POLSKA AKADEMIA NAUK
KOMITET STATYSTYKI I EKONOMETRII

PRZEGLĄD STATYSTYCZNY

STATISTICAL REVIEW

TOM 63

3

2016

WARSZAWA 2016

WYDAWCA

Komitet Statystyki i Ekonometrii Polskiej Akademii Nauk

Wydanie publikacji dofinansowane przez Ministra Nauki i Szkolnictwa Wyższego

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

Nakład 140 egz.

Przygotowanie do druku:
Dom Wydawniczy ELIPSA
ul. Inflancka 15/198, 00-189 Warszawa
tel./fax 22 635 03 01, 22 635 17 85
e-mail: elipsa@elipsa.pl, www.elipsa.pl

Druk:
Wrocławska Drukarnia Naukowa PAN Sp. z o.o.
im. Stanisława Kulczyńskiego
53-505 Wrocław, ul. Lelewela 4

SPIS TREŚCI

CONTENTS

DOI: 10.5604/01.3001.0014.1195
PRZEGLĄD STATYSTYCZNY
R. LXIII – ZESZYT 3 – 2016

ZBIGNIEW ŚWITALSKI[1]

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

1. INTRODUCTION

In their well-known paper, Gale, Shapley (1962) modelled the process of assigning applicants to colleges, where the problem was to match applicants with colleges in some "optimal" way. Such kind of matching process can be treated as a market process, in which applicants are interpreted as "buyers", colleges – as "sellers", and the traded "goods" as seats in particular colleges.

During the last 50 years modelling the so-called markets with two-sided preferences using the idea of Gale and Shapley became very popular. Different kinds of such markets (for example labor markets or auction markets) are described, e.g., by Roth, Sotomayor (1992).

In the simplest version of a market with two-sided preferences we have two disjoint finite sets of buyers and sellers. The buyers have preferences over the sellers, and the sellers have preferences over the buyers (both represented by linear orders). We assume also that each seller owns a certain number of identical objects which he wants to sell, and each buyer wants to buy at most one object (this resembles the "college admissions" market – traditionally called many-to-one market).

In the last 10 years very general market models based on Gale-Shapley theory have been built, for example contract theory of Hatfield, Milgrom (2005). Models with contracts are very intensively used in the modern theory of markets with indivisible goods. Preferences in these models are often represented by the so-called choice functions (see, e.g., Hatfield et al., 2013).

The main theoretical tool used in the Gale-Shapley theory (and hence in the theory of markets with two-sided preferences and in the contract theory) is the notion of stable matching. A matching u assigning buyers to sellers is stable if there is no pair (b, s) such that buyer b and seller s would have simultaneously any incentive to change the matching u.

[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.
[2] The research was financially supported by the Polish National Science Centre (NCN) – grant DEC-2011/01/B/HS4/00812.

It can be proved (see Roth, Sotomayor, 1992) that stable matchings in the GS model form the set of core allocations for the respective cooperative game, i.e. allocations such that no coalition of agents can improve the situation of all members of the coalition. Hence, a stable matching is sometimes interpreted as a kind of cooperative equilibrium for the respective market model (see, e.g., Sotomayor, 2007).

But considering any market model, we can also ask about the notion of competitive market equilibrium (in the sense of Walras). The notion of competitive equilibrium is one of the fundamental notions in economic theory, so for a market model with two-sided preferences it is quite natural to ask two questions:

1. Is there any reasonable way to define competitive equilibrium for such a model?
2. If a competitive equilibrium is defined, what are the relationships between the notion of such equilibrium and the notion of stable matching (cooperative equilibrium) for such a model?

The solution of the problem of relationships between competitive equilibria and stable matchings for such kind of markets can help to solve the problems of existence of such equilibria and to find methods of looking for them. For example, in the simplest version of GS theory the problem of existence of stable matchings and the problem how to find them is fully solved (see Gale, Shapley, 1962). Hence, proving the equivalence between stability and competitive equilibria for the simplest version of GS model (see, e.g., theorem 1 below) means automatically proving the existence of such equilibria and gives the method of finding them.

For traditional, continuous models of market equilibrium it can be often proved that the competitive equilibrium is in the core (see, e.g., Moore, 2007). There are also many discrete matching models, different from the GS model, for which exact relationships between competitive and cooperative equilibria are established. These are models related mainly to the so-called "assignment games" in the sense of Shapley, Shubik (1971/72).

The main difference between the GS models and the SS (= Shapley-Shubik) models is that in the GS models buyers' preferences do not depend on prices (they are exogenously given, as in the neoclassical consumer theory), contrary to the SS models, in which buyers' preferences are determined by quasi-linear utilities depending on prices.

In the simplest case of SS model, it can be proved that the core allocations (which are optimal assignments in this case) are exactly competitive equilibria allocations (see, e.g., Shapley, Shubik, 1971/72; Shoham, Leyton-Brown, 2009, theorem 2.3.5, p. 31). There are many other similar results for different variants of the SS model (see, e.g., Camina, 2006; Sotomayor, 2007).

In the contract theory of Hatfield et al. (2013) the result about strict relationship between stable matchings and competitive equilibria can also be proved (the utility function used by Hatfield et al. (2013) is quasi-linear similarly as in the SS models).

As to the GS models, there was (to the best of our knowledge) no research on this topic until the papers of Świtalski (2008, 2010) and Azevedo, Leshno (2011) appeared. Perhaps the reason for that was the skeptical view of Shapley and Scarf (1974, p. 35) on the problem of introducing the concept of market equilibrium for the GS model:

"*It does not appear to be possible to set up a conventional market for this model* [= GS *model*] *in such a way that a competitive price equilibrium will exist and lead to an allocation in the core*".

In the papers of Świtalski (2008, 2010) different kinds of generalized equilibria (the so-called stable equilibria, order equilibria and boundary equilibria) for the GS models were defined. In Azevedo, Leshno (2011, p. 18) the so-called "supply and demand lemma" was proved. In this lemma it was shown that stable matchings in the college admissions problem can be characterized with the help of families of "cutoffs" (cutoff for a given college is the score of marginal accepted student for this college). If we interpret cutoffs as some kind of prices in the respective market model, then their result can be reformulated in the following way: "a matching u is stable if and only if it is a competitive price equilibrium allocation" (see theorem 1 below).

In the paper of Świtalski (2015) a generalization of Azevedo and Leshno's result was proved (the model described there is one-to-one with preferences of the agents represented by weak orders).

In the presented paper we prove some far-reaching generalization of the result of Azevedo and Leshno. We consider a certain variant of a many-to-many market model, based on GS model, with choice functions representing preferences of buyers and weak orders representing preferences of sellers (and with quotas for all agents). Equilibria in our model (we call them order equilibria, see Świtalski, 2010) are defined by general conditions represented by families of subsets $\{W(s)\}$ (indexed by sellers s) which can be treated as a generalization of "cutoff" or price conditions in the Azevedo, Leshno's (2011) model. For such models we study relationships between equilibria and stability under different assumptions about choice functions. Using the results of Alkan, Gale (2003) we also prove the result on existence of order equilibria for our model (and show that in some cases the so-called strongly order equilibria may not exist).

The simplest one-to-one version of our model (with equilibria defined as usual price equilibria) can be identified with the model of matching markets with budget constraints described by Chen et al. (2014) – with the assumption that the utility functions of the buyers are constant (do not depend on prices). Yet, the stability concept used by Chen et al. differs from the standard one and hence their result (2014, theorem 3.1) cannot be treated as a special case of our results.

Our results cannot also be treated as a special case of the results for trading neworks obtained by Hatfield et al. (2013). Preferences of buyers in their theory depend on prices (they use quasi-linear utility functions) and in our paper we treat buyers' preferences as exogenously given as in the standard GS model.

The paper is organized as follows. In section 2 we describe the simplest one-to-one version of our model and translate Azevedo and Leshno's result in terms of equilibrium theory (theorem 1). In section 3 we construct general model and state the main results of the paper (lemmas 1 and 2, theorem 2). In section 4 we consider the problem of existence of equilibria for our model (theorem 3 and example 1).

2. THE SIMPLEST ONE-TO-ONE MODEL

The model of Gale, Shapley (1962) is concerned with the problem of "optimization" of the process of admitting applicants to colleges. Applicants have preferences over colleges and colleges have preferences over applicants. Gale and Shapley find a method of assigning applicants to colleges which gives stable matching u, i.e. a matching for which there is no pair (b, s) such that the applicant b prefers college s to $u(b)$ ($u(b)$ is the college to which b is admitted) and s prefers b to an applicant admitted to s.

A stable matching may be treated as some kind of equilibrium state in the college admissions market. Yet, it is defined in a completely different way than the classical notion of competitive, Walrasian equilibrium.

We can treat applicants as "buyers" and colleges as "sellers" in the market, but to define the competitive equilibrium we need prices on the sellers' side of the market and budget constraints on the buyers' side.

We can observe that the role of prices can be played by scores with the help of which applicants are classified in many admission systems (see, e.g., Biro, Kiselgof, 2013). In such a system each applicant has a number of scores achieved in different disciplines (maths, physics, biology and so on) and each college ranks students according to the sum of scores achieved in the disciplines which are taken into account by this college.

Let $p(s)$ be the sum of scores of the worst (marginal) applicant admitted to college s (under some assignment u), and let $r(b, s)$ be the sum of scores applicant b achieves in the disciplines required by the college s. Then the inequality $r(b, s) \geq p(s)$ is a necessary and sufficient condition guaranteeing b to be admitted to college s under u ($p(s)$ can be treated as a score-limit in the sense of Biro, Kiselgof (2013) or as a "cutoff" in the sense of Azevedo, Leshno (2011). Hence we can think of $p(s)$ as a kind of price of a seat in college s, and of $r(b, s) \geq p(s)$ as a kind of budget constraint for applicant b when he is interested in being admitted to s (applicant b can be admitted to any college s for which $r(b, s) \geq p(s)$ is satisfied).

Observe that "budget constraints" in the college admissions market depend on both the "buyers"-side and the "sellers"-side of the market.

Given prices and budget constraints, we can now easily define competitive equilibrium in the Walras sense.

We start from the simplest one-to-one model (resembling the marriage model of Gale, Shapley, 1962). Using this model, we explain the main idea of the paper. Let B be a finite n-element set of buyers and S – a finite n-element set of sellers. Each seller owns exactly one indivisible object which he wants to sell (objects can be houses, cars, horses, paintings and so on). Each buyer wants to buy exactly one object.

We identify sellers with objects which they own, hence the phrase "object s" should be understood as a shortened version of "object owned by seller s". We assume that buyers have preferences over sellers (equivalently – over the objects) and sellers have preferences over buyers. Preferences are represented by strict linear orders, i.e. to each agent (buyer b or seller s) an ordered list of agents from the opposite set, indicating preferences of b or s, is assigned (there are no indifferences). We use notation $b >_s c$ meaning that buyer b is better than buyer c for seller s, and $s >_b t$ meaning that seller s is better than seller t for buyer b.

For each buyer b and each seller s we define a reservation price $r(b, s)$ interpreted as maximal price that b is willing to pay for object s (in the college market the role of reservation price is played by the sum of scores applicant b achieves in the disciplines required by college s).

We assume that preferences of the sellers are determined by the reservation prices of the buyers, i.e. for any buyers b and c and seller s we have

$$b >_s c \iff r(b, s) > r(c, s) \tag{1}$$

(to avoid indifferences, we assume here that, for a given s, the numbers $r(b, s)$ are different).

Formula (1) is obvious for the college market (in the score system college s ranks applicants according to the sums of scores). For general markets formula (1) says that seller s prefers a buyer who can pay more over a buyer who can pay less for the object owned by s (hence s can sell this object to b at a higher price than to c).

A *matching* of buyers with sellers (or an allocation of objects among buyers) is a set of pairs (b, s) such that each agent (b or s) occurs in exactly one pair. If u is a matching and $(b, s) \in u$, then we write also $u(b) = s$ or $u(s) = b$.

A matching u is *stable* if there is no pair (b, s) satisfying the condition:

$$s >_b u(b) \quad \text{and} \quad b >_s u(s). \tag{2}$$

A pair satisfying (2) is called a *blocking pair* for u.

Assume now that each seller announces a price $p(s)$ for the object he owns. A sequence of prices $p(s)$ ($s \in S$) is called the *price vector* p. Prices p are called equilibrium prices if there is a matching u such that each buyer b gets object s (i.e. $u(b) = s$) which is the best object for him among all objects satisfying the inequality $r(b, s) \geq p(s)$ (i.e. all feasible objects for b). Such a matching is called equilibrium allocation associated with p. The following result is a kind of reformulation of Azevedo and Leshno's result (2011, see also Świtalski, 2015).

□ **Theorem 1.** A matching u is stable if and only if it is an equilibrium allocation associated with some price vector p.

In the next section we study a market model which generalizes the presented above one-to-one model. Namely, we consider a many-to-many model in which preferences of buyers are represented by choice functions and preferences of sellers are weak orders. The notion of equilibrium which we use is very general. The budget constraints $r(b, s) \geq p(s)$ are replaced by certain conditions defined by families $\{W(s)\}$ of subsets of the set B (the set of buyers). For such defined models we study relationships between stable matchings and equilibria allocations (lemmas 1 and 2). Theorem 2 formulated at the end of section 3 includes theorem 1 as a special case.

3. THE GENERAL MODEL

Gale, Shapley (1962) defined a college admission model which is traditionally called many-to-one model (there can be many applicants that a fixed college wants to admit, but each applicant wants to be admitted to only one college). Many authors (e.g., Echenique, Oviedo, 2006; Klaus, Walzl, 2009; Kominers, 2012), starting from Gale-Shapley model, described many-to-many market models, especially for different kinds of labor markets. For example, Echenique, Oviedo (2006) consider a market consisting of firms and consultants, where each firm wants to hire a set of consultants and each consultant wants to work for a set of firms. Other examples (mentioned by Echenique, Oviedo, 2006) are the markets for medical interns in the U.K. or teacher (university professor) markets in some countries (where teachers (professors) can work in more than one school (university)).

In our paper we also consider a many-to-many model. To start with the formal description of this model, we define two finite and non-empty sets: a set of buyers (e.g. firms) B and a set of sellers (e.g. consultants) S.

The symbol $B \times S$ denotes Cartesian product of B and S, i.e. the set of ordered pairs (b, s) such that $b \in B$ and $s \in S$.

For any relation $u \subset B \times S$ and for any $b \in B$, $s \in S$, we define the sets of "neighbouring" elements:

$$u(b) = \{s \in S\colon (b, s) \in u\}, \tag{3}$$

$$u(s) = \{b \in B\colon (b, s) \in u\}. \tag{4}$$

We assume that a non-empty set of acceptable pairs $F \subset B \times S$ is defined. A pair (b, s) belongs to F if buyer b is acceptable for s and seller s is acceptable for b. Hence, according to (3) and (4), the sets $F(b)$ and $F(s)$ can be defined. The set $F(b)$ can be interpreted as the set of acceptable sellers for buyer b, and $F(s)$ – as the set of acceptable buyers for seller s.

From the point of view of contract theory (Hatfield et al., 2013) acceptable pairs can be interpreted as possible transactions (trades) which can be realized in the market. In other words, $(b, s) \in F$ means that buyer b can sign a contract with seller s. A contract is signed if b and s agree to the conditions of the contract (e.g. price). 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 other buyers, but only one contract with a given buyer b).

In our model we introduce quotas for buyers and sellers. Let $q(b) \geq 1$ be the quota for b, which is a maximum number of contracts which b can sign with different sellers and $q(s) \geq 1$ – the quota for s, which is a maximum 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 the cardinality of a set A).

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

Preferences of buyers are represented by choice functions. Choice functions are a standard tool in economic and decision theory (see, e.g., Aizerman, Aleskerov, 1995; Aleskerov, Monjardet, 2002) and they are very often used for the models of markets with two-sided preferences, especially for the labor markets (see, e.g., Echenique, 2007; Klaus, Walzl, 2009; Hatfield et al., 2013). Defining a choice function means that we know what choice will be made by a decision maker when she is confronted with a given set of decision alternatives. Formally, a choice function is a mapping C from the family T of all subsets of a given set (set of all possible alternatives) to the same family, assigning a set $C(X) \subset X$ to every $X \in T$. The set $C(X)$ is interpreted as the set of elements chosen from X by a decision maker.

Usually, in the papers on many-to-many markets or contract theory (see, e.g., Echenique, Oviedo, 2006; Klaus, Walzl, 2009; Kominers, 2012), choice functions are generated by preferences over the subsets from the family T. In our paper we do not assume a priori that there is some order relation in T and that $C(X)$ is some "best" set (with respect to this order) in the family of all subsets of X.

In our model we assume that a choice function is defined for every feasible set $F(b)$ (for a given buyer b). Hence, 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 a certain set of feasible sellers X. Then her decision will be to choose the set $C(b, X)$ as the set of sellers, with whom she will sign a contract. Of course the number of such sellers should not exceed $q(b)$, hence we assume that:

$$\text{(i)} \quad C(b, X) = X, \qquad \text{if} \quad \# X < q(b),$$

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

(we consider here the so-called quota-filling choice functions in the sense of Alkan, Gale, 2003).

In what follows we consider the following properties of the function C:

The outcast property (see, e.g., Aizerman, Aleskerov, 1995, p. 20; Aleskerov, Monjardet, 2002, p. 39; Echenique (2007) defines an equivalent property called independence of irrelevant alternatives):

For every $b \in B$ and X, $Y \subset F(b)$ we have

$$Y \subset X \setminus C(b, X) \;\Rightarrow\; C(b, X \setminus Y) = C(b, X). \tag{5}$$

The outcast property means that if we delete a set, consisting of not chosen elements, from the set X, then the resulting choice will remain the same.

We note that the outcast property implies the following properties:

$$C(b, C(b, X)) = C(b, X), \tag{6}$$

$$C(b, C(b, X) \cup \{s\}) = C(b, X) \tag{7}$$

for any $s \in X$ (to prove (6) we take $Y = X \setminus C(b, X)$ in (5), and to prove (7) we take $Y = X \setminus (C(b, X) \cup \{s\})$ in (5)).

So, we can add an element s to the set of chosen elements and this operation does not change the set of chosen elements.

The heritage property (see, e.g., Aizerman, Aleskerov, 1995, p. 18; Aleskerov, Monjardet, 2002, p. 36; in the matching literature this kind of property is sometimes called substitutability, see, e.g., Echenique, 2007):

For every $b \in B$ and X, $Y \subset F(b)$ we have

$$Y \subset X \Rightarrow Y \cap C(b, X) \subset C(b, Y). \tag{8}$$

The heritage property means that any element (seller) chosen from X, which belongs to a smaller set Y, should also be chosen from Y.

Choice functions satisfying both 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 a linear order (then $C(b, X)$ is the set of $q(b)$ best sellers in X (if $\# X \geq q(b)$)).

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 $s \in S$).

For a given generalized GS-model (B, S, F, C, P, q) we define now the notion of matching and (strongly) stable matching (definitions 1–5). We define matching as a set of pairs and it is easy to see that our definition is equivalent to the standard definition of Echenique, Oviedo (2006) (they define matching as some mapping from $B \cup S$ into the set of all subsets of $B \cup S$). Our definition of blocking pair is a combination of a standard definition for a many-to-one model (see Roth, Sotomayor, 1992,

p. 129) with the definition of Echenique, Oviedo (2006, p. 240) for a many-to-many model (for the buyers' side of the market). To unify the definitions we introduce the non-standard notion of "improving the situation of an agent" (definitions 2 and 3). Strongly stable matchings are defined similarly as in Manlove (2002).

Definition 1. A relation $u \subset B \times S$ is a matching if

(i) $u \subset F$,

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

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

A matching u can be interpreted as a set of actual contracts signed by agents from the sets B and S (transactions realized in the market). Obviously, according to (i), such contracts should be taken from F – the set of all possible (potential) contracts.

Definition 2. 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 $s >_{\mathrm{b}} u(b)$) if $s \in C\,(b, u(b) \cup \{s\})$.

Definition 3. 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 $b >_{\mathrm{s}} u(s)$ or $b \ge_{\mathrm{s}} u(s)$ respectively) if at least one of the following conditions holds:

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

(ii) $\exists\, c \in u(s), \quad b >_{\mathrm{s}} c \quad (b \ge_{\mathrm{s}} c)$.

Definition 4. 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) $s >_{\mathrm{b}} u(b)$,

(iii) $b >_s u(s) \quad (b \ge_{\mathrm{s}} u(s))$.

Definition 5. A matching $u \subset B \times S$ is stable (strongly stable) if there are no blocking pairs (weakly blocking pairs) for u.

Now we can define a generalized equilibrium in a generalized GS-model (B, S, F, C, P, q). In the simplest one-to-one model the set of feasible sellers (objects) for a given buyer b was defined with the help of inequality

$$r(b, s) \ge p(s). \tag{9}$$

Inequality (9) is a necessary and sufficient condition for buyer b to obtain object s (or to sign a contract with seller s). There can be markets in which sellers can state some other conditions for buyers needed to sign contracts (for example there can be some law requirements needed to buy some products). In general, we can assume that the conditions required by seller s determine a set of buyers $W(s) \subset F(s)$ such that being in the set $W(s)$ is for buyer b a necessary and sufficient condition to sign a contract with s (for example, the price condition (9) determines the set $W(s) = \{b \in F(s): r(b, s) \geq p(s)\}$). The sets $F(s)$ are fixed, but the sets $W(s)$ can vary and we can define equilibrium in the generalized models with respect to the families $W = \{W(s)\}$ $(s \in S)$. A family $W = \{W(s)\}$ $(s \in S)$ is called a system of conditions. In definitions 6 and 7 (below) we define generalized equilibria without imposing any special assumptions on the sets $W(s)$.

First, we define the set of feasible sellers under the system $W = \{W(s)\}$ for buyer b as:

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

Obviously, $F(W, b) \subset F(b)$ and hence we can define the set of the "best" sellers (contracts) for b under the system W as

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

Buyer b demands objects from the set $M(W, b)$ (wants to sign contracts with the sellers from $M(W, b)$) and hence the demand set for seller s (under the conditions W) can be defined as:

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

Demand set $D(W, s)$ is the set of buyers for whom s is among the "best" sellers. It is easy to see (from the above definitions) that $D(W, s) \subset W(s)$.

Definition 6. A system of conditions $W = \{W(s)\}$ is an *equilibrium system* 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 $\# D(W, s) < q(s)$.

Inequality (i) guarantees that, under the conditions W, each buyer can sign all the contracts which are best for her without exceeding the supply limits $q(s)$. Condition (ii) states that it is not possible to weaken the conditions $\{W(s)\}$ in order to increase the demand of the buyers in the situation when supply limits are not reached (it is a generalization of the standard condition of zeroing the prices of unassigned goods for one-to-one matching models – such condition guarantees that we cannot decrease prices of these goods to increase the demand for them, see e.g., Mishra, Talman, 2010; Chen at al., 2014; Świtalski, 2015).

Let $W = \{W(s)\}$ be an equilibrium system. We define a matching associated with W as:

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

Definition 7. An *equilibrium* (in 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 say that $u = u(W)$ is an equilibrium allocation associated with W.

To state our results about relationships between stability and equilibria we need to restrict the notion of equilibrium to the so-called order equilibrium which is an equilibrium for which the sets $W(s)$ are "compatible" with the order relations $>_s$.

Definition 8. A system of conditions $W = \{W(s)$ is compatible (strongly compatible) with the sellers' preferences if

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

Hence, the system W is compatible with the sellers' preferences if all the buyers who are preferred over a certain buyer satisfying $W(s)$, also satisfy $W(s)$ (in other words the conditions which determine the set of "possible" buyers are ordinal). Observe that strong compatibility implies compatibility. It is also easy to see that the price conditions (9) in a one-to-one model with preferences of the sellers defined by (1) are compatible with the sellers' preferences.

Definition 9. An equilibrium (u, W) in 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.

We note here that the notion of order equilibrium needs representation of the sellers' preferences by weak orders. This is the reason why in our model we have an asymmetry in the representation of preferences (choice functions for the preferences of the buyers and weak orders for the preferences of the sellers).

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

In the next two lemmas we show that if C satisfies both the outcast and heritage properties, then a matching $u \subset B \times S$ is stable (strongly stable) iff it is an order (strongly order) equilibrium allocation (i.e. iff there exists a system of conditions W such that (u, W) is an order (strongly order) equilibrium allocation). The final result is formulated as theorem 2 below.

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

Proof. Assume that u is not stable (not strongly stable), so there is a pair $(b, s) \in F \setminus u$ such that $s >_b u(b)$ and $b >_s u(s)$ ($b \geq_s u(s)$). First we prove that $s \in F(W, b)$. Consider two cases:

1. $\# u(s) = q(s)$. Hence $b >_s c$ ($b \geq_s c$) for some $c \in u(s) = D(W, s) \subset W(s)$, and so $c \in W(s)$.
   Thus, by definition 8, $b \in W(s)$, and so $s \in F(W, b)$.
2. $\# u(s) < q(s)$. By definition 6 we have $W(s) = F(s)$, and hence $b \in W(s)$ (because $(b, s) \in F \setminus u$ implies $b \in F(s)$). Thus $s \in F(W, b)$.

We also have (by $s >_b u(b)$):

$$s \in C\ (b, u(b) \cup \{s\}) \tag{10}$$

and, by the definition of $u(W)$:

$$u(b) = \{s\colon (b, s) \in u\} = \{s\colon b \in D(W, s)\} = \{s\colon b \in F(s) \wedge s \in M(W, b)\}$$
$$= M(W, b) = C(b, F(W, b)).$$

Hence, by (10) and (7) (we take $X = F(W, b)$), we have:

$$s \in C(b, C(b, F(W, b)) \cup \{s\}) = C(b, F(W, b)) = u(b).$$

This fact contradicts the assumption $(b, s) \in F \setminus u$ and implies that u is stable (strongly stable). ■

Lemma 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.

Proof. For any matching $u \subset B \times S$ we can define a system of conditions:

$$W(s) = \begin{cases} u(s) \cup \{b \in F(s)\colon \exists\ c \in u(s),\ b >_{\mathrm{s}} c\ (b \geq_{\mathrm{s}} c)\}, & \text{if } \#\, u(s) = q(s), \\ F(s), & \text{if } \#\, u(s) < q(s). \end{cases} \tag{11}$$

In the case when $\#\, u(s) = q(s)$, $W(s)$ consists of all the buyers matched with s and all the buyers which are better (not worse) than at least one buyer matched with s. It is easy to see that $W(s)$ are compatible (strongly compatible) with the sellers' preferences.

To prove that (u, W) is an order (strongly order) equilibrium, we should show (by definitions 6 and 7) that $\#\, D(W, s) \leq q(s)$ for all $s \in S$, $\#\, D(W, s) < q(s)$ implies $W(s) = F(s)$ and that $u(W) = u$. Observe that $u(W) = u$ is equivalent to $u(W)(s) = u(s)$ for all $s \in S$ and $u(W)(s) = D(W, s)$. Hence, to show that (u, W) is an order (strongly order) equilibrium, it is sufficient to show that $u(s) = D(W, s)$ for all $s \in S$ (by (11) and by condition (iii) in the definition 1).

1. First we prove that $u(s) \subset D(W, s)$. Let $b \in u(s)$. We want to prove that $b \in D(W, s)$. By the definition of $W(s)$ (11), $b \in W(s)$. Hence, by the definition of $F(W, b)$, $s \in F(W, b)$. Obviously, $b \in F(s)$ ($u \subset F$). Thus, to state that $b \in D(W, s)$, it suffices to show that $s \in M(W, b) = C(b, F(W, b))$.

Assume that $s \notin C(b, F(W, b))$. Hence $\#\, F(W, b) > q(b)$ (if $\#\, F(W, b) \leq q(b)$, then $\#\, C(b, F(W, b)) = \#\, F(W, b)$, and so $s \in F(W, b)$ would imply $s \in C(b, F(W, b))$).

Summing up, we have the following facts:

(i) $u(b), M(W, b) \subset F(W, b)$,
(ii) $\#\, M(W, b) = q(b)$,
(iii) $\#\, F(W, b) > q(b)$,
(iv) $s \in u(b)$, $s \notin M(W, b)$,
(v) $\#\, u(b) \leq q(b)$ (see (ii), definition 1).

The facts (i)−(v) imply the existence of $t \in M(W, b)$ such that $t \notin u(b)$. Obviously, $u(b) \cup \{t\} \subset F(W, b)$ and hence, by heritage property (8), we have:

$$(u(b) \cup \{t\}) \cap C(b, F(W, b)) \subset C(b, u(b) \cup \{t\}).$$

Hence, because $t \in M(W, b) = C(b, F(W, b))$, we have $t \in C(b, u(b) \cup \{t\})$. Thus $t >_{\mathrm{b}} u(b)$.

Now we will prove that $b >_{\mathrm{t}} u(t)$ ($b \geq_{\mathrm{t}} u(t)$). Consider two cases:

1. $\#\, u(t) = q(t)$. We have $t \in M(W, b) \subset F(W, b)$, hence $b \in W(t)$. We also have $b \notin u(t)$ (because $t \notin u(b)$), hence (by the definition of $W(t)$) there exists $c \in u(t)$ such that $b >_{\mathrm{t}} c$ ($b \geq_{\mathrm{t}} c$). Thus $b >_{\mathrm{t}} u(t)$ ($b \geq_{\mathrm{t}} u(t)$).
2. $\#\, u(t) < q(t)$. Then $b >_{\mathrm{t}} u(t)$ ($b \geq_{\mathrm{t}} u(t)$) by definition 3.

Hence we have obtained $t >_b u(b)$ and $b >_t u(t)$ ($b \geq_t u(t)$). It is easy to see that $(b, t) \in F \setminus u$ (because $t \in F(W, b) \subset F(b)$ and $t \notin u(b)$). Thus, by definition 4, (b, t) is a (weakly) blocking pair for u, a contradiction to the stability of u.

2. Now we prove that $D(W, s) \subset u(s)$ for all $s \in S$. Let $b \in D(W, s)$. Then $b \in F(s)$ (hence $s \in F(b)$) and $s \in M(W, b)$. Assume that $b \notin u(s)$ (hence $s \notin u(b)$). Thus $(b, s) \in F \setminus u$. We will prove that $b >_s u(s)$ ($b \geq_s u(s)$) and $s >_b u(b)$, thus showing that (b, s) is a (weakly) blocking pair for u, a contradiction to the stability of u.

(i) To show that $b >_s u(s)$ ($b \geq_s u(s)$), we can use the same reasoning as in the second part of the proof in p. 1 (t should be changed by s).

(ii) To show that $s >_b u(b)$ observe that $u(b) \subset M(W, b)$ (if $t \in u(b)$, then $b \in u(t)$, and, by proof in p. 1, $u(t) \subset D(W, t)$, hence $b \in D(W, t)$, and so $t \in M(W, b)$). By the definition of $M(W, b)$, $\# M(W, b) \leq q(b)$. Hence, by $s \notin u(b)$ and $u(b) \subset M(W, b)$, we have $\# u(b) < q(b)$ and so, by definition 2 and the definition of choice function, $s >_b u(b)$. ■

From lemmas 1 and 2 we obtain the following result.

□ **Theorem 2.** If $M = (B, S, F, C, P, q)$ is a generalized GS-model such that C is a Plott choice function, then $u \subset B \times S$ is stable (strongly stable) if and only if it is an order (strongly order) equilibrium allocation.

4. THE EXISTENCE OF EQUILIBRIA

Let $M = (B, S, F, C, P, q)$ be a generalized GS-model. We can ask about the existence of an order (strongly order) equilibrium for such a model. Using theorem 2 and some existence results from GS theory we can prove the following

□ **Theorem 3.** If $M = (B, S, F, C, P, q)$ is a generalized GS-model such that C is a Plott choice function, then there exists an order equilibrium (u, W) for M.

Proof. For any preference relation $P(s)$ (weak preference order for a seller s in the model M) take a linear extension $L(s)$ (i.e. a linear order $L(s)$ such that $P(s) \subset L(s)$). Linear orders $L(s)$ determine, in an obvious way, Plott choice functions for the sellers (we take, for any $X \subset F(s)$, the set of $q(s)$ best buyers in the set X, or the set X, if $\# X \leq q(s)$). Having Plott choice functions on both sides of the market, we can use Alkan and Gale theory (2003) to deduce that there exists a stable matching u in the model $N = (B, S, F, C, L, q)$ (Alkan, Gale, 2003, theorem 1, p. 298 – the properties of consistency and persistency used by Alkan and Gale are equivalent to the outcast and heritage properties, respectively). It is easy to see that u is also stable for the model M. Hence, by lemma 2, we can find a system of conditions W, compatible with the seller' preferences P, such that (u, W) is an order equilibrium, and this completes the proof. ■

Unfortunately, similar result cannot be proved for strongly order equilibria. Namely, there can be generalized GS models, for which there are no strongly stable matchings, and hence, by theorem 2, no strongly order equilibria. The following example shows such a situation.

Example 1. Let $B = \{b, c\}$, $S = \{s\}$, $F = \{(b, s), (c, s)\}$, $C(b, \{s\}) = \{s\}$, $C(c, \{s\}) = \{s\}$. Let s be indifferent between b and c and let all quotas be equal to 1. The only possible matchings for M (according to definition 1) are $u = \{(b, s)\}$, $v = \{(c, s)\}$ and empty matching $\varnothing$. It is easy to see that (c, s) is a weakly blocking pair for u ($s >_c u(c)$, because $s \in C(c, u(c) \cup \{s\}) = C(c, \varnothing \cup \{s\}) = C(c, \{s\}) = \{s\}$, and $c \geq_s u(s)$, because $c \geq_s b$ and $b \in u(s) = \{b\}$), (b, s) is a weakly blocking pair for v, and both (c, s) and (b, s) are weakly blocking pairs for $\varnothing$. Hence we have no strongly stable matching for the model $M = (B, S, F, C, P, q)$ (although, as it is also easy to see, both u and v are stable for M).

5. CONCLUDING REMARKS

In our paper we have investigated relationships between the concept of stability and the concept of generalized competitive equilibrium (called here order equilibrium) for some variant of a many-to-many Gale-Shapley market model. The results we have obtained can help to prove existence results for equilibria for such kind of models. In the existing literature there are many similar results for market models of the Shapley-Shubik type, but little has been proved till now for models of the GS type (with preferences not depending on prices). Hence our paper fills a gap in this area.

Preferences in our model are represented by choice functions (for the buyers) and weak orders (for the sellers), so there is an asymmetry here. The reason is that to define an order equilibrium (which is a generalization of price equilibrium) it is necessary to have an ordering relation on the sellers' side of the market. An interesting question could be: can we avoid such an asymmetry by introducing a concept of equilibrium which would not depend on special representation of preferences by weak orders? Another question is the possibility of using in our model choice functions without quota restrictions (as in the model of Echenique, Oviedo, 2006). It could be also interesting to study in detail relationships between stability and price equilibria (a special case of order equilibria) for many-to-many GS models. We leave these problems for further research.

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.

Azevedo E. M., Leshno J. D., (2011), A Supply and Demand Framework for Two-Sided Matching Markets, *working paper*, *www.aeaweb.org* (last access: 25.11.2015).

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, *arxiv.org*.

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

Hatfield J. W., Milgrom P. L., (2005), Matching with Contracts, *American Economic Review*, 95, 913–935.

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.

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.

Manlove D. F., (2002), The Structure of Stable Marriage with Indifference, *Discrete Applied Mathematics*, 122, 167–181.

Mishra D., Talman D., (2010), Characterization of the Walrasian Equilibria of the Assignment Model, *Journal of Mathematical Economics*, 46, 6–20.

Moore J. C., (2007), *General Equilibrium and Welfare Economics*, Springer, Berlin.

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

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

Shapley L. S., Scarf H., (1974), On Cores and Indivisibility, *Journal of Mathematical Economics*, 1, 23–37.

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

Shoham Y., Leyton-Brown K., (2009), *Multiagent Systems – Algorithmic, Game-Theoretic and Logical Foundations*, Cambridge University Press.

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*, *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 248, 222–232.

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

Streszczenie

W artykule zdefiniowano, dla pewnego wariantu modelu rynku Gale'a-Shapleya (typu „many-to--many"), pojęcie uogólnionej równowagi konkurencyjnej i pokazano że, przy odpowiednich założeniach, skojarzenia stabilne w tym modelu mogą być reprezentowane jako alokacje równowag konkurencyjnych (i vice versa). Przedstawione wyniki są daleko idącymi uogólnieniami „lematu o podaży i popycie" z pracy Azevedo, Leshno (2011) dotyczącego modelu rekrutacji kandydatów do szkół.

Wykorzystując wyniki Alkana, Gale'a (2003) udowodniono również twierdzenie o istnieniu uogólnionych równowag dla podanego modelu.

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

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

Abstract

We define, for some variant of a many-to-many market model of Gale-Shapley type, a concept of generalized competitive equilibrium and show that, under suitable conditions, stable matchings in such a model can be represented as competitive equilibria allocations (and vice versa). Our results are far-reaching generalizations of the "discrete supply and demand lemma" of Azevedo, Leshno (2011) for the college admissions market.

Using the results of Alkan, Gale (2003), we also prove a theorem on existence of generalized equilibria in our model.

Keywords: stable matching, Gale-Shapley theory, many-to-many model, competitive equilibrium, discrete market model, contract theory

DOI: 10.5604/01.3001.0014.1203

PRZEGLĄD STATYSTYCZNY
R. LXIII – ZESZYT 3 – 2016

ŁUKASZ LENART[1], BŁAŻEJ MAZUR[2]

ON BAYESIAN INFERENCE FOR ALMOST PERIODIC IN MEAN AUTOREGRESSIVE MODELS[3]

1. INTRODUCTION

In this paper we discuss Bayesian approach in case of autoregressive model with time-varying mean function. The focus is on providing an effective numerical method for posterior inference in a rather specific, highly non-linear case. Our discussion of general prior assumptions and model specification issues is therefore somewhat limited.

