BUSINESS ANALYTICS AND OPTIMIZATION 1

BUSINESS ANALYTICS AND OPTIMIZATION

MEASURING SOCIAL PREFERENCES USING EXPERIMENTAL GAMES: LITERATURE REVIEW

Marija Kuzmanović*[1], Tijana Blažić[1]
[1]University of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: marija.kuzmanovic@fon.bg.ac.rs

Abstract: *In the last few decades, there is an increased interest in social preferences. They point to the fact that individuals are not rational all the time, although it may be desirable. For measuring the social preferences a variety of experimental games can be used. This paper points to the possibility of measuring the social preferences using the experimental games and focuses on two experimental games: ultimatum and dictator game. The results of experiments which use the ultimatum game, have shown that social preferences influence on decision making, as well as factors such as cultural norms, gender have effect to the social preferences. As far as the dictator game, the results indicates on the fact that factors (altruism, social pressure, empathy and others) have influence on the giving in this game.*

Keywords*: social preferences, behavioral game theory, experimental games, ultimatum game, dictator game*

1. INTRODUCTION

One of the standard assumptions in economics is that individuals are motivated by only their material self – interest (Baran et al, 2010). Although Adam Smith in his *Theory of Moral Sentiments* (1759) recognized the possibility that some individuals exhibit other types of interests, it is only in the last few decades have seen a strong course of interest in what are now widely known in economics as "social preferences". Social preferences refer to the phenomena that people seem to care about certain ~~-s~~ocial" goals, such as well-beings of other individuals, or a ~~-f~~air" allocation among members in society, in addition to their own material benefits (Li, 2008).

In their seminal paper, Fehr and Schmidt (1999) suggested that while many people might be motivated only by pure self-interest, this need not be the case for all people. They theorized that some people may be basically motivated by fairness concerns, reciprocity (reward kindness and punish unkindness) and even pure (unmotivated) altruism (Baran et al, 2010) but also they sometimes act in the entire group's interest, even if it hurts some individuals in the group (Li, 2008).

While the study of social behaviors has a long tradition in economics, psychology and sociology, recently, the term social preferences has come to be associated with a more specific program of research originated mainly in experimental and behavioral economics (Galizzi, & Navarro Martinez, 2015).

A considerable body of experiments has examined the effects of various demographic variables on the social preferences (see Wilkinson, & Klaes, 2012). The variable most frequently studied are gender (Heinz et al, 2011), age (Gummeru et al, 2010; Wittig et al, 2013), academic major (Kahneman et al, 1986; Kuzmanovic et al, 2012), culture (Henrich et al, 2005), social distance (Charness & Gneezy, 2008) and religious identity (Brañas-Garza et al, 2014; Gomes & McCullough, 2015).

Other factors whose effects on social preferences experimentally examined are: anonymity (Dufwenberg, & Muren, 2006; List, 2007; Bardsley, 2008; Franzen & Pointer, 2012; Winking & Mizer, 2013); the context and framing (Branas-Garza, 2007; Stoop, 2014) the size of the stakes (Carpenter et al., 2005; List and Cherry, 2008); the subject pool (Cleave et al, 2013; Exadaktylos et al, 2013; Stoop, 2014). One of the recent studies (Arfer et al, 2015) suggests that people may behave prosocially not only because they value the welfare of others, but also to protect their own reputation.

In this paper will be shown some of the experimental games which can be used for measuring the social preferences. It will also be shown an overview of the results of experiments in which have been used the experimental games.

2. SOCIAL PREFERENCE MODELS

There are essentially two types of social preferences: distributive and reciprocal. Distributive preferences assume that individuals have preferences over final payoff allocations (Li, 2008), while reciprocal preferences mean that individuals desire to reward or punish others beyond mere consequences. Accordingly, the existing literature on social preferences distinguishes three main classes of models (Schmidt, 2011): (1) outcome-based social preferences, (2) intention-based social preferences, and (3) type dependent social preferences.

In models of *outcome-based social preferences,* the utility of each player depends directly on the material payoff of other players in his reference group. For example, the theories of unconditional altruism (Andreoni & Miller, 2002) assume that the utility of a player not only depends on his own material payoff but it is also monotonically increasing in the monetary payoff of other players. Charness and Rabin (2002) show that subjects are more concerned with increasing social welfare (sacrificing to increase the payoffs for all recipients, especially low-payoff recipients) than with reducing differences in payoffs. The Fehr-Schmidt model is consistent with observed behavior in many experimental games but it assumes that people care only about outcomes and not about intentions.

Models of *intention-based social preferences* try to capture intention effects and incorporate the idea that players benefit from reciprocating (or punishing) acts to which they attribute good (or bad) intentions (Netzer, & Schmutzler, 2014). In these models, preferences are defined not only over outcomes but also over beliefs about why a player has chosen certain actions. Namely, players reward kind and punish unkind intentions. Rabin (1993) assumes that the social preferences are driven by players‘ kindness towards each other: if a player believes the opponent‘s action is motivated by kindness toward him, then he prefers to react kindly; and vice versa. According Falk et al. (2003) players are punishing much more in case that can be inferred that exist a greedy intention or personality. The *intention-based* model successfully accounts for retaliatory and altruistic behavior (Li, 2008) and requires the use of psychological game theory (Dufwenberg and Kirchsteiger 2004, Falk and Fischbacher 2006).

Finally, in type-based models, the level of the payoff to another player depends on one‘s own (distributional) basic attitude as well as the perceived basic attitude of the other player (Kerschbamer, 2015). In other words, in type-based social preferences models (Rotemberg, 2008), a player behaves kindly to a ~~„g~~ood" person and hostile towards a ~~„b~~ad" person. For instance, in Levine (1998) people are altruistic (or spiteful) to various degrees and – given their own basic altruism – they place more positive (negative) weight on the money received by an opponent who is believed to be more altruistic (spiteful). Andreoni and Bernheim (2009) have a similar model in which others‘ other-regarding concerns are evaluated in the terms of inequality aversion (and not in the terms of altruism).

3. BEHAVIORAL GAME THEORY

Game theory is the study of mathematical models of conflict and cooperation between intelligent rational decision-makers (Myerson, 1991). It is multiplayer decision theory where the choices of each player affect the payoffs to other players, and the players take this into account in their choice behavior (Gintis, 2009). Game theory provides the conceptual and procedural tools for studying the social interaction, including the characteristics of the players, the rules of the game, the informational structure, and the payoffs associated with particular strategic interactions.

Game theory is designed to be flexible enough to be used at many levels of detail in a broad range of sciences. Players may be people, groups, companies or nations. Strategies may be either an action taken by a player at some point during the play of a game or a complete plan of actions i.e. algorithm for playing the game. Outcomes can be anything players value – money, prestige, corporate profits or captured territory (Camerer, & Fehr, 2002).

The most central concept in game theory is a solution concept of non-cooperative games, known as Nash equilibrium. A set of strategies (one for each player) form an equilibrium if each player is choosing the strategy which is a best response to the other players‘ strategies. A common assumption underlies an equilibrium is that players act rationally. However, in practice (in the real world situations), human behavior often deviates from this model.

Behavioral game theory is about what players *actually* do (Camerer, 2003). It expands analytical theory by adding emotion, mistakes, limited foresight, doubts about how smart others are. Actually, behavioral game theory analyzes interactive strategic decisions and behavior using the methods of game theory, experimental economics, and experimental psychology. It attempts to explain decision making using experimental data,

and allows for rational and irrational decisions because both are examined using real-life experiments. Specifically, behavioral game theory attempts to explain factors that influence real world decisions (Colman, 2003).

Several experimental games, particularly the ultimatum game, dictator game, trust game, gift exchange game, and public goods game, are widely used to measure the impact and magnitude of the social preferences. In this paper will be shown two types of experimental games, the ultimatum and dictator game.

3.1. Ultimatum game

Using results of experiments with Ultimatum game, some studies confirmed that the social preferences certainly influence the decision-making process as well as the interactions that people have with each other (Ciampaglia et al, 2014). Also, some authors used the Ultimatum game to examine whether social norms and beliefs affect the distribution of money (Camerer & Fehr, 2002). It is certainly of great importance because social norms determine the type of social preferences, specifically directed individual towards pro-social type of social preferences. If it turns out the fact that the individuals affected by social norms, individuals will be less selfish or completely unselfless during the resource allocation (Azar et al, 2015).

Ultimatum games represent a form of take-it-or-leave-it bargaining (Güth et al, 1982). The first player (proposer) receives a sum of money (S) proposes how to divide the sum between the proposer and the other player. The second player (responder) chooses to either accept offer x, or reject it, in which case neither player receives any money. If the Responder accepts he earns x and the Proposer earns $S - x$. The game is typically played only once (see Figure 1).

In theory, rational (self-interested) responders will accept any positive offer, and proposers who anticipate this should offer the smallest possible positive amount because for the responder something is better than nothing at all. The ultimatum game measures whether responders will negatively reciprocate, sacrificing their own money to punish a proposer who has been unfair.

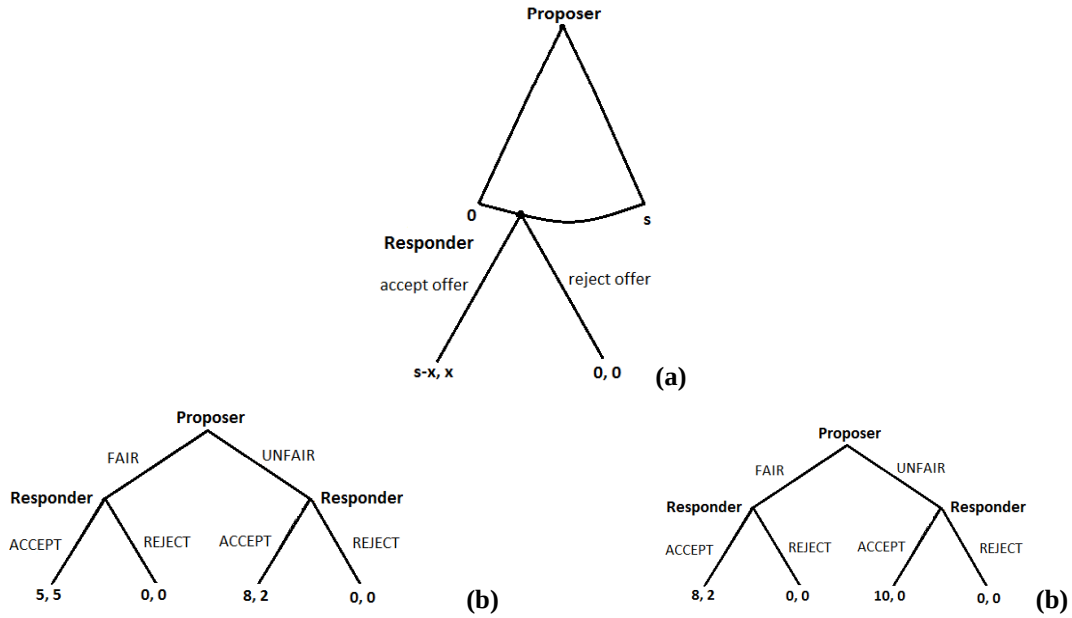

Figure 1: (a) Classic ultimatum game, (b) mini ultimatum games

The experimental results are in stark contrast to the rational solution. On average, proposers offer about 40% of the initial amount (S) to the responder, while Responders reject offers less than 20% of S about half the time (Camerer, 2003). It has also been found that delaying the responder's decision makes people accept "unfair" offers more often (Oechssler et al, 2015).

As Falk and Fischbacher (2006) pointed out, two types of fairness considerations are at work when playing the ultimatum game: outcome-based fairness, in which one compares one‘s own gains and losses relative to those of others, and intention-based fairness, in which one also takes into account the motives lying behind the outcomes (Wittig et al, 2013). Falk et al. (2003) consider a mini-ultimatum game which directly addresses these two possibilities. In this form of game, the proposer is restricted to choose between two different

allocations. In one treatment he can choose between (8,2) and (5,5), in another treatment he can choose between (8,2) and (10,0). In the first treatment the choice of (8,2) may be considered greedy, because he could have offered the equal split (5,5). In the second treatment the choice of (8,2) may be considered generous, because the only alternative would have been (10,0). (see Fugure 1, Schmidt, 2011).

Wittig et al, 2013 used mini-ultimatum game to determine the nature of fairness sensitivity at a period before formal schooling begins (five years olds). They found that preschool children seem to understand „fair = equal" in this task, but not much more, and they are not yet skillful at anticipating what others will find fair beyond 50/50 splits.

The responders took into account the options the proposer had available when he proposed a particular division of resources, including ones that were not equal (Falk et al., 2003). On the other hand, proposers take responder's perspective into account in that they tend to choose the fairest possible alternative and never make offers of zero. For example, when the proposer chooses between offering 20% or 75%, an offer of 20% is rejected 33% of the time. On the other hand, when the proposer's choice set is changed to (20%, 87.5%), the rejection rate for an offer of 20% drops to 16% (Charness and Rabin 2002; Li, 2008).

Some authors suggest that making an fair offer is indeed a rational behavior on the part of proposers because they should take responders' perspective into account (assuming responders are not rational in the standard economic sense) and make an "optimal" offers, that is, those that have a high probability of being accepted while still giving a high payoff to the proposer (Wittig et al, 2013). Harbaugh et al. (2007) suggested that in the practice, people make "optimal" offers, that is, those that have a high probability of being accepted while still giving a high payoff to the proposer.

Many studies have shown that both the proposer's and responder's behavior relies on what they consider to be a fair distribution, sometimes is influenced by general cultural norms (Henrich et al., 2005). The first detailed study in 1991 by Roth et al. examined subjects in four countries: the USA, Japan, Israel and Yugoslavia. They find out that the highest offers are made in the USA and Yugoslavia, and the lowest offers are made in Israel, with Japan in the middle. The main finding was that offers were on average about 10% in Japan and Israel. Interestingly, countries with lower offers did not have higher rates of disagreement. Instead, somewhat lower offers were accepted in countries where low offers were made. Henrich (2000) reports differences between 18 - 30 olds Machiguenga men of the Peruvian Amazon and students at UCLA.

Wilkinson and Klaes (2012) emphasize that although there are a number of significant differences between the genders in terms of fairness and social preferences, there is no simple pattern. Eckel and Grossman, (2001) experimentally demonstrated that in ultimatum game both sexes make similar offers, while woman reject less often. In addition, some authors are using the Ultimatum game pointed to the preferences of people towards the so-called "round" amounts of money, namely that people are choosing a specific amount, for example, half, quarter, and so on (Azar et al, 2015).

3.2. Dictator Game

The dictator game is similar to the ultimatum game. In this game, the first player, (dictator), determines how to split an endowment (such as a cash prize) between himself and the second player (recipient). The recipient simply receives the remainder of the endowment left by the dictator and his role is entirely passive and has no input into the outcome of the game.

Since the dictator's outcome depends only on his own actions, Dictator game is considered as degenerate game and has been used to test the "*Homo economicus*" model of individual behavior. Model "*Homo economicus*" means that individual is concerned with their own economic well being, and dictators would allocate the entire good to themselves and give nothing to the recipient. However, experimental results offer evidence against the rationally self-interested individual concept (Bardsley, 2008).

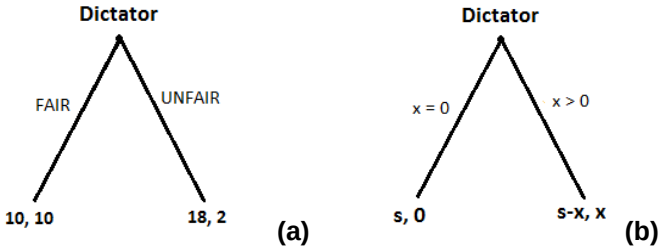

Figure 2: a) Dictator game (Kahneman et al. 1986), b) general form of the Dictator game

The first dictator game experiment was conducted by Kahneman et al, (1986). Authors gave subjects a choice either to split of $20 ($10 each) with another student or an uneven split ($18, $2), favoring themselves. Majority of the students (75%) chooses the equal split. Forsythe et al. (1994) later carried out the dictator game with real stakes and obtained similar results. They find that 70% of the dictators actually do not choose to leave entire endowment for themselves, leaving the recipient with on average 24% of the total amount. Similar patterns have been found in many different dictator game settings. As Camerer (2003) points out, a plenty of subsequent dictator experimental studies replicate these results, finding that usually more than 60 percent of subjects give away a positive amount of money, with the mean transfer roughly 20 percent of the endowment. This has induced numerous researchers to investigate the motivation for such behavior.

The most common explanations for giving in experimental dictator game are altruism and social pressure (Akerlof & Kranton, 2000). Eckel and Grossman (1996) proposed a variant of the dictator game in which a subject decides how to allocate a sum of money between herself and a charity. Moreover, Fong (2007) has shown that an empathy is a more important motivation for dictator game giving when recipients are perceived to be in great need. Additional experiments have shown that subjects maintain a high degree of consistency across multiple versions of the dictator game in which the cost of giving varies (Andreoni, & Miller, 2002).

Most dictator game studies involve only an anonymous recipient. In a few studies, the participants are informed about one or two characteristics of the recipients like the gender (Dufwenberg and Muren 2006) or the name of the organization (DellaVigna et al, 2012). Some studies suggest that Dictator giving tends to increase when weakening anonymity across subjects (Burnham, 2003; Engel, 2011) or when adding an additional observer (Cason and Mui, 1997). However, Bolton et. al (1998) find no significant experimenter anonymity effect.

It has been also argued that giving in dictator games might arise from a desire of subjects to appear acting in a socially appropriate way because their individual behavior is observed by others (Hoffman et al. 1994). Since dictators‘ behavior might also be influenced by expectations about the experimenter‘s judgement, Frequency of positive offers and average donations drop significantly relative to Forsythe et al. (1994). On the other hand, Koch & Normann (2008) suggest that when strategic considerations are eliminated it does not matter to an individual how his or her actions are perceived by others.

Razzaque (2009) experimentally demonstrated that receiver's gender is an important factor of influence on the decision on giving. Some researcher suggests that women are more generous than men (Aguiar et al, 2009) even when recipients have to expert effort to earn rewards as opposed to receiving them as gifts (Heinz et al, 2011). In modified versions of the dictator game, children also tend to allocate some of a resource to a recipient and most five-year-olds share at least half of their goods (Gummerum et al, 2010).

Shariff and Norenzayan (2007) investigated the effect of religious identity on generosity in a dictator game. They found a significant positive effect of religious identity on generosity; the religiously primed participants allocated more money to anonymous strangers than did participants who were not primed. Ahmed and Salas (2011) provide more evidence for religious identity being positively related to generosity in a dictator game.
Brañas-Garza et al, (2014) also explored the relationship between several personal religion-related variables and social behavior, using dictator and ultimatum games. They found that individuals with ―no religion" made decisions closer to rational selfish behavior compared to those who affiliate with a ―standard" religious denomination. However, some studies suggest that the effect of religious priming on generosity in the dictator game may be more limited than previously thought (Gomes & McCullough, 2015).

Forsythe et al. (1994) were the first to compare offers in the ultimatum game to offers in the dictator game. They find out that offers in the dictator game are considerably less generous. A significant fraction of dictators give nothing. This shows that many proposers are generous only for strategic reasons in the ultimatum game.

4. CONCLUSION

Economics argues that each individual is rational, or whether it is really so. Even when it is desirable to be rational does not mean that this is actually happening, and that all people are motivated by self-interests (Fehr and Schmidt, 1999). The experimental results offer evidence against the rationally self-interested individual concept. In recent decades, it appears increased interest in social preferences. The reason is that

they help to understand the behavior of individuals and to discover the factors of influence on the decision-making process.

The review literature suggests ways of measuring the social preferences using the experimental games. As reported throughout the current review, using the ultimatum game, has been discovered that social preferences influence on decision-making. It have been also discovered factors which have influence such as social norms, beliefs (Camerer & Fehr, 2002), cultural norms and gender. It has also been found that delaying the responder's decision makes people accept "unfair" offers more often. According to some authors, people are choosing a specific amount, so-called "round" amounts of money.

Observing the review, using the dictator game, as results in experiments appear different factors for giving: altruism, social pressure, gender, religious identity, and empathy. It has been also argued that giving might arise from a desire of subjects to appear acting in a socially appropriate way because their individual behavior is observed by others.

This paper presents the results of experiments where the experimental games are used for measuring the social preferences. However, researchers should not be limited to the measurement of social preferences using the experimental games only. The direction of future experiments, in this area, can relate to new methods of measurement of the social preferences. Also, some subsequent experiments may relate to new factors and examine their impact on social preferences.

REFERENCES

Aguiar, F., Brañas-Garza, P., Cobo-Reyes, R., Jimenez, N., & Miller, L.M. (2009). Are women expected to be more generous?. *Experimental Economics*, *12*(1), 93-98.

Ahmed, A., & Salas, O. (2011). Implicit influences of Christian religious representations on dictator and prisoner's dilemma game decisions. *The Journal of Socio-Economics , 40* (3), 242-246.

Akerlof, G.A., & Kranton, R.E. (2000). Economics and identity. *Quarterly journal of Economics*, 715-753.

Andreoni, J., & Bernheim, B.D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607-1636.

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737-753.

Arfer, K.B., Bixter, M.T., & Luhmann, C.C. (2015). Reputational concerns, not altruism, motivate restraint when gambling with other people's money. *Frontiers in psychology*, *6*.

Azar, O.H., Lahav, Y., & Voslinsky, A. (2015). Beliefs and social behavior in a multi-period ultimatum game. *Frontiers in behavioral neuroscience*, *9*.

Baran, Carpenter N.M., Sapienza, P., & Zingales, L. (2010). *Can we infer social preferences from the lab? Evidence from the trust game* (No. w15654). National Bureau of Economic Research.

Bardsley, N. (2008). Dictator Game giving: Altruism or artefact? Experimental Economics, 11(2), 122-133.

Brañas-Garza, P. (2007). Promoting helping behavior with framing in dictator games. *Journal of Economic Psychology*, *28*(4), 477-486.

Brañas-Garza, P., Espín, A.M., & Neuman, S. (2014). Religious pro-sociality? Experimental evidence from a sample of 766 Spaniards. PloS one, 9(8), e104685.

Brañas-Garza, P., Espín, A.M., Exadaktylos, F., & Herrmann, B. (2014). Fair and unfair punishers coexist in the Ultimatum Game. *Scientific reports*, *4*.

Burnham, T., 2003. Engineering altruism: a theoretical and experimental investigation of anonymity and gift-giving. Journal of Economic Behavior and Organization 50, 1–13.

Bolton, G. E., R. Zwick, and E. Katok (1998): –Dictator game giving: Rules of fairness versus acts of kindness," *International Journal of Game Theory*, 27, 269–299.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, C.F., & Fehr, E. (2002). Measuring social norms and preferences using experimental games: A guide for social scientists.

Carpenter, J., Verhoogen, E., & Burks, S. (2005). The effect of stakes in distribution experiments. *Economics Letters*, *86*(3), 393-398.

Cason, T. N., & Mui, V. L. (1997). A Laboratory Study of Group Polarisation in the Team Dictator Game. *The Economic Journal*, *107*(444), 1465-1483.

Charness, G., & Gneezy, U. (2008). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization*, *68*(1), 29-35.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly journal of Economics*, 817-869.

Ciampaglia, G.L., Lozano, S., & Helbing, D. (2014). Power and fairness in a generalized ultimatum game. *PloS one*, *9*(6), e99039.
Cleave, B.L., Nikiforakis, N., & Slonim, R. (2013). Is there selection bias in laboratory experiments? The case of social and risk preferences. *Experimental Economics*, *16*(3), 372-382.
Colman, A.M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and brain sciences*, *26*(02), 139-153.
DellaVigna, S., List, J. A., & Malmendier, U. (2009). *Testing for altruism and social pressure in charitable giving* (No. w15629). National Bureau of Economic Research.
Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and economic behavior*, *47*(2), 268-298.
Dufwenberg, M., & Muren, A. (2006). Generosity, anonymity, gender. *Journal of Economic Behavior & Organization*, *61*(1), 42-49.
Eckel, C.C., & Grossman, P.J. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry*, *39*(2), 171-188.
Eckel, C.C., & Grossman, P.J. (1996). Altruism in anonymous dictator games. *Games and economic behavior*, *16*, 181.
Engel, C. (2011). Dictator games: a meta study. Experimental Economics, 14(4), 583-610, doi:10.1007/s10683-011-9283-7.
Exadaktylos, F., Espín, A.M., & Branas-Garza, P. (2013). Experimental subjects are not different. *Scientific reports*, *3*.
Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and economic behavior*, *54*(2), 293-315.
Falk, A., Fehr, E., & Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry*, *41*(1), 20-26.
Fehr, E., & Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *Quarterly journal of Economics*, 817-868.
Fong, C.M. (2007). Evidence from an experiment on charity to welfare recipients: Reciprocity, altruism and the empathic responsiveness hypothesis. *The Economic Journal*, *117*(522), 1008-1024.
Forsythe, R., Horowitz, J.L., Savin, N.E., and Sefton, M. (1994). Fairness in Simple Bargaining Experiments. *Games and Economic Behavior*, 6, 347-369.
Franzen, A., and Pointner, S. (2012). Anonymity in the dictator game revisited. *Journal of Economic Behavior and Organization*, 81(1), 74-81.
Galizzi, M.M., & Navarro Martinez, D. (2015). On the external validity of social-preference games: A systematic lab-field study.
Gintis, H. (2009). The bounds of reason: Game theory and the unification of the behavioral sciences. Princeton University Press.
Gomes, C.M., & McCullough, M.E. (2015). The effects of implicit religious primes on dictator game allocations: A preregistered replication experiment. *Journal of Experimental Psychology: General*, 144(6).
Gummerum, M., Hanoch, Y., Keller, M., Parsons, K., & Hummel, A. (2010). Preschoolers‘ allocations in the dictator game: The role of moral emotions. *Journal of Economic Psychology*, *31*(1), 25-34.
Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, *3*(4), 367-388.
Harbaugh, W.T., Krause, K., & Vesterlund, L. (2007). Learning to bargain. *Journal of Economic Psychology*, 28(1), 127-142.
Heinz, M., Juranek, S., & Rau, H.A. (2011). Do Women Behave More Reciprocally than Men. *Gender Differences in Real Effort Dictator Games*, 105-110.
Henrich, J. (2000). Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *The American Economic Review*, *90*(4), 973-979.
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. (2005). „Economic man“ in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28, 795–855.
Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, *7*(3), 346-380.
Kahneman, D., Knetsch, J.L., & Thaler, R.H. (1986). Fairness and the assumptions of economics. *Journal of business*, S285-S300.
Kerschbamer, R. (2015). The geometry of distributional preferences and a non-parametric identification approach: The Equality Equivalence Test. *European economic review*, *76*, 85-103.
Koch, A.K., & Normann, H.T. (2008). Giving in dictator games: Regard for others or regard by others?. *Southern Economic Journal*, 223-231.
Kuzmanovic, M., Djurovic, M., & Martic, M. (2012). Factors affecting cooperation in social dilemmas: Experimental evidence from one-shot prisoner's dilemma. *Technics Technologies Education Management-TTEM*, *7*(1), 384-394.
Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of economic dynamics*, *1*(3), 593-622.

Li, J. (2008). The power of conventions: A theory of social preferences. *Journal of Economic Behavior & Organization*, *65*(3), 489-505.
List, J.A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, *115*(3), 482-493.
List, J.A., and Cherry T.L. (2008), Examining the role of fairness in high stakes allocation decisions. *Journal of Economic Behavior & Organization* 65(1), 1-8.
Myerson, R. (1991). Game Theory: Analysis of ConflictHarvard Univ. Press, Cambridge.
Netzer, N., & Schmutzler, A. (2014). EXPLAINING GIFT- EXCHANGE—THE LIMITS OF GOOD INTENTIONS. *Journal of the European Economic Association*, *12*(6), 1586-1616.
Oechssler, J., Roider, A., & Schmitz, P.W. (2015). Cooling-Off in Negotiations-Does It Work?. *Journal of Institutional and Theoretical Economics JITE* ***171*** *(4): 565–588.*
Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American economic review*, 1281-1302.
Razzaque, S. (2009). The ultimatum game and gender effect: Experimental evidence from pakistan. *The Pakistan Development Review*, 23-46.
Rotemberg, J.J. (2008). Minimally acceptable altruism and the ultimatum game. *Journal of Economic Behavior & Organization*, *66*(3), 457-476.
Roth, A.E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study. *The American Economic Review*, 1068-1095.
Schmidt, K.M. (2011). Social preferences and competition. *Journal of Money, credit and Banking*, *43*(s1), 207-231.
Shariff, A.F., & Norenzayan, A. (2007). God is watching you priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological science*, *18*(9), 803-809.
Stoop, J. (2014). From the lab to the field: envelopes, dictators and manners. *Experimental Economics*, *17*(2), 304-313.
Wilkinson, N., & Klaes, M. (2012). *An introduction to behavioral economics*. Palgrave Macmillan.
Winking, J., & Mizer, N. (2013). Natural-field dictator game shows no altruistic giving. *Evolution and Human Behavior*, *34*(4), 288-293.
Wittig, M., Jensen, K., & Tomasello, M. (2013). Five-year-olds understand fair as equal in a mini-ultimatum game. *Journal of experimental child psychology*, *116*(2), 324-337.

AHP GROUP DECISION MAKING: EXAMPLE OF SELECTING EVALUATORS FOR EVALUATION OF PERFORMANCE OF MILITARY MOTOR VEHICLES DRIVERS

Vesko Lukovac*[1], Dragan Pamučar[1], Ljubomir Gigović[1]
[1] University of Defense, Military Academy
*Corresponding author, e-mail: lukovacvesko@yahoo.com

Abstract: *This paper deals with group decision making based on complete information. Known method for decision support, Analytical Hierarchy Process (AHP), was used individually by twenty participants in the task: selecting evaluators for evaluation of performance of military motor vehicles drivers. The deciding group consisted of twenty officers of the traffic department of the Army of Serbia (VS). Synthesis of individual evaluations was carried out afterwards using geometric averaging method. Additional analysis of the individual performance of decision makers was performed using the conformity index and Spirman's rank correlation coefficient.*

Keywords: AHP, group decision-making, evaluators, drivers, military vehicles

1. INTRODUCTION

The final score of evaluating the performance of driver of military motor vehicles (m/v) in VS according to the current approach is the result of the evaluation of a person **–** his direct superior. However, from the standpoint of successful human resource management, one of the measures recommended for the reduction of errors of evaluators in the process of evaluating the performance of employees is the inclusion of several evaluators in the evaluation process! Accordingly, there is the question: To whom the role of evaluators in the process of evaluating the performance of drivers of military m/v would be given? For the selection of persons who may be given the role of evaluators in the process of evaluating the performance of employees there are several options at disposal. Evaluator, in principle, can be any person who knows and follows the work process. Five primary sources that are recommended for the selection of persons who may be involved in the process of evaluation of the performance of employees (Noe, Hollenbeck, Gerhart & Wright, 2006):

1) managers, and direct supervisors;
2) co-workers;
3) subordinates;
4) employees (self-evaluation);
5) customers-consumers;

Starting from one of the five primary sources, which are recommended for the selection of evaluators in the process of evaluating the performance of employees, the final score of evaluating the performance of drivers of military m/v may be the result of the evaluation by direct superior (superior officer), by those who are on the same organizational level (co-worker), by personal evaluation (self-evaluation), as well as by clients (service user). Performance evaluation of drivers of military m/v can not only be the result of the evaluation by a direct subordinate, considering that drivers of military m/v represent the lowest - executive level in the organizational structure of the transportation service of VS. On the other hand again, based on the assumptions evaluation of 360 degrees, this process could include dispatchers. Therefore, pursuant to the recommendation for inclusion of more evaluators in the process of evaluating the performance of employees, the final score of evaluating the performance of drivers of military m/v may be the result of evaluation: direct his supervisor, dispatchers, colleagues, service users and their own evaluation.

However, the question of whether final score of evaluating the performance of drivers of military m/v should be the result of these evaluations, and which criteria should be taken into account when selecting persons as evaluators who can get involved in the process of evaluating the performance of military driver m/v? In order to define the criteria that should be considered in the selection of these persons, a group of experts, which consisted of twenty officers of the traffic service, hereinafter EKSG, was interviewed. The survey was conducted in two rounds. In the first round of the survey a

questionnaire, by which they are supposed to provide proposal criteria for selection of evaluator, was sent to the experts. Analysis of the results obtained by examination of the questionnaires showed that proposals of EKSG members regarding the criteria for the selection of persons who can, as evaluators, be involved in process of evaluating the performance of drivers of military m/v mainly relate to:

- Knowing the work that driver carryout;
- Better insight into chores drivers perform;
- Objectivity (impartiality) during the evaluation process;

In accordance with these proposals three so-called "selection criteria" were defined:

- K1 - Knowing the work;
- K2 - Insight into the work;
- K3 - Objectivity of the evaluation;

To validate the proposed selection criteria, in the second round of the survey, the members of EKSG received questionnaire where they were supposed to give their opinion on the validity of the proposed criteria for the selection of evaluators. The results of the survey by this questionnaire fully confirmed (100%) consistency of members of EKSG that persons who participate in the role of evaluator in the process of evaluating the performance of drivers of military m/v should be selected on the based on these 3 criteria.

As the decision on the selection of persons who should be involved, as evaluator, in the process of evaluating the performance of drivers of military m/v, should be made taking into account the multiple (three) criteria, in this paper decision implemented methodology for multi-criteria analysis and ranking known as the AHP-Analytical Hierarchy process (Analytic Hierarchy process) is used and at group level.

2. METHODOLOGY AHP GROUP DECISION MAKING

Analytic Hierarchy Process (Saaty, 1980) is the worlds most used multi-criteria analysis method for assistance in individual and group decision making (Subramanian & Ramanathan, 2012; Bernasconi, Choirat & Sherry, 2014). The method is "analytical" and "hierarchical "because the decision-maker breaks down a complex decision making problem into the several elements establishes a hierarchy between them. Popularity AHP method in the world originates primarily from its fundamental quality: "Always compare two elements of decision-making in pairs and use a simple semantic scale with only five basic graduation degrees of importance (equal, little more important, more important, much more important, absolutely important)". In this way, the method is approximated to person who basically does not have to have an interest in "mathematics" on which the AHP scientifically based on and recognized.

Example below applies Saaty’s group decision making principles in AHP method:

- Number of group members is $K \geq 2$;
- The problem hierarchy is pre-defined with the global aim of the top level, the criteria on the lower level and a set of alternatives at the bottom level;
- Standard Saaty's scale given in Table 1 is used;
- Individually obtained weight values of alternatives are geometrically synthesized into the group value according to the consistency demonstrated by the group members in the individual AHP evaluation hierarchy. Consistency is calculated according to the Saaty’s methodology, and a weight of the group members is determined by the normalization of individual reciprocal coefficients of consistency.

Table1: Saaty's scale of relative importance

Definition	**Numerical equivalent (a_{ij})**
The absolute dominance of the element i over the element j	9
Very strong dominance of the element i over the element j	7
Strong dominance of the element i over the element j	5
Poor dominance of the element i over the element j	3
The same importance of the element i and the element j	1
Poor dominance of the element j over the element i	1/3
Strong dominance of the element j over of the element i	1/5
Very strong dominance of the element j over the element i	1/7
The absolute dominance of the element j over the element i	1/9
(Intermediate values)	(2,4,6,8)

Since all decision makers (DM) made their evaluation, information base is complete, and a group synthesis was performed by consolidating individual priorities (Aggregation of Individual Priorities - AIP) using geometric averaging method (Geometric Mean Method - GMM):

(1)

where:

- weighted value of alternative to A_i awarded by k[th] member of the group (k = 1,2, ..., m)
- overall composite weight of alternative Ai
- t - weight of k[th]

Single weigh of $\alpha_k -$ t group members needed to be previously additively normalized, where supplementary analysis of the performance of individual decision makers, was conducted by using conformity index and Spirman's rank correlation coefficient. When calculating conformity index (C^k) group weight of alternatives obtained by synthesis from individual weights, is adopted as a reference, and then conformity index is calculated for each group member as follows:

... (2)

In the equation (2) k is index group member, K is the number of group members, a superscript ref represents a group weight.

When calculating Spirman's rank correlation coefficient (S) so-called reference list of ranks is used, and any other list is compared with it according to the equation:

———— ... (3)

is the difference between U_a i V_a and where U_a i V_a are ranges for an alternative a according to the reference list and according to the list that compares to the reference list, and n is the number of alternatives. In a group context, the equation (3) is applied to each of the combinations (group list, a list of the k[th] group member) i.e. one need to calculate as many Spirman's coefficients as there are members of the group, same as with conformity index. Value of Spirman's coefficient can vary between theoretical values -1 and 1. When the value is closer to 1, the indication is that the ranks are similar or the same, and when the value is less than zero and closing - 1 ranks are reversed, i.e. negatively correlated.

3. PROBLEM SETTING AND THE INDIVIDUAL AHP EVALUATION

Problem decided by EKSG consists of three hierarchical levels. Five persons who participate in the role of evaluators in the process of evaluating the performance of drivers of military m/v were assessed according to 3 criteria, Figure 1.

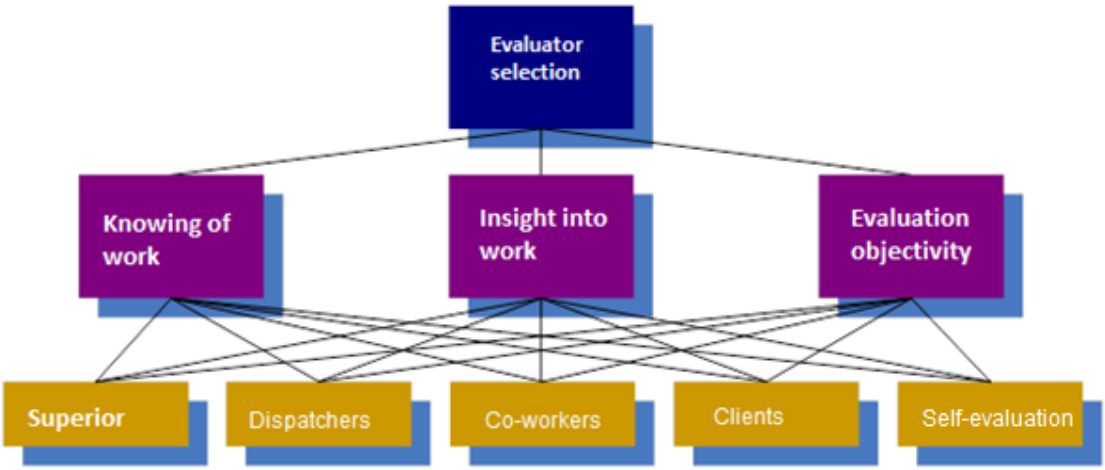

Figure 1: The hierarchy of decision problems

The aim is to "select the evaluator," and it is on top of the hierarchy. Selection criteria are the following - the middle level, and alternative (persons who participate in the role of evaluators), which is subject to selection are at the lowest level of the hierarchy:

- A1 - superior
- A2 - Dispatchers;
- A3 - Colleagues;
- A4 - Clients;
- A5 - Self-evaluation;

With each DM a complete personalized AHP stage consisting of the following steps (Lukovac, 2010) was conducted:

1. The hierarchy in Figure 1 was shown in a short interview and the elements of the hierarchy were explained;
2. The procedure of AHP evaluating was explained i.e. instructions on how to evaluate elements of the hierarchy using Saaty 's Scale in Table 1, were given;
3. The form with comparison matrix for AHP comparison of hierarchy elements in pairs using Saaty 's Scale was delivered;
4. Participants were informed that the results from the comparison matrixes would be processed by the software Expert Choice 2000 (EC 2000), which automatically calculates reciprocal value, and consequently in the comparison matrixes only elements in so-called upper triangles of the comparison matrixes should be evaluated;
5. It was also announced that after the completion of the AHP evaluation, DM would be informed on the final result of decision he made on the problem and about the degree of consistency that he demonstrated, which was also automatically calculated by the software EC 2000.

