values of the theoretical probability were determined:

$$P_{in} = A_n + \frac{i}{100}, \text{ where } i = 0,1,...,10. \tag{49}$$

From equation (48) the empirical probability p_i was determined. For the values of the theoretical probability obtained from equation (49), a random number

generator with a binomial distribution was used, which provided the values of the empirical probability p_i. For these values, the logit, probit and raybit methods were applied. For the obtained estimations $\hat{p}_i$ the error measures were calculated (equtions (43),(44),(45),(46)).

During the computer simulations, for each value i and n (equation (49)), $K = 16000$ repetitions were made. The repetitions consisted in restarting the random number generator. The error measures were calculated as a mean from K repetitons.

The results of the computer simulations are presented in table 4. Rows 1 and 7 contain the values of the theoretical probability P (equation (49)).

In rows 2 and 8 next to the names of the models, in brackets, the numbers of cases for which a given model yielded the smallest errors are provided. The total number of cases for particular probability sub-intervals $[A; A+0.1]$ is 16 – four methods (p_0, p_1, p_2, p_{ML}) multiplied by four criteria of an error. The total number of resaults, for 10 probability sub-intervals equals 160. By adding up the figures in brackets the number of cases with the smallest error is obtained: LOG 30 (17.5%), PRO 74 (46.25%), RAY 58 (36.3%).

Table 4.

Errors of models: logit, probit and raybit obtained through computer simulations

1	Probability	$P \in [0; 0.1]$	$P \in [0.1; 0.2]$	$P \in [0.2; 0.3]$	$P \in [0.3; 0.4]$	$P \in [0.4; 0.5]$
	1	2	3	4	5	6
2	Model	LOG (1) PRO (2) RAY (13)	LOG (4) PRO (2) RAY (10)	LOG (4) PRO (3) RAY (9)	LOG (3) PRO (5) RAY (8)	LOG (5) PRO (3) RAY (8)
3	MAE	p_1	p_{ML}	p_{ML}	p_0	p_0
4	MSE	p_{ML}	p_{ML}	p_{ML}	p_{ML}	p_{ML}
5	WMAE	p_1	p_2	p_0	p_2	p_0
6	WMSE	p_0	p_0	p_2	p_0	p_2
7	Probability	$P \in [0.5; 0.6]$	$P \in [0.6; 0.7]$	$P \in [0.7; 0.8]$	$P \in [0.8; 0.9]$	$P \in [0.9; 1.0]$
8	Model	LOG (4) PRO (11) RAY (1)	LOG (1) PRO (11) RAY (4)	LOG (3) PRO (10) RAY (3)	LOG (3) PRO (12) RAY (1)	LOG (0) PRO (15) RAY (1)
9	MAE	p_0	p_0	p_{ML}	p_{ML}	p_{ML}
10	MSE	p_{ML}	p_{ML}	p_{ML}	p_{ML}	p_{ML}
11	WMAE	p_0	p_1	p_2	p_2	p_1
12	WMSE	p_2	p_0	p_0	p_0	p_0

Source: own elaboration.

It shows that in terms of the goodness of fit, the probit model is the best one, the raybit model is worse and the logit model is the worst. Furthermore it should be noticed that the raybit model is substantially better (in fact twice as good) compared with the logit model.

The following rows (from 3 to 6) contain the information about which equation that defines the probability leads to the smallest value of a selected error measure. In the case of MAE, it is as follows: p_{ML} (five times), p_0 (four times) and p_1 (once). In the case of MSE, there is a clear advantage of the probability determined by applying MLE (p_{ML}) – all ten cases.

In the case of WMAE the following was observed: $p_0(3)$, $p_1(3)$ and $p_2(4)$. WMSE takes the smallest value for: $p_0(7)$ and $p_2(3)$.

Table 5 was compiled on the basis of the results included in table 4. The only difference are the intervals, which now take the form $[0 ; A]$, where $A = 0.1, 0.2, 0.3 ... 1$. Rows 2 and 4 in table 5 contain the sum of subsequent columns in rows 2 and 8 in table 4.

Table 5.

Errors of the models: logit, probit and raybit obtained through computer simulations (cont.)

1	Probability	$P \in [0; 0.1]$	$P \in [0; 0.2]$	$P \in [0; 0.3]$	$P \in [0; 0.4]$	$P \in [0; 0.5]$
	1	2	3	4	5	6
2	Model	LOG (1) PRO (2) RAY (13)	LOG (5) PRO (4) RAY (23)	LOG (9) PRO (7) RAY (32)	LOG (12) PRO (12) RAY (40)	LOG (17) PRO (15) RAY (48)
3	Probability	$P \in [0; 0.6]$	$P \in [0; 0.7]$	$P \in [0; 0.8]$	$P \in [0; 0.9]$	$P \in [0; 1.0]$
4	Model	LOG (21) PRO (26) RAY (49)	LOG (22) PRO (37) RAY (53)	LOG (25) PRO (47) RAY (56)	LOG (28) PRO (59) RAY (57)	LOG (28) PRO (74) RAY (58)

Source: own elaboration.

The data shown in table 5 shows the advantage of the raybit model for $P \in [0; A]$ where $A = 0.1, 0.2, ... 0.8$. It is only for $P \in [0; 0.9]$ and $P \in [0; 1.0]$ that the probit model gains the advantage.

The logit model performs worst of all analyzed models for any value from the interval $P \in [0; A]$.

The data shown in table 4 demonstrates a variability in the number of cases when a given method yields the smallest error in relation to the value of the probability. In order to explain this phenomenon, the following numerical experiment was conducted. A random number generator was replaced by the values of the theoretical probability $P_i = 0.01 \cdot (1 + i)$, where $i = 0,1,...,98$, which were used in place of the values of the empirical probability. Applying equations (27)−(29) and (38)−(40) the values of the

probability p_0 and p_2 were determined. The results of the calculations are shown in figures 1–4, where a dashed line represents the raybit model and a solid line – the linear model. Figure 1 proves that the results for the raybit model for $P_i \in [0.01; 0.5]$ are very similar to the results for the linear model, which results in very small values of the error. This is the reason why the raybit model has a clear advantage over other models for this probability interval.

The probability $pPRO_{0,i}$ obtained for the probit model for the same interval shows much larger nonlinearity. However for the interval $P_i \in [0.5; 0.99]$ the probit model fits well with the linear model.
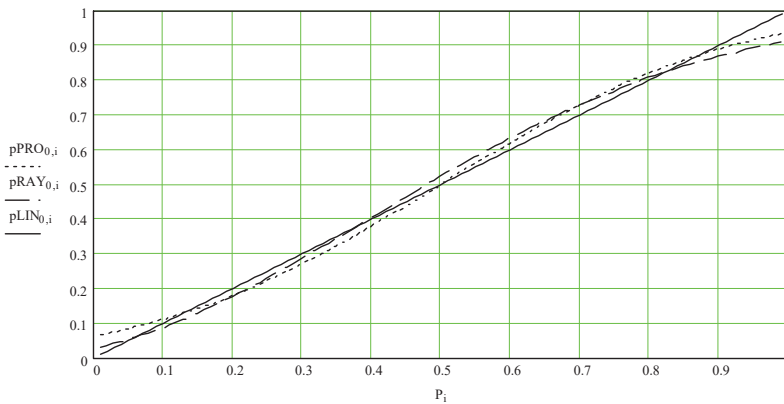

Figure 1. The results of the probability p_0 calculations for $P_i \in [0.01; 0.99]$
Applied labeling: dotted line $pPRO_{0,i}$ (probit model), dashed line $pRAY_{0,i}$ (raybit model),
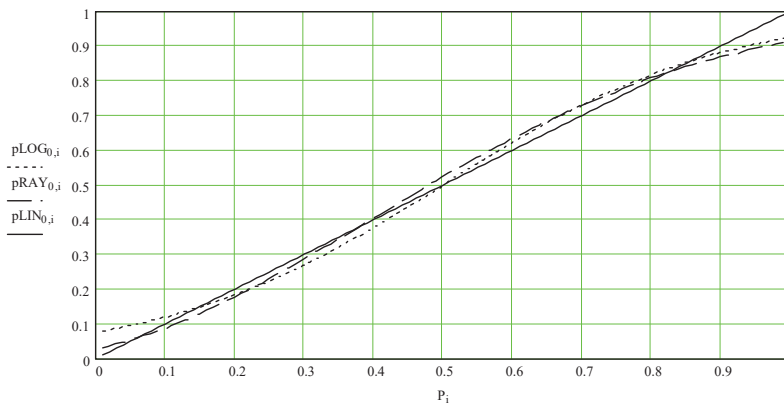solid line $pLIN_{0,i}$ (linear model).
Source: own elaboration.

Figure 2. The results of the probability p_0 calculations for $P_i \in [0.01; 0.99]$
Applied labeling: dotted line $pLOG_{0,i}$ (logit model), dashed line $pRAY_{0,i}$ (raybit model),
solid line $pLIN_{0,i}$ (linear model).
Source: own elaboration.

On the basis of figure 2 it can be noticed that the probability $pLOG_{0,i}$ obtained for the logit model shows much larger nonlinearity (than the raybit model), especially for $P_i \in [0.01; 0.2]$ and $P_i \in [0.6; 0.99]$.
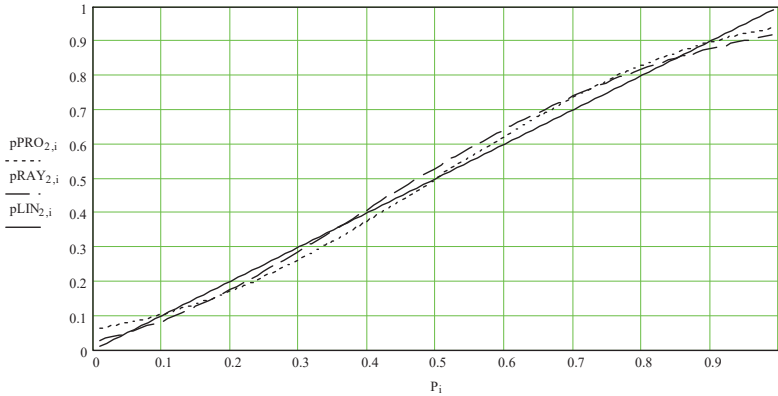

Figure 3. The results of the probability p_2 calculations for $P_i \in [0.01; 0.99]$
Applied labeling: the same as in figure 1.
Source: own elaboration.

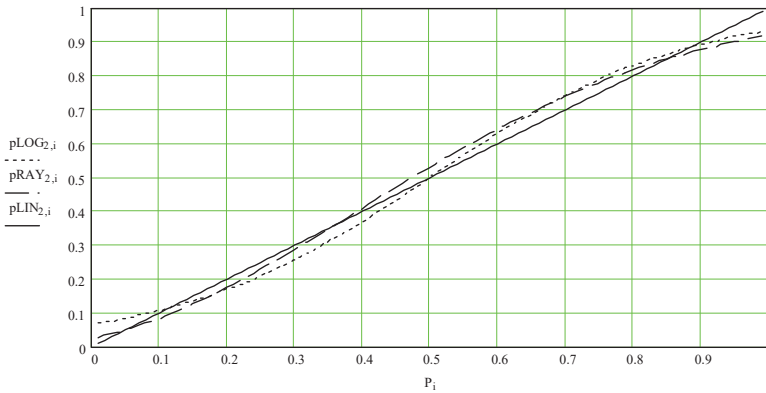The situation described in relation to figure 1 can be also observed in figure 3.

Figure 4. The results of the probability p_2 calculations for $P_i \in [0.01; 0.99]$
Applied labeling: the same as in figure 2.
Source: own elaboration.

The results observed in figure 2 can be also observed in figure 4.