We make use of the idea of almost periodic time series (used in non-parametric statistics) and consider its parametric counterpart in which e.g. unconditional mean is represented by so-called Flexible Fourier Form of Gallant (1981). Models based on Fourier form with unknown set of frequency parameters are highly nonlinear and therefore difficult to estimate in case when the number of frequencies (characterizing the fluctuations) is greater than one, which is exactly the case of empirically interesting specifications.

Models of this kind are often referred to as deterministic cycle models (see for example Harvey, 2004). However, within a Bayesian approach and with non-trivial number of estimated frequencies the resulting pattern of fluctuations is quite complicated and the models can be considered competitive to stochastic cycle specifications, especially for relatively short series of data. The problems of Bayesian inference stem from the fact that the resulting posterior distribution can be multimodal and therefore difficult to explore by standard MCMC methods. One might also notice that the multimodal posterior (resulting from multimodal likelihood function) results in substantial differences between results obtained by Maximum Likelihood (ML in short) and Bayesian methods, as the multimodal posterior cannot be accurately approximated by multivariate Gaussian distribution.

Our suggestion on how to explore the posterior distribution with MCMC methods is actually two-fold. Firstly, following by the results presented in Bretthorst (1988) we

[1] Cracow University of Economics, Faculty of Finance, Department of Mathematics, 27 Rakowicka St., 31-510 Cracow, Poland, corresponding author – e-mail: lenartl@uek.krakow.pl.

[2] Cracow University of Economics, Faculty of Management, Department of Econometrics and Operations Research, 27 Rakowicka St., 31-510 Cracow, Poland.

[3] This research was supported by the Polish National Science Center based on decision number DEC-2013/09/B/HS4/01945.

make use of a non-parametrically motivated estimator to construct a proposal density for the frequency parameters. Secondly, we demonstrate how the standard conjugate results (with respect to other model parameters) can be used to reduce dimensionality of the problem. The latter step is quite interesting as it takes precisely the opposite direction compared to the usual augmentation strategy that expands the parameter space.

The remaining part of the paper has the following structure: we begin by introducing the idea of almost periodicity and recall basic results on non-parametric models with non-periodicity in mean. We also indicate some relationships between parametric Bayesian and non-parametric estimates in a very simple case. Subsequently we develop a parametric counterpart to a model representing almost-periodicity in mean which makes use of a Flexible Fourier Form. Eventually we consider two parametric models representing the process of interest. The first model, labelled "approximate" allows for taking full advantage of the standard conjugate results in Bayesian partially linear (or conditionally linear) models. In the model it is possible to obtain the kernel of marginal posterior density for frequency parameters using analytical integration only, with generates a closed-form solution (up to a normalizing constant).

However, the approximation model is not satisfactory being quite restrictive as to the way the prior information can be introduced. It does not allow for clear elaboration of prior knowledge as to the unconditional men of the process without interference with information on its autocovariance structure. Moreover, the stationarity restriction of the autoregressive part is somewhat more difficult to handle in the setup.

We therefore consider another Bayesian model based on modified parametrization, labeled "final", which is free of such inconveniences. The two Bayesian models (the "approximate" and the "final") are built upon sampling models (likelihoods) that are observationally equivalent, however only the latter has desirable overall properties. Our ultimate goal is to develop a practical MCMC algorithm for estimation of the "final" model.

We claim that the standard MCMC approaches applied to the final model are very likely to fail to explore the full posterior (and the failure is not easy to detect based just on the MCMC output). We make use of the approximate model to demonstrate the problematic structure of the posterior distribution (in particular its multimodality). The demonstration is not contaminated by possible numerical inaccuracies since it is based on analytical results.

After discussing the reasons that are likely to make the standard algorithms impractical we introduce two ideas that alleviate the problem. The first one amounts to indicating that certain non-parametric results can be used to create an efficient proposal for one group of parameters that display multimodality. The second one is based on the fact that for some other vector of parameters a standard full conditional distribution is available. The fact is often used to build a Gibbs sampler exploring the posterior, but in the case considered here such a strategy would lead to numerical inefficiency. Instead, we use the analytical results to integrate out a sub-vector of parameters from the posterior.

Our amended numerical method therefore targets a marginalized posterior kernel for a sub-vector of all the remaining model parameters. The marginalized posterior kernel is likely to be less irregular compared to the full kernel. The remaining parameters (that have been integrated out) can be sampled outside the MCMC by direct sampling, which has no negative effect on numerical efficiency. We show that using the amended algorithm for the final model we obtain the results that are in line with the analytical results from the approximate model (and the models differ only by the priors). The above problems are illustrated using both simulated and real data.

2. NON-PARAMETRIC APPROACH

The models with periodic mean or autocovariance function are broadly used in econometrics (see for example: Parzen, Pagano, 1979; Osborn, Smith, 1989; Franses, 1996; Franses, Dijk, 2005; Bollerslev, Ghysels, 1996; Burridge, Taylor, 2001; Mazur, Pipień, 2012; Lenart, Pipień, 2013a; 2013b). Formally we say that a second order real valued time series $\{Y_t\colon t \in \mathbb{Z}\}$ is periodically correlated (in short PC) if the mean function $\mu(t) = E(Y_t)$ and the autocovariance function $B(t, \tau) = \mathrm{cov}(Y_t, Y_{t+\tau})$ exists (for any $T \in \mathbb{Z}$) and are periodic functions at variable t with period T. In this parer we consider broader class, the class of almost periodically correlated time series (in short APC). In this class of time series the mean function and the autocovariance function are assumed to be almost periodic in time (see Corduneanu, 1989). This class of time series was applied in business fluctuations analysis in Lenart, Pipień (2013a) with subsampling application. Mazur, Pipień (2012) used this class of time series in modeling volatility of daily financial returns. In ACP case the mean function and the autocovariance function has Fourier representation (see for example Hurd, 1989; Hurd, 1991; Dehay, Hurd, 1994):

$$\mu(t) \sim \sum_{\varphi \in \Psi} m(\varphi) e^{i\varphi t}, \tag{1}$$

$$B(t,\tau) \sim \sum_{\lambda \in \Lambda_\tau} a(\lambda,\tau) e^{i\lambda t}, \tag{2}$$

where the Fourier coefficients $m(\varphi)$ and $a(\lambda,\tau)$ are given by the limits:

$$m(\varphi) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mu(t) e^{-i\varphi t},\ a(\lambda,\tau) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} B(j,\tau) e^{-i\lambda j}, \tag{3}$$

and the sets $\Psi = \{\varphi \in [0,2\pi) \colon |m(\varphi)| \neq 0\}$ and $\Lambda_\tau = \{\lambda \in [0,2\pi) \colon |a(\lambda,\tau)| \neq 0\}$ are countable.

In non-parametric approach the natural estimator based on sample $\{X_1, X_2, \ldots, X_n\}$ of Fourier coefficients $m(\varphi)$ has the following form

$$\widehat{m}_n(\varphi) = \frac{1}{n} \sum_{j=1}^{n} X_j e^{-ij\varphi}, \tag{4}$$

where $\varphi \in [0,2\pi)$. As was shown in Lenart (2013) this estimator after appropriate normalizing is asymptotically normal distributed with zero mean and variance-covariance matrix that depends on spectral density function. Unfortunately in non-parametric approach the spectral density estimation is still an open problem in the case of unknown set of frequency Ψ. Therefore it is not passible to use plug in technique in statistical inference. Therefore authors use subsampling method to estimate asymptotic distribution, where knowledge about exact parameters is not necessary. An applications of the non-parametric methodology to business cycle analysis was presented by Lenart (2013) and Lenart, Pipień (2013a). Details concerning subsampling methodology in general problems are discussed e.g. by Politis et al. (1999).

In our future consideration we weaken the assumption concerning the set Ψ. For the set Ψ we assume that is finite. Therefore the equivalent representation for $\mu(t)$ takes the form:

$$\mu(t) = \delta_0 + \sum_{f=1}^{F} a_f \sin(t\varphi_f) + \sum_{f=1}^{F} b_f \cos(t\varphi_f), \tag{5}$$

where F is an unknown nonnegative integer, $\delta_0 \in \mathbb{R}$, $\mathbf{a} = (a_1,a_2,\ldots,a_F) \in \mathbb{R}^F$, $\mathbf{b} = (b_1,b_2,\ldots,b_F) \in \mathbb{R}^F$ and $\boldsymbol{\varphi} = (\varphi_1,\varphi_2,\ldots,\varphi_F) \in (0,\pi]^F$. Parameters $\mathbf{a}$ and $\mathbf{b}$ are below referred to as amplitudes, whereas elements of $\boldsymbol{\varphi}$ are labeled frequencies.

3. PARAMETRIC BAYESIAN APPROACH

In what follows we confine our attention to parametric models with time-varying unconditional mean given by (5), which (for known F) corresponds to a special case of Flexible Fourier Form discussed by Gallant (1981). One might notice that without further assumptions the parameters in (5) are not identified (due to so-called label switching). This is one source of multimodality of the joint posterior kernel and can be relatively easily eliminated by introducing a restriction of the form $0 < \varphi_1 < \varphi_2 < \ldots < \varphi_F \leq \pi$. However, here we do not impose it, though it can be easily be done in post-processing of MCMC output if desired. Our point is that there exists another source of multimodality driven by properties (5) and typical features of macroeconomic data, and it can be seen even in the case of $F = 1$, where no identification issues arise (as discussed below).

Moreover, here we do not discuss how one choses the value of F. However, within the Bayesian paradigm the models representing whole sequence with $0 < F < F_{\max}$ can be compared and the inference on regular fluctuations in mean or prediction can be based on the pooled results taking into account various values of F.

Here we assume that the deviations from the mean take the autoregressive form with J lags:

$$L(B)(y_t - \mu(t)) = \varepsilon_t, \tag{6}$$

where the function $\mu(t)$ is given by (5) and $L(B) = 1 - \eta_1 B - \eta_2 B^2 - \ldots - \eta_J B^J$ with backshift operator B: $B^k y_t = y_{t-k}$, $\{\varepsilon_t\}$ being a Gaussian (therefore strict) white noise process with precision $\tau > 0$. Notice that the observable series y_t is non-stationary in mean, though its covariance structure (under standard assumptions for coefficients of polynomial $L(B)$) corresponds to that of a covariance stationary process.

The sampling model (6) is observationally equivalent to:

$$L(B)y_t = \mu^*(t) + \varepsilon_t, \tag{7}$$

though of course in (7) $\mu^*(t) = L(B)\mu(t)$ is no longer an unconditional mean of y_t. We refer to (6) as to a final model, whereas (7) is labeled approximate.

Bretthorst (1988) has shown that in a simple case with $F = 1$ and $J = 0$ the posterior distribution of φ_1 (under uniform priors for amplitudes and Jeffreys prior for τ) can be approximated by:

$$p(\varphi_1|y) \propto \left[1 - \frac{2n|\hat{m}_n(\varphi_1)|^2}{\sum_{t=1}^{n} y_t^2}\right]^{\frac{2-n}{2}}. \tag{8}$$

It is easy to see that distribution with kernel (8) is generally a multimodal distribution. The number of modes is the same as the number of local maxima of the periodogram. The kernel (8) is a differentiable function on variable φ_1 and the derivative can be express as a product of derivative of the function $|\hat{m}_n(\varphi_1)|$ and some function with positive values on considered interval.

4. POSTERIOR DISTRIBUTION OF FREQUENCY PARAMETERS IN THE APPROXIMATE MODEL

In this section we obtain explicit formulation of marginal distribution for vector of frequency parameters in the approximate model (7) using the results based on the use of conjugate priors in conditionally linear models (see. e.g. Osiewalski, 1991). Note that the approximate model can be equivalently written as:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon, \tag{9}$$

where $\mathbf{y} = (y_1\ y_2\ \ldots\ y_T)'$,

$$\mathbf{X} = \begin{bmatrix} 1 & y_0 & y_{-1} & \ldots & y_{-J+1} & \sin(\varphi_1) & \cos(\varphi_1) & \sin(\varphi_2) & \cos(\varphi_2) & \ldots & \sin(\varphi_F) & \cos(\varphi_F) \\ 1 & y_1 & y_0 & \ldots & y_{-J+2} & \sin(2\varphi_1) & \cos(2\varphi_1) & \sin(2\varphi_2) & \cos(2\varphi_2) & \ldots & \sin(2\varphi_F) & \cos(2\varphi_F) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & y_{T-1} & y_{T-2} & \ldots & y_{T-J} & \sin(T\varphi_1) & \cos(T\varphi_1) & \sin(T\varphi_2) & \cos(T\varphi_2) & \ldots & \sin(T\varphi_F) & \cos(T\varphi_F) \end{bmatrix},$$

$$\beta = (\delta_0\ \eta_1\ \eta_2\ \ldots\ \eta_J\ a_1\ b_1\ a_2\ b_2\ \ldots\ a_F\ b_F)',$$

$$\varepsilon = (\varepsilon_1\ \varepsilon_2\ \ldots\ \varepsilon_T)', \varepsilon_t \sim N(0, \tau^{-1}), \text{ for } t = 1, 2, \ldots, T.$$

The $N(0,\tau^{-1})$ denotes Gaussian distribution with zero mean and variance τ^{-1}. In the above $\mathbf{X}$ depends on $\boldsymbol{\varphi}$'s being model parameters, but we suppress that to keep notation simple and to highlight relationships with standard conjugate results obtained in linear regression. Denote $\theta = (\beta,\tau,\varphi)$. Then the likelihood function has the following form:

$$p(\mathbf{y}|\theta) = \frac{1}{\sqrt{(2\pi)^n}}\tau^{\frac{n}{2}}\exp\left\{-\frac{\tau}{2}(\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}\beta)\right\}. \tag{10}$$

Following the standard conjugate approach we assume the following prior structure:

$$p(\theta) = p(\beta,\tau)p(\boldsymbol{\varphi}) = p(\beta|\tau)p(\tau)p(\boldsymbol{\varphi}),$$

with $\beta|\,\tau\sim N(\mathbf{c},(\tau\mathbf{B})^{-1})$ and $\tau\sim G\left(\frac{n_0}{2},\frac{s_0}{2}\right)$, where $G\left(\frac{n_0}{2},\frac{s_0}{2}\right)$ denotes the Gamma distribution with expectation $\frac{n_0}{s_0}$ and variance $\frac{2n_0}{s_0^2}$ and $\mathbf{c}$, $\mathbf{B}$, n_0, s_0 are hyperparameters. This implies:

$$p(\beta|\tau) = (2\pi)^{-k/2}(\det(\mathbf{B}))^{1/2}\tau^{k/2}\exp\left\{-\frac{\tau}{2}(\beta-\mathbf{c})'\mathbf{B}(\beta-\mathbf{c})\right\},$$

$$p(\tau) = \frac{(s_0/2)^{\frac{n_0}{2}}}{\Gamma\left(\frac{n_0}{2}\right)}\tau^{\frac{n_0}{2}-1}\exp\left(-\frac{s_0\tau}{2}\right).$$

For the frequency parameter we assume uniform prior distribution:

$$p(\boldsymbol{\varphi}) = \prod_{i=1}^{F} p(\varphi_i) \text{ and } \varphi_i \sim U(0,\pi),$$

where $U(0,\pi)$ denotes uniform distribution on interval $(0,\pi)$. The above implies that:

$$p(\theta|\mathbf{y}) \propto \tau^{\frac{n+k+n_0}{2}-1}\exp\left\{-\frac{\tau}{2}[(\beta-\mathbf{d})'\mathbf{D}(\beta-\mathbf{d})]\right\}\exp\left\{-\frac{\tau}{2}[-\mathbf{d}'\mathbf{D}\mathbf{d}+\mathbf{c}'\mathbf{B}\mathbf{c}+\mathbf{y}'\mathbf{y}+s_0]\right\},$$

where $\mathbf{D} = \mathbf{X}'\mathbf{X} + \mathbf{B}$ and $\mathbf{d} = \mathbf{D}^{-1}(\mathbf{X}'\mathbf{y} + \mathbf{B}\mathbf{c})$. Integrating over β and over τ we get[4]

$$\begin{aligned} &p(\varphi|\mathbf{y}) \\ &\propto (\det(\mathbf{D}))^{-1/2}(\mathbf{y}'\mathbf{y}-\mathbf{d}'\mathbf{D}\mathbf{d}+\mathbf{c}'\mathbf{B}\mathbf{c}+s_0)^{-\frac{n+n_0}{2}} \\ &\propto (\det(\mathbf{X}'\mathbf{X}+\mathbf{B}))^{-1/2} \\ &\cdot(\mathbf{y}'\mathbf{y}-(\mathbf{X}'\mathbf{y}+\mathbf{B}\mathbf{c})'(\mathbf{X}'\mathbf{X}+\mathbf{B})^{-1}(\mathbf{X}'\mathbf{y}+\mathbf{B}\mathbf{b})+\mathbf{c}'\mathbf{B}\mathbf{c}+s_0)^{-\frac{n+n_0}{2}}. \end{aligned} \tag{11}$$

[4] Note that we assume here that the parameters $\eta_1\ \eta_2\ \ldots\ \eta_J$ are unrestricted so we do not consider stationarity restrictions due to complexity of the issue.

Assuming $\mathbf{c} = \mathbf{0}$ for simplicity we obtain the analytical solution:

$$p(\boldsymbol{\varphi}|\mathbf{y}) \propto (\det(\mathbf{X}'\mathbf{X} + \mathbf{B}))^{-1/2}(\mathbf{y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X} + \mathbf{B})^{-1}\mathbf{X}']\mathbf{y} + s_0)^{-\frac{n+n_0}{2}}. \quad (12)$$

Note that distribution with kernel (12) is bounded on the set $(\varphi_1, \varphi_2, \ldots, \varphi_F) = [0, \pi)^F$, hence all posterior moments exist and it is symmetric, which follows directly from model equation (7). Unfortunately, the kernel (12) does not characterize any known distribution in the literature. In addition, contrary to the result (8) of Bretthorst (1988), the direct theoretical relation to periodogram in the case is not obvious (see distribution (12)). Hence, to illustrate the linkages between (14) and a periodogram we consider a short simulation study.

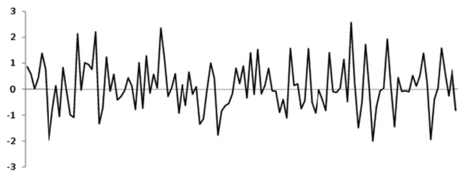
5. A SIMULATION STUDY

We restrict the attention only to the case $J = 0$ to examine the relation of (12) to usual periodogram function without additional relation to autoregressive part. We consider three cases with $F = 0, 1, 2, 3$. At each case we generate $n = 120$ realizations from considered model and we determine the distribution (12). In practice we try to choose the best F, therefore in simulation study at each case we compare the periodogram with the univariate distribution (12) (under model assumption $F = 1$) and bivariate distribution (12) (under model assumption $F = 2$). To make the results visible we use additionally the logarithmic scale. For the hyperparamiters we take $\mathbf{B}^{-1} = 100\ \mathbf{I}$, $n_0 = 2.1$ and $s_0 = 1.05$.

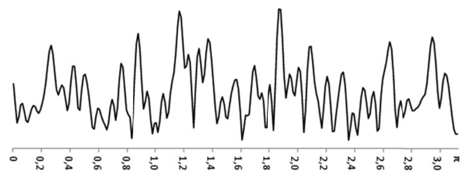
When sample is generated in the case $F = 0$ (see figure 1), the distribution (12) turns out to be multimodal under assumption of $F = 1$ and $F = 2$. Two peaks for posterior distribution (11) with $F = 1$ (see figure 1(c)) and four peaks with $F = 2$ (see figure 1(e)) correspond clearly to two dominant peaks on periodogram (see figure 1(b)).

Figure 2(a) shows a sample from model with one frequency ($\varphi_1 = 0.15$) with relatively large amplitude as compared to the variance of the white noise (see figure 2(b)). In this case the mass of probability in the posterior distribution (12) is strongly concentrated around the point where $\varphi = 0.15$ (under model assumption $F = 1$ and around sets: $0.15 \times (0, \pi)$ and $(0, \pi) \times 0.15$ (under model assumption $F = 2$).

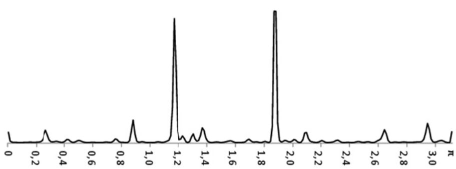
If we consider sample obtained from the model with two different frequencies ($\varphi_1 = 0.15$, $\varphi_2 = 0.5$) with different amplitudes (see figure 3(a–b)), the posterior distribution (12) with $F = 1$ (see figure 3(c)) has only one dominating peak around the frequency with larger amplitude (in this case: φ_1). The probability mass concentrated around the second frequency (φ_2) is much lower (see figure 3(d)). The posterior distribution (12) under assumption of $F = 2$ (see figure 3(e)) has two symmetric peaks that clearly correspond to points $(\varphi_1, \varphi_2) = (0.15, 0.5)$ and $(\varphi_1, \varphi_2) = (0.5, 0.15)$.



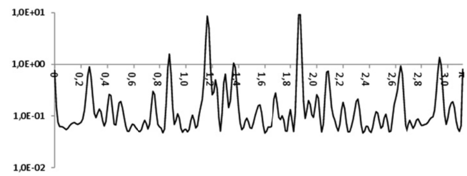
(a) Realizations from model (6)



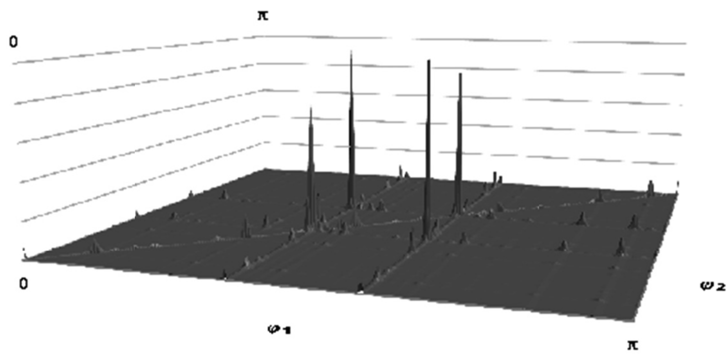
(b) The value of normalized periodogram: $|\hat{m}_n(\varphi)|$ for $\psi \in (0, \pi)$



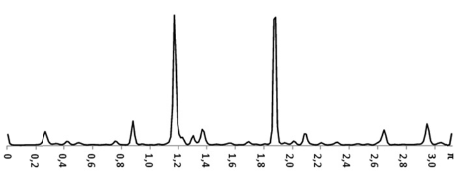
(c) Posterior distribution (12) for $\varphi \in (0, \pi)$ under model assumption $F = 1$



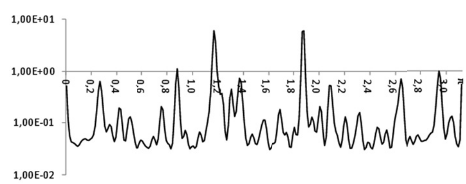
(d) Posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under model assumption $F = 1$



(e) Posterior distribution (12) for $(\varphi_1, \varphi_2) \in (0, \pi)^2$, under assumption $F = 2$



(f) Marginal posterior distribution (12) for $\varphi \in (0, \pi)$ under assumption $F = 2$

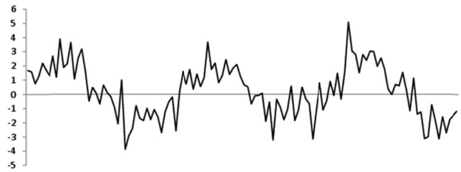


(g) Marginal posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under assumption $F = 2$

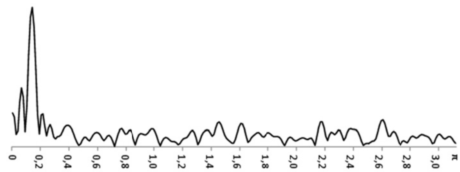
Figure 1. Posterior distributions in case of sample with length $n = 120$ generated from considered model (6) with $F = 0$ and $\tau = 1$

Source: own calculations.

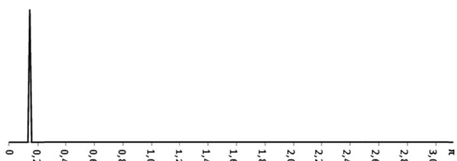
The last case, where sample is generated from model with three frequencies: $\varphi_1 = 0.15$, $\varphi_2 = 0.5$ and $\varphi_3 = 2.2$ is presented on figure 4. The amplitude for the first frequency is the biggest, while for second and third frequencies are equal (see figure 4(b)). Univariate distribution (under model assumption $F = 1$) has only one peak that clearly corresponds to frequency with the highest amplitude (φ_1). Two peaks that correspond to second and third frequency are visible only in the logarithmic scale (on marginal distribution). The bivariate distribution for frequency (under model assumption $F = 2$) has four peaks that clearly corresponds to the points (0.15, 0.5), (0.5, 0.15), (0.15, 2.2) and (2.2, 0.15).



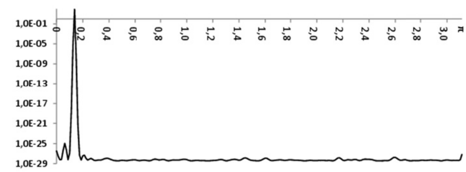
(a) Realizations from model (6)



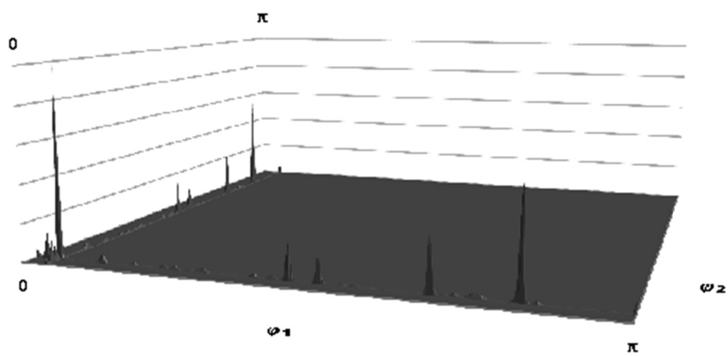
(b) The value of normalized periodogram: $|\hat{m}_n(\varphi)|$ for $\psi \in (0, \pi)$



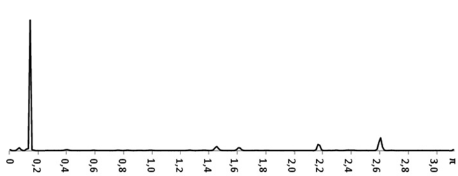
(c) Posterior distribution (12) for $\varphi \in (0, \pi)$ under model assumption $F = 1$



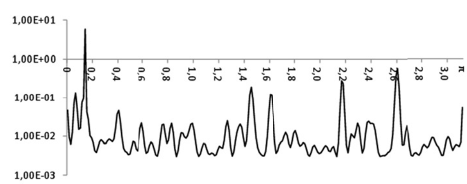
(d) Posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under model assumption $F = 1$



(e) Posterior distribution (12) for $(\varphi_1, \varphi_2) \in (0, \pi)^2$, under assumption $F = 2$



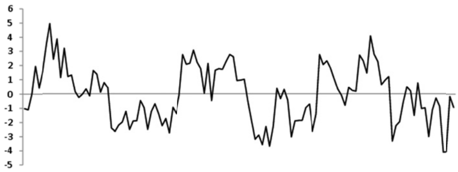
(f) Marginal posterior distribution (12) for $\varphi \in (0, \pi)$ under assumption $F = 2$



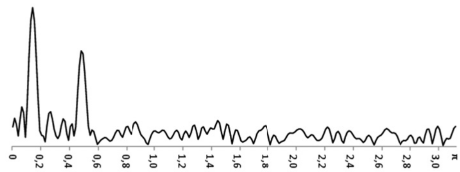
(g) Marginal posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under assumption $F = 2$

Figure 2. Posterior distributions in case of sample with length $n = 120$ generated from considered model (6) with $F = 1$, $\varphi_1 = 0.15$, $a_1 = 2$, $b_1 = 0$, $\tau = 1$

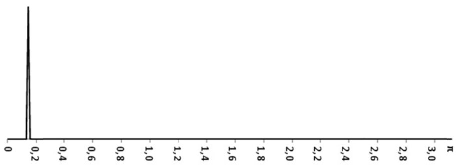
Source: own calculations.



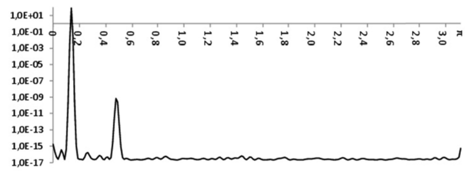
(a) Realizations from model (6)



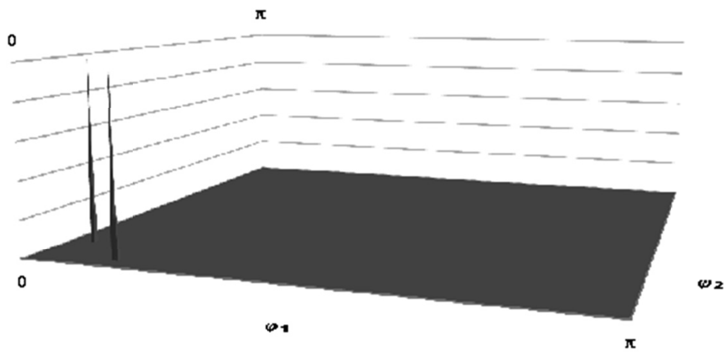
(b) The value of normalized periodogram: $|\hat{m}_n(\varphi)|$ for $\psi \in (0, \pi)$



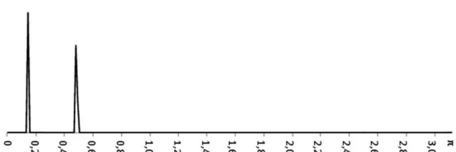
(c) Posterior distribution (12) for $\varphi \in (0, \pi)$ under model assumption $F = 1$



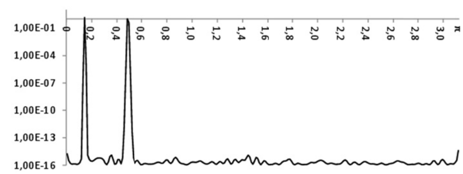
(d) Posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under model assumption $F = 1$



(e) Posterior distribution (12) for $(\varphi_1, \varphi_2) \in (0, \pi)^2$, under assumption $F = 2$



(f) Marginal posterior distribution (12) for $\varphi \in (0, \pi)$ under assumption $F = 2$

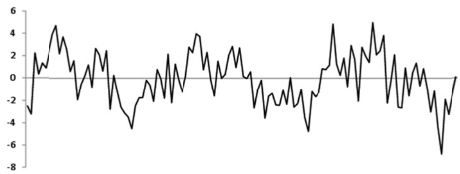


(g) Marginal posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under assumption $F = 2$

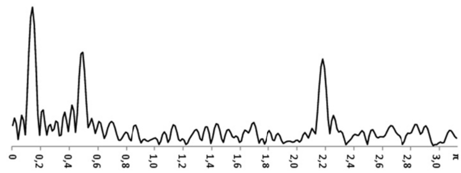
Figure 3. Posterior distributions in case of sample with length $n = 120$ generated from considered model (6) with $F = 2$, $\varphi_1 = 0.15$, $\varphi_2 = 0.5$, $a_1 = 2$, $b_1 = 0$, $a_2 = -1.5$, $b_2 = 0$, $\tau = 1$

Source: own calculations.

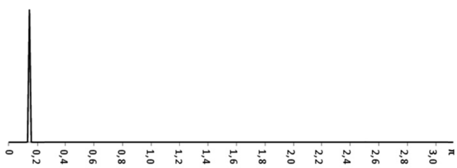
The above simulation study strongly exposes the relationship between shape of the periodogram and related posterior distributions for frequency parameters. Most importantly we demonstrate that in cases corresponding to the number of observation that characterizes typical macroeconomic applications, the resulting posterior might have highly irregular shape. Sources of the multimodality go well beyond the non-identification issue arising from the label switching. Moreover, the lack of global identification generates no theoretical problems within the Bayesian approach and can be easily resolved without any change of the MCMC algorithm discussed here.



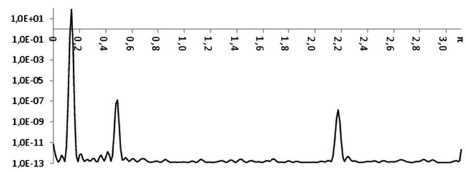
(a) Realizations from model (6)



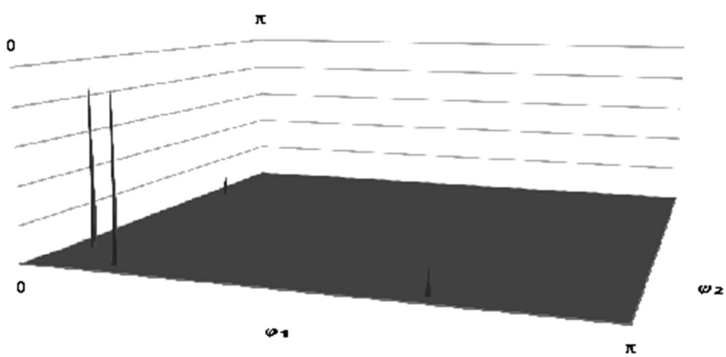
(b) The value of normalized periodogram: $|\hat{m}_n(\varphi)|$ for $\psi \in (0, \pi)$



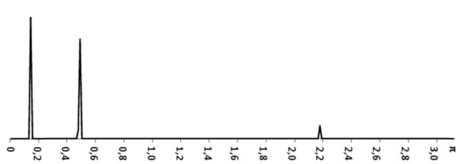
(c) Posterior distribution (12) for $\varphi \in (0, \pi)$ under model assumption F = 1



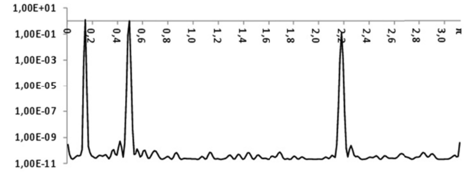
(d) Posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under model assumption F = 1



(e) Posterior distribution (12) for $(\varphi_1, \varphi_2) \in (0, \pi)^2$, under assumption F = 2



(f) Marginal posterior distribution (12) for $\varphi \in (0, \pi)$ under assumption F = 2



(g) Marginal posterior distribution (12) in logarithmic scale for $\varphi \in (0, \pi)$ under assumption F = 2

Figure 4. Posterior distributions in case of sample with length n = 120 generated from considered model (6) with $F = 2$, $\varphi_1 = 0.15$, $\varphi_2 = 0.5$, $\varphi_3 = 2.2$, $a_1 = 2$, $b_1 = 0$, $a_2 = -1.5$, $b_2 = 0$, $a_3 = 0$, $b_3 = 1.5$, $\tau = 1$

Source: own calculations.

6. MCMC SAMPLER FOR POSTERIOR INFERENCE

In the following discussion we assume basic knowledge of MCMC algorithm used in Bayesian inference – for an accessible review see e.g. Osiewalski (2001). An obvious approach to Bayesian estimation of the final model would be to sample from the full posterior $p(\theta|\mathbf{y})$ using a Gibbs sampler. Such a sampler would be based on factorization of $p(\theta|\mathbf{y})$ into full conditionals for sub-vectors of θ, of which at least some have a standard form (as conjugate-type priors are used). In particular a sampler consisting of four steps, for linear parameters of the mean $(\delta_0, a_1, a_2, \ldots, a_F, b_1, b_2, \ldots, b_F)$, τ, $\boldsymbol{\eta} = (\eta_1\ \eta_2\ \ldots\ \eta_J)$ and $\boldsymbol{\varphi}$ respectively, is an obvious solution. However, we point out that two difficulties would arise. Firstly, one would need a good proposal for the frequency parameters sampled within a Mertopolis-Hastings (M-H in short) step, as the full conditional posterior is definitely not a standard one in this case. Secondly, even after addressing that, such a sampler could fail to achieve convergence to true posterior within a finite and practical timespan. This is because in practical cases the joint posterior would be multimodal and a move from one mode to another would require a change in parameters belonging to two separate Gibbs blocks (namely frequencies and amplitudes). Under fairly weak conditions such a change has a very low chance and this arises just from the conditioning inherent in such a sampler, which therefore would fail to visit all the relevant modes.[5]

In order to solve the issue we first introduce the idea of posterior marginalization. Consider the following general factorization of a posterior distribution:

$$p(\theta|\mathbf{y}) = p(\theta^{(1)} \,|\, \mathbf{y},\theta^{(2)})p(\theta^{(2)} \,|\, \mathbf{y}) \propto k(\theta^{(1)} \,|\, \mathbf{y},\theta^{(2)})k(\theta^{(2)} \,|\, \mathbf{y}) = k(\theta|\mathbf{y}),$$

where $\theta' = [\theta^{(1)\prime}\ \theta^{(2)\prime}]$ We assume that $p(\theta^{(1)} \,|\, \mathbf{y},\theta^{(2)})$ represents full conditional posterior for $\theta^{(1)}$ that has a known form. Its kernel is $k(\theta^{(1)} \,|\, \mathbf{y},\theta^{(2)})$ and the normalizing constant is known (and depends on $\theta^{(2)}$). Consequently, $\theta^{(1)}$ can be integrated out from the posterior using analytical techniques, resulting in a closed form of marginal posterior kernel for $\theta^{(2)}$ only. The resulting marginal kernel usually retains all the terms from $k(\theta|\mathbf{y})$ not included in $k(\theta^{(1)} \,|\, \mathbf{y},\theta^{(2)})$ and inverse of the normalizing constant of $p(\theta^{(1)} \,|\, \mathbf{y},\theta^{(2)})$ being a function of $\theta^{(2)}$.