Final results of the individual evaluations of decision makers are shown in Table 2. Results are obtained by processing data from the comparing matrix by the EC in 2000. In the far right column overall degree of consistency (CR) is given, by which, in an indirect way, according to established methodology AHP (Saaty, 1980), the consistency of decision-making in respect of transitive rules such as "If A has strong dominance over B (numerical preference equal to 5; see Table 1), and B is of

equal importance as the C (numerical preference is 1; see Table 1) then A has strong dominance over C (numerical preference equal to 5)", was presented. Mathematically, complete consistency is achieved if a_{AB}=5 and a_{BC}=1 than a_{AC}=$a_{AB}a_{BC}$=5x1=5. If this rule is violated in any way, it results inconsistency. If anywhere in AHP within the entire hierarchy the smallest inconsistency occurs, the level of CR is greater than zero. According to (Saaty, 1980), it is considered that for the CR less than 0.10 consistency is satisfactory.

Table 2: Individual evaluation of DM

EKSG	Alternatives					CR
	0.351	0.351	0.165	0.066	0.068	0.02
	0.352	0.334	0.176	0.052	0.085	0.05
	0.433	0.257	0.171	0.055	0.083	0.05
	0.39	0.287	0.172	0.068	0.082	0.03
	0.424	0.253	0.181	0.045	0.098	0.03
	0.341	0.313	0.193	0.054	0.099	0.04
	0.336	0.378	0.15	0.053	0.083	0.05
	0.307	0.307	0.233	0.064	0.089	0.03
	0.303	0.303	0.203	0.067	0.127	0.02
	0.312	0.268	0.217	0.067	0.135	0.02
	0.33	0.33	0.17	0.068	0.101	0.02
	0.34	0.307	0.181	0.073	0.099	0.02
	0.286	0.286	0.224	0.076	0.128	0.01
	0.301	0.301	0.198	0.076	0.124	0.01
	0.317	0.317	0.213	0.066	0.088	0.02
	0.329	0.292	0.218	0.054	0.107	0.02
	0.319	0.319	0.211	0.067	0.084	0.01
	0.36	0.316	0.168	0.071	0.085	0.02
	0.405	0.284	0.16	0.057	0.094	0.02
	0.388	0.301	0.177	0.054	0.079	0.02

Ranges of individual evaluation of alternatives by the members EKSG are shown in Table 3.

Table 3: Ranges of individual evaluation of alternatives

EKSG	Alternatives				
	1-2	1-2	3	5	4
	1	2	3	5	4
	1	2	3	5	4
	1	2	3	5	4
	1	2	3	5	4
	1	2	3	5	4
	2	1	3	5	4
	1-2	1-2	3	5	4
	1-2	1-2	3	5	4
	1	2	3	5	4
	1-2	1-2	3	5	4
	1	2	3	5	4
	1-2	1-2	3	5	4
	1-2	1-2	3	5	4
	1-2	1-2	3	5	4
	1	2	3	5	4
	1-2	1-2	3	5	4
	1	2	3	5	4
	1	2	3	5	4
	1	2	3	5	4

4. SYNTHESIS OF INDIVIDUAL AHP EVALUATION

Group synthesis of individual weight vectors listed in Table 2 was performed AIP GMM method according to the form (1). DM weights were obtained by normalizing the reciprocals of the individual degrees of consistency (CR) (Srđević, 2008), and are shown in table 4.

Table 4: Weights decision makers

EKSG	CR	Weights (α_k)
	0.02	0.0508
	0.05	0.0203
	0.05	0.0203
	0.03	0.0338
	0.03	0.0338
	0.04	0.0254
	0.05	0.0203
	0.03	0.0338
	0.02	0.0508
	0.02	0.0508
	0.02	0.0508
	0.02	0.0508
	0.01	0.1015
	0.01	0.1015
	0.02	0.0508
	0.02	0.0508
	0.01	0.1015
	0.02	0.0508
	0.02	0.0508
	0.02	0.0508

When the weights listed in Table 4 are used for the group synthesis of individual vectors DM from Table 2, by using AIP GMM method according to the equation (1), a group decision is obtained as shown in Table 5.

Table 5: Weights alternatives and ranks according to the AIP GMM

AIP GMM	Alternatives				
Weight	0.336	0.305	0.194	0.065	0.100
Rank	1	2	3	5	4

In order to analyze the individual performance of DM, values of the conformity index (C^k) and Spirman's rank correlation coefficient (S) were calculated. Conformity index was calculated according to equation (2) and Spirman's rank correlation coefficient by the software IBM SPSS Statistics 22. The values of conformity index and Spirman's DM correlation coefficient are shown in Table 6.

The values of conformity index of (0.026-0.195) show a small deviation of individual decisions from group decisions. The lowest level of conformity (0.026) was at which means that his deviation of weight values of alternative from their group weight value is slightest. , DM_{16}, DM_{15} follow and so on up to with the highest conformity index (0.195). Based on the values Spirman's rank correlation coefficient (0.900-1.000), it can be concluded that the individual decisions are in correlation with the group decision. Even eleven DM has equally ranked all alternatives (comparing to a group decision) ($S = 1.000$).

Table 6: Conformity index and Spirman's rank correlation coefficient

EKSG		
	0.123	0.975
	0.091	1.000
	0.195	1.000
	0.115	1.000
	0.175	1.000
	0.026	1.000
	0.146	0.900
	0.082	0.975
	0.073	0.975
	0.121	1.000
	0.059	0.975
	0.028	1.000
	0.138	0.975
	0.078	0.975
	0.063	0.975
	0.062	1.000
	0.066	0.975
	0.082	1.000
	0.138	1.000
	0.105	1.000

5. CONCLUSION

Each possible choice of those who should be assigned the role of evaluator in the process of evaluating the performance of employees has its advantages and disadvantages with regard to evaluation quality. For whom would the organization decide, depends on a number of factors: organizational design, management style, goals set and information that are desired. This paper shows how known method of multi-criteria analysis- AHP could be used for this purpose. The method is interesting because in addition to individual analysis it allows easy expansion to the group context.

This paper considers situation where the problem of selection of persons who should be, as evaluators, involved in the process of evaluating the performance of drivers of military m/v, was decided by 20 individuals. Identified as $\mathrm{DM}_1 - \mathrm{DM}_{20}$ decision makers individually used AHP, and then the context of decision-making was extended to the group. Since the group decision making allows numerous variants of handling the process, one of the most important is that the individual DM was assigned relative weight and thus possibly favoring one over the others in forming a group decision. It should be noted the difference between the terms "common" and "group" decision. In the first case consensus is implied, but it is not necessarily in another. Group context that is considered in this paper entirely corresponds to the second case, i.e. there was no harmonization or consultations among the participants, and the results of individual evaluations were summarized later. In order to maximally objectify group context in the synthesis of individual decision, decision makers were assigned weights that were corresponded with manifested levels of the individual consistency.

The results of the process of aggregation of the individual to the group decision and measurement in each case of conformity and rank – distance according to Spirman in relation to the derived group decision, shows that AHP methods produce logical results in a multi-criteria tasks directly expressed or hidden conflicts. Additional analysis would allow, for example, drawing of conclusions about the effects of classification of DM into subgroups according to the consistency and then the subsequent synthesis sub-group decisions in a group decision, or the elimination of some DM (DMs) from the decision-making process because of the reported inconsistencies, etc.

REFERENCES

Bernasconi, M., Choirat, C., & Seri, R. (2014). *Empirical properties of group preference aggregation methods employed in AHP: Theory and evidence.* European Journal of Operational Research 232 (3), 584-592.

Lukovac, V. (2010). *A model for evaluating the quality of military drivers,* thesis, Faculty of Traffic Engineering, Belgrade.

Noe, AR, Hollenbeck, RJ, Gerhart, B., & Wright MP (2006). *Human Resource Management,* Third Edition, MATE doo, Zagreb, 2006.

Saaty, TL (1980). *The Analytic Hierarchy Process,* McGraw-Hill, New York, in 1980.

Srđević, B., Suvočarev, B., & Srđević, Z. (2008). *AHP group decision without consensus: Example planning segmentation wet fields.* Waterpower engineering 40 (231-233), 51-58.

Subramanian, N., & Ramanathan, R. (2012). *A review of applications of Analytic Hierarchy Process and operations management.* International Journal of Production Economics 138 (2), 215-241.

MULTICRITERIA DECISION MAKING - PRACTICAL EXAMPLE OF AHP BASED APPROACH FOR SUPPLIER SELECTION

Jovana Mihailovic*[1]
[1]Vip mobile d.o.o.
*Corresponding author, e-mail: j.mihailovic@vipmobile.rs

Abstract: *Paper presents multi-criteria decision making problem of supplier selection. When evaluating and choosing the best supplier, decision maker faces complex task that usually consists of analysing many qualitative and quantitative conflicting performance measures. On practical example applied in one company from Belgrade, this paper describes how decision maker can chose best supplier by using Analytical Hierarchical Process and software Expert Choice. This study identifies criteria used for decision making, alternatives, and presents Expert Choice calculation and sensitivity analysis of solution.*

Keywords*: multi-criteria decision making, AHP method, supplier selection, Expert Choice*

1. INTRODUCTION

The goal of multi-criteria decision making (MCDM) techniques is to help decision maker choose the best option among set of alternatives, by analysing them from different angles or attributes. Ideal situation would be if one had alternative best by all criteria. However, almost every real life example consists of intangible and conflicting criteria and evaluating alternatives is rather complex task.

Since 1970s, MCDM has undergone rapid development and has become a hot research topic. Today there are more than hundred different MCDM techniques and their variations, and this lead to conclusion that choosing right MCDM method has become a decision problem itself (Saaty & Ergu, 2015). Some of the most popular MCDM techniques are the Analytic Hierarchy Process (AHP), the Analytic Network Process (ANP), Technique for Order Preference by Similarity to ideal Situation (TOPSIS), ELimination Et Choix Traduisant la REalité (ELECTRE), Preference Ranking Organization METHod for Enrichment Evaluations (PROMETHEE), VIKOR method, Simple Multiattribute Rating Technique (SMART) and Decision-Making Trial and Evaluation Laboratory (DEMATEL).

Different authors have presented comparisons of different decision making techniques in their researches (Velasquez & Hester, 2013, Chai, Liu & Ngai, 2013, Hodgett. 2015). It is hard to estimate which MCDM technique is the best since many methods can lead to different results when dealing with identical problem. It is difficult to classify, evaluate, and compare different MCDM methods because they are based on a variety of assumptions about the decision maker's preferences and use different types of preference information (Saaty & Ergu, 2015).

AHP appeared to be one of the most prominent methodologies used to address the problem of multi-criteria decision making (Bruno et al, 2012, Chai Liu & Ngai, 2013, Hodgett. 2015, Bentes et al, 2012). Literature review has showed wide use and numerous applications of AHP method. AHP was used to solve problems in: business planning, resource allocation, organisation priority setting, and selection (Bhushan & Rai, 2004), total quality management, business process reengineering, quality function deployment. The Ford Motor Company used the AHP (1999) to establish priorities for criteria that improve customer satisfaction. IBM used AHP in the design and marketing of its successful midrange AS 400 computer (1991). AHP had significant role in creating agreement called the Pittsburgh Principles in Israeli-Palestinian conflict (2011). Some of the reasons AHP method has wide use are its simplicity (decision maker compare only two objects at the time), it is adaptable (this is possible because of hierarchical structure) and it is supported by user friendly software package commercially available. Additionally, it can be used in combination with many other approaches: Data Envelopment Analysis- DEA (Liu & Hai, 2005), fuzzy set theory (was investigated by Chan & Kumar, 2007), genetic algorithms (Rao, 2007), Balanced Scorecard - BSC (nice example of applying AHP and BSC in measuring organisational performance was presented in Bentes et al, 2012).

1.1. Supplier selection- multi-criteria decision making problem

Supplier selection is process that includes acquisition of required material, services and equipment for all types of businesses. Evaluating vendors is complex procedure that requires analysis of multiple tangible and intangible criteria and many alternatives. Almost every company deals with supplier selection problem and criteria used for evaluation are usually based on cost, quality and vendor ability to meet the deadline and keep up with schedule. MCDM techniques can be used to support decision makers in supplier evaluation and their final judgment of choosing the best alternative. AHP method was evaluated as the most used methodology in supplied selection problem (Bruno et al, 2012, Chai Liu & Ngai, 2013). This can be explained by observing some of the strengths of AHP: it has ability to structure problem hierarchically which facilitates decision maker analysis since he investigates each level of hierarchy separately, AHP can handle both tangible and intangible attributes and characteristics, it is possible to monitor consistency decision maker has when making judgement and it can be used in combination with many other approaches.

This paper should highlight the need to use MCDM techniques in corporate environment. The goal is to solve supplier selection problem by using AHP method and software package Expert Choice. This is shown on practical example applied in one company in Belgrade. Paper is organised in four sections. Following introduction, section 2 describes theoretical framework of AHP method. Section 3 presents application of AHP on supplier selection problem: criteria for supplier evaluation and alternatives are defined and structured hierarchically (3), attributes and alternatives are evaluated in 3.1 and 3.2, results, sensitivity analysis and discussion is shown in 3.3.

2. ANALYTICAL HIERARHICAL PROCESS

AHP is decision making approach that helps decision maker solve complex problem consisting of multiple conflicting criteria both quantitative and qualitative and multiple alternatives. Modeling process consists of following steps:

- Problem structuring
- Comparative judgements and weights evaluation
- Calculation of global score
- Sensitivity analysis

Like in all MCDM processes, first challenge is to structure the problem, which means to define goal, criteria and alternatives. One of the advantages of AHP method is the fact that it structures attributes in hierarchy. This provides decision maker better focus on specific criteria and sub criteria and simplifies allocating the weights.

Second phase begins with data collection and measurement. Decision maker does pairwise comparisons at each node of hierarchy. Psychologist argue that it is easier and more accurate to express one's opinion on only two alternatives than simultaneously on all alternatives (Ishizaka & Labib, 2009). To compare two attributes and characteristics *i* and *j* the ratio scale illustrated in Table 1 is used.

Table 1: Ratio scale of comparative judgements (Saaty, 1977)

Intensity of importance	Definition	Explanation
1	Equal importance	*i* and *j* are equally important
3	Weak importance of one over another	*i* is slightly more important than *j*
5	Strong importance	*i* is much more important than *j*
7	Very strong importance	*i* is by far much more important than *j*
9	Absolute or extreme importance	*i* is definitely much more important than *j*
2, 4, 6, 8	Intermediate values between 2 adjacent judgements	
Reciprocals of above nonzero	If activity *i* has one of the above nonzero numbers assigned to it when compared to activity *j,* then *j* has the reciprocal value when compared to *i*	

Next step is determining relative weights of attributes. Assume *n* attributes $A_1, A_2 \dots A_n$ with known worth $w_1, w_2 \dots w_n$. Matrix of ratios is shown on Table 2. To recover the scale *w* from the matrix of ratios *A*, one must solve the eigenvalue problem shown in Equation 1:

$$Aw = nw \text{ or } (A - nI)w = 0 \tag{1}$$

Table 2: Matrix of ratios

$$\begin{array}{c} \\ A_1 \\ A_2 \\ \\ A_n \end{array} \begin{array}{c} \begin{array}{cccc} A_1 & A_2 & & A_n \end{array} \\ \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & & \ddots & \vdots \\ \vdots & & & \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \end{array}$$

Table 3: Equation to get back the scale

$$\begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & & \ddots & \vdots \\ \vdots & & & \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

In general case when only judgements but not the numbers are available, precise value of $\frac{w_i}{w_j}$ is not known, instead only an estimate of it can be given as a numerical judgement from the scale in Table 1 (Saaty, 2013). The maximum eigenvalue is no longer equal to *n* but is replaced by the maximum eigen value of matrix of judgements, λ_{max} (Table 4 and Table 5). The solution is obtained from the Equation 2:

$$A' w' = \lambda_{max}\, w' \tag{2}$$

Table 4: Matrix of ratios

$$A' = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{12}} & & & \\ \vdots & & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1 \end{bmatrix}$$

Table 5: Equation to get the scale of priorities

$$\begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{12}} & & & \\ \vdots & & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \lambda_{max} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

The consistency index (CI) and the consistency ratio (CR) measure the degree to which judgements of decision maker are not coherent. Saaty (1977) has proposed a CI which is related to eigenvalue method:

$$CI = \frac{\lambda_{max} - n}{n-1} \tag{3}$$

$$CR = \frac{CI}{RI} \tag{4}$$

RI stands for random index (average CI of 500 randomly filled matrices (Saaty, 1977)). Satisfactory degree of consistency is when CR is less than 10%.

Once local priorities are calculated they should be synthetized across all criteria to determine the global priority. This is done by additive aggregation.

The last step in this model is the sensitivity analysis. The aim of it is to observe the results when input data are slightly modified. If the ranking does not change, the results are said to be robust.

3. SUPPLIER SELECTION – PRACTICAL USE OF AHP METHOD

Practical use of AHP method is demonstrated on example in company MBT Concept. Best supplier was chosen with the help of software Expert Choice. Problem solving steps follow the steps of AHP method: defining goal, defining criteria and sub-criteria, defining alternatives, pairwise comparison of attributes, and pairwise comparison of alternatives according to each criteria, aggregation of local results and sensitivity analysis.

Company is the consultant for procurement of FF&E (furniture, fixtures and equipment) and at the time of study company had big project going on which included choosing vendors for hotel furniture. Choosing best supplier is challenging task that requires observation and evaluation of many qualitative and quantitative criteria, so commercial management team was proposed to use AHP method in order to help them in selection process. Several meetings were organised with commercial management team. In the beginning they were familiarised with AHP method, how it can be applied and all the parts they have to participate in. First challenge was to define criteria and arrange them in hierarchical structure. Some authors have proposed set of attributes that can be chosen for supplier selection (Sung & Ramayya, 2008). This paper was used as starting point for identifying criteria, however the final set of attributes was defined according to the needs of the company. Most important criteria this company values when choosing supplier were recognized together with commercial director and procurement manager. In this study agreement based approach was used, since only two representatives from the company were participating in analysis. Constructive dialogue conducted during the meetings lead to common understanding and decisions. Criteria used for evaluation were divided in three groups: Service level, Quality, Financial conditions. Indicators for Service level were: delivery time (time supplier needs to produce, transport and install furniture), job knowledge (how well informed and professional are vendor representatives), flexibility (how fast can supplier respond to changes in specifications), previous experience (if company cooperated with vendor in the past how satisfied they were). Quality child attributes were: design (similarity between offered and designed furniture), warranties (duration of guarantee period), quality of material (quality of fabrics, woods and other material used in construction of furniture), references (big projects company worked on, especially hotels). Financial conditions were divided into: product price (fee of offered furniture), advance payment (the required amount of advanced payment), other conditions (at which point supplier required full payment). Hierarchical structure of supplier selection problem is presented on figure 1.

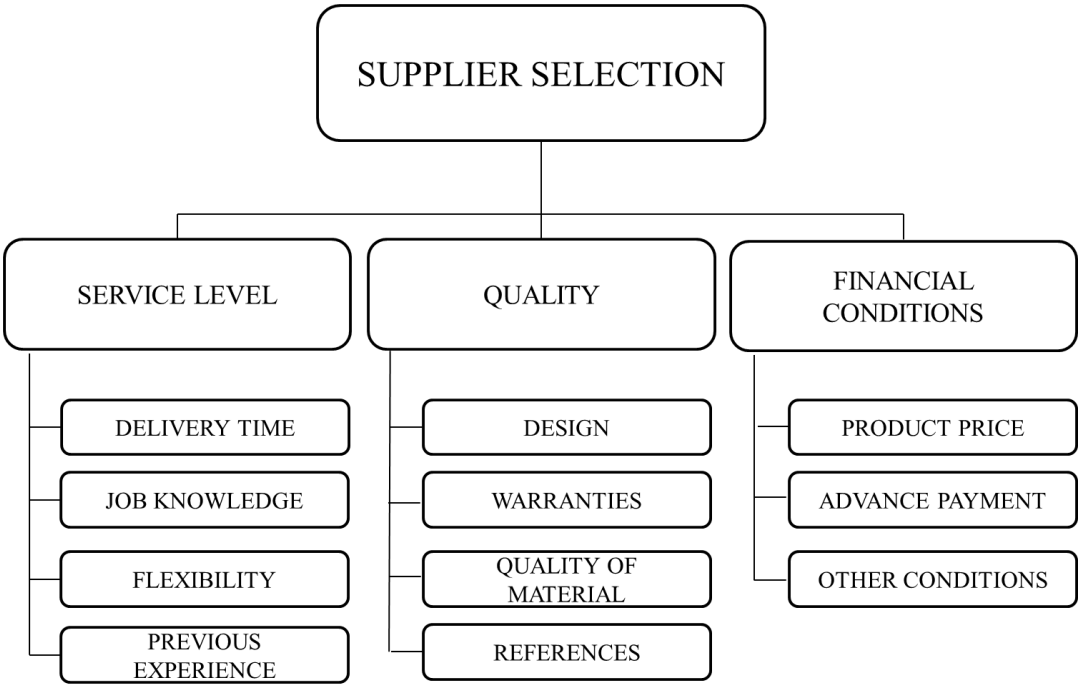

Figure 1: Hierarchical structure of supplier selection problem

In this case study three suppliers were evaluated. Due to data privacy vendors are defined as: Supplier A, Supplier B, Supplier C. Authors are familiar with companies names as well as with their proposals.

3.1. Supplier selection criteria evaluation

After attributes are defined and placed in hierarchical structure pairwise attribute comparison is performed on each level of hierarchy. Ratio scale from table 1 was used as shown on figure 2

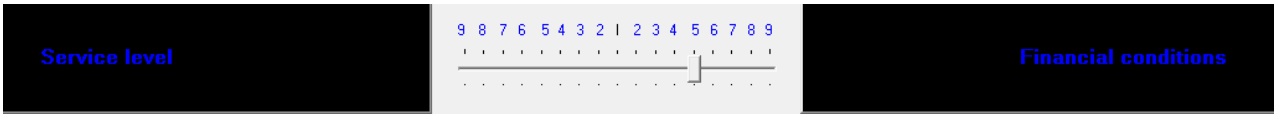

Figure 2: Numerical scale in Expert Choice

Decision makers were advised to use scales: 1, 3, 5, 7, 9, and in the case they were hesitating between two to pick number in the middle. Comparison values and derived local and global priorities are shown on table 6. Consistency ratio was checked in each comparison matrix, and its level was always less or equal to 0.1.

Table 6: Criteria evaluation

SUPPLIER SELECTION	Service level	Quality	Financial conditions	Priority
Service level	1	1/4	1/5	0.093
Quality	4	1	1/3	0.280
Fin. conditions	5	3	1	0.627

IC=0.08

SERVICE LEVEL	Delivery time	Job knowledge	Flexibility	Previous experience	Local priority	Global priority
Delivery time	1	4	5	1/2	0.354	0.033
Job knowledge	1/4	1	1/2	1/3	0.091	0.008
Flexibility	1/5	2	1	1/4	0.112	0.011
Prev.experience	2	3	4	1	0.443	0.041

IC=0.09

QUALITY	Design	Warranties	Quality of material	References	Local priority	Global priority
Design	1	1	1	1	0.250	0.070
Warranties	1	1	1	1	0.250	0.070
Quality of m.	1	1	1	1	0.250	0.070
References	1	1	1	1	0.250	0.070

IC=0

FINANCIAL CONDITIONS	Product price	Advance payment	Other conditions	Local priority	Global priority
Product price	1	8	7	0.777	0.487
Advance pay.	1/8	1	1/3	0.070	0.044
Other condition	1/7	3	1	0.153	0.096

IC=0.1

One can notice that financial conditions are by far most important evaluation criteria, influencing the final decision by 62.7%. As expected most important second level attribute is product price and it influences 48.7% total ranking.

3.2. Supplier evaluation

Supplier evaluation is performed after pairwise comparison of criteria is finished and their local and global priorities are calculated. Comparison is done by each criteria (Figure 3). Results are shown in table 7.

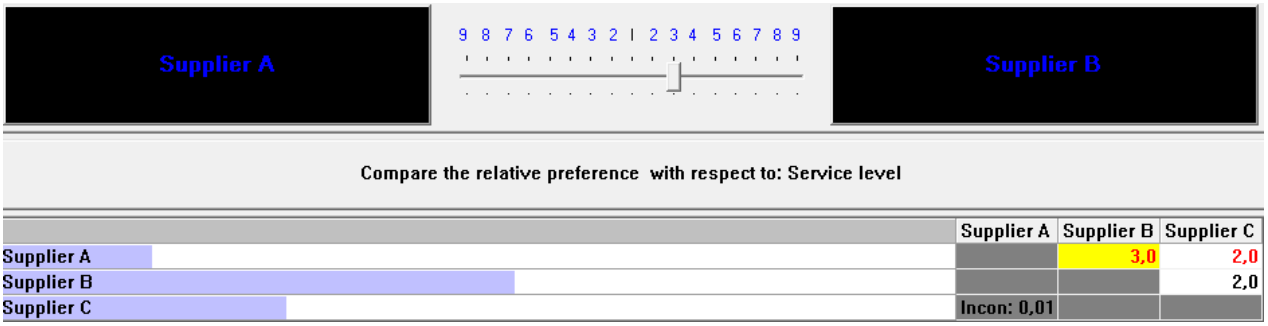

Figure 3: Supplier comparison in Expert Choice

Table 7: Supplier evaluation

Deliv.t.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1/3	1/2	0.163
Sup. B	3	1	2	0.540
Sup. C	2	1/2	1	0.297

IC=0.1

Flex.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1/2	2	0.297
Sup. B	2	1	3	0.540
Sup. C	1/2	1/3	1	0.163

IC=0.08

Know.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1	3	0.429
Sup. B	1	1	3	0.429
Sup. C	1/3	1/3	1	0.143

IC=0

P. exp.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1	1	0.333
Sup. B	1	1	1	0.333
Sup. C	1	1	1	0.333

IC=0

Design	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	2	5	0.570
Sup. B	1/2	1	4	0.333
Sup. C	1/5	1/4	1	0.097

IC=0.02

Warran.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1	1	0.333
Sup. B	1	1	1	0.333
Sup. C	1	1	1	0.333

IC=0

Q. mat.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	2	4	0.558
Sup. B	1/2	1	3	0.320
Sup. C	1/4	1/3	1	0.122

IC=0.02

Ref.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	2	5	0.582
Sup. B	1/2	1	3	0.309
Sup. C	1/5	1/3	1	0.109

IC=0

Price	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1/3	1/4	0.143
Sup. B	3	1	1/2	0.286
Sup. C	4	2	1	0.571

IC=0.02

Adv. p.	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1	1	0.333
Sup. B	1	1	1	0.333
Sup. C	1	1	1	0.333

IC=0

Other c	Sup. A	Sup. B	Sup. C	Priority
Sup. A	1	1/4	1/2	0.143
Sup. B	4	1	2	0.571
Sup. C	2	1/2	1	0.286

IC=0

3.3. Results and discussion

In previous chapter one can observe supplier order by each examined attribute. However it is very hard to tell which supplier is the best in total. Overall best supplier, or final result of evaluation is presented on Figure 4.

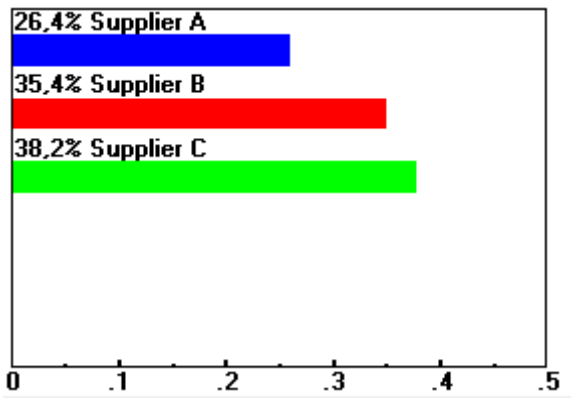

Figure 4: Supplier evaluation results

Supplier with the best total grade is Supplier C with 0.382, second is Supplier B with 0.354 and last is Supplier A with 0.264. Figure 5 demonstrates supplier ranking in more detail and one can observe supplier order by service level, quality and financial conditions. Supplier C is by far better than any other supplier

regarding financial conditions. On the other hand regarding quality and service level its results are rather poor. As previously mentioned, financial conditions were identified as most influential criteria (62.7%) and only thanks to good price offer they defeated their competition. Supplier B on the contrary was evaluated as best regarding service level, and second regarding quality and financial conditions. The difference between Supplier C and B is only 0.028 or 2.8%.

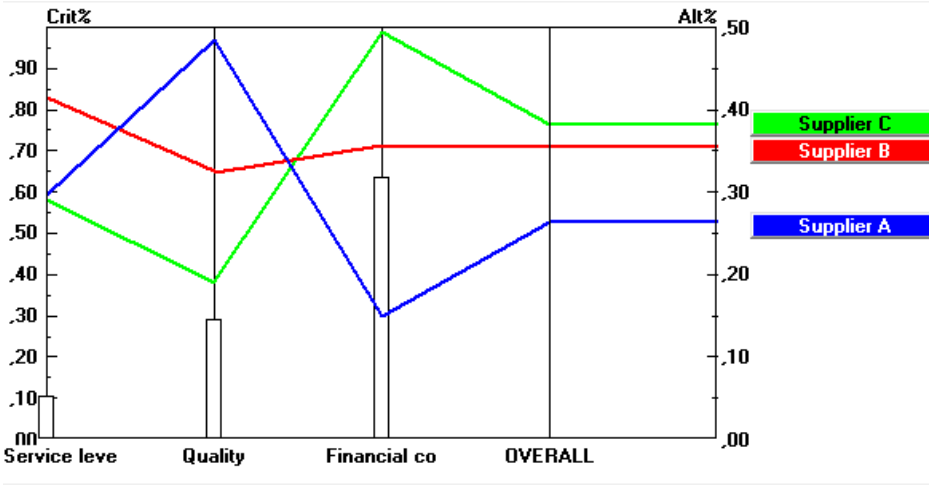

Figure 5: Supplier order by each group of criteria

It is always recommended to perform sensitivity analysis and check if overall ranking changes when the input data changes. Supplier B is better than C by both service level and quality. Expert Choice is very convenient for sensitivity analysis. One can generate Gradient sensitivity graph and observe how supplier ranking can change when priority of criteria changes. Vertical line represents current state and by moving it, variation in alternative order can be observed. It is logical to check result in the case financial priority decreases since financial conditions have highest impact. If the impact of financial conditions drops by 0.1, or by 10%, ranking changes and supplier B is ranked as first (Figure 6). When one criteria ponder drops, some others must be increased, in this case impact of quality and service level increase equally, 5% each.

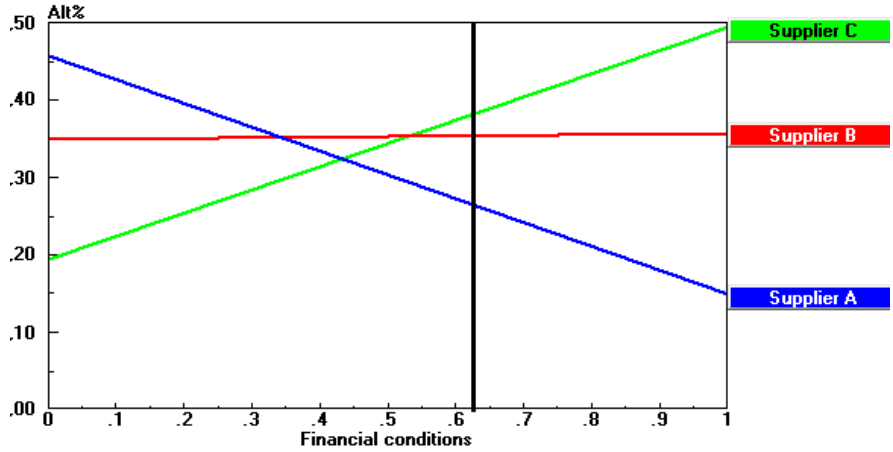

Figure 6: Supplier ranking according to financial conditions ponder

Previous analysis have showed that even though Supplier C was ranked first in overall result, Supplier B seemed to have better performance by more criteria and was serious candidate for winning the job especially after sensitivity analysis. Before showing results, Commercial director was asked to present his feeling about which vendor should be chosen just based on his experience and by analysing offers. He had chosen Supplier B. After seeing results he was more certain in his decision.

Expert Choice was proven to be very useful tool to support judgements when evaluating suppliers. It is very intuitive and easy to use. On the other hand some imperfections were noted. Hierarchical structure in many ways simplifies problem, however when forming it one must be very careful since different structure may lead to different result. Pairwise comparison even though is simple, can be tiring, because one must make a lot of judgements. Additionally it is not easy to choose numbers from Saaty scale. In order to be sure which grade

to grant one must precise criteria for evaluation for each attribute. If compared values are numerical, Expert Choice offers possibility to calculate priority based on their value. In the case of minimization (for example attribute price), priority is calculated based on reciprocal value. In supplier selection example, when Expert Choice was calculating with exact prices, supplier final order was different, Supplier B was first with 35.9%, Supplier C had 34.4%.

4. CONCLUSION

Supplier selection is complex task almost every company faces, and solving it requires analysis of many qualitative and quantitative criteria, usually conflicting, and many alternatives. Companies require finding solution in a quick and simple way. Literature review has showed big interest for supplier selection problem and AHP method appeared to be mostly used and investigated. Due to that fact, in this paper AHP method for supplier evaluation was suggested and applied on practical example in company MBT Concept. First, important criteria were identified and were organised in hierarchical structure. Criteria were divided in three groups: Service level, Quality and Financial conditions. Three suppliers and their offers were analysed and evaluated. Expert Choice proved to be very intuitive and simple to use and was convenient for sensitivity analysis. Final result proved that AHP method can be easily applied and be very helpful in corporate environment to support manages in their decisions. In the last years researchers are trying to combine AHP with other mathematical programming methods such as DEA, genetic algorithm, fuzzy sets, as well as with strategic framework such as SWOT and BSC analysis, so it is assumed that the number of applications of AHP method in different areas in corporate world will grow in the future.

REFERENCES

Bentes et al.. (2012). Multidimensional assessment of organizational performance Integrating BSC and AHP. *Journal of Business Research*, 65(12), 1790-1799.

Bhushan, N & Rai, K. (2004). *Strategic Decision Making: Applying the Analytic Hierarchy Process*. England: Springer-Verlag.

Bruno et al.. (2012). AHP-based approaches for supplier evaluation: Problems and perspectives. *Journal of Purchasing & Supply Management*, 18(3), 159–172.

Chai, J, Liu, J & Ngai, E. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications*, 40(10), 3872–3885.

Chan, F & Kumar, N. (2007). Global supplier development considering risk factors using fuzzy extended AHP-based approach. *Omega*, 35(4), 417–431.

Hodgett, R. (2015). Comparison of Multi-Criteria Decision-Making Methods for Equipment Selection. *The International Journal of Advanced Manufacturing Technology*, 1-13.

Ishizaka, A & Labib, A. (2009). Analytic Hierarchy Process and Expert Choice: Benefits and Limitations. *ORInsight*, 22(4), 201–220.

Liu, F & Hai, L. (2005). The voting analytic hierarchy process method for selecting supplier. *International Journal of Production Economics*, 97(3), 308–317.

Rao, V. (2007). Vendor selection in a supply chain using analytic hierarchy process and genetic algorithm methods. *International Journal of Services and Operations Management*, 3(3), 355-369.

Saaty, T & Ergu, D. (2015). When is a Decision-Making Method Trustworthy? Criteria for Evaluating Multi-Criteria Decision-Making Methods. *International Journal of Information Technology & Decision Making*, 14(6).

Saaty, T. (2013). The Modern Science of Multicriteria Decision Making and Its Practical Applications: The AHP/ANP Approach. *Operations research* , 61(5), 1101–1118.

Saaty, T. (1977). A Scaling Method for Priorities in Hierarchical Structures. *Journal of mathmematical pshychology*, 15, 234–281.

Sung, H & Ramayya, K. (2008). A hybrid approach to supplier selection for the maintenance of a competitive supply chain. *Expert Systems with Applications*, 34(2), 1303–1311.

Velasquez, M & Hester, T. (2013). An Analysis of Multi-Criteria Decision Making Methods. *International Journal of Operations Research*, 10(2), 56-66.

EFFICIENCY EVALUATION OF ABA LEAGUE TEAMS FOR THE SEASON 2012-13 THROUGH DMU CLUSTERING AND DEA

Maja Mitrović*[1]
[1]University of Belgrade, Faculty of Organizational Sciences, Serbia
* Corresponding author: mitrovicmaja4@gmail.com

Abstract: *In this work, Data Envelopment Analysis (DEA) approach measures the efficiency of basketball teams, which played the 2012–13 Adriatic (ABA) League regular season. To apply such methodology, 14 teams were considered as Decision Making Units (DMU), and the algorithm K-means Clustering is used to group teams into homogenous clusters. Then, DEA is applied separately to evaluate efficiency of each team in a specific cluster. The proposed model in the paper is basic BCC DEA model, oriented to outputs, which considered a single input and two outputs. The advantages of the work rest on the fact that presented methodology should help ensuring a greater efficiency of each team, and getting more precise prognoses of sports results.*

Keywords*: DEA methodology, K-Means clustering, efficiency, basketball, ABA-league*

1. INTRODUCTION

As one of the well-known sports in the United States (ReadingTree, 2016), basketball provides a great portfolio for a business platform, which contributes to the great result in finance (Labonté, 2016). Looking at National Basketball Association (NBA), or at the most famous one from Spain - ACB league (ENBR, 2016), sports statistic is of great importance for different groups of people, since it allows many types of analyses. From those matches, as well as from other less popular leagues, available data enhance the chance of better results in the next season if it"s used in a right way.