6. CONCLUSION

In the paper the estimation of the parameters of qualitative econometric models was discussed including: the logit model, the probit model and the raybit model. The following methods of estimation were considered. The generalized least squares (equation (23)), where the elements of a diagonal covariance matrix are determined on the basis of the empirical probability. The method leads to the estimation of a theoretical probability labelled as p_0 (equations (27)−(29)). The next method is two-step. As a first step, using OLS, the estimation of the probability p_1 was determined (equations (31)−(33)). As a second step, GLS was used, where the elements of a diagonal covariance matrix were determined on the basis of the probability p_1. Consequently, the estimation of the probability labelled as p_2 was obtained (equations (38)−(40)). Although the probability p_1 was used to calculate the probability p_2, it was also treated as yet another value of the theoretical probability estimation. The last method of estimation of a qualitative econometric model was the maximum likelihood method, where the probability estimation was labelled as p_{ML}.

With reference to the computational examples (tables 2 and 4), the raybit model, proposed in this paper, proved to be the best out of the three models under study. In computer simulations this model showed clear advantage for probability $P \in [0; 0.8]$ (table 5). Only for $P \in [0; 0.9]$ and $P \in [0; 1.0]$ the probit model performs best. Despite the fact that for the above mentioned probability intervals the raybit model is worse than the probit model, it still has its advantages, namely, the analytical forms of the cumulative distribution function as well as the inverse function to the cumulative distribution function.

It should be noticed that in the whole probability interval the raybit model yields a smaller error than the logit model.

It means that while analyzing a binomial qualitative variable, along with the classic logit and probit models, it is worth taking into account the results of the raybit model.

REFERENCES

Amemiya T., (1981), Qualitative Response Models: A Survey, *Journal of Economic Literature*, December, 1483–1536.

Budżety gospodarstw domowych w 2012 r., (2013), Household Budget Survey in 2012, GUS, Warsaw.

Celik A. N., (2003), A Statistical Analysis of Wind Power Density Based on the Weibull and Rayleigh Models at the Southern Region of Turkey, *Elsevier, Renewable Energy,* 29, 593–604.

Chow G. C., (1983), *Econometrics*, Mc-Graw-Hill Inc, New York.

Devroye L., (1986), *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.

Finney D. J., (1983), *Probit Analysis*, Cambridge University Press.

Gruszczyński M., (2012), *Empiryczne finanse przedsiębiorstw. Mikroekonometria finansowa,* Difin, Warszawa.

Guzik B., Appenzeller D., Jurek W., (2005), *Prognozowanie i symulacje. Wybrane zagadnienia*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, nr 168, Poznań.

Jajuga K., (1989), Modele z dyskretną zmienną objaśnianą, in: Bartosiewicz S. (ed.), *Estymacja modeli ekonometrycznych*, PWE, Warszawa.

Judge G. G., Griffiths W. E., Hill R. C., Lee T. C., (1980), *Theory and Practice of Econometrics*, Wiley, New York.

Koutsoyiannis D., (2003), On the Appropriateness of the Gumbel Distribution in Modelling Extreme Rainfall, *Proceedings of the ESF LESC Exploratory Workshop held at Bologna*, Italy, October 24–25, 303–319.

Maddala G. S., (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

Maddala G. S., (1992), *Introduction to Econometrics*, Macmillan Publishing Company, New York.

McFadden D., (1984), Econometric Analysis of Qualitative Response Models, in: Gribliches Z., Intriligator M. D., (eds.), *Handbook of Econometrics*, Vol. II, Elsevier Science Publishers BV.

Nerlove M., Press S. J., (1973), *Univariante and Multivariante Log-Linear and Logistic Models*, R-1406-EDA/NIH, Dec., RAND Santa Monica, CA, 90406.

Polakow D. A., Dunne T. T., (1999), Modelling Fire-Return Interval *T*: Stochasticity and Censoringin the Two-Parameter Weibull Model, *Elsevier, Ecological Modelling*, 121, 79–102.

Purczyński J., Porada-Rochoń M., (2015), Ocena jakości modeli ze zmienną dychotomiczną, *Logistyka*, 3, 4064-4073.

Rinne H., (2009), *The Weibull Distribution. A Handbook*, CRC PressTaylor & Francis Group, LLC.

Train K., (2009), *Discrete Choice Methods with Simulation*, Cambridge University Press.

MODEL RAYBITOWY I OCENA JEGO JAKOŚCI W PORÓWNANIU Z MODELEM LOGITOWYM I PROBITOWYM

Streszczenie

W pracy zaproponowano nowy model dla zmiennej objaśnianej zero-jedynkowej (binarnej, dychoto-micznej). Nazwa modelu raybit wynika stąd, że prawdopodobieństwo odpowiada dystrybuancie rozkładu Rayleigha. Ocenę jakości modeli przeprowadzono z wykorzystaniem 4 definicji błędu: MSE, MAE, WMSE, WMAE. Rozpatrzono dwa przykłady obliczeniowe, które wykazały, że model raybitowy prowa-dzi do mniejszych wartości błędu, niż model logitowy i probitowy. Wykonano symulacje komputerowe z wykorzystaniem generatora liczb losowych o rozkładzie dwumianowym. Przeprowadzone symulacje wykazały, że dla wartości prawdopodobieństwa teoretycznego z przedziału $P_i \in [0; 0,8]$ model raybitowy przewyższa pozostałe dwa modela prowadząc do mniejszej wartości błędu.

Słowa kluczowe: jakościowe modele ekonometryczne, model logitowy, model probitowy

THE RAYBIT MODEL AND THE ASSESSMENT OF ITS QUALITY IN COMPARISON WITH THE LOGIT AND PROBIT MODELS

Abstract

A new model for a dependent variable taking the value 0 or 1 (binary, dichotomous) was proposed. The name of the proposed model – the raybit model – stems from the fact that the probability corresponds to the Rayleigh cumulative distribution function. The assessment of the quality of selected models was conducted with the use of four definitions of error: MSE, MAE, WMSE, WMAE. Two computational examples were considered, which proved that the raybit model yields smaller values of error than the logit and probit models. Computer simulations were conducted using a random number generator with a binomial distribution. They proved that for the values of the theoretical probabilityfor the interval $P_i \in [0; 0.8]$ the raybit model outperforms the other two models yielding a smaller value of error.

Keywords: qualitative econometric models, logit model, probit model

ALICJA OLEJNIK[1], JAKUB OLEJNIK[2]

AN ALTERNATIVE TO PARTIAL REGRESSION IN MAXIMUM LIKELIHOOD ESTIMATION OF SPATIAL AUTOREGRESSIVE PANEL DATA MODEL[3] [4]

1. INTRODUCTION

Partial regression, developed by Frisch, Waugh (1933), is a popular method of elimination of nuisance slope parameters. It is widely used in *inter alia* panel data analysis. One of its special cases is commonly used to estimate regression parameters in, so called, fixed effects models. Partial regression allows one to find slope parameters without the need of estimating actual levels of fixed effects. In this form it is referred to as the *demeaning* procedure (cf. Baltagi, 2005).

As the Maximum Likelihood (ML) estimation procedure is one of the most popular estimation methods for Spatial Autoregressive Model (SAR) many researchers have also used the technique of demeaning in ML estimation of the SAR model. However, validity of this approach has been occasionally subjected to doubt (e.g. Anselin et al., 2006) on the grounds that the demeaning procedure yields singular variance of the error term. As Pace (2014) rightly points out, maximising demeaned likelihood can still produce consistent estimates[5] of regression parameters, as the demeaned likelihood can be interpreted as a concentrated likelihood. However, estimates of their variances are likely to be invalid.

A procedure to overcome this problem was first proposed by Lee, Yu (2010) for a reasonably general spatial fixed time/individual fixed effect model. They noticed that applying certain transformation of data, prior to conducting ML procedure, can effectively eliminate fixed effects and, at the same time, properly account for the singularity. In this paper we generalise this approach and show that, contrary to a statement included in Lee, Yu (2010), ML estimation with demeaning of SAR model is feasible in a larger class of settings than originally described.

Our invariant subspace framework allows, under some assumptions, to effectively deal with large class of fixed effects designs in panel and non-panel data models. Designs handled by the framework range from group-specific fixed effects with non-uniform cardinality to multiple levels of group-specific effects with possibly overlapping groups and non-constant (yet known) effect sizes within those groups. This can be done under the assumption that the Krylov subspace[6] for spatially lagged fixed effects is of incomplete dimension. The crucial requirement expresses certain degree of compatibility of the fixed effects design with assumed spatial weight matrix. In the original paper of Lee, Yu (2010) the considered model specification also includes spatially correlated error term, however in our paper, for simplicity of presentation, we employ only autoregressive scheme. Therefore, the aim of our paper is to develop extension of the fixed effect eliminating transformation of Lee, Yu (2010), so that effectively a larger class of fixed effect designs can be handled.

Unless specified differently, throughout the paper we use the short term SAR model to actually describe the panel-data SAR model. All statements applicable to non-panel data SAR model are also valid in the panel case. Whenever n is used to denote sample size, it can be read $n = NT$, moreover $\mathbb{R}^n = \mathbb{R}^N \otimes \mathbb{R}^T$, $I_n = I_N \otimes I_T$ etc. This notation can also cover the case of either spatial unit unbalanced or time unbalanced panel data set, that is if $N = n_1 + \cdots + n_T$ and $\mathbb{R}^n = \mathbb{R}^{n_1} \oplus ... \oplus \mathbb{R}^{n_T}$, or if $T = t_1 + \cdots + t_N$, etc., respectively.

The rest of this paper is structured as follows. Section 2 introduces the concept of partial regression by Frisch, Waugh (1933). Section 3 introduces Spatial Autoregressive Model specification and describes the well-known naive approach of demeaning in ML estimation. Section 4 presents our original approach. Section 5 formulates statements on asymptotic behaviour of our estimator. Finally, section 6 presents a summary and conclusions.

2. PARTIAL REGRESSION

Although for our purposes it is enough to consider the basic form of the Frisch-Waugh (F-W) theorem, it is worthwhile to mention some of its interesting extensions. In particular, Fiebig, Bartels (1996) develop an extension of the F-W procedure that is able to handle model specifications with non-spherical disturbances, that is where the variance covariance matrix of the error term is not proportional to identity matrix.

Another interesting extension to partial regression has been recently developed in Yamada (2016). It has been shown that the F-W theorem is invariant under certain modification of the least squares objective function. Namely, if instead of the usual least squares optimization problem

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^k} \|Y - X\beta\|^2$$

---

[6]   To be defined in section 4, can be found also in e.g. Liesen, Strakoš (2013).

we consider the LASSO (least absolute shrinkage and selection operator) regression. This can be described as solution to the modified problem $\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^k} \|Y - X\beta\|^2 + \lambda\|\beta\|_1$, where λ is a tuning parameter and $\|\beta\|_1 = \sum_{i<k}|\beta_i|$. Similarly, the F-W theorem still holds if the usual least squares is replaced with ridge regression i.e. $\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^k} \|Y - X\beta\|^2 + \lambda\|\beta\|^2$.

Those results suggest that partial regression might be a technique applicable in a variety of estimation schemes. The maximum likelihood estimation procedure is one of them. This is implied by fact of equivalence of the estimates form OLS and ML approaches under normality of error term. However, the question of applicability of partial regression becomes far more difficult if one considers the spatially autoregressive term in model specification. In our paper we show that, under some assumptions, the Maximum Likelihood estimation procedure in case of a spatially autoregressive DGP can also benefit from virtues of F-W theorem.