By the virtue of marginalization, $p(\theta^{(2)} \,|\, \mathbf{y})$ is likely to have more regular shape compared to $p(\theta|\mathbf{y})$. Essentially, we aim to improve properties of the MCMC algorithm by adjusting its target distribution, replacing $p(\theta|\mathbf{y})$ with $p(\theta^{(2)} \,|\, \mathbf{y})$ being potentially more regular. Of course finally we draw from $p(\theta|\mathbf{y})$, but this can be achieved by

[5] One might imagine the example of sampling from a bivariate target (with one variable in each Gibbs step) when the target distribution is a mixture of two bivariate normal densities with modes that are separated in both dimensions and variances that are small relative to the difference in modes. Conditionally on MCMC chain visiting one mode, a move to the other one would require occurrence of a very particular tail event in the first step.

additional direct sampling since $p(\theta^{(1)} \mid \mathbf{y},\theta^{(2)})$ has known, standard form. In the final model (6), $\theta^{(1)}$ corresponds to parameters that appear in the unconditional mean (5) linearly, i.e. $\delta_0, a_1, a_2, \ldots, a_F, b_1, b_2, \ldots, b_F$. The resulting marginal posterior kernel is $p(\theta^{(2)} \mid \mathbf{y})$ where $\theta^{(2)}$ includes $\boldsymbol{\eta} = (\eta_1\ \eta_2\ \ldots\ \eta_J)$, τ and φ and it does not include amplitude parameters, as these are integrated out.

In order to sample from $p(\theta^{(2)} \mid \mathbf{y})$ one might construct another Gibbs sampler with M-H steps for $\boldsymbol{\eta} = (\eta_1\ \eta_2\ \ldots\ \eta_J)$, τ and $\boldsymbol{\varphi}$ (with stationary restrictions imposed on $\boldsymbol{\eta}$). Here again a crucial problem to be solved would be the one of sampling frequency parameters $\boldsymbol{\varphi}$. We suggest using a M-H step with a proposal density being a product of identical (normalized) magnitude of periodogram functions (4) restricted to interval of interest for the frequencies. The one-dimensional problem with finite support can be handled numerically in an effective way (using one-dimensional numerical representation of the univariate density generated with an arbitrary precision). The distribution would allow for a simple design of a M-H step with an independent proposal.

7. REAL DATA EXAMPLE

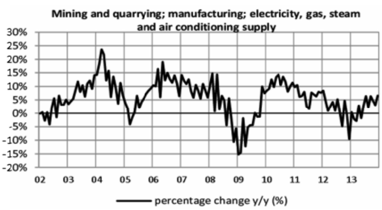
In this section we consider two data sets from the Polish economy concerning growth rates of monthly production in industry (percentage change compared to corresponding period of the previous year, y-o-y in short): *Mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply* and *MIG – Non-durable consumer goods*[6]. The samples start at January 2002 and end at December 2013. These two economic processes belong to main cyclical indicators of economy. In comparison analysis we take $J = 0$ and the same prior distributions as in section 4, since under the assumptions our “approximate” and “final” formulations coincide exactly, therefore the MCMC output[7] (see detailed algorithm in the previous section) can be compared to an exact, analytical benchmark (12).

The two real data examples show that in the bivariate case the shape of the distribution for the frequency parameters obtained by proposed MCMC sampler is comparable with theoretical distribution (12) (see results on figures 5–6). This demonstrates the efficiency of the sampler proposed here in a real data example. For extreme cases with really high number of unknown frequencies (that are unlikely to be encountered within macroeconomic applications) the approach could be refined by taking a proposal for frequency parameters based on the marginal posterior (12) obtained analytically from the approximate model (this would also require a numerical approximation of a marginalized univariate version of (12) instead of periodogram).

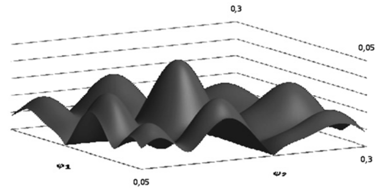
---

[6] Source: Eurostat.

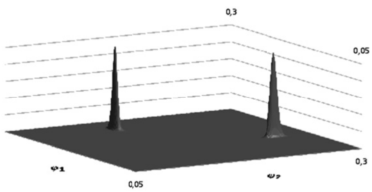
[7] 50 000 burn-in cycles and 1 000 000 final cycles.



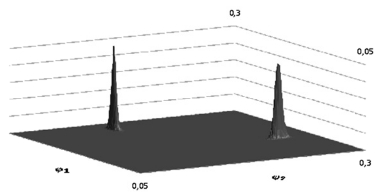
The data



The value of $|\hat{m}_n(\varphi_1)\ \hat{m}_n(\varphi_2)|$ – proposal kernel of distribution in Metropolis-Hastings step

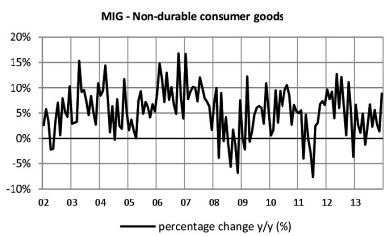


Analytical marginal posterior for frequency parameters (12) under assumption $F = 2$

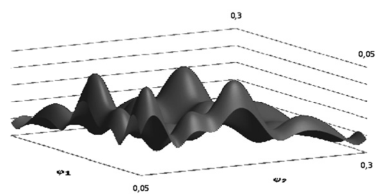


Histogram (marginal posterior) for frequency parameters based on MCMC sample from the algorithm proposed in the paper

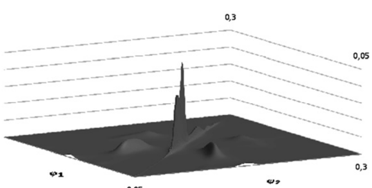
Figure 5. The comparison of posterior distribution (12) with obtained MCMC sample
Source: own calculations.



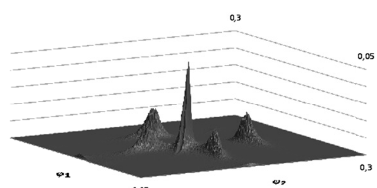
The data



The value of $|\hat{m}_n(\varphi_1)\ \hat{m}_n(\varphi_2)|$ – proposal kernel of distribution in Metropolis-Hastings step



Analytical marginal posterior for frequencies (12) under assumption $F = 2$



Two dimensional histogram (marginal posterior) for frequency based on MCMC sample from the algorithm proposed in the paper

Figure 6. The comparison of posterior distribution (12) with obtained MCMC sample
Source: own calculations.

8. CONCLUSIONS

In the paper we highlight some problems that arise in Bayesian estimation of parametric time-series model with fluctuations (corresponding to e.g. business cycle) are modelled using Flexible Fourier Form of Gallant (1981). The problems appear in empirically appealing cases with more than one unknown frequency parameter. We demonstrate that the resulting posterior is likely to be highly multimodal. This cast doubts on applicability of ML estimation, but can also result in problems within the Bayesian approach, as standard MCMC methods might fail to explore the whole posterior, especially when the modes are separated.

We demonstrate that the multimodality is actually an issue using the exact solution (i.e. an analytical marginal posterior) in an approximate model. The approximate model differs from our target (final) specification by the prior assumptions only. The posterior multimodality seems to be most severe within the joint space of amplitude and frequency parameters.

We address that problem using two essential steps. Firstly, we integrate the posterior with respect to amplitude parameters, which can be carried out analytically. Secondly, we propose a non-parametrically motivated proposal for the frequency parameters. This allows for construction of an improved MCMC sampler that effectively explores the space of all the model parameters, with the amplitudes sampled by the direct approach outside the MCMC chain.

Using the improved algorithm we are able to estimate our target specification which allows prior information to be introduced in a reasonable way: parameters characterizing unconditional mean are separated from those describing autocovariances. In particular one can express prior knowledge on possible amplitudes of regular fluctuations in mean (by setting prior precision of amplitude parameters) or cycle length (by specifying φ_L and φ_U). The approach can be therefore used to "filter" cyclical fluctuations characterized by cycle lengths within a given range. The "extracted" pattern of regular fluctuations would be described by posterior distribution of unconditional mean (as a function of model parameters given by (5)).

Moreover, the causality restriction can be imposed on autoregressive parameters (so that e.g. explosive paths are ruled out *a priori*). In our experience the approach is feasible even with quite high lag order of the autoregressive process.

REFERENCES

Bollerslev T., Ghysels E., (1996), Periodic Autoregressive Conditional Heteroscedasticity, *Journal of Business and Economic Statistics*, 14 (2), 139–152.

Bretthorst G. L., (1988), *Bayesian Spectrum Analysis and Parameter Estimation*, Springer-Verlag, Germany.

Burridge P., Taylor A. M., (2001), On Regression-based Tests for Seasonal Unit Roots in the Presence of Periodic Heteroscedasticity, *Journal of Econometrics*, 104, 91–117.

Corduneanu C., (1989), *Almost Periodic Functions*, Chelsea, New York.

Dehay D., Hurd H., (1994), Representation and Estimation for Periodically and Almost Periodically Correlated Random Processes, in: Gardner W. A., (ed.), *Cyclostationarity in Communications and Signal Processing, IEEE Press*, 295–329.

Franses P. H., Dijk D., (2005), The Forecasting Performance of Various Models for Seasonality and Nonlinearity for Quarterly Industrial Production, *International Journal of Forecasting*, 21, 87–102.

Franses P. H., (1996), *Stochastic Trends in Economic Time Series*, Oxford University Press, New York.

Gallant A. R., (1981), On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Flexible Fourier Form, *Journal of Econometrics*, 15, 211–245.

Harvey A., (2004). Tests for cycles, in: Harvey A. C., Koopman S. J., Shephard N., (eds.), *State space and unobserved component models, CUP*, pages 102–119.

Hurd H., (1989), Representation of Strongly Harmonizable Periodically Correlated Process and their Covariances, *Journal of Multivariate Analysis*, 29, 53–67.

Hurd H., (1991), Correlation Theory of Almost Periodically Correlated Processes, *Journal of Multivariate Analysis*, 37, 24–45.

Lenart Ł., (2013), Non-parametric Frequency Identification and Estimation in Mean Function for Almost Periodically Correlated Time Series, *Journal of Multivariate Analysis*, 115, 252–269.

Lenart Ł., Pipień M., (2013a), Almost Periodically Correlated Time Series in Business Fluctuations Analysis, *Acta Physica Polonica A*, 123 (3), 567–583.

Lenart Ł., Pipień M., (2013b), Seasonality Revisited - Statistical Testing for Almost Periodically Correlated Processes, *Central European Journal of Economic modelling and Econometrics*, 5, 85–102.

Mazur B., Pipień M., (2012), On the Empirical Importance of Periodicity in the Volatility of Financial Returns - Time Varying GARCH as a Second Order APC(2) Process, *Central European Journal of Economic modelling and Econometrics*, 4, 95–116.

Osborn D. R., Smith J. P., (1989), The Performance of Periodic Autoregressive Models in Forecasting Seasonal U.K. Consumption, *Journal of Business & Economic Statistics*, 9, 117–127.

Osiewalski J., (1991), Bayesowska estymacja i predykcja dla jednorównaniowych modeli ekonometrycznych, *Zeszyty Naukowe / Akademia Ekonomiczna w Krakowie. Seria Specjalna, Monografie nr 100.*

Osiewalski J., (2001), *Ekonometria bayesowska w zastosowaniach*, wyd. AE w Krakowie.

Parzen E., Pagano M., (1979), An Approach to Modeling Seasonally Stationary Time Series, *Journal of Econometrics*, 9, 137–153.

Politis D., Romano J., Wolf M., (1999), *Subsampling*, Springer-Verlag, New York.

WNIOSKOWANIE BAYESOWSKIE DLA ZMIENNEJ W CZASIE PRAWIE OKRESOWEJ FUNKCJI WARTOŚCI OCZEKIWANEJ W MODELU AUTOREGRESJI

Streszczenie

Artykuł ma na celu przedstawienie problematyki bayesowskiej estymacji klasy jednowymiarowych modeli dla danych charakteryzujących się występowaniem skomplikowanych wahań cyklicznych w średniej. Koncentrujemy się na zagadnieniach powstających w estymacji parametrycznych modeli dla szeregów czasowych wykorzystujących tzw. giętką formę Fouriera (Flexible Fourier Form, zob. Gallant, 1981), której parametry opisują amplitudę i częstotliwość wahań. Wskazujemy, iż w takich modelach łączny rozkład *a posteriori* charakteryzuje się silną wielomodalnością, przez co standardowe metody numeryczne typu MCMC mogą okazać się raczej zawodnym narzędziem wnioskowania. Ma to miejsce, gdy próbnik MCMC nie odwiedza (w praktyce) wszystkich modalnych badanego rozkładu. Wykorzystując dokładne rozwiązanie analityczne w bardzo zbliżonym modelu wykazujemy, iż wzmiankowana wielomodalność faktycznie ma miejsce. Proponujemy dwa rozwiązania szczegółowe. Po pierwsze wycałkowujemy analitycznie z rozkładu *a posteriori* parametry odpowiadające za amplitudę wahań. Po drugie przedstawiamy specjalnie dobrany rozkład proponujący dla parametrów częstotliwości wyspecyfikowany z wykorzystaniem wyników otrzymanych na gruncie podejścia nieparametrycznego. Tak otrzymany próbnik MCMC w ramach praktycznie użytecznej liczby losowań jest w stanie skutecznie przemieszczać się w (zredukowanej) przestrzeni parametrów. Wycałkowane parametry są dolosowywane poza algorytmem MCMC poprzez losowanie bezpośrednie ze standardowego rozkładu warunkowego. Ilustrujemy omawianą problematykę wykorzystując dane symulacyjne a także dwa przykłady danych rzeczywistych.

Słowa kluczowe: wnioskowanie bayesowskie, funkcja prawie okresowa wartości oczekiwanej, model autoregresji, próbnik MCMC

ON BAYESIAN INFERENCE FOR ALMOST PERIODIC
IN MEAN AUTOREGRESSIVE MODELS

Abstract

The goal of the paper is to discuss Bayesian estimation of a class of univariate time-series models being able to represent complicated patterns of “cyclical” fluctuations in mean function. We highlight problems that arise in Bayesian estimation of parametric time-series model using the Flexible Fourier Form of Gallant (1981). We demonstrate that the resulting posterior is likely to be highly multimodal, therefore standard Markov Chain Monte Carlo (MCMC in short) methods might fail to explore the whole posterior, especially when the modes are separated. We show that the multimodality is actually an issue using the exact solution (i.e. an analytical marginal posterior) in an approximate model. We address that problem using two essential steps. Firstly, we integrate the posterior with respect to amplitude parameters, which can be carried out analytically. Secondly, we propose a non-parametrically motivated proposal for the frequency parameters. This allows for construction of an improved MCMC sampler that effectively explores the space of all the model parameters, with the amplitudes sampled by the direct approach outside the MCMC chain. We illustrate the problem using simulations and demonstrate our solution using two real-data examples.

Keywords: Bayesian inference, almost periodic mean function, autoregressive model, MCMC sampler

DOI: 10.5604/01.3001.0014.1209

PRZEGLĄD STATYSTYCZNY
R. LXIII – ZESZYT 3 – 2016

JAN GADOMSKI[1]

ASSESSMENT OF THE IMPACT OF THE REDUCTION OF THE GASEOUS EMISSIONS ON GROWTH IN POLAND. ASSUMPTIONS AND PRELIMINARY RESULTS

1. INTRODUCTION

The presented research addresses the course of economic transformations induced by the technology conversion forced upon a country by the policy of the abatement of the greenhouse gases emission (GHG).

Most research on this topic present in the literature has been performed using Computable General Equilibrium (CGE) models. In the Polish case such models are, for example, the PLACE model, see Antoszewski et al. (2015), Boratyński (2012), Roberts (1994), and others.

The development of CGE models involves large teams and detailed structure of the models. However, not all research is concerned with very detailed questions and not all assumptions of the research using the CGE models are relevant. For example, the energy sector does not adhere to the model of the perfect competition, on which CGE models are based. A monopoly (or oligopoly) can operate in the range of technical inefficiency. Such a situation is not accounted for in the model of perfect competition. This is why the neoclassical production functions such as, for example, Cobb-Douglass or CES, commonly used in the CGE modelling, cease to be adequate for this task. Moreover, a significant part of the energy sector consists also of the integrated networks (electricity), where it is necessary, out of the strategic reasons, to maintain larger reserves of the unused production capacities than it is common in other sectors. This also makes simplification assumptions applied in the CGE models hard to accept.

Far-reaching simplification commonly used in CGE models is micro-rationality of producers, who maximize profits and are not concerned with market shares or other long-term factors affecting the behavior of firms. Macroeconomic policy in these models is expressed in the values of such parameters as the turnover, personal and corporate taxes, custom duties, interest rates etc. This property makes it possible to investigate the response of the national economy, or more economies linked via economic exchange, to different variants of the economic policy.

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

Another problem concerning common assumptions of the CGE models is that there exists a continuum of available technologies. We doubt that, because it is hard to imagine a complex technology combining, for example, the nuclear technology and the renewable one. These technologies coexist, but develop separately and remain separated.

As to the utilization of the production capacities; reserves of unused capacities persists in long periods. This feature is common not only in the network monopolies.

The above discussion indicating some weaker aspects of the CGE modelling does not dismiss this technique but it shows that there is still space for other approaches.

In this paper we propose a method based on the simpler model, and thus much less work-intensive, able to generate no-nonsense results. This model has been developed in the Systems Research Institute of the Polish Academy of Sciences and evolved from an earlier version with the addition of a separate energy sector; see Gadomski et al. (2014).

The concept of the proposed model is based, contrary to that of CGE models, on the assumption of the macroeconomic rationality and a perfect ability of the macroeconomic policy to pursue its goals by optimal allocation of resources. Such approach provides a benchmark. Similarly to CGE models, all changes preserve sectoral equilibria in real terms at every step, without assuming that prices clear the markets. Quantitative equilibria are maintained in such a way that surpluses/deficits of the domestic markets are cleared via the foreign trade. Producers react to the changes in demand by increasing utilization rates of the production capacities and by increasing production capacities, by purchases of the investment goods. In the long run, without the technical progress, the sector output structure and the country's GDP are determined by the amount of the final allotted amount of the emission allowances. This is equivalent to the zero growth economy. In the presence of favorable technical changes, such as a beneficial evolution of the technological parameters or the emergence of a new economically more efficient technology, economy would start growing with the rate determined by the improvement of the relevant parameters.

Following this introduction, the paper is divided into three sections. The first one describes the method of analysis including the construction of the model. The next section describes the simulation results, and the final one contains conclusions.

2. METHOD OF ANALYSIS

The process of the macroeconomic technological conversion is analyzed with the support of the macroeconomic long-term model embracing four production sectors, each having a limited number of available production technologies. The sectors exchange their products at both the domestic and international markets. The focus is on modelling a small-country economy, a price-taker of international prices. The analysis is simplified by assuming that a change in emission levels does not affect productivities of the production factors. It is an optimization model, and its result indicates a perfect reaction of the national economy to the changes in its conditions/

rules. In the variant considered in this investigation, the overall economic goal of the national economy is the maximization of the present value of the total consumption over the whole simulation period.

In developing this model we do not point to tools and channels of the economic policy. Instead, this model is to serve as a benchmark showing ideal, but feasible in real terms, long-term behavior.

Two options are considered: an economic development without impediment to growth in the form of the emission limits, and another one with the emission limits imposed. It is reasonable to consider, in both options, the impact of the long – term technical progress expressed by evolving values of the parameters that define a given technology.

In the first option we assume that the economy described by the model develops along the long-term growth path using a single technology in each sector, maintaining in all sectors both domestic and external equilibria. The rate of growth is determined by the propensity to invest. This type of growth is characterized by constant proportions of the sectors' outputs, fixed assets, balances of foreign trade, and a certain rate of the utilization of the production capacities. The concept of the long-term equilibria allows another assumption: constant proportions of prices and their real values.

A variant of this option with evolving technology parameters is also worth considering. However, one should be aware that in certain cases there may emerge a possibility of rising economic competitiveness of technologies, which previously did not exist.

In the second option the sectors come across the emission limits, which force adoption of cleaner, previously unconsidered, economically inefficient technologies. Technology conversion influences both levels as well as the output and costs structures. Consider a case without the long-term technological progress. If economic agents are able to coordinate their activities in order to pursue the common goal of welfare/consumption maximization, then after the adjustment period, economy attains a new steady state and the equilibrium at the level determined by the admissible emission level and the structure of the foreign excha nge.

Also in this option a variant with evolving technology parameters can be considered. Such a solution considerably complicates the analysis, therefore it is reasonable to consider only simple hypotheses, such as, for example, one with gradually improving technology parameters reflecting a long-term technical progress.

Model

The letter t, $t = t_0,\ldots,T$, denotes the year. The numbering of years starts with the year 2010, so that t_0 corresponds to the year 2010. The following convention of indexing the model parameters has been applied in this paper: The letter $i = M, E, C, I$, denotes the sector, the letter $j = 1, 2, 3$, denotes technology. M stands for the sector producing non-energy intermediate inputs used in all producing sectors, E denotes the sector producing energy used in all producing sectors as well as the consuming sector,

C stands for the sector producing consumption goods consumed by households and the public sector, and I denotes the sector producing the investment goods supplying the stocks of fixed assets in the production sectors. It is assumed that the number of the available technologies is limited to two in the sectors M, C, I, and three in the energy sector E.

Technology of production

Technology of production in all sectors is described by the following set of parameters in i-th sector, $i = M, E, C, I$; in j-th technology, $j = 1, 2, 3$; in year t, $t = 1,\ldots,T$:

γ_{ijt} – productivity of fixed assets in year t in i-th sector and j-th technology, it is assumed that in the long term in each year the technical progress increases the productivity of the fixed assets by a constant ratio r_γ:

$$\gamma_{ijt} = \gamma_{ijt_0}(1 + r_\gamma)^{t-t_0};$$

where γ_{ijt_0} denotes productivity of the fixed assets in the year t_0;

δ_{ij} – depreciation rate of fixed assets in i-th sector and j-th technology;

α_{ij} – use of goods produced in sector M in producing the unit of the gross product of the i-th sector and j-th technology;

β_{ij} – use of goods produced in sector E in producing the unit of the gross product of the i-th sector and j-th technology;

μ_{ijt} – emission per unit in producing the gross product of the i-th sector and j-th technology in year t, it is assumed that in the long term in each year the technical progress decreases the unit emission by a constant ratio r_μ:

$$\mu_{ijt} = \mu_{ijt_0}(1 + r_\mu)^{t-t_0}.$$

where μ_{ijt_0} denotes unit emission in the year $t_{0,}$ while r_μ denotes the rate of the decrease of the emission unit.

In the current version of the model in all non-energy production sectors (M, C, I) two competing technologies are assumed: the old one, economically more efficient but emitting more GHG, and the costlier but cleaner one. In the energy sector E three technologies are available: the old one, economically more efficient but emitting more GHG; the costlier but cleaner one; and the preferred one, the cleanest of them all but economically inefficient (of which the second can be interpreted as modernized conventional technology, and the latter can be interpreted as renewable energy).

Production capacity

Production capacity defined as the potential gross output Q_{ijt} of the sector i, $i = E, M, C, I$; using j-th technology, $j = 1, 2, 3$; in the year t, $t = 1,\ldots,T$; is described by the following one factor production function:

$$Q_{ijt} = \gamma_{ijt} K_{ijt-1}, \tag{1}$$

where K_{ijt} stands for stock of the fixed assets in sector i and j-th technology at the beginning of the year t. In this paper, the potential gross output (1) will be also called the production capacity of the j-th technology in the sector E in year t.

Actual gross output X_{ijt} cannot exceed the production capacity

$$0 \leq X_{ijt} \leq Q_{ijt}, \quad j = 1, 2, 3;\ t = 1,\ldots,T, \tag{2}$$

and it can be expressed in the following form:

$$X_{ijt} = \varphi_{ijt} Q_{ijt}, \quad j = 1, 2, 3;\ t = 1,\ldots,T, \tag{3}$$

where φ_{ijt} stands for the coefficient of the production capacity utilization in the i-th sector, $i = E, M, C, I$; using j-th technology, $j = 1, 2, 3$; in year t, assuming values from the range [0;1]. (In particular, $\varphi_{ijt} = 0$ indicates fully idle capital and $\varphi_{ijt} = 1$ represents full utilization of the production capacity of j-th technology in i-th sector in the year t).

Total actual output of the i-th sector, $i = E, M, C, I$; is the sum of outputs produced using available technologies:

$$X_{it} = X_{i1t} + X_{i2t} + X_{i3t}, \quad t = 1,\ldots,T. \tag{4}$$

Stock of the fixed assets K_{ijt} using j-th technology, $j = 1, 2, 3$; in the i-th sector, $i = E, M, C, I$; at the end of year t is given by the relationship:

$$K_{ijt} = K_{ijt-1}(1 - \delta_{ij}) + I_{ijt}, \quad j = 1, 2, 3;\ t = 1,\ldots,T, \tag{5}$$

where I_{ijt} denotes investment in the j-th technology, $j = 1, 2, 3$; in the i-th sector, $i = E, M, C, I$; in the year t. (Note that the term $K_{ijt-1}\delta_{ij}$ denotes depreciation of the capital in i-th sector). For simplicity one year lag between the investment and its contribution to the stock of fixed assets is assumed.

Production of the i-th sector using j-th technology in year t causes the emissions S_{ijt} of GHG:

$$S_{ijt} = \mu_{ijt} S_{ijt}, \quad i = E, M, C, I;\ j = 1, 2;\ t = 1,\ldots,T. \tag{6}$$

The total emission of GHG by the i-th sector in the year t equals:

$$S_{it} = S_{i1t} + S_{i2t} + S_{i3t}, \quad i = E, M, C, I;\ t = 1,\ldots,T. \tag{7}$$

Gross income GI_t is defined as the sum of incomes generated in the sectors E, M, C and I:

$$\begin{aligned} GI_t = & \left[1-(\alpha_{E1}+\beta_{E1})\right]X_{E1t}+\left[1-(\alpha_{E2}+\beta_{E2})\right]X_{E2t}+\left[1-(\alpha_{E3}+\beta_{E3})\right]X_{E3t}+ \\ & +\left[1-(\alpha_{M1}+\beta_{M1})\right]X_{M1t}+\left[1-(\alpha_{M2}+\beta_{M2})\right]X_{M2t}+ \\ & +\left[1-(\alpha_{C1}+\beta_{C1})\right]X_{C1t}+\left[1-(\alpha_{C2}+\beta_{C2})\right]X_{C2t}+ \\ & +\left[1-(\alpha_{I1}+\beta_{I1})\right]X_{I1t}+\left[1-(\alpha_{I2}+\beta_{I2})\right]X_{I2t}. \end{aligned} \tag{8}$$

Each year country is endowed with certain number N_t of the emission permits and its trajectory is determined by the following relationship:

$$N_t = f_N(t, N_{td}), \quad t = 1,\ldots,T, \tag{9}$$

where N_{td} denotes the yearly number of the emission permits in the last considered period. Two variants of the function N_t considered in this paper are presented in figure 1d. The mild variant assumes decreasing numbers of the emission permits till 2030, after which it attains steady value of 57% of the 2005 emission level, and the restrictive variant with decreasing numbers of the emission permits till 2050, after which it attains steady value of 45% of the 2005 emission level.

Disposable income DI_t equals the defined above gross income GI_t, decreased/ increased by the debt servicing/income from foreign assets:

$$DI_t = GI_t - r \cdot D_{t-1} + P(N_t - S_t), \tag{10}$$

where:
r – interest rate;
D_t – foreign debt (if positive)/ foreign assets (if negative) at the end of the year t:

$$D_t = D_{t-1} - (F_{Et} + F_{Mt} + F_{Ct} + F_{It}), \tag{11}$$

where P stands for the price of the emission permit, N_t denotes the number of the emission permits in the year t, defined above, and S_t denotes actual total emission:

$$S_t = S_{Et} + S_{Mt} + S_{Ct} + S_{It}. \tag{12}$$

Trade balance of all sectors (the sum in parentheses in (11)) increases debt if it is negative; and decreases debt if it is positive. Negative debt is interpreted as foreign assets, which in the year t generate an income equal to $- r \cdot D_{t-1}$. Note also that the excessive emission above the number of the emission permits has to be purchased in the international market at the emission unit price P, thus decreasing disposable income. In the opposite situation a country's disposable income is supplemented by the sale of the excessive emission permits in the international market.

Below, the balance equations for each sector are presented. The left hand sides of these equations denote domestic supply and the right hand sides represent domestic demand supplemented by the balances of foreign exchange in given good.

The balance equation of the *E* sector is expressed by the following equation:

$$
\begin{aligned}
X_{E1t} + X_{E2t} + X_{E3t} = \beta_{E1t}X_{E1t} + \beta_{E2t}X_{E2t} + \beta_{E3t}X_{E3t} + \beta_{M1t}X_{M1t} + \beta_{M2t}X_{E2t} + \\
+ \beta_{C1t}X_{C1t} + \beta_{C2t}X_{C2t} + \beta_{I1t}X_{I1t} + \beta_{I2t}X_{I2t} + \rho_t \lambda DI_t + F_{Et},
\end{aligned}
\tag{13}
$$

where the term

$$
\begin{aligned}
&\beta_{E1t}X_{E1t} + \beta_{E2t}X_{E2t} + \beta_{E3t}X_{E3t} + \beta_{M1t}X_{M1t} + \beta_{M2t}X_{E2t} + \beta_{C1t}X_{C1t} + \\
&+ \beta_{C2t}X_{C2t} + \beta_{I1t}X_{I1t} + \beta_{I2t}X_{I2t}
\end{aligned}
$$

denotes consumption of energy in year t in the sectors M, E, C, I; using all technologies available in those sectors, and the term F_{Et} stands for the net balance of the foreign trade of the sector E (if $EXP_{Et} - IMP_{Et} = F_{Et} \geq 0$, then export EXP_{Et} exceeds import IMP_{Et} in the foreign trade of goods produced by the sector E, and if $F_{Et} < 0$ then import IMP_{Et} exceeds export EXP_{Et} in the foreign trade in energy). The term $\rho_t DI_t$, $0 < \rho_t \leq 1$, denotes part of the disposable income DI_t in the year t designed for the purchases of the consumption goods, of which $\lambda\rho_t DI_t$ stands for the part of the total consumption expenditures directed for the purchases of energy. Note that the part $(1 - \rho_t)DI_t$ of the disposable income equals the total investment expenditures. Coefficient ρ_t is not a constant as it depends on the propensity to invest. Constant coefficient λ, $0 < \lambda \leq 1$, denotes assumed constant share of the energy expenditures in the total consumption expenditures.

Supply of goods produced by the sector M is supplemented by import, while some part of its output can be directed to export. The gross output of the sector M is distributed in the way expressed by the following balance equation:

$$
\begin{aligned}
X_{Mt} = \alpha_{M1}X_{M1t} + \alpha_{M2}X_{M2t} + \alpha_{E1}X_{E1t} + \alpha_{E2}X_{E2t} + \alpha_{E3t}X_{E3t} + \\
+ \alpha_{C1}X_{C1t} + \alpha_{C2}X_{C2t} + \alpha_{I1}X_{I1t} + \alpha_{I2}X_{I2t} + F_{Mt}, \\
t = 1,\ldots,T;
\end{aligned}
\tag{14}
$$

where the term

$$
\begin{aligned}
&\alpha_{M1}X_{M1t} + \alpha_{M2}X_{M2t} + \alpha_{E1}X_{E1t} + \alpha_{E2}X_{E2t} + \alpha_{E3t}X_{E3t} + \alpha_{C1}X_{C1t} + \\
&+ \alpha_{C2}X_{C2t} + \alpha_{I1}X_{I1t} + \alpha_{I2}X_{I2t}
\end{aligned}
$$

denotes consumption of the non-energy intermediate inputs in year t in the sectors M, E, C, I, and F_{Mt} stands for the net balance of the foreign trade of the sector M

($EXP_{Mt} - IMP_{Mt} = F_{Mt} \geq 0$ means that export EXP_{Mt} exceeds import IMP_{Mt} in the foreign trade of goods produced by the sector M, and when $F_{Mt} < 0$, the opposite).

Supply of goods produced by the sector I is supplemented by import, while some part of its output can be directed to export. The gross output of the sector I is distributed as described by the following balance equation:

$$X_{It} = I_t + F_{It}, \quad t = 1,\ldots,T; \tag{15}$$

where the term I_t

$$I_t = (1-\rho_t)DI_t = I_{M1t} + I_{M2t} + I_{E1t} + I_{E2t} + I_{E3t} + I_{C1t} + I_{C2t} + I_{I1t} + I_{I2t}$$

denotes total investment in the sectors M, E, C, I, and all technologies in year t, and F_{It} stands for the net balance of the foreign trade of the sector I (if $EXP_{It} - IMP_{It} = F_{It} \geq 0$, export EXP_{It} exceeds import IMP_{It} in the foreign trade of goods produced by the sector I, and if $F_{It} < 0$, the opposite).

Supply of goods produced by the sector C is supplemented by import, while some part of its output can be directed to export. The balance equation of the sector C is as follows:

$$X_{Ct} = \rho_t \cdot (1 - \lambda) \cdot DI_t + F_{Ct}, \quad t = 1,\ldots,T; \tag{16}$$

showing that the domestic supply (left-hand side of the above equation) of the non--energy consumption goods is equal to the demand generated by the part of the disposable income directed at purchasing non-energy consumption goods and the balance of the foreign trade in those goods (right hand side of the equation (16)). It is worth noting that the variable ρ_t can be interpreted as the propensity to consume. The term F_{Ct} stands for the net balance of the foreign trade of the sector C (if $EXP_{Ct} - IMP_{Ct} = F_{Ct} \geq 0$, export EXP_{Ct} exceeds import IMP_{Ct} in the foreign trade of goods produced by the sector C, and if $F_{Ct} < 0$, the opposite).

Households and the public sector belong to the same sector called the consuming sector, where decisions being made concern: utilization of the production capacities in sectors and technologies; distribution of the disposable income between consumption and investment; technology choice; and the role of the foreign trade. Constant proportion between the household and public consumption is assumed.

Decision variables of the model include: the actual gross outputs in sectors and technologies; investment in the capital assets in sectors and technologies; and the foreign trade balances of all production sectors:

$$\begin{gathered} X_{E1t}, X_{E2t}, X_{E3t}, X_{M1t}, X_{M2t}\ X_{C1t}, X_{C2t}\ X_{I1t}, X_{I2t}, I_{E1t}, I_{E2t}, I_{E3t}, \\ I_{M1t}, I_{M2t}\ I_{C1t}, I_{C2t}\ I_{I1t}, I_{I2t}, F_{Et}, F_{Mt}, F_{Ct}, F_{It}. \end{gathered} \tag{17}$$

The inequality constraints are as follows.

Non-negative outputs and investments:

$$X_{E1t}, X_{E2t}, X_{E3t}, X_{M1t}, X_{M2t}\ X_{C1t}, X_{C2t}\ X_{I1t}, X_{I2t}, I_{E1t}, I_{E2t}, I_{E3t}, \\ I_{M1t}, I_{M2t}\ I_{C1t}, I_{C2t}\ I_{I1t}, I_{I2t} \geq 0. \tag{18}$$

Note that the foreign trade balances $F_{Et}, F_{Mt}, F_{Ct}, F_{It}$ can be either positive or negative.

Propensity to invest, defined as a ratio I_t / DI_t, cannot exceed the maximum propensity to invest:

$$I_t / DI_t \leq \sigma_{I / DI}, \tag{19}$$

where $\sigma_{I / DI}$ denotes the maximum value of the investment to income ratio.

The above constraint reflects social resistance to the exceedingly high propensity to invest. The propensity to consume ρ_t is also constrained from beneath:

$$\rho_t \leq \sigma_{cons / DI}, \tag{20}$$

where coefficient $\sigma_{cons / DI}$ denotes the minimum value of the consumption to income ratio.

Another set of constraints deals with the feasible shares of foreign trade in the output of sectors. The following constraints:

$$\sigma_{IMP / X} \leq \frac{IMP_t}{X_t} \leq \sigma_{IMP / X},\ j = M, E, C, I; \tag{21}$$

$$\sigma_{EXP / X} \leq \frac{EXP_t}{X_t} \leq \sigma_{EXP / X},\ j = M, E, C, I; \tag{22}$$

impose maximum proportion of import and export respectively, in the national supply of the given product, where coefficients $\sigma_{IMP / X}$ and $\sigma_{EXP / X}$, $j = M, E, C, I$; denote respectively the maximum ratio of import and export of a given product to its national gross output.

The following two constraints:

$$-r^{(-)}_{INV\, i\, j} \leq \frac{I_{i\, j\, t} - I_{i\, j\, t-1}}{I_{i\, j\, t-1}} \leq r^{(+)}_{INV\, i\, j},\ j = 1, 2, 3; j = M, E, C, I; \tag{23}$$

$$-r^{(-)}_{cons} \leq \frac{\rho_t DI_t - \rho_{t-1} DI_{t-1}}{\rho_{t-1} DI_{t-1}} \leq r^{(+)}_{cons}, \tag{24}$$

limit relative increases and decreases of investments in sectors and total consumption, respectively, where parameters $r^{(-)}_{INV\, i\, j}$ and $r^{(+)}_{INV\, i\, j}$ stand for the lowest and highest admissible rate of increase of the investment in technology j, j = 1, 2, 3; $i = M, E, C, I$; while $r_{cons}^{(-)}$ and $r_{cons}^{(+)}$ denote the lowest and highest admissible rate of the consumption change respectively. In particular, the constraint (24) reflects social sensitivity to the changes in consumption and a possible resistance to them.

The following constraint reflects policy decisions concerning the desired share of a certain technology in the total output of a certain sector. In the current version of the model this constraint is the consequence of the requirement that in the energy sector the share of the renewable technology should be at least equal to 20% from the year 2030:

$$\frac{X_{E3t}}{X_{E1t}+X_{E2t}+X_{E3t}} \geq 20\%;\ t \geq 2030. \tag{25}$$

The last constraint limits the possibility of the excessive debt/credit relative to gross income

$$-0.60 \cdot GI_t \leq D_t \leq 0.60 \cdot GI_t. \tag{26}$$

Macroeconomic goal of economic development

The overall goal of the economic development, which is considered in this paper, is maximization of the discounted future consumption given by the following expression:

$$PVC = \sum_{t=t_0}^{T} \rho_t DI_t (1+r_d)^{-(t-t_0)} \tag{27}$$

subject to the constraints (1)–(26), where r_d denotes the discounting rate and $\rho_t DI_t$, $t = t_0, t_0+1, t_0+2, \ldots, T$, denote future consumption rates (note that the total consumption in the year t is equal to $\rho_t DI_t$).