Using Data Envelopment Analyse (DEA) methodology becomes very popular in recent years. DEA is widely used to measure the efficiency of decision-making units (DMUs). It is possible to find many works which describe how the methodology has been used for performance evaluation in widespread areas in the last decades, from non-profit sector like education and hospitals evaluation (Jeremic *et al*. 2011, Jeremic *et al.* 2012, Savic *et al*. 2012) to profit sector like banks evaluation (Bulajic *et. al*, 2011; Savic *et al*. 2013; Radojicic *et. al*, 2015). Variety DEA studies are also done for different sports, such as football (Tiedemann *et. al*. 2011, Petrovic *et al*. 2015), baseball (Lewis, 2014), and tennis (Ramón & Sirvent, 2012; Ruiz & Pastor, 2013). Many authors used the benchmark methodology with the DEA to show improvements in productivity in sports (Anderson, 2004; Barros & Santos, 2014). For the sake of this work, the most important is efficiency assessment in sports, in particular to basketball.

De Mello (Angulo-Meza & Da Silva, 2009) used basketball as official Olympic discipline, to rank clusters by aggregation of the partial sports. Cooper *et al*. (2009) demonstrate how multiplier values are used to identify major characteristics of individual players. For one of the most popular competitions – NBA, DEA approach is used by Radovanović, Radojičić & Savić (2014). Furthermore, Moreno and Lozano (2014) used DEA to assess the efficiency of NBA teams for the regular season 2009-2010, with the results which show the differences between network DEA and conventional DEA approach. Aizemberg *et al*. (2014) present an analysis of the NBA team for period of five years, using a multi-objective approach, while Alves, et al. (2013) show the efficiency of teams just for one season and ranks them according to composed efficiency score of each DMU, using the inverted frontier method.

For the purpose of this work, the author will analyze a regional professional basketball league, commonly known as the **Adriatic League**. In the recent years, the basic countries which took part were: Serbia, Croatia, Montenegro, Bosnia and Herzegovina and Slovenia. The body that organizes the league is **Adriatic Basketball Association** and the league is also considered as a local version of the Europe-wide Euro league (Aba-league, 2016).

The organization of the league is next: in the regular part of the season, 14 teams are playing. Each team plays against every other team once at home and once away. After 26 played rounds, the best placed 4 teams (also known as „Final Four") are qualified to the playoffs.

The semi-finals are played in the best of three games format (1-1-1), where the first placed team after regular part of the season will face the fourth-placed team in the first semi-final pair and the second placed team after regular part of the season will face the third placed team. The first and the second placed team will have the home court advantage in the semi-finals.

Winners of both semi-final pairs will advance to the finals, which will be played in the best of five games format (2-2-1). There will be no games for the third place played. The finalist, that is the better placed after the regular season, will have the home court advantage (Aba-league, 2016).

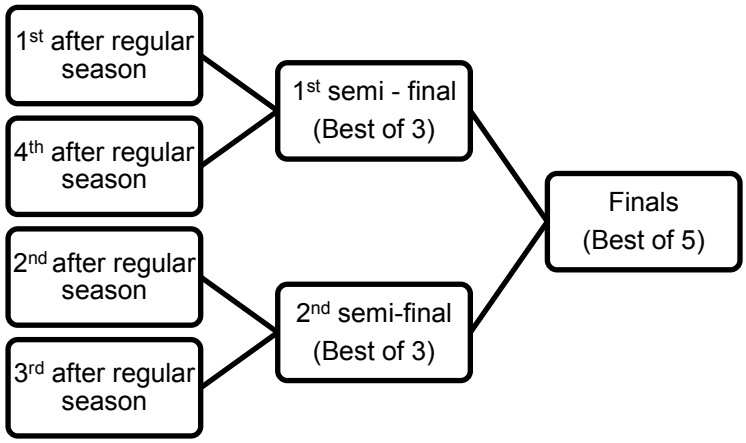

Figure 1: Competition system

This study proposes an efficiency analysis of 14 teams that played the 2012-13 regular season. Moreover, it will be shown clustering via K-Means Clustering algorithm and evaluation of the teams in each cluster using DEA methodology. The concept of clustering consists of grouping the DMUs with similar characteristics. Various authors have used different clustering methods in order to obtain homogeneous clusters (Alves *et al.,* 2013; Ma, Wang & Wang, 2014) and then, apply DEA for assessing efficiency index (Yu & Lee, 2013; Alves *et al.*, 2013; Moreno & Lozano, 2014). In order to improve the efficiency of each team, the author applied DEA methodology, observing the classic efficiency of teams in each cluster using the BCC DEA model (Banker *et al*., 1984) oriented to output. The goal is to grade DMUs. The obtained results can act as indicators for management, players, but also for civil society to realize what to improve, and to predict the future results in coming seasons in order to get a higher rank of selected team.

The paper is organized as follows. First, the basic concept of the data used is described. Then, in **Section 2** proposed methodology, including the K-means Clustering and fundamental concept of DEA model is given. In **Section 3** the result of the study is shown, as well as its analysis. Finally, concluding remarks of this work are presented.

2. METHODOLOGY

2.1. K-means Clustering

At the beginning of this paper, the author introduces that the matches took place in different Balkan cities, which can intensify the home advantage. Anderson *et al.* (2012) suggest that this factor could significantly affect the performance of teams in different sports. In order to set a homogeneous assessment, the DMUs were grouped in clusters based on circumstances of the cities that hosted their matches.

One of the most widely used algorithm for grouping units is K-means clustering. Tan, Steinbach and Kumar (2013) gave the definition: *"Cluster analysis groups data objects based only on information found in the data that describes the objects and their relationships"(p.490)*. They also argue that, if the similarity within a group is greater, as well as the difference between groups, the clustering is better *(p.490)*.

This work uses the K-means Clustering algorithm for the purpose of grouping teams in homogenous clusters. As a partitioning clustering technique, it uses centroids to specify the number of clusters (K). Centroids are usually the mean of a group of points (Tan, Steinbach, & Kumar, 2013) while K is the user specified number. Each point is assigned to the closest centroid, and the collection of points assigned to a centroid is a cluster. The basic algorithm is given in the following:

- Select K points as initial centroids

- Repeat
    - Form K clusters by assigning each point to its closest centroid
    - Recompute the centroid of each cluster
- Until centroids do not change

In the first iterations, the conversation is the fastest. A measure used to assign a point to the closest centroid is Euclidean (L_2) distance:

$$dist(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}\ , \tag{1}$$

where *p* and *q* are points and the distance represents the length of their line segment connecting them (Deza & Deza, 2009). In order to minimize the squared distance of each point to its centroid, it"s necessary to express objective function. The sum of squared error is a measure of the quality of clustering. The goal is next: to get the smallest total sum of squared errors of each data point:

$$SSE = \sum_{i=1}^{K}\sum_{x \in C_i} dist^2(m_i, x)\,, \tag{2}$$

where *x* is the point from the specific cluster $C_{i,}$ while m_i is the most representative point for the cluster *C* (Vazirgiannis, Halkidi, & Gunopulos, 2012). It"s also possible to use other measures as Proximity function Manhattan (L_1), where Centroid is the median of the points and also minimize the sum of the L_1 distance between an object and its median. Squared Euclidean use means as centroids and minimizes the sum of the squared L_2 distance (Deza & Deza, 2009).

The most followed issues for the previously described measure are Handling Empty Clusters and Outliers (Tan, Steinbach, & Kumar, 2013). Empty clusters can be obtained if no points are allocated to a cluster. Improving the methods allows overcoming the issues (Delibasic *et. al*, 2009; Delibasic *et. al*, 2012). Wang *et. al* (2015) described a new K-Means algorithm to reduce the computational complexity in the assignment step, using multiple random spatial partition trees. An adaptive initialization method (IM) can be used to initialize other partitional clustering algorithms such as fuzzy c-means and its variants (Celebi, Kingravi & Vela, 2013).

2.2. DEA – BCC model

DEA was first developed by Farrel (1957) and improved by Charnes, Cooper & Rhodes (1978) as a non-parametric method that compares a decision making-units directed to frontiers rather than central tendencies, using performance indicators. This tool allows identifying which units are efficient, besides defining benchmarks between such units. DEA has two classical models, which are the CRS (Charnes et al, 1978) and BCC (Banker et al, 1984). The CRS model works with constant returns to scale while the BCC assumes that the DMUs can exhibit variable returns to scale. Both models can work with input orientation, where inputs are minimized while outputs remain unchanged, or can work with output orientation when the objective is to maximize the outputs while inputs do not change.
In this paper, the DEA BCC model was chosen due to the fact that increasing input variable – in this study *Total budget* won"t increase the same quantity of output variable. Therefore, the variables do not vary proportionally. The output orientation is preferred since the main purpose of the basketball teams is to have high scores and to win the matches.

Let x_{ij} - denote the observed magnitude of *i* - type input for entity *j* ($x_{ij} > 0$, $i = 1, 2, ..., m$, $j = 1, 2, ..., n$) and y_{rj} - the observed magnitude of *r*-type output for entity *j* ($y_{rj} > 0$, $r = 1, 2, ..., s$, $j = 1, 2, ..., n$). Then, the basic DEA model is formulated in the following form for the selected entity k:

$$(\max)\, h_k = \sum_{r=1}^{s} u_r y_{rk} \tag{3}$$

$$\sum_{i=1}^{m} v_i x_{ik} = 1 \tag{4}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, (j = 1,...,n) \quad (5)$$

$$u_r \geq \varepsilon, r = 1,2,..., s$$
$$v_i \geq \varepsilon, i = 1,..., m) \quad (6)$$

Where:

- v_i is the weights to be determined for input *i*;
- *m* is the number of inputs;
- u_r is the weights to be determined for output *r*;
- *s* is the number of outputs;
- h_k is the relative efficiency of DMU_k;
- *n* is the number of entities;
- *ε* is a small positive value.

The relative efficiency h_k, of one decision-making unit *k*, is defined as a ratio of the weighted sums of their outputs (virtual output) and the weighted sums of their inputs (virtual input). As for the decision-making unit *k*, for which a maximum in the objective function (3) is sought, the condition (4) is true, meaning that it is obviously $0 < h_k \leq 1$, for each DMU_k. The weights v_i and u_r show the importance of each input and output and are determined in the model so that each DMU is efficient as much as possible. Given that the condition (6) is true for every DMU, it means that each of them lies on the efficiency frontier or beyond it. If Max $h_k = h_k^* = 1$, it means that efficiency is being achieved, so we can tell that DMUk is efficient. Efficiency is not achieved for $hk^* < 1$ and DMUk is not efficient in that case. The mathematical model given above is linear and can be solved using any of the familiar programs packages for LP (Martic et al., 2009).

Variable Description

Table 1: ABA league 2012-13 Teams and Variables used for analyses

Team information		***K-Means Clustering***			***DEA***		
		Inputs			*Inputs*	*Outputs*	
Home country	**Team**	**Average home attendance**	**Total kilometres travelled**	**GDP per capita (€)**	**Total budget (mil €)**	**Victories**	**Average score**
Bosnia and Herzegovina	Igokea	2160,00	18072	11,1	2,5	20	46
Bosnia and Herzegovina	Siroki WWin	2807,69	22944	7,963	2,5	9	35
Croatia	Cedevita	1076,92	19160	13,493	7	15	41
Croatia	Zadar	4076,92	24456	17,8	2,5	9	35
Croatia	Cibona	1916,67	19160	13,493	3,7	9	35
Croatia	Split	1861,54	23976	10,47	1,9	8	34
Hungary	Szolnoki Olaj	1600,00	30264	12,39	3	9	35
Macedonia	MZT Skopje	3080,00	35068	3,688	2,5	14	40
Montenegro	Aerodrom Buducnost VOLI	1960,00	26896	5,211	2,8	16	42
Serbia	Crvena Zvezda	5097,07	20896	7,136	4	18	44
Serbia	Telekom Partizan mt:s	5178,57	20896	7,136	4,5	16	42
Serbia	Radnicki	3040,00	24232	7,4	3,25	17	43
Slovenia	Union Olympic	2100,00	24264	10,1	4	13	39
Slovenia	Krka	980,00	21244	17,379	3,7	9	35

ABA league teams are grouped based on the following set of inputs: *Average attendance* at home matches, *Total kilometres* travelled by the team during the regular season and *Average GDP per capita* of host cities. For the purpose of DEA analyses, the author had used *Total budget* (in million euros), as an input variable, and *Number of Victories* followed by obtained *Average Score* as output variables. All of the mentioned variables are shown in **Table 1.**

Average home attendance – defines the level of public participation in the games that occurred in the own city. Higher *home attendance* is correlated with higher *home advantage*; teams with a relatively high *home advantage* perform better in front of larger crowds, while teams with a relatively low *home advantage* perform better in front of smaller crowds (Willoughby, 2014). As a half of the matches were played in team's hall, that increased *home advantage* attendance.

Total kilometres travelled – expresses the distance travelled by each team during the regular season. This variable can also be interpreted as the effort that fans of a team would make in order to attend the match; it should be lower the participation of its fans at the games, as farther the game is played. Scheduling tournament is also important for this point of view because it's directly connected with the team budget (Ribeiro, 2012).

The last variable chosen for clustering – *Average GDP per capita* – is the average GDP per capita of all cities that hosted the games that each team participated. This variable reflects the economic situation of the city's public – how much people would pay for the tickets for the games occurred in their town. Also, it has the direct impact on future participation in sport (Humphreys, Maresova, & Ruseski, 2012).

Total budget - represents the amount of money that each team had for organizing tournament during the season. The budget expresses in which condition is team – if the team has better results, players will be desired by another team, the performance of the team will increase (Lera-López *et. al*, 2013).

The number of victories - shows the main purpose of a team, which is winning the game. Also, another variable which has almost the same meaning is the score reached. Observing this, the author used DEA BSS model oriented to output, in order to maximize score with the minimum of available input.

3. RESULTS AND DISCUSSION

This section presents the results obtained from the application of the proposed approach described in the previous section of this paper. **Table 2** presents the clusters found and the efficiency scores assessed for teams in each cluster.

Table 2: Team country, main cluster, and team efficiency

Home country	City	Team	Cluster	Efficiency
Bosnia and Herzegovina	Široki Brijeg	Siroki WWin	C1	1
Slovenia	Ljubljana	Union Olimpija	C1	0,932
Slovenia	Novo Mesto	Krka	C1	1
Croatia	Zagreb	Cedevita	C2	0,848
Croatia	Zadar	Zadar	C2	0,784
Croatia	Zagreb	Cibona	C2	0,913
Croatia	Split	Split	C2	1
Hungary	Szolnok	Szolnoki Olaj	C2	1
Macedonia	Skopje	MZT Skopje Aerodrom	C2	0,784
Montenegro	Podgorica	Buducnost VOLI	C2	0,77
Bosnia and Herzegovina	Aleksandrovac	Igokea	C3	1
Serbia	Beograd	Crvena Zvezda Telekom	C3	0,875
Serbia	Beograd	Partizan mt:s	C3	0,875
Serbia	Kragujevac	Radnicki	C3	0,85

The methodology of clustering enabled the author to makes groups of units which have some similar characteristics. That provides fair analyses of efficiency since the teams were compared with other teams that played in similar circumstances.

From Cluster 1 teams it can be seen that Krka has the smallest *home attendance*, as well as *total kilometres travelled* during the season, but has the biggest *GDP per capita* in front of other countries. These opposite values put it next to the teams Union Olympia and Siroki WWIn. Interesting is that even Krka has the highest GDP per capita, home matches are not enough known. Higher GDP means better health and education outcomes, better sporting infrastructure and more resources spent for sports (Economictimes, 2016). Siroki WWin has more fans than Union Olympic, but they were less active.

Most of the Cluster 2 teams present low values for the *Average home attendance* and high values related to the average *GDP per capita* of home cities. Climate can play a significant role in home advantage. All Croatian cities have Mediterranean climate with a lot of rain during the year. That involves people to watch the match at home. Skopje had the lowest *GDP per capita* for the year 2013, but urbanization of that city influenced *home attendance*.

On the other hand, Cluster 3 teams played under favourable conditions, since they mostly played in cities with a high average *GDP per capita*, and had a high home advantage. Furthermore, these teams present a middle value looking at *total kilometres travelled*. Played matches between teams in this Cluster were in top 10 attendance (Wikipedia, 2013), since Belgrade is an attractive turist destination.

The second part of the methodology is to calculate the efficiency indexes. The obtained results have shown interesting values about the performance of the teams. The author got in two clusters pairs of teams which have efficiency equal to 1. As the efficient teams from Cluster 1 author can observe Siroki WWin and Krka from Slovenia. Union Olympic considers as an inefficient team because of the amount of budget (4 million Euros) and just 4 matches more victories, as well as 4 points greater average score than Krka. For the season 2012-13 players didn't make game very well, they bypass the top 10 players.

In the Cluster 2, Split and Szolnoki have the efficiency equal to 1. Split had the minimum budget for the season, even 3 times smaller than the highest one, but also had just eight victories more and 4 points more than others. Split and Szolnoki round by round became at the top of the best player. That ends with high scores.

Figure 2 shows Final rank of the teams, as well as achieved points in the final games. It can be seen that all "Final Four" teams are grouped in the same cluster.

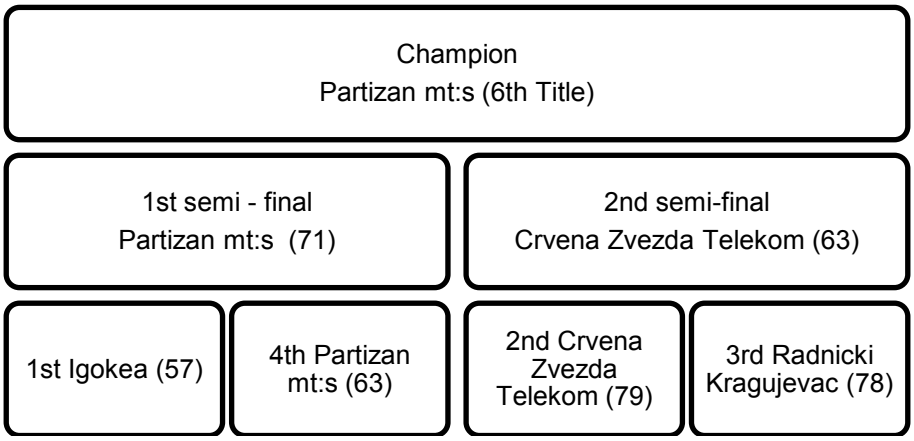

Figure 2: Results and Final rank

From the strongest teams, only Igokea was ranked 1st in its cluster. It's pretty surprisingly because Igokea ranked 2nd in value for variable *Total budget* (in a group with Siroki WWinn, Zadar and MZT Skopje), and ranked 1st based on the *Number of victories* and *Average score*. Even Igokea had the greatest efficiency, that wasn't enough for the victory. Going further, two other teams in the same cluster (Crvena Zvezda Telekom and Partizan mt:s) had almost 2 times the greater budget for the season, but a smaller number of victories and scores. With the same efficiency, the both of teams played in Semi-Final and the champion of the season (Partizan mt:s) was neither first in its cluster. The study also found that Radnicki Kragujevac, the team with the smallest efficiency in it's cluster ranked 3rd looking at the *Number of victories* and *Average scores.*

4. CONCLUSION

In this paper, the author shows through DEA methodology efficiency measures of 14 basketball teams belonging to the ABA league 2012/13 regular season. Since the approach has never been applied on the ABA league before, the author has developed a framework for teams analyses, by clustering the teams into effective DMUs. Another main feature of the proposed approach is to consider the importance of obtained results – how to plan and allocate the budget, or how to enforce the public to support their teams on the upcoming matches.

Regarding teams analysis, the author can observe the performance of teams in each group - Cluster 2 presents the greater difference between maximum and minimum values for each efficiency score, while Cluster 1 presents the lowest difference. All of the teams in Cluster 1 are efficient, and teams with the highest and lowest composed efficiency were grouped in Cluster 2. In this context, the recommendation for the Team management is to consider greater budget for the next regular season, in order to allow better conditions, as well as better performances of the team players.

For the future research work, the author suggests additional analyses of team players, considering the fact that each team belongs to a specific cluster. Another topic of further research can be to evaluate the change in efficiency for every team in a specific period of time. Than, the application of statistical methods could be significant, in order to reduce the number of related problems, such as a high degree of correlation between proposed variables; or to show the most important variables, applying the I-distance methodology.

REFERENCES

Aba-league. Retrieved March 2, 2016, from http://www.aba-liga.com/Home.aspx.

Aizemberg, L., Roboredo, M. C., Ramos, T. G., de Mello, J. C. C. S., Meza, L. A., & Alves, A. M. (2014). Measuring the NBA Teams" Cross-Efficiency by DEA Game. *American Journal of Operations Research*, *4*(03), 101.

Alves, L. A., Kramer, H. H., Tschaffon, P. B., & de Mello, J. S. (2013). Assessing efficiency and setting benchmarks for NBA teams through DEA and DMU clustering. In *Proceedings of the 4th International Conference on Mathematics in Sport*.

Anderson, M., Wolfson, S., Neave, N., Moss, M. (2012) Perspectives on the home advantage: A comparison of football players, fans and referees. *Psychological of Sport and Exercise* **13**, 311-316.,

Anderson, T. R. (2004). Benchmarking in Sports. In *Handbook on data envelopment analysis* (pp. 443-454). Springer US.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis.*Management science*, *30*(9), 1078-1092.

Barros, C. P., & Santos, A. (2003). Productivity in sports organisational training activities: a DEA study. *European Sport Management Quarterly*,*3*(1), 46-65.

Bulajic, M., Savic, G., Savic, S., Mihailovic, N., & Martic, M. (2011). Efficiency assessment of banks in Serbia. *Technics Technologies Education Management-TTEM*, *6*(3), 657-662.

Celebi, M. E., Kingravi, H. A., & Vela, P. A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*, *40*(1), 200-210.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444.

Cooper, W. W., Ruiz, J. L., & Sirvent, I. (2009). Selecting non-zero weights to evaluate effectiveness of basketball players with DEA. *European journal of operational research*, *195*(2), 563-574.

De Mello, J. C. C. S., Angulo-Meza, L., & Da Silva, B. P. B. (2009). A ranking for the Olympic Games with unitary input DEA models. *IMA Journal of Management Mathematics*, *20*(2), 201-211.

Delibašić, B., Kirchner, K., Ruhland, J., Jovanović, M., & Vukićević, M. (2009). Reusable components for partitioning clustering algorithms. *Artificial Intelligence Review*, *32*(1-4), 59-75.

Delibašić, B., Vukićević, M., Jovanović, M., Kirchner, K., Ruhland, J., & Suknović, M. (2012). An architecture for component-based design of representative-based clustering algorithms. *Data & Knowledge Engineering*,*75*, 78-98.

Deza, M. M., & Deza, E. (2009). *Encyclopedia of distances* 1-583. Springer Berlin Heidelberg.

Economictimes. Retrieved April 7, 2016, from http://articles.economictimes.indiatimes.com/2011-04-03/news/30149956_1_higher-gdp-population-medals.

European National Baketball Ranking (ENBR). Retrieved March 2, 2016, from https://en.wikipedia.org/wiki/European_national_basketball_league_rankings#Current_league_rankings

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, *120*(3), 253-290.

Humphreys, B. R., Maresova, K., & Ruseski, J. E. (2012). *Institutional factors, sport policy, and individual sport participation: An international comparison* (No. 2012-1). University of Alberta, Department of Economics.

Jeremic, V., Bulajic, M., Martic, M., & Radojicic, Z. (2011). A fresh approach to evaluating the academic ranking of world universities. *Scientometrics, 87* (3), 587-596. doi:10.1007/s11192-011-0361-6

Jeremic, V., Bulajic, M., Martic, M., Markovic, A., Savic, G., Jeremic, D., et al. (2012). An Evaluation of European Countries" Health Systems through Distance Based Analysis. *Hippokratia, 16* (2), 170-174.

Labonté, R. (2016). From the Myth of Level Playing Fields to the Reality of a Finite Planet; Comment on'A Global Social Support System: What the International Community Could Learn from the United States" National Basketball Association's Scheme for Redistribution of New Talent'. *Int J Health Policy Manag*, *5*(2), 137-139.

Lera-López, F., Lizalde-Gil, E., Hallmann, K., & Petry, K. (2013). 12.2 Sport System and the Structure of Organized Sport. *Comparative Sport Development: Systems, Participation and Public Policy*, *8*, 149.

Lewis, H. F. (2014). Performance measurement of major league baseball teams using network DEA. In *Data Envelopment Analysis* 475-535. Springer US.

Ma, T. L., Wang, L., & Wang, D. (2014). Fuzzy C-Mean Clustering Image Segmentation Algorithm Research for MRI Sport Graphics Based on Genetic Algorithm. In *Applied Mechanics and Materials* . Vol. 513, 3711-3714.

Martić, M., Novaković, M., & Baggia, A. (2009). Data envelopment analysis-basic models and their Utilization. *Organizacija*, *42*(2), 37-43.

Moreno, P., & Lozano, S. (2014). A network DEA assessment of team efficiency in the NBA. *Annals of Operations Research*, *214*(1), 99-124.

Petrović Djordjević, D., Vujošević, M., & Martić, M. (2015). MEASURING EFFICIENCY OF FOOTBALL TEAMS BY MULTI-STAGE DEA MODEL.*Tehnicki vjesnik/Technical Gazette*, *22*(3).

Radojicic, M., Savic, G., Radovanovic, S., & Jeremic, V. (2015). A NOVEL BOOTSTRAP DBA-DEA APPROACH IN EVALUATING EFFICIENCY OF BANKS. *Scientific Bulletin" Mircea cel Batran" Naval Academy*, *18*(2), 375.

Radovanović, S., Radojičić, M., & Savić, G. (2014). Two-phased DEA-MLA approach for predicting efficiency of NBA players. *Yugoslav Journal of Operations Research ISSN: 0354-0243 EISSN: 2334-6043*, *24*(3).

Ramón, N., Ruiz, J. L., & Sirvent, I. (2012). Common sets of weights as summaries of DEA profiles of weights: with an application to the ranking of professional tennis players. *Expert Systems with Applications*, *39*(5), 4882-4889.

ReadingTree. Retrieved March 16, 2016, from http://www.readingtree.org/top-10-most-popular-sports-in-the-world/.

Ribeiro, C. C. (2012). Sports scheduling: Problems and applications.*International Transactions in Operational Research*, *19*(1-2), 201-226.

Ruiz, J. L., Pastor, D., & Pastor, J. T. (2013). Assessing professional tennis players using data envelopment analysis (DEA). *Journal of Sports Economics*, *14*(3), 276-302.

Savic, G., Makajic-Nikolic, D., & Suknovic, M. (2012). AHP-DEA Measure for study program selection. *Symorg* (pp. 1217-1223). Zlatibor: University of Belgrade, Faculty of Organizational Sciences.

Savic, G., Radosavljevic, M., & Ilijevski, D. (2013). DEA Window Analysis Approach for Measuring the Efficiency of Serbian Banks Based on Panel Data. *Management.* doi:10.7595/management.fon.2012.0028.

Tan, P. N., Steinbach, M., & Kumar, V. (2013). Data mining cluster analysis: Basic concepts and algorithms.

Tiedemann, T., Francksen, T., & Latacz-Lohmann, U. (2011). Assessing the performance of German Bundesliga football players: a non-parametric metafrontier approach. *Central European Journal of Operations Research*,*19*(4), 571-587.

Vazirgiannis, M., Halkidi, M., & Gunopulos, D. (2012). *Uncertainty handling and quality assessment in data mining*. Springer Science & Business Media.

Wang, J., Wang, J., Ke, Q., Zeng, G., & Li, S. (2015). Fast approximate k-means via cluster closures. In *Multimedia Data Mining and Analytics* 373-395. Springer International Publishing.

Wikipedia. Retrieved April 11, from https://en.wikipedia.org/wiki/2012%E2%80%9313_ABA_League.

Willoughby, Jack. (2014). Attendance, Home Advantage, and the Effect of a City on its Professional Sports Teams. Duke University Durham, North Carolina.

Yu, P., & Lee, J. H. (2013). A hybrid approach using two-level SOM and combined AHP rating and AHP/DEA-AR method for selecting optimal promising emerging technology. *Expert Systems with Applications*, *40*(1), 300-314.

APPLICATION OF GENETIC ALGORITHM IN AUTOMATED TRADING

Pavle Jovanović*[1], Aleksandar Rakićević[1], Bratislav Petrović[1]
[1]University of Belgrade, Faculty of organizational sciences
*Corresponding author, e-mail: pavlej93@gmail.com

Abstract: *This paper describes the use of genetic algorithm in the automated trading. Automated trading makes up a high percentage of the total trading volume which makes it very promising for future research. The proposed trading system is based on genetic algorithm which is here used to optimize parameters of a trading strategy in order to obtain the most profitable combination. The proposed trading strategy uses technical indicators to extract trading signals from market prices. Further, we use several performance measures to evaluate trading successfulness. The proposed trading system is tested in US stock market data.*

Keywords: *automated trading, genetic algorithm, optimization, stock market, technical indicators, trading strategy.*

1. INTRODUCTION

Automated trading (AT), also known as algorithmic trading, represents electronic trading in computerized manner, using computer programs and algorithms. Following development of information-communication technologies, AT raised as the new area of modern finance. In order to speed-up their reactions to new information and profit opportunities, market players started introducing algorithms in market trading at the end od 1980s and during 1990s. At the first time, these algorithms were meant to accelerate trading order execution, but with time the algorithms became more complex and intelligent. During the 2000s, AT has experienced extensive expansion in use and popularity. Kissel (2013) states that the total market volume made by automated trading systems raised from 25% in 2005 to almost 75% in 2009.

The main advantage of an automated trading system, if its trading algorithm is well written, is its speed. Today's market is very dynamic and unpredictable. Due to increasingly rapid changes in the market, it often happens that a man is not able to react to new information or profit opportunity in a timely manner, even if those are foreseen in advance. In order “to compete the market”, one needs to use some kind of computerized support. Computer algorithms have ability to react to a rapidly changing market conditions and execute transactions before another change occurs.

There are two main approaches to financial market analysis. Fundamental analysis (FA) relies on financial statement data, analyze company"s business performances. On the other side, technical analysis (TA) uses historical market data such as price, volume, bid, ask etc. to build technical indicators – models that attempt to mathematically and graphically describe the movement on the market. Many traders and other market participants use these indicators for trading and for building automated trading systems.

The first work in which the genetic algorithm (GA) was used to optimize a trading strategy was done by Allen and Karjalainen (1999). They used GA to learn technical trading rules for the S&P 500 index using daily prices from 1928 to 1995. Outside of training data (out-of-sample) their algorithm was not giving much of a good results, but it was very simple and its parameters were not necessarily optimized. Their work represented the basis for further research of many authors.

Deng *et al.* (2012) reported that trading profit is a discrete and multimodal function with many parameters and that therefore genetic algorithms and genetic programming may be used to optimize the trading rules in order to obtain profitable trading strategy. Their trading environment was currency market (FOREX). Their trading strategy was based on technical indicators, while the GA was used to optimize parameters of the strategy. Cheung and Kaymak (2007) were among those who tried to optimize fuzzy-rule-based trading strategy. They used GA to set the membership functions of fuzzy sets within the if-then rules in fuzzy inference system.

In this study a stock market trading strategy based on TA was created. We use three technical indicators to obtain buy/sell signals. Further, we apply GA to optimize input parameters for these indicators in order to discover set of parameters that will ensure the profitable outcome of stock market trading. We use profit factor as a performance measure of strategy successfulness. Profit factor is also used as a fitness function

for genetic algorithm. To test our model, we create virtual trading engine to simulate stock market trading. We use real data from American stock market.

The remainder of this paper is organized in four sections. In Section 2 we will give a theoretical basis of genetic algorithm. The proposed model is presented in Section 3. It gives necessary details about trading strategy and technical indicators being used to build it. In this section, we also explain the proposed mechanism for GA which is used to optimize trading strategy. In Section 4 we explain stock market data, experimental settings and performance measures used for testing the proposed model. Further, we present results and briefly discuss them. Finally, the paper is concluded with Section 5.

2. GENETIC ALGORITHM

Genetic algorithms (GAs) were invented by John Holland in the 1960s. Mitchell (1999) report that „ Holland's 1975 book *Adaptation in Natural and Artificial Systems* presented the genetic algorithm as an abstraction of biological evolution and gave a theoretical framework for adaptation under the GA. Holland's GA is a method for moving from one population of „chromosomes" (e.g., strings of ones and zeros, or „bits") to a new population by using a kind of „natural selection" together with the genetics−inspired operators of crossover, mutation, and inversion. Each chromosome consists of „genes" (e.g., bits), each gene being an instance of a particular „allele" (e.g., 0 or 1)." Accordingly we can say that GAs mimic the process of natural selection. Each individual in a population is characterized by a chromosome that represents its genetic code. Those individuals showing better adaptation (better fitness) in problem environment get the greater opportunity to reproduce themselves. This is the mechanism that allows "better genes" to spread across the population. For this reason, the new generation of individuals is better adapted from the previous one.

One of the main strengths of the GAs lie in the fact that they are able to determine the position of the global optimum in the area with several local extremes, in the so-called multimodal space. This is the reason why they are widely and extensively practiced for optimization purpose. However, as with other heuristics, there is no guarantee for optimal solution (global optimum).

There are three most important features, also known as genetic operators, which largely determine the success of a genetic algorithm. Those features are:

- reproduction (multiplication, crossover)
- mutations and
- selection.

The reproduction starts when two parents are selected to participate in the process. The most important characteristic of reproduction process is that the children inherit the properties of their parents. If the parents possess good genetic material then it is most likely that the child will posses good, if not better genetic material than the parents. Reproduction operator is binary and it can be defined as: uniform reproduction, reproduction with one breakpoint and reproduction with multiple breakpoints.

The mutation represents a random change of one or more alleles in a gene sequence. It is an unary operator, since it acts on one unit. The result of mutations is a changed unit. Mutation operator is very important since it helps GA to avoid local extremes and to restore lost genetic material.

The purpose of the selection process is the preservation and transfer of good traits to the next generation of units. It imitates the famous mechanism of "natural selection". The selection is used to choose best fitted units in a current population to enable them to participate in reproduction process. In this way the good genetic material is given opportunity to spread in next generation, while the bad material is rejected. However, continuous selection of the most fitted units and discarding others (so called elitism strategy) will lead to premature convergence of genes in a population. This effect usually ends with GA stucked in local optimum.

3. THE PROPOSED MODEL

The proposed model is consisted of input parameters, trading strategy and GA. In the following parts of this section we will give detail explanation of these components.

3.1. Technical indicators

In the following Table 1 we give details about technical indicators used for our trading strategy. These indicators are used to explain underlying model of a market and to signal trading opportunities for our trading system.

Table 1: Technical indicators

Indicator	Formula	Description
MACD line	$MACD(N,M) = EMA(N) - EMA(M)$	Moving Average Convergence/Divergence line tracks a difference between longer and shorter trend of the price
ROC	$ROC(N) = \frac{Price_t - Price_{(t-N)}}{Price_{(t-N)}} \cdot 100$	Rate of Change is designed to show the relative difference between today's closing price and the close N days ago
PSAR	$SAR_{(i)} = SAR_{(i-1)} + \alpha \cdot (ExtrimePrice - SAR_{(i-1)})$	Parabolic stop and reversal try to determine the end of current trend.

3.2. Strategy

Input parameters for the strategy are as follows:

- Input data are daily data consisted of opening, highest, lowest and closing prices.
- The initial trading account value
- Stop-loss and take-profit parameters
- Input parameters of previously explained technical indicators

When strategy is activated it first calculates indicator values for input price data using parameters provided with GA. Further, the strategy monitors indicators" values in order to check whether the trading conditions are met. If conditions are met, the strategy further checks if the position is already opened by the system or not. The strategy is written so that at one point it cannot manage more than one opened position. When position is finally opened, than the stop-loss and take-profit orders are set as closing conditions. Take-profit order is a trading order type that allows one to close out a profitable trade at a pre-determined price. On the other side, a stop-loss order is a protective order that closes out a losing position at some pre-determined price. When position is opened, the system monitors market price and if the price hits level of either one of these two orders, system closes position. When position is closed system recalculates account balance along with performance measures.

3.3. Model of genetic algorithm

The goal of this study is to apply GA to optimize trading strategy for automated trading. To optimize trading strategy we use profit factor as criteria function. Profit factor is a ratio of trading profits and losses. It is a discrete and multimodal but unknown function of many parameters. For this reason, GA represents an adequate tool for solving this problem. Since our trading model is based on technical indicators, the profit factor will be a function of indicators parameters.

The model of GA is probably the most important piece of work. Figure 1 shows the proposed model of GA used to optimize the proposed trading strategy.

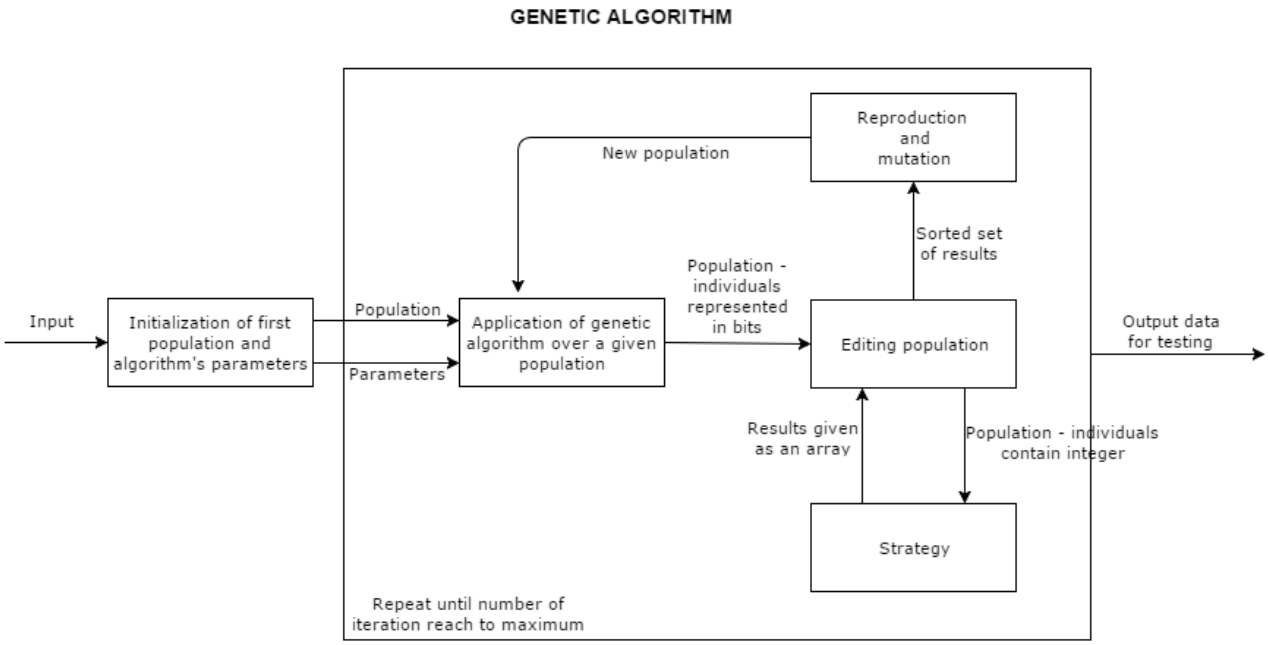

Figure 1: Model of the described genetic algorithm

In this study, we use binary representation to code genes in our GA. Consequently, genetic operators are also binary. One individual is represented with genetic sequence consisted of 5 genes and total of 27 alleles. Population is consisted of 64 individuals. Initial population is created randomly. Further, the selection procedure is such to retain certain percentage of the best individuals in population. The rest of population enters the reproduction process and creates children that will replace their parents in a new population. It is important to be noted that the better individual is fitted the greater chance for reproduction is assigned to it. In order to achieve this, we rank the rest of population by fitness and then split it in three parts. The first third of individuals is then duplicated before entering the reproduction pool. In that way we increase probability for their selection. We use multiple breakpoints as reproduction operator, which takes alleles alternately from both parents. The mutation is performed by random selection of alleles to be mutated. The number od alleles to be mutated (from 1 to 9) and their position in gene sequence (from 1 to 27) are determined randomly. Therefore, the mutation rate is not fixed but it changes through generations.