In the reminder of this section we present the concept of partial regression developed by Frisch, Waugh (1933). Let us consider a standard linear model

$$Y = X_1\,\beta_1 + X_2\,\beta_2 + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n),$$

where Y is a $n \times 1$ vector of observations, X_1 and X_2 are respectively $k_1 \times n$ and $k_2 \times n$ design matrices, $\theta = (\beta_1, \sigma^2)$ is the unknown parameter of interest and $\tau = \beta_2$ is a nuisance parameter. The partial regression technique allows us to find θ without actually estimating τ (c.f. Greene, 2008), Section 3.3). Let us denote $M_{X_2} = I_n - X_2(X_2^T X_2)^{-1} X_2^T$. The partial regression estimator $\hat{\theta} = (\hat{\beta}, \widehat{\sigma^2})$ is given by

$$\hat{\beta} = \left(X_1^T M_{X_2} X_1\right)^{-1} M_{X_2} X_1^T Y \tag{1}$$

and asymptotically unbiased[7]

$$\widehat{\sigma^2} = n^{-1}\left(M_{X_2}Y - M_{X_2}X_1\hat{\beta}\right)^T\left(M_{X_2}Y - M_{X_2}X_1\hat{\beta}\right). \tag{2}$$

It turns out that $\hat{\beta}$ coincides with the corresponding element of slope estimator in the full Ordinary Least Squares scheme, i.e. $\hat{\beta} = \widehat{\beta_1}^{OLS}$ with

$$\left[\widehat{\beta_1}^{OLS}, \widehat{\beta_2}^{OLS}\right]^T = ([X_1\ X_2]^T[X_1\ X_2])^{-1}[X_1\ X_2]^T Y.$$

Moreover, the Frisch-Waugh theorem also states that

$$\widehat{\beta_2} = (X_2^T X_2)^{-1} X_2^T\left(Y - X_1\hat{\beta}\right) = \widehat{\beta_2}^{OLS} \tag{3}$$

---

[7] Provided that $k_1 + k_2 = O(1)$.

and the variance of $\hat{\beta}$ can be obtained through the partitioned inverse[8] of the design moment matrix $[X_1 \, X_2]^T[X_1 \, X_2]$, which is $\left(X_1^T M_{X_2} X_1\right)^{-1}$. In the context of panel data model, by substituting a time or individual effect dummy variable for X_2 we obtain the well-known demeaning procedure.

3. DEMEANING IN ML ESTIMATION OF SPATIAL AUTOREGRESSIVE MODEL

In this section we introduce the Spatial Autoregressive Model specification and describe the well-known naive approach of demeaning in ML estimation, as used in e.g. Elhorst and Fréret (2008). Let us consider a standard spatial autoregressive linear model

$$Y = \rho \mathbf{W} Y + X_1 \beta_1 + X_2 \beta_2 + \varepsilon, \; \varepsilon \sim \mathrm{N}(0, \sigma^2 \mathrm{I}_n), \tag{4}$$

where $\mathbf{W}$ is an arbitrary spatial weight $n \times n$ matrix (with zero diagonal) and ρ is the scalar autoregressive parameter. Moreover, as previously, Y is a $n \times 1$ vector of observations, X_1 and X_2 are $k_1 \times n$, $k_2 \times n$ respectively design matrices, $\theta = (\rho, \beta_1, \sigma^2)$ is the unknown parameter of interest and $\tau = \beta_2$ is the nuisance parameter. The $\rho \mathbf{W} Y$ term is referred to as the spatial autoregressive term. The elements $\left(w_{ij}\right)_{ij \leq n}$ of $\mathbf{W}$ have the common interpretation of spatial weights, i.e. a measure of influence of j-th unit on unit i. Since $\rho \mathbf{W} Y = \rho \left(\sum_{j=1}^n w_{ij} Y_j\right)_{i \leq n}$, the spatial autoregressive term conveys information on weighted averages of influences from other spatial[9] units ($w_{ii} = 0$, $i \leq n$) on a given unit.

It is a well-known fact that the specification (4) cannot be estimated with the use of classical Ordinary Least Squares (see Anselin, 1988). Instead, the ML estimation procedure is a commonly suggested feasible alternative. To implement the maximum likelihood estimation procedure for the SAR specification it is enough to notice that, using the form of Gaussian density of ε, we can obtain the following formula for log likelihood function

$$\log L(Y, \theta, \tau) =$$
$$= -\frac{n}{2} \log(2\pi\sigma^2) + \log \det(\mathrm{I}_n - \rho \mathbf{W}) - \frac{1}{2\sigma^2} \|Y - \rho \mathbf{W} Y - X_1 \beta_1 - X_2 \beta_2\|^2, \tag{5}$$

with the assumption that $\det(\mathrm{I}_n - \rho \mathbf{W})$ is positive for all ρ in its parameter space[10]. A straightforward implementation of the idea of partial regression consists in applying

---

[8] I.e. the relevant part of the inverse.

[9] Or spatio-temporal in dynamic panel case.

[10] It is a common practice to assume that the parameter space for spatial autoregressive parameter ρ is an interval $\mathcal{J} \subset \mathbb{R}$ such that $0 \in \mathcal{J}$. The endpoints of $\mathcal{J}$ are established from a condition ensuring invertibility of the spatial lag $\mathrm{I}_n - \rho \mathbf{W}$, for example $\|\mathbf{W}\| < 1$, for a matrix (submultiplicative) norm or

the demeaning operator M_{X_2} to the formula under the norm in (5). This, widely used, approach is supported by the fact that first order differential optimality condition on (5) is consistent with (3), i.e.

$$\tau_{\max} = \tau_{\max}(\rho, \beta, \sigma^2) = (X_2^{\mathrm{T}} X_2)^{-1} X_2^{\mathrm{T}} (Y - \rho \mathbf{W} Y - X_1 \beta)$$

and, as a result of simple algebra, we get the concentrated log likelihood

$$\log L(Y, \theta, \tau_{\max}) =$$

$$= -\frac{n}{2} \log(2\pi\sigma^2) + \log \det(\mathrm{I_n} - \rho \mathbf{W}) - \frac{1}{2\sigma^2} \left\| M_{X_2} Y - \rho M_{X_2} \mathbf{W} Y - M_{X_2} X_1 \beta_1 \right\|^2. \tag{6}$$

Unfortunately, the operator M_{X_2} is not unitary, let alone invertible thus the formula above cannot be interpreted as a regular likelihood function. Nonetheless, the estimation approach of maximising (6) with respect to θ yields reasonably good estimates, provided that $k_2 = O(1)$, c.f. Elhorst (2009). However, the estimate of the asymptotic variance of the resulting ML estimator may be invalid. Moreover, the estimates of θ, when $\limsup_{n\to\infty} k_1/n > 0$, are not consistent.

4. PARTIAL REGRESSION IN ML ESTIMATION OF SAR MODEL

In this section we derive our original approach to the problem of eliminating fixed effects in case of ML estimation of the SAR model by employing an alternative to the idea of partial regression. We will consider two cases. In Case I we assume that the spatial weight matrix is in some sense consistent with the nuisance slope parameter design X_2. In Case II an assumption about dimension of certain invariant subspace is made instead.

One approach to the issue of singularity of the M_{X_2} operator can be to pre-multiply ε by an orthogonal (i.e. transformation of coordinates) matrix $\boldsymbol{E}$ which maps the range of M_{X_2} onto $\mathbb{R}^{n-k_2}$ interpreted as a natural subset of $\mathbb{R}^n = \boldsymbol{E}(\mathbb{R}^n) = \mathbb{R}^{n-k_2} \oplus \mathbb{R}^{k_2}$, where $\oplus$ is the coordinate-wise direct sum of linear spaces. Then, we could integrate over unnecessary degrees of freedom, at the same time eliminating τ from the likelihood function (5). Effectively, this is the same as using the transformation $\pi \boldsymbol{E} M_{X_2}$, with $\pi = \pi_{n-k_2}$ being the natural projection $\mathbb{R}^n \xrightarrow{\pi} \mathbb{R}^{n-k_2}$ preserving $n - k_2$ first coordinates. Indeed, it can be observed that $(\pi \boldsymbol{E})^{\mathrm{T}} \pi \boldsymbol{E} = M_{X_2}$, thus $\pi \boldsymbol{E} M_{X_2} = \pi \boldsymbol{E}$.

Obviously, the transformations $\boldsymbol{E}$ and $\pi \boldsymbol{E}$ are not uniquely defined. In fact any such transformation $\boldsymbol{E}$, as described in previous paragraph can be equally useful. Here, we propose a method of construction of a possible candidate. Let $c_1, \dots, c_n$ be rows of the matrix M_{X_2}. Let $i_1 = 1$. For $j = 2, \dots, n - k_2$, once i_k, for $k < j$, have been defined,

---

max$(\mathcal{I} \cup -\mathcal{I}) < 1/\lambda_{\mathfrak{R}}$, where $\lambda_{\mathfrak{R}}$ is a modulus-maximal real eigenvalue of $\mathbf{W}$ (c.f. Appendix in Olejnik, Özyurt, 2016).

we can set $i_j = \min\left\{k > i_{j-1} : M_{[c_1 \ldots c_{j-1}]}c_k \neq 0\right\}$. Having chosen vectors $c_{i_1}, \ldots, c_{i_{n-k_2}}$ as a basis of the range of M_{X_2}, we can proceed with Gram-Schmidt ortho-normalization process and thus obtain an orthonormal system of vectors $\tilde{c}_i$, for $1 \leq i \leq n - k_2$. We apply the same procedure for matrix $I - M_{X_2}$ and obtain orthonormal system $\tilde{d}_i$, for $1 \leq i \leq k_2$. Finally, it is enough to set $\boldsymbol{E} = \left[\tilde{c}_1 \ldots \tilde{c}_{n-k_2} \tilde{d}_1 \ldots \tilde{d}_{k_2}\right]^{\mathrm{T}}$.

Case I

Let us assume that $\pi \boldsymbol{E} \mathbf{W} = \pi \boldsymbol{E} \mathbf{W} M_{X_2}$. Since $\varepsilon \sim \mathrm{N}(0, \sigma^2 \mathrm{I}_n)$, we can immediately conclude that $\boldsymbol{E}\varepsilon \sim \mathrm{N}(0, \sigma^2 \mathrm{I}_n)$, so that $\pi \boldsymbol{E}\varepsilon \sim \mathrm{N}\left(0, \sigma^2 \mathrm{I}_{n-k_2}\right)$. Notice that, by denoting $\mathbf{W}_{\pi \boldsymbol{E}} = \pi \boldsymbol{E} \mathbf{W} \boldsymbol{E}^{\mathrm{T}} \pi^{\mathrm{T}}$ and observing that $\pi \pi^{\mathrm{T}} = \mathrm{I}_{n-k_2}$, we have

$$\pi \boldsymbol{E}\varepsilon = \pi \boldsymbol{E}Y - \rho \pi \boldsymbol{E} \mathbf{W} Y - \pi \boldsymbol{E}X_1\beta_1 - \pi \boldsymbol{E}X_2\beta_2 =$$

$$= \pi \boldsymbol{E}Y - \rho \pi \boldsymbol{E} \mathbf{W} M_{X_2} Y - \boldsymbol{\pi \boldsymbol{E}X_1\beta_1} = \left(\mathrm{I}_{n-k_2} - \rho\, \mathbf{W}_{\pi \boldsymbol{E}}\right)\pi \boldsymbol{E}Y + \boldsymbol{\pi \boldsymbol{E}X_1\beta_1}.$$

It can be noticed that $\left(\mathrm{I}_{n-k_2} - \rho \mathbf{W}_{\pi \boldsymbol{E}}\right) = \pi \boldsymbol{E}(\mathrm{I}_n - \rho \mathbf{W})\boldsymbol{E}^{\mathrm{T}} \pi^{\mathrm{T}}$ is invertible if $(\mathrm{I}_n - \rho \mathbf{W})$ is invertible. Indeed, it is enough either to observe that, by a simple algebra, we have $\left(\mathrm{I}_{n-k_2} - \rho \cdot \mathbf{W}_{\pi \boldsymbol{E}}\right)\pi \boldsymbol{E}(\mathrm{I}_n - \rho \cdot \mathbf{W})^{-1}\boldsymbol{E}^{\mathrm{T}} \pi^{\mathrm{T}} = \mathrm{I}_{n-k_2}$, since $\boldsymbol{E}\mathbf{W}(\mathrm{I}_n - (\pi \boldsymbol{E})^{\mathrm{T}} \pi \boldsymbol{E}) = \mathbf{0}$. Thus, for each value of ρ in its parameter space we can properly define the transformation

$$T(\epsilon) = \left(\mathrm{I}_{n-k_2} - \rho \mathbf{W}_{\pi \boldsymbol{E}}\right)^{-1}\epsilon - \pi \boldsymbol{E}X_1\beta_1, \text{ for } \epsilon \in \mathbb{R}^{n-k_2}.$$