Another tool worth considering is the multicriteria optimization, which aims at the harmonization of two conflicting objectives: maximization of the discounted future consumption and minimization of the cumulated GHG emissions. Such an approach was applied in Gadomski et al. (2016), and is suitable in the negotiations or training.

Data

In order to perform computations it was necessary to transform available data into a relevant form. The main source of the data was the Head Statistical Office (2011).

The method of reaggregation of the original input-output table was as follows. The energy sector E has been created by aggregating the following products: (i) Coal and lignite; (ii) Crude petroleum and natural gas; (iii) Coke, refined petroleum products; (iv) Electricity, gas, steam and air conditioning. Product of the sector E is interpreted further as the energy produced for the needs of the sector E and all other sectors, as well as tradable goods in the foreign trade. Products of other sectors were classified respectively as: M – the non-energy intermediary inputs in other production sectors, C – non-energy goods used in the consuming sector (consisting of households and the public sector), and I – investment goods serving for creation of the fixed assets exploited in the production sectors. The structure of the end uses of goods served also as a structure for decomposition of exports and imports of the original sectors. The new sectors were obtained by summing up all similarly classified parts of the original sectors; the same procedure was used in determining the exports and imports of the new sectors.

The initial values of variables were taken from the reaggregated input-output table and data concerning fixed assets.

In particular, the productivities of the fixed assets were estimated on the basis of the input-output data and the additional assumption that the utilization rates in sectors equaled 90%.

3. SIMULATION RESULTS

Two types of the simulation scenarios have been considered. The first one, called the static one, is based on the assumption that the number of available technologies in each sector is given and that they do not evolve. The second type is also based on the assumption that the number of available technologies in each sector is given and that there exists a technical progress, which improves technology parameters.

In each type of the simulation scenario two variants are considered. The first one (mild variant) assuming that the number of the emission allowances from the year 2030 on settles at the level of 57% of the initial emission level in 2010. The second (restrictive variant) assumes further reduction of the number of allowances from the level of 57% of the initial emission level in 2010 achieved from 2030 to 2050, when it settles at the level of 45% of the 2010 level.

Static scenarios

In all simulation scenarios a simplifying assumption has been adopted that before 2010 only old technologies had been in use so that the choice of technology starts in 2010. Also the initial level of foreign debt has been assumed to be equal to zero (simulation results were insensitive to that quantity). In all variants, solutions of the model converged to the steady state so that it was sufficient to present the development of variables till 2070.

The development of GDP, consumption, investment and emissions paired with relevant allowances are presented in figure 1.

In all sectors but sector I (having negligibly low emissions in both technologies), new technologies replaced old ones in the investment outlays. It is necessary to note that in the energy sector the most expensive technology has been chosen (the one interpreted as the renewable). This can be explained by the severity of the end-period emission constraints. However, because of the volatility of supply from this source of energy, it is worth considering additional constraint setting the maximum share of the third technology in the total energy output.

A necessity to adjust to the lowest emission levels at the end period forces the economic system to cumulate consumption at the beginning period, figure 1, panel (b), with the similar impact on investment, figure 1, panel (c), and GDP, figure 1, panel (a). As a result, after the initial growth period lasting to 2013, there comes recession and then stagnation, both determined by the low admissible level of emission.

Having in mind that the commented results were based on the assumption of fixed price relation and the absence of the technical progress, these results indicate that in such conditions it would be more effective to build considerable surplus in foreign trade, figure 4b, supporting the level of consumption in the end period.

As could be expected, investment and foreign exchange are the most volatile variables with variability concentrated in the beginning period.

The results described above explain the behavior of the economic system without the technical progress.

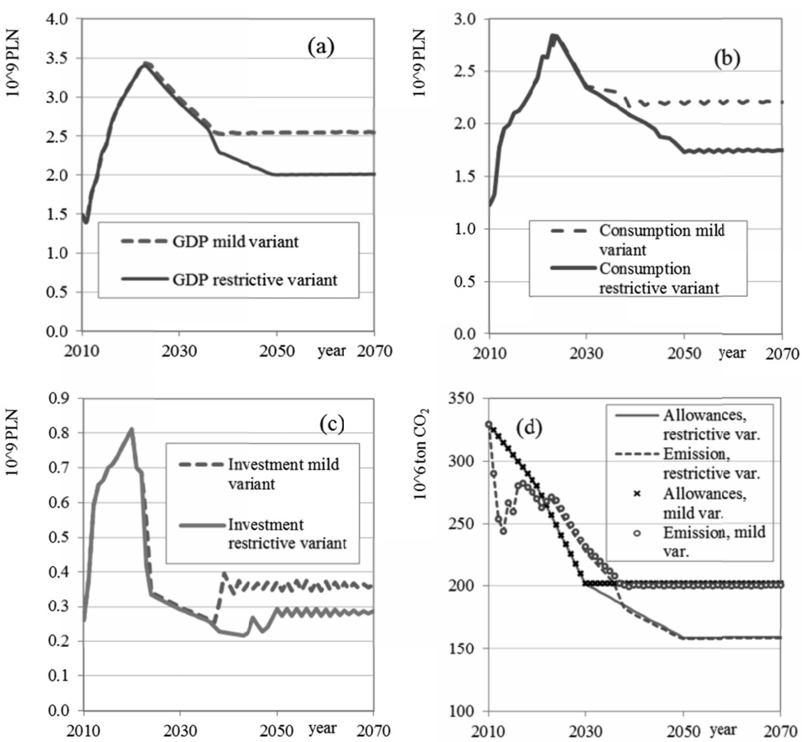


Figure 1. GDP in the mild and restrictive variants, panel (a), consumption in the mild and restrictive variants, panel (b), investment in the mild and restrictive variants, panel (c), emission allowances and emissions in the mild and restrictive variants, panel (d)

Source: own calculations.

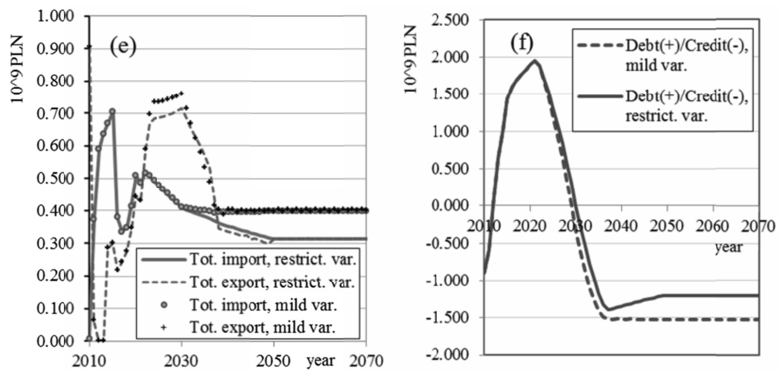


Figure 2. Total exports and imports, left hand panel, debt (if positive) or foreign assets (if negative), right hand panel, in the mild and restrictive variants

Source: own calculations.

4. CONCLUSIONS

Results obtained by the proposed model confirm its applicability in the analysis of the impact of the policy of curbing the GHG emissions on the national economy. This model should not be treated as the substitute but as a supplementary analytical tool used along the CGE models. One has to keep in mind the fact that the results are presented in constant prices, and that exogenous evolution in prices can be considered, given a credible scenario.

The technological conversion significantly affects the sectoral structure of the economy. The development of the shares of the gross output of each sector in the total gross output is presented in figure 3. One can observe that an increased share of the energy sector achieves the second position in the end period (not because of increased production but because of the high cost of the cleaner technologies).

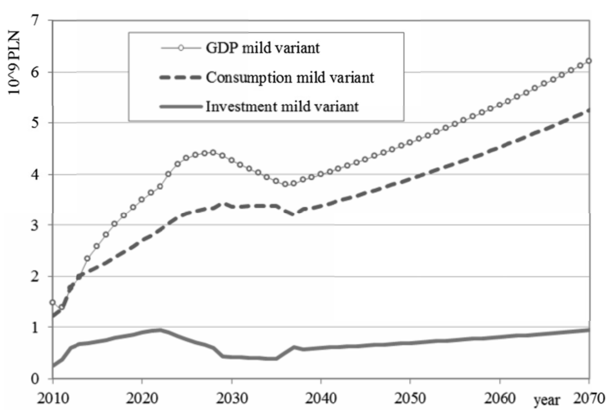


Figure 3. Development of GDP, consumption and investment according to the model with assumed yearly 1.5% decreases of the unit emissions of all technologies, mild variant, constant prices

Source: own calculations.

Economics would remain a dismal science, if the technical progress did not exist. In the presence of the technical progress expressed in the form of yearly 1.5% improvement (decrease) of the unit emission rates, main results with such technical progress accounted for are presented in figure 4.

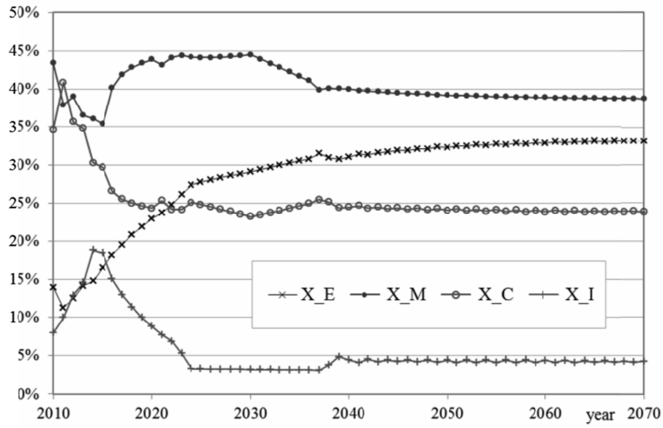


Figure 4. Changing shares of sectors in total output in constant prices
Source: own calculations.

The results presented in figure 4 show that the technical progress slightly extends the initial growth period, however it is also succeeded by a shorter recession period. Its depth is obviously determined by the rate of the technical progress.

REFERENCES

Antoszewski M., Boratyński J., Zachłod-Jelec M., Wójtowicz K., Cygler M., Jeszke R., Pyrka M., Sikora P., Böhringer C., Gąska J., Jorgensen E., Kąsek L., Kiuila O., Malarski R., Rabiega W. (2015), CGE Model PLACE, MF Working Paper Series, Ministry of Finance, Republic of Poland. No 22-2015.

Boratyński J., (2012), Historical Simulations with a Dynamic CGE Model: Results for an Emerging Economy, Ecomod 2012 Conference, Seville.

Central Statistical Office, Republic of Poland, (2011), Input-Output Table at Basic Prices in 2010.

Gadomski J., Kruś L., Nahorski Z., (2014), A Multicriteria Model for Analysis of the Impact of GHG Limiting Policies on Economic Growth. The Case of Poland. Systems Research Institute of PAS Working Paper No RB/37/2014. Warszawa.

Gadomski J., Nahorski Z., (2010), Impact of the Greenhouse Gases Limiting Policy on The Economic Growth and Structure, 69th International Atlantic Economic Society Conference, Prague, Czech Republic, 24–27 March 2010.

Roberts B. M., (1994), Calibration procedure and the Robustness of CGE Models: Simulations with a Model for Poland, *Economics of Planning*, 27 (3), 189–210.

OCENA WPŁYWU REDUKCJI EMISJI GAZOWYCH NA WZROST GOSPODARCZY POLSKI. ZAŁOŻENIA I WSTĘPNE WYNIKI

Streszczenie

W pracy przedstawiono model służący do oceny procesu konwersji technologicznej będącej następstwem ograniczania emisji gazów cieplarnianych. Limity emisji są wprowadzane w celu ograniczenia ocieplenia klimatu, czego skutkiem jest ograniczenie wzrostu gospodarczego. Konwersja technologiczna oznacza wybór czystszych, lecz ekonomicznie mniej sprawnych technologii. W rezultacie, długookresowy wzrost gospodarczy zmienia charakter: ze wzrostu względnie swobodnego ograniczonego przez dostępność czynników produkcji, zasobów oraz tempa postępu technicznego, na wzrost ograniczany ponadto przez dodatkowe ograniczenie – limit emisji. Analizę przeprowadzono przy pomocy modelu opartego na założeniach różniących się od stosowanych w budowie modeli CGE. Model składa się z następujących sektorów: a) konsumujący (obejmujący gospodarstwa domowe i sektor publiczny), b) wytwarzający dobra (z wyłączeniem energii) kupowane przez sektor konsumujący, c) wytwarzający nakłady pośrednie (bez energii) zużywane przez wszystkie sektory produkcyjne, d) wytwarzający energię zużywaną przez wszystkie sektory, e) wytwarzający dobra inwestycyjne kupowane przez wszystkie sektory produkcyjne. Wszystkie sektory produkcyjne realizują wspólny cel maksymalizacji zdyskontowanej wartości konsumpcji dla całego okresu optymalizacji, przy czym wielkości produkcji, inwestycje w poszczególne technologie w sektorach oraz salda wymiany zagranicznej stanowią zmienne decyzyjne. Model jest rozwiązywany jako zadanie optymalizacji liniowej. Rozwiązanie modelu jest traktowane jako wielkość referencyjna, nie obejmuje narzędzi polityki gospodarczej służących realizacji celu.

Słowa kluczowe: modelowanie ekonomiczne, polityka ekonomiczna, zmiana technologii, polityka ochrony środowiska

ASSESSMENT OF THE IMPACT OF THE REDUCTION OF THE GASEOUS EMISSIONS ON GROWTH IN POLAND. ASSUMPTIONS AND PRELIMINARY RESULTS

Abstract

The paper presents a model aimed at assessing the process of technology conversion imposed by limits of the greenhouse gas (GHG) emission. These limits are being introduced in order to stop climate warming, but by themselves they also inevitably curb economic growth. The change signifies choosing cleaner but economically less efficient technologies. In effect, the nature of the long-term economic growth is thus changed from a relatively free growth constrained by the availability of resources, production factors and technical progress, to that codetermined by the new constraint: the emission limit. The analysis is performed by using a model based on assumptions different from those applied in the CGE modelling. The model consists of the following sectors: a) consuming (both households and public); b) producing non-energy goods purchased by the consuming sector; c) producing intermediary non-energy inputs used in all producing sectors; d) producing energy consumed in all sectors; and e) producing investment (capital) goods purchased by all producing sectors. All economic agents pursue a common goal of achieving maximum total discounted consumption over the whole period of analysis, while the outputs in sectors and technologies, investment in sectors and technologies, as well as net foreign trade in sectors are decision variables. The model is solved using linear optimization. The model results constitute a benchmark; no economic tools are indicated for achieving the optimum.

Keywords: economic modelling, economic policy, technological change, environmental policy

DOI: 10.5604/01.3001.0014.1210

PRZEGLĄD STATYSTYCZNY
R. LXIII – ZESZYT 3 – 2016

JOANNA WYSZKOWSKA-KUNA[1]

FINANCIAL SERVICES INPUT AS A SOURCE OF ECONOMIC GROWTH IN THE EUROPEAN UNION COUNTRIES

1. INTRODUCTION

In the wide-ranging literature on the relationship between financial development and economic growth, different approaches can be identified with respect to the role of financial institutions and markets in stimulating economic growth. Lucas (1988, p. 6) dismissed finance as an "over-stressed" determinant of economic growth. At the other extreme, Miller (1998, p. 14) argued that "[the idea] that financial markets contribute to economic growth is a proposition too obvious for serious discussion." Between these two diametrically opposed approaches one can find three other lines of research: (1) Finance follows enterprises (Robinson, 1952, p. 86) – finance does not cause growth but responds to changing demands from the "real sector", so a faster economic development results in higher demand for financial services, which stimulates the development of financial institutions and markets (the demand-following view); (2) Financial development has a positive impact on economic growth, as credit is the basic source for enabling business, including innovative activities. Thus, a business cycle depends on financial activity (Fisher, 1933), and well-functioning banks support technological innovation by identifying those entrepreneurs who have the greatest chances of implementation of innovative products or processes (this approach was initiated by Schumpeter, 1912, and later developed by Minsky, 1982, 1990, as well as by a wide range of other research); (3) There are dynamic interactions between finance and growth, as the financial system influences growth, and growth transforms the operation of the financial system (the theoretical literature in this line of research is comparatively less well-developed).

An extensive survey of the literature can be found in Levine (2005)[2]. Based on different theoretical models he defined financial development[3] as involving improvements in financial functions that may influence savings and investment decisions and hence economic growth, i.e. in the (i) production of *ex ante* information about possible

---

[1] University of Lodz, Faculty of Sociology and Economics, Department of World Economy and European Integration, 41/43 Rewolucji 1905 St., 90-214 Łódź, Poland, e-mail: jkuna@uni.lodz.pl.

[2] A review of the existing literature can also be found in Kasprzak-Czelej (2010).

[3] It is measured by different indicators, among them: the ratio of credit to the private sector to GDP; the ratio of stock markets' size to GDP; the ratio of broad money to GDP; the margin between lending and deposit interest rates and the EBRD transition index of financial institutional development.

investments; (ii) monitoring of investments and the implementation of corporate governance; (iii) trading, diversification, and management of risk; (iv) mobilization and pooling of savings; and (v) exchange of goods and services. Summarizing the bulk of the existing research Levine stated that it is not just a question of finance following industry, but neither it is just industry following finance, which means that additional inquiry into the co-evolution of finance and growth is required.

In recent years some new empirical studies have proven the positive effect of financial development on economic growth in emerging markets (Africa: Ncube, 2007; India: Krishnan, 2011; North Africa: Kouki, 2013; Asia: Bayar, 2014; 42 emerging markets: Masoud, Hardaker, 2012 – bi-directional relations with respect to stock market; South Africa: Sunde, 2012 – bi-directional relations), as well as in economies after transition, i.e. the new EU member states (Caporale et al., 2014, 2015).

All these works examine the relationship between the development of financial institutions and markets and economic growth. However, to the best of the author's knowledge there are no studies on the impact of financial services input on output and productivity growth. This paper contributes to the research literature by presenting how the methodology of decomposition of output growth can be used to calculate the contributions of financial services input to gross output (GO) volume growth (in different industries and in the whole economy). What is worth stressing, this methodology can be also used to calculate the contributions of other components of intermediate input. This is shown in the paper, as FS input contribution is compared with the contribution of knowledge-intensive business services (KIBS), which have been already recognized as affecting output and productivity growth (to find out more on KIBS input contribution, see Wyszkowska-Kuna, 2016).

The goal of the paper is also to compare the results of the decomposition of GO volume growth for two periods: 1995–2007 and 2008–2009, to find out how the recent financial crisis affected economic growth in the EU countries, and how FS input contributed to the growth or decline in GO volume when the crisis started. For both periods I calculated the EU weighted averages for the results of the decomposition of GO volume growth, with the weights assigned based on each country's share in the total EU's GO. On the basis of the results of the decomposition of GO volume growth, one can also analyse whether and how FS input affects productivity.

Finally, one should note that the indicator proposed in this paper can be used in further research on the relation between financial services development and economic growth and productivity improvement.

2. REVIEW OF THE LITERATURE ON THE ROLE OF PRODUCER SERVICES IN ECONOMIC DEVELOPMENT

The division of services into intermediate and final was first introduced by Greenfield (1966, p. 11), and then developed by Browning, Singelman (1978, p. 489–90). Browning and Singelman distinguished two groups of intermediate services, i.e.:

(1) distributive services: transport and storage, communication, wholesale and retail trade and (2) producer services: financial services, insurance, real estate and business services.

In the literature one can find various papers studying the impact of services supporting economic activities on output and productivity growth in companies using these services. Stigler (1956) was the first to note that a company's development stimulates its demand for producer services, which in turn contributes to the development of external service providers. A decade later, Greenfield (1966, p. 11) noted that services input may have an impact on production conditions, comparable with those of the physical inputs.

Increased interest in the role of producer services has been visible only since the 1980s, but they were analysed in the context of final, not intermediate, consumption. This led to the belief that the economies where services dominates over industry and agriculture may experience slower growth in terms of output and productivity, because service activities have a lower potential for productivity growth than industrial and even agricultural activities (the model of unbalanced growth: Baumol, 1967; Baumol et al., 1989). Thus service prices may relatively increase,[4] which could limit demand for them and eventually also economic development (this phenomenon is called the "cost disease").

A new approach was presented by Oulton (2001, p. 606), who saw that demand for producer services has characteristics of intermediate consumption. Thus it should not decline in the long run, and what's more, if producer services contribute to output growth in companies using them, it should rather accelerate economic growth. Among the studies showing positive effects of producer services on output and productivity growth the following should be mentioned:[5] Windrum, Tomlinson (1998, 1999), Antonelli (1999, 2000), Tomlinson (2000), Katsoulacos, Tsounis (2000), Drejer (2002), Baláž (2003, 2004), Cagno di, Meliciani (2005), Baker (2007), Camacho, Rodriguez (2007), Desmarchelier et al. (2013), Wyszkowska-Kuna (2016). One should note, however, that none of these studies separately analysed the impact of financial services input on output and productivity growth.

3. METHODOLOGY

In order to assess the contribution of the various inputs to aggregate economic growth, the growth accounting framework can be applied. This methodology was theoretically motivated by Jorgenson, Griliches (1967) and put in a more general input-output framework by Jorgenson et al. (1987).

[4] A relative increase in service prices is a result of wage growth in service industries (not experiencing productivity growth) due to wage growth in other industries (experiencing productivity growth).

[5] Antonelli, Katsoulacos and Tsounis studied the impact of communications and business services; Drejer and Baker of business services; Camacho and Rodriguez of high-tech knowledge-intensive services (telecommunications, computer and R&D); and the others of aggregated values of communication, financial and business services.

The starting point for the analysis is production possibility frontiers, where industry gross output (GO) is a function of capital, labour, intermediate inputs and technology, which is indexed by time (T). Each industry (indexed by j) can produce a set of products and purchases a number of distinct intermediate inputs, capital and labour inputs to produce its output. The production function is given by:

$$Y_j = f_j(X_j, L_j, K_j, T), \tag{1}$$

where: Y – is output; X – is an index of intermediate inputs, either purchased from domestic industries or imported; L – is an index of labour service flows; K – is an index of capital service flows.

Output is expressed in producer prices, and the costs – in purchasers' prices. Under the assumptions of competitive factor markets, full input utilization and constant returns to scale, the growth of output in the period between any two discrete points of time, say t and t–1, can be expressed as the cost-share weighted growth of inputs and technological change A^Y (Jorgenson et al., 1987, p. 32–40; O'Mahony, Timmer, 2009, p. 376):

$$\Delta lnY_j = \bar{v}_j^X \Delta lnX_j + \bar{v}_j^L \Delta lnL_j + \bar{v}_j^K \Delta lnK_j + \Delta lnA_j^Y, \tag{2}$$

where v^i denotes the two period average share of input i in nominal output defined as follows:

$$\bar{v}_j^X = \frac{1}{2}\left[\frac{P_{jt}^X X_{jt}}{P_{jt}^Y Y_{jt}} + \frac{P_{jt-1}^X X_{jt-1}}{P_{jt-1}^Y Y_{jt-1}}\right], \tag{3}$$

$$\bar{v}_j^L = \frac{1}{2}\left[\frac{P_{jt}^L L_{jt}}{P_{jt}^Y Y_{jt}} + \frac{P_{jt-1}^L L_{jt-1}}{P_{jt-1}^Y Y_{jt-1}}\right], \tag{4}$$

$$\bar{v}_j^K = \frac{1}{2}\left[\frac{P_{jt}^K K_{jt}}{P_{jt}^Y Y_{jt}} + \frac{P_{jt-1}^K K_{jt-1}}{P_{jt-1}^Y Y_{jt-1}}\right], \tag{5}$$

and: $j = (1, 2, \ldots, n)$, and $\bar{v}^X + \bar{v}^L + \bar{v}^K = 1$.

Each element on the right side of equation (2) indicates the proportion of output growth accounted for by growth in intermediate inputs, capital services, labour services and technical change. Technical change is measured by total factor productivity (TFP).[6]

Jorgenson et al. (1987) pointed to the possibility of calculating the volume growth of labour, capital, and intermediate inputs with taking into account not only the volume growth (e.g. hours worked in the case of labour input), but also the changes in input's composition (e.g. in hours worked by different types of labour), which are

[6] Jorgenson et al. used the term "changes in productivity", whereas O'Mahony and Timmer "multifactor productivity", but they both mean the same as "total factor productivity".

referred to also as changes in the quality of input. Then the growth of output in the period between two points of time (t and t–1) is expressed also by equation (2), but the components ΔlnX_j, ΔlnL_j, ΔlnK_j have the following form (Jorgenson et al., 1987, p. 92–94, 130–131, 160–161; O'Mahony, Timmer, 2009, p. 377):

$$\Delta lnX_j = \sum_x \overline{w}_{x,j}^X \Delta lnX_{x,j}, \tag{6}$$

$$\Delta lnL_j = \sum_l \overline{w}_{l,j}^L \Delta lnL_{l,j}, \tag{7}$$

$$\Delta lnK_j = \sum_k \overline{w}_{k,j}^K \Delta lnK_{k,j}, \tag{8}$$

where:

$$\overline{w}_{x,j}^X = \frac{1}{2}\left[\frac{P_{x,jt}^X X_{x,jt}}{\sum P_{x,jt}^X X_{x,jt}} + \frac{P_{x,jt-1}^X X_{x,jt-1}}{\sum P_{x,jt-1}^X X_{x,jt-1}}\right], \tag{9}$$

$$\overline{w}_{l,j}^L = \frac{1}{2}\left[\frac{P_{l,jt}^L L_{l,jt}}{\sum P_{l,jt}^L L_{l,jt}} + \frac{P_{l,jt-1}^L L_{l,jt-1}}{\sum P_{l,jt-1}^L L_{l,jt-1}}\right], \tag{10}$$

$$\overline{w}_{k,j}^K = \frac{1}{2}\left[\frac{P_{k,jt}^K K_{k,jt}}{\sum P_{k,jt}^K K_{k,jt}} + \frac{P_{k,jt-1}^K K_{k,jt-1}}{\sum P_{k,jt-1}^K K_{k,jt-1}}\right], \tag{11}$$

and: (j = 1, 2, …, n; $l/k/x$ = 1, 2, …, q).

Sectoral quality remains unchanged if all components of intermediate, labour and capital inputs within an industry j are growing at the same rate. Sectoral quality rises if components with higher productivity are growing more rapidly, otherwise quality falls.

Taking into account both these methods of decomposition of output growth, it is possible to allocate output growth not only to intermediate, labour and capital inputs, but also with respect to different components of these three main types of input. In the EU KLEMS database intermediate inputs are subdivided into three components: energy, materials and services. For the purpose of the present study financial services input (herein after called FS input) is split of services inputs and the decomposition of output growth is made also with the allocation into FS input contribution.

This method can be applied to the decomposition of output growth not only in each industry, but also with respect to total industries, as in the present study. To assign GO volume growth in the EU countries (WIOD, 2014) to the contributions of intermediate, labour, capital inputs and TFP, average annual growth rates of each input volume should first be calculated, and then they should be weighed by average shares of their costs in GO value.

Intermediate inputs (II) are calculated by summing firms' expenditures on all raw and manufacturing materials, as well as services (values are taken from input-output tables), while FS input is calculated by summing firms' expenditures on services purchased from three industries, i.e.: Financial intermediation services, except insurance and pension funding services (65 – industry codes according to NACE Rev. 1.1); Insurance and pension funding services, except compulsory social security services (66); and Services auxiliary to financial intermediation (67) (WIOD, 2013). To calculate the average annual growth rates of II and FS input volume, it is necessary to deflate the values of II and FS input components. II values are deflated by deflators for intermediate inputs, while the components of FS input (i.e. X_{65}, X_{66}, X_{67}) by deflators for GO for industries "Financial services" (65–67) (WIOD, 2014).[7] KIBS input (compared with FS input in figure 2) is calculated by summing firms' expenditures on services purchased from the following industries: Computer and related services – 72, Research and development services – 73; Other business services – 74) (Wyszkowska--Kuna, 2016, p. 82).

Labour input is the number of hours worked by persons engaged (WIOD, 2014). The category "persons engaged" is broader than the category "employees", because it includes, in addition to employees, self-employed workers (Timmer et al., 2007, p. 25).

Capital input is the value of real fixed capital assets in 1995 prices multiplied by the number of hours worked per person engaged (WIOD, 2014). The number of hours worked per person engaged is used as an indicator showing the shift-factor, i.e. the degree to which capital assets are used in the analysed period, depending on the economic situation.

Capital stocks have been constructed on the basis of the Perpetual Inventory Method (PIM) in which the capital stock (K) in year *t* is estimated as the sum of the depreciated capital stock in year *t*–1 plus real investment (I) in year *t*:

$$K_t = (1 - d)K_{t-1} + I_t \tag{12}$$

with *d* the depreciation rate. The depreciation rates are taken to be geometric and industry-specific (from less than 4% in e.g. Education and Public Administration to more than 10% in financial and business services) (Erumban et al., 2012, p. 6–7).

For the majority of the EU countries long time-series of investments are available and there is no need to have information on an initial stock estimate. However, for some countries (Bulgaria, Cyprus, Estonia, Latvia, Lithuania, Luxembourg, Malta, Poland, Romania and Slovak Republic) no investment data before 1995 was available, and thus the ICVAR method was used[8]. In the ICVAR method, industry specific ratios of value added to capital stocks were used of a country at a similar stage of development (often Spain). These industry-specific ratios (averaged over 5 years to smooth out business cycle fluctuations) were applied to the 1995 value added to derive the

[7] In the WIOD database (as in the EU KLEMS database) there is no data on the values of deflators for particular components of II. Thus, the components of FS input for total industries (i.e.: X_{65}, X_{66}, X_{67}) are deflated by GO deflator for industries 65–67, which have delivered FS input. The same method is applied to the KIBS input's deflation. One should also note some weaknesses in data showing the values of deflators, as the same values of deflator are used for industries 65, 66, 67, and 72, 73, 74. What's more there are some differences in the values of deflators in the WIOD and the EU KLEMS databases.

[8] Only in the case of Belgium the Harberger method was used (Erumban et al., 2012, p. 7).

1995 capital stock. For years after 1995 the PIM method was used based on this 1995 estimate (Erumban et al., 2012, p. 6–8).

Labour compensation is the compensation of all persons engaged, while capital compensation (WIOD, 2014) is derived as gross value added minus labour compensation (O'Mahony, Timmer, 2009, p. 380).

4. DATA SOURCES AND ANALYSED PERIOD

The data needed for the decomposition of GO volume growth are available in two databases, i.e.: the EU KLEMS and the WIOD, both developed by the European Commission as a part of the EU 7th Framework Programme. In the present study the WIOD data are used, due to the availability of data on capital investments for all the EU countries (in the EU KLEMS such data are available only for some of the EU countries) and of more recent data (the WIOD usually contains data till 2009, whereas the EU KLEMS only till 2007). Data on capital investments are available only till 2007, and therefore a complete decomposition of GO volume growth is possible only for the period 1995–2007, but for the next two years GO volume growth and the contribution of intermediate inputs, including financial services input, to this growth have been calculated. Analysis of the subsequent years is not possible due to the lack of relevant data.

The creation of the EU KLEMS and the WIOD databases gave the opportunity to work on more complete and comparable data between countries (O'Mahony, Timmer, 2009, p. 396), which has created new opportunities for research on the decomposition of output volume growth. However, one should keep in mind that in both cases the data for some years have been created by interpolation, and haven't been derived directly from statistical sources. Thus their completeness should be treated with a fairly significant degree of approximation, which leads to caution when interpreting the results of the studies based on them. One should also note the risk of lower reliability of data on service industries than on manufacturing industries. This is due to the fact that when constructing these databases a variety of additional data sources were used, which are generally less numerous and often more incomplete in the case of service industries (O'Mahony, Timmer, 2009, p. 390). Finally the problems with measuring service output, especially in areas such as financial or business services (O'Mahony, Timmer, 2009, p. 390–391), should be mentioned.

5. RESULTS OF THE DECOMPOSITION OF GROSS OUTPUT VOLUME GROWTH INCLUDING THE ALLOCATION INTO FINANCIAL SERVICES INPUT CONTRIBUTION TO THIS GROWTH

Table 1 shows average annual growth rates of GO volume in the period 1995–2007 for total industries in the EU countries (column 2) and their decomposition into the contributions of: labour inputs (column 4); capital inputs (column 5); intermediate inputs (II – column 6) and changes in TFP (column 3). For the purposes of the research conducted in the present paper, FS input contributions (column 7) were calculated as a part of II contributions. They have been calculated for aggregated values of FS input in each country, which means they do not include changes in the composition of FS input. Therefore, their values are not equal to summed values of: FIS input contribution (Financial intermediation services, except insurance and pension funding services input contribution – column 8); I&PFS input contribution (Insurance and pension funding services, except compulsory social security services input contribution – column 9), and SAtFI input contribution (Services auxiliary to financial intermediation input contribution – column 10), which include changes in the composition of FS input. The values of both FS input contributions are compared in figure 2.

Table 1.

Gross output volume growth[a] in 1995–2007, and its decomposition into the contributions of: labour, capital and intermediate inputs, including financial services input[b], and changes in TFP, in the EU countries

Country	GO	TFP	Labour input	Capital input	II	**FS input**	FIS input	I&PFS input	SAtFI input
AUT	3.52	0.71	0.28	0.37	2.16	**0.05**	-0.02	0.02	0.07
BEL	2.64	0.18	0.33	0.45	1.68	**0.07**	-0.07	0.03	0.14
DNK	3.18	0.29	0.36	0.40	2.13	**0.13**	0.09	0.01	0.03
FIN	4.53	1.17	0.38	0.35	2.63	**0.03**	0.01	0.01	0.02
FRA	3.24	0.84	0.19	0.21	1.99	**0.17**	0.07	0.03	0.08
DEU	2.35	0.60	-0.04	0.30	1.49	**0.07**	0.00	0.03	0.04
GBR	3.26	0.71	0.24	0.50	1.81	**0.09**	0.01	0.06	0.02
GRC	3.58	0.31	0.37	1.34	1.56	**0.28**	0.28	0.02	0.00
IRL	7.73	0.46	0.95	1.42	4.90	**0.57**	0.27	0.19	0.13
ITA	2.11	0.02	0.31	0.31	1.47	**0.11**	0.06	0.01	0.05
LUX	8.32	0.51	0.83	0.61	6.37	**5.43**	1.88	0.04	3.59
NLD	3.00	0.69	0.35	0.36	1.60	**0.07**	0.05	0.00	0.02
PRT	2.64	-0.03	0.25	0.95	1.48	**0.23**	0.17	0.04	0.03
ESP	4.17	0.14	0.91	0.63	2.49	**0.20**	0.04	0.06	0.13
SWE	3.35	0.93	0.20	0.52	1.70	**0.01**	-0.02	0.02	0.01
BGR	4.27	0.68	0.04	0.31	3.24	**0.05**	0.02	0.03	0.00
CYP	5.17	1.28	0.79	0.46	2.64	**0.36**	0.30	0.02	0.04
CZE	5.63	0.90	-0.03	0.45	4.31	**0.07**	0.04	0.02	0.01
EST	7.85	2.00	0.07	1.19	4.60	**0.37**	0.23	0.03	0.11
HUN	6.63	1.72	0.05	0.18	4.68	**-0.05**	-0.05	0.00	0.00
LAT	6.74	1.63	0.30	1.27	3.54	**0.18**	0.04	0.19	0.01

Country	GO	TFP	Labour input	Capital input	II	**FS input**	FIS input	I&PFS input	SAtFI input
LTU	5.76	1.39	0.23	1.55	2.60	**0.04**	0.03	0.00	0.01
MLT	3.73	0.51	0.21	0.58	2.43	**0.51**	0.43	0.04	0.31
POL	6.44	1.91	-0.07	0.30	4.30	**0.18**	0.12	0.02	0.04
ROU	4.49	0.75	0.08	0.57	3.09	**0.05**	0.03	0.01	0.01
SVK	6.88	1.34	-0.02	0.90	4.66	**-0.12**	-0.11	-0.02	0.01
SVN	4.61	1.16	0.05	0.64	2.76	**0.10**	0.05	0.04	0.01
EU[c]	3.12	0.57	0.24	0.40	1.91	**0.13**	0.04	0.03	0.06

[a] Average annual growth rate for total industries. [b] FS input contributions to GO volume growth (total and with respect to its components) have been calculated on the basis of formulas 6 and 9, while intermediate inputs contributions on the basis of formula 3. FS input contributions calculated for aggregated values of FS input in each country. [c] The EU(27) weighted average, with weights assigned based on each country's share in the EU's gross output.

Source: own calculations based on: WIOD, 2013, National Input-Output Tables: *Time Series Supply and Use Tables, Use Tables at Purchasers' Prices,* WIOD database; WIOD, 2014, *Basic Data on Output and Employment*, WIOD database.

The highest value of FS input contribution to GO volume growth, at much higher level than in any other EU country, took place in Luxembourg. FS input contribution amounted there to 5.43, which accounted for 85% of total II contribution and 62% of GO volume growth in this country, which means that FS input was by far the most important source of GO volume growth (the highest among the EU countries). However, one should note that Luxembourg is a special case – it is a small economy, specific in terms of its sectoral structure and position within the EU, recognized as a tax haven and an offshore financial centre (OFC), and characterised by very favourable regulations, political stability, financial security and its location in the centre of Europe (Tax Justice Network, 2007; Mainelli, Yeandle, 2007, 2009).[9] Therefore, it does not seem reasonable to compare Luxembourg with other EU countries.