Our genetic code is consisted of 27 alleles which are used to represent 5 genes. Each gene represents one parameter in one of the three selected indicators. Parameters are coded in the following way:

- The first gene represents time window parameter of fast exponential moving average (EMA) within MACD indicator. It is coded as binary sequence consisted of 4 alleles which represent, when converted to integers, values between 0 and 15.
- The second gene is time window parameter of slow EMA within MACD indicator. It is coded with binary sequence of 5 alleles representing integer values between 0 and 31.
- The third gene is a binary sequence consisted of 5 alleles, representing time window in ROC indicator with values between 0 and 31.
- Fourth gene is a take-profit parameter described with a binary sequence of 7 alleles, representing integer values between 0 and 127.
- Finally, the fifth gene is a stop-loss parameter. It is a binary sequence of 6 alleles representing integer values between 0 and 63.

Evaluation of individuals is done in the following manner. Each individual is a set of parameters that are incorporated in trading strategy and tested against stock market data to obtain profit factor. The profit factor is adopted as a fitness function. Each individual is evaluated within population, while the process is repeated for each generation.

4. EXPERIMENTAL RESULTS

In this chapter we will present test data, experimental settings, performance measures and experimental results with the corresponding discussion.

4.1. Data

The data used in this experiment is a real US stock market data. We focus on two companies: Microsoft corporation (MSFT) and Nike Inc. (NKE). The data includes price series from 1 January 1994 to 31 December 2013. Data is separated in two samples: training data (in-sample) and test data (out-of-sample).

Table 2: *Microsoft data*

Date	Open	High	Low	Close
3.1.1994	2,52	2,52	2,48	2,51
4.1.1994	2,51	2,52	2,48	2,52
...	...	...	...	...
11.5.2000	33,31	24,06	32,88	33,94
12.5.2000	34,22	34,88	34,12	34,41
...	...	...	...	...
4.1.2007	29,7	29,97	29,44	29,81
5.1.2007	29,63	29,75	29,45	29,64
....	....	...	...	...
30.12.2013	37,22	37,38	36,9	37,29
31.12.2013	37,4	37,58	37,22	37,41

In Table 2 we present the structure of data. Data contains daily prices of shares including:
- price at the beginning of the trading day (open price)
- price at the end of the trading day (close price)
- the highest price for the trading day (high price)
- the lowest price for the trading day (low price)

4.2. Experimental settings

Here, we present the basic experimental settings of our model. The values of parameters used in the algorithm:
- Initial trading funds are 100.000 monetary units,
- Maximum number of iterations is 200 (which means GA will have 200 generations),
- Maximum population size is 64,
- Number of parents that are transferred to a new generation is 8 (12,5%),
- Number of children obtained is 56
- The number of alleles that will mutate in a selected individual is a random number between 1 and 9,
- Trading commission is accounted for 3%

4.3. Performance measures

The following measures are used to evaluate trading strategy performances:
- **Profit factor –** the ratio of total gains and total losses made by the strategy

$$Profit\ Factor = \frac{Total\ Gains}{Total\ Losses}; \tag{1}$$

- **Drawdown –** the biggest drop in trading funds relatively to the maximum achieved level;
- **Profit –** total net earnings made by the strategy; difference between final and initial size of trading funds.

4.4. Results and discussion

Having in mind the structure of genetic operators and the fact that initial population is chosen randomly, it is obvious that each new startup of our model will give different final solution. However, the solutions vary a little which indicates the obtained final solutions are relatively stable. In the following Tables 3, 4 and 5, we present results for best individuals (parameter combinations) obtained for MSFT in training period from 1 January 1994 to 31 December 2006.

Table 3: Best 5 individuals in training results for the first generation

Profit factor	Profit	Drawdown	Number of transaction	Short	Long	Roc period	Take profit	Stop loss
2,41	158,81%	-62,74%	13	3	13	10	13%	25%
1,26	51,33%	-72,52%	11	3	24	24	112%	21%
1,02	6,94%	-87,25%	14	13	20	22	47%	22%
0,94	-13,44%	-84,57%	11	4	19	27	64%	22%
0,89	-33,41%	-90,83%	14	14	27	16	114%	23%

Table 4: Best 3 individuals in training results for the 100th generation

Profit factor	Profit	Drawdown	Number of transaction	Short	Long	Roc period	Take profit	Stop loss
144.584,81	144,58%	-66,51%	14	11	30	14	5%	44%
144.584,81	144,58%	-66,51%	14	11	30	14	5%	46%
137.699,97	137,7%	-62,84%	14	11	12	12	5%	44%

Table 5: Best 3 individuals in training results for the 200th generation

Profit factor	Profit	Drawdown	Number of transaction	Short	Long	Roc period	Take profit	Stop loss
356.839,98	356,84%	-61,62%	10	10	30	10	15%	58%
322.377,43	322,37%	-61,62%	10	10	30	10	14%	60%
248.681	248,68%	-58,69%	18	10	30	10	6%	28%

As we can see in Table 3, there are only three profitable combinations of parameters in the initial population. Results for 100th generation shows that population contains 14 individual that are evaluated almost without losses. From the best three ones that are shown in Table 4, we can conclude that parameters begin to converge to certain values. This mean they are located in the vicinity of an optimum (local or global). In Table 5 we present results for the last 200th generation. Individuals in this generation are even more successful (with higher profit factor) than those in 100th generation.

Optimization of strategy using genetic algorithm took 2.348 seconds or a little less than 40 minutes. Compared to, for example, sequential optimizations, it takes much less time and the obtained final solutions are good enough. It is worth mentioning that the profit factor of the best strategy obtained by using genetic algorithm optimization is not much lower than the one obtained by sequential optimization. For this reason the time aspect is crucial. In order to react adequately and fast enough to changes in the market the data should be processed as quickly as possible. This information shows the importance and potential of using genetic algorithm to optimize trading strategies.

Tables 6 and 7 indicate the test results of the best strategies for MSFT and NKE in the period from 3 January 2007 to 31 December 2013.

Table 6: Test results for the best 3 individuals in training set (MSFT)

Profit factor	Profit	Drawdown	Number of transaction	Short	Long	Roc period	Take profit	Stop loss
53.415,46	53,42%	-62,59%	3	10	30	10	15%	58%
73.114,59	73,11%	-62,59%	4	10	30	10	14%	60%
0,84	-6,11%	-66,58%	6	10	30	10	6%	28%

Table 7: Test results for the best 3 individuals in training set (NKE)

Profit factor	Profit	Drawdown	Number of transaction	Short	Long	Roc period	Take profit	Stop loss
53.286,44	53,29%	-70,13%	3	10	30	10	15%	58%
1,52	17,92%	-50,21%	9	10	30	10	14%	60%
0,419	-34,71%	-68,74%	8	10	30	10	6%	28%

As can be seen from these tables, the first two strategies tested on the MSFT data operated positively. However, the gains were reduced. Contrary to the first two, test simulation of the third strategy shows negative profitability. The similar is with NKE data, though it should be noted that the second strategy has seen a big drop in profitability even though it remains positive. The reason for the fall in profitability of these strategies can be found in the fact the year of great economic crises in 2008 was covered by the test data sample. Such market behavior was not seen before that in train data which is the possible reason why the best fitted individuals in training experienced big drop in performance in test. The way in which this could be

overcome is to monitor the accounts and to stop the algorithm when losses hit a certain percentage of the maximum resources available, and then to have it re-learn from the new dataset including this period. This is one of the ideas for the future extensions of this study.

4. CONCLUSION

This paper describes stock market trading using genetic algorithm to optimize trading strategy parameters. The strategy employs technical indicators to obtain trading signals. The task of genetic algorithm is to optimize the parameters of technical indicators in order to find the most profitable trading strategy.

On the basis of results obtained in this paper, we can conclude that the genetic algorithm was successful in finding a strategy that will operate positively. We also conclude that one cannot absolutely rely on trading using the proposed algorithm, since the final solutions experience big drop in performances during test period. Further, the number of profitable trading strategies is not great. However, a good knowledge of financial instruments and properly written algorithm may lead to the creation of profitable trading strategies. Generally, the use of algorithms for trading on the stock market represents the future of automated trading. As automated trading makes up over 50% of the total trading on the stock exchange, we conclude that the application of genetic algorithm trading can be very useful and that it has a promising future.

Possible future directions of work on this topic may relate toward increasing the number of indicators used in the strategy and introducing fuzzy or some kind of multi-valued logic to model trading rules.

REFERENCES

Cheung, W. M., & Kaymak, U. (2007). A fuzzy logic based trading system. *Proceedings of the Third European Symposium on Nature-inspired Smart Information Systems* (Vol. 59).

Colby, R. W., & Meyers, T. A. (1988). *The encyclopedia of technical market indicators*. New York: Irwin.

Franklin A., & Karjalainen, R. (1999). Using genetic algorithms to find technical trading rules. *Journal of Financial Economics*, 51(2), 245-271. doi:10.1016/S0304-405X(98)00052-X

Kim, K. (2010). *Electronic and algorithmic trading technology: the complete guide*. Academic Press.

Kissell, R. (2013). *The science of algorithmic trading and portfolio management*. Academic Press.

Leshik, E. A., & Cralle, J. (2011). *An Introduction to Algorithmic Trading*. Padstow: John Wiley and Sons.

Mitchell, M. (1998). *An introduction to genetic algorithms*. MIT press. Cambridge, Massachusetts., England: London

Ozturk, M., Toroslu, I. H., & Fidan, G. (2016). Heuristic based trading system on Forex data using technical indicator rules. *Applied Soft Computing*, *43*, 170-186. doi:10.1016/j.asoc.2016.01.048

Papadamou, S., & Stephanides, G. (2007). Improving technical trading systems by using a new MATLAB-based genetic algorithm procedure. *Mathematical and computer modelling*, *46*(1), 189-197. doi:10.1016/j.mcm.2006.12.033

Shangkun, D., Yizhou, S., & Akito, S. (2012). Robustness test of genetic algorithm on generating rules for currency trading. *Procedia Computer Science*,*13*, 86-98. doi:10.1016/j.procs.2012.09.117

Straßburg, J., Gonzàlez-Martel, C., & Alexandrov, V. (2012). Parallel genetic algorithms for stock market trading rules. *Procedia Computer Science*, *9*, 1306-1313. doi:10.1016/j.procs.2012.04.143

ON A PARAMETRIC APPROACH TO IMPROVE THE SOLUTION TO A FUZZY BI-CRITERIA TRANSPORTATION PROBLEM

Bogdana Stanojević[1], Milan Stanojević*[2]
[1]Mathematical Institute of the Serbian Academy of Sciences and Arts
[2]Faculty of Organizational Sciences, University of Belgrade
*Corresponding author, e-mail: milans@fon.bg.ac.rs

Abstract: *Keshavarz and Khorram (2011) developed a new approach to solve a fuzzy bi-criteria transportation problem, based on parametric linear programming. They stated four fuzzy goals, used the general non-linear programming model suggested by Bellman and Zadeh (1970) to aggregate them, and derived and solved two bi-level optimization problems. In this paper we propose an improved approach, such that the obtained solution reaches the highest level of satisfaction of the stated fuzzy goals. We recall an illustrative example to show the effectiveness of the new methodology.*

Keywords*: fuzzy programming, bi-criteria transportation, fuzzy interval, bi-level programming, parametric programming*

1. INTRODUCTION

The classical transportation problem is a particular case of a linear programming problem. Due to its particularity, the transportation problem is commonly solved using efficient specific algorithms, for instance simplified versions of the simplex technique. Despite the fact that the transportation problem is very old, nowadays researchers continue to pay attention to it. In the recent literature many heuristics are proposed to solve large-scale transportation problems. Also, generalized models involving fuzzy or stochastic quantities are introduced and discussed (see for instance Keshavarz and Khorram, (2011), Kundu et al. (2015), Li et al. (2014), Ojha et al. (2014), Sabbagh et al. (2015), Sterzik and Kopfer (2013), Wang et al. (2014)).

Kocken and Sivri (2015) proposed a new method to generate all optimal solutions to a fully fuzzy version of the solid transportation problem, i.e. a solid transportation problem with fuzzy costs coefficients, fuzzy supply, fuzzy demand, and fuzzy conveyance. They developed auxiliary programs to construct the parametric form of the solutions. Kaur and Kumar (2012) proposed a new algorithm for solving a special type of fuzzy transportation problems. They assumed that a decision maker is uncertain about the precise values of the transportation cost, but he is certain about the supply and demand of the product. They considered the uncertain quantities expressed by generalized trapezoidal fuzzy numbers; and developed their approach as direct extension of the classical one.

Ebrahimnejad (2014), by choosing a ranking function for the Kaur and Kumar's method to solve the fuzzy transportation problem, converted the original problem into a crisp one, thus reducing the computational complexity of the solving method. Keshavarz and Khorram (2011) solved a fuzzy bi-objective transportation problem using Bellman-Zadeh principle for aggregating the objective functions and the fuzzy goals. We address the same problem, and involve decisively the Pareto optimality concept in the solving approach.

The paper is organized as follows. In Section 2 we set up notation and terminology related to the fuzzy bi-criteria transportation problem (FBCTP). In Section 3 we briefly sketch the approach introduced by Keshavarz and Khorram (2011) to solving the FBCTP. In Section 4 we introduce an extended solving method, while in Section 5, we recall Example 2 from (Keshavarz and Khorram (2011) to illustrate the effectiveness of the new method. Section 6 provides the final remarks and suggests directions for the further research.

2. PROBLEM FORMULATION

In this section we present the general bi-criteria balanced transportation problem introduced by Keshavarz and Khorram (2011). The problem has *m* supply nodes and *n* destination nodes. Let S_i, $i=\overline{1,m}$, and D_j,

$j=\overline{1,n}$ denote the quantities that are supplied by the source node *i*, and the quantities that are required by the destination node *j*, respectively.

Let t_{ij} and p_{ij}, $i=\overline{1,m}$, $j=\overline{1,n}$ be the coefficients of the objective functions. Let x_{ij}, $i=\overline{1,m}$, $j=\overline{1,n}$ be the number of units transferred along the link (i,j). The problem is to determine a feasible way of transporting the available amount of commodities from sources to destinations, such that it minimizes the first objective function and maximizes the second one. The mathematical formulation of the general bi-criteria balanced (i.e. $\sum_{i=1}^{m} S_i = \sum_{j=1}^{n} D_j$) transportation problem is given by (1).

$$\begin{aligned}
&\min \quad \sum_{i=1}^{m}\sum_{j=1}^{n} t_{ij}x_{ij},\\
&\max \quad \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}x_{ij},\\
&\text{s.t.} \quad \sum_{j=1}^{n} x_{ij} = S_i, \qquad i=\overline{1,m},\\
&\qquad \sum_{i=1}^{m} x_{ij} = D_j, \qquad j=\overline{1,n},\\
&\qquad x_{ij} \geq 0, \qquad i=\overline{1,m}, \quad j=\overline{1,n}.
\end{aligned} \tag{1}$$

For the FBCTP in (Keshavarz and Khorram (2011)), the coefficients t_{ij} and p_{ij}, $i=\overline{1,m}$, $j=\overline{1,n}$ were defined as fuzzy intervals, while the supply and demand quantities were exact numbers. For their problem, the desired solution must be a good compromise between four fuzzy goals defined as follows:

- given the coefficients α_{ij} and β_{ij}, the membership function of the first goal, on the link (i,j), is defined by

$$\mu_{ij}^{1}(t_{ij}) = \begin{cases} 1, & t_{ij} \geq \beta_{ij}, \\ (t_{ij}-\alpha_{ij})/(\beta_{ij}-\alpha_{ij}), & \alpha_{ij} < t_{ij} < \beta_{ij}, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

- the membership function of the second goal is defined by

$$\mu^{2}(T(x)) = \begin{cases} 1, & T(x) \leq \alpha, \\ (\beta - T(x))/(\beta-\alpha), & \alpha < T(x) < \beta, \\ 0, & \text{otherwise}, \end{cases} \tag{3}$$

  where $T(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} t_{ij}x_{ij}$, $\alpha = \min_{x\in X} \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}x_{ij}$, and $\beta = \max_{x\in X} \sum_{i=1}^{m}\sum_{j=1}^{n} \beta_{ij}x_{ij}$ (note that the second goal is inverse related to the first one, through the coefficients α_{ij} and β_{ij} used in defining both of them);
- given the coefficients a_{ij} and b_{ij}, the membership function of the third goal, related to the transportation on the link (i,j), is defined by

$$\mu_{ij}^{3}(p_{ij}) = \begin{cases} 1, & p_{ij} \leq a_{ij}, \\ (b_{ij}-p_{ij})/(b_{ij}-a_{ij}), & a_{ij} < T(x) < b_{ij}, \\ 0, & \text{otherwise}, \end{cases} \tag{4}$$

- the membership function of the fourth goal is defined by

$$\mu^{4}(P(x)) = \begin{cases} 1, & P(x) \geq b, \\ (P(x)-a)/(b-a), & a < P(x) < b, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where $P(x)=\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}x_{ij}$, $a=\min_{x\in X}\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_{ij}$, and $b=\max_{x\in X}\sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij}x_{ij}$ (note that the forth goal is inverse related to the third one, through the coefficients a_{ij} and b_{ij} used in defining both of them).

3. KESHAVARZ AND KHORRAM'S SOLVING APPROACH

In this section we briefly present Keshavarz and Khorram's (2011) solving approach.

Keshavarz and Khorram (2011) chose the Bellman-Zadeh principle to solve the FBCTP. More precisely, the membership functions of the fuzzy goals of the FBCTP were aggregated using the max-min operator, as shown in (6).

$$\max_{x\in X}\left\{\min\left\{\min_{(i,j)\|x_{ij}\neq 0}\left(\mu_{ij}^{1}(t_{ij})\right),\min_{(i,j)\|x_{ij}\neq 0}\left(\mu_{ij}^{3}(p_{ij})\right),\mu^{2}(T(x)),\mu^{4}(P(x))\right\}\right\}, \tag{6}$$

where μ_{ij}^{1}, μ^{2}, μ_{ij}^{3}, and μ^{4} are defined by (2), (3), (4), and (5), respectively; and X is the feasible set of Problem (1). More precisely, for each $x\in X$, only those membership functions μ_{ij}^{1} and μ_{ij}^{3} that have indexes (i,j) such that $x_{ij}\neq 0$ were considered in (6).

Since the max-min problem (6) is generally nonlinear, Keshavarz and Khorram (2011) proposed a way to simplify it for the particular case of a FBCTP. They first formulated a bi-level problem with two parametric objective functions in the lower level (Model (23) in Keshavarz and Khorram (2011)); and announced that the optimal transportation assignment should be a Pareto optimal solution to the lower level problem.

Further, they proceeded to solve separately two bi-level problems, namely (7) and (8)

$$\begin{aligned}&\max && \lambda\\ &\text{s.t.} && \mathrm{f}(\lambda,\bar{\mathrm{x}})\le 0,\ g(\lambda,\bar{x})\le 0,\ \mathrm{f}(\lambda,\bar{\mathrm{x}})\cdot g(\lambda,\bar{x})=0\\ &&& \bar{x}\in\arg\min\left\{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\alpha_{ij}+(\beta_{ij}-\alpha_{ij})\lambda\right)x_{ij}\mid x\in X\right\}\end{aligned} \tag{7}$$

$$\begin{aligned}&\max && \lambda\\ &\text{s.t.} && \mathrm{f}(\lambda,\bar{\mathrm{x}})\le 0,\ g(\lambda,\bar{x})\le 0,\ \mathrm{f}(\lambda,\bar{\mathrm{x}})\cdot g(\lambda,\bar{x})=0\\ &&& \bar{x}\in\arg\min\left\{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(b_{ij}-(b_{ij}-a_{ij})\lambda\right)x_{ij}\mid x\in X\right\}\end{aligned} \tag{8}$$

where

$$f(\lambda,x)=\lambda-\frac{\beta-\sum_{(i,j)\|x_{ij}\neq 0}\left(\alpha_{ij}+(\beta_{ij}-\alpha_{ij})\lambda\right)x_{ij}}{\beta-\alpha},\quad g(\lambda,x)=\lambda-\frac{\sum_{(i,j)\|x_{ij}\neq 0}\left(b_{ij}-(b_{ij}-a_{ij})\lambda\right)x_{ij}-a}{b-a}. \tag{9}$$

Finally, they compared the two solutions, and chose as final solution the one with the maximal value of λ.

Note that Keshavarz and Khorram solved two bi-level problems instead of the original bi-level problem with a bi-objective function on lower level; and obtained a Pareto optimal solution to the original bi-objective problem, but not that solution that defines the transportation assignment with the highest level of satisfaction of all four fuzzy goals. An improved solution, with respect to the satisfaction level of the original fuzzy goals, may be obtained by analyzing the entire set of Pareto optimal basic solutions. The extended method presented in Section 4 supports this statement. The numerical example presented in Section 5 illustrates it.

4. THE EXTENDED SOLVING METHOD

Our extended method decisively relays on the efficient set of the parametric bi-objective optimization problem

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\alpha_{ij}+\left(\beta_{ij}-\alpha_{ij}\right)\lambda\right)x_{ij} \\
\max \quad & \sum_{i=1}^{m}\sum_{j=1}^{n}\left(b_{ij}-\left(b_{ij}-a_{ij}\right)\lambda\right)x_{ij} \\
\text{s.t.} \quad & \sum_{i=1}^{m} x_{ij} = D_j, \quad j=\overline{1,n} \\
& \sum_{j=1}^{n} x_{ij} = S_i, \quad i=\overline{1,m} \\
& x_{ij} \geq 0, \quad i=\overline{1,m}, \quad j=\overline{1,n}
\end{aligned}
\tag{10}
$$

In order to solve Problem (10), we combine a parametric linear programming analysis with a bi-objective simplex method, and find all basic feasible solutions that are efficient solutions for at least one value of $\lambda \in [0,1]$.

A parametric simplex algorithm for bi-objective linear programs can be found in Ehrgott (2005). This algorithm aggregates the two objective functions using a parameter $w \in (0,1)$, and provides a sequence of efficient basic feasible solutions to the bi-objective problem. For a bi-objective problem with coefficients depending on a parameter λ, in addition, we apply an efficiency test (for instance Benson's method, see Benson (1978)) to each basic feasible solution to determine whether it is efficient or not, and, if it is efficient, to find for which values of the parameter λ it remains efficient.

In what follows we propose a general method – Algorithm 4.1 – to find the efficient basic feasible solutions to the parametric bi-objective linear programming problem (10).

Algorithm 4.1.

- **Step 1.** Set $h \leftarrow 0$, and $\lambda^h \leftarrow 0$.
- **Step 2.** While $\lambda^h < 1$ do
    - solve Problem (10) with $\lambda \leftarrow \lambda^h$, using the parametric simplex algorithm (Ehrgott (2005)), and obtain the set of all efficient basic feasible solutions X_{eff}^h;
    - for each $x \in X_{eff}^h$, using the efficiency test (Benson (1978)) find the maximal interval $[\lambda^h, \lambda_x]$ such that *x* is efficient solution to (10) for every $\lambda \in [\lambda^h, \lambda_x]$;
    - compute $\lambda_{\min} = \min\{\lambda_x \mid x \in X_{eff}^h\}$;
    - set $h \leftarrow h+1$;
    - if $\lambda_{\min} \geq 1$, then set $\lambda^h \leftarrow 1$ otherwise $\lambda^h \leftarrow \lambda_{\min}$;

  end-while.
- **Step 3.** Set $s \leftarrow h$.

Algorithm 4.1 provides a sequence of sets X_{eff}^r, $r=\overline{0,s-1}$ together with a sequence of intervals $[\lambda^r, \lambda^{r+1}]$, $r=\overline{0,s-1}$.

Further, we construct

$$
X_{eff} = \bigcup_{r=0}^{s-1} X_{eff}^r , \tag{11}
$$

and for each $x \in X_{eff}$ we derive the set

$$
\Lambda_x = \bigcup_{r \in \{h \mid 0 \leq h < s, x \in X_{eff}^h\}} [\lambda^r, \lambda^{r+1}] \tag{12}
$$

Algorithm 4.2 solves Problem (6); thus finding the transportation assignment with the highest λ-level of satisfaction of all four original fuzzy goals. It uses Algorithm 4.1 to solve the problem on the lower level. The steps of Algorithm 4.2 are summarized below.

Algorithm 4.2.

- **Step 1.** Find all sets X_{eff}^{r}, and the corresponding intervals $\left[\lambda^{r}, \lambda^{r+1}\right], r = \overline{0, s-1}$ that solve the bi-objective problem (10) using Algorithm 4.1.
- **Step 2.** Construct X_{eff} and Λ_x, $x \in X_{eff}$ using (11) and (12), respectively.
- **Step 3.** For each $x \in X_{eff}$
  - find λ_x^f, the unique solution to $f(\lambda, x) = 0$, *f* defined by (9);
  - find λ_x^g, the unique solution to $g(\lambda, x) = 0$, *g* defined by (9);
  - compute $\lambda_x = \min\left\{\lambda_x^f, \lambda_x^g\right\}$. If $\lambda_x \notin \Lambda_x$, then remove *x* from X_{eff}^{r}.
- **Step 4.** Find $\lambda^* = \max\left\{\lambda_x \mid x \in X_{eff}\right\}$ and x^* such that $\lambda^* = \lambda_{x^*}$.

5. NUMERICAL RESULTS

In this section, we recall the numerical example from (Keshavarz and Khorram, 2011), and use it to show that our extended solving method yields a solution that is better than Keshavarz and Khorram's solution, from the point of view of the λ-level of satisfaction of all four original fuzzy goals.

Practically, we solve the FBCTP with parameters given in Table 1. The intervals that describe the fuzzy coefficients of the criteria are given as pairs $< \alpha_{ij}, \beta_{ij} > / < a_{ij}, b_{ij} >$, $i = \overline{1,3}, j = \overline{1,4}$.

Table 1: Fuzzy intervals $< \alpha_{ij}, \beta_{ij} > / < a_{ij}, b_{ij} >$ for the coefficients t_{ij} and p_{ij}, $i = \overline{1,3}$, $j = \overline{1,4}$,

	***j* = 1**	***j* = 2**	***j* = 3**	***j* = 4**	**Supply**
i = 1	<5,10>/<2,7>	<4,9>/<3,8>	<5,8>/<6,9>	<6,7>/<3,5>	10
i = 2	<3,9>/<8,11>	<3,6>/<7,9>	<6,10>/<6,12>	<6,15>/<4,10>	20
i = 3	<2,12>/<7,12>	<5,10>/<5,11>	<7,9>/<7,7>	<8,10>/<3,4>	40
Demand	20	10	15	25	

The left and right endpoints of the intervals for defining the second and the forth goals, computed by Keshavarz and Khorram, are α =335, β =815, a =310, and b =695 respectively.

Following the extended methodology, we generated all basic feasible solutions $\overline{x^k}$, $k = \overline{1,6}$, that are efficient solutions to the parametric Problem (10) for at least one value of λ, hence handling the compromise among the second and forth goals. For each solution $\overline{x^k}$ we computed the λ-level of the satisfaction of the first and third goals, (λ_f^k and λ_g^k, respectively), and derived $\lambda_{\min}^k$ as $\min\left\{\lambda_f^k, \lambda_g^k\right\}$ (see Table 2). As final solution, we chose the solution $\overline{x^k}$ with the greatest value of $\lambda_{\min}^k$.

Our optimal solution is $\overline{x^3}$ =(0,0,0,10,0,5,0,15,20,5,15,0) with the corresponding $\lambda_{\min}^k$ =0.5275591.

Table 2: The efficient basic feasible solutions

k	$\overline{x^k}$	λ_f^k	λ_g^k	$\lambda_{\min}^k$	$T\left(\lambda_{\min}^k, \overline{x^k}\right)$	$P\left(\lambda_{\min}^k, \overline{x^k}\right)$
1	(0,0,5,5,0,0,0,20,20,10,10,0)	0.5052632	0.5507246	0.5052632	572.47368	535.8947
2	(0,0,0,10,0,0,5,15,20,10,10,0)	0.5191257	0.5401460	0.5191257	565.81967	524.2623
3	(0,0,0,10,0,5,0,15,20,5,15,0)	0.5363128	0.5275591	0.5275591	553.93701	513.1102
4	(0,0,0,10,0,10,0,10,20,0,15,5)	0.5647059	0.5	0.5	520	502.5
5	(0,0,0,10,0,10,10,0,20,0,5,15)	0.5875	0.475	0.475	497	492.875
6	(0,0,0,10,10,10,0,0,10,0,15,15)	0.6081081	0.4326923	0.4326923	477.5	476.5865

Optimizing the objective functions distinctly, Keshavarz and Khorram obtained only two Pareto optimal solutions to Problem (10), the marginal ones, $\overline{x^1}$ and $\overline{x^2}$ in Table 2. For each of these two solutions, they computed the λ-level of the satisfaction of the original fuzzy goals, and chose the solution with the greatest value of λ. Their maximal λ =0.5052632 was obtained for $\overline{x^1}$ =(0,0,5,5,0,0,0,20,20,10,10,0).

The values of the objective functions for each Pareto optimal solution $\overline{x^k}$, $k=\overline{1,6}$, are also given in Table 2. Comparing our solution to the solution of Keshavarz and Khorram we found $T\left(\lambda^3_{\min}, \overline{x^3}\right) < T\left(\lambda^1_{\min}, \overline{x^1}\right)$ and $P\left(\lambda^3_{\min}, \overline{x^3}\right) < P\left(\lambda^1_{\min}, \overline{x^1}\right)$. The value of the first objective function at our solution is better (smaller) than the old one, while the value of the second objective at our solution is worse (smaller) than the old one. Both transportation assignments $\overline{x^1}$ and $\overline{x^3}$ are efficient solutions from the point of view of the original objective functions, thus they cannot be compared without looking to an additional attribute. Since the additional attribute for the FBCTP is the λ-level of the satisfaction of the original fuzzy goals, we conclude that our solution $\overline{x^3}$ to the transportation problem is better than the one proposed by Keshavarz and Khorram (2011).

6. CONCLUSIONS

In this paper we addressed the FBCTP introduced by Keshavarz and Khorram (2011). Keshavarz and Khorram solved the FBCTP via Bellman-Zadeh principle and desired to obtain the highest level of satisfaction of four original fuzzy goals.

In order to solve the FBCTP, we involved the Pareto optimality concept in our solving approach; and developed an algorithm that chooses, among the basic feasible Pareto optimal solutions to a parametric bi-criteria transportation problem, the solution with the highest level of satisfaction of all four fuzzy goals. Using the same numerical example as in Keshavarz and Khorram (2011), and solving it using the new algorithm, we obtained improved results, thus illustrating the effectiveness of our methodology.

The FBCTP may be solved using the new concepts of fuzzy solution to fuzzy optimization problems. In our future work we want to research the possibility to find the values of a membership function related to a certain fuzzy set solution to the FBCTP.

ACKNOWLEDGEMENTS

This research was partially supported by the Ministry of Education and Science, Republic of Serbia, Project numbers TR36006 and TR32013.

REFERENCES

Bellman, R., & Zadeh, L. (1970). Decision-making in a fuzzy environment, *Management Science, 17B,* 141-164.

Benson, H., (1978). Existence of efficient solutions for vector maximization problems, *Journal of Optimization Theory and Applications 26*(4), 569- 580.

Ebrahimnejad, A. (2014). A simplified new approach for solving fuzzy transportation problems with generalized trapezoidal fuzzy numbers, *Applied Soft Computing*, *19*, 171-176. doi:10.1016/j.asoc.2014.01.041.

Ehrgott, M., (2005). *Multicriteria Optimization*, Springer Verlag.

Kaur, A., & Kumar, A. (2012). A new approach for solving fuzzy transportation problems using generalized trapezoidal fuzzy numbers, *Applied Soft Computing 12*(3), 1201-1213. doi:10.1016/j.asoc.2011.10.014.

Keshavarz, E., & Khorram, E. (2011). A fuzzy bi-criteria transportation problem, *Computers & Industrial Engineering*, *61*(4), 947- 957. doi:10.1016/j.cie.2011.06.007.

Kocken, H.G., & Sivri, M. (2015). A simple parametric method to generate all optimal solutions of fuzzy solid transportation problem, *Applied Mathematical Modelling* (in press). doi:10.1016/j.apm.2015.10.053.

Kundu, P., Kar, S., & Maiti, M. (2015). Multi-item solid transportation problem with type-2 fuzzy parameters, Applied Soft Computing *31*, 61- 80. doi:10.1016/j.asoc.2015.02.007.

Li, F., Jin, C., & Wang, L. (2014). Quasi-linear stochastic programming model based on expectation and variance and its application in transportation problem, *Applied Mathematical Modelling, 38*(7-8), 1919- 1928. doi:10.1016/j.apm.2013.10.024.

Ojha, A., Das, B., Mondal, S. K., & Maiti, M. (2014). A transportation problem with fuzzy-stochastic cost, *Applied Mathematical Modelling, 38*(4), 1464- 1481. doi:10.1016/j.apm.2013.08.022.

Sabbagh, M. S., Ghafari, H., & Mousavi, S. R. (2015). A new hybrid algorithm for the balanced transportation problem, *Computers & Industrial Engineering 82,* 115- 126. doi:10.1016/j.cie.2015.01.018.

Sterzik, S., & Kopfer, H. (2013). A tabu search heuristic for the inland container transportation problem, *Computers & Operations Research, 40*(4), 953- 962. doi:10.1016/j.cor.2012.11.015.

Wang, G., Jia, N., Ma, S., & Qi, H. (2014). A rank-dependent bi-criterion equilibrium model for stochastic transportation environment, *European Journal of Operational Research, 235*(3), 511- 529. doi:10.1016/j.ejor.2014.01.023.

OPTIMAL ADJUSTMENTS OF ORDER QUANTITIES DUE TO PACKAGING AND TRUCK CONSTRAINTS

Aleksa Đorđević[1], Ana Belić[2], Dragana Makajić Nikolić*[3]
[1]Nelt Co. d.o.o. Dobanovci
[2]Alti d.o.o. Čačak
[3]University of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail:gis@fon.bg.ac.rs

Abstract *This paper present one problem that appear in supply chain from a perspective of a distributor in FMCG (fast–moving consumer goods), where goods are being ordered and sold every day in big quantities. Using predictive models for demand supply planning, planner calculate optimal quantities of products that should be delivered to retailers. However, these quantities rarely match with packaging units and truck weight limits and therefore must be adjusted. For solving the observed problem, mathematical model was formulated where we discuss two different ways of its application: obtaining of optimal weights and examination of the influence of packaging sizes on optimal vehicles loading. The optimization results show to decision maker how to avoid the decrease of DOC (days of coverage) and on the other hand to avoid increase of DOC for poorly performing products especially if they have high costs of stock-keeping. Second scenario shows that both proposed types of packaging are better in all performances than the results for pallet as packaging unit.*

Keywords*: supply chain*, *distributor*, *order quantities*, *mathematical model*, *optimization*

1. INTRODUCTION

A modern company must constantly respond to the huge challenges of the market and improve their business. To ensure optimal functioning, company has to react preventively and plan each of its activities. Distribution systems, which will be discussed in this paper as one of the most complex systems require constant analysis of supply chains. The challenge of coordinating operations across all facets of a business has become known as supply chain management (SCM) (Stank et al. 2001). Supply chain management can achieve significant savings by reducing the cost of inventory and planning, affect other business processes. SCM (Supply chain management) can be defined as integrated process wherein a number of various business entities (i.e., suppliers, manufacturers, distributors, and retailers) work together in an effort to: (1) acquire raw materials, (2) convert these raw materials into specified final products, and (3) deliver these final products to customers (Beamon, 1998). Another definition notes a supply chain is the network of organizations that are involved, through upstream and downstream linkages, in the different processes and activities that produce value in the form of products and services delivered to the ultimate consumer (Christopher 1992). It involves two basic processes tightly integrated with each other: the production planning and inventory control process, which deals with manufacturing, storage, and their interfaces, and the distribution and logistics process, which determines how products are retrieved and transported from the warehouse to retailers (Tsiakis et al. 2001). Supply chain management is not just another name for logistics. It includes elements that are not typically included in a definition of logistics, such as information systems integration and coordination of planning and control activities (Cooper et al. 1997).

The term supply chain management has risen to prominence over the past ten years (Cooper et al. 1997) and has grown in importance, although the approach was introduced in early 1980 (Oliver and Webber 1982). This concept received considerable attention in the popular business press and in some academic literatures (Cooper et al. 1997).

Uncertainty plays an important role in the modern supply chain systems (Xu et al. 2003). Underestimating uncertainty and its impact can lead to planning decisions that might jeopardize a company. One of the key sources of uncertainty in any production-distribution system is the product demand. (Gupta et al. 2003). Having the competitive product and right supply chain is not enough thus supply chain has to be right for the individual customer as well (Hoover et al. 2001). Demand chain design is based on a thorough market understanding and has to be managed in such a way as to effectively meet differing customer needs (Treville

et al. 2004). Demand integration includes increased access to demand information throughout the supply chain to permit rapid and efficient delivery, coordinated planning, and improved logistics communication. (Treville et al. 2004). In a broader sense, DCM (Demand Chain Management) can be defined as "a set of practices aimed at managing and co-ordinating the whole demand chain, starting from the end customer and working backward to raw material supplier (Selen et al. 2002).