Since $\pi \boldsymbol{E}Y = T(\varepsilon)$ and $\frac{\partial}{\partial \epsilon}T^{-1} = \mathrm{I}_{n-k_2} - \rho \mathbf{W}_{\pi \boldsymbol{E}}$ we obtain the following form of logarithm of likelihood function for θ, based on observable values of $\pi \boldsymbol{E}Y$

$$\log L(\pi \boldsymbol{E}Y, \theta) = -\frac{n - k_2}{2}\log(2\pi\sigma^2) +$$

$$+ \log \det\left(\mathrm{I}_{n-k_2} - \rho \mathbf{W}_{\pi \boldsymbol{E}}\right) - \frac{1}{2\sigma^2}\|\pi \boldsymbol{E}Y - \rho \mathbf{W}_{\pi \boldsymbol{E}}\pi \boldsymbol{E}Y - \pi \boldsymbol{E}X_1\beta_1\|^2. \tag{7}$$

Now, we can differentiate $\log L(\pi \boldsymbol{E}Y, \theta)$ with respect to (β, σ^2) and equate the result to zero, thus get the optimal relations between β, σ^2 and ρ

$$\beta_1 = (X_1 E^T \pi^T \pi E X_1)^{-1} X_1 E^T \pi^T (\pi E Y - \rho \mathbf{W}_{\pi \boldsymbol{E}}\pi E Y), \tag{8}$$

$$\sigma^2 = \frac{1}{n-k_2}\|\pi E Y - \rho \mathbf{W}_{\pi \boldsymbol{E}}\pi E Y - \pi E X_1\beta_1\|^2. \tag{9}$$

Relations (8) and (9) are SAR counterparts of partial regression estimators (1), (2). In order to obtain the ML estimates β_1 and σ^2 we need to evaluate the above formulas a the maximum likelihood estimate of ρ.

Substituting the above equations, (8) and (9), into (7) we get the concentrated log likelihood function $\log L_{\text{Conc}}$

$$\log L_{\text{Conc}}(\pi EY, \rho) = -\tfrac{n-k_2}{2}\log(a\rho^2 + b\rho + c) + \log \det\left(I_{n-k_2} - \rho W_{\pi E}\right) + \text{Const},$$

where the coefficients of the quadratic polynomial in ρ are given by

$$a = \|e(\pi EY, \pi EX_1)\|^2,$$

$$b = 2e(\pi EY, \pi EX_1)^{\mathrm{T}} e(W_{\pi E}\pi EY, \pi EX_1),$$

$$c = \|e(W_{\pi E}\pi EY, \pi EX_1)\|^2,$$

and $e(V_1, V_2)$ is the column vector of OLS residuals obtained by regressing V_1 on V_2. Lastly, it is clear that simply maximising (presumably numerically) $\log L_{\text{Conc}}(\pi EY, \rho)$ with respect to its single parameter ρ gives the desired value of $\rho_{\max}$, which can be further substituted into (8) and (9).

Case II

Now, let us assume that $\pi EW - \pi EWM_{X_2} \neq 0$. The approach we present further is based on the concept of Krylov subspace. The Krylov subspace for X_2 with respect to W is the minimal W-invariant subspace containing X_2. We will denote it by H, i.e.

$$H = \text{span}(W^k X_2 \beta : k = 0,1,\dots,n \wedge \beta \in \mathbb{R}^{k_2}).$$

We will assume that H is a proper subspace of $\mathbb{R}^n$, so that $n_* = n - k_* > 0$, with $k_* = \dim H$.

With the notation of M_H being orthogonal projection on orthogonal complement of H, as previously, we define linear isometry F to be an operator that takes range of M_H onto $\mathbb{R}^{n_*} \simeq F(H^\perp)$ with coordinate-wise $\mathbb{R}^n = F(H^\perp) \oplus \mathbb{R}^{k_*}$. Furthermore, let $\pi_*: \mathbb{R}^n \to \mathbb{R}^{n_*}$ be orthogonal projection preserving first n_* coordinates. We have $\pi_* F = \pi_* F M_H$ and $\pi_* FW(I - M_H) = \pi_* F(I - M_H) = 0$.

Again, considering the fact that $\varepsilon \sim N(0, \sigma^2 I_n)$, we can immediately conclude that $F\varepsilon \sim N(0, \sigma^2 I_n)$, and further that $\pi_* F\varepsilon \sim N(0, \sigma^2 I_{n_*})$. Notice that, denoting $W_{\pi_* F} = \pi_* FWF^{\mathrm{T}}\pi_*^{\mathrm{T}}$ and observing that $\pi_* \pi_*^{\mathrm{T}} = I_{n_*}$ we have

$$\pi_* \boldsymbol{F} \varepsilon = \pi_* \boldsymbol{F} Y - \rho \pi_* \boldsymbol{F} \mathbf{W} Y - \pi_* \boldsymbol{F} X_1 \beta_1 - \pi_* \boldsymbol{F} X_2 \beta_2 =$$

$$= \pi_* \boldsymbol{F} Y - \rho \pi_* \boldsymbol{F} \mathbf{W} (M_H Y + (I - M_H) Y) - \pi_* \boldsymbol{F} X_1 \beta_1 =$$

$$= \left(I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \right) \pi_* E Y - \pi_* \boldsymbol{F} \mathbf{W} (I - M_H) Y + \pi_* \boldsymbol{F} X_1 \beta_1 =$$

$$= \left(I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \right) \pi_* E Y + \pi_* \boldsymbol{F} X_1 \beta_1.$$

It can be also noticed that $\left(I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \right) = \pi_* \boldsymbol{F} (I_n - \rho \mathbf{W}) \boldsymbol{F}^{\mathrm{T}} \pi_*^{\mathrm{T}}$ is invertible if $(I_n - \rho \mathbf{W})$ is invertible, since $\left(I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \right) \pi_* \boldsymbol{F} (I_n - \rho \mathbf{W})^{-1} \boldsymbol{F}^{\mathrm{T}} \pi_*^{\mathrm{T}} = I_{n_*}$, since $\boldsymbol{F} \mathbf{W} (I_n - M_H) = \mathbf{0}$.

As a result, for each value of ρ in its parameter space we can properly define the transformation

$$T(\epsilon) = \left(I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \right)^{-1} \epsilon - \pi_* \boldsymbol{F} X_1 \beta_1, \text{ for } \epsilon \in \mathbb{R}^{n_*}.$$

Since $\pi_* \boldsymbol{F} Y = T(\varepsilon)$ and $\frac{\partial}{\partial \epsilon} T^{-1} = I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}}$ we obtain the following form of logarithm of likelihood function for θ based on observable values of $\pi_* \boldsymbol{F} Y$

$$\log L(\pi E Y, \theta) =$$

(10)

$$= -\frac{n_*}{2} \log(2\pi\sigma^2) + \log \det\left(I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \right) - \frac{1}{2\sigma^2} \left\| \pi_* \boldsymbol{F} Y - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \pi_* \boldsymbol{F} Y - \pi_* \boldsymbol{F} X_1 \beta_1 \right\|^2.$$

Now, we can differentiate with respect to (β, σ^2) and equate to zero to get optimal relations between β, σ^2 and ρ

$$\beta_1 = (X_1 \boldsymbol{F}^T \pi_*^T \pi_* \boldsymbol{F} X_1)^{-1} X_1 \boldsymbol{F}^T \pi_*^T \left(\pi_* \boldsymbol{F} Y - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \pi_* \boldsymbol{F} Y \right), \tag{11}$$

$$\sigma^2 = \frac{1}{n_*} \left\| \pi_* \boldsymbol{F} Y - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \pi_* \boldsymbol{F} Y - \pi_* \boldsymbol{F} X_1 \beta_1 \right\|^2, \tag{12}$$

which are SAR counterparts of partial regression estimators (1), (2). In order to be able to evaluate the above formulas we need to obtain maximum likelihood estimate of ρ.

Substituting the above equations, (11) and (12), to (10) we get the concentrated likelihood function L_{Conc}

$$\log L_{\text{Conc}}(\pi E Y, \rho) = -\frac{n_*}{2} \log(\boldsymbol{a} \rho^2 + \boldsymbol{b} \rho + \boldsymbol{c}) + \log \det\left(I_{n_*} - \rho \mathbf{W}_{\pi_* \boldsymbol{F}} \right) + \text{Const},$$

where the coefficients of quadratic polynomial in ρ are given by

$$a = \|\mathbb{e}(\pi_* FY, \pi_* FX_1)\|^2,$$

$$b = 2\mathbb{e}(\pi_* FY, \pi_* FX_1)^{\mathrm{T}} \mathbb{e}\big(\mathbf{W}_{\pi_* F}\pi_* FY, \pi_* FX_1\big),$$

$$c = \big\|\mathbb{e}\big(\mathbf{W}_{\pi_* F}\pi_* FY, \pi_* FX_1\big)\big\|^2$$

and $\mathbb{e}(V_1, V_2)$ is the column vector of OLS residuals obtained by regressing V_1 on V_2. By maximising $\log L_{\mathrm{Conc}}(\pi_* FY, \rho)$ with respect to its single parameter ρ we obtain the desired value of $\rho_{\max}$, which can be further substituted into (11) and (12).

Finally, let us notice that Case II simplifies to Case I if $M_H = M_{X_2}$ thus the remaining part of the paper considers Case II only. Still, in this section, we have retained the presentation of both cases separately since Case I is considerably simpler in application and should be used instead of Case II whenever possible.

5. ASYMPTOTICS OF THE PARAMETER ESTIMATES

In this section we formulate two statements on asymptotic behaviour of the ML estimator presented in section 3. First of those statements concerns consistency, second concerns limiting variance of the estimates from our ML estimator.

Large sample theory for maximum likelihood estimation establishes, under some assumptions, two important facts about the ML estimator $\hat{\theta} = \hat{\theta}_n$. Firstly, it is the consistency of ML estimates and secondly the limiting distribution for the quantity $\sqrt{n}(\hat{\theta} - \theta_0)$, where θ_0 is the true parameter value. Clearly, in the case of the SAR model, given by equation (4), the observed sample $Y = (Y_1, \dots, Y_n)^{\mathrm{T}}$ is not independent thus the classical textbook results are not applicable. Nonetheless, it has been a commonplace since the early days of applied spatial econometrics (c.f. Anselin, 1988) to assume that $\hat{\theta}$ is consistent and the its deviation $\sqrt{n}(\hat{\theta} - \theta_0)$ is asymptotically normal with zero mean and variance $\left[-\frac{1}{n} \mathbb{E} \frac{\partial^2}{\partial \theta^2} \log L(Y, \theta) \right]^{-1}$. This popular belief was supported by the fact that any sensible asymptotic theory (covering at least the increasing domain scheme) would definitely have to give asymptotics of the form mentioned. This is because, such a theory would have to include the simple asymptotic setting, in which there exists a sequence of parallel spatial domains, independent and unrelated to one another, being included in the sample as n increases. Notice that this simple setting is subject to vector-valued independent sample ML asymptotics theorem. If one further assures identifiability and uniqueness of the maximiser, the above-mentioned asymptotics follow.

With the papers of Kelejian and Prucha (2001) as well as Lee (2004) it became apparent that an asymptotic theory covering more sophisticated asymptotic settings is possible. Using general tools for consistency and asymptotic normality proofs (described in e.g. Pötscher, Prucha, 1997) one can construct asymptotic theory for ML estimates covering both infill and increasing domain schemes. The crucial assumptions

that have to be made concern the spatial weight $\mathbf{W}$ and the design matrix X. From those assumptions identifiable uniqueness of parameters and a certain uniform law of large numbers for the log likelihood function can be deduced. Those two elements allow one to utilize the theory of general M-estimators from Pötscher, Prucha (1997) to obtain the desired results.

Below we describe (after Lee, Yu, 2010) a set of possible assumptions which assure fairly general statement about asymptotics of ML estimates. Apart from the natural postulates of the zero diagonal of spatial weight matrix $\mathbf{W} = \mathbf{W}(n)$ we mention the following.