The second highest value of FS input contribution to GO volume growth was reached by Ireland (0.57, however it was 9.5 times lower than in Luxembourg), followed by Malta (0.51), Estonia (0.37), Cyprus (0.36), Greece (0.28) and Portugal (0.23). Three of them (Malta, Cyprus and Ireland) have been also recognized as tax havens and OFCs.[10] Among the abovementioned countries only Ireland, Estonia and Cyprus recorded high rates of GO volume growth, which indicates that FS input was

[9] Tax havens are low-tax jurisdictions that provide investors with opportunities for tax avoidance or paying lower taxes (Desai et al., 2004, p. 1). OFCs are located in tax havens and they exploit the structures that can be created using the tax haven's legislation for the benefit of those residents elsewhere. They combine some of the following characteristics: a high number of financial institutions that mainly serve non-residents, financial systems out of proportion with the domestic economy`s need, low or no taxes, light financial supervision and regulation, flexible use of different company structures, and high levels of bank secrecy and anonymity (Levin, 2002, p. 2).

[10] In recent years Luxembourg and other EU countries perceived as tax heavens have taken some actions to change their image, which is in line with the EU policy to eliminate regulations supporting tax avoidance within its member states (Blomeyer, Sanz, 2013). However, the elimination of all differ-

an important, but not the main, source of GO volume growth. In Poland, FS input contribution also had a relatively high value (0.18), which was accompanied by a high rate of GO volume growth. In Slovakia and Hungary FS input contribution to GO volume growth recorded negative values, with relatively high rates of GO volume growth.

In the last row in table 1, the weighted averages for the EU(27) are presented, with the weights assigned based on each country's share in the EU's GO (in 1995 prices). They show that the EU average FS input contribution to GO volume growth was at a medium level (0.13), which accounted for 4.2% of the EU average GO volume growth. Eleven countries (Luxembourg, Ireland, Malta, Estonia, Cyprus, Portugal, Spain, Greece, Latvia, Poland and France) reached values above the EU(27) average.

In percentage terms (in relation to GO volume growth – figure 1) FS input most significantly contributed to GO volume growth (excepting Luxembourg) in Malta (13%), Portugal (approx. 8.5%), Greece and Ireland (almost 8%), and Cyprus (approx. 7%). In these countries, as well as in four other (Italy, France, Spain and Estonia), the importance of FS input contribution for GO volume growth was above the EU(27) average (4.2%).

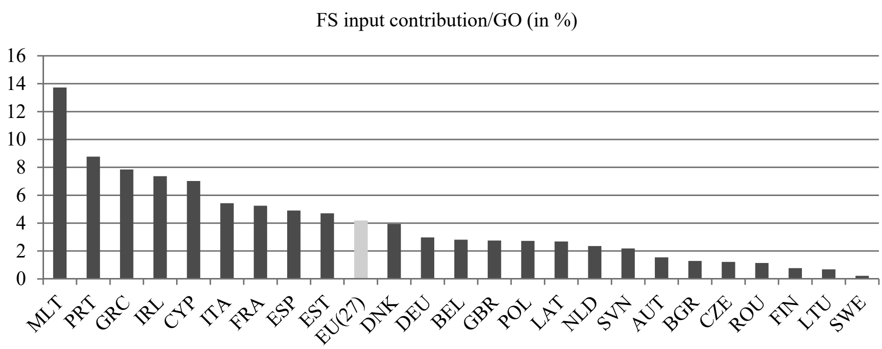


Figure 1. The ratio of financial services input contribution to gross output volume growth and gross output volume growth, in 1995–2007 in the EU countries[a]

[a] Except for Luxembourg (because of much higher value of the ratio in comparison with other EU countries), as well as except for Hungary and Slovakia (because of negative values of FS input contribution).

Source: own calculations based on the values of GO volume growth and FS input contribution from table 1.

Among the countries with a surprisingly high importance of FS input contribution to GO volume growth Portugal and Greece should be mentioned. Greece (similarly as Austria, Finland, France and Sweden) was recognized by the OECD as a potentially harmful tax regime, whereas Madeira, being a part of Portugal (similarly as Belgium, Frankfurt in Germany, Campione d'Italia & Trieste in Italy, the Netherlands and Hungary) were recognized as tax havens, although none of them was recognized

ences in tax regulations is not possible, and thus some EU countries remain more attractive for foreign businesses than others (Parietti, 2016).

as an OFC. This may lead to the conclusion that the importance of FS input contribution may also depend on some other factors, e.g. the level of competition on the market (the methodology used in the paper assumes perfect competition), or some others. One should also bear in mind that there may also be some differences between the countries covered by the study in the quality of relevant data, which may have an impact on these results. Thus it seems advisable to continue research in this field in order to identify the factors that determine the importance of FS input contribution in different countries.

The countries with high FS input contribution usually recorded TFP change on the medium level (except for Cyprus and Poland). On the contrary, relatively high growth of TFP can be noticed in Slovakia and Hungary.

In figure 2 there are values of FS input contributions calculated in two ways: (1) for aggregated values of FS input in each country (FS input1 – as in table 1) and (2) for summed values of the contribution of each type of FS input – i.e. summed values of the contribution of: FIS input, I&PFS input, and SAtFI input (FS input2). The values of FS input2 contributions include changes in the composition of FS input (Jorgenson et al., 1987). In the case of those countries where higher values were reached for FS input2 contribution, one can speak of positive changes in the composition (quality) of FS input. These positive changes are a result of a relative increase in the importance of new products based on more advanced technologies and knowledge, which in turn results in their higher productivity. The highest differences between the two values (26 percentage points – pp) are visible in Malta, where changes in the composition are due to the high increase in SAtFI input contribution. It should be noted that in Malta these services recorded a very low value of GO (0.002 million) in the base year, which later resulted in its very high average annual growth rate (the increase to 29 million euro meant that average annual growth rate was 125%). Large differences are also visible in Luxembourg (9 pp; changes in the composition due to the increasing importance of FIS and SAtFI inputs contribution), Latvia (6 pp, changes in the composition due to the increaseing importance of I&PFS input contribution and to a lesser degree of SAtFI), and Spain (4 pp, changes in the composition due to the increasing importance of SAtFI and I&PFS input contribution).

For comparison, the values of KIBS input contribution are presented in figure 2. FS input contribution was generally lower than KIBS input contribution, with the exception of Luxembourg (where FS input contribution was 6 times higher than KIBS input contribution), Greece and Cyprus (more than twice higher), as well as Malta and Portugal.

In 2007–2008, most countries maintained GO volume growth and positive values of FS input contribution. The exceptions were Estonia, Ireland and Latvia, which recorded a decline in GO volume and negative values of FS input contribution. In turn, Luxembourg, Malta, and France recorded negative values of FS input contribution with GO volume growth (the opposite situation took place in the UK and Denmark, i.e. positive values of FS input contribution while GO volume declined). In 2008–2009, all countries, including Poland, recorded a decline in GO volume and all countries (except Bulgaria) negative values of II contribution. The highest negative values of FS input contribution can be noticed in Luxembourg (-2.75 in 2008 and -4.05 in 2009), and then in Ireland (-0.53 and -0.46), Estonia (-0.27 and -0.46) and Latvia (-0.27 and -0.34).

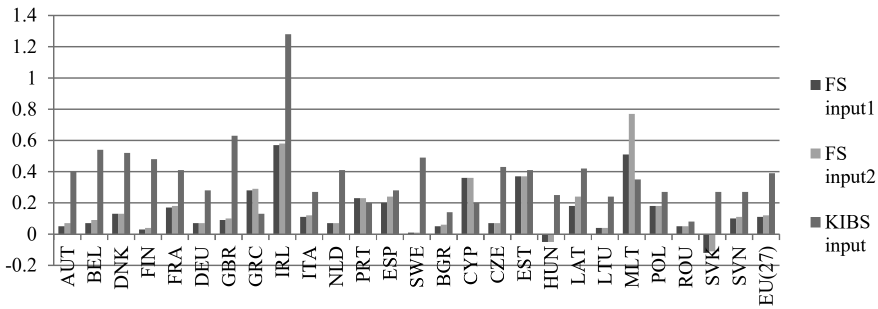


Figure 2. The contributions of financial services input and knowledge-intensive business services input to gross output volume growth, in 1995–2007 in the EU countries

FS input1 calculated for aggregated values of FS input. FS input2 calculated by summing the contribution of each type of FS input. EU(27) – the EU(27) weighted average, with weights assigned based on each country's share in the EU's gross output.

Source: own calculations based on the sources as in table 1.

Table 2.

Gross output volume growth in 2007–2008 and 2008–2009, and intermediate inputs contribution – including financial services input contribution – to this growth, in the EU countries

Country	GO	II	**FS input**	FIS input	I&PFS input	SAtFI input	GO	II	**FS input**	FIS input	I&PFS input	SAtFI input
	2007–2008						2008–2009					
AUT	2.67	1.99	**0.09**	0.06	0.02	0.01	-4.66	-2.68	**0.11**	0.06	0.01	0.04
BEL	1.13	0.60	**0.14**	0.03	0.02	0.08	-3.54	-2.35	**-0.13**	-0.04	-0.03	-0.07
DNK	-0.02	0.04	**0.08**	0.06	0.01	0.01	-6.91	-4.65	**-0.23**	-0.17	-0.03	-0.02
FIN	2.11	1.78	**0.01**	0.00	0.00	0.00	-9.24	-5.57	**0.05**	0.04	0.01	0.00
FRA	0.64	0.39	**-0.01**	-0.01	0.00	0.00	-4.99	-3.69	**-0.44**	-0.26	-0.06	-0.12
DEU	0.73	0.21	**0.07**	0.03	0.02	0.02	-7.37	-4.74	**-0.09**	-0.04	-0.03	-0.01
GBR	-0.50	-0.45	**0.15**	0.60	-0.21	-0.12	-5.25	-2.85	**-0.28**	-0.18	-0.06	-0.03
GRC	0.54	-0.36	**0.28**	0.27	0.03	-0.02	-3.61	-2.73	**0.22**	0.26	-0.02	-0.02
IRL	-3.37	-2.44	**-0.53**	-0.26	-0.18	-0.09	-5.06	-3.04	**-0.46**	-0.26	-0.15	-0.05
ITA	-1.87	-1.34	**0.04**	0.01	0.00	0.02	-8.19	-5.62	**-0.17**	-0.12	-0.02	-0.04
LUX	0.22	-0.27	**-2.75**	-1.59	-0.10	-1.04	-6.47	-5.30	**-4.05**	-1.62	-0.11	-2.32

Country	GO	II	**FS input**	FIS input	I&PFS input	SAtFI input	GO	II	**FS input**	FIS input	I&PFS input	SAtFI input
	2007–2008						2008–2009					
NLD	2.10	1.13	**0.25**	0.16	0.03	0.06	-4.18	-2.73	**-0.07**	-0.04	-0.02	-0.02
PRT	0.12	-0.20	**0.03**	0.03	0.00	0.01	-4.21	-2.88	**0.06**	0.05	0.01	0.01
ESP	0.45	-0.05	**0.02**	0.01	0.00	0.00	-5.72	-4.07	**-0.13**	-0.09	-0.03	0.00
SWE	0.06	0.25	**0.03**	0.02	0.01	0.00	-8.16	-5.60	**-0.02**	-0.01	0.00	0.00
BGR	2.41	0.84	**0.70**	0.47	0.20	0.03	-5.03	0.68	**0.42**	0.31	0.09	0.02
CYP	6.42	4.42	**0.42**	0.33	0.04	0.05	-2.88	-1.93	**-0.11**	-0.10	-0.01	-0.01
CZE	3.25	2.10	**0.07**	0.05	0.01	0.02	-7.92	-6.23	**0.00**	0.00	0.00	0.00
EST	-5.53	-3.68	**-0.27**	-0.15	-0.03	-0.09	-17.23	-10.71	**-0.46**	-0.29	-0.05	-0.11
HUN	2.13	1.68	**0.05**	0.03	0.01	0.01	-11.74	-8.78	**-0.01**	-0.01	0.00	-0.01
LAT	-2.99	-1.78	**-0.27**	-0.14	-0.12	-0.01	-17.22	-10.08	**-0.34**	-0.16	-0.17	-0.01
LTU	7.72	6.38	**0.23**	0.17	0.04	0.02	-20.36	-12.79	**0.07**	0.06	0.00	0.01
MLT	3.61	0.61	**-0.29**	-0.23	-0.04	-0.02	-4.45	-2.76	**0.74**	0.61	0.07	0.06
POL	5.10	2.93	**0.25**	0.15	0.03	0.07	-3.39	-4.07	**-0.27**	-0.17	-0.04	-0.06
ROU	8.72	5.20	**0.34**	0.16	0.06	0.11	-6.04	-3.05	**-0.16**	-0.07	-0.04	-0.05
SVK	7.18	4.66	**0.08**	0.04	0.02	0.02	-9.68	-7.78	**-0.10**	-0.03	-0.04	-0.02
SVN	3.06	1.60	**0.04**	0.03	0.01	0.00	-10.96	-7.34	**-0.05**	-0.04	-0.01	0.00
EU[a]	0.42	0.13	**0.06**	0.10	-0.02	-0.01	-6.23	-4.17	**-0.20**	-0.11	-0.04	-0.05

[a] The EU(27) weighted average, with weights assigned based on each country's share in the EU's gross output.
Source: own calculations based on the sources as in table 1.

It should be noted, that generally FS input only marginally contributed to the decline in GO volume in the EU countries. In 2007–2008, the EU(27) average value of FS input contribution decreased less than the EU(27) average GO volume, and in a result 14% of GO volume growth could be assigned to FS input contribution. The following year, when the EU (27) GO volume declined, the EU(27) average FS input contribution to this decline accounted only for 3%. The analysis at a country level also shows that in countries recording the highest decline in their output negative values of FS input contribution were relatively low, and interestingly in Lithuania, where GO declined the most (-20.4%), the contribution of FS input was positive (a similar situation took place in several other countries, i.e. in Austria, Bulgaria, Finland, Greece, Malta and Portugal, and a particularly high positive value of FS input contribution with very high output decline took place in Malta). On the contrary, Luxembourg, Ireland, France, Poland and the United Kingdom recorded relatively high negative values of FS input contribution in relation to the decline in GO volume.

In countries with negative values of FS input contribution all its components were negative. In both periods the most important contribution to both GO volume growth and decline can usually be assigned to FIS input, then to I&PFS input, and finally to SAtFI input.

In table 3 the results of more standard economic growth accounting methods are presented to compare them with the results of the decomposition of GO volume growth (table 1). In 1995–2007, all the EU countries recorded a growth of total value added (VA) and value added in Financial intermediation (VAFI)[11]. In most countries the growth rates of VAFI were higher than that of VA. Only in Finland, Germany, Hungary, Latvia, Lithuania, Romania and Slovakia was the situation reversed, and Hungary was the only country where VAFI declined. The highest growth rates of VAFI took place in Estonia, Poland and Cyprus, and in the case of these countries one can note the highest differences in growth rates of both values. As far as the shares of VAFI in VA are concerned, the highest values were reached by Luxembourg (23%), followed by Portugal (9%), Cyprus and Belgium, Great Britain, Ireland (8%) and Austria (7%), whereas the lowest shares of VAFI in VA were recorded by Slovakia and Hungary (2%). Finally, the ratio of intermediate consumption of FI services (ICFI) to the global output of this sector (GOFI) shows the extent to which FI services constituted intermediate input, and the extent to which they constituted final output, in each country. The ratio was the highest in Luxembourg (77%), followed by Germany, but with Germany's index being lower by 20 percentage points. In other countries the ratio ranged between 50% (Great Britain, France) and 27% (Romania, Cyprus).

Table 3.

The importance of value added and intermediate consumption of Financial intermediation services, in 1995–2007 and 2007–2009 in the EU countries

Country	VA(G)[a]	VAFI(G)[b]	VAFI(S)[c]	ICFI(R)[d]	VA(G)[a]	VAFI(G)[b]	VAFI(S)[c]	ICFI(R)[d]
	1995–2007				2007–2009			
AUT	2.60	5.98	7.16	35.02	-1.33	9.28	9.56	30.99
BEL	2.21	4.45	7.62	41.21	-0.75	-2.07	8.06	43.71
DNK	1.99	7.67	6.63	35.34	-2.38	-1.32	10.12	30.62
FIN	3.89	2.03	3.58	39.61	-3.81	3.20	3.61	45.52
FRA	2.21	3.22	4.85	50.72	-1.02	1.39	5.33	52.65
DEU	1.68	0.57	4.34	57.27	-2.11	1.94	4.11	65.05
GBR	2.91	5.38	7.54	50.98	-2.45	-1.44	8.68	52.30
GRC	3.75	5.41	4.35	30.89	0.09	7.76	5.17	30.19
IRL	6.90	8.17	7.56	47.05	-3.52	-4.54	8.77	53.88

[11] "Financial intermediation" is the name of section J comprising all financial divisions (65–67). The terms "Financial intermediation services" and "FI services" refer to all services delivered by this section.

Country	VA(G)[a]	VAFI(G)[b]	VAFI(S)[c]	ICFI(R)[d]	VA(G)[a]	VAFI(G)[b]	VAFI(S)[c]	ICFI(R)[d]
	1995–2007				2007–2009			
ITA	1.42	3.15	4.98	38.89	-3.32	-0.83	5.92	38.96
LUX	4.92	6.08	22.68	77.41	-0.97	-5.06	23.49	81.97
NLD	2.83	3.94	6.67	47.07	-0.44	2.70	7.55	47.64
PRT	2.63	8.54	9.22	32.27	-1.06	1.74	12.67	31.45
ESP	3.59	6.48	4.97	37.01	-1.22	-2.09	6.54	36.15
SWE	3.30	4.42	4.95	31.72	-2.95	0.51	5.31	30.37
BGR	2.33	8.06	3.47	36.53	-0.24	17.34	8.11	39.67
CYP	3.69	10.00	8.31	26.59	0.98	3.78	11.10	24.94
CZE	3.25	7.23	4.13	51.60	-0.77	14.09	5.59	42.68
EST	7.39	23.26	5.78	45.18	-9.51	-16.90	10.23	43.92
HUN	3.51	-3.41	2.41	46.01	-3.34	1.84	1.79	49.49
LAT	7.35	6.62	4.34	35.15	-9.21	-10.92	4.30	32.91
LTU	6.61	4.78	1.91	39.11	-5.91	-2.13	1.75	43.89
MLT	2.95	3.77	4.69	43.90	1.60	17.63	4.30	63.36
POL	4.38	13.80	4.75	40.63	3.47	-6.55	6.50	45.90
ROU	3.09	2.75	6.59	27.32	0.57	1.76	6.78	36.55
SVK	5.09	-3.67	2.22	42.79	0.97	4.50	1.39	46.58
SVN	4.47	8.92	6.91	31.53	-2.37	6.67	10.25	23.21
EU[e]	2.44	3.74	5.54	47.18	-1.82	0.72	6.25	49.41

[a] The average growth rates of gross value added (VA). [b] The average growth rates of VA in Financial intermediation. [c] The average shares of VA in Financial intermediation in total VA (in %). [d] The average ratios of intermediate consumption of Financial intermediation services and gross output of Financial intermediation sector (in %). [e] The EU(27) weighted average, with weights assigned based on each country's share in the EU's gross output.

Source: own calculations based on the sources as in table 1.

In 2007–2009, most countries recorded a decline in VA, but only 11 experienced a decline in VAFI. The highest decline in VAFI took place in Estonia (-17%) and Latvia (-11%), whereas some countries maintained high growth rates of VAFI (Malta and Bulgaria +17% and Czech Republic +14%). In 2007–2009, the share of VAFI in VA generally increased in comparison with the period 1995–2007 (it declined only in Germany, Hungary, Latvia, Lithuania). The same can be said about the ratio of ICFI and GOFI, but in this case more countries (eleven) experienced decline, with the greatest decline taking place in the Czech Republic and Slovenia.

One can note that most countries with the highest shares of VAFI in VA (Austria, Belgium, Denmark, Great Britain, the Netherlands, Portugal, Romania and Slovenia) recorded relatively low values of FS input contribution to GO volume growth, as well as of the ratio of ICFI to GOFI (except Great Britain, where it was above the EU(27) average). Based on this it can be concluded that FI services were to a greater extent final output, not intermediate input, in these countries. The same can be said about Cyprus, where FS input contribution was relatively high, but the ratio of ICFI to GOFI reached the lowest value. The opposite situation took place in Luxembourg, where FI services were mainly intermediate input, as well as in Ireland and Estonia, although to a lesser extent than in Luxembourg. In Poland, FS input contribution to GO volume growth, as well as the share of VAFI in VA were both above the EU(27) average, whereas the ratio of ICFI to GOFI was below the EU(27) average.

In 1995–2007, in the EU countries GO volume growth and FS input contribution to this growth were positively correlated with each other, as the correlation coefficient for both variables achieved a value of 0.43. In 2007–2008, the correlation between the analysed variables decreased to 0.31, and in the following year it vanished (0.006). For the entire analysed period there was no correlation between FS input contribution to GO volume growth and TFP. The estimation of the regression equation shows that the relationship between FS input contribution to GO volume growth and GO volume growth in the period 1995–2007 was bi-directional. It should be noted, however, that FS input is a part of GO (it is a part of total production costs), therefore its growth automatically leads to an increase in GO. The share of FS input in total costs, however, is small, so the direct impact of FS input volume growth on GO volume growth is also low. If, therefore, there is a correlation between GO volume growth and FS input contribution to this growth, it can be assumed that the role of FS input in driving GO volume growth is greater than is apparent from its small share in GO.

6. CONCLUSIONS

1. Previous research examined the relationship between financial development and economic growth, but there are no studies on the impact of financial services input on output and productivity growth. The literature review shows that services input should be treated as a contribution to output growth in the same way as raw materials and manufacturing inputs.

2. The methodology of decomposition of GO volume growth, implemented by Jorgenson et al., and the availability of data in the WIOD database (as well as in the EU KLEMS database) has made it possible to calculate the contributions of different components of intermediate inputs to GO volume growth. This indicator captures both the size and the dynamics of intermediate expenditures and it can be used in further research studying the impact of FS input on output and productivity growth.

3. In 1995–2007, all the EU countries recorded GO volume growth and almost all (except for Slovakia and Hungary) had positive values for FS input contribution. In most countries the growth rates of VA in Financial intermediation were higher than of total VA, with Hungary being the only one country where VA in Financial intermediation declined. In 2008–2009, all the EU countries recorded a decline in GO volume (some already in 2007–2008) and usually negative values of FS input contribution, but only a few countries experienced a decrease in VA in Financial intermediation. As a result, the share of VA in Financial intermediation in total VA, as well as the ratio

of intermediate consumption of financial services to GO in Financial intermediation, both increased in most countries.

4. In 1995–2007, the EU weighted average FS input contribution to GO volume growth reached a medium value (0.13), which accounted for 4.2% of the EU weighted average GO volume growth. When the crisis started the values of FS input contribution decreased less than GO volume.

5. FS input was by far the main source of GO volume growth, and later decline in Luxembourg. Among the other EU countries, the importance of FS input to GO volume growth was much lower, although Malta, Estonia and Cyprus stood out. The EU policy to remove favourable tax regulations among its members may decrease the GO growth and the FS input contribution to this growth in European offshore financial centres, but some differences between countries will probably remain, although of a lower scale.

6. In Luxembourg, as well as in Ireland and Estonia, Financial intermediation services were mainly intermediate input, whereas in other countries where they recorded their highest contribution to value added they were final output to a larger extent.

7. In the entire group of EU countries a positive correlation between GO volume growth and FS input contribution to this growth was found, and this relation have appeared to be bi-directional. It should be noted, however, that while FS input contributed positively to GO volume growth, it had no significant impact on GO volume decline. In both periods covered by the study, FS input had no impact on productivity growth.

8. In general, the most important contribution to GO volume growth can be assigned to FIS input, then to SAtFI input and finally to I&PFS input. When the world financial crisis began FIS input contributed most to GO volume decline, but I&PFS input had higher contribution to this decline than SAtFI input.

REFERENCES

Antonelli C., (1998), Localized Technological Change, New Information Technology and the Knowledge-based Economy: the European Evidence, *Journal of Evolutionary Economics*, 8 (2), 177–198.

Antonelli C., (2000), New Information Technology and Localized Technological Change in the Knowledge Based Economy, in: Boden M., Miles I., (eds.), *Services and the Knowledge-based Economy*, Continuum, London, 170–191.

Baker D., (2007), The Impact of Business-services Use on Client Industries: Evidence from Input Output Data, in: Rubalcaba L., Kox H., (eds.), *Business Services in European Economic Growth*, Palgrave MacMillan, New York, 97–115.

Baláž V., (2003), Knowledge Intensive Business Services in a Transition Economy, *Ekonomický časopis*, 51 (4), 475–488.

Baláž V., (2004), Patterns of Intermediate Consumption and Productivity in the Knowledge Intensive Services in Transition Economies, *Ekonomický časopis*, 52 (3), 298–314.

Baumol W. J., (1967), Macroeconomics of Unbalanced Growth: the Anatomy of Urban Crisis, *The American Economic Review*, 57 (3), 415–426.

Baumol W. J., Blackman S. A. B., Wolff E. N., (1989), *Productivity and American Leadership: The Long View*, MIT Press, Cambridge.

Bayar Y., (2014), Financial Development and Economic Growth in Emerging Asian Countries, *Asian Social Science*, 10 (9), 8–17.

Blomeyer & Sanz, (2013), *European Initiatives on Eliminating Tax Havens and Offshore Financial Transactions and the Impact of these Constructions on the Union's own Resources and Budget*, European Union, Brussels, http://www.europarl.europa.eu/RegData/etudes/etudes/join/2013/490673/IPOL-JOIN_ET(2013)490673_EN.pdf, (31.05.2016).

Browning H., Singelmann J., (1978), The Transformation of the U. S. Labor Force: The Interaction of Industry and Occupation, *Politics and Society*, 8 (3–4), 481–509.

Cagno di D., Meliciani V., (2005), Do Inter-sectoral Flows of Services Matter for Productivity Growth? An Input/Output Analysis of OECD Countries, *Economics of Innovation and New Technology*, 14 (3), 149–171.

Camacho J. A., Rodriguez M., (2007), Integration and Diffusion of KIS for Industry Performance, in: Rubalcaba L., Kox H., (eds.), *Business Services in European Economic Growth*, Palgrave MacMillan, New York, 128–143.

Caporale G. M., Rault C., Sova A. D., Sova R., (2014) Financial Development and its Effects on Economic Growth: A Dynamic Analysis, in: Arouri M., Boubaker S., Nguyen D. K., *Emerging Markets and the Global Economy: A Handbook*, Academic Press, Oxford (UK), Walthman (USA), San Diego (USA), 811–824, Science Direct, (15.05.2016).

Caporale G. M., Rault C., Sova A. D., Sova R., (2015), Financial Development and Economic Growth: Evidence from 10 New European Union Members, *International Journal of Finance and Economics*, 20 (1), 48–60.

Desai M. A., Foley C. F., Hines J. R., (2004), Economic Effects of Regional Tax Havens, *NBER Working Paper Series*, 10806, http://www.nber.org/papers/w10806, (31.05.2016).

Desmarchelier B., Djellal F., Gallouj F., (2013), Knowledge Intensive Business Services and Long Term Growth, *Structural Change and Economic Dynamics*, 25 (C), 188–205.

Drejer I., (2002), *Business Services as a Production Factor*, Paper presented at the DRUID Conference on "Industrial Dynamics of the New and Old Economy – Who is Embracing Whom?", Copenhagen-Elsinore, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.197.2243&rep=rep1&type=pdf, (10.11.2013).

Erumban A. A., Gouma R., de Vries G., de Vries K., Timmer M., (2012), *WIOD Socio-Economic Accounts (SEA): Sources and Methods*, http://www.wiod.org/publications/source_docs/SEA_Sources.pdf, (13.07.2016).

Fisher I., (1933), The Debt-Deflation Theory of Great Depressions, *Econometrica*, 1(4), 337–357.

Greenfield H. I., (1966), *Manpower and the Growth of Produces Services*, Columbia University Press, New York.

Jorgenson D. W., Griliches Z., (1967), The Explanation of Productivity Change, *The Review of Economic Studies*, 34 (3), 249–283.

Jorgenson D., Gollop F. M., Fraumeni B., (1987), *Productivity and U.S. Economic Growth*, Harvard University Press, Cambridge, MA.

Kasprzak-Czelej A., (2010), Wpływ rozwoju system finansowego na wzrost gospodarczy, *Annales Universitatis Mariae Curie-Skłodowska*, Lublin, XLIV(2), 17–30, http://docplayer.pl/12102395-Wplyw-rozwoju-systemu-finansowego-na-wzrost-gospodarczy.html, (10.05.2016).

Katsoulacos Y., Tsounis N., (2000), Knowledge-intensive Business Services and Productivity Growth: the Greek Evidence, in: Boden M., Miles I., *Services and the Knowledge-based Economy*, Continuum, London-New York, 192–208.

Kouki I., (2013), Financial Development and Economic Growth in the North African Region, *African Development Review*, 25 (4), 551–562.

Krishnan K. P., (2011), Financial Development in Emerging Markets: The Indian Experience, in: Kawai M., Prasad E. S., (eds.), *Financial Market Regulation and Reforms in Emerging Markets*, Asian Development Banks, Tokyo and Brookings Institution Press, Washington, 226–262.

Levin M., (2002), *The Prospects for Offshore Financial Centres in Europe*, Centre for European Policy Studies, https://www.ceps.eu/publications/prospects-offshore-financial-centres-europe, (10.05.2016).

Levine R., (2005), *Finance and Growth: Theory and Evidence*, in: Aghion P., Durlauf S. N., (eds.), *Handbook of Economic Growth*, ELSEVIER, Amsterdam, 865–934, http://faculty.haas.berkeley.edu/ross_levine/papers/Forth_Book_Durlauf_FinNGrowth.pdf, (10.05.2016).

Lucas R. E., (1988), On the Mechanics of Economic Development, *Journal of Monetary Economics*, 22 (1), 3–42.

Mainelli M., Yeandle M., (2007, 2009), *The Global Financial Centres Index*, Z/Yen Group for the City of London, http://www.zyen.com/PDF/GFCI_1_March_2007.pdf, (15.05.2016).

Masoud N., Hardaker D., (2012), The Impact of Financial Development on Economic Growth: Empirical Analysis of Emerging Market Countries, *Studies in Economics and Finance*, 29 (3), 148–173.

Miller M. H., (1998), Financial Markets and Economic Growth, *Journal of Applied Corporate Finance*, 11 (3), 8–15.

Minsky H. P., (1982), Can "It" Happen Again? Essays on Instability and Finance, *Hyman P. Minsky Archive*, Paper 314, http://digitalcommons.bard.edu/hm_archive/314, (2.08.2016).

Minsky H. P., (1990), Schumpeter: Finance and Evolution, in: Heertje A., Perlman M., (eds.), *Evolving Technology and Market Structure: Studies in Schumpeterian Economics*, the University of Michigan Press, Ann Arbor, MI, 51–76.

Ncube M., (2007), Financial Services and Economic Development in Africa, *Journal of African Economies*, 16 (1), 13–57.

O'Mahony M., Timmer M. P., (2009), Output, Input and Productivity Measures at the Industry Level: the EU KLEMS Database, *Economic Journal*, 119 (538), 374–403.

Oulton N., (2001), Must the Growth Rate Decline?, *Oxford Economic Papers*, 53 (4), 605–627.

Parietti M., (2016), *The Top 10 European Tax Havens*, 4.04.2016, Investopedia, http://www.investopedia.com/articles/wealth-management/121515/top-10-european-tax-havens.asp, (1.06.2016).

Robinson J., (1952), The Generalization of the General Theory, in: *The Rate of Interest and Other Essays*, MacMillan, London.

Schumpeter J. A., (1912, 1934), *The Theory of Economic Development* (translated by R. Opie), Harvard University Press, Cambridge, MA.

Sunde T., (2012), Financial Sector Development and Economic Growth Nexus in South Africa, *International Journal of Monetary Economics and Finance*, 5 (1), 64–75.

Tax Justice Network, (2007), *Identifying Tax Havens and Offshore Financial Centres*, http://www.taxjustice.net/cms/upload/pdf/Identifying_Tax_Havens_Jul_07.pdf, (31.05.2016).

Timmer M., Moergastel van T., Stuivenwold E., Ypma G., O'Mahony M., Kangasniemi M., (2007), *EU KLEMS Growth and Productivity Accounts*, *Part I Methodology*, European Commission, Luxembourg.

Tomlinson M., (2000), Information and Technology Flows from the Service Sector: a UK–Japan Comparison, in: Boden M., Miles I., (eds.), *Services and the Knowledge-based Economy*, Continuum, London, 209–221.

Windrum P., Tomlinson M., (1998), *The Impact of KIBS on International Competitiveness: a UK-Netherlands Comparison*, SI4S Topical Paper, 10, http://survey.nifu.no/step/old/Projectarea/si4s/papers/topical/si4s10.pdf, (3.02.2012).

Windrum P., Tomlinson M., (1999), Knowledge-intensive Services and International Competitiveness: a Four Country Comparison, *Technology Analysis and Strategic Management*, 11 (3), 391–408.

WIOD, (2013), National Input-Output Tables: *Time Series Supply and Use Tables, Use Tables at Purchasers' Prices*, WIOD database, www.wiod.org, (5.11.2013).

WIOD, (2014), Socio-Economic Accounts: *Basic Data on Output and Employment*, WIOD Database, www.wiod.org, (1.12.2014).

Wyszkowska-Kuna J., (2016), *Usługi biznesowe oparte na wiedzy. Wpływ na konkurencyjność gospodarki na przykładzie wybranych krajów Unii Europejskiej*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

WYDATKI PRZEDSIĘBIORSTW NA USŁUGI FINANSOWE JAKO ŹRÓDŁO WZROSTU GOSPODARCZEGO W KRAJACH UNII EUROPEJSKIEJ

Streszczenie

Celem pracy jest zbadanie i porównanie znaczenia wydatków przedsiębiorstw na usługi finansowe dla wzrostu produkcji w krajach Unii Europejskiej. W badaniu wykorzystano metodę dekompozycji wzrostu produkcji według Jorgensona et al. (1987), która zakłada, iż zmiany produkcji wynikają ze zmian wielkości wydatków przedsiębiorstw na zakup surowców, materiałów, usług i czynników produkcji (pracy i kapitału) oraz łącznej produktywności czynników produkcji. Zaletą tej metody jest możliwość obliczenia wkładów wydatków na zakup materiałów lub usług (ogółem lub dla poszczególnych kategorii) we wzrosty produkcji w całej gospodarce oraz w poszczególnych działach. Badanie przeprowadzono w odniesieniu do usług finansowych, jednakże znaczenie usług finansowych dla wzrostu gospodarczego porównano ze znaczeniem usług biznesowych opartych na wiedzy, które postrzegane są jako mające wpływ na wzrost produkcji i produktywności. Dane wykorzystane w badaniu pochodzą z WIOD (World Input-Output Database). Okres badawczy to lata 1995–2009, z uwagi na dostępność danych.

Słowa kluczowe: usługi finansowe, wzrost gospodarczy, dekompozycja wzrostu produkcji, Unia Europejska

FINANCIAL SERVICES INPUT AS A SOURCE OF ECONOMIC GROWTH IN THE EUROPEAN UNION COUNTRIES

Abstract

The aim of this paper is to study and compare the importance of intermediate demand for financial services for the growth of production in the European Union countries. In the study the methodology introduced by Jorgenson et al. (1987) is used. This assumes that changes in the production (in real terms) result from changes in intermediate inputs of raw and manufacturing materials and services, as well as in factor inputs (labour and capital) and in total factor productivity. The advantage of this method is the ability to calculate the contributions of different components of intermediate inputs (including service inputs – total or with respect to particular service categories) to production growth in the whole economy and in individual industries. The study is carried out with respect to financial services, but their contribution to economic growth is compared with the contribution of knowledge-intensive business services that have been already recognized as affecting economic and productivity growth. The data used in the study come from the World Input-Output Database. The analysed period covers the years 1995–2009, owing to the availability of relevant data.

Keywords: financial services, economic growth, the decomposition of economic growth, European Union

DOI: 10.5604/01.3001.0014.1211

PRZEGLĄD STATYSTYCZNY
R. LXIII – ZESZYT 3 – 2016

ALICJA WOLNY-DOMINIAK[1]

THE COPULA-BASED TOTAL CLAIM AMOUNT REGRESSION MODEL WITH AN UNOBSERVED RISK FACTOR

1. INTRODUCTION

The basic characteristic of an insurance portfolio is its heterogeneity, which means that individual risks generate different claim amounts. In view of this, assigning a single premium to each risk is unfair. Therefore, a common practice of any insurance company is ratemaking, which is defined as the process of classification of the risk portfolio into risk groups where the same premium corresponds to each risk. The grouping is done based on what is referred to as risk factors, which cause the portfolio homogeneity. The risk factors may be divided into:

– observed factors (observed at the conclusion of an insurance contract) – these are the factors that describe an insured person and an insurance subject, as well as a spatial variable (in the sense of the geographical region),
– unobserved factors – such as a driver's skills, the safety of a district where a property is located, a factor specific to each risk treated as a random variable with a certain distribution.

The current practice of insurance companies is to carry out ratemaking in two stages determined by the risk factors that are taken into consideration (cf. Dionne, 1989). The first stage is *a priori* ratemaking, which means dividing the risk portfolio into groups of risks that are homogeneous in terms of the observed factors. Then *a posteriori* ratemaking is carried out, when the unobserved risk factor is taken into account individually for each risk.

The ratemaking problem comes down to determining a premium for a homogeneous risk group, where a premium is understood as the expected total claim amount for a single risk. In the estimation, two separate models – the average value of claims (called a claim severity model) and the number of claims (called a claim frequency model) – are applied to a single risk. Due to the character of risk portfolios and insurance data, a common practice applied by insurance companies is to use generalized linearized models (GLM's – cf. De Jong, Heller, 2008; Frees, 2009; Ohlsson, Johansson, 2010; Antonio, Valdez, 2012; Wolny-Dominiak, Trzpiot, 2013; Wolny-

[1] University of Economics in Katowice, Faculty of Economics, Department of Statistical and Mathematical Methods in Economics, 50 1th May St., 40-287 Katowice, Poland, e-mail: woali@ue.katowice.pl.

-Dominiak, 2014). Owing to the progress in numerical algorithms for finding maximum values of the log-likelihood function and their numerical implementation in commercial and non-commercial software, GLM's have become a common practice in the Polish insurance market as well.