Demand Supply chain problem became attractive in recently years. Authors analyzed a one-supplier–one-retailer model (Weng, 1995), one-supplier–two-retailer case (Zheng et al. 2001), generalized the model to one-supplier–many retailer case. (Zheng et al. 2001) and studied the situation in which the price–demand relationship varied from period to period, but was known in each period. (Zhao et al. 2002).

This paper deals with one particular problem recognized at FMCG Company. The observed problem appears in operational planning at supply chains distributors. Based on demand data and information about retailer's inventory, planner should decide which quantities of products to deliver to retailer. The goal is to provide desired days of stock coverage. However, such quantities often do not correspond to the packaging units of products as well as the total products quantities often do not correspond to the desired capacities utilization of provided vehicles. In this paper, we formulate this operational decision as an optimization problem. This particular problem directly arise from real life and, according to our knowledge based on available literature, it is not previously formulated in that manner.

The paper consists of four sections. Section 2 is related to detailed description of the observed problem and its mathematical model formulations. Section 3 gives numerical results of a group of scenarios on real-life example of distributer of fast-moving consumer goods. Conclusions and directions of the further research are described in Section 4.

2. PROBLEM DESCRIPTION AND MATHEMATICAL MODEL FORMULATION

From the perspective of a distributor in FMCG (Fast-moving consumer goods), where goods are being ordered and sold every day in big quantities, demand and supply planner encounters many problems in day-to-day business. One of the issues which we are trying to improve using this model is the problem of fulfillment of a truck when ordering from the producer. As the biggest part of the transportation costs are fixed, therefore the distributor and the producer agreed on ordering at least 97% (q) of the full truck (Q).

However, using predictive models for demand supply planning, optimal ordering quantities rarely match with the truck weight limits (volume is not the issue in this problem as all the products are of a high density). Therefore, a planner has to modify optimal quantities he got using predictive models and adapt them to fill each truck with at least 97% of its maximum weight capacity.

Every product has defined its perfect inventory period (ip) which serves as a target when modeling ordering quantities, hence every change in suggested optimal ordering quantity will modify the actual days of coverage (DOC) which represents the number of days for which our goods on stock (iap) can fulfill the needs of forecasted sales (sf). The ideal quantity of each product which is used as a target during every order planning represents inventory theoretical (it) which is calculated as a sales forecast (sf) for the targeted inventory period (ip). Having in mind that between two orders (daily, weekly, monthly) some sales have happened which is commonly followed by uncertainty, a planner uses a formula where from inventory theoretical subtracting actual inventory at the moment of ordering and adding sales forecast for the period until the next moment of ordering to get the optimal ordering quantities (tr). Later in the model explanation, these optimal ordering quantities will be called the estimated weight of the product. Additionally, those optimal ordering quantities have to fulfill the packaging unit condition which means that the ordered quantity of every particular product has to be rounded to nearest integer of packaging units.

Following changes of optimal quantities to satisfy the packaging units and the truck fulfillment conditions can have two possible negative outcomes; if the quantity is decreased from optimal, there is a chance there will be not enough goods to satisfy the sales, hence being out of stock and miss the sales opportunity; if the quantity is increased, negative outcome is increasing costs of stock-keeping and having non-optimal cash flow.

First negative outcome is often considered more undesirable and that is why we are going to give a bigger penalty in our model if the order quantities are to be decreased from optimal (pd). Penalty for increasing the ordered quantity is defined as (pi).

The notation used in the mathematical model of the formulated problem is as follows.

N – set of products,

Parameters:

Q – capacity of the vehicle,

q - minimum required vehicles' charge status (%),

tr_j - the estimated weight of the j -th product to load into vehicles, $j \in N$,

p_j - packaging unit of the j -th product, $j \in N$,

iap_j - inventory actual

sf_j - sales forecast of the j -th product (weight), $j \in N$,

it_j - inventory theoretical of the j -th product (weight), $j \in N$,

ip_j - preferred DOC, inventory coverage of the j -th product (days), $j \in N$,

pi - penalty for increasing the preferred inventory coverage,

pd - penalty for decreasing the preferred inventory coverage.

Variables:

x_j - weight of the j -th product that should be to loaded into vehicles, $j \in N$,

y_j - number of packages of the j -th product that should be to loaded into vehicles, $j \in N$,

z - number of vehicles needed for transportation,

di_j - increasing the preferred inventory coverage for the j -th product (days), $j \in N$,

di_j - negative deviation variables for the tr_j , increasing of the estimated weight of the j -th product to load into vehicles, $j \in N$,

dd_j - positive deviation variables for the tr_j , decreasing of the estimated weight of the j -th product to load into vehicles, $j \in N$,

$doci_j$ - negative deviation variables for the ip_j , increasing of the preferred DOC for the j -th product (days), $j \in N$,

$docd_j$ - positive deviation variables for the ip_j , decreasing of the preferred DOC for the j -th product (days), $j \in N$,

Using the introduced notation, the following mathematical model was formulated:

$$(\min)\ \ pi \cdot \sum_{j \in N} doci_j + pd \cdot \sum_{j \in N} docd_j \tag{1}$$

s.t.

$$x_j - di_j + dd_j = tr_j,\ j \in N \tag{2}$$

$$\sum_{j \in N} x_j \geq Q \cdot q \cdot z \tag{3}$$

$$\sum_{j \in N} x_j \leq Q \cdot z \tag{4}$$

$$\frac{ip_j \cdot (iap_j - sf_j + x_j)}{it_j} - doci_j + docd_j = ip_j,\ j \in N \tag{5}$$

$$x_j - p_j \cdot y_j = 0,\ j \in N \tag{6}$$

$$x_j \geq 0, di_j \geq 0, dd_j \geq 0, doci_j \geq 0, docd_j \geq 0, y_j \in \square^+, j \in N$$

Objective function (1) minimize total penalized increasing and decreasing the prefered DOC. The constraint (2) is related to the estimated weights of the products. These weights can not always be satisfied because of the products packaging units and vehicles capacities. Therefore, deviations are allowed but not penalized. The constraints (3) and (4) refer to the minimum and maximum vehicle capacity that must be reached and must not be exceeded. The constraint (5) is related to the preferred DOC which cannot always be reached for the same reasons stated in the constraint (2). Since the main goal of the transport planning is to achieve preferred DOCs, negative and positive deviations are penalized in the objective function. The constraint (6) conects two types of variables (weight and the number of packages of the j -th product) and ensure that the optimal weight of the product corresponds to the integer number of packages.

Note: Variable z that appears in constraints (3) and (4) can be defined as parameter if desired number of vehicle is known in advance.

3. NUMERICAL RESULTS

In order to evaluate the proposed mathematical model (1-6) we discuss two different ways of its application. First one refers to obtaining of optimal weights i.e. number of packages of products that should be loaded into vehicles in order to fulfill requested vehicle capacity with minimal disturbance of required days of coverare. The second application is related to the examination of the influence of packaging sizes on optimal vehicles loading.

Both applications will be ilustrated on a range of 234 products in the field of cosmetics and household chemicals. Total estimated weight of the products is 123947.621kg. Since the capacity of one vehicle is 23500kg this weight requires 5.2744 vehicles. It is obvious that the estimated weights of the products should be adjusted to fulfill whole number of vehicles. In addition, estimated weights of the products in most cases do not correspond whole numbers of packages. Therefore, this type of adjustment is also needed.

In determining the optimal weights of products that should be loaded into vehicles, the following four scenarios was tested:

1. Decreasing the preferred DOC is penalized five times more than its increasing; z is a variable.
2. Penalties for decreasing and increasing are the same; z is a variable.
3. Decreasing the preferred DOC is penalized five times more than it's increasing; z is a parameter equal to 5.
4. Penalties for decreasing and increasing are the same; z is a parameter equal to 5.

The model (1-6) has been implemented and solved using GNU Linear Programming Kit (GLPK, 2016). The optimal solutions for all four scenarios are shown in Table 1. In addition to total penalized increasing and decreasing (Objective), number of vehicles and total weight of optimal products quantities, Table 1 gives maximal inreasing and decreasing of DOC for each solution. The right part of the table shows the number of products for which increasing of DOC is between 2 and 5 days i.e. up to one week (column 1w), between 5 and 10 days i.e. 1 to 2 weeks (column 2w) and more than 10 days i.e. 2 weeks (column >2w). Analogue values are given in columns named DOC decreasing.

Table 1. Optimal solutions

				max DOC		DOC increasing			DOC decreasing		
Scenario	Objective	Number of vehicles	Total weight	inc	dec	1 w	2 w	> 2 w	1 w	2 w	> 2 w
1	188.581	6	136779	11.069	-1.048	37	5	1	1	0	0
2	130.537	6	136771	9.091	-8.660	19	3	0	10	1	0
3	222.168	5	117496	11.069	-5.372	34	5	1	0	1	0
4	138.381	5	117491	9.091	-8.660	17	3	0	12	1	0

Although the lower value of the objective function is achieved in the scenario 2, it requires one vehicle more than scenarios 3 and 4. In practice, the decision-maker would decide to fill 5 rather than 6 vehicles considering that the total estimated weight of products required 5.27 vehicles. Therefore, only scenarios 3 and 4 should be considered. Values of objective functions in these two scenarios are incomparable due to

the different penalties for DOC decreasing. In addition, none of these two scenarios dominating in all deviations of DOCs. Scenario 3 gives lower maximal DOC decreasing but higher maximal DOC increasing and more products with DOC deviation than scenario 4. On the other hand, scenario 4 gives 12 products with DOC decreasing more than scenario 3. Therefore, the final decision is on the decision maker. However, the decision maker should always have in mind prioritized products with high importance to avoid the decrease of DOC and on the other hand to avoid increase of DOC for poorly performing products especially if they have high costs of stock-keeping.

The packaging unit for about the half of the product portfolio in the previous example is the pallet. This is the main reason for the inability to obtain solutions with 5 vehicles in the scenarios 1 and 2, although the required number of vehicles based on estimated weight of the products is closer to 5 than 6. Hence, we have investigated how smaller packaging units influence on optimal weights of products and consequently optimal number of vehicles. Table 2 gives optimal solution for four previously defined scenarios when packaging units are layers (P1) and cases (P2) instead of pallets.

Table 2. Optimal solutions with smaller packaging units

Packaging unit	Scenario	Objective	Number of vehicles	Total weight	max DOC		DOC increasing			DOC decreasing		
					inc	dec	1 w	2 w	> 2 w	1 w	2 w	> 2 w
P1	1	106.541	6	136773	6.258	-0.409	19	2	0	0	0	0
	2	69.867	5	117494	4.080	-3.148	8	0	0	7	0	0
	3	126.648	5	117500	6.258	-3.787	18	2	0	1	0	0
	4	69.867	5	117494	4.080	-3.148	8	0	0	7	0	0
P2	1	92.8418	6	136772	6.258	-0.409	16	3	0	0	0	0
	2	60.1435	5	117500	4.080	-3.361	8	0	0	5	0	0
	3	107.128	5	117499	6.258	-3.558	16	2	0	1	0	0
	4	60.1435	5	117500	4.080	-3.361	8	0	0	5	0	0

The results obtained for both types of packaging are better in all performances than the results for pallet as packaging unit. However, if high penalty is given to the DOC decreasing (scenario 1), optimal solution with 5 vehicles cannot be obtained, regardless packaging unit. Finally, using layers as packaging units is more appropriate than cases due to reception issues in a warehouse where cases of the same product could be dispersed among different pallets and with layers that's not of concern.

4. CONCLUSION

The subject of this paper is one optimization problem that is recognized in real life FMCG distribution company. Planners are faced ones a week with the problem of determining the quantities of products that should be delivered to retailers. Based on demand and retailer's inventory data, planners calculate the quantities of products that provide the desired days of stock coverage. Since these quantities usualy do not correspond to packaging units and vehicle capacities, they should be adjusted. In this paper, we have observed that problem as an optimization problem and formulated corresponding mixed integer mathematical model. The proposed models can be used at two level. The first use is for obtaining the optimal weights of products that should be loaded into vehicles. The second use is related to the examination how packaging sizes influence on optimal vehicles loading. Both application were illustrated on FMCG company. Obtained results show that, as expected, package units smaller than pallets provide better adjustment. However, the final decision about packaging units should be made by responsible decision maker, which is enabled by the proposed approach.

The proposed mathematical model includes determination of the number of vehicles assuming they are of the same capacity. Since, the capacities of vehiclel can differ, observed problem could be modified by introducing the assignment of products to individual vehicles. The inclusion of this feature into the problem would require the modification of the mathematical model and, probably, development of appropriate heuristic.

REFERENCES

Beamon, B. M. (1998). Supply chain design and analysis:: Models and methods. *International journal of production economics*, *55*(3), 281-294.

Chen, F., Federgruen, A., & Zheng, Y. S. (2001). Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Management Science*, *47*(5), 693-708.

Christopher, Martin L. (1992), Logistics and Supply Chain Management, London: Pitman Publishing.

Cooper, M. C., Ellram, L. M., Gardner, J. T., & Hanks, A. M. (1997). Meshing multiple alliances. *journal of Business Logistics*, *18*(1), 67.

Cooper, M. C., Lambert, D. M., & Pagh, J. D. (1997). Supply chain management: more than a new name for logistics. *The international journal of logistics management*, *8*(1), 1-14.

De Treville, S., Shapiro, R. D., & Hameri, A. P. (2004). From supply chain to demand chain: the role of lead time reduction in improving demand chain performance. *Journal of Operations Management*, *21*(6), 613-627.

GLPK (2016), available at: http://www.gnu.org/software/glpk/ (accessed 15 March 2016).

Gupta, A., & Maranas, C. D. (2003). Managing demand uncertainty in supply chain planning. *Computers & Chemical Engineering*, *27*(8), 1219-1227.

Hoover, W. E., Eloranta, E., Holmstöm, J., & Huttunen, K. (2001). Managing the Demand Chain.

Jüttner, U., Christopher, M., & Baker, S. (2007). Demand chain management-integrating marketing and supply chain management. *Industrial marketing management*, *36*(3), 377-392.

Lambert, D. M., & Cooper, M. C. (2000). Issues in supply chain management.*Industrial marketing management*, *29*(1), 65-83.

Oliver, R. K., & Webber, M. D. (1982). Supply-chain management: logistics catches up with strategy. *Outlook*, *5*(1), 42-47.

Selen, W., & Soliman, F. (2002). Operations in today's demand chain management framework. *Journal of Operations Management*, *20*(6), 667-673.

Stank, T. P., Keller, S. B., & Daugherty, P. J. (2001). Supply chain collaboration and logistical service performance. *Journal of Business logistics*, *22*(1), 29-48.

Tsiakis, P., Shah, N., & Pantelides, C. C. (2001). Design of multi-echelon supply chain networks under demand uncertainty. *Industrial & Engineering Chemistry Research*, *40*(16), 3585-3604.

Weng, Z. K. (1995). Channel coordination and quantity discounts. *Management science*, *41*(9), 1509-1522.

Weng, Z. K., & Zeng, A. Z. (2001). The role of quantity discounts in the presence of heterogeneous buyers. *Annals of Operations Research*, *107*(1-4), 369-383.

Xu, M., Qi, X., Yu, G., Zhang, H., & Gao, C. (2003). The demand disruption management problem for a supply chain system with nonlinear demand functions. *Journal of Systems Science and Systems Engineering*, *12*(1), 82-97.

Zhao, W., & Wang, Y. (2002). Coordination of joint pricing-production decisions in a supply chain. *IIE transactions*, *34*(8), 701-715.

OPTIMAL HETEROGENEOUS STAFF SCHEDULING IN CASE OF PLANNING A SPECIAL EVENT

Maja Hadžiahmetović*, Minja Marinović,
Universtiy of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: hadziahmetovic.maja@fon.bg.ac.rs

Abstract: *In this paper a scheduling staff problem with multiple work skills in case of planning a special event is presented. This problem can be quite challenging because each special event has different conditions regarding days of preparation, activities which need to be done, the number of staff required and their qualification. However, not all staff members, whether they are a part of organizational management, interns or volunteers possess all of the required skills, and they cannot be available at all times which can bring to the scheduling problem. This paper deals with such problem, introducing an original mathematical model used on a real-life example of scheduling staff for the Knight Fest held each year in Belgrade, Serbia. The problem is formulated and solved using goal programming model.*

Keywords: *staff scheduling, goal programming, optimization, heterogeneous staff, multi-skill workforce*

1. INTRODUCTION

Scheduling staff represents a problem where organization assigns working schedules to its employees in order to meet the demand for the goods or services they provide. (Ernst et al. 2004) These schedules have a purpose of maximizing the employees' productivity or their preferences towards shifts, days off or holidays or minimize the costs of their assignment or even both. However, it is rather complicated to find the optimal solution to this problem since there are many conditions to meet in the organization itself as well outside of it. The cause of this complexity is that there are many government rules regarding staff deployment which vary from one country to another.

Staff scheduling as an optimization problem has become an extremely important and popular area in the field of operations research. The reason for that is that there are always new problems which occur in different branches of business where each requires special approaches, methods and applications. This is why there are many ways of solving scheduling problems, especially because there is no universal solution because each organization no matter what its' core businesses, has unique ways of functioning, deployment politics, or government regulations which need to be taken into account. However, what all of these organizations have in common is to use their staff in the most efficient way. Personnel scheduling is applicable at most for scheduling medical staff, see (Hall, 2012), (Cheang et al. 2003), (Azaiez and Al Sharif 2005), call center operators, see Avramidis et al. (2010), Stefanov et al. (2014) and airline companies (Brusco et al. 1995), (Barnhart et al. 2003) .

Assignment staff problem started to occur in the fifties of the XX century. The interest of science for this area and the subject continues to this day, even though the problem had some significant changes from the Dantzig (1954) and Edie`s (1954) problem that has been related to traffic delays at toll booths. One of the first classification methods for assigning staff was presented by Baker (1976), in which he presented three types of problems of scheduling staff. The first type covers scheduling shifts “shift scheduling”, the second type is related to “days-off scheduling” while the third type presents the combination of first two types (tour scheduling). Alfares (2004) proposes ten categories for staff allocation: (1) manual solution, (2) integer programming, (3) implicit modeling, (4) decomposition, (5) goal programming, (6) working set generation, (7) LP-based solution, (8) construction and improvement, (9) metaheuristics, (10) other methods. When it comes to literature review on this subject, comprehensive literature review of classical studies on this problem is given in Ernst et al. (2004) while the extensive amount of covered papers on this subject is presented by Van Den Bergh et al. (2012) Present scheduling staff problem can be divided according to belonging areas, with the largest number of published papers focused on allocating staff employed in airlines, medical personnel and call center’s operators as a new form of staff allocation. A similar problem of scheduling heterogeneous staff with part-time employees with limited availability is presented by Hojati and Patil (2010). Scheduling employees method that combines two-step procedure in one which iterates between

a schedule evaluator and a schedule generator is presented by Ingolfsson et al. (2011). Mohan (2008) suggests another approach for maximizing employees` satisfaction with certain restrictions considering their availability, preferences for the shifts and seniority level.

This paper presents a heterogeneous staff scheduling problem in case of planning a special event. This is a special form of staff deployment, since it is limited in time, requires a great number of staff with different skills, the tasks which need to performed have multiple level of difficulty. Consequently, it is possible to employ volunteers or interns since the days of preparation usually last much more than the event itself and they are carried out on various locations. The suggested solution would provide an optimal schedule of all staff depending on their skills and availability, respecting conditions concerning a minimal and a maximum number of staff allowed in order to minimize the costs of their assignment for each day required for the organization and preparation of the event itself. The problem was formulated using 0-1 goal programming model and solved exactly. Therefore, it would be needed to change only the input data about staff availability and to get a new optimal solution for each day necessary until the event is over.

The paper is organized as follows. After the introductory part, Section 2 describes a problem of scheduling employees in case of special events. Section 3 describes model formulation and presents an original mathematical model for the solving problem described in Section 2. Numerical results are presented in Section 4 and discussed in the real-life example of staff scheduling for the manifestation named Knight Fest. Conclusions and suggestions regarding future work are given in Section 5.

2. PROBLEM DESCRIPTION

The paper discusses a problem of determining the optimal schedule of staff members with different hierarchy levels when planning a special event. Each special event has different organizational goals and a specific way of deployment of its employees. This means that there is no common solution when planning events such as these since each event differs from other significantly. For that reason, it can be rather challenging to schedule employees where there is not a common pattern of work to follow or shifts to change on a daily or weekly level. These types of scheduling often require a large number of staff and recourses for the sake of everything to look the most authentic and go as planned. That is why the organization has to be done days, usually months in advance which leads to the fact that plenty of staff needs to be hired in order to everything go well. The staff is usually divided into groups, depending on their wages, position, availability and the scope of work they do. Since there is a vast number of activities which have to be done, it is expected that not all staff member are qualified to do them all. Especially because most special events demand to hire a large number of volunteers and interns to do less complex jobs, which do not require special skills, for a small or no fee included. This is why staff member with experience are eligible to work on all activities, where interns and volunteers do not have to be familiar with all of the activities which need to be done.

Besides the diversity in staff members, there are many activities which contribute to the organization itself, before the main event occurs. These activities, among many, may include marketing, promotion, organizational and IT skills and they have to be held in various locations depending on the prepared timeline. At each location, the activities are carried out in several terms during the day.

For each of these terms, there is a minimal number of available staff required, where it is necessary that, at least, one of the staff with experience (neither a volunteer nor an intern) is present during each term and location. Even though the volunteers do not receive a salary, a certain amount of money has to be set aside for their meals, transportation etc. Interns, however, do receive a salary, but it is significantly lower than managers' due to the fact they do not have as much experience. Since there is a difference in working hours for each of the group of staff, their availability during the day may vary. Depending on their function for each group the assigned number of given terms is different, including the fact that not all of them are available every day and every term. For example, the volunteers may be assigned to not more than two terms and locations during one day, interns have a maximum number of three terms and locations, where experienced staff should work four terms during one day (and four locations). The break period usually lasts one hour between terms and it could be used for rest, meals and switching locations if they were assigned so.

Taking everything into account, it is necessary to create an optimal schedule and assign all activities to staff members for each day, respecting the given constraints in order to minimize costs needed to pay staff members.

3. MODEL FORMULATION

In the formulation of a mathematical model for optimal staff scheduling using goal programming, the following notation has been used:

Sets:

I – set of all staff that need to be scheduled, $i \in I$,
I_v – set of volunteers, $I_v \subset I$
I_o – set of organizational managers, $I_o \subset I$,
I_p – set of interns, $I_p \subset I$, $I_v \cup I_o \cup I_p = I$,
J – set of activities, $j \in J$,
K – set of locations, $k \in K$
T – set of terms during the day, $t \in T$

The variable X can be described as follows:

$$x_{ijkt} = \begin{cases} 1 & \text{if staff member } i \text{ is assigned to the activity } j \text{ on the location } k \text{ in the term } t \\ 0 & \text{otherwise} \end{cases}$$

The matrix A describes whether the certain staff member is qualified to do the given activity. The elements of the matrix A are:

$$a_{ij} = \begin{cases} 1 & \text{if staff member } i \text{ is qualified to do the activity } j \\ 0 & \text{otherwise} \end{cases}$$

. .

The elements of the matrix A could also be real numbers from 0 to 1, giving the fact that perhaps, some staff members are more or less familiar with a certain activity. In this case, number one would mean that the staff member *i* is fully familiar with the activity *j,* and zero would mean that he is not familiar at all.

As mentioned above, not all members are available to work every term of the day on each location. The matrix B shows if a staff member *i* is available to work on the given term and location. The elements of matrix B are:

$$b_{ikt} = \begin{cases} 1 & \text{if staff member } i \text{ is available on the location } k \text{ in the term } t \\ 0 & \text{otherwise} \end{cases}$$

Mathematical model of staff scheduling is described as follows:

$$\min \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \sum_{t \in T} c_i \cdot x_{ijkt} \tag{1}$$

s.t.

$$\sum_{i \in I} \sum_{j \in J} x_{ijkt} \le p_{kt},\ k \in K,\ t \in T \tag{2}$$

$$\sum_{i \in I} \sum_{j \in J} x_{ijkt} \ge d_{kt},\ k \in K,\ t \in T \tag{3}$$

$$\sum_{i \in I_o} x_{ijkt} \ge 1,\ j \in J,\ k \in K,\ t \in T \tag{4}$$

$$\sum_{i \in I_p}' x_{ijkt} \ge mip_{jkt},\ j \in J,\ k \in K,\ t \in T \tag{5}$$

$$\sum_{i \in I_v}' x_{ijkt} \ge miv_{jkt},\ j \in J,\ k \in K,\ t \in T \tag{6}$$

$$\sum_{j \in J} x_{ijkt} \leq b_{ikt},\ i \in I, k \in K, t \in T \quad (7)$$

$$x_{ijkt} \leq a_{ij},\ i \in I, j \in J, k \in K, t \in T \quad (8)$$

$$\sum_{j \in J} \sum_{k \in K} \sum_{t \in T} x_{ijkt} \leq mv,\ i \in I_v \quad (9)$$

$$\sum_{j \in J} \sum_{k \in K} \sum_{t \in T} x_{ijkt} \leq mp,\ i \in I_p \quad (10)$$

$$\sum_{j \in J} \sum_{k \in K} \sum_{t \in T} x_{ijkt} \leq mo,\ i \in I_o \quad (11)$$

$$\sum_{j \in J} \sum_{k \in K} x_{ijkt} \leq 1,\ i \in I,\ t \in T \quad (12)$$

The goal function (1) minimizes the costs which need to be paid for each worker. The parameter c_i represents the cost for the staff member *i*, regardless of the skill level they possess. The majority of these costs include wages, but these costs include expenses such as food or transportation since volunteers do not receive wages for the activities they do. Constraints (2) and (3) refer to the maximum and minimal number of staff on each location and term of the day, where parameters p_{kt} and d_{kt} represent these numbers respectively. Constraints (4)-(6) represent a minimal number for each staff group for every activity, term and on each location. Parameters miv_{jkt}, mip_{jkt}, mio_{jkt} are integer numbers for each group. The seventh constraint assigns location and term for each available staff member defined by matrix B. Constraint (8) provides that only qualified members for certain activities are assigned to those activities. Constraints (9)-(11) describe a maximum number of locations which need to be assigned. For volunteers that number is two, for interns three, and for organizational managers, the maximum number of locations is four. Finally, constraint (12) provides that it is impossible to assign the same term for different location when it comes to the staff member *i.*

4. NUMERICAL RESULTS

Each year a traditional manifestation, the Knight Fest, is held in Belgrade, Serbia. The Knight Fest is a multimedia knight spectacle which is held on a great children holiday – Lazarus' Saturday/Palm Sunday on Belgrade fortress ("Knight Fest", 2016). The event falls into one of the most interesting events of this type, and it is a major tourist attraction.

The main event, the festival itself, is held on Belgrade fortress while the activities concerning preparation and promotion of the festival are held in various locations in Belgrade such as city center, shopping malls and other manifestations and events. Activities consist of festival organization, marketing activities, human resources activities (HR), knowing knight skills (such as fencing, horsemanship, and medieval warfare), logistics and IT skills, and all of these activities are held in locations mentioned above. The staff is divided into three groups, in total of 98 persons, and they all consist of eight organizational managers, 20 interns and 70 volunteers, where interns and volunteers do not have the full-time commitment; however, there is a difference in working hours for both of these groups. The whole preparation lasts for 15 days (including the two days of the festival), and each day consists of five terms which last two hours and between the terms there is a break included. The first term starts at 8:00 AM, and the last finishes at 10:00 PM. The maximum number of terms for each of these groups varies. The volunteers should not work more than two terms per day, the interns three, while the organizational managers not more than four terms a day. However, there should always be present at least one member of the organizational management to supervise and give additional help to interns and volunteers. There is as well a maximum and a minimal number of all staff members for each location and each term, and a minimal number of interns and volunteers for each activity, location and term. For example, if a minimal number for each term in on the certain location for interns and volunteers is 10 for all activities given above, and there is at least one member of organizational managers included that would be a total of at least 11 staff members. In spite of this, the minimal number for that location and term is 15. Therefore, four additional members must be assigned depending only on their availability and costs.

These minimal numbers for terms and locations excluding the qualifications are eight for all shopping malls, 15 for the city center and 20 for each place on the Belgrade fortress, while the minimal number for each activity, location and term varied from two to five depending on the term itself, size and attractiveness of the locations and number of activities which could be performed there.

The problem was observed on one day, and it could be applicable to all days of the festival, including the preparation and the manifestation itself, changing only the input data about availability of the staff and the quotes for the maximum and minimum number regarding the terms, activities and locations.

The problem was successfully solved using the GNU *Linear Programming Kit* (GLPK) (GLPK, 2016), open source software for solving problems of linear and mixed integer programming, on the computer with the following specifications: *Intel(R) Core (TM) i7-2600 CPU* (3.40 GHz), *RAM* 12 GB. The process of finding the optimal integer solution took 0.3 seconds. This problem contains 123480 variables. The goal function value is 13650 monetary units. The optimal solution for the each term is shown in Tables 1-5.

Table 1: Optimal solution for staff scheduling problem shown for the first term

	Location1	**Location2**	**Location3**	**Location4**	**Location5**	**Location6**
Organization	1o,3v	1v	1p	1v	2p,2v	3v
Logistics	2p,1v	1o,1v	2v	1p	3v	1o,1p,2v
Marketing	1v	1p	1o	1v	3v	2p,2v
Knight skills	2v	2v	2v	3v	1o,5v	3v
HR	1p,3v	1v	1p	1o	2v	2v
IT	2v	1v	1p	1v	2p	2p,2v

Table 2: Optimal solution for staff scheduling problem shown for the second term

	Location1	**Location2**	**Location3**	**Location4**	**Location5**	**Location6**
Organization	2v	1v	1p	1v	3v	2p,1v
Logistics	3p,1v	1p,	2v	1o,	3v	1p,2v
Marketing	2v	1o,1v	1o,	1v	3v	3v
Knight skills	2v	2v	2v	1v	1o,1p,2v	5v
HR	1o,1v	1p	1v	1p	4p	1o,1p
IT	1p,2v	1v	1p	3v	3v	4v

Table 3: Optimal solution for staff scheduling problem shown for the third term

	Location1	**Location2**	**Location3**	**Location4**	**Location5**	**Location6**
Organization	4v	1v	1v	1v	1o,3v	1o,1p
Logistics	1o,1v	1p	1o,1v	1p,1v	4v	3v
Marketing	4v	1v	1p	1o,	1p,2v	1p,3v
Knight skills	1p,2v	2v	2v	1p	1p,3v	5v
HR	1v	1o,1v	1v	1v	3v	1p,2v
IT	1v	1p	1v	2v	1p,1v	3v

Table 4: Optimal solution for staff scheduling problem shown for the fourth term

	Location1	**Location2**	**Location3**	**Location4**	**Location5**	**Location6**
Organization	6v	2p	1o,	1o	4v	3v
Logistics	1o,1v	1o	2p	1p	1o,4v	1p,3v
Marketing	2v	1v	1v	1v	3v	4v
Knight skills	1p,2v	1p	2v	1p	1v	2p
HR	1v	1v	1v	3v	1o	1o,1p,2v
IT	1v	2v	1p	1v	2p,2v	3v

Table 5: Optimal solution for staff scheduling problem shown for the fifth term

	Location1	**Location2**	**Location3**	**Location4**	**Location5**	**Location6**
Organization	2p,1v	1p	1o,	1p	1o,3v	2p,1v
Logistics	1o,1v	1v	2v	1o,	2v	1o,3v
Marketing	2v	1v	1p	1v	1v	3v
Knight skills	2v	2v	2v	1v	2p,3v	1p,2v
HR	1o,2v	1o,	1v	3v	4v	5v
IT	2p,1v	1p,1v	1p	1p	1p,3v	1p,1v

Tables 1-5 present the optimal schedule for the each term of one day in the organization the Knight Fest. The integer number before the letter shows how many of certain type of staff is assigned to given activities and locations, where the letter *o* stands for the organizational manager, letter *v* for volunteers and the letter *p* stands for interns. From the Table 1, it is possible to conclude that there is a total of 6 organizational managers, 56 volunteers and 17 interns assigned in total for the mentioned activities and locations. The minimal number of all staff members in the first term and location1 was 15 and the minimal number for each activity was 4, 2, 1, 1, 2, 1 respectfully as given in the table. This means that there were four staff members

to assign additionally in order to fulfill the constraint regarding the minimal number. This was achieved by assigning four extra volunteers who were available regardless of their qualification. The same observation can be applied to other four tables. One interesting example can be found in term 5 on location1 where there was a need for a minimum of one organizational manager, three interns, and seven volunteers, but because of unavailability of volunteers software found a solution with only two additional volunteers, one additional organizational manager and one additional intern.

5. CONCLUSION AND FURTHER RESEARCH

The subject of this paper is the problem of scheduling heterogeneous multi-skill staff in case of organizing a special event. Although the staff scheduling problem, in general, was recognized more than sixty years ago, most organizations still perform this type of work manually, without the help of software which causes more time, energy and there is no warranty that the solution found this way would be even acceptable, nevertheless optimal. The paper proposed an optimal solution to a certain problem of this type – staff scheduling in case of planning and organizing a special event, as a specific type of staff scheduling. Since the schedule is applicable only for the time of the event whereas its preparation is limited in time, its organization requires multiple locations where different activities are carried out. Therefore, additional staff can be deployed and for less complex activities it is possible to hire volunteers or students as a part of their internship program. This is why it can be quite complex to create an optimal schedule which would provide minimal costs of deployment and respect all the conditions which were given. The problem was observed on the real-life example and translated into a mathematical model using 0-1 goal programming model and successfully solved using GLPK software.

The problem was scheduling multi-skill heterogeneous staff in order to plan a manifestation named The Knight Fest held each year in Belgrade, Serbia. The festival required around two weeks in total for preparation and the manifestation itself. For each day there were several locations to work on, and various activities had to be carried out. A single day was observed for this problem, where each day had five terms where actual work had to be done, and four breaks between. The staff assigned to this event consisted of organizational managers, interns and volunteers, unevenly distributed. Each of these groups had different wages and limitations concerning the minimal number of each group, the maximum number of the locations and terms which needed to be assigned to them and their availability and qualifications were taken into account.

Since the problem was observed on a single day of the festival organization, the future research could include planning for every day of the organization. This would probably lead to the fact that staff could have certain shifts and maybe a fixed working schedule. Also, it could lead to assigning the same activities to the same people during whole 15 days. As said before, their qualifications could be real numbers from 0 to 1 so it should be aimed to maximize their qualifications as well as minimizing costs which would lead to multi-criteria optimization. One of the ideas could be to decide whether is it better to deploy one organizational manager or two volunteers since in this proportion their wages would be very similar but the amount of work, knowledge and decision-making process would be different when it could be possible to choose between these two options.

6. REFERENCES

Alfares, H. K., (2004). Survey, categorization, and comparison of recent tour scheduling literature. *Annals of Operations Research*, 127(1-4), 145-175.

Avramidis, A. N., Chan, W., Gendreau, M., L'ecuyer, P., &Pisacane, O. (2010).Optimizing daily agent scheduling in a multiskill call center. *European Journal of Operational Research*, *200*(3), 822-832.

Azaiez, M., & Al Sharif, S. (2005). A 0-1 goal programming model for nurse scheduling. *Computers & Operations Research, 32*, 491-507.

Baker, K. R., (1976). Workforce allocation in cyclical scheduling problems: A survey. *Operational Research Quarterly*, 27(1), 155-167.

Barnhart, C., Cohn, A. M., Johnson, E. L., Klabjan, D., Nemhauser, G. L., & Vance, P. H. (2003).Airline crew scheduling.In *Handbook of transportation science* (pp. 517-560).Springer US.

Brusco, M. J., Jacobs, L. W., Bongiorno, R. J., Lyons, D. V., & Tang, B. (1995). Improving personnel scheduling at airline stations. *Operations Research*, *43*(5), 741-751.

Cheang, B., Li, H., Lim, A., Rodrigues, B. (2003). Nurse rostering problems—a bibliographic survey, European Journal of Operational Research, 151, 447–460.

Dantzig, G. B., (1954). Letter to the Editor-A Comment on Edie's "Traffic Delays at Toll Booths". *Journal of the Operations Research Society of America*, 2(3), 339-341.

Edie, L. C., (1954). Traffic delays at toll booths. *Journal of the operations research society of America*, 2(2), 107-138.

Ernst, A., Jiang, H., Krishnamoorthy, M., &Sier, D., (2004). Staff scheduling and rostering: A review of applications, methods and models. European Journal of Operational Research, 153 (1), 3–27 .

GNU *Linear Programming Kit* GLPK (2016). Retrieved from http://www.gnu.org/s/glpk/

Hall, R. (2012). *Handbook of healthcare system scheduling*.Springer Science+ Business Media, LLC.

Hojati,M.,&Patil, S. A., (2011). An integer linear programming based heuristic for scheduling heterogeneous, part-time service employees, *European Journal of Operational Research*, 209(1), 37-50.

Ingolfsson, A.,Campello, F., Wu, X., & Cabral, E., (2010).Combining integer programming and the randomization method to schedule employees, *European Journal of Operational Research*, 202(1), 153-163.

Knight fest organization (2016).Retrieved from http://www.vitezfest.org

Mohan.S.,(2008).Scheduling part-time personnel with availability restrictions and preferences to maximize employee satisfaction,*Mathematical and Computer Modeling,* 48(1),1806–1813.

Stefanov, D, Kostić-Stanković, M, Makajić Nikolić D, (2014) Raspoređivanje operatera u službi podrške kontakt centra komercijalne banke AD Beograd. Zbornik radova XLI SYMOPIS, Divčibare, 355-360

Van Den Bergh, J., Beliën, J., De Bruecker, P., Demeulemeester, E., & De Boeck, L., (2012). Personnel scheduling: A literature review. European Journal of Operational Research, 26 (3), 367–385 .

AGENT-BASED MODELS: A REVIEW OF APPLICATIONS

Nikola Zornic*[1], Aleksandar Markovic[1]
[1]University of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: nikola.zornic@fon.bg.ac.rs

Abstract: *Agent-based models (ABMs) have been used in various fields of science with growing interest. After a description of the approach, some interesting bibliometric data on ABM usage in numerous fields is presented, including a citation network analysis. Almost 3000 papers from Thomson Reuters Web of Science (Science Citation Index Expanded and Social Sciences Citation Index) databases have been scrutinized in terms of keywords and field of application, as well as publishing frequency in the observed journals for the period 2000-2015. Examples of thought-provoking usage in solving optimization problems and analysis of customer behaviour are also presented to depict the vast applicability.*

Keywords*: agent-based model, agent-based simulation, application review, bibliometrics, literature review*

1. INTRODUCTION

Agent-based models (ABMs) consist of a set of elements (agents) characterized by specific attributes, which interact with each other through the definition of appropriate rules in a given environment. ABMs can be used to reproduce systems related to ecology (Gerst et al., 2013), economy (Shafiei et al., 2012), agriculture (Rebaudo & Dangles, 2013), medicine (Kasaie, Andrews, Kelton, & Dowdy, 2014), to solve optimization problems (Nikolopoulou & Ierapetritou, 2012), and many others. The use of ABMs can help in understanding properties and patterns concerning the whole scenario (Billari, Fent, Prskawetz, & Scheffran, 2006) that could not be anticipated nor predicted by the observation of agents individually due to the complexity of the interactions among the elements of the system (Barbati, Bruno, & Genovese, 2012). Wooldridge and Jennings (1995) defined agent as computational system interacting with an environment, that can be described with the following features:

- *Independence* - Each agent acts without the direct control of human beings or other devices.
- *Social ability* - Interactions occur among agents through a communication language with the aim to satisfy the objectives.
- *Re-activeness* - Agents answer in a precise way to signals coming from the environment.
- *Pro-activeness* - Agents are capable of acting goal-directed. They take the initiative in order to satisfy their objectives.