Assumption 0. The error term ε in (4) follows multivariate normal distribution with uncorrelated, homoscedastic components. In particular, this implies that all moments of ε are finite.

Assumption 1[11]. For elements ρ of its parameter space[12] $\mathcal{R}$ the spatial lag operator $I_n - \rho\mathbf{W}$ is invertible and the true value of ρ_0 is an interior point of $\mathcal{R}$.

Assumption 2[13]. There exists a constant C such that for any rows or columns, say v, of any of the matrices $\mathbf{W}$, $\mathbf{W}_{\pi_*F}$ and $(I_n - \rho\mathbf{W})^{-1}$, $\left(I_{n_*} - \rho\mathbf{W}_{\pi_*F}\right)^{-1}$, $\rho \in \mathcal{R}$, $n \in \mathbb{N}$, its ℓ_1-norm[14] $\|v\|_1$ does not exceeds C.

Assumption 3[15]. The elements of non-stochastic design matrix $X = X(n)$ are bounded and the sequence $\frac{1}{n}X^T M_H X$ converges to a non-singular limit.

Assumption 4. The ratio n_*/n converges as $n \to \infty$ and $\alpha_* = \liminf_{n\to\infty} \frac{n_*}{n} > 0$.

Assumption 5[16]. Estimated parameters are uniquely identified[17].

Let us note that the natural Assumption 1 is crucial not only for identifiability of the parameter ρ but it is also necessary for our ability to present a closed form of Y from (4), thus for effective interpretation of the model. Assumption 2 limits spatial dependence to 'manageable degree'. This means, in particular, that the amount of information obtained from a larger sample is sufficient to decrease variance of estimates. Assumption 3 assures that the design matrix is well-behaved and in particular through non-singularity of the corresponding limit conveys sufficient information on the slope parameters of interest.

Assumption 4 guarantees that the dimension of appropriate Krylov space does not reduce the number of available degrees of freedom excessively. Assumption 5 assures that the hypothetical probability distributions for different parameter values remain clearly distinguishable by ML estimation procedure as sample increases. This assumption is typically expanded into a highly technical statement involving terms from log likelihood function, so that it implies unique identification. In our paper, to

---

[11] C.f. Lee, Yu, page 167, Assumption 3 therein.
[12] See footnote 10.
[13] C.f. Lee, Yu, page 167, Assumption 5 therein.
[14] For either a column or row vector $v = (v_1, ..., v_n)$ its ℓ_1-norm is $\|v\|_1 = \sum_{i=1}^{n}|v_i|$.
[15] C.f. Lee, Yu, page 167, Assumption 4 therein.
[16] C.f. Lee, Yu, page 168, Assumption 7 therein.
[17] In the sense of Definition 3.1 in Pötscher, Prucha (1997).

avoid unnecessary complexity we decide to readably assume the unique identification of parameters.

Lastly, for completeness of presentation, we conclude with the asymptotic distribution of $\hat{\theta}$. Namely, under the assumptions 1–5 we can state that $\sqrt{n_*}(\hat{\theta} - \theta_0)$ converges in distribution to $\mathrm{N}(0, n_* I(\theta_0)^{-1})$. More precisely, conducting differentiation and applying expectation in the score matrix we obtain a formula for the Fisher information $I(\theta, n)$ for parameter $\theta = (\beta, \rho, \sigma^2)$. Namely, denoting $\overline{\mathbf{W}} = \mathbf{W}(\mathrm{I}_n - \rho\mathbf{W})^{-1}$ and $\overline{\mathbf{W}}_* = \mathbf{W}_{\pi_* F}(\mathrm{I}_{n_*} - \rho\mathbf{W}_{\pi_* F})^{-1}$ we obtain[18]

$$I(\theta, n) = \begin{bmatrix} \sigma^{-2} X_1^T M_H X_1 & I_{\beta,\rho}(\theta) & 0 \\ I_{\beta,\rho}(\theta)^{\mathrm{T}} & I_{\rho,\rho}(\theta) & \sigma^{-2}\mathrm{trace}(\overline{\mathbf{W}}_*) \\ 0 & \sigma^{-2}\mathrm{trace}(\overline{\mathbf{W}}_*) & \dfrac{n_*}{2}\sigma^{-4} \end{bmatrix},$$

$$I_{\beta,\rho}(\theta) = \sigma^{-2} X_1^{\mathrm{T}} M_H \overline{\mathbf{W}} M_H X_1 \beta,$$

$$I_{\rho,\rho}(\theta) = \mathrm{trace}\left(\overline{\mathbf{W}}_*\overline{\mathbf{W}}_* + \overline{\mathbf{W}}_*^{\mathrm{T}}\overline{\mathbf{W}}_*\right) + \beta^{\mathrm{T}} X_1^{\mathrm{T}} M_H \overline{\mathbf{W}}^{\mathrm{T}} M_H \overline{\mathbf{W}} M_H X_1 \beta.$$

Setting $\Sigma = \lim_{n \to \infty} n_* I(\theta_0, n)^{-1}$, we have convergence in distribution of $\sqrt{n_*}(\hat{\theta} - \theta_0)$ to $\mathrm{N}(0, \Sigma)$ provided that the limit Σ exists.

6. DISCUSSION OF THE ADOPTED ASSUMPTION

The Assumptions 1–3 and 5 are well known in spatial econometric literature. Extensive discussion on the topic has been given in numerous papers e.g. Kelejian, Prucha (2001), Lee (2004), Lee, Yu (2010). The new assumption introduced in this paper is the Assumption 4, which connects n – the increasing sample size, with the amount of degrees of freedom lost due to the use of the generalized demeaning procedure. In terms of a standard fixed effects setting, where $n = N \cdot T$, Assumption 4 is equivalent to requiring that $T \to \infty$, when N spatial fixed effects are present, and requiring that $N \to \infty$, whenever T temporal fixed effects are included in the model.

Obviously, in a fully general case it cannot be guaranteed that the requirement in Assumption 4 is satisfied. Then, a natural question arises: is Assumption 4 often met in practice? It turns out that some "rules of thumb" can be formulated which imply affirmative answer in many practical settings. We will present them in the following examples as well as in an empirical illustration described in next section. For simplicity, we consider the case of $n = N \cdot T$, that is a standard balanced panel

---

[18] C.f. Elhorst (2014), page 46.

data set. Moreover, the spatial $n \times n$ weight matrix is purely spatial, i.e. it does not contain any dynamic references.

For arbitrary $m \in \mathbb{N}$ let us denote $\mathbf{1}_m = (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^m$. If the spatial weight matrix $\mathbf{W}$ is constant in time, i.e. $\mathbf{W} = \bar{\mathbf{W}} \otimes \mathrm{I}_N$, and the fixed effect design X_2 is constant in time (i.e. each column of X_2 is of the form $v \otimes \mathbf{1}_T$, for some $v \in \mathbb{R}^N$) then $\mathbf{W}X_2$ is also time-constant. By induction, we infer that the Krylov space H for X_2 contains only time-constant vectors. Finally, $\dim H \leq N$, thus $\frac{n_*}{n} = 1 - \frac{\dim H}{NT}$, which converges to 1 as $T \to \infty$.

Another example is when, as it is very often found in practice, $\mathbf{W}$ is row-standardized, and when the matrix X_2 contains purely temporal effect of arbitrary shape. That is, each column of X_2 is of the form $\mathbf{1}_N \otimes v$, for some $v \in \mathbb{R}^T$, then $\mathbf{W}(\mathbf{1}_N \otimes v) = \mathbf{1}_N \otimes v$. This implies that $\dim H = k_2$, thus $\alpha_* = 1$. Even if k_2 is not bounded, Assumption 4 is satisfied if $N \to \infty$, since obviously $k_2 \leq T$.

7. AN EMPIRICAL ILLUSTRATION

The background for this empirical illustration is a theoretical model developed by Fingleton (2001, 2004), which is based on the NEG theory and Verdoorn's law (c.f. Verdoorn, 1949). This law links the increase in labour productivity with an increase in production. More precisely, the Verdoorn's law states that in a long run productivity grows proportionally to the square root of output. According to e.g. Fingleton (2001), the exponential growth rate of productivity can be modelled by the use of the following specification

$$p = \alpha_0 + \rho \mathbf{W} p + \alpha_1 H + \alpha_2 G_0 + \alpha_3 q + \varepsilon, \ \varepsilon \sim N(0, \sigma^2)$$

where: p represents the exponential growth rate of productivity, $\mathbf{W}$ is a spatial weight matrix, H refers to human capital, G_0 is the initial level of technology, and q is the exponential growth rate. As described in Olejnik and Olejnik (2017) the specification can be further transformed into the following Spatial Panel Durbin Model

$$p = \rho \mathbf{W} p + \pi_1 q + \pi_2 \mathbf{W} q + \eta_1 H + \mathbf{FE} + \varepsilon, \tag{13}$$

with α_0, ρ, π_1, π_2, η_1 being model parameters, $\mathbf{W}$ is spatial weight matrix. The term G_0 does not appear in (13) as they have been incorporated into fixed effects $\mathbf{FE}$. In our example the fixed effects are $2N$ dummy variables of the form $e_i \otimes v_{2004}$ and $e_i \otimes v_{2008}$, $i = 1, \dots, N$, where v_{2004} and v_{2008} are fixed effects distinguishing periods after EU enlargement and global financial crises in 2008, respectively. Notice that the groups of observations distinguished by this fixed effect design are overlapping, thus the standard demeaning procedure cannot be used. Moreover, the additional term of spatially lagged exogenous variable q has been introduced into (13) to account for additional externalities.

The data for the example covers 261 regions of EU for the years 2000–2013. The productivity growth p for the years 2001–2013 is approximated by the exponential rate of change of regional productivity (quotient of regional production over the number of economically active population) related to regional productivity in the initial year 2000. Similarly, the exponential growth rate is approximated by logarithm of the ratio of regional production in years 2001–2013 to the base year 2000. The matrix **W** is a row-standardised spatial weight matrix of three nearest neighbours (c.f. Anselin, 1988). The human capital H is approximated by employment in technology and knowledge-intensive sectors expressed as a percentage of economically active population, expressed in logarithms.

For the purpose of empirical comparison we apply both standard Maximum Likelihood estimation procedure using dummy variables and our modified approach. Results are presented in table 1.

Table 1.

Comparison of standard ML and the augmented ML approach

Parameter	Corresponding variable	Standard ML		New ML	
		Coeff.	t-stat	Coeff.	t-stat
ρ	**W**p	0.64	49.44	0.64	45.48
π_1	q	0.74	56.50	0.74	51.97
π_2	**W**q	-0.45	25.66	-0.45	23.61
η_1	H	0.09	9.69	0.09	8.92
σ^2	Error variance	1.0542		1.2459	
$\overline{R^2}$	Goodness of fit	0.9517		0.9429	

Source: own calculation.

Table 1 shows that both estimation procedures yield virtually the same values for both autoregressive and regressive parameters (ρ and β_1 respectively in notation from previous sections). However, there is a considerable difference in estimates of the σ^2 parameter. As expected, our procedure yields a consistent estimates of the error variance, which turns out to be rather cautious. This is because the standard ML estimate does not properly reflect the loss of degrees of freedom related to the use of general fixed effect dummies. In contrast our estimation scheme, through consideration of the dimension of Krylov space for fixed effects design matrix, allows one to estimate σ^2 and also goodness-of-fit measures more reliably. Moreover, if the size of fixed effect design grows with sample size (e.g. in our example $2N$ might grow with n), then the standard ML estimate of σ^2 might even turn out to be inconsistent.

8. SUMMARY

Since the early days of spatial econometrics it has been known that ordinary least squares procedure for estimating model parameters in the case of spatial autoregressive specification leads to inconsistent estimates. This is because, the specification incorporates the lagged dependant variable term as one of the regressors. Maximal likelihood procedure has been long considered a remedy for this endogeneity problem. Although, new alternatives to ML have been found (e.g. generalized method of moments) the original procedure of maximal likelihood remains widely used by practitioners.