The above approach to ratemaking requires the independence between an average value of claims and the number of claims. The reason for this is that the expected total claim amount is understood as the product of the expected claim frequency and the expected claims severity. However, in the literature this assumption is called into question, as in Krämer et al. (2013) or Shi et al. (2015). The dependence between two random variables is accommodated by the copula and the authors propose a copula-based regression model in order to estimate the total claim amount. The interest of this paper is to extend this model taking into account an unobservable risk factor in the claim frequency model. This factor, called also unobserved heterogeneity, is treated as a random variable influencing the number of claims. Typically, in such a situation a mixed Poisson distribution is assumed, but for our purposes we propose to apply the zero-truncated distribution. The goal is then to estimate the expected value of the product of two random variables: the average value of claims and the number of claims for a single risk assuming the dependence between the average value of claims and the number of claims for a single risk and the dependence between the number of claims for a single risk and the unobservable risk factor. In the model, we construct the bivariate distribution, which gives us the opportunity to estimate this expected value using the Monte Carlo (MC) simulation.

In the paper we give the details of the theoretical aspects of the model as well as the empirical example. To acquaint the reader with the model operation, every step of the process of the expected value estimation is described and the **R** code is available for download (see http://web.ue.katowice.pl/woali/ and R code Team, 2014).

2. TOTAL CLAIM AMOUNT MODEL UNDER INDEPENDENCE

A starting point for *a priori* ratemaking is the total claim amount model, in which the random variables – the average value of claims and the number of claims for a single risk – are independent. Consider a portfolio of n property risks where the risk is understood as a random variable with a certain distribution, hereinafter denoted as S_i, $i = 1,\dots,n$, representing the total claim amount for the i-th risk. If the number of claims for the i-th risk in the portfolio is marked as N_i and if i denotes the value of a single claim, the variable S_i may be expressed in the following form:

$$S_i = Y_{i1} + \dots + Y_{iN_i},\ S_i = 0 \text{ if } N_i = 0. \tag{1}$$

The considerations presented below take into account only the risks for which at least one claim has occurred. Assuming that variables $Y_{i1},\dots,Y_{iN_i}$, are independent and have identical distributions, and that they are independent of N_i, the expected value and the variance of variable S_i may be expressed as follows:

$$E[S_i] = E[Y_i N_i] = E[Y_i]E[N_i], \qquad (2)$$
$$Var[S_i] = E^2[Y_i]Var[N_i] + E[N_i]Var[Y_i].$$

The expected value $E[S_i]$ corresponds to the so-called *pure premium* for a single risk. This is the premium covering the risk, without any additional costs of insurance. If an insurance company has a mass portfolio of risks, which is the case for example in motor third part liability (MTPL) and motor own damage (MOD) insurance or in immovable property insurance, the claim frequency model and the claims severity model are used to estimate the pure premium $E[S_i]$. The parameters of the models are estimated using data included in insurance policies. This practice is described in detail in works authored by, for example, De Jong, Heller (2008), Frees (2009), Ohlsson, Johansson (2010), Cizek et al. (2011).

Modelling the total claim amount (not the pure premium) for a single risk, the following assumptions are commonly made in this approach:

– In the claim frequency model the number of claims for a single risk has the Poisson distribution $N_i \sim Pois(\lambda_i)$,
– In the claim severity model variables Y_{ik} have identical distributions coming from the exponential dispersion family of distributions with the same dispersion parameter $Y_i \sim EDM(\mu_i, \varphi_Y)$.

The heterogeneity of an insurance portfolio is described by regression coefficients introduced to the mean of both models:

$$\mu_i = \exp(\mathbf{x}_i^Y \boldsymbol{\beta}^Y), \ \lambda_i = E_i \exp(\mathbf{x}_i^N \boldsymbol{\beta}^N), \qquad (3)$$

where $\boldsymbol{\beta}^Y = (\beta_0^Y, \beta_1^Y, ..., \beta_k^Y)^T$, $\boldsymbol{\beta}^N = (\beta_0^N, \beta_1^N, ..., \beta_k^N)^T$ are fixed-effect vectors corresponding with observed risk factors; $\mathbf{x}_i^Y$, $\mathbf{x}_i^N$ are i-th rows of the matrix of models $\mathbf{X}^Y$ and $\mathbf{X}^N$, respectively. E_i denotes the risk exposure (typically – the time of the policy duration). Then the total claim amount for a single risk is simply:

$$E[S_i] = \mu_i \lambda_i = E_i \exp(\mathbf{x}_i^Y \boldsymbol{\beta}^Y) \exp(\mathbf{x}_i^N \boldsymbol{\beta}^N). \qquad (4)$$

It should be noticed that if no claim has occurred for the i-th risk, the number of claims $N_i = 0$, which means, naturally, that the value of variable Y_i should also be zero. However, only the average claim non-zero value is assumed in the claims severity model. Therefore, the zero-truncated distribution of the number of claims is assumed in the case under analysis. Assuming the Poisson distribution for the number of claims, the probability mass function with deleted zero values has the following form:

$$\Pr^{ZTPois}(N_i = k_i \mid k_i > 0, \lambda_i) = \frac{\Pr^{Pois}(N_i = k_i \mid \lambda_i)}{1 - \Pr^{Pois}(N_i = 0 \mid \lambda_i)} = \frac{\lambda_i^{k_i}}{[\exp(\lambda_i) - 1]k_i!} \tag{5}$$

where $\Pr^{Pois}(N_i = 0 \mid \lambda_i) = \exp(-\lambda_i)$. The expected value and the variance are $E(N_i) = \frac{\lambda_i \exp(\lambda_i)}{\exp(\lambda_i) - 1}$ and $Var(N_i) = \frac{\lambda_i \exp(\lambda_i)}{\exp(\lambda_i) - 1}[1 - \frac{\lambda_i}{\exp(\lambda_i) - 1}]$ respectively (cf. Cruyff, van der Heijden, 2008).

The parameters of frequency and severity models are usually estimated separately, using the maximum likelihood method. Finally, the estimated value of the expected total claim amount is obtained in the point estimation by plugging in coefficient estimators into formula (4). The same strategy can be used with respect to the variance value taking the formula (2).

Example 1 – total claim amount model under independence

In order to demonstrate the current practice, the insurance portfolio taken from (Wolny-Dominiak, Trzesiok, 2014) is investigated herein. The data comes from the former Swedish insurance company Wasa and concerns partial *casco* insurance for motorcycles in the period of 1994–1998. The frequency and severity models are assumed as $Y_i \sim Gamma(\mu, \varphi_Y)$ and $N_i \sim ZTPois(\lambda)$ without regressors. We use the maximum likelihood method (MLE) in the estimation. The fitted parameters are presented in table 1 below.

Table 1.

Estimates of parameters in claim severity-frequency model

Model	Parameters	Mean	Variance
Severity	$\hat{\mu} = 25437$ $\hat{\varphi}_Y = 2.03$	$\hat{E}[Y_i] = 25437$	$\sqrt{V\hat{a}r[Y_i]} = 38651$
Frequency (without exposure)	$\hat{\lambda} = 0.04$	$\hat{E}[Y_i] = 1.024$	$\sqrt{V\hat{a}r[Y_i]} = 0.16$

Source: own calculations.

Plugging values from table 1 into formula (4), estimated characteristics of the total claim amount are obtained. The quantiles of $\hat{E}[S_i]$ are presented in figure 1, taking into account the exposure to each risk.

The left-hand figure displays quantiles of the order from 0 to 0.95, while the right-hand one – quantiles of the order from 0.95 to 1.

Insurance companies use the above-described practice only if an assumption is made that the claim amount value Y_i is independent of the claim number N_i for the risk. If this assumption is rejected, a dependence between variables has to be accommodated. And this could be done using a copula.

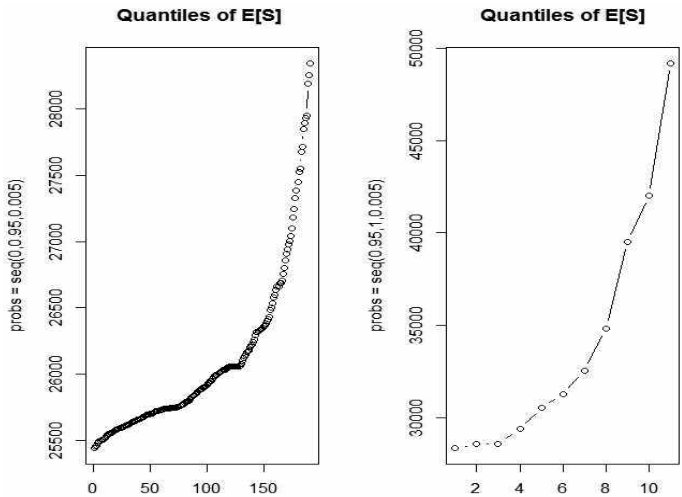


Figure 1. Quantiles of total claim amount
Source: own calculations.

3. DEPENDENCE WITH BIVARIATE COPULAS

The theory of copulas is frequently referred to in literature as in Joe (1997), Nielsen (1999), Wanat (2012). Here we give a short introduction for those who are not familiar with the subject. A bivariate copula $C(\cdot)$ is a two-dimensional cumulative distribution function (cdf) $C : [0, 1]\ x\ [0, 1] \to [0, 1]$ whose univariate margins are uniform on $[0, 1]$. For continuous random variables (X_1, X_2) with marginal cdf's $F_1(\cdot)$, $F_2(\cdot)$ and densities $f_1(\cdot)$, $f_2(\cdot)$, random variables of the form $U_1 = F_1(X_1)$, $U_2 = F_2(X_2)$ are also uniform on $[0, 1]$. According to the Sklar theorem (1959):

$$F_{X_1,X_2}(x_1, x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)). \tag{6}$$

Hence, the joint distribution $F(\cdot)$ is decomposed into marginal distributions and the copula $C(u_1, u_2)$, which captures the structure of the relation between X_1 and X_2. The corresponding joint density $f_{X_1,X_2}(\cdot)$ is then as follows:

$$f_{X_1,X_2}(x_1, x_2) = c(F_{X_1}(x_1), F_{X_2}(x_2)) f_{X_1}(x_1) f_{X_2}(x_2), \tag{7}$$

where $c(\cdot)$ denotes the copula density.

Generally, if a bivariate cdf of (X_1, X_2) exists, also a bivariate copula $C(\cdot)$ exists, and in the case of continuous random variables the copula is unique. However, the model proposed herein assumes mixed continuous and discrete variables.

Let us assume N is the count variable with a density function $f_N(\cdot)$ and consider a continuous-discrete random variable (Y, N). Let us focus on the parametric bivariate copula with one parameter θ, such as the Gauss, Clayton or Frank copulas, which separates the dependence structure from margins. Denoting the partial derivative of the copula with respect to variable Y as $D(u_1,u_2)=\frac{\partial}{\partial u_1}C(u_1,u_2)$, u_1, $u_2 \in (0, 1)$, according to the formula (7), as is shown in Krämer et al. (2013) in the case of mixed outcomes, the joint density function $f_{Y,N}(\cdot)$ may be expressed as follows:

$$f_{Y,N}(y,k)=f_Y(y)\big(D(F_Y(y),F_N(k))-D(F_Y(y),F_N(k-1))\big). \tag{8}$$

In order to construct the above density function, the parameter vector of marginal distributions has to be estimated as well as the copula parameter θ. The inference functions for margins (IFM) method is used in this paper. It consists in estimating univariate parameters from separately maximized univariate likelihoods, and then estimating the copula parameter θ. Like in the above-described formula (8), only the margin of the first variable appears as the proper log-likelihood function giving the estimated value of θ in the following form:

$$l(\theta)=\sum_{i=1}^{n}\log\big(D(F_Y(y_i),F_N(k_i))-D(F_Y(y_i),F_N(k_i-1))\big). \tag{9}$$

Hence, the IFM method consists of three main steps (A1):
1. obtaining estimates of the vector parameters of margins,
2. transforming (y_i, k_i) to (u_{1i}, u_{2i}) as
   $u_{1i}=F_Y(y_i\,|\,\varphi_Y)$, $u_{2i}=F_N(k_i\,|\,\varphi_N)$, $u_{3i}=F_N(k_i-1\,|\,\varphi_Y)$,
3. optimizing $l(\theta)=\sum_{i=1}^{n}\log\big(D(u_{1i},u_{2i})-D(u_{1i},u_{2i})\big)$.

The example below illustrates the construction of density function $f_{Y,N}(y, k)$ for different types of one-parameter copulas $C(\cdot|\theta)$.

Example 2 – copula-based bivariate density construction

This example makes use of simulated data. The margins are taken as: $Y \sim Gamma(\mu,\phi_Y)$ with a mean μ and a dispersion ϕ_Y and $N \sim Poisson(\lambda)$ with a mean λ. Data (y_i, k_i), $i = 1,\ldots,100$ are drawn from $Gamma(\mu = 300, \phi = 1.5)$ and $Poisson(\lambda = 1)$. Assuming the parameter vector of margins as (300, 1.5, 1), observations (y_i, k_i) are transformed into (u_{1i}, u_{2i}) in the following way:

$$u_{1i}=F_Y(y_i\,|\,300{,}1.5),\; u_{2i}=F_N(k_i\,|\,1),\; u_{3i}=F_N(k_i-1\,|\,1) \tag{10}$$

assuming that $k_i = 0$ for $k_i < 0$. Finally, the copula parameter θ is estimated using the Gauss and Frank copulas and the copula-based density function $f_{Y,N}(\cdot)$ is constructed.

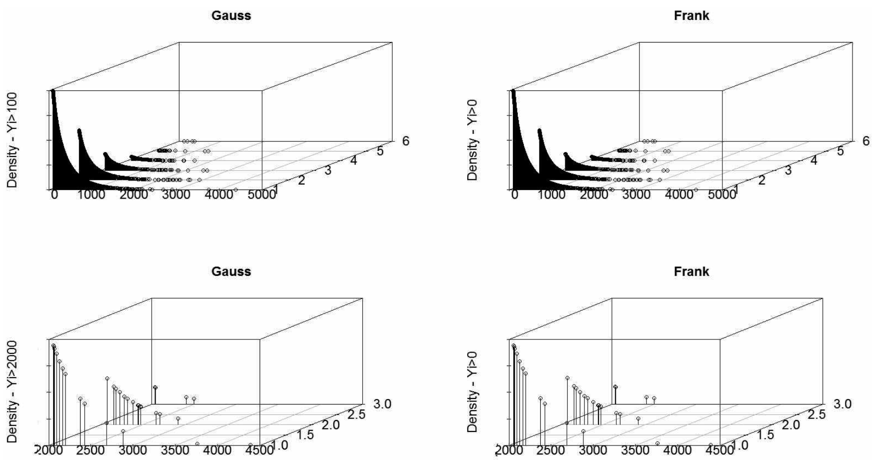


Fig. 2. Bivariate density of the random variable (Y_i, N_i)
Source: own calculations.

The IFM method is useful for models with the closure property of parameters being expressed in lower-dimensional margins. In addition, due to the fact that each inference function derives from a log-likelihood of a marginal distribution, the inference does not have to be obtained explicitly and numerical optimizations can be carried out for the log-likelihoods of margins. For this purpose, the BFGS algorithm implemented in **R** is used in this paper (see optim function).

4. COPULA-BASED TOTAL CLAIM AMOUNT MODEL

If it is assumed that the average claim value Y_i and the number of claims N_i are dependent random variables, the total claim amount S_i is defined as the following product:

$$S_i = Y_i N_i, \quad i = 1,\ldots,n. \tag{11}$$

The variable obtained in this way is a continuous variable with positive values. Due to the occurrence of interrelations between random variables Y_i, N_i, the expected value of variable S_i has the following form:

$$[S_i] = E[Y_i N_i], \tag{12}$$

which means that the frequency-severity model does not apply here. Using therefore the basic formula for the expected value, the following is obtained:

$$E[S_i] = E[Y_i N_i] = \int_0^{+\infty} s_i f_S(s_i \mid \varphi_Y, \varphi_N, \theta) ds_i, \tag{13}$$

where $s_i = y_i$, k_i, $y_i > 0$, $k_i = 1, 2, 3,\ldots$, and $f_S(\cdot)$ is the density function of the variable S_i. If it is assumed that the relation between variables Y_i, N_i is described by the copula $C(\cdot|\theta)$, then according to theorem 6 in Krämer et al. (2013) the distribution of the total claim amount is given by the following density function:

$$f_S(s_i) = \\ = \sum_{k_i=1}^{\infty}[D(F_Y(\frac{s_i}{k_i}|\varphi_Y), F_N(k_i|\varphi_N)) - D(F_Y(\frac{s_i}{k_i}|\varphi_Y), F_N(k_i - 1|\varphi_N))]\frac{s_i}{k_i} f_Y(\frac{s_i}{k_i}|\varphi_Y) \tag{14}$$

for $s_i > 0$. It can be seen that the function has a complex form and the expected value $E[S_i]$ cannot be determined analytically and a numerical procedure has to be used. This paper puts forward the following algorithm (A2):

1. obtaining the vector parameters of margins and the copula parameter $C(\cdot|\theta)$ using the IFM method $(\varphi_Y, \varphi_N, \theta)'$ under the assumption of the family of copulas,
2. obtaining the value of $f_S(s_i|\hat{\varphi}_Y, \hat{\varphi}_N, \hat{\theta})$ according to (14).

It gives the opportunity to obtain the value of expectation $E[S_i]$ and the value of variance $Var(S_i) = \int_0^{+\infty} s_i^2 f_S(s_i)ds_i - \hat{E}^2[S_i]$ through numerical integration.

The advantage of the proposed procedure is its flexibility. Any model can be used to determine the initial values needed to estimate the copula parameters in point 1. In the case of insurance applications, it is convenient to adopt the frequency and severity model with the independence assumption (cf. Section 2 above). Unfortunately, the downside of the algorithm is its relatively slow operation, which is the effect of the need to sum up in step 2 and perform numerical integration in steps 3 and 4.

Example 3 – estimation of the total claim amount expectation using the copula-based model without unobserved heterogeneity

The model is illustrated using the same portfolio as in Example 1, but the structure of the relation between Y_i and N_i changes. It is accommodated by the two-dimensional copula C with the parameter θ. Assuming margins $Y_i \sim Gamma(\mu_i,\varphi_Y)$ and $N_i \sim ZTPois(\lambda)$, the algorithm (A2) is run in the case of four families of parametric bivariate copulas: the Gauss, Clayton, Gumbel and Frank copulas. As the IFM method is applied, the parameters of margins are the same as in Example 1. Using these values, the copula parameters and the corresponding Kendall coefficient τ are obtained. The results are listed in table 2.

Table 2.

Estimation of Kendall's tau and copula parameter

Copula	Gauss	Clayton	Gumbel	Frank
$\hat{\theta}$	0.48	2.8)	1.21	4.71
$\hat{\tau}$	0.32	0.58	0.17	0.44

Source: own calculations.

Based on the estimators presented above and using formula (14), the copula-based density of the total claim amount is constructed. Next, the expected values $E[S_i]$, $i = 1,\ldots,666$ are estimated through numerical integration. Figure 3 displays histograms of $\hat{E}[S_i]$ for different copulas.

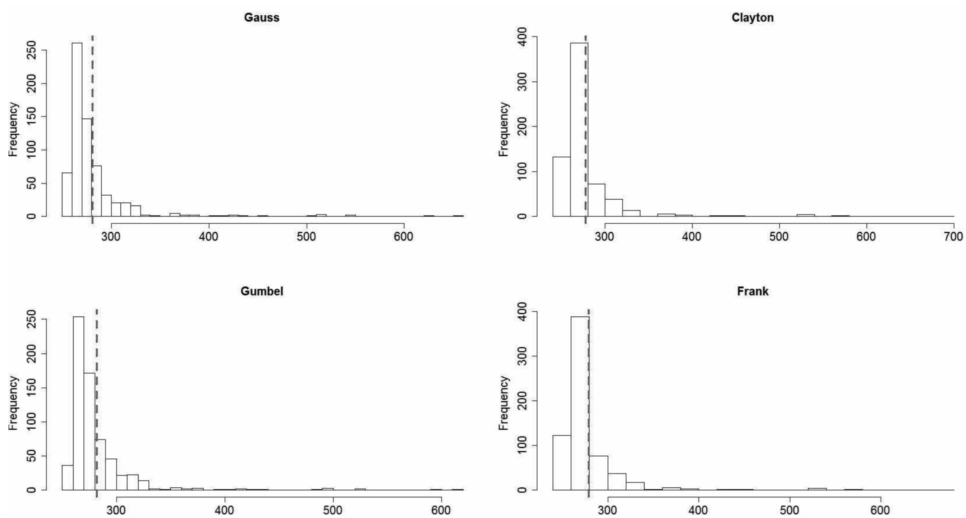


Figure 3. Histogram of expected total claim amount

Source: own calculations.

5. THE COPULA-BASED TOTAL CLAIM AMOUNT MODEL WITH AN INDIVIDUAL UNOBSERVABLE RISK FACTOR

Another starting point for *a posteriori* ratemaking are total claim amount models where the individual unobserved factor for the i-th risk, referred to as the risk profile (cf. Bühlmann, Gisler, 2005), is taken into account. This risk profile is usually taken into consideration in the claim frequency model using cross-sectional data (cf. Dimakos, Rattalma, 2002; Denuit et al., 2007; Boucher et al., 2007; Wolny-Dominiak, 2014) or longitudinal data (cf. Boucher et al., 2009; Wolny-Dominiak, 2014). It is well-known that this quantity is also affected by individual unobserved factors. One example is motor insurance, where the unobserved factor is equated with a driver's (an insured person's) individual features that have an impact on a given risk loss burden. A driver with a strong aversion to driving fast, with little children etc., will display a weaker tendency towards causing claims to arise than a daring driver. Most frequently, the unobserved risk factor is treated as a realization of a certain random variable with a pre-set probability distribution.

5.1. MARGINAL CLAIM FREQUENCY

Let us assume that the unobserved risk factor corresponding to unobserved heterogeneity defines the continuous random variable V with the density function $f_V(\cdot)$ with the parameter vector φ_V. In the copula-based total claim amount model, a proposal is made to introduce the factor into the marginal frequency model as a random effect V. Consequently, as in the mixed Poisson model (cf. Denuit et al., 2007), the parameter λ_i of the model $ZTPois(\lambda_i)$ is randomized by $\lambda_i V$, which gives a conditional distribution of the number of claims $N_i \mid V \sim ZTPois(\lambda_i V)$ with the mass probability function defined by the following formula:

$$P[N_i = k_i \mid V] = \frac{(\lambda_i V)^{k_i}}{[\exp(\lambda_i V) - 1]k_i!},\ k_i > 0. \tag{15}$$

The claim number distribution requires a transition from the conditional distribution to the marginal one. One possibility is the direct use of the conditional distribution and a formula for the infinite mixture of distributions of the number of claims and the unobserved factor:

$$P[N_i = k_i \mid k_i > 0] = \int_o^{+\infty} P^{ZTPois}[N_i = k_i \mid u] f_V(v \mid \varphi_V) du. \tag{16}$$

As it can be seen, for any density function $f_V(\cdot)$ the estimation of the distribution parameters is a complex task due to the occurrence of the random effect $\lambda_i V$. The direct use of formula (16) then requires numerical integration, which involves considerable lengthening of the computation time. Another possibility is to use the Expectation-Maximization (EM) method, which is also rather time-consuming (cf. Karlis, 2001; Trzęsiok, Wolny-Dominiak, 2015). On the other hand, the probability function (16) can sometimes be determined analytically. One example is the popular negative-binomial (NB) distribution, which is a *Poisson-Gamma* distribution mixture. Assuming that $V \sim Gamma(\alpha)$ and $N_i \mid V \sim ZTPois(\lambda_i V)$, the marginal distribution of the number of claims is a first-order NB distribution.

The zero-truncated distribution with an unobserved factor can be obtained easily in the same way as in the case of the ZTPois distribution.

$$\Pr^{ZT}(N_i = k_i \mid k_i > 0) = \frac{\Pr(N_i = k_i)}{1 - \Pr(N_i = 0)}. \tag{17}$$

For example, the zero-truncated NB (ZTNB) distribution has the following probability mass function:

$$\Pr^{ZTNB}(N_i = k_i \mid k_i > 0, \lambda_i, \alpha) = \frac{\Pr^{NB}(N_i = k_i \mid \lambda_i, \alpha)}{1 - \Pr^{NB}(N_i = 0 \mid \lambda_i, \alpha)}, \tag{18}$$

where $\alpha > 0$ is a dispersion parameter. The probability of the occurrence of zero is then: $\Pr^{NB}(N_i = 0 \mid \lambda_i, \alpha) = (1 + \alpha\lambda_i)^{-\alpha^{-1}}$ and the expectation $E_{ZTNB}[N_i] = \dfrac{\lambda_i}{1 - (1 + \alpha\lambda_i)^{-\alpha^{-1}}}$.

In order to estimate ZTNB parameters one can use the MLE method. The log-likelihood function is defined as follows:

$$\begin{aligned} l(\boldsymbol{\beta}^N, \alpha) = \sum_{i=1}^{n} [&\log\Gamma(k_i + \frac{1}{\alpha}) - \log\Gamma(\frac{1}{\alpha}) - \log k_i! - (k_i + \frac{1}{\alpha})\log(1 + \alpha\lambda_i) + \\ &+ k_i \log \alpha\lambda_i - \log(1 - (1 + \alpha\lambda_i)^{-\alpha^{-1}})], \end{aligned} \tag{19}$$

where regression coefficients are introduced into the model through the parameter $\lambda_i = E_i \exp(\mathbf{x}_i^N \boldsymbol{\beta}^N)$.

Example 4 – parameter estimation and construction of a ZTNB distribution

To illustrate the ZTP model with unobserved heterogeneity Gamma distributed, which gives a ZTNB distribution, we simulate the sample $n = 500$ of the numbers of claims distributed as $N_i \sim NB(\lambda = 2, \alpha = 0.67)$. Then, we truncate the sample receiving zero-truncated data. Maximizing the log-likelihood (19) with the BFGS method, the estimated parameters are $\hat{\lambda} = 1.91$, $\hat{\alpha} = 0.64$. Figure 4 provides the probability function and the cdf of the constructed ZTNB (equivalent to ZTP-Gamma) based on the NB with parameters $(\hat{\lambda}, \hat{\alpha})$.

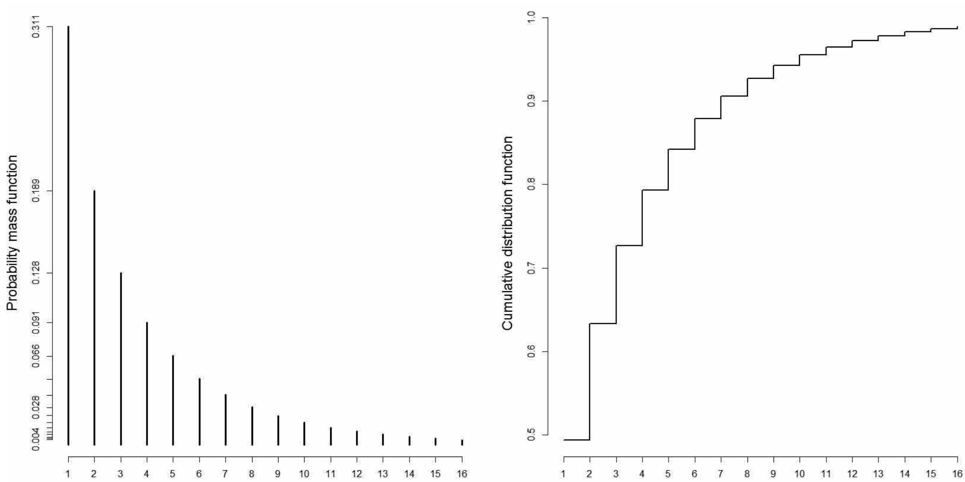


Figure 4. The probability function and the cumulative distribution function of ZTNB
Source: own calculations.

5.2. TOTAL CLAIM AMOUNT

Proceeding to the copula-based total claim amount model with the unobserved factor, three random variables are considered: the average claim value Y_i, the number of claims N_i and the unobserved factor V. The total claim amount is defined according to formula (2), except that the distribution of variable N_i is the marginal distribution of the two-dimensional variable (N_i, V). A proposal is made in this paper to determine the expected value of the total claim amount using the ZTNB distribution. It means that the unobserved factor is taken into account in the margin of the number of claims. The new procedure (A3) has the following steps:

1. obtaining the vector parameters of the number of claims (λ_i, α) assuming $N_i \sim ZTNB(\lambda_i, \alpha)$ and the regression component $\lambda_i = E_i \exp(\mathbf{x}_i^N \boldsymbol{\beta}^N)$,
2. obtaining the vector parameters of the average value of claims φ_Y assuming $Y_i \sim EDM(\mu_i, \varphi_Y)$ and the regression component $\mu_i = \exp(\mathbf{x}_i^Y \boldsymbol{\beta}^Y)$,
3. obtaining the copula parameter $C(\cdot|\theta)$ using the IFM method under the assumption of the copula type,
4. obtaining the value of $f_S(s_i \mid \hat{\mu}_i . \hat{\varphi}_Y, \hat{\lambda}_i, \hat{\alpha}, \hat{\theta})$ according to (14).

The constructed density of the total claim amount for a single risk gives the opportunity to estimate a pure premium. In the example below the proposed model is illustrated using real data from a Polish insurance company. As data is confidential, one can use another database in the **R** code.

Example 5 – total claim amount model with unobserved heterogeneity

We consider the portfolio that consists of 1,276 MOD (Motor Own Damage) policies insured in 2010 with the observed average value of claims Y_i and the number of claims N_i for every policy. The exposure E_i is taken as the duration of the policy. The histograms of random variables are shown in figure 5. The right-hand side is generated by the product Y_iN_i. The red lines represent the means.

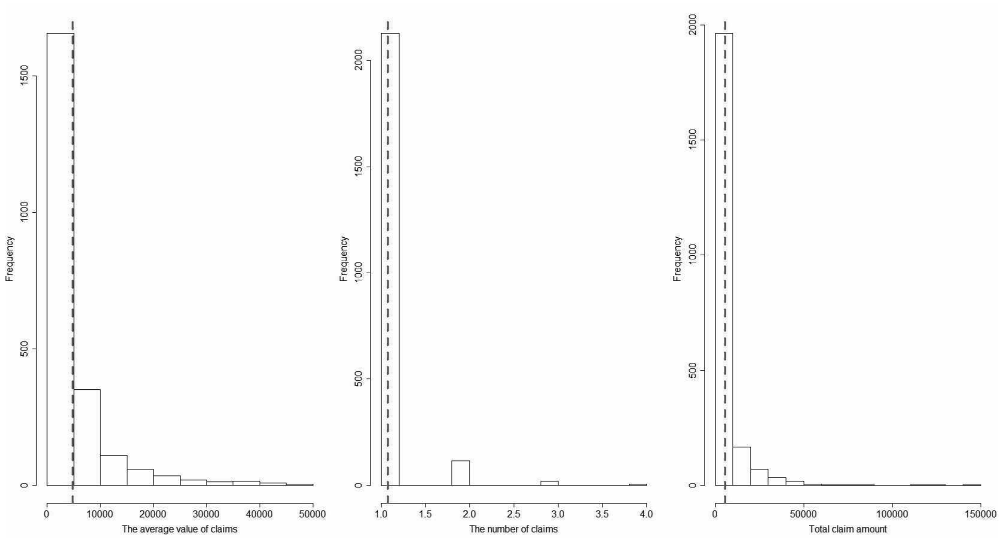


Figure 5. Histograms of the average value of claims, the number of claims and total claim amount for single risk

Source: own calculations.

The portfolio consists of three categorical covariates. Details on the factors are given in table 3.

Table 3.

Details on rating factors

Rating Factors	Categories/Number of observations		
POWER RANGE	0–66 269	67–124 803	125+ 187
GENDER	0 (Female) 416	1 (Male) 843	
PREMIUM_SPLIT	0 (No split) 754	1 (Split) 505	

Source: own calculations.

First, we analyze marginal models. As we see, the skew histogram of the average value of claims in the figure 4, the gamma distribution $Y_i \sim Gamma(\mu_i,\varphi_Y)$ is assumed. Figure 6 provides the boxplot divided according to the factor GENDER.

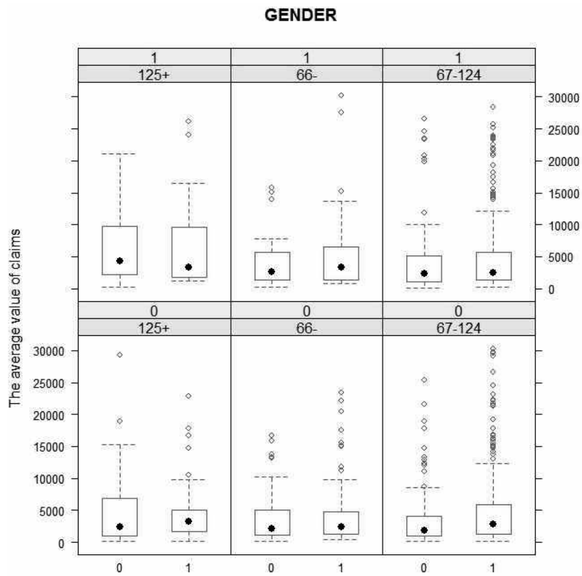


Figure 6. Boxplots of the average value according to the factor GENDER

Source: own calculations.

The Gamma assumption gives the opportunity to estimate the model parameters using the IWSL algorithm as in the standard practice in the GLM. As no claims policies are observed, the number of claims is modelled using the $N_i \sim ZTNB(\lambda_i,\alpha)$ distribution. It allows us to take into account unobserved heterogeneity in the total claim amount estimation. In order to estimate the model parameters and fit the claim frequency, we use the numerical optimization in the MLE method taking the log-likelihood function as in the formula (19).

All three coefficients are statistically significant according to the Wald test, but only in the GLM Gamma. For the number of claims no factors have significant coefficients on a level of 0.05. Therefore, we estimate λ parameter to be the same for every policy. The regression coefficient estimators in GLM Gamma are presented in table 4.

Table 5 shows fitted values of the average value of claims for all combinations of regression coefficients.

We observe a relatively high variability in the fitted claims amount. The lowest value is given by the cars with low power and a female driver, who pays the premium without splitting the payment, while the highest value is generated by high-power cars and a male driver paying in instalments.

Table 4.

GLM Gamma parameters

Rating Factors	$\hat{\beta}_j$	Standard error
Intercept	8.65	0.13
POWER RANGE (0–66)	-0.52	0.15
POWER RANGE (67–124)	-0.40	0.13
GENDER (1)	0.27	0.10
PREMIUM_SPLIT (1)	0.27	0.10
Dispersion parameter	$\hat{\phi}$ = 1.37	-

Source: own calculations.

Table 5.

The fitted average value of claims $\hat{Y}_i$ in groups

POWER RANGE.GENDER.PREMIUM SPLIT	Fitted value
125+.0.0	4582.11
66-.0.0	**3650.85**
67-124.0.0	3957.34
125+.1.0	5278.62
66-.1.0	4205.80
67-124.1.0	4558.88
125+.0.1	5132.54
66-.0.1	4089.42
67-124.0.1	4432.72
125+.1.1	**5912.72**
66-.1.1	4711.03
67-124.1.1	5106.52

Source: own calculations.

Afterwards we analyze the number of claims for a single risk. In order to take into account the unobserved factor, the distribution is assumed as $N_i \sim ZTNB(\lambda_i,\alpha)$. No factors have significant coefficients on a level of 0.1. Therefore, we estimate λ parameter, the same for every policy, receiving $\hat{\lambda} = 0.0003$, $\hat{\alpha} = 461.95$ with standard errors equal to 120.1 and 0.23 respectively. Thus, plugging this values into the $E_{ZTNB}[N_i]$ and multiplying by the exposure E_i the expected number of claims for a single risk is obtained. In the portfolio only 35 risks are not covered in the whole period ($E_i < 1$). Hence, most risks have $\hat{E}_{ZTNB}[N_i] = 1.08$ with $E_i = 1$.

Using the received estimated values of parameters we consider four type of copulas: Gaussian, Clayton, Gumbel and Frank. Maximizing the log-likelihood (9) we choose the Gumbel copula with fitted $\hat{\theta} = 1.19$, which is equivalent to Kendall's $\tau = 0.16$. This type of the copula gives the smallest AIC value.

Finally, we construct the copula-based density of the total claim amount $f_S(\cdot)$ according to the formula (14) and using estimated parameters $\hat{\mu}_i$, $\hat{E}_{ZTNB}[N_i]$, $\hat{\theta}$. It gives us full information about this random variable and the possibility of estimating the expected value of the total claim amount. Figure 7 on the left-hand side provides the plots of values of the density for risks from the analyzed portfolio. For comparison, we also present the density plot based on the kernel estimation (cf. Sheather, Jones, 1991).

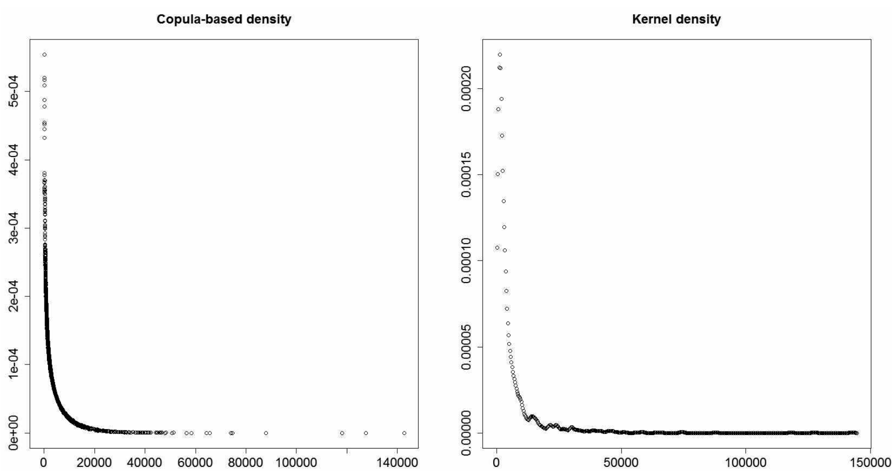


Figure 7. The density of total claim amount

Source: own calculations.