Furthermore, it can be concluded from Billari et al. (2006) and Weiss (1999), that the development of an ABM needs a set of basic building blocks to be defined:

- *The object of the simulation* - It has to be specified what is the phenomenon/problem to be reproduced. Also, space, where the simulation takes place, has to be defined.
- *The agents' population* - Agents can be grouped into different categories with common characteristics, reproducing the various components of the reproduced system.
- *The adaptive capability of each agent category* - Agents of each category present a specific adaptive capability, i.e. the degree of reactiveness and pro-activeness.
- *The interaction paradigm among agents* - Each agent can interact with agents of the same or another category. In this context, agents can be competitive if they have conflicting objectives and non-competitive otherwise. In the literature, various interaction paradigms have been defined, which can be mainly classified as *Cooperative paradigms*, when a coordination from among non-competitive agents exists and *Competitive paradigms*, in which some rules for the behavior of competitive agents are defined (Weiss, 1999).

In order to follow their own objectives, agents are interacting and communicating among themselves. Communication capabilities consist of the ability to send and receive messages. This is required to ensure a coordination mechanism among agents, in order to prevent and avoid conflicts among agents' accessing resources or achieving objectives.

Implementing a cooperation-based paradigm means employing planning approaches to reduce resource contention and to ensure the achievement of global objectives. No mediator is involved. In this case, each

agent is self-interested, meaning that the final solution may be the best for the single agent involved, but not for the group as a whole. These approaches can be distinguished into two main categories (Barbati et al., 2012):

- *Distributed approaches*, in which agents are endowed with self-organizing rules for resource sharing and goal pursuing.
- *Centralized approaches*, in which a mediator agent is assigned with the task of regulating and supervising agents' behaviors.

Since its first appearance in early 1970s, the exploration of the ABM literature has seen vigorous development, especially in the last decade owing to availability of various ABM software. This paper employs a bibliometric methodology that is geared towards a review of literature productivity, and an observation of the trends in ABM.

After defining agent-based models more closely, the paper is organized as follows. A bibliometric review of ABMs usage in scientific papers is presented in section 2. Then, the application of ABMs in chosen fields is exemplarily demonstrated in section 3. Finally, conclusions and some prospects for the future research are provided in section 4.

2. BIBLIOMETRIC REVIEW OF ABMs APPLICATION

In order to analyze the presence of ABMs in the scientific literature we have performed a search through the Thomson Reuters Web of Science (Thomson Reuters, 2016b) in Science Citation Index Expanded (SCI-EXPANDED) and Social Sciences Citation Index (SSCI) databases and collected data. The keywords we were looking for are:

"agent-based simulation" OR "agent-based model"

for the time interval 2000-2015, as there is an inconsiderable amount of published papers in previous years. The search was conducted for the *topic*, which includes: title, abstract, author and keywords in Web of Science (Thomson Reuters, 2016b).

As displayed in Table 1, we have collected data for 2902 papers. The increase in a number of papers per year is evident, which means the method is becoming more acknowledged. The number of ABM-related papers increased almost 30 times from 2000. until 2015, while total number of SCI-EXPANDED and SSCI papers increased only 1.77 times in the same time period. The biggest leap in ABM-related published papers occurred in period 2012-2013, from 294 to 380.

Table 1: Number of ABM-related and total SCI-EXPANDED & SSCI papers

Year	ABM Papers	Total SCI-EXPANDED & SSCI Papers
2000	15	1,099,828
2001	33	1,087,477
2002	30	1,132,596
2003	61	1,177,644
2004	72	1,272,284
2005	92	1,340,592
2006	116	1,399,016
2007	118	1,471,075
2008	137	1,553,662
2009	173	1,613,263
2010	227	1,650,449
2011	290	1,736,725
2012	294	1,820,086
2013	380	1,917,193
2014	415	1,942,001
2015	449	1,945,858
Total	**2902**	**24,159,749**

Those 2902 papers are published in 908 journals. In Table 2 we present 19 journals with at least 20 papers published in the observed period. The top contributor is the *Journal of Artificial Societies and Social Simulation* with 208 papers published. Since its first issue in 1998, it has been one of the leading references for readers interested in social simulation and the application of computer simulation in the social sciences. The list of journals shows a multidisciplinary interest in ABMs. In addition to journals that are expected to be on the list, journals specialized in Simulation and Modeling (*Journal of Artificial Societies and Social Simulation*, *Simulation Modelling Practice and Theory*, etc.), there are journals specialized in other scientific fields, like: Ecology (*Ecological Modelling*, *Environmental Modelling & Software*, etc.), Biology (*Journal of Theoretical Biology*), Economics (*Journal of Economic Dynamics and Control*, *Journal of Economic Interaction and Coordination*), etc.

Table 2: Journals accounting for at least 20 papers

Journal	Papers	Journal IF (2014)	IF Without Journal Self Cites (2014)	% IF Change Without Self Cites
Journal of Artificial Societies and Social Simulation	208	0.941	0.811	-13.82%
Physica A: Statistical Mechanics and its Applications	88	1.732	1.181	-31.81%
PLoS ONE	88	3.234	2.885	-10.79%
Journal of Theoretical Biology	62	2.116	1.868	-11.72%
Ecological Modelling	61	2.321	2.085	-10.17%
Environmental Modelling & Software	49	4.420	2.636	-40.36%
Advances in Complex Systems	45	0.968	0.920	-4.96%
SIMULATION: Transactions of The Society for Modeling and Simulation International	40	0.818	0.631	-22.86%
Simulation Modelling Practice and Theory	33	1.383	1.064	-23.07%
Computers, Environment and Urban Systems	32	1.537	1.349	-12.23%
Computational and Mathematical Organization Theory	24	0.840	0.740	-11.90%
Expert Systems With Applications	24	2.240	1.793	-19.96%
Journal of Economic Dynamics and Control	22	1.018	0.851	-16.40%
Journal of Economic Interaction and Coordination	22	0.962	0.807	-16.11%
Energy Policy	21	2.575	2.106	-18.21%
Physical Review E	21	2.288	1.778	-22.29%
Environment and Planning B: Planning and Design	20	0.983	0.941	-4.27%
Journal of Economic Behavior & Organization	20	1.297	1.154	-11.03%
Journal of Simulation	20	0.580	0.500	-13.79%

Table 2 also reports impact factor (IF), IF without journal self-citations, and percentage of change in IF after the self-citations have been excluded. Data about journals and IF is collected from Tomson Reuters InCites™ (Thomson Reuters, 2016a). We can see that the journal with the highest IF (4.420) is *Environmental Modelling & Software*, which is also the journal with the highest decrease in IF after the self-citations were excluded (-40.36%). Journal that stands out with high IF (3.234) and a small percentage of change after the self-citations have been excluded (-10.79%) is *PloS ONE*. We can also see that there are journals with IF lower than 1 (*Environment and Planning B: Planning* and *Design and Advances in Complex Systems*) whose IF did not drastically change after excluding self-citations - change is lower than 5%.

Additional analysis of citations is conducted using network visualization and exploration software Gephi (Gephi, 2016). Total citation count for the selected journals has been collected from the Tomson Reuters InCites™ (Thomson Reuters, 2016a) for the period until 2014. The obtained results are displayed in Figure 1, where node size is determined by the percentage of self-citations specific journal has (number self-citations divided by a total number of citations). For example, it is noticeable that *Computers, Environment and Urban Systems* has lowest (5.56%) and *Environmental Modelling & Software* highest (20.48%) percentage of self-citations. The weight of the line is defined as the sum of percentages of citations acquired from the other observed journal. For example, *Physica A: Statistical Mechanics and its Applications* got 1015 citations from *Physical Review E* and has 16,974 citations in total, which means *Physica A: Statistical Mechanics and its Applications* got 5.98% of citations from *Physical Review E*. In the other direction, *Physical Review E* got 2948 citations from *Physica A: Statistical Mechanics and its Applications* and has

81,746 in total, which means *Physical Review E* got 3.61% of citations from *Physica A: Statistical Mechanics and its Applications.* The weight of line is determined by the sum of these percentages (5.98% + 3.61%). The size of the arrow is defined separately for each side in a similar manner. For example, the two-sided arrow between *Physical Review E* and *Physica A: Statistical Mechanics and its Applications* means that journals have cited each other quite often. Connection between *Physica A: Statistical Mechanics and its Applications* and *Advances in Complex Systems* stands out. Namely, *Advances in Complex Systems* is very frequently cited by *Physica A: Statistical Mechanics and its Applications* (8.19% of total citations *Advances in Complex Systems* received are from *Physica A: Statistical Mechanics and its Applications*). Similarly, we can analyze the remaining connections between journals.

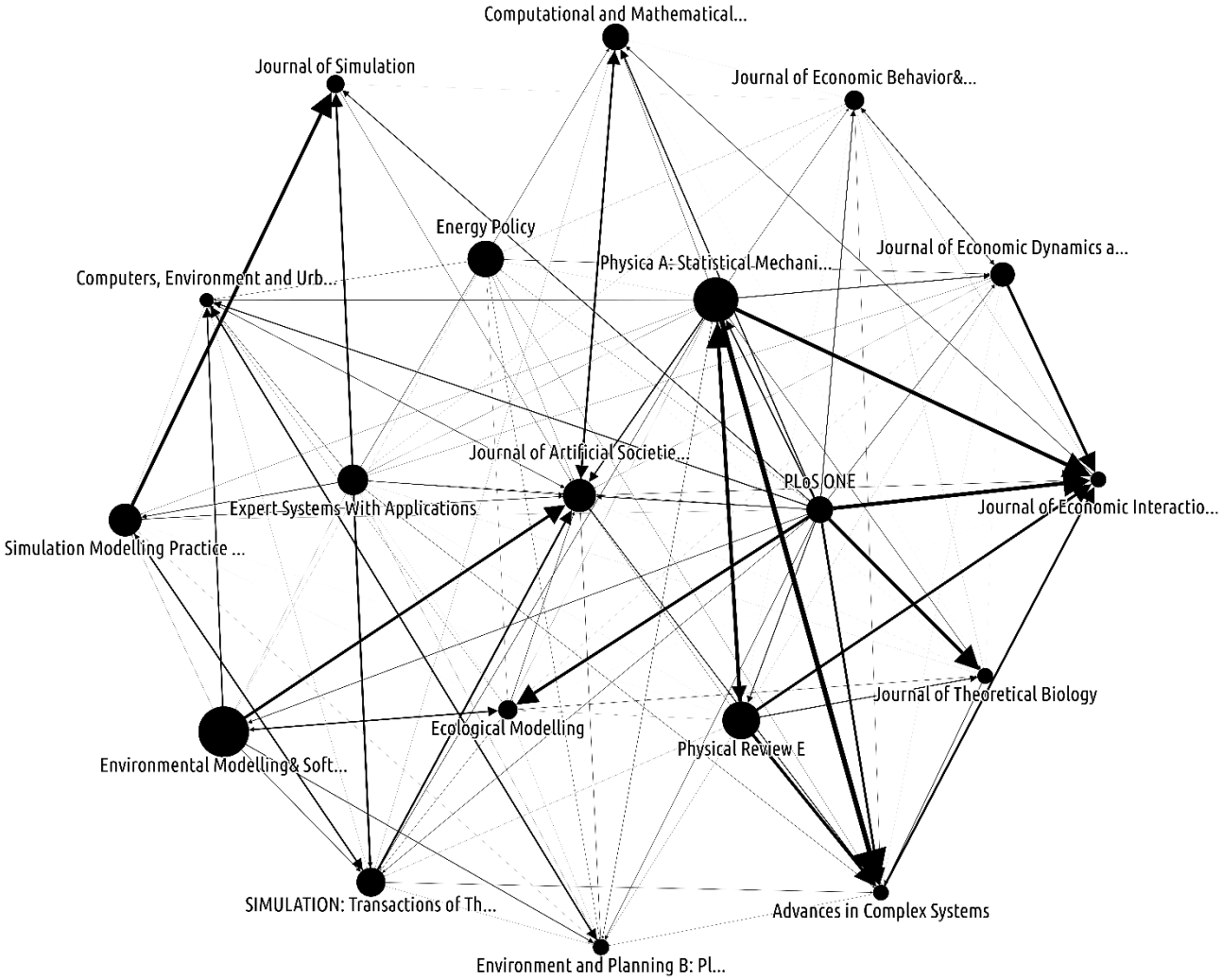

Figure 1: Network of citations among selected journals

Analyzing the research area retreived from Web of Science (Thomson Reuters, 2016b), ABMs are applied in (Table 3), we can provide one more evidence that ABMs are suitable for solving various problems. It is particularly interesting that *Environmental Sciences & Ecology* and *Business & Economics* are among areas with most frequent ABMs usage, just behind IT related areas. In the next paper section, we presented some of the applications in these fields.

Table 3: Research area

Research Area	Papers	Research Area	Papers
Computer Science	635	Life Sciences & Biomedicine - Other Topics	140
Environmental Sciences & Ecology	441	Geography	101
Engineering	396	Transportation	86
Business & Economics	364	Mathematical Methods In Social Sciences	63
Social Sciences - Other Topics	254	Public, Environmental & Occupational Health	63
Science & Technology - Other Topics	219	Biochemistry & Molecular Biology	55
Physics	212	Energy & Fuels	55

Research Area	Papers	Research Area	Papers
Mathematics	202	Sociology	53
Operations Research & Management Science	185	Information Science & Library Science	52
Mathematical & Computational Biology	165	Psychology	51

We also wanted to place attention on the keywords authors chose to best represent their work (Table 4), the most frequent keywords used in the observed papers. Disregarding the presumed ones (such as *model*, *systems*, *simulation*, etc.), other keywords are describing the field of application (such as *management*, *social networks*, etc.). Another keyword which is very popular in recent years is *optimization* (Barbati et al., 2012).

Table 4: Most used keywords

Keyword	Papers	Keyword	Papers
Dynamics	321	Agent-based simulation	65
Model	301	Strategies	65
Systems	268	Performance	63
Simulation	214	Information	62
Behavior	212	Transmission	60
Management	141	Social networks	59
Evolution	138	Impact	58
Networks	135	System	56
Models	123	Competition	55
Agent-based model	96	Diffusion	55
Cooperation	79	Decision-making	54
Patterns	79	Optimization	51
Growth	73	Framework	50
Land-use	69	Market	50
Design	67		

3. APPLICATION EXAMPLES

The applicability of the agent-based models and agent-based simulation is exemplarily demonstrated in two fields which were found interesting: optimization problems and analysis of customer behaviour.

3.1. Optimization problems

Agent-based approaches are successfully employed to tackle a broad class of scheduling, transportation, supply chain planning, and other optimization problems (Barbati et al., 2012). Some of the attention-grabbing examples are presented in the following research papers.

Böcker, Lind, and Zirkler (2001) presented the scheduling system for railway transport system train coupling and sharing (TCS). This scheduling system includes the planning of travel units and the optimization of solutions. The usage of ABM allows planning in a real-time scenario. The initially defined solution is then optimized by a simulated trading approach.

Fischer and Gehring (2005) used ABM as a supporting system for planning transshipments of imported finished vehicles via a seaport. The focus was on storage allocation, i.e. allocation of parking areas for the temporary storage of vehicles and on deployment scheduling, i.e. the allocation of drivers to the vehicles that have to be moved in the terminal area. Planning tasks, which is usually carried out by different operators, are assigned to two different agent types: a further agent and the coordinator agent, who is responsible for combining the local sub-plans into a global plan in such a way that the demand for drivers is minimized and balanced.

An ABM, which can simulate a supply chain, was developed by Liang and Huang (2006). The main purpose of their model is to coordinate all entities of the supply chain and minimize total cost while every entity may

use different inventory system. Two types of agents are employed: control agents (each echelon has agents of this type) and demand forecast agents. The control agent collects historical demand data and strategies from managers and aims at building rule-base for supply chain management. The demand forecast agent, which communicates with control agents, is using a genetic algorithm (GA) to forecast each echelon's demand and minimize the total cost for the entire supply chain.

3.2. Analysis of customer behaviour

Consumer behavior research is a multidisciplinary phenomenon which involves various areas: psychology, marketing, sociology, economics, and engineering (Zhang & Zhang, 2007). The process in which consumers make their purchase decisions has been of great interest to researchers and practitioners for a long time (Burnett & Lunsford, 1994). Interactions among consumers and the purchase decision-making process generate market dynamics, such as the decoy effect and lock-in. Research into consumers' purchase decision-making and interactions increases the understanding of market dynamics. Traditional studies on consumer decision patterns, such as data mining, focus on using static equilibrium-based mathematical and statistical techniques to model consumers' socio-demographic and behavioral data (Zhang & Zhang, 2007). However, three types of data exist regarding consumer purchase decisions: 1) demographic data, 2) behavioral data, and 3) psychographic data (Rud, 2001). ABMs take a stage exactly here as they can incorporate all three types of data to analyze customers' behavior in more detail.

Cavoski and Markovic (2015) applied ABM to analyze customers' behaviour and online retailers' strategies. They identified three key segments of the observed system: the seller segment, the consumer segment, and the communication channel segment. Each of these segments are afterwards modeled as one type of agent population. Namely, they used empirical and theoretical data retrieved from the B2C online market for modelling agents. As a result of this modelling process, market managers are provided with the tool for investigating impacts of implementing their own business strategies (Marković, Čavoški, & Novović, 2016).

Another interesting application of ABM is in predicting electric vehicles market share evolution (Shafiei et al., 2012). The model has potential to analyze the effects of fuel prices, vehicles' attributes, consumer preferences, and social influences. The market share evolution of electric vehicles is observed in different scenarios. The results showed that electric vehicles can be successfully adopted in moderate and favorable scenarios. In adverse scenarios, like the ones with low gasoline price and the one with constantly high electronic vehicle price, support policies are needed (such as reduction in import taxes and infrastructure development for easier recharging).

4. CONCLUSION

In this work, the application of agent-based models is reviewed. Firstly, the comprehensive bibliometric data has been presented, including IF, IF without journal self-citations, and citation network analysis. We can conclude just at a glance that ABMs are successfully used in various fields of science. Moreover, their usage is growing with the course of time, from 15 papers published in 2000, this number increased to 449 in 2015. This proves that researchers from numerous scientific fields recognized the power of ABMs and are utilizing that power for solving the problems they run into.

Further in the paper, ABMs' applications are presented. The focus of the chosen papers was on solving optimization problems and analyzing the customer behaviour. It is proved that ABMs can be successfully employed to cope with scheduling, transportation, and supply chain optimization problems. Also, we provided insights on the implementation of ABMs for customer's behaviour and market share prediction.

The growing interest in agent-based modeling and agent-based simulation seems to be a very promising field of study. For this reason, future researches will be devoted to recognize and model other classes of problems using ABMs.

REFERENCES

Barbati, M., Bruno, G., & Genovese, A. (2012). Applications of agent-based models for optimization problems: A literature review. *Expert Systems with Applications*, *39*(5), 6020–6028. http://doi.org/10.1016/j.eswa.2011.12.015

Billari, F. C., Fent, T., Prskawetz, A., & Scheffran, J. (Eds.). (2006). *Agent-Based Computational Modelling*. Heidelberg: Physica-Verlag HD. http://doi.org/10.1007/3-7908-1721-X

Böcker, J., Lind, J., & Zirkler, B. (2001). Using a multi-agent approach to optimise the train coupling and sharing system. *European Journal of Operational Research*, *131*(2), 242–252.

http://doi.org/10.1016/S0377-2217(00)00124-7
Burnett, M. S., & Lunsford, D. A. (1994). Conceptualizing Guilt in the Consumer Decision-making Process. *Journal of Consumer Marketing*, *11*(3), 33–43. http://doi.org/10.1108/07363769410065454
Cavoski, S., & Markovic, A. (2015). Analysis of Customer Behaviour and Online Retailers Strategies Using the Agent-Based Simulation. *Management - Journal for Theory and Practice of Management*, *20*(77), 13–24. http://doi.org/10.7595/management.fon.2015.0031
Fischer, T., & Gehring, H. (2005). Planning vehicle transhipment in a seaport automobile terminal using a multi-agent system. *European Journal of Operational Research*, *166*(3), 726–740. http://doi.org/10.1016/j.ejor.2004.03.042
Gephi. (2016). Gephi - The Open Graph Viz Platform. Retrieved April 20, 2016, from https://gephi.org
Gerst, M. D., Wang, P., Roventini, A., Fagiolo, G., Dosi, G., Howarth, R. B., & Borsuk, M. E. (2013). Agent-based modeling of climate policy: An introduction to the ENGAGE multi-level model framework. *Environmental Modelling & Software*, *44*, 62–75. http://doi.org/10.1016/j.envsoft.2012.09.002
Kasaie, P., Andrews, J. R., Kelton, W. D., & Dowdy, D. W. (2014). Timing of Tuberculosis Transmission and the Impact of Household Contact Tracing. An Agent-based Simulation Model. *American Journal of Respiratory and Critical Care Medicine*, *189*(7), 845–852. http://doi.org/10.1164/rccm.201310-1846OC
Liang, W.-Y., & Huang, C.-C. (2006). Agent-based demand forecast in multi-echelon supply chain. *Decision Support Systems*, *42*(1), 390–407. http://doi.org/10.1016/j.dss.2005.01.009
Marković, A., Čavoški, S., & Novović, A. (2016). Agent-Based Model for Analysis of Stakeholders behaviour on B2C E-Markets. In *35th International Conference on Organizational Science Development* 576–589. Portorož.
Nikolopoulou, A., & Ierapetritou, M. G. (2012). Hybrid simulation based optimization approach for supply chain management. *Computers & Chemical Engineering*, *47*, 183–193. http://doi.org/10.1016/j.compchemeng.2012.06.045
Rebaudo, F., & Dangles, O. (2013). An agent-based modeling framework for integrated pest management dissemination programs. *Environmental Modelling & Software*, *45*, 141–149. http://doi.org/10.1016/j.envsoft.2012.06.014
Rud, O. (2001). *Data Mining Cookbook: Modeling Data for Marketing, Risk, and Customer Relationship Management*. New Yourk: John Wiley & Sons, Inc.
Shafiei, E., Thorkelsson, H., Ásgeirsson, E. I., Davidsdottir, B., Raberto, M., & Stefansson, H. (2012). An agent-based modeling approach to predict the evolution of market share of electric vehicles: A case study from Iceland. *Technological Forecasting and Social Change*, *79*(9), 1638–1653. http://doi.org/10.1016/j.techfore.2012.05.011
Thomson Reuters. (2016a). InCites™. Retrieved April 20, 2016, from https://incites.thomsonreuters.com
Thomson Reuters. (2016b). Web of Science. Retrieved April 14, 2016, from https://apps.webofknowledge.com
Weiss, G. (1999). *Multiagent systems: a modern approach to distributed artificial intelligence*. Cambridge: MIT Press.
Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: theory and practice. *The Knowledge Engineering Review*, *10*(02), 115–152. http://doi.org/10.1017/S0269888900008122
Zhang, T., & Zhang, D. (2007). Agent-based simulation of consumer purchase decision-making and the decoy effect. *Journal of Business Research*, *60*(8), 912–922. http://doi.org/10.1016/j.jbusres.2007.02.006

THE INFLUENCE OF EDUCATIONAL DEVELOPMENT INDICATORS ON GRADUATES' EMPLOYMENT RATES

Ana Bjelaković[1], Marina Dobrota*[1]
[1]University of Belgrade, Faculty of Organizational Sciences, Serbia
*Corresponding author, e-mail: dobrota.marina@fon.bg.ac.rs

Abstract: *Education as one of the fundamental factors of the development improves the quality of life and leads to broad social benefits to individuals and society. It has a major impact on employment prospects. Economic development is impossible without a substantial investment in human capital. This paper investigates the role of the education system in employment. We determined a set of non-financial and financial indicators of educational development as explanatory factors. To examine the correlation between proposed factors and the employment rate of recent graduates, backwards multiple linear regression analysis was conducted. Our findings suggest that we have both financial and non-financial indicators influencing the employment rates in the final model. The most influential non-financial indicators were those that measure a number of young people that continued their education and foreign languages spoken by the nation while the financial factors were public expenditure on education and annual expenditure per student.*

Keywords*: education, development, employment, human capital*

1. INTRODUCTION

The population is continuously progressing and boosting throughout observation and learning, and this fact makes education and lifelong learning the major segments of human development (Dobrota & Benkovic, 2014; Dobrota, Martic, Bulajic, & Jeremic, 2015; Dobrota, Savic, & Bulajic, 2015). Being as important as it is, education has been a main topic of a number of scientific research (Andjelkovic Labrovic, Bijelic, & Milosavljevic, 2014; Benkovic & Dobrota, 2012; Dobrota, Savic, et al., 2015; Radenkovic, Despotovic Zrakic, Bogdanovic, Barac, & Labus, 2014). According to Hirsto, Lampinen, and Syrjakari (2013), in the field of educational development, many different concepts have either been used interchangeably or to focus on different aspects of development. Early research originally examined key developments affecting higher education worldwide and patterns in higher education development (Altbach, 2003). More than ever before, education has become a lifelong course in which individuals proceed to learn in formal and informal environments throughout their lives (Blossfeld, Roßbach, & von Maurice, 2011; Dobrota, Jeremić, Bulajić, & Radojičić, 2015; Dobrota, Savic, et al., 2015; Fleming, 2011; Nguyen & Walker, 2016; Radenkovic et al., 2014; Roche, 2015; Volles, 2016).

Education correlates strongly with most important social and economic outcomes (Hout, 2012). The belief that education is an engine of growth rests on the quality and quantity of education in any country (Horvat, Dobrota, Krsmanovic, & Cudanov, 2015; Horvat, Krsmanovic, Dobrota, & Cudanov, 2013; Lutz & KC, 2011). Further policy reforms and changes in higher education systems have become top national priorities, as governments in several countries embrace the notion that higher education is a key component of their efforts towards economic and social development (Chapman & Miric, 2009). Educational systems are expected to boost the society in various ways through the deployment of different educational models, on primary and secondary, but especially tertiary level. Higher education models are various, and for a new model to contribute something novel, it must have the potential to move educational practice and research forward in ways that improve teaching and learning experiences and outcomes (Benkovic & Dobrota, 2012; Dobrota & Benkovic, 2014; Knewstubb, 2016).

Theories have proved that formal education is necessary to improve the production capacity of a nation and discusses the rationality behind investment in human capital (Olaniyan & Okemakinde, 2008). Education has a substantial impact on labor market outcomes such as earnings and employment, as well as non-market outcomes such as health, longevity, civic participation, and even criminal activity (Oreopoulos & Salvanes, 2011). The primary focus of the paper is examining the relations between the educational system features on one hand, and earnings and employment as a major factor of social development and well-being on the other.

It is well-known that the education has a major influence on earnings and employment (Blundell, Dearden, Goodman, & Reed, 1997; P. Brown, 1995; Evans, 2016; Purcell & Quinn, 1996; Waters, 2009; Weiss, Klein,

& Grauenhorst, 2014). For example, Brown and Sessions (1998) debate that, regardless of whether schooling signals or augments productivity, the correlation between education and earnings renders the enhancement of lifetime earnings and, as such, may represent a good investment for individual workers. However, the path from a pupil or a student to an employee is not always so mere. According to Charles (2015), the transition from higher education to employment requires a pre-professionalization, which is characterized by the dominant role of professional skills and their incorporation into the structure of initial higher education. This mechanism is explained by the persistence of an idealized conception of "matching" that still profoundly marks the relations between education and employment (Charles, 2015). The question finally arises as to whether or not a student simply needs a Ph.D. diploma in order to be a dominant employee candidate in a private sector (Pedersen, 2016).

In the course of previous studies on education and employment, the goal of this study was to show which indicators of the educational development have a significant impact on the employment rate of recent graduates. The research data were obtained from an accessible dataset Eurostat (Eurostat, 2016) that includes levels of educational development and structure based on individual criteria. In the following section, we defined the methodology and indicators that were used. This is followed by a presentation and discussion of the results. Finally, a conclusion was given.

2. METHODOLOGY

In this article, the main goal was to determine whether there exists and what is the strength and form of the relationship between various educational indicators and employment rates of recent graduates. For this purpose, twelve different criteria were chosen to determine a country's educational structure and development. These are (Eurostat, 2016):

Non-financial indicators of the educational structure:

- *Employment rates of recent graduates* - the employment rates of persons aged 20 to 34.The indicator is calculated based on data from the EU Labour Force Survey.
- *School expectancy* - expected years of education over a lifetime (this type of estimate is accurate if current patterns of enrolment continue on the same trend in the future).
- *18-year-olds in education* - the percentage of all 18-year-olds who are still in any kind of school (all ISCED levels). It gives an indication of the number of young people who have not abandoned their efforts to improve their skills through initial education, and it includes both those who had a regular education career without any delays as well as those who are continuing even if they had to repeat some steps in the past.
- *Mobility of students in Europe* - the incoming students and outgoing students for each country, using the figures provided by the host country on foreign students enrolled in tertiary education by nationality. The lack of data on the distribution of students by nationality in some countries leads to underestimation of the values.
- *Share of women among tertiary students* - the percentage of women among all students in tertiary education, regardless of the field of education they are in.
- *Pupil teacher ratio in primary education* **-** calculated by dividing the number of full-time equivalent pupils by the number of full-time equivalent teachers teaching at ISCED level 1. Only teachers in service (including special education teachers) are taken into account.
- *Foreign languages learnt per pupil* **-** average number of foreign languages learned per pupil in secondary education. A foreign language is recognized as such in the curriculum or other official documents relating to education in the country.
- *Pupils learning English* - the percentage of all pupils in upper secondary education who are learning English as a foreign language. It only covers general and not vocational education in countries where English is described as a foreign language in the curriculum or other official document relating to education in the country.

The next five variables represent financial indicators of the educational structure

- *Annual expenditure on public and private educational institutions per student*
- *Private expenditure on education*
- *Annual expenditure on public and private educational institutions compared to GDP per capita*
- *Total public expenditure on education*
- *Public expenditure on education*

These variables show education funding of the public sector either by directly bearing the current and capital expenses of educational institutions (direct expenditure for educational institutions) or by supporting students and their families with scholarships and public loans, as well as by transferring public subsidies for educational activities to private firms or non-profit organizations (transfers to private households and firms). These financial indicators demonstrate the importance given to education in any particular country and the potential for its further development.

The main idea was to determine whether the employment rates of recent graduates depends on other listed variables, i.e. financial and non-financial indicators of educational development (Dobrota, Savic, et al., 2015). We used backwards multiple LS linear regression (Narula & Wellington, 1982; Rencher & Christensen, 2012; Seal, 1967) to create a model, which would automatically exclude non-significant variables from the observation (Gujarati, 2002). The original multivariate regression model is given by the following formula:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i \quad (1)$$

where Y_i is an i-th observation od the dependent variable Y, β_j is a j-th regression coefficient (j=1,…k), X_{ji} is an i-th observation for the j-th independent variable, and ε is a residual.

Employment rates of recent graduates was chosen as a dependent variable while other listed variables were chosen as independent explanatory variables.

3. RESULTS

Before presenting the multiple LS linear regression model, we examined and established the possible relationships among variables used in this research. First, we wanted to identify is whether there exists and what is the form of the dependency of Employment rates from some of the most specific indicators of educational development.

Figure 1 shows the dependency of the Employment rates of recent graduates from an explanatory variable Annual expenditure on public and private educational institutions per student.

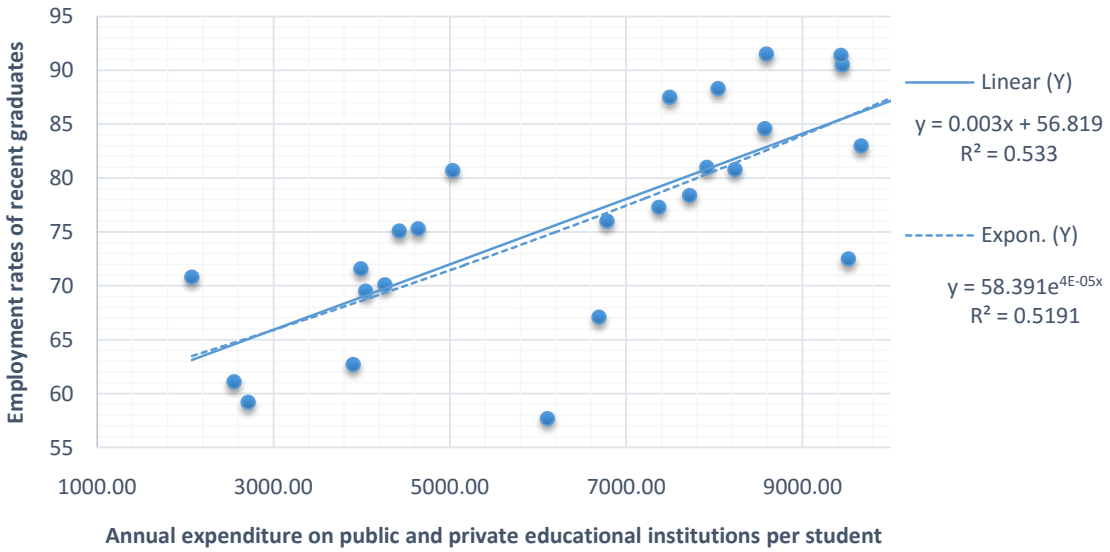

Figure 1: Relationship between Employment rates of recent graduates and Annual expenditure on public and private educational institutions per student

As can be seen in Figure 1, there is a mutual dependency between these two variables. We were able to extract a positive linear relationship from the given data, and the linear model is given in Figure 1. The Pearson's correlation coefficient r=0.73 (p<0.001), and the coefficient of determination R^2=0.533, indicating that 53.3% of the variability of Employment rates is explained by the model. For comparison, we also showed an exponential relationship between these two variables. Model is showed in Figure 1, where the coefficient of determination R^2=0.519, indicating a bit lower percent of explained variability.

Figure 2 shows the dependency of the Employment rates of recent graduates from an explanatory variable Public expenditure on education.

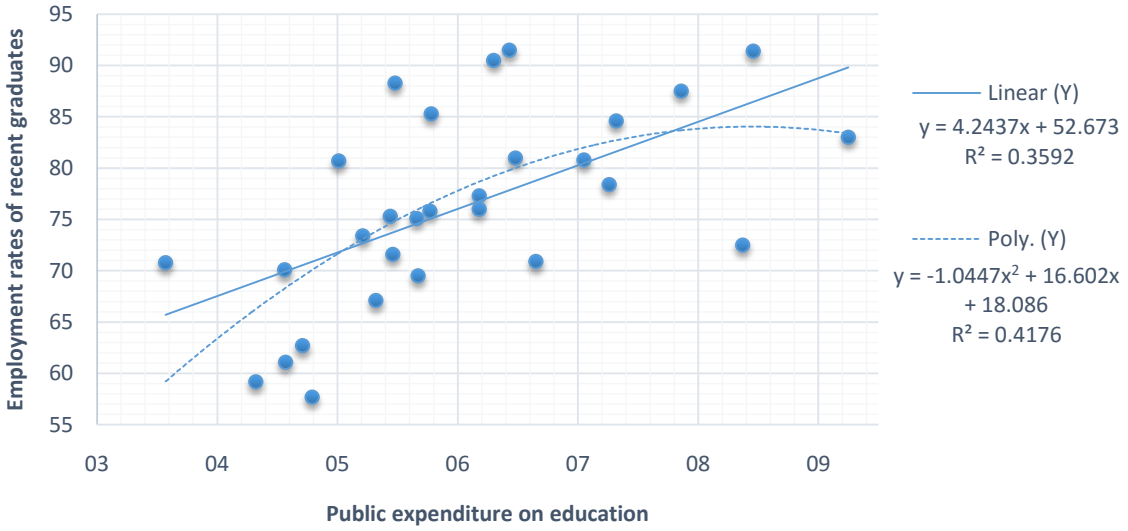

Figure 2: Relationship between Employment rates of recent graduates and Public expenditure on education

From Figure 2, we can see that there is a mutual dependency between these two variables and the linear model is given in the figure. We also extracted a positive linear relationship from the data and the Pearson's correlation coefficient r=0.599 (p=0.001). The coefficient of determination R^2=0.359, indicating that 35.9% of the variability of Employment rates is explained by the model. For comparison, here we showed a polynomial relationship between these two variables, given in Figure 2. In this model, the coefficient of determination R^2=0.417, indicating a higher percent of explained variability, having that the polynomial model could be a better fit for these data.

Figure 3 shows the dependency of the Employment rates of recent graduates from an explanatory variable Annual expenditure on public and private educational institutions compared to GDP per capita.

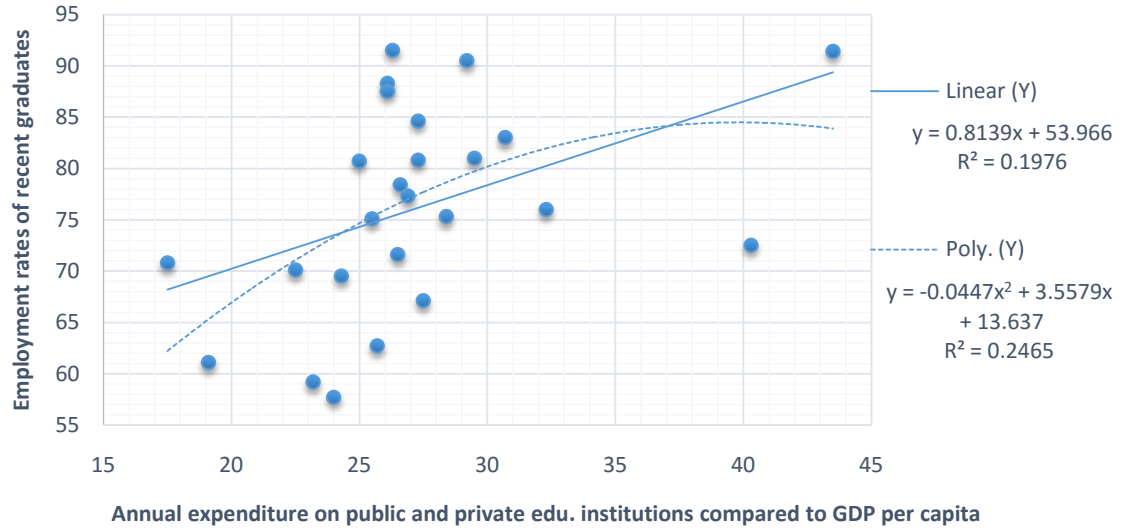

Figure 3: Relationship between Employment rates of recent graduates and Annual expenditure on public and private educational institutions compared to GDP per capita

From the Figure 3, a larger dispersion can be seen, which is also the characteristic of most of the other independent variables, especially non-financial. For the majority of financial independent variables, positive dependence is found, while this is not the case for non-financial variables (Table 1). In this model, we extracted a positive linear relationship where the Pearson correlation coefficient r=0.445 (p=0.026). The coefficient of determination R^2=0.197, indicating that only 19.7% of the variability of Employment rates is

explained by the model. For comparison, we also showed a polynomial relationship, given in Figure 3, for which coefficient of determination R^2=0.246, indicating a bit higher percent of explained variability.