In this paper we have proposed an alternative to partial regression in a spatial autoregressive econometric model when the maximum likelihood procedure is used. Under certain assumptions on the dimension of some invariant space associated with spatial weight matrix we have managed to formulate a feasible procedure, which can be used to handle large class of fixed effect designs. This can be done at the expense of possibly decreased number of degrees of freedom in Gaussian log likelihood function.

Our result contradicts the conjecture, expressed in a celebrated paper by Lee, Yu (2010) on bias correction in the case of incidental parameter problem, that such scheme would not be possible except cases of individual fixed effects and time fixed effects with row standardized spatial weight matrix. As in our reasoning we carefully manage the degrees of freedom at the step of sample transformation (demeaning), estimator bias in the case of incidental parameter problem does not occur in our setting, cf. Elhorst (2014).

REFERENCES

Anselin L., (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Anselin L., Le Gallo J., Jayet H., (2006), *Spatial Panel Econometrics*, in: Matyas L., Sevestre P., (eds.), *The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice*, 3rd edition, Kluwer, Dordrecht.

Baltagi B. H., (2005), *Econometric Analysis of Panel Data*, John Wiley & Sons, Chichester West Sussex, England.

Elhorst J. P., (2014), *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, Springer, Heidelberg.

Elhorst J. P., Fréret S., (2009), Evidence of Political Yardstick Competition in France Using a Two-Regime Spatial Durbin Model with Fixed Effects, *Journal of Regional Science*, 49 (5), 931–951.

Fiebig D. G., Bartels R., (1996), The Frisch-Waugh Theorem and Generalized Least Squares, *Econometric Reviews*, 15 (4), 431–443.

Frisch R., Waugh F. V., (1933), Partial Time Regressions as Compared with Individual Trends, *Econometrica*, 1 (4), 387–401.

Greene W. H., (2008), *Econometric Analysis*, 6th edition, Prentice Hall, Upper Saddle River, N. J.

Kelejian H. H., Prucha I. R., (2001), On the Asymptotic Distribution of the Moran I Test Statistic with Applications, *Journal of Econometrics*, 104, 219–257.

Lee L.-F., (2004), Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models, *Econometrica*, 72 (6), 1899–1925.

Lee L.-F., Yu J., (2010), Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects, *Journal of Econometrics*, 154 (2), 165–85.

Liesen J., Strakoš Z., (2013), *Krylov Subspace Methods: Principles and Analysis*, Oxford University Press, Oxford UK.

Olejnik A., Olejnik J., (2017), Increasing Returns to Scale, Productivity and Economic Growth – A Spatial Analysis of the Contemporary EU Economy, submitted to *Argumenta Oeconomica*.

Olejnik A., Özyurt S., (2016), Multi-Dimensional Spatial Auto-regressive Models: How Do They Perform in an Economic Growth Framework?, working paper.

Pace R. K., (2014), *Maximum Likelihood Estimation*, in: Fisher M. M., Nijkamp P., (eds.), *Handbook of Regional Science*, Springer-Verlag, Berlin Heidelberg.

Pötscher B. M., Prucha I., (1997), *Dynamic Nonlinear Econometric Models*, Springer-Verlag, Berlin Heidelberg.

Yamada H., (2016), The Frisch–Waugh–Lovell Theorem for the Lasso and the Ridge Regression, *Communications in Statistics – Theory and Methods*, Available online at: http://dx.doi.org/10.1080/03610926.2016.1252403.

PROCEDURA ALTERNATYWNA DO REGRESJI CZĘŚCIOWEJ W ESTYMACJI MODELU PRZESTRZENNEGO AUTOREGRESYJNEGO METODĄ NAJWIĘKSZEJ WIAROGODNOŚCI

Streszczenie

W niniejszej pracy wprowadzono procedurę alternatywną do procedury regresji częściowej. Opisane postępowanie może być zastosowane w przypadku estymowania parametrów modelu przestrzennego autoregresyjnego metodą największej wiarogodności. Przy pewnych założeniach dotyczących wymiaru pewnej przestrzeni niezmienniczej związanej z macierzą wag przestrzennych sformułowany jest schemat postępowania obejmujący szeroką klasę macierzy efektów stałych. W pewnych przypadkach opisana procedura może eliminować efekty stałe kosztem obniżonej liczby stopni swobody. Dodatkowo, przedstawiono własności asymptotyczne zaprezentowanego estymatora.

Słowa kluczowe: regresja częściowa, metoda największej wiarogodności, model przestrzenny autoregresyjny, model z efektami stałymi

AN ALTERNATIVE TO PARTIAL REGRESSION IN MAXIMUM LIKELIHOOD ESTIMATION OF SPATIAL AUTOREGRESSIVE MODEL

Abstract

In this paper an alternative procedure to partial regression is introduced. The presented procedure can be used in maximum likelihood estimation of spatial autoregressive model. Under certain assumptions on the dimension of certain invariant space associated with spatial weight matrix a feasible procedure is formulated, which can be used to handle large class of fixed effect designs. This is done at the expense of possibly decreased number of degrees of freedom in the Gaussian log likelihood function. Additionally, a statement on asymptotic behaviour of presented estimator is given.

Keywords: partial regression, maximum likelihood estimation, spatial autoregressive model, fixed effects model

WITOLD RZYMOWSKI[1], AGNIESZKA SUROWIEC[2]

MODELLING POPULATION GROWTH
WITH DIFFERENCE EQUATION METHOD

1. INTRODUCTION

The problem to create the demographic model to predict the world or countries population in a fixed time range has been investigated extensively for many years. The two oldest models: an exponential model by the British economist Thomas Malthus and a logistic one by the Belgian mathematician, Pierre Verhulst are the most known models in case of the world human population. Malthus framed a model based on the observation that biological populations, including human ones, tended to increase at rate proportional to the population size (Malthus, 1798). The model of Verhulst (Verhulst, 1838) differs from the Malthusian model by changing some assumption. This model was rediscovered and popularized in the 1920's by Pearl, Reed (1924). Both of these models, the Malthusian and the Verhulst model, really do not work in the longer periods of time (Murray, 1989). Malthus reasoned that an exponential growth of the world's population could not go on infinitely and therefore one must interrupt the inexorable working of model by artificially reducing the size of the population. The Verhulst model does not adequately describe either short-range changes or very long-range trends in human population growth. Pearl, Reed (1924) predicted a maximum world population of about 2 billion, which was exceeded by 1930.

In the work Smith (1977), in addition to above mentioned models, author examines the Doomsday model (Foerster et al., 1960). The unreliability of the Doomsday model consisting in the fact that the actual world population is slightly ahead of the Doomsday projection, nearly a generation after it was made, was observed in Austin, Brewer (1971) and again in Serrin (1975).

In the paper Rzymowski, Surowiec (2012) the authors propose a pseudologistic model of world population with three parameters estimated by the method which minimizes relative error. This model gives a better description of the world human population than the ones mentioned above.

[1] Lublin University of Technology, Fundamentals of Technology Faculty, Department of Applied Mathematics, 38 Nadbystrzycka St., 20-618 Lublin, Poland.
[2] Lublin University of Technology, Management Faculty, Department of Quantitative Methods in Management, 38 Nadbystrzycka St., 20-618 Lublin, Poland, corresponding author – e-mail: a.surowiec@pollub.pl.

All of these models of living beings are time dependent models and they tend to have the form:

$$L_t = f(t) + \varepsilon_t, \ t = 1,2,\ldots N. \tag{1}$$

L_t is the function of time t and it is the size of the population under consideration. L_t can be considered as a dynamic system.

Therefore in this article we propose a new approach – the difference equation method – to obtain the time dependent models of the human population in countries or groups of countries[3]. We also consider the data representing the population in the world. We take into account the data from the years 1950–2011 ($N = 62$), the data from the years 1950–2012 ($N = 63$) and the data from the years 1950–2013 ($N = 64$) depending on the country. The list of the countries and number of data corresponding them is presented in table 1.

Table 1.

The list of countries or group of countries and corresponding them the number of data N

Country	N	Country	N
Australia	64	New Zealand	64
Austria	64	Norway	64
Belgium	63	Poland	63
Canada	63	Portugal	64
Chile	64	Slovak Republic	64
Czech Republic	64	Slovenia	64
Denmark	64	Spain	64
Estonia	64	Sweden	64
Finland	64	Switzerland	62
France	64	Turkey	64
Germany	64	United Kingdom	64
Greece	63	United States	64
Hungary	64	G7	64
Iceland	62	OECD – Total	64
Ireland	64	World	64
Israel	64	Brazil	63
Italy	64	China	64
Japan	64	India	64
Luxembourg	63	Indonesia	64
Mexico	64	Russia	64
Netherlands	64	South Africa	63

Source: own elaboration.

---

[3]   Available at http:/stats.oecd.org (04 Mar 2016).

The difference equations have not been studied well to describe dynamic systems but many dynamic systems in many fields of science including physics, engineering, economics and biomedical sciences can be described by means of the second order differential equations (Li et al., 2002; Huang et al., 2006; Ramsay et al., 2007; Chen, Wu, 2008; Miao et al., 2009; Liang, Wu, 2008; Huang, 2010). The studies of differential equations in literature have mainly focused on the so-called forward problem, i.e., simulation and analysis of the behavior of state variables for a given system. The inverse problem, using the measurements of state variables to construct the econometric model and estimate the parameters that characterize the system, has not been studied well. Within the scope of investigation in this field an application example from human population dynamic study is tested in this article. The model of the number of living beings in chosen country in year t, $t = 1,2,\ldots N$ (see eq. (1)) we consider as a discrete time model.

The aim of this work is to find the best model of living beings in chosen country. To estimate the parameters in our model of human population dynamic we use the least squares principle that are used comparatively often by mathematicians (Hemker, 1972; Bard, 1974; Li et al., 2005), computer scientists (Varah, 1982), and chemical engineers (Ogunnaike, Ray, 1994; Poyton et al., 2006). We also use the generalized Least Squares Method (Nowak, 2006; Rao, 1982).

The measure applied to verify the model of human population is the relative error. However, it should be pointed out that the empirical verification is only the correct final evaluation of the quality of the model.

2. THE DIFFERENCE EQUATIONS METHOD

Use of the second order differential equation is very popular to describe the dynamic systems in many fields of science. For example, the model:

$$x'' + 2\zeta\omega_0 x' + \omega_0^2 x = F_{ext}, \qquad (2)$$

is the model of a damped harmonic oscillator, where x is mass's position, $\dfrac{dx}{dt}$ is mass's velocity, $\dfrac{d^2 x}{dt^2}$ is mass's acceleration, $\omega_0 = \sqrt{\dfrac{k}{m}}$ is called the undamped angular frequency of the oscillator and $\zeta = \dfrac{c}{2\sqrt{mk}}$ is called the damping ratio. In the model of a damped harmonic oscillator the coefficients ω_0 and ξ are positive. The equation (2) describes the behavior of the system where friction (frictional force $F_f = -cv$) or damping (damping force $F_d = -kx$) slows the motion of the system with known parameters m, k, c, where m is mass, k is spring constant and c is called the viscous damping coefficient.

Let L_t represent the number of living beings in chosen country in year t, $t = 1, 2, \ldots N$. Then the derivatives L'_t, L''_t do not exist for our model (1). Therefore, we replace the differential equation (2) by the difference equation (3).

In analogy to the equation (2) we can consider the following equation for L_t:

$$L''_t = aL'_t + bL_t + c \, , \ t = 5, 6, \ldots, \tau \, , \tag{3}$$

where a, b, c are the unknown parameters. $\tau = 10, 11, \ldots, N - 4$.

Assuming (Lanczos, 1964):

$$L'_t = -\frac{1}{5}L_{t-2} - \frac{1}{10}L_{t-1} + \frac{1}{10}L_{t+1} + \frac{1}{5}L_{t+2}, \ \ t = 3, 4, \ldots, N - 2 \, , \tag{4}$$

$$L''_t = -\frac{1}{5}L'_{t-2} - \frac{1}{10}L'_{t-1} + \frac{1}{10}L'_{t+1} + \frac{1}{5}L'_{t+2}, \ \ \ t = 5, 6, \ldots, N - 4 \, , \tag{5}$$

we can obtain the following difference model for L_t:

$$L_t = \alpha L_{t-1} + \beta L_{t-2} + \gamma L_{t-3} + \delta L_{t-4} + \zeta + \varepsilon_t, \ t = 5, 6, \ldots, \tau \, , \tag{6}$$

where α, β, γ, δ, ζ are the unknown parameters and ε_t are the residuals. The model (6) is the discrete time model of L_t.