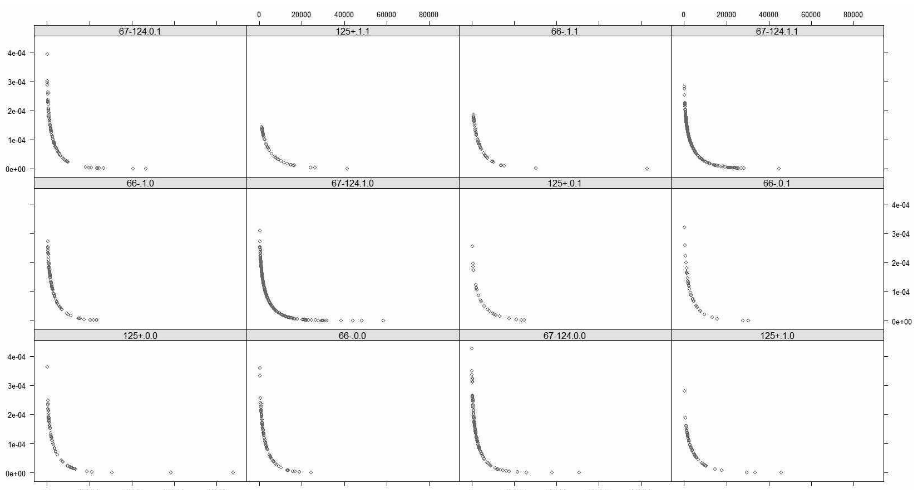


Figure 8. Copula-based density of total claim amount in groups

Source: own calculations.

In figure 8, we notice that the distributions in all groups are generally left-skewed. This is natural, as the margins of the average value of claims are Gamma distributed.

After that the copula-based expected total claim amount is determined using the MC simulation. This simulation provides values $\hat{E}[S_i]$ received via numerical integration. Figure 9 provides the summary.

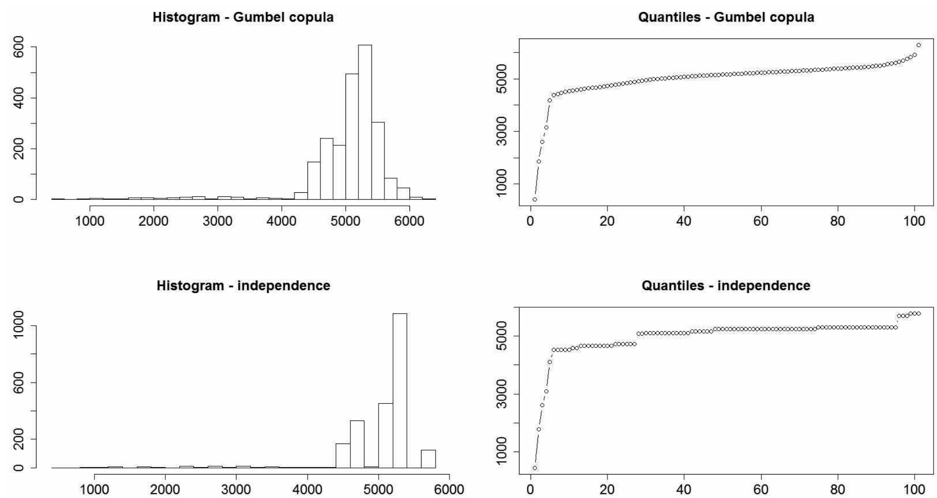


Figure 9. Expected total claim amount – MC integration

Source: own calculations.

The results show that values received in the copula-based model are slightly higher than the values in the model under the independence assumption. This fact is observed in the histograms as well as in the quantile plots. It can suggest that models commonly applied by insurance companies underestimate total claim amounts and hence pure premiums for a single risk. To visualize the variability of the expected total claim amount in groups according to the combinations of regressors taken in the Gamma GLM, the boxplot is displayed in figure 10.

It shows low variability in all groups appearing rather for risks with the low value of the claim amount. Except that the means (the black dots) are decisively higher for males with power 125+ than for females with any power, which is the intuitive result.

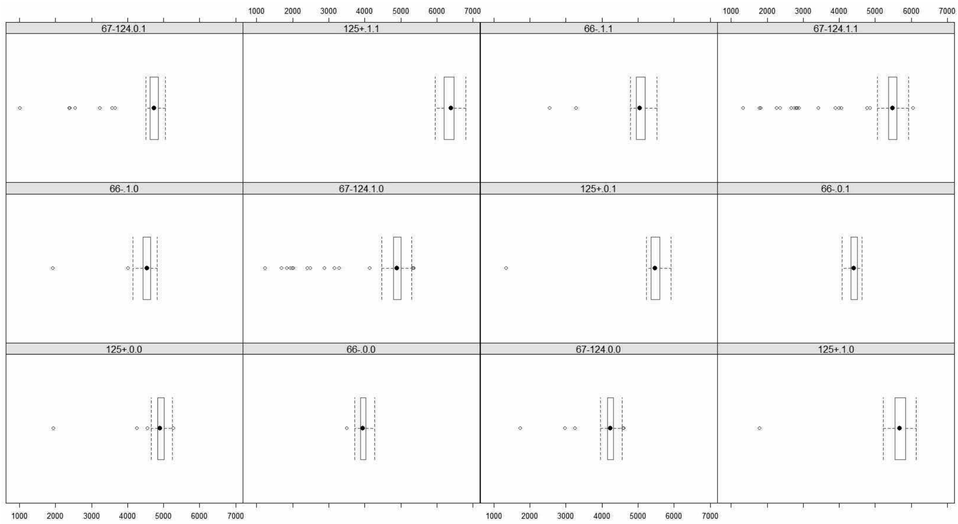


Figure 10. Copula-based expected total claim amount in groups
Source: own calculations.

6. CONCLUSIONS

In this paper, we model an average value of claims and the number of claims in the case of dependence between both random variables. The proposed model provides exact distributions of individual total claim amounts, which tend to be left-skewed. Moreover, we also show how to numerically construct the density of the bivariate random variable. This gives the possibility of estimating the expected total claim amounts in the portfolio using e.g. MC integration in pricing. As we use the ZTNB distribution, heterogeneity is taken into account. It corresponds to credibility representing the unobservable factor influencing the number of claims for a single risk. However, there are no obstacles to use another mixed Poisson model (cf. Karlis, 2001; Wolny-Dominiak, Trzęsiok, 2015). Nowadays the statistical modelling cannot do without computation, so the numerical examples discussed in this paper required strong programming work. Therefore, the full **R** code with a complete description is available for download.

REFERENCES

Antonio K., Valdez E. A., (2012), Statistical Concepts of A Priori and A Posteriori Risk Classification in Insurance, *Advances in Statistical Analysis*, 96 (2), 187–224.

Boucher J. P., Denuit M., Guillén M., (2007), Risk Classification for Claim Counts: A Comparative Analysis of Various Zeroinflated Mixed Poisson and Hurdle Models *North American Actuarial Journal*, 11 (4), 110–131.

Boucher J. P., Denuit M., Guillén M., (2009), Number of Accidents or Number of Claims? An Approach with Zero-Inflated Poisson Models for Panel Data, *Journal of Risk and Insurance*, 76 (4), 821–846.
Bühlmann H., Gisler A., (2005), *A Course in Credibility Theory and its Applications*, Springer.
Cizek P., Härdle W. K., Weron R., (2011), *Statistical Tools for Finance and Insurance*, Springer Science & Business Media.
De Jong P., Heller G. Z., (2008), *Generalized Linear Models for Insurance Data*, Cambridge University Press.
Denuit M., Maréchal X., Pitrebois S., Walhin J. F., (2007), *Actuarial Modelling of Claim Vounts: Risk Classification, Credibility and Bonus–malus Systems*, John Wiley & Sons, Chichester.
Dimakos X. K., Di Rattalma A. F., (2002), Bayesian Premium Rating with Latent Structure, *Scandinavian Actuarial Journal*, 2002 (3), 162–184.
Frees E. W., (2009), *Regression Modeling with Actuarial and Financial Applications*, Cambridge University Press.
Joe H., (1997), *Multivariate Models and Multivariate DSependence Concepts*, CRC Press.
Karlis D., (2001), A General EM Approach for Maximum Likelihood Estimation in Mixed Poisson Regression Models, *Statistical Modelling*, 1 (4), 305–318.
Krämer N., Brechmann E. C., Silvestrini D., Czado C., (2013), Total Loss Estimation using Copula-based Regression Models, *Insurance: Mathematics and Economics*, 53, 829–839.
Nelsen R. B., (1999), *An Introduction to Copulas*, Springer Science & Business Media.
Ohlsson E., Johansson B., (2010), *Non-life Insurance Pricing with Generalized Linear Models*, Springer.
R Core Team (2014), R: A *Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R–project.org/.
Sheather S. J., Jones M. C., (1991), A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation, *Journal of the Royal Statistical Society*, Series B (Methodological), 683–690.
Shi P., Feng X., Ivantsova A., (2015), Dependent Frequency–severity Modeling of Insurance Claims, *Insurance: Mathematics and Economics*, 64, 417–428.
Sklar M., (1959), *Fonctions de Répartition à n Dimensions et Leurs Marges*, Université de Paris 8.
Sun J., Frees E. W., Rosenberg M. A., (2008), Heavy-tailed Longitudinal Data Modeling Using Copulas, *Insurance: Mathematics and Economics*, 42 (2), 817–830.
Trzęsiok M., Wolny-Dominiak A., (2015), Złożony Mieszany Rozkład Poissona – Zastosowania Ubezpieczeniowe, *Studia Ekonomiczne*, 227, 85–95.
Wanat S., (2012), *Modele Zależności w Agregacji Ryzyka Ubezpieczyciela*. Zeszyty Naukowe, Uniwersytet Ekonomiczny w Krakowie, Seria Specjalna, Monografie.
Wolny-Dominiak A., (2014), *Taryfikacja w Ubezpieczeniach Majątkowych z Wykorzystaniem Modeli Mieszanych*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
Wolny-Dominiak A., Trzęsiok M., (2014), *insuranceData: A Collection of Insurance Datasets Useful in Risk Classification in Non-life Insurance*, R package version 1.0.
Wolny-Dominiak A., Trzpiot G., (2013), GLM and Quantile Regression Models in A Priori Ratemaking, *Insurance Review*, 4, 49–57.

REGRESYJNY MODEL ŁĄCZNEJ WARTOŚCI SZKÓD Z UWZGLĘDNIENIEM NIEOBSERWOWALNEGO CZYNNIKA RYZYKA

Streszczenie

W masowych portfelach ryzyk zakłady ubezpieczeń przeprowadzają tzw. taryfikację, której celem jest wyznaczenie składki czystej dla pojedynczego ryzyka. Modele statystyczne stosowane obecnie w praktyce należą najczęściej do klasy uogólnionych modeli liniowych (GLM), w których szacuje się w osobnych modelach wartości oczekiwane dwóch zmiennych losowych: średniej wartości szkody oraz liczby szkód dla ryzyka. Składka czysta definiowana jest wtedy jako iloczyn uzyskanych wartości. Takie podejście wymaga założenia niezależności pomiędzy rozpatrywanymi dwoma zmiennymi losowymi. Jednak w literaturze to założenie jest podważane. Celem tego artykułu jest zaproponowanie modelu z kopulą uwzględniającego nieobserwowalny czynnik ryzyka w modelowaniu liczby szkód. Model ten służy do oszacować oczekiwanej wartości iloczynu dwóch zmiennych losowych: średniej wartości szkody oraz liczby szkód dla pojedynczego ryzyka przy założeniu zależności oraz występowaniu czynnika nieobserwowalnego. W pracy szczegółowo opisano aspekty teoretyczne związane z budową modelu oraz szacowaniem wartości oczekiwanej. Ponadto w licznych przykładach przedstawiono numeryczne rozwiązania obliczeniowe w programie **R**. Dodatkowo udostępniono kody programu **R** na stronie internetowej http://web.ue.katowice.pl/woali/.

Słowa kluczowe: taryfikacja, GLM, nieobserwowalny czynnik ryzyka, kopula

THE COPULA-BASED TOTAL CLAIM AMOUNT REGRESSION MODEL WITH AN UNOBSERVED RISK FACTOR

Abstract

Nowadays a common practice of any insurance company is ratemaking, which is defined as the process of classification of the mass risk portfolio into risk groups where the same premium corresponds to each risk. As generalised linear models are usually applied, the process requires the independence between the average value of claims and the number of claims. However, in literature this assumption is called into question. The interest of this paper is to propose the copula-based total claim amount model taking into account an unobservable risk factor in the claim frequency model. This factor, called also as unobserved heterogeneity, is treated as a random variable influencing the number of claims. The goal is to estimate the expected value of the product of two random variables: the average value of claims and the number of claims for a single risk assuming the dependence between the average value of claims and the number of claims for a single risk and the dependence between the number of claims for a single risk and the unobservable risk factor. We give details of the theoretical aspects of the model as well as the empirical example. To acquaint the reader with the model operation, every step of the process of the expected value estimation in described and the **R** code is available for download, see http://web.ue.katowice.pl/woali/.

Keywords: ratemaking, GLM, unobserved factor, copula

DOI: 10.5604/01.3001.0014.1212

PRZEGLĄD STATYSTYCZNY
R. LXIII – ZESZYT 3 – 2016

MARCIN CHLEBUS[1]

CAN LOGNORMAL, WEIBULL OR GAMMA DISTRIBUTIONS IMPROVE THE EWS-GARCH VALUE-AT-RISK FORECASTS?

1. INTRODUCTION

International regulations established by the Basel Committee on Banking Supervision impose the obligation to manage the market risks, which are regarded as one of the three main risks in banking. Essential part of the risk management is its measurement. It has to be based on a Value-at-Risk in order to satisfy the basic requirement for an internal model.

According to the results obtained by researchers, it is not possible to determine the best method of measuring Value-at-Risk that would allow to achieve the best forecasts of Value-at-Risk in every situation. Therefore, the analysis of the quality of a Value-at-Risk forecasts generated on the basis of different models is a topic widely discussed in the literature (among others, in Engle, 2001; 2004; Tagilafichi, 2003; Alexander, Lazar, 2006; Angelidis et al., 2007; Engle, Manganelli, 2001; McAleer et al., 2009; Marcucci, 2005; Ozun et al., 2010; Dimitrakopoulos et al., 2010; Brownlees et al., 2011; Degiannakis et al., 2012 and Abad et al., 2014).

Moreover, McAleer et al. (2009) and Degiannakis et al. (2012) showed that different models may be better during tranquil or turbulent periods. In both cases, simple GARCH model was good for Value-at-Risk forecasting during a pre-crisis 2007–2009 period, but its quality significantly decreased during and after the crisis. McAleer et al. (2009) showed that RiskMetrics™ was the best model during the crisis but EGARCH-t model was better after the crisis. Whereas in the study of Degiannakis et al. (2012) APARCH with a skewed Student's t distribution was the best model during the crisis. These results show that less conservative models are best in tranquil periods, while during the crisis models that consider the distributions of returns with fatter tails are better. Degiannakis et al. (2012) stated that these claims are valid for both developed and developing countries.

Despite the conclusions drawn from the aforementioned articles, the use of regime switching models in Value-at-Risk forecasting has a rather niche character; it has been considered, among others, by Hamilton, Susmel (1994), Cai (1994), Gray (1996), Alexander, Lazar (2006) and McAleer, Chan (2002). A characteristic trait of the pro-

[1] University of Warsaw, Faculty of Economic Sciences, Department of Quantitative Finance, 44/50 Długa St., 00-241 Warsaw, Poland, e-mail: mchlebus@wne.uw.edu.pl.

posed models is that losses come from the same distribution but with different parameters, in all states. This feature contradicts the findings stated in McAleer et al. (2009) and Degiannakis et al. (2012), where models with different distributions were found to be the best in different states.

In order to fill this gap, an EWS-GARCH models were presented in Chlebus (2016b). In these models, the Value-at-Risk forecasts are calculated in two steps. First, a state of the portfolio is forecasted (a state of tranquillity or a state of turbulence – the approach is analogous to Early Warning System (EWS) models for crisis prediction) and then, depending on the forecasted state, a different model is used to forecast the Value-at-Risk. The EWS-GARCH models give the opportunity to use models to forecast Value-at-Risk in the state of tranquillity assuming a distribution of returns with relatively thinner tails, and in the state of turbulence, models with much more conservative assumptions.

In the aforementioned study, a GARCH(1,1), or a GARCH(1,1) with the amendment to empirical distribution of random error, were considered as a Value-at-Risk forecasting model in the state of tranquillity; whereas exponential, empirical or Pareto distributions were considered in the state of turbulence. The obtained results were promising and showed that the EWS-GARCH models concept may provide Value-at-Risk forecasts of very good quality. However, a lot of aspects remain in which the EWS-GARCH models may be improved.

The aim of the study is to examine whether incorporation of lognormal, Weibull or Gamma distributions in the Value-at-Risk forecasting model (in the state of turbulence), instead of distributions used previously, may increase a quality of the Value-at-Risk forecasts. The use of these distributions in Value-at-Risk forecasting is a practice met in an operational risk measurement (see Panjer, 2006). They may be considered as distributions in the state of turbulence, as all of them may have tail shape (when specific values of parameters assumed).

The lognormal, the Weibull or the Gamma distribution were compared to each other and with benchmark models: the GARCH(1,1), the GARCH(1,1) with the amendment to empirical distribution of random error, an EGARCH(1,1), a GARCH-t (1,1) (model was parametrised assuming unit variance and the number of the degrees of freedom greater than 2), and the EWS-GARCH(1,1) models with the exponential or the empirical distributions; in order to assess the quality of the Value-at-Risk forecasts obtained from the EWS-GARCH models. The evaluation of the quality of the Value-at-Risk forecasts was based on the Value-at-Risk forecasts adequacy (an excess ratio, a Kupiec test, a Christoffersen test, an asymptotic test of unconditional coverage and a backtesting criteria defined by the Basel Committee – both for Value-at-Risk and Stressed Value-at-Risk) and the analysis of loss functions (a Lopez quadratic loss function, an Abad & Benito absolute loss function, a 3rd Caporin loss function and an excessive cost function).

The paper is organized as follows: in the first section an EWS-GARCH models framework is discussed, in the second section a testing framework is presented, and in the third section an empirical verification of the Value-at-Risk forecasts obtained from the EWS-GARCH models with the lognormal, the Weibull or the Gamma distribution is analysed.

2. EWS-GARCH MODELS

At the beginning a brief definition of Value-at-Risk ($VaR_\alpha(t)$) should be presented. The Value-at-Risk may be defined as a value that a loss would not excess with a certain probability α within a specified period of time in normal market situation. Value-at-Risk can be defined as follows (Engle, Manganelli, 1999):

$$P(r_t < VaR_\alpha(t)|\Omega_{t-1}) = \alpha, \tag{1}$$

where r_t is a return at time t, $VaR_\alpha(t)$ is Value-at-Risk at time t and Ω_{t-1} is a set of information available at time t–1.

A Value-at-Risk forecasting procedure based on the EWS-GARCH models consists of two steps. In the first step, the state of time series for the next day is forecasted, then in the second step a Value-at-Risk for the next day is forecasted. The Value-at-Risk forecast is provided from an appropriate model regarding the state forecasted in the first step.

In the EWS-GARCH models it is proposed that the prediction of the state should be carried out by a model for binary dependent variable: logit, probit or cloglog models. Each of these models can be defined in a similar manner differing only in regard of a random error distribution. The logit model assumes a logistic distribution, the probit model a normal distribution, and the cloglog – a Gompertz distribution of random errors. These models can be defined as follows (Allison, 2005):

$$y_t^* = \boldsymbol{\beta X}_t + \varepsilon_t, \tag{2}$$

$$y_t = \begin{cases} 1 & y_t^* > 0, \\ 0 & y_t^* \leq 0, \end{cases} \tag{3}$$

where y^* is a latent dependent variable, $\boldsymbol{\beta}$ is a vector of parameters describing the relationship between independent variables and unobserved dependent variable, $\boldsymbol{X_t}$ is a vector of observations of independent variables that have an impact on an unobservable dependent variable, ε_t is a random error coming from the relevant distribution, and y_t is observable result of the modelled phenomenon. All aforementioned models are estimated using maximum likelihood estimators.

In the process of forecasting the state of turbulence, the y_t is equal to 1 for a certain percentage of the lowest observed returns (5% or 10%). Independent variables in the model describe a current situation on stock, exchange rates and short-term interest rates markets (prices and returns, 15-day moving averages of prices and returns and

15-days moving variances of prices and returns of Warsaw Stock Exchange Indices – WIG & WIG20, of most important to polish market exchange rates – EUR/PLN, USD/PLN and CHF/PLN, and of short-term interest rates – overnight and 3-month WIBOR). Moreover, a selection of an optimal cut-off point for the event forecast is considered (set up to 5% and 10% for the 5% and 10% definitions of y_t relevantly) to achieve the best possible forecasts quality. The choice of models for binary variable, the definition of the observable dependent variable, the choice of independent variables and the optimal cut-off threshold have been established in accordance with the results obtained in the study of Chlebus (2016a). Additionally, a set of independent variables will be limited only to variables statistically significant at the 5% significance level selected by a stepwise selection method.

The model to predict a state gives the opportunity to distinguish two states (the state of tranquillity and the state of turbulence) in a time series, which can vary considerably in their nature (with respect to expected returns, volatility etc.). In each state different models to forecast Value-at-Risk should be used in order to take into account different specificities of these two states. In the EWS-GARCH models a tail distribution is used only when the state of turbulence is forecasted, otherwise the entire distribution is used. During the study it is assumed that the dependent variable in the Value-at-Risk models is a continuous one-day rate of return, which may be expressed as $r_t = (\ln(p_t) - \ln(p_{t-1})) * 100$.

In the state of tranquillity, the considered Value-at-Risk forecasting models were: the GARCH(1,1) and the GARCH(1,1) with amendment to empirical distribution of random error. The GARCH(1,1) model can be written as:

$$r_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t \xi_t, \tag{4}$$

where r_t is a return on assets analysed at time t, μ_t is a conditional mean (assumed in the study to be constant – no independent variables included), ε_t is a random error in time t and ε_t can be expressed as the product of the conditional standard deviation σ_t and standardized random error ξ_t at time t, which satisfies the assumption $\xi_t \sim$ i.i.d.(0, 1). The equation of conditional variance in the GARCH(1,1) can be written as:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{5}$$

where ω is a constant which satisfies the assumption $\omega > 0$, α_1 and β_1 are parameters that satisfy the assumptions $\alpha_1 \geq 0$ and $\beta_1 \geq 0$. The GARCH(1,1) model is estimated using the maximum likelihood method.

For the GARCH(1,1) Value-at-Risk for the long position is estimated based on the following formula (Abad, Benito, 2013):

$$\mathrm{VaR}_\alpha(t) = \widehat{\mu_t} + k_\alpha * \sqrt{\hat{\sigma}_t^2}, \tag{6}$$

where $VaR_{\alpha}(t)$ is a forecast of Value-at-Risk on α tolerance level at time t, $\widehat{\mu_t}$ is a forecast of conditional mean at time t, k_{α} is a value of quantile α from assumed random error distribution and $\hat{\sigma}_t^2$ is an forecast of conditional variance at time t.

The Basel Committee requirements state that the Value-at-Risk should be estimated with the 99% confidence level (the α is assumed to be equal to 1%). The Value-at-Risk forecast from the GARCH(1,1) with the amendment to empirical distribution of random error (Engle, Manganelli, 2001) is obtained in a similar manner as in the GARCH(1,1); the difference lies in the use of a quantile from the empirical distribution of residuals instead of a quantile from the normal distribution.

In the state of turbulence, the lognormal, the Weibull or the Gamma distributions are considered. The lognormal distribution is uniquely determined by two parameters. A cumulative distribution function (cdf) can be written as:

$$F_{LN}(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right), \tag{7}$$

where $\Phi(x)$ is the standard normal distribution cdf, μ is a location parameter and σ is a shape parameter.

The second possible distribution is the Weibull distribution, which is a generalization of the exponential distribution. It is extended by a scaling parameter τ. In case where the parameter τ is equal to 1, the Weibull distribution reduces to the exponential distribution. The cdf can be written as:

$$F_{WEI}(x) = 1 - e^{-(x/\theta)^{\tau}}, \tag{8}$$

where θ is a scale parameter and τ is a shape parameter.

Last considered distribution is the Gamma distribution. The cdf can be written as:

$$F_{GAM}(x) = \frac{\gamma(\alpha; x/\theta)}{\Gamma(\alpha)}, \tag{9}$$

where $\gamma(\alpha; x/\theta)$ is the incomplete Gamma function, $\Gamma(\alpha)$ is the Gamma function, θ is a scale parameter and α is a shape parameter. In case when the α is a natural number, the Gamma distribution can be interpreted as the sum of exponentially distributed random variables. The formulation of the exponential distribution may be found in Chlebus (2016b). All aforementioned distributions are fitted using maximum likelihood estimators.

For the tail distributions Value-at-Risk is forecasted simply as a value of the quantile of the distribution. A problem in this case is the determination which quantile of the distribution provides the confidence level equal to 99%. Two quantiles are considered: the 99$^{\text{th}}$ percentile of the tail returns distribution (conservative assumption) and for the 10% definition of the state of turbulence the 90$^{\text{th}}$ percentile of the tail distribution and accordingly, for the 5% definition of the state of turbulence the 80$^{\text{th}}$ percentile of the tail distribution (liberal assumption).

The two-stage nature of the EWS-GARCH models forecasts two elements: the state of turbulence, and the Value-at-Risk. Forecasts of the state and the Value-at-Risk at time $t+1$ are based on data available at time t. A data set to forecast the states is prepared using the recursive window approach. Data set for Value-at-Risk forecasting is prepared using the rolling window approach (the window width was set to 1004 observations, which corresponds to about 4 years of one day returns).

3. TESTING FRAMEWORK

Performing a thorough analysis of the quality of EWS-GARCH models requires the development of multi-aspect testing process. Tests of the adequacy of the Value-at-Risk forecasts and the loss functions analysis were carried out in order to confirm the quality of Value-at-Risk forecasts and comparisons of the models in terms of their quality.

As a part of the Value-at-Risk forecasts adequacy assessment, analyses of the following were performed: the excess ratio comparison, the Kupiec test, the Christoffersen test, the asymptotic test of unconditional coverage, and the backtesting criterion specified by the Basel Committee (see BCBS; 2006). The excess ratio and the backtesting criterion was analysed for the Value-at-Risk and the Stressed Value-at-Risk (a measure defined by the Basel Committee in the BCBS (2011)).

The excess ratio may be calculated as:

$$ER = \frac{\sum_{t=1}^{N} 1_{r_t < VaR_t}}{N}, \tag{10}$$

where N is a number of the Value-at-Risk forecasts and $1_{r_t < VaR_t}$ is a number of cases when a realized rate of return is smaller than a forecasted Value-at-Risk.

Using the excess ratio each of the Value-at-Risk models can be assigned to one of the Basel backtesting criterion zones – green, yellow or red. The Basel Committee requires comparing the quality of the models based on the Value-at-Risk forecasts results, however it is also worth to consider the quality of the models with regards to the Stressed Value-at-Risk. For this purpose, the worst excess ratio (from the set of 250 consecutive days with the highest excess ratio) from the out-of-sample was calculated. The result shows how the model works in the worst possible conditions observed. Analogously to the Value-at-Risk forecasts, in this case the excess ratio can be attributed as well to one of the backtesting zones defined by the Basel Committee.

The analysis of the backtesting zones has a one-tailed character. An important issue missing from this analysis is the negative assessment of the model forecasts due to excessive conservatism. In the backtesting, a model that does not identify any exceedances of the Value-at-Risk is assessed as very good (the green zone), although the expected and observed number of exceedances differ significantly. In order to assess the quality of forecasts from the perspective of both underestimation and overestimation of Value-at-Risk forecasts, among other, coverage tests are used.

The most popular test of this type is the Kupiec test (also called the unconditional coverage test) (see Kupiec, 1995). The idea of the test is based on a comparison of expected and observed numbers of Value-at-Risk exceedances. The test statistic comes from the asymptotic distribution of χ^2 with 1 degree of freedom and can be written as:

$$LR_{UC} = 2[ln(\hat{\alpha}^X(1-\hat{\alpha})^{N-X}) - ln(\alpha^X(1-\alpha)^{N-X})] \sim \chi^2_a(1), \tag{11}$$

where α is an expected excess ratio (according to the Basel Committee requirements it should be 1%), $\hat{\alpha}$ is an observed excess ratio, X is an observed number of Value-at-Risk exceedances and N is a number of Value-at-Risk forecasts. In the null hypothesis it is assumed that the expected and observed excess ratio is equal to each other. In contrast to the backtesting criterion, the Kupiec test identify models that both underestimate and overestimate Value-at-Risk, however there is no straightforward method to assess whether the analyzed model tends to overestimate or underestimate Value-at-Risk forecasts. Such an analysis is possible based on a backtesting criterion statistics, also called an asymptotic test of unconditional coverage (see Abad et al., 2014). The backtesting criterion statistics come from the asymptotic standard normal distribution. This test is two-tailed. Strongly negative values of the test statistics indicate overestimation of the Value-at-Risk forecasts, while strongly positive, underestimation of these forecasts. The test statistic can be calculated according to the following formula:

$$z_{BT} = \frac{(N\hat{\alpha} - N\alpha)}{\sqrt{N\alpha(1-\alpha)}} \sim N(0,1), \tag{12}$$

where α is an expected excess ratio, $\hat{\alpha}$ is an observed excess ratio and N is a number of Value-at-Risk forecasts.

The Christoffersen test (the conditional coverage test) proposed by Christoffersen (1998) is an extension of the Kupiec test. This test extends the Kupiec test by inclusion of an independency of Value-at-Risk exceedances testing. The test statistic comes from the asymptotic χ^2 distribution with 2 degrees of freedom and can be formulated as:

$$LR_{CC} = LR_{UC} + LR_{IND} \sim \chi^2_a(2), \tag{13}$$

where LR_{UC} is the Kupiec test statistics and LR_{IND} is an independency of exceedances statistics. The LR_{IND} is equal to

$$2[\ln((1-\pi_{01})^{N_{00}}\pi_{01}{}^{N_{01}}(1-\pi_{11})^{N_{10}}\pi_{11}{}^{N_{11}}) - \ln((1-\hat{\alpha})^{N_{00}+N_{10}}\hat{\alpha}^{N_{01}+N_{11}})],$$

where $\hat{\alpha}$ is an observed excess ratio, N_{ij} is a number of observation for which a state j (exceedance or not exceedance) is observed under condition that a state i (exceedance or not exceedance) was observed in the previous period, π_{01} is a probability of observing Value-at-Risk exceedances conditional on not observing them in the previous period and π_{11} is a probability of observing Value-at-Risk exceedances conditional on observing them in the previous period.

The tests presented above allow to evaluate Value-at-Risk models based on the adequacy of its forecasts. Additionally, an analysis of the cost (loss) compares on the one hand the losses resulting from exceeding the Value-at-Risk, and on the other hand, accuracy and cost efficiency of the used models. The cost (loss) functions analysis are not formal tests, during the analysis the score is calculated which allows to compare the Value-at-Risk models with each other.

The first cost (loss) function considered is the quadratic Lopez function (see Lopez, 1999), which may be defined as:

$$CL_t = \begin{cases} 1 + (r_t - VaR_\alpha(t))^2 & for \ r_t < VaR_\alpha(t), \\ 0 & for \ \ r_t \geq VaR_\alpha(t), \end{cases} \tag{14}$$

where r_t is a realised rate of return at the moment t and VaR_t is a Value-at-Risk forecast for the same moment t. The score is calculated as $\sum_{t=1}^{N} CL_t$ (where N is a number of Value-at-Risk forecasts). The Lopez function considers two aspects of Value-at-Risk forecasts: a number and a severity of exceedances. Each exceedance increase a score by at least 1, where the excess over 1 is calculated with respect to its severity and is calculated as $(r_t - VaR_t)^2$. The main disadvantage of the Lopez quadratic function is that it does not give an easy interpretation. The solution may be a function proposed by Abad, Benito (2013), which can be written as:

$$CA_t = \begin{cases} |r_t - VaR_\alpha(t)| & for \ r_t < VaR_\alpha(t), \\ 0 & for \ \ r_t \geq VaR_\alpha(t). \end{cases} \tag{15}$$

In this case a score is calculated as an average of severity of exceedances with respect to a number of Value-at-Risk forecasts considered, which can be calculated as $\sum_{t=1}^{N} CA_t/N$. This loss function differs from the previous one in two basic dimensions. Firstly, an average is minimized instead of the sum, therefore the number of exceedances is not taken into account. This may cause models with a larger number of exceedances to be preferred. Secondly, absolute deviation is analyzed, which makes the interpretation easier.

Both aforementioned functions consider non-zero values only in the case of exceedance. From a perspective of use of Value-at-Risk models in a financial institutions, it is reasonable to consider also cost (loss) functions that take into account the costs associated with both exceedances and lack of exceedances (opportunity costs). First considered function of this type is a function presented by Caporin (2008). In his study, he proposed three different cost functions, which assume that a cost of deviations of a forecasted Value-at-Risk from a realized rate of return is equally important regardless of whether the exceedance was observed or not. In the study the following cost function is considered:

$$CC_t = \begin{cases} |r_t - VaR_\alpha(t)| & for \ r_t < VaR_\alpha(t), \\ |r_t - VaR_\alpha(t)| & for \ r_t \geq VaR_\alpha(t). \end{cases} \tag{16}$$

Caporin proposes that in order to compare the Value-at-Risk forecasts, a sum of all CC_t should be used, however in the study the average of these values is considered. Both analyzes lead to similar conclusions, but the average can be interpreted as the average absolute error of the Value-at-Risk forecasts.

Additionally, an absolute excessive cost functions proposed in Chlebus (2016b) were analysed. The absolute excessive cost function, like the Caporin loss function, includes costs either in the case of the Value-at-Risk exceedance or lack of exceedance. The difference is that the analysis is focused rather on the excessive cost of the use of the model than precision of the forecast. Therefore, the process of assigning point values is divided into three cases and focuses precisely on the costs made by the model:

$$CAE_t = \begin{cases} |VaR_\alpha(t)| & for \ r_t \geq VaR_\alpha(t) \ and \ r_t \geq 0, \\ |VaR_\alpha(t) - r_t| & for \ r_t \geq VaR_\alpha(t) \ and \ r_t < 0, \\ |r_t| & for \quad r_t < VaR_\alpha(t). \end{cases} \tag{17}$$

Value-at-Risk models should be compared in terms of mean value of excessive cost function for the analysed number of forecasts $\overline{CAE} = \frac{\sum_{t=1}^{N} CAE_t}{N}$. The $\overline{CAE}$ may be interpreted as a measure of excessive model conservatism. The higher the $\overline{CAE}$ is, the more conservative the model is, which means that the model predicts on average more conservative Value-at-Risk than needed to cover losses arising from changes in a value of analysed assets.

The variety of Value-at-Risk forecast quality methods gives an opportunity to assess models form many different perspectives and thoroughly compare them. The empirical assessment of the quality of Value-at-Risk forecasts based on EWS-GARCH models with lognormal, Weibull and Gamma distribution are presented in the next section.

4. EMPIRICAL RESULTS

4.1. DATA

The quality of Value-at-Risk forecasts obtained from the EWS-GARCH models was analysed for 79 time series of returns of companies listed on the Warsaw Stock Exchange (a detailed list available upon request). Assets were selected randomly. Only one condition was imposed on the drawing process, that the shares have been listed on the Warsaw Stock Exchange since at least January 2006. It is a technical requirement intended to ensure the best possible quality of data used for modelling and similarity of sample for each company.

The empirical study was performed for the series of returns from the 1st January 2006 to 31st January 2012. The period from the beginning of 2006 to the end of 2009 constituted the original estimation sample; the forecast sample started from the beginning of 2010 and ended in 2012, thereby giving 525 predictions of the Value-at-Risk for each asset.

All considered models used to forecast the Value-at-Risk have been developed in such a way as to meet the requirements set by the Basel Committee for internal models of the market risk measurement. The measure of market risk is based on the one-day Value-at-Risk predictions satisfying the 99% confidence level. For the quality of Value-at-Risk forecasts only one-day predictions are required and sufficient. The assessment was carried out for 525 observations, which is more than expected in the Basel regulations of the minimum equal to 250 observations.

4.2. RESULTS

In the study, analogously to the practice used in the literature, the EWS-GARCH models are evaluated and compared on the basis of the Value-at-Risk forecasts quality, so the quality of states forecasts is not discussed in detail. Nevertheless, it is worth noting that the models for state of turbulence estimated in accordance with the procedure discussed earlier provide a good quality forecasts, as confirmed by the results obtained by Chlebus (2016a).

The discussion of the results for the EWS-GARCH model was divided into two parts. In the first part, results for the EWS-GARCH model with the GARCH(1,1) were presented, and in the second part were the results for the GARCH(1,1) with the amendment to empirical distribution of random error as a model in the state of tranquillity. In order to maintain transparency of the results, a crossover comparison between models of different EWS-GARCH groups (with different state of tranquillity models) was omitted. Additionally, results in this paper for an EWS-GARCH model with particular state of tranquillity and particular state of turbulence VaR forecasting models are presented only for one (with the lowest excess ratio) state of turbulence model. It means that even though in every case Probit, Logit and Cloglog with and without stepwise selection process were considered only best results are presented. All calculations and estimations were performed in SAS 9.4.

4.2.1. VALUE-AT-RISK FORECASTS QUALITY – THE EWS-GARCH(1,1) MODELS

The evaluation of the Value-at-Risk forecasts quality for the EWS-GARCH models began with the EWS-GARCH(1,1) models. Results for the EWS-GARCH(1,1) are presented in two tables. In table 1, results of the Value-at-Risk exceedances and the cost functions are presented, in table 2 results of the coverage tests are presented (same division was made for EWS-GARCH(1,1) with the amendment to empirical error distribution). In the tables only results for models that have lower excess ratio than the GARCH(1,1) are presented.