Table 1 presents the Pearson's correlation coefficients between all the given variables. These are given in order better to understand the relationships and dependencies between data.

Table 1: Pearson's correlation coefficients between variables

Employment rates of recent graduates	ERR	SE	18ED	MEU	SWT	PTR	FL	PLE	AEE PP	PREE	AEE GDP	TPEE
School expectancy	.221											
18-year-olds in education	-.051	.770**										
Mobility of students in Europe	-.104	-.008	.097									
Share of women among tertiary students	-.118	.439*	.407*	-.402*								
Pupil teacher ratio in primary education	-.211	-.156	-.123	.346	-.121							
Foreign languages learnt per pupil	.092	-.294	-.039	-.129	-.194	-.277						
Pupils learning English	-.179	-.056	.121	.121	-.077	.490*	-.023					
Annual expenditure on public and private educational institutions per pupil student	.730**	.034	-.342	-.070	-.287	-.413*	.043	-.428*				
Private expenditure on education	.195	-.308	-.542**	.057	-.230	.107	-.088	.000	.276			
Annual expenditure on public and private educational institutions compared to GDP per capita	.445*	-.285	-.598**	-.158	-.094	-.344	.100	-.200	.603**	.664**		
Total public expenditure on education	.093	.017	-.127	.512**	-.473**	.364	-.262	.184	.186	.218	-.097	
Public expenditure on education	.599**	-.022	-.277	-.126	-.071	-.439**	.278	-.304	.839**	.284	.700**	-.114

*p<0.05, **p<0.01

Table 2 presents the original multiple linear regression model.

Table 2: The original multiple linear regression model

Predictors[1]	M	SD	B	SE	Beta	95%CI B	
Intercept			64.188	43.204		-33.546	161.923
School expectancy	17.573	1.5848	-.048	1.846	-.008	-4.224	4.128
18-year-olds in education	81.227	16.2219	.250	.229	.441	-.268	.768
Mobility of students in Europe	18.56	21.100	-.049	.108	-.112	-.294	.196
Share of women among tertiary students	55.805	3.6624	.095	1.074	.038	-2.333	2.524
Pupil teacher ratio in primary education	14.041	2.8811	.382	.687	.120	-1.173	1.937
Foreign languages learnt per pupil	1.655	.3334	10.423	5.530	.378	-2.088	22.933
Pupils learning English	93.877	8.7026	-.422	.254	-.399	-.996	.153
Annual expenditure on public and private educational institutions per student	6483.514	2407.889	.006*	.002	1.701	.002	.011
Private expenditure on education	.6323	.36810	-3.880	7.929	-.155	-21.815	14.056
Annual expenditure on public and private educational institutions compared to GDP per capita	27.668	5.4817	.089	1.002	.053	-2.178	2.356
Total public expenditure on education	21183.282	32893.378	-3.922E-6	.000	-.014	.000	.000
Public expenditure on education	5.675	1.4605	-6.293	3.515	-1.000	-14.245	1.658
R^2	0.553						
F	3.162*						

*p<0.05, **p<0.01
[1]Dependent variable: Employment rates of recent graduates

In the original model, all of the variables listed in section 2 were included in the analysis. From Table 2 it can be seen that, when combining independent variables in this way, only Annual expenditure on public and

private educational institutions per student significantly influences the dependent variable Employment rates of recent graduates. All other variables have significance larger than 0.05. With the coefficient of determination R^2=0.553, the model reveals that 55.3% of the variability of Employment rates of recent graduates is explained by the given combination of explanatory variables. The whole model is significant at 0.05 level of significance. The model also exhibits a large level of multicollinearity.

Table 3 presents the reduced backwards multiple linear regression model. This model was used to exclude the non-significant variables from the model, step by step automatically, and thus create the model that fits the given data in a best possible way. The method performed eight continuous iterations until the final model was obtained.

Table 3: The reduced backwards multiple linear regression model

Predictors[1]	M	SD	B	SE	Beta	95%CI B	
Intercept			63.335	18.131		24.899	101.772
18-year-olds in education	81.227	16.2219	.246**	.078	.434	.081	.411
Foreign languages learnt per pupil	1.655	.3334	10.416*	4.181	.378	1.552	19.280
Pupils learning English	93.877	8.7026	-.313*	.157	-.297	-.646	.019
Annual expenditure on public and private educational institutions per student	6483.514	2407.889	.006**	.001	1.573	.004	.008
Public expenditure on education	5.675	1.4605	-5.752*	2.028	-.913	-10.051	-1.452
R^2	0.705						
F	11.059**						

*p<0.05, **p<0.01
[1]Dependent variable: Employment rates of recent graduates

The reduced model includes only five, of originally eleven independent variables. Among them, we have two financial and three non-financial indicators influencing the Employment rate of recent graduates. The most influential non-financial indicators were 18-year-olds in education, Foreign languages learnt per pupil, and Pupils learning English; the most influential financial indicators are Annual expenditure on public and private educational institutions per student and Public expenditure on education. Among them, 18-year-olds in education and Annual expenditure on public and private educational institutions per student are significant on 0.01 level of significance while others are on 0.05 level of significance. The most significant absolute influence has Annual expenditure on public and private educational institutions per student, followed by Public expenditure on education, and then 18-year-olds in education, Foreign languages learnt per pupil, and Pupils learning English respectively, which can be seen from the column Beta in Table 3. The estimated model is presented as follows in the given formula:

$$\hat{Y} = 63.335 + 0.246X_1 + 10.416X_2 - 0.313X_3 + 0.006X_4 - 5.752X_5 \quad (2)$$

where Y is the dependent variable Employment rates of recent graduates and X_1 to X_5 are 18-year-olds in education, Foreign languages learnt per pupil, Pupils learning English, Annual expenditure on public and private educational institutions per student, and Public expenditure on education, respectively. With the coefficient of determination R^2=0.705, the model reveals that 70.5% of the variability of Employment rates of recent graduates is explained by the given combination of explanatory variables. The whole model is significant at 0.01 level of significance.

4. CONCLUSION

Under the influence of the Human Capital theory, which elaborated on the functions of education in promoting employment and increasing employment incomes (Schultz, 1961), many countries devoted major efforts to educational development in the 1960s and 1970s (Chen & Wu, 2007). This paper investigates the influence of educational development indicators on the employment rate of recent graduates.

We extracted the 12 financial and non-financial indicators that served as explanatory factors in the least squares linear regression model. We used a backwards method to exclude non-significant variables from the model automatically. The final model was obtained after eight iterations. Our findings suggest that we have both financial and non-financial indicators influencing the employment rate of recent graduates in the final model. The most influential non-financial indicators were those that measure a number of young people that continued their education and foreign languages spoken by the nation while the financial factors were public expenditure on education and annual expenditure per student.

The given model explains 70.5% of the variability of graduates' employment rates. Accordingly, it is important to recognise that the employment does not depend solely on the educational development factors but from the series of other factors (Barnes & Brown, 2016), such as the area of employability, social conditions, but their individual features as well. For example, it is discussed that graduates of entrepreneurial education should be better prepared for the employment market and the rapidly changing economic environment, and thus, more productive, even in wage employment (Kucel, Róbert, Buil, & Masferrer, 2016). Future research directions might imply the further and more detailed analysis of influential factors on the employment.

REFERENCES

Altbach, P. G. (2003). Introduction to Higher Education theme issue on the Academic Profession in Central and Eastern Europe. *Higher Education*, *45*(4), 389–389. http://doi.org/10.1023/A:1023949808905

Andjelkovic Labrovic, J., Bijelic, A., & Milosavljevic, G. (2014). Mapping Students' Informal Learning Using Personal Learning Environment. *Management - Journal for Theory and Practice of Management*, *19*(71), 73–80. http://doi.org/10.7595/management.fon.2014.0009

Barnes, S.-A., & Brown, A. (2016). Stories of learning and their significance to future pathways and aspirations. *British Journal of Guidance & Counselling*, *44*(2), 233–242. http://doi.org/10.1080/03069885.2016.1145194

Benkovic, S., & Dobrota, M. (2012). Application Of Teaching Methods And Techniques AT Serbian Universities: Progress Over Time. *Management - Journal for Theory and Practice of Management*, *17*(63), 35–44. http://doi.org/10.7595/management.fon.2012.0007

Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (2011). *Education as a Lifelong Process*. New York: Springer.

Blundell, R., Dearden, L., Goodman, A., & Reed, H. (1997). *Higher education, employment and earnings in Britain*. London: The Institute for Fiscal Studies. http://doi.org/10.1920/re.ifs.1997.0052

Brown, P. (1995). Cultural Capital and Social Exclusion: Some Observations on Recent Trends in Education, Employment and the Labour Market. *Work, Employment & Society*, *9*(1), 29–51. http://doi.org/10.1177/095001709591002

Brown, S., & Sessions, J. (1998). Education, Employment Status and Earnings: A Comparative Test of the Strong Screening Hypothesis. *Scottish Journal of Political Economy*, *45*(5), 586–591. http://doi.org/10.1111/1467-9485.00113

Chapman, D. W., & Miric, S. L. (2009). Education Quality in the Middle East. *International Review of Education*, *55*(4), 311–344. http://doi.org/10.1007/s11159-009-9132-5

Charles, N. (2015). When Training Is Not Enough: Preparing Students for Employment in England, France and Sweden. *Sociologie Du Travail*, *57*, e1–e21. http://doi.org/10.1016/j.soctra.2015.09.008

Chen, Z., & Wu, Y. (2007). The relationship between education and employment: A theoretical analysis and empirical test. *Frontiers of Economics in China*, *2*(2), 187–211. http://doi.org/10.1007/s11459-007-0010-4

Dobrota, M., & Benkovic, S. (2014). Comparing 'Ex-Cathedra' and IT-Supported Teaching Methods and Techniques: Policy of Teaching Practice. *Croatian Journal of Education*, *16*(Sp.Ed.3), 91–108.

Dobrota, M., Jeremić, V., Bulajić, M., & Radojičić, Z. (2015). Uncertainty and Sensitivity Analyses of PISA Efficiency: Distance Based Analysis Approach. *Acta Polytechnica Hungarica*, *12*(3), 41–58.

Dobrota, M., Martic, M., Bulajic, M., & Jeremic, V. (2015). Two-phased composite I-distance indicator approach for evaluation of countries' information development. *Telecommunications Policy*, *39*(5), 406–420. http://doi.org/10.1016/j.telpol.2015.03.003

Dobrota, M., Savic, G., & Bulajic, M. (2015). A New Approach to the Evaluation of Countries' Educational Structure and Development: The European Study. *European Review*, *23*(04), 553–565. http://doi.org/10.1017/S1062798715000277

Eurostat. (2016). European statistics. Retrieved March 10, 2016, from http://ec.europa.eu/eurostat/web/main/home

Evans, C. (2016). Moving away or staying local: the role of locality in young people's "spatial horizons" and career aspirations. *Journal of Youth Studies*, *19*(4), 501–516. http://doi.org/10.1080/13676261.2015.1083955

Fleming, T. (2011). *Models of Lifelong Learning: An Overview*. Oxford University Press. http://doi.org/10.1093/oxfordhb/9780195390483.013.0018

Gujarati, D. (2002). *Basic Econometrics*. McGraw-Hill/Irwin.

Hirsto, L., Lampinen, M., & Syrjäkari, M. (2013). LEARNING OUTCOMES OF UNIVERSITY LECTURERS FROM A PROCESS-ORIENTED UNIVERSITY PEDAGOGICAL COURSE. *Trames. Journal of the Humanities and Social Sciences*, *17*(4), 347. http://doi.org/10.3176/tr.2013.4.03

Horvat, A., Dobrota, M., Krsmanovic, M., & Cudanov, M. (2015). Student perception of Moodle learning management system: a satisfaction and significance analysis. *Interactive Learning Environments*, *23*(4), 515–527. http://doi.org/10.1080/10494820.2013.788033

Horvat, A., Krsmanovic, M., Dobrota, M., & Cudanov, M. (2013). Students` Trust in Distance Learning:

Changes in Satisfaction and Significance. *Management - Journal for Theory and Practice of Management*, *18*(69), 47–54. http://doi.org/10.7595/management.fon.2013.0026
Hout, M. (2012). Social and Economic Returns to College Education in the United States. *Annual Review of Sociology*, *38*(1), 379–400. http://doi.org/10.1146/annurev.soc.012809.102503
Knewstubb, B. (2016). The learning–teaching nexus: modelling the learning–teaching relationship in higher education. *Studies in Higher Education*, *41*(3), 525–540. http://doi.org/10.1080/03075079.2014.934802
Kucel, A., Róbert, P., Buil, M., & Masferrer, N. (2016). Entrepreneurial Skills and Education-Job Matching of Higher Education Graduates. *European Journal of Education*, *51*(1), 73–89. http://doi.org/10.1111/ejed.12161
Lutz, W., & KC, S. (2011). Global Human Capital: Integrating Education and Population. *Science*, *333*(6042), 587–592. http://doi.org/10.1126/science.1206964
Narula, S. C., & Wellington, J. F. (1982). The Minimum Sum of Absolute Errors Regression: A State of the Art Survey. *International Statistical Review / Revue Internationale de Statistique*, *50*(3), 317. http://doi.org/10.2307/1402501
Nguyen, T. T. H., & Walker, M. (2016). Sustainable assessment for lifelong learning. *Assessment & Evaluation in Higher Education*, *41*(1), 97–111. http://doi.org/10.1080/02602938.2014.985632
Olaniyan, D. A., & Okemakinde, T. (2008). Human Capital Theory: Implications for Educational Development. *Pakistan Journal of Social Sciences*, *5*(5), 479–483.
Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, *25*(1), 159–184. http://doi.org/10.1257/jep.25.1.159
Pedersen, H. S. (2016). Are PhDs winners or losers? Wage premiums for doctoral degrees in private sector employment. *Higher Education*, *71*(2), 269–287. http://doi.org/10.1007/s10734-015-9901-y
Purcell, K., & Quinn, J. (1996). Exploring the education-employment equation in hospitality management: a comparison of graduates and HNDs. *International Journal of Hospitality Management*, *15*(1), 51–68. http://doi.org/10.1016/0278-4319(96)00002-3
Radenkovic, B., Despotovic Zrakic, M., Bogdanovic, Z., Barac, D., & Labus, A. (2014). IT Education as an Opportunity for Uprising of Serbian Economy. *Management - Journal for Theory and Practice of Management*, *19*(71), 57–72. http://doi.org/10.7595/management.fon.2014.0014
Rencher, A., & Christensen, W. (2012). Chapter 10, Multivariate regression – Section 10.1, Introduction. In *Methods of Multivariate Analysis* (p. 19). John Wiley & Sons.
Roche, S. (2015). Of integration and disintegration: The “hows” and “whys” of lifelong learning in a rapidly changing world. *International Review of Education*, *61*(5), 577–583. http://doi.org/10.1007/s11159-015-9521-x
Schultz, T. W. (1961). Investment in human capital. *American Economic Review*, *51*(1), 1–17.
Seal, H. (1967). Studies in the History of Probability and Statistics. XV The historical development of the Gauss linear model. *Biometrika*, *54*(1-2), 1–24. http://doi.org/10.1093/biomet/54.1-2.1
Volles, N. (2016). Lifelong learning in the EU: changing conceptualisations, actors, and policies. *Studies in Higher Education*, *41*(2), 343–363. http://doi.org/10.1080/03075079.2014.927852
Waters, J. L. (2009). In Pursuit of Scarcity: Transnational Students, “Employability”, and the MBA. *Environment and Planning A*, *41*(8), 1865–1883. http://doi.org/10.1068/a40319
Weiss, F., Klein, M., & Grauenhorst, T. (2014). The effects of work experience during higher education on labour market entry: learning by doing or an entry ticket? *Work, Employment & Society*, *28*(5), 788–807. http://doi.org/10.1177/0950017013506772

IMPACT OF LARGE EXPORTERS' OWNERSHIP STRUCTURE ON THE GROWTH OF EXPORT IN SERBIA

Zoran Jeremić[1]

[1]Singidunum University, Faculty of Business Economics
*Corresponding author, e-mail: zjeremic@singidunum.ac.rs

Abstract: *In this paper, we will analyse 100 largest exporters in Serbia in 2014 primarily from the standpoint of ownership structure i.e. whether the exporting company was founded by a foreign or domestic legal entity/ person as well as whether it operates as a limited liability company or a joint stock company whose shares are traded on the Belgrade Stock Exchange. The ownership structure is analysed in order to determine the extent of foreign i.e. domestic private or state-owned capital involved in these companies. The aim is identify largest exporters among domestic companies in order to determine what industries they operate in and the group of companies they belong to, as it provide insight into which domestic companies that are competitive on the international market.*

Keywords: *export, industrial sectors, competitiveness, ownership structure*

1. INTRODUCTION

The significance of export in a small country with an impoverished market is undisputable along with strategic ambition to create a suitable environment for further development of export. Even though there is a general consensus among domestic public and in strategic documents with this statement, the practice shows the presence ofa series of questions which haven't been resolved and different views on how to achieve the proclaimed goal of strengthening export. High rate of consent on this goal is also evident among international organisations such as World Bank and IMF whose publications and documents prioritise growth of export in Serbia. The subject matter of this paper is analysis of the ownership structure of 100 largest exporters in Serbia, with special focus on exporters founded by domestic legal entities and natural persons in the form of joint stock companies, limited liability companies and sole proprietorships.

For the purpose of this analysis, a database was created on actual export and import of 100 largest exporting companies based on data available to Customs Administration and we processed their balance sheets and income statements based on data obtained from the Serbian Business Registers Agency. In order to ensure comparability of export data with data contained in balance sheets and income statements, amounts expressed in dinars were converted into euros by applying average annual exchange rate for income statement items and the exchange rate prevailing on the last day of the year for balance sheet items. Ownership structure was determined by using websites of Serbian Business Registers Agency and of the companies which are the subject matter of this analysis as well as the website of Belgrade Stock Exchange for the listed joint stock companies. Database covers the period from 2008 to 2014, with 2008 being regarded as the base year for analysis since it is the last year for which comprehensive data on companies' balances is still available, though the focus of analysis will be placed on 2014.

The group of top 100 exporters is relevant for our research because it had more than half of the share in the commodity export in 2014. Moreover, the share which top 100 exporters have in total export is continuously on the rise. While this share was 41% in 2008, in 2014 it grew to 55%. This group achieved total export equal to €6,203 million euros, total assets of €11,341 million euro, capital worth €4,745 million euro, operating income of €11,629 million euro, net profit of €716 million euro while the total number of their employees amounted to 68,690 people. These indicators show that this group certainly is not a relevant sample only in terms of exporting companies but also for analysing the state of Serbian economy. Aggregate data on 100 largest exporters are divided into three groups according to who is the majority shareholder: foreign company, domestic legal entity or natural person, or the state. After this, we analysed the share that each of these subgroups had in the total export of 100 largest exporters. What is particularly important is to identify the industrial sectors and groups they operate in, ratio of export and import and business performances of these companies over the observed period.

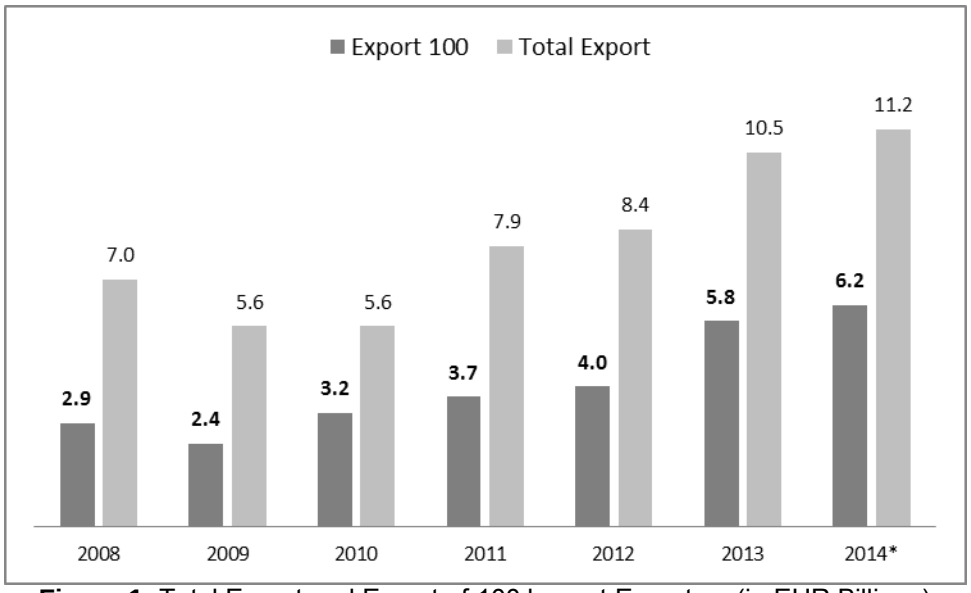

Figure 1: Total Export and Export of 100 largest Exporters (in EUR Billions)
Source: Republic of Serbia, Customs Administration, summarised by author

By observing the period between 2008 and 2014, we noted positive changes in the growth of export, export structure and the import coverage ratio. Export has recorded faster growth than import with a slightly positive change in terms of its structure. A certain number of new companies emerged on the market in the observed period primarily export-oriented with positive effect on the reindustrialisation, mainly drawn to this market by government subsidies and low labour costs. Simultaneously, some of the leading exporting companies recorded a dramatic downfall in their business at the time of economic crisis which resulted in the drop of export as well.

However, when it comes to export in 2014, the positive trends observed in the previous years were continued. By analysing the fluctuation of export among the 30 largest net exporters, it was observed that net exporters are mainly companies in the processing industry, primarily metal processing industry and electric power industry followed by agricultural production (maize, sugar) and by military industry on a smaller scale. Having in mind the rising trend in export recorded by the largest exporters, special attention should be paid to their analysis by answering the following question: Are there any potential "stars" among Serbian exporters which could contribute to faster growth of the country's export as had already happened in several other countries (see: Export stars...) particularly if such growth would imply integration of smaller domestic companies into their production chain instead of providing incentive for each company individually to enhance its export. What is worrying when it comes to large exporters is the fact that they are simultaneously large importers as well, and for this reason, it would be interesting to find out the group in which net export is predominant. It could be reasonably assumed that domestic companies are in advantage in respect of net export and thus their total effect on the country's economy is more useful,which is why this hypothesis will be tested here as well.

2. FOREIGN COMPANIES AMONG 100 LARGEST EXPORTERS

By dividing the largest exporters according to their country of origin into group of companies owned by foreign companies and exporters whose majority shareholders are domestic privately-owned or state-owned companies, we note a strong dominance of foreign countries both in terms of their number (63) and their share in the total export (77.1%). Foreign companies' share in the export realised by 100 largest exporters in 2014 was 77.1% i.e. € 4.8 billion euro, employing 49,313 workers i.e. 71.8% of the group's employees. This group also includes large net exporters, for example in the area of car spare parts industry, which came to Serbia primarily because of high government subsidies and low labour costs and partially due to presence of FIAT company. The next group includes home appliances and machine manufacturers – three Gorenje factories.

Table 1: Ownership structure of 100 largest exporters

Ownership structure	Number of exporters	of which: net exporters
Total number of exporters	100	67
a. Owned by foreign companies	63	40
b. Owned by domestic companies	30	20
c. Owned by state	7	7
of which in restructuring	4	4

Source: Republic of Serbia, Customs Administration, summarised by author

Though there had been plenty of off-target and unnecessary government subsidies, the effects are visible in cases where foreign investors oriented to exports and new technologies were attracted to Serbia. However, the issue here is that the economic climate was not further improved but, on the contrary, macroeconomic flows and the business conditions were dramatically aggravated so the subsidies were necessary to attract the investors because without such subsidies we would have obtained lesser level of foreign investments. However, attracting foreign investors with subsidies, as a compensation for poor economic environment, with employment of unqualified and poorly paid labour force, cannot be used as a permanent model to increase export and the solution should be sought in improvement of the business and investment climate and reduction of subsidies can then go hand in hand with such improvement.

Table 2: Export, Import (in EUR Millions) and Number of Employees

Companies	Export	Import	Net export	No. of empl.
a. Owned by foreign companies	4.780,1	5.828,7	-1.048,6	49.313
b. Owned by domestic companies	836,1	573,8	262,3	10.422
c. Owned by state	587,1	251,8	318,1	8.955
Total	6.203,3	6.654,3	-468,2	68.690

Source: Republic of Serbia, Customs Administration, and Agency for Business Registers, data processed by author

In addition to their high concentration among top 100 exporters, the growth of the group of 5 to 7 largest companies is conspicuous and they also record the largest growth of export. The general trend in export shows that the concentration of large companies is growing as they are more competitive and better prepared to face the international competition whereas smaller companies have issues with emerging on the foreign markets and maintaining their market position. Large exporters are also large importers, with third of exporters having higher import than export. These are vital facts. For this reason, we will identify the most significant net exporters, from the standpoint whether they are owned by foreign or domestic companies, as well as what industry they operate in. Net exporters namely have significantly higher contribution to the newly added value of gross domestic product and more significantly improve the country's balance of trade and the balance of payments.

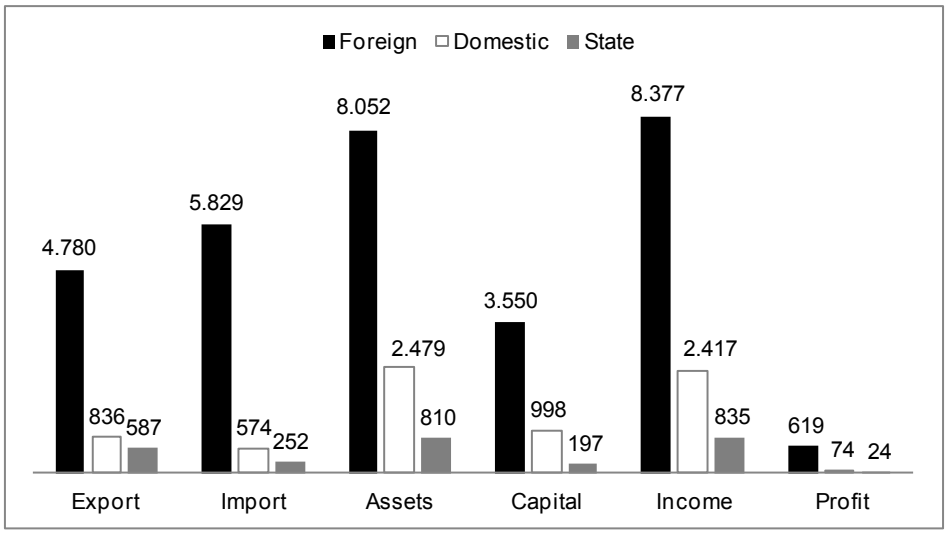

Figure 2: Foreign, Domestic and State-Owned Exporters (in EUR Millions)

Source: Republic of Serbia, Customs Administration, and Agency for Business Registers, data processed by author

3. COMPANIES OWNED BY DOMESTIC LEGAL ENTITIES AND NATURAL PERSONS

Out of 30 companies owned by domestic legal entities and natural persons which are among 100 largest exporters, 15 are agricultural produce processing companies (sugar beet, soya, oils, meat, animal feeds, and dairy products). This group also includes one company which is a successful metal-processing company, one furniture company and one bankrupt company while the rest are mainly involved in the export of primary agricultural products and in food import and processing for their own needs. So, export of primary and secondary agricultural products is dominant while export of other industrial products is on a minimum level. Largest agricultural produce processing companies include 4 sugar-processing companies, 4 fruit-processing companies, 2 animal-feed processing companies and one soya and one milk-processing company.The largest exporters were 4 sugar processing companies which are also the largest net exporters as import plays a very small role in their business. The structure dominated by agricultural produce manufacturing and processing companies reveals that it is difficult to be competitive and to penetrate the foreign markets in other non-agricultural industrial sectors. This observation particularly applies to the medium and small size companies, with 90% of them having recorded no export at all and having no potential to expand to foreign markets. They can find their chance for export if connected to exporting companies which are competitive on foreign markets.

4. STATE-OWNED EXPORTERS

The group of large exporters owned by the state is dominated by factories operating in military industry such as *Prvi Partizan*, Užice and *Zastava oružje*, Kragujevac. The broad list of exporters includes several factories which carry out their export via *Jugoimport SDRP*. Having in mind their export potential, we may expect further growth of export in military industry. HIP-Petrohemija is also an important exporter but for the time being there is no clear solution on sight for its strategic partner.

When it comes to export, *Železara Smederevo* is an ever-present topic. If the agreement concluded with the Chinese partner were to be realised, it would certainly be a huge encouragement for Serbian export and great relief on the country's budget. Železara can be profitable only if investments are oriented to highly processed products as was announced by the potential strategic partner. The balance sheet and income statement data of *RTB Bor* were not included in this analysis, but only export achieved by RTB Invest. The analysis of this company is far too complex since the data are controversial and non-transparent. However, it is certain that the recovery of *Železara Smederevo* and *RTB Bor* would largely depend on the future recovery and growth of the global economy i.e. on the increase of prices of export products since these companies cannot stay profitable with the prices remaining at the current level. Out of 7 companies owned by the state,

four are in the restructuring process and it remains uncertain whether a solution would be found for them before the deadline expires.

Table 3: Balance Sheet Items of 30 Largest Domestic Exporters (in EUR Millions)

Num.	Rank		Total Income		Net Income		No. of employees		ROE	
	in 100	Company name	2014	2013	2014	2013	2014	2013	2014	2013
1	23	SOJAPROTEIN AD BEČEJ	99	101	2.1	0.0	387	416	2.3%	0.0%
2	26	VICTORIAOIL AD ŠID	108	117	0.0	0.0	282	294	0.0%	0.0%
3	27	SUNOKO DOO NOVI SAD	154	206	6.0	45.5	469	478	4.3%	33.3%
4	38	MK COMMERCE DOO NOVI SAD	134	125	3.2	0.4	100	79	11.9%	18.3%
5	46	TE-TO, Senta	59	76	0.0	5.4	174	190	8.7%	1.9%
6	51	DOO ĆIRIĆ I SIN, SAKULE	32	25	1.7	1.2	19	18	13.8%	2.0%
7	33	UMKA D.O.O. UMKA	55	57	3.6	5.3	324	298	17.5%	14.2%
8	48	DELTA AGRAR, Beograd	168	180	2.6	6.6	273	278	0.0%	18.9%
9	63	FARMAKOM FINANCE, Šabac	0	5	0.0	0.5	0	43	6.6%	9.8%
10	79	MAT-AGRO, Novi Sad	58	41	1.4	0.3	13	10	2.1%	5.3%
11	71	ŠEĆERANA CRVENKA	53	63	0.0	6.4	207	210	10.9%	8.6%
12	89	ALTIVA DOO, BEOGRAD	0	17	0.0	0.2	0	7	24.5%	29.0%
13	93	AD FABRIKA ŠEĆERA, ŽABALJ	42	39	0.0	1.3	189	189	13.0%	7.3%
14	96	ELIXIR FOOD, Šabac	28	13	0.3	0.0	125	121	0.0%	5.8%
15	69	FORMA IDEALE,KRAGUJEVAC	56	56	3.4	3.4	1121	1042	15.7%	36.1%
16	47	AGROGLOBE, NOVI SAD	142	117	1.5	2.0	131	114	16.1%	16.7%
17	60	ALFA-PLAM AD, Vranje	40	36	6.5	3.5	718	742	3.2%	10.4%
18	88	GEBI DOO, PO ČANTAVIR	62	62	1.3	0.7	214	209	0.0%	21.9%
19	70	NECTAR DOO, BAČKA PALANK	68	69	1.5	5.1	638	667	12.7%	7.1%
20	57	DRENIK ND DOO BEOGRAD,	57	54	6.6	6.5	462	446	0.0%	0.0%
21	86	NOVKABEL AD, NOVI SAD	30	28	0.0	0.0	558	588	0.0%	0.0%
22	82	VICTORIA LOGISTIC DOO, NOVI	307	285	0.0	0.0	257	291	28.2%	60.9%
23	91	AD NEOPLANTA, NOVI SAD	72	60	3.2	2.0	742	720	0.0%	0.0%
24	36	SWISSLION DOO, BEOGRAD	75	83	5.9	1.3	873	879	0.0%	0.0%
25	43	PEŠTAN DOO, BUKOVIK	54	54	9.1	6.3	644	608	8.8%	4.9%
26	72	ELIXIR ZORKA, ŠABAC	88	46	4.0	2.0	221	201	0.0%	13.6%
27	90	DOO ALMEX, PANČEVO	85	88	6.0	8.4	272	233	12.2%	20.7%
28	66	PROMIST DOO,NOVI SAD	153	112	3.9	7.8	66	57	5.8%	3.6%
29	74	CENTRO-SPICE, Surčin	87	54	0	0	715	508	0.0%	8.8%
30	75	DELTA SPORT,BEOGRAD	51	58	0	0	228	348	1.3%	0.0%
		Total	2416.9	2326.6	73.8	122.3	10422.0	10284.0	7.2%	11.8%
		Source: Agency for Business Registers, data processed by autor								

Table 4: State-Owned Largest Exporters (in EUR Millions)

Num.	Rank	Company name	Export	Import	Income		Net Profit		No. of empl.	
	in 100		2014	2014	2014	2013	2014	2013	2014	2013
1	4	Železara, Smederevo	201	177	248	189	0	0	5.025	5.036
2	7	Hip-Petrohemij, Pančevo	148	21	257	377	0	0	1.738	1.787
3	18	RTB Invest	81	15	124	153	0	0	3	3
4	21	Jugoimport-SDPR JP Bgd.	67	12	73	83	6	5	379	362
5	29	Prvi Partizan, Užice	44	19	47	54	5	5	948	935
6	50	SMATSA DOO	29	8	86	82	12	9	862	882
7	94	Zastava oružje, Krag.	17	6	27	34	0	0	1.958	2.189
		TOTAL	587	258	863	973	24	19	10.913	11.194

Source: Republic of Serbia, Customs Administration, and Agency for Business Registers, data processed by author

5. CONCLUSION

Analysis of 100 largest exporters revealed several of their important characteristics: the export is dominated by companies in foreign ownership both in terms of their number and the value of their export. High concentration of export is also evident since 100 largest exporters had a share of 55% in total export in 2014. Even more important is the concentration of 5 to 7 largest exporters which continuously increase their export. Large exporters founded by foreign companies are also large importers with only 67 companies characterised by export exceeding import. In this respect, companies with domestic capital are more useful from the perspective of newly added value, with the most significant group of companies being agricultural product processing companies.

Out of 30 companies owned by domestic legal entities and natural persons which are among 100 largest exporters, 15 are agricultural products processing companies (sugar beet, soya, oils, meat, animal feeds, and dairy products). This group also includes one company which is a successful metal-processing company, one furniture company and one bankrupt company, while the rest are mainly involved in the export of primary agricultural products and in food import and processing for their own needs. So, export of primary and secondary agricultural products is dominant whereas export of other industrial productsison a minimum level. When we add 7 exporting companies owned by the state and operating in the ore-processing and military industry, we practically summarise the total export potential in domestic ownership. If Železara, RTB Bor and Petrohemija find their strategic partners in the future, which is a pre-condition for their survival, the state will retain the ownership only over the exporting companies operating in military industry. It should certainly be noted that the state is the co-owner of two large companies whose majority shareholder is a foreign company i.e. NIS a.d. Novi Sad and Air Serbia, and this model should not be excluded as a solution for the rest of the state-owned companies which are large exporters.

When it comes to private companies with domestic capital, they have a one third share in the top 100 exporters but their export value ranks them at the bottom portion of the list. In addition to that, they are mainly characterised by poorer business performance than the companies owned by well-known global companies. Even though some of our domestic companies have managed to come to the spotlight (Alfa plam), for domestic exporters it is an increasingly difficult battle to fight off foreign competition as those foreign companies bring with them knowledge, captured foreign markets to which they export their products, better corporate governance, bigger budgets for research and development and expert personnel. Domestic companies, however, are frequently exposed to unfair competition as their foreign competitors receive government subsidies just to come to Serbia. Domestic medium and small size enterprises are facing even greater difficulties to fight off comparative advantages of large foreign companies. What is an even bigger issue for them is to have sufficient demand and obtain necessary guarantees while they do not have neither financial nor development capacities to conquer foreign markets and oppose the mighty competitors. For this reason, further enlargement of companies may be expected in the upcoming period via takeover of some of current exporters by the stronger domestic companies but more so by foreign companies. The example of sugar processing companies that are large exporters sheds a lot of light on the issue. Due to fiercer competition on the European market, which is the main export market for Serbian sugar processing companies, further consolidation will be necessary because small sugar processing companies which are far removed from raw material base, falling behind in respect of technology while using more expensive energy sources, will not be able to survive on the market. Fruit and vegetable processing industry showed great possibility for export growth but it also requires well-thought through incentive for connecting primary production with domestic processing companies. This would greatly contribute to the growth of export, more even regional development and strengthening of domestic processing industry in the segment in which it can be competitive since Serbia is the global leader only in the quality and production of certain agricultural products such as raspberry, blackberry, apple... This sort of incentive does not mean that Serbia should not pay greatest attention to information industry and to the strengthening of high technology sectors but in this instance,the changes to the system of education would be of greater significance than financial incentives. Having in mind that export is a precondition for development and survival of a small and poor economy such as Serbian, it is vital to support exporting companies in all shapes and forms: by creating better business climate, by targeted incentives and assistance to industrial sectors in which Serbia can be competitive. The example of 100 largest exporters showed that the choice of competitive industrial sectors is very narrow. The most realistic chances lie in the development of processing industry based on high quality primary agricultural raw material. This is precisely where the added value is the highest as the share of import in this segment is the smallest, which makes 15 food processing companies also the largest net exporters. The wood and metal processing industries also have good chances for increasing export. In addition to food-

processing industry, we also have information industry where we can be competitive while the mining industry and tourism also provide good basis for further growth of export.