We estimate the parameters in model (6) with use the generalized Least Squares Method (Nowak, 2006; Rao, 1982).

To solve the equation (6) we consider the characteristic equation of this equation (6) in the form:

$$\lambda^4 - \alpha\lambda^3 - \beta\lambda^2 - \gamma\lambda - \delta = 0 \, . \tag{7}$$

Depending on the values of parameters α, β, γ, δ the equation (6) can have one of the ten forms of model of L_t (Koźniewska, 1972; Sierpiński, 1946) but only three forms of model of L_t:

1. M1: $\hat{L}_t = C_1\lambda_1^t + C_2\lambda_2^t + C_3\lambda_3^t + C_4\lambda_4^t + C_5,$ (8)

    where $\lambda_1, \ \lambda_2, \lambda_3, \lambda_4 \in R$,

2. M2: $\hat{L}_t = C_1\lambda_1^t + C_2\lambda_2^t + C_3r^t \cos(\varphi t) + C_4r^t \sin(\varphi) + C_5,$ (9)

    where $\lambda_1, \ \lambda_2 \in R$ and $\lambda_3 = a + bi$, $\lambda_4 = a - bi$ and $b > 0$ are complex numbers and $r = \sqrt{a^2 + b^2}$, $\varphi = \arccos\dfrac{a}{r}$,

3. M3: $\hat{L}_t = C_1 r_1^t \cos(\varphi_1 t) + C_2 r_1^t \sin(\varphi_1 t) + C_3 r_2^t \cos(\varphi_2 t) + C_4 r_2^t \sin(\varphi_2 t) + C_5,$ (10)

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are complex numbers: $\lambda_1 = a_1 + b_1 i, \lambda_2 = a_1 - b_1 i, \lambda_3 = a_2 + b_2 i,$ $\lambda_4 = a_2 - b_2 i,$ where $b_1, b_2 > 0$ and $r_1 = \sqrt{a_1^2 + b_1^2}, r_2 = \sqrt{a_2^2 + b_2^2}, \varphi_1 = \arccos \frac{a_1}{r_1},$ $\varphi_2 = \arccos \frac{a_2}{r_2}$

can be achieved numerically because it is very difficult to obtain zero or one numerically.

The parameters C_1, C_2, C_3, C_4, C_5 can be estimated by using the Least Squares Method for chosen form of model of L_t for $t = 5,6,\ldots,\tau$.

To verify the models obtained with the difference equations method we calculate for every $\tau = 10,11,\ldots, N - 4$:

– the theoretical values of human population $\hat{L}_t(\tau)$, $t = 5,6,\ldots, N$ (a full range of data from the years 1950–2011 or 1950–2012, or 1950–2013) according to the country or group of countries (see table 1) with use M1, M2 or M3 model by equations (8), (9) or (10),
– the relative errors $\delta_t(\tau)$:

$$\delta_t(\tau) = \frac{\left| L_t - \hat{L}_t(\tau) \right|}{L_t} 100\%, \; t = 5,6,\ldots, N,$$ (11)

– the maximum relative errors $\delta_N(\tau)$ and $\delta_p(\tau)$ for the total and prediction range respectively:

$$\delta_N(\tau) = \max_{t=5,6,\ldots N} \delta_t(\tau),$$ (12)

$$\delta_p(\tau) = \max_{t=\tau+1,\tau+2,\ldots,N} \delta_t(\tau),$$ (13)

and
– τ^* that corresponds to

$$\delta_N(\tau^*) = \min_{\tau=10,11,\ldots,N-4} \delta_N(\tau),$$ (14)

where $\delta_N(\tau)$ is given by equation (12). The measure to verify the obtained model is $\delta_p(\tau^*)$. We find also τ^{**} such that $\delta_p(\tau) \leq 5\%$ for $\tau \geq \tau^{**}$.

We tested the models defined by equation (6) for all the analyzed countries from table 1.

3. THE RESULTS FOR POPULATION MODELS

The relative errors

Table 2 shows values of τ^*, models corresponding to them (see eq. (8), (9) and (10)), the relative errors $\delta_N(\tau^*)$ (see eq. (12) and (14)) and $\delta_p(\tau^*)$ (see eq. (13) and (14)) and τ^{**} for the human population for all analyzed countries, group of countries and for the world.

Table 2.

Values of τ^*, models corresponding to them, the relative errors $\delta_N(\tau^*)$ and $\delta_p(\tau^*)$ and τ^{**} for the human population in analysed country

Country	τ^*	Model	$\delta_N(\tau^*)$ [%]	$\delta_p(\tau^*)$ [%]	τ^{**}
Australia	60	M2	2.92	2.92	58
Austria	59	M2	3.20	3.20	55
Belgium	24	M2	3.76	3.76	59
Canada	59	M1	2.18	2.18	47
Chile	60	M1	1.95	1.95	39
Czech Republic	60	M2	1.56	1.28	24
Denmark	59	M2	2.61	2.61	53
Estonia	60	M3	8.68	8.68	-
Finland	60	M2	1.83	1.83	43
France	60	M2	2.08	2.08	38
Germany	50	M2	2.53	2.45	29
Greece	40	M2	2.24	2.16	52
Hungary	60	M3	4.11	4.11	54
Iceland	15	M2	4.70	4.70	-
Ireland	60	M2	4.87	3.52	59
Israel	57	M2	7.45	1.96	48
Italy	60	M2	4.41	4.41	59
Japan	58	M3	2.65	2.40	50
Luxembourg	59	M2	3.08	3.08	59
Mexico	56	M1	2.53	2.16	47
Netherlands	54	M2	0.87	0.78	38
New Zealand	57	M2	4.13	4.06	54
Norway	59	M2	0.76	0.54	58

Country	τ^*	Model	$\delta_N\left(\tau^*\right)$ [%]	$\delta_p\left(\tau^*\right)$ [%]	τ^{**}
Poland	58	M3	1.68	1.65	45
Portugal	60	M2	6.92	0.91	47
Slovak Republic	60	M2	2.07	1.57	42
Slovenia	54	M2	3.07	2.83	48
Spain	59	M2	3.99	3.99	58
Sweden	60	M2	2.45	2.45	52
Switzerland	58	M2	5.93	5.93	-
Turkey	54	M2	1.23	1.11	46
United Kingdom	60	M1	4.66	4.66	59
United States	51	M2	1.48	1.48	42
G7	58	M2	0.88	0.82	39
OECD - Total	59	M2	0.37	0.37	30
World	53	M3	0.55	0.55	44
Brazil	24	M2	1.95	1.95	49
China	60	M2	2.88	2.87	53
India	51	M3	0.41	0.32	44
Indonesia	60	M3	0.92	0.89	47
Russia	58	M2	3.07	3.07	50
South Africa	56	M3	1.21	1.21	54

Source: own elaboration.

The shaded cell in table 2 corresponds to the minimal value of $\delta_N\left(\tau^*\right)$ or $\delta_p\left(\tau^*\right)$ for all analyzed countries, group of countries and for the world. The minimal values of $\delta_N\left(\tau^*\right)$ and $\delta_p\left(\tau^*\right)$ are equal to 0.37 for OECD – Total and 0.32 for India, respectively.
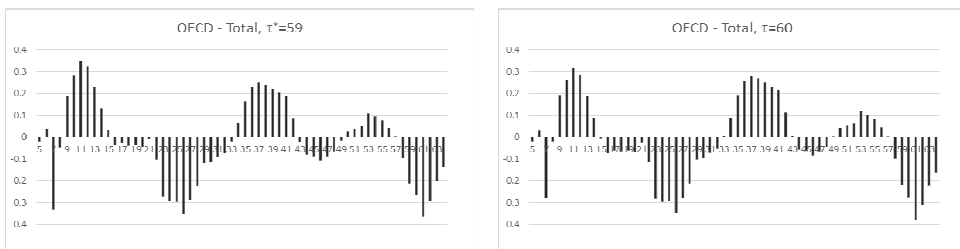

Figure 1. Comparison the δ_t, $t = 5,6,\ldots, N$ for OECD-Total population for τ^* (left) and for $\tau = N - 4$ (right)
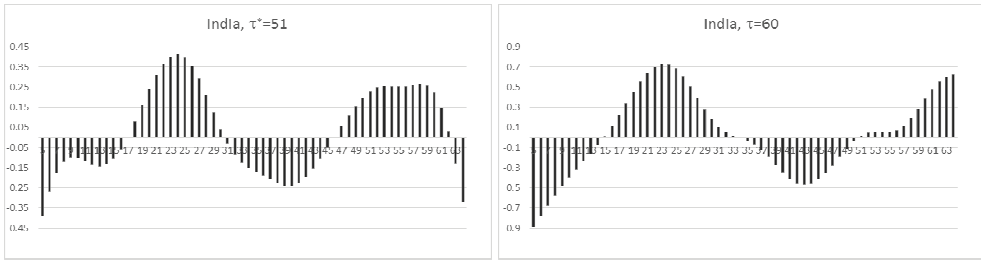Source: own elaboration.

Figure 2. Comparison the δ_t, $t = 5,6,\ldots, N$ for India population for τ^* (left) and for $\tau = N - 4$ (right)

Source: own elaboration.

Figure 1 and figure 2 presents the comparison between $\delta_t\left(\tau^*\right)$ and δ_t ($\tau = N - 4$) for OECD – Total and for India, respectively. The results presented in figure 1 for $\tau^* = 59$ and for $\tau = N - 4$ for OECD – Total are very similar. The models obtained for this group of countries are not very sensitive to change the τ. OECD – Total is third in order after Czech Republic and Germany with small number of τ^{**}. The figure 3 presents the values of $\delta_N\left(\tau\right)$ for $\tau = 30,31,\ldots, N - 4$ for OECD – Total. The best model for OECD – Total population is model of M2 type (see eq. (9)).

Comparing the results presented in figure 2 for India population we can find that the model obtained for $\tau^* = 51$ is better than the model obtained for $\tau = N - 4$ in spite of the longer period of estimation. The best model for India population is model of M3 type (see eq. (10)).
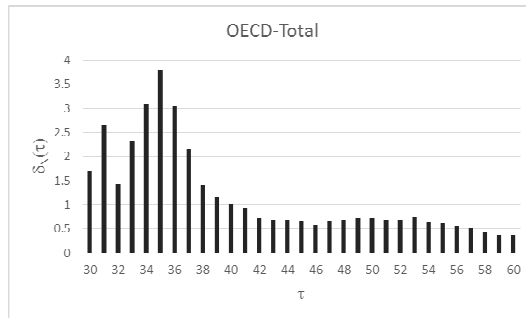

Figure 3. $\delta_N\left(\tau\right)$ for $\tau = 30,31,\ldots, N - 4$ for OECD – Total

Source: own elaboration.

Taking into account the value of $\delta_p\left(\tau^*\right)$ we can obtain 8 models out of 42 with $\delta_{\tau^*} < 1\%$. Only 2 models out of 42 have the value of $\delta_p\left(\tau^*\right)$ greater than 5%. In 30 countries out of 42 $\tau^{**} \leq 54$ what guarantees 10 years at least forecast with $\delta_p\left(\tau\right) \leq 5\%$ (see table 2).

Only in 3 country out of 42 (Estonia, Iceland and Switzerland, see table 2) τ^{**} do not exist.