The GARCH-t(1,1) model is the model with the lowest excess ratio: it has the excess ratio equal to 0.24%, much below expected 1%. After this model, a group of models with the excess ratio smaller (between 0.84% and 0.96%) than 1% may be identified. Those models are: the GARCH(1,1) with the amendment to empirical distribution of random error, and the EWS-GARCH(1,1) models with conservative definition of Value-at-Risk quantile in the state of turbulence. Among the aforementioned EWS-GARCH(1,1) models more conservative are: models assuming the exponential or the empirical distribution, than models assuming the lognormal, the Weibull or the Gamma distributions; and models with the 10% definition of the state of turbulence then models with the 5% definition.

The EWS-GARCH(1,1) models with the liberal definition of Value-at-Risk quantile are generally less conservative and have the excess ratio higher or equal to 1%; the only exception is the EWS-GARCH(1,1) with the exponential distribution, which is rather conservative (the excess ratio equal to 0.89%).

Among the EWS-GARCH models with the lognormal, the Weibull or the Gamma distribution, the most conservative are models with the Weibull distribution; the only exception is the model with a conservative approach defining quantile to forecast Value-at-Risk and the 5% definition of the state of turbulence.

It can also be seen that the Lopez and the Abad and Benito loss functions generally decrease with lowering excess ratio. The EWS-GARCH(1,1) models with the lognormal, the Weibull or the Gamma distributions have higher values of these functions in comparison to models with the exponential or the empirical distributions.

Improvement in the excess ratio and the costs associated with the occurrence of exceeding (expressed by the Lopez and the Abad and Benito cost functions), is associated with an increase in the costs of the model used (expressed by the values of the Caporin and the excess costs functions). The increase in the cost of use of models is growing steadily along with the decrease of the excess ratio. Exceptions are models in which the Value-at-Risk was calculated as the 99th percentile of the exponential, or the Gamma distributions at the 5% definition of the state of turbulence, in which case the increase of the cost of model is significant. It is also worth mentioning that EWS-GARCH models with the lognormal, the Weibull or the Gamma distributions cost less in comparison to models with the exponential or the empirical distributions used to forecast Value-at-Risk in the state of turbulence.

Regarding the Basel Committee backtesting procedure, it can be seen that all models characterized by the lower excess ratio than 1% were assigned to the green zone more than in 90% of cases. Most often the GARCH-t(1,1) (in 98.7% cases) and the GARCH(1,1) with the amendment to empirical distribution of random error (94.9%) were assigned to the green zone. The EWS-GARCH(1,1) models with conservative definition of Value-at-Risk quantile in the state of turbulence and the lognormal, the Weibull or the Gamma distribution were assigned to the green zone in 92.4% cases (the only exception is model with the Gamma distribution and the 10% definition of the state of turbulence). Slightly different results may be found when analysing assignation to at least the yellow zone. In this case, not only the GARCH-t(1,1) has the highest rate (equal to 98.7%), but the EWS-GARCH(1,1) models with the 10% definition of the state of turbulence and the exponential distribution and the EWS-GARCH(1,1) models with the 5% definition of the state of turbulence for any distribution, including the lognormal, the Weibull or the Gamma distribution have it as well. This result is interesting, because models with the lognormal, the Weibull or the Gamma distributions (which are less conservative) are of the same quality (regarding being at least in the yellow zone) as the GARCH-t(1,1) and better than the GARCH(1,1) with the amendment to empirical distribution of random error.

Analysing results for the Stressed Value-at-Risk, again the GARCH-t(1,1) model is the most often assigned to the green and at least the yellow zone (97.5% and 98.7% respectively). Rest of the models drop its quality in terms of the green zone assignment, but keep its quality in terms of being assigned to at least the yellow zone. Again, models with the 5% definition of the state of turbulence, including models with the lognormal, the Weibull or the Gamma distribution are of good quality and are assigned to at least the yellow zone in 93.7% cases.

Analysing results of the coverage tests it can be seen that the smallest rejection rate in the Kupiec test have the EWS-GARCH(1,1) models with the 5% definition of the state of turbulence, the conservative definition of Value-at-Risk quantile and with one out of the lognormal, the Weibull or the Gamma distributions. According to the results of the Christoffersen test, they are not the best but still of good quality (the best is the GARCH(1,1) with the amendment to empirical error distribution).

Very interesting conclusion may be drawn from the asymptotic unconditional coverage test, as this test is two-tailed, and because of that both the overestimation and the underestimation of the Value-at-Risk forecasts may be considered as a reason of rejection of the null hypothesis. According to the obtained test results, it may be stated that for the models with the 5% definition of the state of turbulence, the conservative definition of Value-at-Risk quantile and one out of the lognormal, the Weibull or the Gamma distributions rejections of the null hypothesis due to either the overestimation or the underestimation are on similar level and close to expected (5% for each tail). The Value-at-Risk forecasts from the EWS-GARCH models with the 10% definition of the state of turbulence, the conservative definition of Value-at-Risk quantile and one out of the lognormal, the Weibull or the Gamma distributions are rejected slightly more often, mainly because of the overestimation of forecasts. The models with the liberal definition of Value-at-Risk quantile in the state of turbulence are too liberal and lead to rejection rate due to the underestimation of Value-at-Risk much more often than expected. It should be also stated, that the GARCH-t(1,1) model is far too much conservative and rejected by all the formal tests in most of the cases.

The results obtained for the EWS-GARCH(1,1) with the lognormal, the Weibull or the Gamma distributions in the state of turbulence show that this models provides the Value-at-Risk forecasts of good quality. Taking all the results into account, it seems that the most appropriate are models with the 5% definition of the state of turbulence and the conservative definition of Value-at-Risk quantile in the state of turbulence. They maintain a good balance between conservatism (relatively low excess ratio, low values of the Lopez function and the Abad and Benito function, and relatively high qualification rate to the green zone, and at least the yellow zone in the backtesting procedure) and adequacy (the coverage tests) of the Value-at-Risk forecast. Additionally, regarding the Caporin and the excess cost functions using aforementioned models is relatively not expensive (an exception is the model with the Gamma distribution assumed). All three models (either with the lognormal, the Weibull or the Gamma distributions) exhibit similar quality of the Value-at-Risk forecasts, however among them the most appropriate seems to be the model with the lognormal distribution: it is relatively conservative, with relatively small cost of use.

In the end it is also worth mentioning that the GARCH-t(1,1) model is far too conservative, and in contrast the GARCH(1,1) with the amendment to empirical distribution of random error is very good and in many aspects the best from the analysed models. The GARCH(1,1) model seems to be too liberal, even if used only in the state of tranquillity (it leads to slightly too excessive number of Value-at-Risk exceedances). According to that, it is worth analysing of what quality the Value-at-Risk forecasts provided by the EWS-GARCH(1,1) with the amendment to empirical distribution models would be, as the GARCH(1,1) with the amendment to empirical distribution of random error model is slightly more conservative than the GARCH(1,1) model.

4.2.2. VALUE-AT-RISK FORECASTS QUALITY – THE EWS-GARCH(1,1) WITH THE AMENDMENT TO EMPIRICAL DISTRIBUTION OF RANDOM ERROR MODELS

The results with respect to the exceedances and the cost functions for the EWS-GARCH(1,1) with the amendment to empirical error distribution models are shown in table 3. Results of the coverage tests are presented in table 4.

For the EWS-GARCH(1,1) with the amendment to empirical error distribution only the results of models that improve (reduce) the excess ratio will be discussed. The GARCH(1,1) with the amendment to empirical error distribution is a conservative model itself – the excess ratio on average is smaller than the expected 1%. According to that, choosing EWS-GARCH(1,1) with the amendment to empirical error distribution models that provide the excess ratio closer to 1% than the GARCH(1,1) with the amendment to empirical error distribution, would lead to the choice of models with smaller conservatism than the GARCH(1,1) with the amendment to empirical error distribution in the state of turbulence, which is not a purpose of the EWS-GARCH models development and, therefore, will not be discussed.

As noted above, the GARCH(1,1) with the amendment to empirical error distribution is on average conservative. The average excess ratio is equal to 0.88%. Therefore, reducing excess ratio requires a relatively conservative approach to be used in the state of turbulences. It is possible for all models, assuming the Value-at-Risk is equal to the 99th percentile of a distribution in the state of turbulence. Additionally, reduction

Table 1.

The results of the analysis of the quality of Value-at-Risk forecasts models obtained from the EWS-GARCH(1,1) models

SFM	TSVM	TUSVM	VALUE-AT-RISK (WHOLE OUT-OF-SAMPLE)									STRESSED VALUE-AT-RISK (THE WORST 250 DAYS)				
			EN	ER	ABAD	LOPEZ	CAPORIN	EXCOST	GREEN	YELLOW	RED	EN	ER	GREEN	YELLOW	RED
-	GARCH-t	-	1.25	0.24%	6.3%	2.46	12.5%	11.6%	98.7%	98.7%	1.3%	1.03	0.4%	97.5%	98.7%	1.3%
PROBIT	GARCH	EX9_10	4.39	0.84%	8.4%	4.51	11.1%	10.3%	93.7%	98.7%	1.3%	3.56	1.4%	72.2%	93.7%	6.3%
PROBIT	GARCH	EM9_10	4.58	0.87%	8.8%	4.71	8.7%	7.8%	92.4%	97.5%	2.5%	3.71	1.5%	68.4%	92.4%	7.6%
-	GARCH EMP	-	4.61	0.88%	9.2%	4.67	7.2%	6.4%	94.9%	97.5%	2.5%	3.73	1.5%	68.4%	96.2%	3.8%
PROBIT	GARCH	EX0_10	4.67	0.89%	9.0%	4.80	7.9%	7.1%	92.4%	96.2%	3.8%	3.80	1.5%	65.8%	91.1%	8.9%
CLOGLOG	GARCH	EX9_5	4.76	0.91%	9.0%	4.82	17.1%	16.3%	93.7%	98.7%	1.3%	3.81	1.5%	69.6%	93.7%	6.3%
PROBIT	GARCH	WE9_10	4.82	0.92%	9.5%	4.95	7.7%	6.9%	92.4%	96.2%	3.8%	3.92	1.6%	62.0%	89.9%	10.1%
PROBIT	GARCH	LN9_10	4.86	0.93%	9.4%	4.93	7.6%	6.8%	92.4%	96.2%	3.8%	3.99	1.6%	62.0%	89.9%	10.1%
CLOGLOG	GARCH	EM9_5	4.87	0.93%	9.3%	4.94	8.6%	7.8%	92.4%	98.7%	1.3%	3.92	1.6%	64.6%	93.7%	6.3%
PROBIT	GARCH	GM9_10	4.95	0.94%	9.6%	5.02	7.5%	6.7%	89.9%	96.2%	3.8%	4.05	1.6%	60.8%	88.6%	11.4%
CLOGLOG	GARCH	LN9_5	5.01	0.95%	9.6%	5.02	7.7%	6.9%	92.4%	98.7%	1.3%	4.03	1.6%	62.0%	93.7%	6.3%
CLOGLOG	GARCH	GM9_5	5.04	0.96%	9.6%	5.04	13.1%	12.3%	92.4%	98.7%	1.3%	4.05	1.6%	62.0%	93.7%	6.3%
CLOGLOG	GARCH	WE9_5	5.06	0.96%	9.6%	5.07	7.8%	7.0%	92.4%	98.7%	1.3%	4.06	1.6%	62.0%	93.7%	6.3%
CLOGLOG	GARCH	EX8_5	5.23	1.00%	10.1%	5.23	7.3%	6.5%	91.1%	96.2%	3.8%	4.23	1.7%	58.2%	89.9%	10.1%
CLOGLOG	GARCH	WE0_10	5.91	1.13%	12.1%	5.92	6.9%	6.1%	82.3%	94.9%	5.1%	4.92	2.0%	54.4%	78.5%	21.5%
CLOGLOG	GARCH	WE8_5	5.94	1.13%	12.1%	5.94	6.9%	6.1%	86.1%	92.4%	7.6%	4.85	1.9%	48.1%	83.5%	16.5%
CLOGLOG	GARCH	EM0_10	6.01	1.15%	12.6%	6.02	6.8%	6.0%	82.3%	92.4%	7.6%	4.99	2.0%	50.6%	79.7%	20.3%

CLOGLOG	GARCH	GM8_5	6.14	1.17%	12.7%	6.15	6.8%	6.0%	79.7%	92.4%	7.6%	5.04	2.0%	48.1%	81.0%	19.0%
CLOGLOG	GARCH	LN8_5	6.20	1.18%	13.0%	6.21	6.8%	6.0%	79.7%	92.4%	7.6%	5.09	2.0%	45.6%	81.0%	19.0%
CLOGLOG	GARCH	GM0_10	6.27	1.19%	13.0%	6.27	6.8%	6.0%	79.7%	91.1%	8.9%	5.22	2.1%	44.3%	75.9%	24.1%
CLOGLOG	GARCH	EM8_5	6.29	1.20%	13.2%	6.30	6.7%	5.9%	81.0%	91.1%	8.9%	5.15	2.1%	43.0%	75.9%	24.1%
CLOGLOG	GARCH	LN0_10	6.35	1.21%	13.3%	6.36	6.7%	5.9%	79.7%	91.1%	8.9%	5.30	2.1%	43.0%	75.9%	24.1%
-	GARCH	-	6.42	1.22%	12.5%	6.42	6.6%	5.8%	78.5%	93.7%	6.3%	5.18	2.1%	39.2%	78.5%	21.5%
-	EGARCH	-	6.53	1.24%	12.5%	6.54	6.7%	5.9%	78.5%	92.4%	7.6%	5.19	2.1%	40.5%	81.0%	19.0%

In the table, white fields refer to the EWS-GARCH models with lognormal, Weibull or Gamma distributions, while grey fields to benchmark models.
The following abbreviations are used: SFM – the state forecasting model, TSVM – the Value-at-Risk forecasting model in a state of tranquillity, TUSVM – the Value-at-Risk forecasting model in a state of turbulence, EN - the average number of exceedances, ER – the average excess ratio, ABAD – the average value of the Abad & Benito cost function, LOPEZ – the average value of the Lopez cost function, CAPORIN – the average value of the Caporin cost function, EXCOST – the average value of the excessive cost function, GREEN – the average frequency of a model being in the green zone, YELLOW – the average frequency of a model being at least in the yellow zone, RED – the average frequency of a model being in the red zone. Short names of the Value-at-Risk models in the state of turbulence are in the form DRQ_CP, where the DR defines a distribution of returns, Q defines the quantile for which Value-at-Risk was forecasted and CP defines the cut-off point that was used to forecast the state of turbulence in the states forecasting model. For the distributions in the state of turbulence following abbreviations are used: EX – exponential distribution, EM – empirical distribution, LN – lognormal distribution, WE – Weibull distribution, GM – Gamma distribution; Q equal to 9 represents the 99th percentile, 0 represents the 90th percentile, and 8 represents 80th percentile; 5% cut-off is denoted by 5 and the cut-off point equal to 10% by 10.

Source: own calculations.

Table 2.

The results of the analysis of the quality of Value-at-Risk forecasts obtained from the EWS-GARCH(1,1) models – coverage tests results

SFM	TSVM	TUSVM	LR_{UC}	LR_{IND}	LR_{CC}	Z_{UC}	Z^{D}_{UC}	Z^{G}_{UC}
CLOGLOG	GARCH	WE9_5	5.06%	13.92%	8.86%	11.39%	3.80%	7.59%
CLOGLOG	GARCH	GM9_5	6.33%	12.66%	7.59%	12.66%	5.06%	7.59%
CLOGLOG	GARCH	LN9_5	6.33%	12.66%	7.59%	12.66%	5.06%	7.59%
CLOGLOG	GARCH	WE0_10	6.33%	16.46%	11.39%	18.99%	1.27%	17.72%
-	GARCH EMP	-	7.59%	8.86%	5.06%	10.13%	5.06%	5.06%
CLOGLOG	GARCH	EX8_5	7.59%	11.39%	6.33%	12.66%	3.80%	8.86%
CLOGLOG	GARCH	EX9_5	7.59%	12.66%	8.86%	12.66%	6.33%	6.33%
CLOGLOG	GARCH	EM9_5	7.59%	12.66%	8.86%	13.92%	6.33%	7.59%
-	GARCH	-	8.86%	8.86%	7.59%	24.05%	2.53%	21.52%
CLOGLOG	GARCH	EM0_10	8.86%	15.19%	11.39%	18.99%	1.27%	17.72%
CLOGLOG	GARCH	WE8_5	10.13%	12.66%	8.86%	16.46%	2.53%	13.92%
PROBIT	GARCH	GM9_10	10.13%	10.13%	8.86%	16.46%	6.33%	10.13%
-	EGARCH	-	10.13%	5.06%	8.86%	24.05%	2.53%	21.52%
PROBIT	GARCH	WE9_10	10.13%	8.86%	10.13%	13.92%	6.33%	7.59%
CLOGLOG	GARCH	LN8_5	10.13%	12.66%	10.13%	22.78%	2.53%	20.25%
CLOGLOG	GARCH	GM0_10	10.13%	16.46%	13.92%	21.52%	1.27%	20.25%
CLOGLOG	GARCH	LN0_10	10.13%	16.46%	13.92%	21.52%	1.27%	20.25%
PROBIT	GARCH	LN9_10	11.39%	8.86%	8.86%	15.19%	7.59%	7.59%
PROBIT	GARCH	EX9_10	11.39%	8.86%	11.39%	16.46%	10.13%	6.33%
CLOGLOG	GARCH	EM8_5	11.39%	11.39%	11.39%	21.52%	2.53%	18.99%
PROBIT	GARCH	EX0_10	12.66%	8.86%	10.13%	16.46%	8.86%	7.59%
PROBIT	GARCH	EM9_10	12.66%	8.86%	11.39%	17.72%	10.13%	7.59%
-	GARCH-t	-	77.22%	2.53%	51.90%	77.22%	75.95%	1.27%

In the table, white fields refer to the EWS-GARCH models with lognormal, Weibull or Gamma distributions, while grey fields to benchmark models.
The following abbreviations are used: SFM – the state forecasting model, TSVM – the Value-at-Risk forecasting model in a state of tranquillity, TUSVM – the Value-at-Risk forecasting model in a state of turbulence, LR_{UC} – the ratio of cases in which the null hypothesis was rejected in the Kupiec test, LR_{IND} – the ratio of cases in which the null hypothesis was rejected in the LR_{IND} part of the Christoffersen test, LR_{CC} – the ratio of cases in which the null hypothesis was rejected in the Christoffersen test, Z_{UC} – the ratio of cases in which the null hypothesis was rejected in the asymptotic test of unconditional coverage, Z^{D}_{UC} – the ratio of cases in which the null hypothesis was rejected in the asymptotic test of unconditional coverage in favour of alternative hypothesis that the actual excess ratio is significantly lower than expected, Z^{G}_{UC} – the ratio of cases in which the null hypothesis was rejected in the asymptotic test of unconditional coverage in favour of an alternative hypothesis that the actual excess ratio is significantly higher than expected. All tests were performed for the 5% significance level, except the asymptotic test of unconditional coverage, where level of significance was set up to 10% (5% for each tail).
Short names of the Value-at-Risk models in the state of turbulence are in the form DRQ_CP, where the DR defines a distribution of returns, Q defines the quantile for which Value-at-Risk was forecasted and CP defines the cut-off point that was used to forecast the state of turbulence in the states forecasting model. For the distributions in the state of turbulence following abbreviations are used: EX – exponential distribution, EM – empirical distribution, LN – lognormal distribution, WE – Weibull distribution, GM – Gamma distribution; Q equal to 9 represents the 99^{th} percentile, 0 represents the 90^{th} percentile, and 8 represents 80^{th} percentile; 5% cut-off is denoted by 5 and the cut-off point equal to 10% by 10.

Source: own calculations.

of excess ratio is possible also by the models which assume the liberal approach to forecast Value-at-Risk using the exponential or the Gamma distribution in the state of turbulence. It is worth mentioning that in most cases the best state of turbulence forecasting model was the probit model, the cloglog model was better only once.

It can be seen, as well, that models with the lognormal, the Weibull or the Gamma distributions are less conservative than models with the exponential or the empirical distributions. Among models with the lognormal, the Weibull or the Gamma distributions, models with the 10% definition of the state of turbulence are slightly more conservative than models with the 5% definition, but the differences are not significant.

Use of any of the EWS-GARCH models presented in table 3 reduces the costs associated with the Value-at-Risk exceedances (both based on the Lopez and the Abad and Benito cost functions). For the models with the lognormal, the Weibull or the Gamma distributions slightly better results with respect to Abad and Benito cost function have models with the 5% definition of the state of turbulence.

The EWS-GARCH models with the conservative definition of Value-at-Risk quantile in the state of turbulence are qualified in 100% of cases to the green zone in the backtesting procedure, which is more frequent than in the case of the GARCH(1,1) with the amendment to empirical error distribution, and the much more conservative GARCH-t(1,1) model.

Regarding the Stressed Value-at-Risk values, the GARCH-t(1,1) was the most often assigned to the green zone. However the EWS-GARCH models with the conservative definition of Value-at-Risk quantile in the state of turbulence the exponential or the empirical distribution for both definitions of the state of turbulence, or with the lognormal, the Weibull or the Gamma distributions and the 5% definition of the state of turbulence, were assigned to at least the yellow zone in 100% cases, which is again even more than for the GARCH-t(1,1).

The improvement of all the discussed measures, as in previous cases, is associated with an increase of excess costs of using the model. Again, the excess cost grows steadily with the reduction of excess ratio (except models in which the excessive cost is inappropriately high – it happened in the models assuming that Value-at-Risk forecasts are calculated as the 99th percentile of the exponential or the Gamma distribution with the 5% definition of the turbulent state). Among the EWS-GARCH models the excess costs of using the model are relatively small for models with the lognormal or the Weibull distributions.

In the results of the coverage tests it can be seen that for the EWS-GARCH models with the lognormal, the Gamma or the Weibull distributions the Kupiec test is rejected more often than for the GARCH(1,1) with the amendment to empirical error distribution, but according to the asymptotic unconditional coverage test, this happened only due to the fact that for these models the excess ratios are lower than expected. Moreover, according to the same tests it may be noted that for the EWS-GARCH models analysed the excess ratio is never higher than expected.

Table 3.

The results of the analysis of the quality of Value-at-Risk forecasts obtained from the EWS-GARCH(1,1) models with the amendment to empirical distribution of random error

SFM	TSVM	TUSVM	VALUE-AT-RISK (WHOLE OUT-OF-SAMPLE)											STRESSED VALUE-AT-RISK (THE WORST 250 DAYS)		
			EN	ER	ABAD	LOPEZ	CAPORIN	EXCOST	GREEN	YELLOW	RED	EN	ER	GREEN	YELLOW	RED
-	GARCH-t	-	1.25	0.24%	6.3%	2.46	12.5%	11.6%	98.7%	98.7%	1.3%	1.03	0.4%	97.5%	98.7%	1.3%
PROBIT	GARCH EMP	EX9_10	3.06	0.58%	6.4%	3.27	11.6%	10.7%	100.0%	100.0%	0.0%	2.56	1.0%	88.6%	100.0%	0.0%
PROBIT	GARCH EMP	EM9_10	3.25	0.62%	6.8%	3.48	9.1%	8.3%	100.0%	100.0%	0.0%	2.72	1.1%	86.1%	100.0%	0.0%
PROBIT	GARCH EMP	EX0_10	3.34	0.64%	6.9%	3.52	8.4%	7.5%	100.0%	100.0%	0.0%	2.80	1.1%	86.1%	100.0%	0.0%
PROBIT	GARCH EMP	EX9_5	3.39	0.65%	6.5%	3.48	22.1%	21.3%	100.0%	100.0%	0.0%	2.81	1.1%	87.3%	100.0%	0.0%
PROBIT	GARCH EMP	EM9_5	3.49	0.67%	6.7%	3.59	9.1%	8.3%	100.0%	100.0%	0.0%	2.89	1.2%	87.3%	100.0%	0.0%
PROBIT	GARCH EMP	WE9_10	3.49	0.67%	7.5%	3.73	8.2%	7.4%	100.0%	100.0%	0.0%	2.91	1.2%	86.1%	98.7%	1.3%
PROBIT	GARCH EMP	LN9_10	3.53	0.67%	7.3%	3.67	8.1%	7.3%	100.0%	100.0%	0.0%	2.97	1.2%	86.1%	98.7%	1.3%
PROBIT	GARCH EMP	GM9_10	3.62	0.69%	7.5%	3.77	8.0%	7.2%	100.0%	100.0%	0.0%	3.04	1.2%	82.3%	97.5%	2.5%
PROBIT	GARCH EMP	LN9_5	3.65	0.69%	7.1%	3.70	8.2%	7.4%	100.0%	100.0%	0.0%	3.01	1.2%	86.1%	100.0%	0.0%
PROBIT	GARCH EMP	GM9_5	3.67	0.70%	7.1%	3.72	18.1%	17.2%	100.0%	100.0%	0.0%	3.04	1.2%	86.1%	100.0%	0.0%
PROBIT	GARCH EMP	WE9_5	3.70	0.70%	7.1%	3.75	8.3%	7.5%	100.0%	100.0%	0.0%	3.05	1.2%	86.1%	100.0%	0.0%
PROBIT	GARCH EMP	EX8_5	3.87	0.74%	7.6%	3.93	7.8%	7.0%	100.0%	100.0%	0.0%	3.19	1.3%	78.5%	100.0%	0.0%
CLOGLOG	GARCH EMP	WE0_10	4.54	0.87%	9.7%	4.55	7.4%	6.6%	93.7%	98.7%	1.3%	3.82	1.5%	73.4%	94.9%	5.1%
-	GARCH EMP	-	4.61	0.88%	9.16%	4.67	7.23%	6.43%	94.9%	97.5%	2.5%	3.73	1.49%	68.4%	96.2%	3.8%
-	GARCH	-	6.42	1.22%	12.48%	6.42	6.58%	5.79%	78.5%	93.7%	6.3%	5.18	2.07%	39.2%	78.5%	21.5%
-	EGARCH	-	6.53	1.24%	12.48%	6.54	6.68%	5.90%	78.5%	92.4%	7.6%	5.19	2.08%	40.5%	81.0%	19.0%

In the table, white fields refer to the EWS-GARCH models with lognormal, Weibull or Gamma distributions, while grey fields to benchmark models.
The same abbreviations as in table 1 are used.

Source: own calculations.

Table 4.

The results of the analysis of the quality of Value-at-Risk forecasts models obtained from the EWS-GARCH(1,1) with the amendment to empirical distribution of random error – coverage tests results

SFM	TSVM	TUSVM	LR_{UC}	LR_{IND}	LR_{CC}	Z_{UC}	Z^{D}_{UC}	Z^{G}_{UC}
-	GARCH EMP	-	7.59%	8.86%	5.06%	10.13%	5.06%	5.06%
-	GARCH	-	8.86%	8.86%	7.59%	24.05%	2.53%	21.52%
PROBIT	GARCH EMP	EX8_5	10.13%	8.86%	3.80%	10.13%	10.13%	0.00%
CLOGLOG	GARCH EMP	WE0_10	10.13%	15.19%	8.86%	15.19%	8.86%	6.33%
-	EGARCH	-	10.13%	5.06%	8.86%	24.05%	2.53%	21.52%
PROBIT	GARCH EMP	WE9_5	11.39%	10.13%	6.33%	11.39%	11.39%	0.00%
PROBIT	GARCH EMP	GM9_5	13.92%	8.86%	5.06%	13.92%	13.92%	0.00%
PROBIT	GARCH EMP	LN9_5	13.92%	8.86%	5.06%	13.92%	13.92%	0.00%
PROBIT	GARCH EMP	EM9_5	16.46%	8.86%	7.59%	16.46%	16.46%	0.00%
PROBIT	GARCH EMP	EX9_5	17.72%	8.86%	7.59%	17.72%	17.72%	0.00%
PROBIT	GARCH EMP	GM9_10	17.72%	8.86%	7.59%	17.72%	17.72%	0.00%
PROBIT	GARCH EMP	WE9_10	17.72%	7.59%	10.13%	17.72%	17.72%	0.00%
PROBIT	GARCH EMP	LN9_10	18.99%	7.59%	7.59%	18.99%	18.99%	0.00%
PROBIT	GARCH EMP	EX0_10	20.25%	6.33%	8.86%	20.25%	20.25%	0.00%
PROBIT	GARCH EMP	EM9_10	21.52%	6.33%	12.66%	21.52%	21.52%	0.00%
PROBIT	GARCH EMP	EX9_10	22.78%	6.33%	12.66%	22.78%	22.78%	0.00%
-	GARCH-t	-	77.22%	2.53%	51.90%	77.22%	75.95%	1.27%

In the table, white fields refer to the EWS-GARCH models with lognormal, Weibull or Gamma distributions, while grey fields to benchmark models.
The same abbreviations as in table 2 are used.

Source: own calculations.

5. CONCLUSIONS

Given all the results, it can be stated that the EWS-GARCH models provide Value-at-Risk forecasts with sufficient quality and can be used as the Value-at-Risk forecasting models. The EWS-GARCH models with the lognormal, the Gamma or the Weibull distributions are sufficient alternatives for the EWS-GARCH models with the exponential or the empirical distributions. The models with the lognormal, the Gamma or the Weibull distributions have a bit higher excess ratios, but they also cost less in terms of the excess cost.

Among the models with the lognormal, the Gamma or the Weibull distributions the best seems to be the model with the conservative definition of Value-at-Risk quantile, the 5% definition of the state of turbulence and the lognormal distributions. This model has in both cases a very good relation between the Value-at-Risk quality and the excessive costs of using the model.

Even though the EWS-GARCH models provide Value-at-Risk of good quality and may be used to measure the market risk, it could be improved in the future. Firstly, the states forecasting models may be extended by considering the use of additional variables or incorporating an autoregressive process into the model. Secondly, different Value-at-Risk models in both states may be considered (other GARCH models for the tranquil or turbulent states).

REFERENCES

Abad P., Benito S., (2013), A Detailed Comparison Of Value At Risk Estimates, *Mathematics and Computers in Simulation*, 94, 258–276.

Abad P., Benito S., López, C., (2014), A Comprehensive Review Of Value At Risk Methodologies, *The Spanish Review of Financial Economics*, 12 (1), 15–32.

Alexander C., Lazar E., (2006), Normal Mixture GARCH(1,1): Applications To Exchange Rate Modelling, *Journal of Applied Econometrics*, 21, 307–336.

Alexander C., (2008), *Market Risk Analysis*, John Wiley & Sons, Chichester.

Allison P., (2005), *Logistic Regression Using SAS: Theory And Application*, John Wiley & Sons, Cary, NC: SAS Inst.

Angelidis T., Benos A., Degiannakis S., (2007), A Robust Var Model Under Different Time Periods And Weighting Schemes, *Review of Quantitative Finance and Accounting*, 28 (2), 187–201.

Basel Committee on Banking Supervision, (2006), *International Convergence of Capital Measurement and Capital Standards. Revised framework. Comprehensive Version.*, Online access at 16 November 2013.

Basel Committee on Banking Supervision, (2011), *Revisions to the Basel II market risk framework – updated as of 31 December 2010*, Online access at 16 November 2013.

Bollerslev T., (1986), Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, 307–327.

Brownlees C. T., Engle R. F., Kelly B. T., (2011), A Practical Guide To Volatility Forecasting Through Calm And Storm, *SSRN Electronic Journal SSRN Journal*, Online access at 16 November 2013.

Cai J., (1994), A Markov Model Of Switching-Regime ARCH, *Journal of Business Economic Statistics*, 12(3), 309–316.

Caporin M., (2008), Evaluating Value-At-Risk Measures In The Presence Of Long Memory Conditional Volatility, *Journal of Risk*, 10, 79–110.

Chan F., Mcaleer M., (2002), Maximum Likelihood Estimation Of STAR and STAR-GARCH Models: Theory And Monte Carlo Evidence, 17 (5), 509–534.

Chlebus M., (2016a), One-Day Prediction Of State Of *Journal Of Applied Econometrics* Turbulence For Portfolio. Models For Binary Dependent Variable, *Working Papers 2016-01*, Faculty of Economic Sciences, University of Warsaw.

Chlebus M., (2016b), EWS-GARCH: New Regime Switching Approach To Forecast Value-At-Risk, *Working Papers 2016-06*, Faculty of Economic Sciences, University of Warsaw.

Degiannakis S., Floros C., Livada A., (2012), Evaluating Value-At-Risk Models Before And After The Financial Crisis Of 2008, *Managerial Finance*, 38 (4), 436–452.

Dimitrakopoulos D. N., Kavussanos M. G., Spyrou S. I., (2010), Value At Risk Models For Volatile Emerging Markets Equity Portfolios, *The Quarterly Review of Economics and Finance*, 50 (4), 515–526.

Engle R., (2001), GARCH 101: The Use Of ARCH/GARCH Models In Applied Econometrics, *Journal of Economic Perspectives*, 15 (4), 157–168.

Engle R., (2004), Risk And Volatility: Econometric Models And Financial Practice, *American Economic Review*, 94 (3), 405–420.

Engle R., Manganelli S., (1999), Caviar: Conditional Autoregressive Value At Risk By Regression Quantiles, *NBER Working Paper 7341*.

Engle R., Manganelli S., (2001), Value-At-Risk Models In Finance, Frankfurt am Main 2001. On Line (available at Social Science Research Network). Online access at 16 November 2013.

Gray S. F., (1996), Modeling The Conditional Distribution Of Interest Rates As A Regime-Switching Process, *Journal of Financial Economics*, 42 (1), 27–62.

Hamilton J. D., Susmel R., (1994), Autoregressive Conditional Heteroskedasticity And Changes In Regime, *Journal of Econometrics*, 64 (1–2), 307–333.

Lopez J. A., (1999), Methods for Evaluating Value-at-Risk Estimates, *Federal Reserve Bank of San Francisco Economic Review*, 2, 3–17.

Marcucci J., (2005), Forecasting Stock Market Volatility With Regime-Switching GARCH Model, *Studies in Nonlinear Dynamics & Econometrics*, 9 (4), 1–55.

McAleer M., Jiménez-Martin J., Amaral T. P., (2009), Has The Basel II Accord Encouraged Risk Management During The 2008-09 Financial Crisis?, *SSRN Electronic Journal SSRN Journal*, Online access at 16 November 2013.

Nelson D., (1991), Conditional Heteroscedasticity In Asset Returns: A New Approach, *Econometrica*, 59, 347–370.

Ozun A., Cifter A., Yılmazer S., (2010), Filtered Extreme-Value Theory For Value-At-Risk Estimation: Evidence From Turkey, *The Journal of Risk Finance*, 11 (2), 164–179.

Panjer H. H., (2006), *Operational Risk: Modeling Analytics*, Wiley Interscience, New York.

CZY ZASTOSOWANIE ROZKŁADÓW LOGNORMALNEGO, WEIBULLA LUB GAMMA MOŻE POPRAWIĆ PROGNOZY WARTOŚCI NARAŻONEJ NA RYZYKO UZYSKIWANE NA PODSTAWIE MODELI EWS-GARCH?

Streszczenie

W badaniu analizie poddane zostały dwustopniowe modele EWS-GARCH służące do prognozowania wartości narażonej na ryzyko. W ramach analizy rozpatrywane były modele EWS-GARCH zakładające rozkłady lognormalny, Weibulla oraz Gamma w stanie turbulencji oraz modele GARCH(1,1) i GARCH(1,1) z poprawką na rozkład empiryczny w stanie spokoju.

Ocena jakości prognoz Value-at-Risk uzyskanych na podstawie wspomnianych modeli została przeprowadzona na podstawie miar adekwatności (wskaźnik przekroczeń, test Kupca, test Christoffersena, test asymptotyczny bezwarunkowego pokrycia oraz kryteria *backtestingu* określone przez Komitet Bazylejski) oraz analizy funkcji strat (kwadratowa funkcja straty Lopeza, absolutna funkcja straty Abad i Benito, 3 wersja funkcji straty Caporina oraz funkcja nadmiernych kosztów). Uzyskane wyniki wskazują, że modele EWS-GARCH z rozkładem lognormalnym, Weibulla lub Gamma mogą konkurować z modelami EWS-GARCH z rozkładem wykładniczym lub empirycznym. Modele EWS-GARCH z rozkładem lognormalnym, Weibulla lub Gamma są nieco mniej konserwatywne, jednocześnie jednak koszt ich stosowania jest mniejszy niż modeli EWS-GARCH z rozkładem wykładniczym lub empirycznym.

Słowa kluczowe: wartość zagrożona (Value-at-Risk), modele GARCH, modele zmiany stanu, prognozowanie, ryzyko rynkowe

CAN LOGNORMAL, WEIBULL OR GAMMA DISTRIBUTIONS IMPROVE THE EWS-GARCH VALUE-AT-RISK FORECASTS?

Abstract

In the study, two-step EWS-GARCH models to forecast Value-at-Risk are analysed. The following models were considered: the EWS-GARCH models with lognormal, Weibull or Gamma distributions as a distributions in a state of turbulence, and with GARCH(1,1) or GARCH(1,1) with the amendment to empirical distribution of random error models as models used in a state of tranquillity.

The evaluation of the quality of the Value-at-Risk forecasts was based on the Value-at-Risk forecasts adequacy (the excess ratio, the Kupiec test, the Christoffersen test, the asymptotic test of unconditional coverage and the backtesting criteria defined by the Basel Committee) and the analysis of loss functions (the Lopez quadratic loss function, the Abad & Benito absolute loss function, the 3rd version of Caporin loss function and the function of excessive costs). Obtained results show that the EWS-GARCH models with lognormal, Weibull or Gamma distributions may compete with EWS-GARCH models with exponential and empirical distributions. The EWS-GARCH model with lognormal, Weibull or Gamma distributions are relatively less conservative, but using them is less expensive than using the other EWS-GARCH models.

Keywords: Value-at-Risk, GARCH models, regime switching, forecasting, market risk