Remarks about the world population model

Applying the difference equations method for the world population model we can obtain the model with very small maximal relative error. The minimal value of the maximal relative error for the world population model is less than 0.6% for $\tau^* = 53$ (see table 2) and for $\tau = 52$ and $\tau = 54$. It guarantees at least 10 years forecast with $\delta_p(\tau) \leq 0.6\%$. In figure 4 we present the comparison between relative errors distributions for $\tau^* = 53$ and for $\tau = 54$. The models for world population for these values of τ are the models of M3 type (see eq. (10)). The models obtained for the world population are not very sensitive to change the τ. It is the same result like this presented in figure 1 for OECD-Total. M3 type of model is obtained for world population for $\tau = 44,45,\ldots,57$. But if we consider the model for world population for $\tau = N - 4$ we will obtain the model of M2 type (see eq. (9)). In figure 5 we present the relative errors distributions for $\tau = N - 4$. It differs from this presented in figure 4.
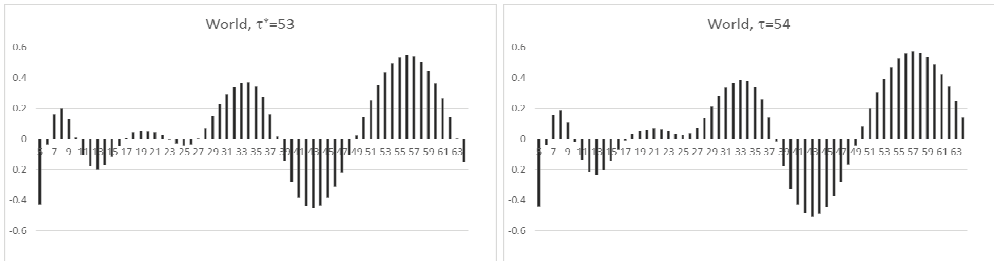

Figure 4. Comparison the δ_t, $t = 5,6,\ldots, N$ for world population for $\tau^* = 53$ (left) and for $\tau = 54$ (right)
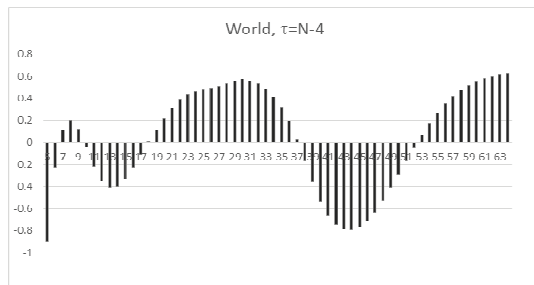
Source: own elaboration.

Figure 5. The distribution δ_t, $t = 5,6,\ldots, N$ for world population for $\tau = N - 4$

Source: own elaboration.

The parameters of the models for $\tau^* = 53$, $\tau = 54$ and for $\tau = N - 4 = 60$ are presented in table 3.

Table 3.

The parameters of M2 and M3 models obtained for world population
for $\tau^* = 53$, $\tau = 54$ and for $\tau = N - 4 = 60$

| Parameters	$\tau^* = 53$, (M3)	$\tau = 54$, (M3)	$\tau = N - 4 = 60$ (M2)
λ_1	-	-	0.993334625
λ_2	-	-	0.978372676
r	-	-	0.898565728
φ	-	-	0.33155998
r_1	1.007269262	1.002126621	-
φ_1	0.025390024	0.023948069	-
r_2	0.900067272	0.896320454	-
φ_2	0.376079832	0.36475173	-
C_1	-1233332453	-2446916724	-26107791799
C_2	2127947086	2099598613	5630888596
C_3	-59462245.29	-65658485.65	-99683889.48
C_4	22589702.33	32577674.82	87080983.82
C_5	3667243839	4875343817	22834463029

Source: own elaboration.
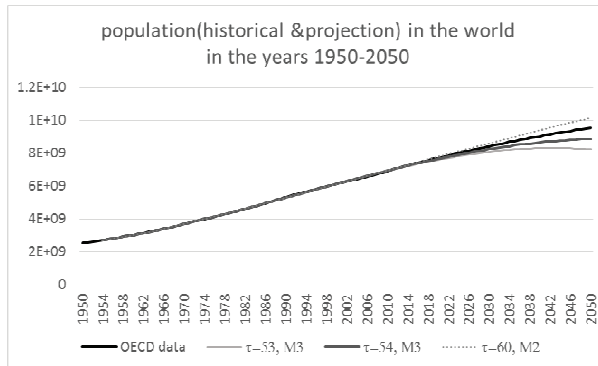
Figure 6. Comparison between the M2, M3 models and the world data for projection
available at http:/stats.oecd.org, 2016

Source: own elaboration.

The comparison between the models obtained for $\tau^* = 53$, $\tau = 54$ and for $\tau = N - 4 = 60$ with parameters presented in table 3 and the world data for projection[4]

---

[4] Available at http://stats.oecd.org, 2016.

is shown graphically in figure 5. The forecast obtained by M3 models satisfies better the statement "There must be an upper limit on the earth's life support capabilities, and therefore the population cannot grow without bound" (Robertson et al., 1961) than OECD model. Our model satisfies another statement "A model predicts that the world's population will stop growing in 2050"[5].

4. FINAL REMARKS

a) The difference equations method provides an easy programmable way to choose the model of the population in particular countries in the world as well as for the world population. The type of the model (see eq. (8), (9) and (10)) is established and depending on the values of parameters of characteristic equation (7) of equation (6).

b) Despite the statement that population forecasting by fitting mathematical curves is notably unreliable because it ignores so many important factors of demography (Dorn, 1962), the model of population obtaining by the difference equations method provides remarkably good fit with nearly all available data. For nearly all analyzed countries (40 out of 42), the $\delta_p\left(\tau^*\right)$ is less than 5%, for the world $\delta_p\left(\tau^*\right) < 0.6\%$.

c) The M1 model for $\left(\tau^*\right)$ (see eq. (8) and table 2) was obtained only for 4 countries. It is not a surprise that the population cannot be modelled by the model of t that estimates the number of people without upper limit (Robertson et al., 1961)[6].

REFERENCES

Austin A. L., Brewer J. W., (1971–1972), World Population Growth and Related Technical Problems, *Technological Forecasting and Social Change,* 3, 23–49.
Bard Y., (1974), *Nonlinear Parameter Estimation*, Academic Press, New York.
Chen J., Wu H., (2008), Efficient Local Estimation for Time-varying Coefficients in Deterministic Dynamic Models with Applications to HIV-1 Dynamics, *Journal of the American Statistical Association*, 103, 369–384.
Dorn H. F., (1962), World Population Growth: An International Dilemma, *Science*, 135, 283–290.
Hemker P. W., (1972), Numerical Methods for Differential Equations in System Simulation and in Parameter Estimation, in: Hemker H. C., Hess B., (eds.), *Analysis and Simulation of Biochemical Systems*, North Holland Publ. Comp, 59–80.
Huang Y., (2010), A Bayesian Approach in Differential Equation Dynamic Models Incorporating Clinical Factors and Covariates. *Journal of Applied Statistics*, 37 (2), 181–199.
Huang Y., Liu D., Wu H., (2006), Hierarchical Bayesian Methods for Estimation of Parameters in a Longitudinal HIV Dynamic System, *Biometrics*, 62 (2), 413–23.
Koźniewska I., (1972), *Równania rekurencyjne,* PWN, Warszawa.

---

[5] http://www.sciencedaily.com/releases/2013/04/130404072923.htm, 2013.
[6] http://www.sciencedaily.com/releases/2013/04/130404072923.htm, 2013.

Lanczos C., (1964), *Applied Analysis,* Prentice Hall. Inc., USA.

Li L., Brown M. B., Lee K. H., Gupta S., (2002), Estimation and Inference for a Spline-Enhanced Population Pharmacokinetic Model, *Biometrics*, 58 (3), 601–611.

Li Z., Osborne M. R., Prvan T., (2005), Parameter Estimation of Ordinary Differential Equations, *IMA Journal of Numerical Analysis*, 25 (2), 264–285.

Liang H., Wu H., (2008), Parameter Estimation for Differential Equation Models Using a Framework of Measurement Error in Regression Models, *Journal of the American Statistical Association*, 103 (484), 1570–1583.

Malthus T. R., (1798), *An Essay on the Principal of Population*, J. Johnson, in St. Paul's Church-Yard: London.

Miao H., Dykes C., Demeter L. M., Wu H., (2009), Differential Equation Modeling of HIV Viral Fitness Experiments: Model Identification, Model Selection, and Multimodel Inference, *Biometrics*, 65 (1), 292–300.

Murray J. D., (1989), *Mathematical Biology*, 19 of Biomathematics Texts, Springer: Berlin.

Nowak E., (2006), *Zarys metod ekonometrii. Zbiór zadań,* PWN, Warszawa.

Ogunnaike B. A., Ray W. H., (1994), *Process Dynamics, Modeling, And Control*, Series: Topics in Chemical Engineering, Oxford University Press, New York.

Pearl R., Reed L. J., (1924), The Growth of Human Population, in: Pearl R., (ed.), *Studies in Human Biology*, Wiliam and Wilkins: Baltimore, 584–637.

Poyton A. A., Varziri M. S., McAuley K. B., McLellan P. J., Ramsay J. O., (2006), Parameter Estimation in Continuous-Time Dynamic Models Using Principal Differential Analysis, *Computer and Chemical Engineering*, 30 (4), 698–708.

Ramsay J. O., Hooker G., Campbell D., Cao J., (2007), Parameter Estimation for Differential Equations: A Generalized Smoothing Approach (with Discussions), *Journal of the Royal Statistical Society*: Series B, 69 (5), 741–796.

Rao C. R., (1982), *Modele liniowe statystyki matematycznej*, PWN, Warszawa.

Robertson J. S., Bond V. P., Cronkite E. P., Hutton W. E., Howland W. E., Shinbrot M., von Foerster H., Mora P. M., Amiot L. W., (1961), Doomsday, *Science*, 133, 936–946.

Rzymowski W., Surowiec A., (2012), Method of Parameters Estimation of Pseudologistic Model, in: Zieliński Z. E., (ed.), *Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne*, 2, WSH, Kielce, 256–265.

Serrin J., (1975), Is 'Doomsday' on target? (Letter), *Science*, 189, 86–88.

Sierpiński W. F., (1946), *Zasady algebry wyższej,* PWN, Warszawa–Wrocław.

Smith D. A., (1977), Human Population Growth: Stability or Explosion? *Mathematics Magazine*, 50 (4), 186–197.

Varah J. M., (1982), A Spline Least Squares Method for Numerical Parameter Estimation in Differential Equations, *SIAM Journal on Scientific and Statistical Computing*, 3 (1), 28–46.

Verhulst P. F., (1838), Notice sur la loi que la population suit dans son d'accroissement. *Correspondance Mathématique et Physique Publiée par A. Quételet*, 10, 113–121.

von Foerster H., Mora P. M., Amiot L. W., (1960), Doomsday: Friday, 13 November, A. D. 2026, *Science*, 132 (3436), 1291–1295.

MODELOWANIE LICZBY LUDNOŚCI ZA POMOCĄ RÓWNAŃ RÓŻNICOWYCH

Streszczenie

W tym artykule przedstawiono metodę równań różnicowych czwartego rzędu do doboru postaci modelu ekonometrycznego. Zastosowanie zaproponowanej metody pozwala uzyskać model dobrze dopasowany do danych empirycznych o małym maksymalnym błędzie względnym.

Metoda równań różnicowych czwartego rzędu w artykule została wykorzystana do modelowania liczby ludności świata jak również do modelowania liczby ludności w wybranych krajach świata.

Słowa kluczowe: równania różnicowe, modele nieliniowe, estymacja parametrów strukturalnych, maksymalny błąd względny modelu, demografia

MODELLING POPULATION GROWTH
WITH SECOND ORDER DIFFERENTIAL EQUATION METHOD

Abstract

In this paper, we present a new method of the econometric model construction: the difference equation method. We illustrate the proposed approach using an application example from human population dynamic study. We find out that proposed method is very useful to find one of the three forms of proposed models of human population satisfying the small maximal relative errors. The maximal relative error is a measure to verify the model of human population.

The proposed method is tested for all available data referring to the human population in the OECD countries as well as in selected non-OECD countries.

Keywords: difference equations, nonlinear models, parameter estimation, relative error, demography