Attracting foreign investments with government subsidies as a compensation for poor economic climate, along with employment of unqualified and poorly paid labour forrce, cannot be used as a permanent model to increase export even though these incentives were necessary in current circumstances. Companies owned by foreign companies are primarily needed to bring new knowledge and new technologies and should employ domestic labour force having highest education and if they do not meet these criteria, any subsidies to such companies would be very questionable. The solution to the problem lies, naturally, in the continuous improvement of the business and investment climate and the reduction in subsidies can then go hand in hand with such improvement.

When it comes to changing the export structure, it is necessary to ensure that the main articles exported abroad are not agricultural produce such as maize, raspberries, apples and copper and such change in the export should be achieved systematically, having long-term goals in mind, in order to increase the share of highly processed products in export, which implies the use of new technologies,as well as to create economic climate which would favour net exporters.

REFERENCES

European Commission (2014*), Progress Report for Serbia*

Hausmann, R., Hwang, J. and Rodrik, D. (2006) *What You Export Matters.*Working Paper, Center for International Development, Harvard University,

Hesse, H. (2008). *Export Diversification and Economic Growth,* The World Bank On behalf of the Commission on Growth and Development,

Hidalgo, C., Klinger, B., Barbasi, A., and Hausmann, R. (2007)*: The Product Space Conditions the development of nations,* Science, 317-329.

Jeremić, Z., Milojevic, M. and Terzć, I. (2015): Business Performance of the Largest Exporters in Serbia During the Period 2008-2014, Ekonomika preduzeća, 5-6, 293-305.

Jeremić, Z., Milojevic, M. (2015): *The Most Successful Exporters and Importers in the Period 2008-2014,* The Economy and Finance Board 2015, Business Info Group, Kopaonik business forum, 6-37

Jeremic, Z., and group of authors. (2008)*: Stock Market in Serbia - Analysis of the Sector and 100 companies from the Belgrade Stock Exchange,* Chamber of Commerce of Serbia & Economist media group.

Serbia's Real Sector performance, CEVES, (2015),

Trbovic,A,Radosavljevic,G&Subotic,J.*,Competitivness Gap in Serbia's Metals and Electronics Industry,* Ekonomika preduzeća,(2012), 301- 315

The strategy to support the development of small and medium enterprises, entrepreneurship and competitiveness for the period from 2015.to 2020, Službeni glasnik RS 35/2015, 10

Food and Agriculture Organization of the United Nations, *The State of Agricultural Commodity Markets 2015/2016,* Rome 2015

World Bank (2012*), Republic of Serbia, The Road to Prosperity: Productivity and Exports,* Country Economic Memorandum*,22-43*

World Bank, (2014),*Rebalansing Serbia's Economy – Improving Competitiveness, Strengthening the Private Sector and Creating Jobs*

Why Invest in Serbia, Serbian Development Agency*, (2016), retrived from www.ras.gov.rs*

THE ALTERATION OF U21 RANKING METHODOLOGY: FROM EXPERT-DRIVEN TO DATA-DRIVEN WEIGHTING SCHEME

Milica Maričić*[1], Milica Bulajić[1], Marina Dobrota[1]
[1]University of Belgrade, Faculty of Organizational Sciences, Laboratory of Statistics
*Corresponding author, e-mail: milica.maricic@fon.bg.ac.rs

Abstract: *Indicator-based measurement of higher educational institutions and higher education systems' performance has started to play a significant role in defining national education policies and reforms. However, the issues surrounding the process of assessment and ranking are complex and far-reaching. Therefore, the controversy about ranking methodologies remains, whereas the critical pitfalls are the weighting scheme and the aggregation method. This study aims at tackling one of the two critical pitfalls by proposing a data-driven weighting scheme of the U21 ranking of national higher education systems (U21) using the Composite I-distance Indicator (CIDI) methodology. The CIDI weights proved to be more balanced and less oriented on the higher education systems' output. The presented approach and results can be a foundation for further academic research on university or higher education systems rankings that will be independent of subjectively formed weighting factors.*

Keywords*: U21 Ranking, I-distance method, CIDI Methodology, composite index, ranking of countries, higher education system*

1. INTRODUCTION

University rankings have long been employed, first on the national scale and later on a global scale (Marginson & van der Wende, 2007). In the last decade, the interest of various stakeholders for university rankings had risen. Firstly, the growing population of students and researchers needed insights on the quality performance of higher educational institutions (HEIs) (Saisana, D'Hombres, & Saltelli, 2011). On the other hand, the results of university rankings became a valuable source of information for policymakers. Ranks of national universities have become a symbol of national achievement and prestige, but also an engine of economic growth (Marginson & van der Wende, 2007). Therefore, university rankings can be used to trigger debates and policy reforms on a national level (Saisana & D'Hombres, 2008). Consequently, the rankings influence the higher educational institution's mission, strategy, recruitment procedures, and teaching process; i.e. nearly every aspect of the educational process (Harzing & Mijnhardt, 2015). Steadily, global university rankings are integrating into higher education, and there are various benefits from such occurrence. Just some of them are a positive impact on the number of potential applicants, opportunities to collaborate with the industry or with international higher educational institutions, additional government and non-government funding, and attraction of prospective students and academics.

Putting aside the potential benefits of the global university rankings, one should bear in mind the concerns the academic community has risen on this matter. Namely, academics have often criticized some of the rankings for their statistical inaccuracy (Huang, 2012), possible unfavourable impact on the overall performance of universities (Cheng, 2012), and for their influence on the educational policy (Hazelkorn, 2014). Saisana and D' Hombres (2008) outlined the fact that many ranking methodologies fail to follow accepted methodologic rules when selecting indicators and afterwards aggregating them into a composite index.

University rankings proved to be very useful, especially for government representatives in decision-making on the national level. However, users of university rankings, and especially the government, should have in mind the fact that although higher education system consists of separate institutions, it is necessary to have a holistic approach to ranking. Therefore, a new direction in the development of educational rankings emerged: the rankings of higher education systems. Martin and Sauvageot (2011) refer to the national higher education system as an integral part of the modern economy and social structure. As such, it should give as many people as possible a chance to gain analytical skills and technical knowledge that will enable them to respond to the competitive conditions in the labour market. Promotion and long-term investments in a small number of institutions will not raise the overall quality of the educational system. Also, a small number of institutions that are advancing faster cannot respond to the overall national needs for highly educated staff. In addition, it is necessary to show the public how their country is positioned on the global map of higher education systems (Williams et al., 2013).

For the foregoing reasons, in the last few years, additional efforts have been placed to devise a ranking of higher education systems. In 2012, one of them appeared: the Universitas Ranking of National Higher Education Systems (U21) devised by Universitas 21. Since then, the index attracted attention and positive and negative critique on its methodology. According to Soh (2014), although it has a clear theoretical base and more trustworthy data sources compared to other ranking methodologies, its weights show discrepancy as they are subjectively assigned. Also, Millot (2014) concludes that U21 has the potential to differentiate itself from university rankings and become a valuable supplement in assessing higher education.

The aim of this paper is to analyze the subjective weighting scheme of the U21 ranking and to propose a new one that will rely on a statistically defined framework. The multivariate statistical analysis used to perform such a task is the Composite I-distance Indicator (CIDI) methodology (Dobrota et al., 2016).

The paper is organized as follows: the second section features the U21 ranking methodology, followed by the statistical method used to perform the analysis: the CIDI methodology. The research results are presented and elaborated in Section 4, while we provide the concluding remarks in the final chapter.

2. UNIVERSITAS U21 RANKING OF NATIONAL HIGHER EDUCATION SYSTEMS

Universitas 21, an international group of 25 world universities, devised in 2012 a composite index for ranking national higher education systems and named it U21. The basic idea of the Universitas 21 group was to draw attention from the ranking of universities and individual institutions on the ranking of the national higher education systems. To achieve its goal, the U21 composite index includes four categories (*Resources, Environment, Connectivity,* and *Output*) and 25 indicators. The index is calculated for 50 world countries, but due to the lack of data, the recreated database, which we used to perform the analysis, contains data for 48 countries; Brazil and Saudi Arabia had to be excluded from further research.

The data for 25 indicators, which are used for the computation of the ranking according to the latest methodology, is collected from reputable sources such as OECD, UNESCO, and the World Economic Forum (WEF) (Universitas 21, 2015). Table 1 provides the categories, indicators, effective weights, and codes of indicators, assigned by the authors for easier presentation in the rest of the manuscript.

The *Resources* category provides information on the material and immaterial pre-conditions for the core transformation processes in a higher education system. This dimension aims at measuring the government's readiness to invest in higher education and at investigating what are the consequences of their investment. Although resources are vital for a system to operate, they are not enough to ensure its efficient and proper functioning (Williams et al., 2013). To ensure the effective use of inputs, a regulatory *Environment* is needed. Lately, the idea of internalization became firmly embedded in the institutional mission statements, strategies, as well as national policy frameworks. Category *Connectivity* aims at measuring the relationship between the national higher education system and the industry and government. Finally, the last category (*Output*) includes the measurement of higher education systems output through the scientific output (research papers), the number of scientific workers, enrolment rate in higher education institutions, as well as the percentage of the population that has a college degree.

Although being a comprehensive ranking methodology obtained using mostly objective data, U21 ranking has its limitations. Namely, U21 methodology has been criticized for several reasons (Rauhvargers, 2014). One of the drawbacks is that the model is, in fact, a production model (Cheng, 2012). Another is that the index does not include measurements of teaching while it gives high significance to output (Millot, 2014). For this research, the most important critiques are the ones regarding the weighting scheme and the aggregation method. As many composite indexes, U21 faces the issue of expert-driven weights. The lack of information and justification of the assigned weights are one of the major drawbacks of this newcomer to the university ranking scene (Soh, 2014).

The aim of this research is the alteration of the U21 weighting scheme, thus making it data-driven. We propose the twofold Composite I-distance Indicator (CIDI) methodology to obtain objective and unbiased weights.

3. METHODOLOGY

3.1 Composite I-distance Indicator (CIDI) Methodology

A need for an unbiased ranking of countries was emphasized in the 1960s. Many ranking methodologies emerged, but one that stood out was the I-distance method devised by Ivanovic (1977). Namely, his method ranked countries according to their level of socio-economic development by using several indicators. What sets the I-distance apart from the others is its ability to bypass the process of weighting individual indicators as it aggregates indicators without assigning weights. Also, recent research shows that the I-distance can be applied with success not only in various fields but for both creation (Zornic et al., 2015) and evaluation of composite indexes (Dobrota et al., 2015).

Table 1 U21 categories, indicators, devised indicator codes, and their effective weights (Universitas 21, 2015)

Category	Indicator	Code	Effective weight
Resources (20%)	Government expenditure on tertiary education as a % of GDP, 2011	R1	5.0%
	Total expenditure on tertiary education as a % of GDP, 2011	R2	5.0%
	Annual expenditure per student (full-time equivalent) by tertiary education institutions at USD purchasing power prices, 2011	R3	5.0%
	Expenditure in tertiary institutions for research and development as a percent of GDP, 2012	R4	2.5%
	Expenditure in tertiary institutions for research and development per head of population at USD purchasing power prices, 2012	R5	2.5%
Environment (20%)	Proportion of female students in tertiary education, 2012	E1	1.0%
	Proportion of academic staff in tertiary institutions who are female, 2012	E2	2.0%
	Data quality	E3	2.0%
	Qualitative measure of the policy environment	E4	10%
	Responses to WEF survey question: “How well does the educational system in your country meet the needs of a competitive economy?”	E5	5.0%
Connectivity (20%)	Proportion of international students in tertiary education, 2012	C1	4.0%
	Proportion of articles with international collaborators	C2	4.0%
	Number of open access full text files on the Web, per head of population, 2009-2013	C3	2.0%
	External links that university web domains receive from third parties, per head of population, 2009-2013	C4	2.0%
	Rating of knowledge transfer between university and companies, 2013	C5	4.0%
	Percentage of university research publications co-authored with industry researchers, 2009-2012	C6	4.0%
Output (40%)	Total number of journal articles produced by higher education institutions, 2012	O1	10.0%
	Total articles produced by higher education institutions per head of population, 2012	O2	3.0%
	Average impact of articles measured by citations in 2012 to articles published in previous years using Karolinska Institute normalized impact factor	O3	5.0%
	Weighted Shanghai ranking scores for universities per head of population, 2014	O4	3.0%
	Average of nation’s best three universities on the 2014 Shanghai Jiao Tong ranking	O5	7.0%
	Tertiary enrolment rates as a percentage of the eligible population, 2012	O6	3.0%
	Percentage of population aged 24-64 with a tertiary qualification, 2012	O7	3.0%
	Number of researchers in the nation per head of population, 2012	O8	3.0%
	Unemployment rate of the tertiary educated compared with unemployment rates of those with only upper or post-secondary non-tertiary education, 2012	O9	3.0%

The I-distance method is based on calculating mutual distances between the entities being processed, whereupon they are compared to one another so as to create a rank (Ivanovic, 1977). In order to rank entities (in this case countries) by using the I-distance method, it is necessary to determine one entity as a referent in the observed set. The referent entity can be the minimal, maximal, or average observed or fictive value (Jeremic et al., 2011). In our analysis, the referent entity was the one with minimal values.

For a selected set of variables $X^T = \left(X_1, X_2, ...X_k\right)$ chosen to characterize the entities, the I-distance between the two entities $e_r = \left(x_{1r}, x_{2r}, ...x_{kr}\right)$ and $e_s = \left(x_{1s}, x_{2s}, ...x_{ks}\right)$ is defined as:

$$D^2\left(r,s\right) = \sum_{i=1}^{k} \frac{d_i^2\left(r,s\right)}{\sigma_i^2} \prod_{j=1}^{i-1}\left(1 - r_{ji.12...j-1}^2\right) \tag{1}$$

where $d_i(r,s)$ is the distance between the values of the variable X_i for e_r and e_s e.g. the discriminate effect

$$d_i\left(r,s\right) = x_{ir} - x_{is} \quad i \in \left\{1, ...k\right\} \tag{2}$$

σ_i is the standard deviation of X_i, and $r_{ji.12...j-1}$ is the partial coefficient of the correlation between X_i and X_j $\left(j < i\right)$ (Ivanovic, 1977).

The construction of the I-distance is an iterative process, which consists of several steps. First, the value of the discriminate effect of the first variable (the most significant variable, which encompasses the highest amount of information on the phenomena upon which the entities will be ranked) is calculated. Then, the value of the discriminate effect of the second variable that is not covered by the first one is calculated. This procedure is repeated for all of the observed variables in the data set (Jeremic & Jovanovic-Milenkovic, 2014).

The values obtained using the I-distance value are not comparable as they represent the distance between entities. Therefore, I-distance results could be employed so as to create a weighting scheme (Dobrota et al., 2016). Namely, besides aggregating variables, I-distance method can determine the relevance of the input criteria and so clearly point out variables which are important for their contribution to the final rank (Jeremic et al., 2011). Namely, Pearson's correlation coefficient between each indicator and the I-distance value provides additional information about the importance of each indicator for the ranking process. Therefore, weights can be determined as the ratio of the indicator's correlation coefficient and the sum of all correlation coefficients with I-distance value. The proposed methodology is called the Composite I-distance Indicator (CIDI) methodology. The equation for determining weights is:

$$w_i = r_i \Big/ \sum_{j=1}^{k} r_j \tag{3}$$

where $r_i(i = 1, ..., k)$ is the Pearson's correlation coefficient of variables with the I-distance value (Dobrota et al., 2016; 2015). The sum of the obtained weights equals 1 at both indicator and domain level. Therefore, applying the newly obtained weights on the original data provides unbiased and comparable results and rankings.

3.2 Twofold CIDI Approach

The twofold approach for evaluation has repeatedly been used in operational research and statistics (for example Simar & Wilson, 2007). However, the twofold I-distance approach has been recently introduced by Maricic and Kostic-Stankovic (2016) and Jovanovic-Milenkovic et al. (2015). The basic idea of the fresh approach was to follow the structure of the composite index and apply the I-distance method on each of the indicator levels. Therefore, the I-distance method is applied several times before obtaining the overall result. Herein we propose a twofold CIDI approach based on the idea of the twofold I-distance approach.

The original twofold I-distance algorithm has been changed, and the latest development of the CIDI methodology has been incorporated into it. The first step remains the same; the I-distance is applied on indicators within each category to obtain the Category I-distance value. Next, correlation coefficients between each indicator and the Category I-distance value it belongs to are calculated to perform the CIDI in the following step. Upon acquiring the CIDI indicator weights, the new CIDI Categories are calculated. Again, the procedure is repeated to obtain the CIDI category weights. Therefore, the weighting scheme is now acquired on two separate levels: first on the indicator level and second on the category level.

4. RESULTS AND FINDINGS

The aim of this paper was to scrutinize the U21 weighting scheme using the twofold CIDI methodology which is based on the previously introduced I-distance method. First, the CIDI methodology was employed on indicators of each category and afterwards on the newly calculated CIDI category values. The obtained CIDI weights and the effective weights are presented in Table 2.

Table 2 Weights of U21 indicators and categories based on the CIDI methodology

Category	Indicator	Weight assigned to indicators within category (a)	Weight assigned to category (b)	Effective weights (a*b)
Resources	R1	20.2%	24.22%	4.89%
	R2	17.9%		4.33%
	R3	20.4%		4.94%
	R4	19.8%		4.79%
	R5	21.7%		5.25%
Environment	E1	21.90%	23.83%	5.21%
	E2	18.70%		4.45%
	E3	17.90%		4.26%
	E4	24.90%		5.93%
	E5	16.60%		3.95%
Connectivity	C1	15.10%	27.31%	4.12%
	C2	15.80%		4.31%
	C3	15.00%		4.10%
	C4	18.40%		5.02%
	C5	19.30%		5.27%
	C6	16.40%		4.48%
Output	O1	7.1%	24.64%	1.75%
	O2	12.1%		2.99%
	O3	13.8%		3.41%
	O4	12.4%		3.06%
	O5	13.6%		3.36%
	O6	9.8%		2.42%
	O7	11.8%		2.91%
	O8	11.9%		2.94%
	O9	7.5%		1.85%

When analysing the new weighting scheme on the category level, we can see that CIDI suggests balanced weights. Such a distribution of weights is in accordance with the attitude held by Hazelkorn (2014), who argues that the balance is needed for the optimal functioning of the higher education system. The biggest difference in the weights assigned to categories is in the case of *Output*, which has been reduced from 40% to 24.64%. It is crucial that any ranking methodology, either ranking of higher education systems, either higher educational institutions, does not give significantly more importance to a particular aspect of education. Putting ahead high-quality outputs at the expense of lecture and teaching quality can lead to a situation in which the quality of the basic university mission significantly drops.

To better present the differences between the newly obtained and the official weighting scheme, the comparison of the official U21 and CIDI effective weights is provided. Namely, within the category *Resources*, the biggest weight changes occurred for indicators R4 and R5. According to the official

methodology, both indicators have effective weights of 2.5%, and now their effective weight increased to 4.79% (R4) and 5.25% (R5). These two indicators measure the level of expenditure on research and development, and an increase in their significance is in accordance with the idea of improving the competitiveness of the higher education system. In the next category, *Environment*, CIDI suggests more evenly distributed weights. Namely, according to the official methodology, the effective weights vary from 1 to 10%, and according to the CIDI they vary from 3.95 to 5.93%. Indicators E1 and E2, which measure the share of the female population in higher education (as a student and as a teacher) gained importance. It can be concluded that the new weighting scheme favours gender equality. The significance of the indicator E4 has reduced from 10% to 5.93%. Though this indicator is comprehensive and consists of sub-indicators, it is based on surveys conducted by Universitas. The validity of data collected using the survey is questionable, therefore, lowering the weight assigned to this indicator can be justified as a tendency to reduce the subjectivity of the overall index. All indicators of the category *Connection* gained importance, whereas indicators C3 and C4 stand out as their effective weights increased by 2.1 and 3.02 percentage points respectively. These indicators measure the presence of higher education institutions on the Internet and number of open access articles. Both indicators are oriented to measure the implementation of modern technologies in higher education, so the rise of their importance is in accordance with global trends. Weights of the fourth category, *Output*, went through drastic changes; their values either decreased or remained almost unchanged. For example, indicator O1, which measures the total number of published scientific papers, plummeted from 10% to 1.75%. Giving high significance to just one aspect of the output of higher education limits the overall metric, so the proposed weight reduction is justified. Also, we should mention the reduction of significance indicator O5, which is defined as the average value of the top three universities in the Shanghai ranking. As several authors raised the question of the validity of the Shanghai ranking list (for example Docampo & Cram, 2015; Dobrota & Dobrota, 2016), reducing the impact of its results should not be interpreted as a pitfall of the CIDI weighting scheme.

When the new weighting scheme was obtained, it was possible to calculate the values of the CIDI index using the official data and the CIDI weights. In Table 3 are given CIDI index values, CIDI ranks, and their comparison with the official U21 scores for the top 15 observed countries.

Table 3 Twofold CIDI scores, ranks, and comparison with the official U21 scores and ranks for 2015, 15 top ranked countries according to twofold CIDI values

Country	Twofold CIDI score	Twofold CIDI rank	U21 score	U21 rank
Switzerland	75.95	1	87.1	2
Sweden	74.17	2	84.7	5
Denmark	74.09	3	85.3	3
Finland	73.90	4	85.2	4
United States	73.63	5	100	1
Canada	69.60	6	82.8	6
Netherlands	68.90	7	81.6	7
Singapore	67.70	8	80.3	9
Norway	65.92	9	75.3	12
Austria	65.20	10	74.6	13
Australia	64.86	11	77.1	10
United Kingdom	64.50	12	80.6	8
Belgium	64.44	13	76.0	11
Germany	60.20	14	72.1	14
New Zealand	59.83	15	69.6	16

The ranks of the countries covered in the research largely coincide; however, differences between U21 and CIDI results can be observed. The measured Pearson's correlation coefficient is $r=0.862$ ($p<0.01$), which means that the correlation between the two overall scores is strong. Countries in the CIDI top 15 who dropped the most compared to the official U21 methodology, are the United States and the United Kingdom, both for 4 places. The reasons for such a difference in ranks can be found in the new weighting scheme. Namely, both countries are ranked exceptionally by the *Output* category, which is, according to the proposed methodology, less important for the ranking process. On the other hand, Sweden, Norway, and Austria have advanced for 3 places, as the proposed weighting scheme is more balanced and favors egalitarian societies and systems. Other countries are ranked similarly by both methods, with little rank changes. In order to thoroughly assess the validity of CIDI indicator, in addition to the comparison of the CIDI and U21 scores, we

compared their ranks. The obtained Spearman's correlation coefficient is also very high, and statistically significant, $r_s = 0.867$ ($p<0.01$).

5. CONCLUSION

In the recent years, a new type of rankings appeared: rankings of national higher education systems. Together with global university rankings, they are becoming a valuable tool for framing higher education on a global scale. However, their results should be used with caution due to their possible statistical inaccuracy and their influence to push the academic world into becoming more oriented on its assessment than its primary mission - teaching (Shin, Toutkoushian, & Teichler, 2011). Nevertheless, there are many drivers fueling further development and application of rankings, so it is exceptionally important to provide rankings as accurately as possible.

One of the newcomers to the rankings market is the U21 ranking of national higher education systems. The creators of this ranking acknowledged the priority of devising a tool for comparing national higher education systems across countries. They have taken courageous steps to deal with such a daunting task, but there is still place for improvement (Millot, 2014). Major concerns have arisen regarding its subjective weighting scheme. Therefore, the aim of this paper was to analyze and suggest a novel, data-driven weighting scheme based on the CIDI methodology. Namely, the methodology has previously been applied with success for assessing ranking methodologies of universities (Dobrota et al., 2016).

The twofold CIDI approach was implemented on the U21 index. Research findings show that among the categories, *Connectivity* stands out as the most important for the ranking system, followed by *Output*. This is a change compared to the official weighting scheme, as the focus is moved from *Output* to other categories.

In the case of indicators, the analysis outlined that *Output* category needs refinement as some indicators were given significantly lower importance compared to the official ranking methodology, especially the indicator related to the number of published papers. What can also be noted is that the U21 is missing complex bibliometric indicators such as percentile-based indicators which take into account the citations count (for example Bornmann & Mutz, 2014). Also, differences in citing behavior between sciences should be taken into account. A study conducted by Bornmann, de Moya Anegón, and Mutz (2013) proved that certain subject-specific types of institutions are in advantageous position when it comes to ranking in terms of outcome performance. Therefore, one of our suggestions for further development of the U21 ranking is the inclusion of more sophisticated bibliometric indicators.

Two other directions of further research could be additionally defined. First, the question of the level of objectivity of metric arises. Performing the CIDI methodology (either one-fold or two-fold) creates an objective metric, and eliminates the experts' knowledge. However, such an approach can raise certain concerns regarding the possibility of creating a measurement that may not be in accordance with the nature of the observed phenomenon. Therefore, we suggest a hybrid approach (Maricic et al., 2015). Namely, weights on one indicator level could be redefined using the CIDI methodology, while the weights on the other level would be left as suggested by the experts. Such an approach would take into account benefits of both subjective and data derived weighting methods. Next, the presented results suggest that the number of indicators can be refined. To tackle this, the I-distance post hoc analysis could be employed (Markovic et al., 2016). This analysis is an iterative process where the number of indicators in each iteration is reduced for the indicator which has the lowest correlation coefficient with previously obtained I-distance value.

International academic rankings are faced with methodological difficulties which affect the validity of their results. Herein we proposed a transparent approach for redesigning a ranking methodology that is primarily oriented on the process of assigning weights. We believe our paper might act as an incentive for further academic research on the topic of national higher education systems rankings and their weighting schemes.

REFERENCES

Bornmann, L, de Moya Anegón, F., & Mutz, R. (2013). Do universities or research institutions with a specific subject profile have an advantage or a disadvantage in institutional rankings? *Journal of the Association for Information Science and Technology, 64*(11), 2310-2316. doi: 10.1002/asi.22923

Bornmann, L., & Mutz, R. (2014). From P100 to P100': A new citation-rank approach. *Journal of the Association for Information Science and Technology, 65*(9), 1939-1943. doi: 10.1002/asi.23152

Cheng, S. K. (2012). The U21 rankings of national higher education systems: re-analyze to optimize. *European Journal of Higher Education, 2*(2-3), 205-217. doi: 10.1080/21568235.2012.697695

Dobrota, M., & Dobrota, M. (2016). ARWU ranking uncertainty and sensitivity: What if the award factor was excluded? *Journal of the Association for Information Science and Technology 67*(2), 480-482. doi: 10.1002/asi.23527

Dobrota, M., Bulajic, M., Bornmann, L., & Jeremic, V. (2016). A new approach to the QS university ranking using the composite I-distance indicator: Uncertainty and sensitivity analyses. *Journal of the Association for Information Science and Technology, 67*(1), 200-211. doi: 10.1002/asi.23355

Dobrota, M., Martic, M., Bulajic, M., & Jeremic, V. (2015). Two-phased composite I-distance indicator approach for evaluation of countries' information development. *Telecommunications Policy, 39*(5), 406-420. doi: 10.1016/j.telpol.2015.03.003

Docampo, D., & Cram, L. (2015). On the effects of institutional size in university classifications: the case of the Shanghai ranking. *Scientometrics, 102*(2), 1325-1346. doi: 10.1007/s11192-014-1488-z

Harzing, A. W., & Mijnhardt, W. (2015). Proof over promise: towards a more inclusive ranking of Dutch academics in Economics & Business. *Scientometrics, 102*(1), 727-749. doi: 10.1007/s11192-014-1370-z

Hazelkorn, E. (2014). Reflections on a Decade of Global Rankings: what we've learned and outstanding issues. *European Journal of Education, 49*(1), 12-28. doi: 10.1111/ejed.12059

Huang, M. H. (2012). Opening the black box of QS World University Rankings. *Research Evaluation, 21*(1), 71-78. doi: 10.1093/reseval/rvr003

Ivanovic, B. (1977). Classification Theory. Institute for Industrial Economic, Belgrade.

Jeremic, V., & Jovanovic-Milenkovic, M. (2014). Evaluation of Asian university rankings: Position and perspective of leading Indian higher education institutions. *Current Science, 106*(12), 1647-1653.

Jeremic, V., Bulajic, M., Martic, M., & Radojicic, Z. (2011). A fresh approach to evaluating the academic ranking of world universities. *Scientometrics, 87*(3), 587-596. doi: 10.1007/s11192-011-0361-6.

Marginson, S., & van der Wende, M. C. (2007). Globalization and higher education. Education Working Paper No. 8. Paris: OECD.

Maricic, M., & Kostic-Stankovic, M. (2016). Towards an impartial Responsible Competitiveness Index: a twofold multivariate I-distance approach. *Quality & Quantity, 50*(1), 103-120. doi: 10.1007/s11135-014-0139-z.

Maricic, M., Bulajic, M., Martic, M., & Dobrota, M. (2015). Measuring the ICT Development: The Fusion of Biased and Objective Approach. Proceedings of the XII Balkan Conference on Operational Research (BALCOR 2015), September 9-13, Constanta, Romania.

Marković, M., Zdravković, S., Mitrović, M., & Radojičić, A. (2016). An Iterative Multivariate Post Hoc I-Distance Approach in Evaluating OECD Better Life Index. *Social Indicators Research, 126*(1), 1-19. doi: 10.1007/s11205-015-0879-8.

Martin, M., & Sauvageot, C. (2011). Constructing an indicator system or scorecard for higher education: A practical guide. Paris: International Institute for Educational Planning

Millot, B. (2014). Top Universities or Top Higher Education Systems? *International Higher Education*, (75), 7-8.

Rauhvargers, A. (2014). Where are the global rankings leading us? An analysis of recent methodological changes and new developments. *European Journal of Education, 49*(1), 29-44. doi: 10.1111/ejed.12066

Saisana, M., & D'Hombres, B. (2008). Higher Education Rankings: Robustness Issues and Critical Assessment; How much confidence can we have in Higher Education Rankings? Luxembourg: JRC.

Saisana, M., D'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research policy, 40*(1), 165-177. doi: 10.1016/j.respol.2010.09.003.

Shin, J. C., Toutkoushian, R. K., & Teichler, U. (Eds.). (2011). University rankings: Theoretical basis, methodology and impacts on global higher education. Springer Science & Business Media.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics, 136*(1), 31-64. doi: 10.1016/j.jeconom.2005.07.009

Soh, K. (2014). Nominal versus attained weights in Universitas 21 Ranking. *Studies in higher education, 39*(6), 944-951. doi: 10.1080/03075079.2012.754866

Universitas 21. (2015). U21 Ranking of National Higher Education Systems 2015. Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Melbourne, Australia

Williams, R., de Rassenfosse, G., Jensen, P., & Marginson, S. (2013). The determinants of quality national higher education systems. *Journal of Higher Education Policy and Management, 35*(6), 599-611. doi: 10.1080/1360080X.2013.854288

Zornic, N., Bornmann, L., Maricic, M., Markovic, A., Martic, M. & Jeremic, V. (2015). Ranking institutions within a university based on their scientific performance: A percentile-based approach. *El Profesional de la informacion, 24*(5), 551-566. doi: 10.3145/epi.2015.sep.05.

SELECTING AN APPROPRIATE METHOD FOR MISSING DATA IMPUTATION: A CASE OF COUNTRIES RANKING

Jovana Rakićević*[1], Gordana Savić[1], Milica Bulajić[1]

[1]University of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: jovana.rakicevic@fon.bg.ac.rs

Abstract: *One of the general problems we face when crating composite indices is the missing data. It is almost impossible to have complete datasets for all the observed entities according to different indicators. Firstly, the paper discusses what to do when having incomplete datasets. The easiest ways are to exclude the indicators or the cases for which a large amount of data is missing. None of these solutions is the best, since neither we want to lose any aspect of the observed phenomenon, nor the cases we observe. Therefore, we need to opt for an imputation technique. The paper further provides an overview of several ideas and explanations of imputation, and addresses the next problem that arises on how to select one appropriate technique among the available ones. Bearing in mind the extent to which the data can be manipulated when imputing, we provide an analysis oriented to find an answer to the posed question. The analysis tests six single imputation techniques applied to several countries for two observed indicators – R&D expenditures and R&D researchers.*

Keywords*: missing data, imputation, single imputation, mean imputation, regression imputation, cluster analysis, composite index*

1. INTRODUCTION

Composite indices have evolved and become a very popular way of assessing large phenomena that cannot be captured by one single indicator. They serve as a tool for measuring countries performance, and ranking accordingly. Scholars and international organizations create different methodologies for constructing composite indices to meet a wide range of purposes, primarily to "fairly" rank the countries and enable the cross-country comparisons, and to facilitate the strategic decision-making processes. Examples include assessments of human development, ICT development, technological readiness, e-readiness, innovativeness, competitiveness, sustainability, wellbeing, quality of life, gender equity, ease of doing business and many others (see Bandura, 2008).

However, although having dozens of new or improved composite indices every year that are proven to have many positive aspects, we should be aware of their shortages and different difficulties we come across while constructing. There are several general problems we face: from selecting the indicators that will be included in the final index, through finding the right sources for collecting the data, dealing with the missing data and standardizing the data, to finding the right methods for weighting and aggregating the selected indicators. Our decisions in each step will highly affect the final result. For example, Grupp and Mogee (2004) showed how innovation scoreboards can be manipulated in the policymaking system, depending on the selected indicators. Hurdlikova (2013) showed seven different scenarios of the EU countries ranking according to the selected indicators for Europe 2020 with the aim to highlight the sensitivity of results which vary and depend on choosing both the normalization methods and the weighting and aggregation schemes. One of the mostly criticized aspects of constructing indices is the choice of weighting method, since weights are subjective by nature and can considerably affect the final results. Cherchye et al. (2006), and Filippetti and Peyrache (2011) propose the use of Data Envelopment Analysis (DEA) for this purpose. However, there is one crucial problem which precedes, and that is the missing data.

The authors have been creating a composite index of technological capability of countries and experienced how difficult it could be to deal with the missing data problem, since different imputations can highly affect the changes of ranks, i.e. the final scores of countries. As Foa and Turner (2012) highlight, one of the key decisions we have to make when designing composite indices is how to deal with this problem in such a manner to ensure that the final index values reflect a meaningful image of the available information. This paper firstly provides an overview of existing imputation techniques, and then addresses the issue of missing data imputation in such a manner to provide explanations on how to select the most appropriate imputation technique. We restrict the analysis to the single imputation techniques implemented in IBM SPSS Statistics 22. In some cases cold desk imputation is the most appropriate, sometimes the regression analysis provides better results than the simple mean value, in some cases it is better to firstly classify the countries and then to impute the missing values by using an average cluster value or regression analysis over the selected cluster, etc. This will be elaborated later in the text.

The rest of the paper is organized as follows. Section 2 explains the possible ways to solve the problem of missing data and provides an overview of several imputation techniques. Section 3 provides an analysis on choosing the "right" imputation technique among six applied on the sample of several countries observed according to two indicators – R&D expenditures and R&D researchers; root mean square error is calculated in order to test decisions by finding accuracy of the applied techniques on different datasets for two of other indicators used for assessing countries' technological capability: ICT goods imports and Electric power consumption. This section offers the results and discussion of the study. Section 4 concludes the paper.

2. SOLVING THE PROBLEM OF MISSING DATA

After deciding which indicators to include in an index, we collect the data. At this point we become aware of the amount of the missing data, and whether we will be able to encompass the selected group of indicators or not. If a lot of data is missing that cannot be imputed meaningfully, we have to exclude some indicators from the analysis, i.e. to drop out the variables or instances for which the data is missing. So, the **case deletion** is the first and the most radical way of dealing with this problem. This method is used in the case when, for example, among the observed countries we miss the values for the majority of countries for the larger part of the observed period of time, so that we cannot estimate and impute the missing values. For example, the famous Human Development Index deals with the missing data problem in this way, selecting the variables for which complete data across the domain of countries is relatively easy to obtain, restricting to a small set of indicators for which complete data exists. On one hand this is good since it provides objective and more reliable results, but on the other hand this way we cannot include all important aspects of one phenomenon, but have to restrict to the basic ones. The case deletion approach is, of course, appealing because of its simplicity. However, this is not applicable in cases when missing values cover a lot of instances, or their presence in essential attributes is large (Little & Rubin, 1987).

Generally, if this radical exclusion of indicators is not needed or wanted, composite indices can deal with the problem of missing data in three ways (Foa &Tanner, 2012). The first and the simplest solution is to drop out any country for which complete data does not exist; the second solution is to impute missing values using different methods; and the third is to use only existing data in the estimation of the index, but supplementing with an estimated margin of error. The first solution of reducing the sample is not always acceptable, especially if we want to make the global cross-country comparisons. However, there are indices, as the Doing Business Indicators whose authors want to avoid the methodological problems and obtain an objective result (Foa & Turner, 2012). This is why the country domain is smaller than in the case of e.g. ArCo index of technological capability of countries (Archibugi & Coco, 2004), which is measured for 162 countries. So, in order to enlarge the sample, we need to find the "right way" of imputing the missing values, always having in mind that the selection of methods manipulates the results. We will discuss this in detail in the Section 2.1. The third proposed solution refers to using only the existing data in the estimation of the index, but supplementing the results with an estimated margin of error, based on the number of missing items, among other criteria observed. This approach is used in a number of recent indices such as the Corruptions Perceptions Index (CPI) where the confidence range indicates the reliability of the country scores. It tells that allowing for a margin of error we can be 90 percent sure that the true score for countries lies within the given range.

2.1. Missing data imputation techniques – an overview

The literature on the analysis of missing data is extensive and in rapid development. OECD & EC-JRC (2008) published the *Handbook on Constructing Composite Indicators – Methodology and User Guide*, which provides help in creating indices. The Handbook aims to contribute to a better understanding of the complexity of composite indicators and to an improvement in the techniques currently used to build them. Among others, the Handbook deals with the problem of missing data imputation, and suggests single and multiple imputations as possible solutions. As defined in the Handbook, "imputations are means or draws from a predictive distribution of missing values." The predictive distribution must be generated by using the observed data.

Single imputation refers to both implicit and explicit modelling. The implicit techniques are simple. ***Hot desk imputation*** refers to filling in the blanks cells with individual data which are drawn from the unit that has similar characteristics (for example if we observe units according to four indicators and miss a value for one unit for indicator x, we will fill in that missing value with the value of indicator x for the unit which is the most similar to the one observed according to the other three indicators). ***Substitution*** means the replacement of non-responding units with the unselected units in the sample, while ***Cold desk imputation*** is the replacement of missing values with the values from an external source (for example from the previous

realization of the same survey or the value of an indicator from the previous year in the case of assessing the countries performance). Additionally, we propose the ***Simple mean imputation*** which refers to imputing the missing values considering only one instance and its' dataset (separate from the sample), imputing the missing value by finding average of the previous and the next value to the one missing (it is important to have in mind that assessing countries performance we deal with time-series datasets).

Explicit modelling is more complex and demands more detailed explanations. Using ***unconditional mean imputation*** means that we impute the missing values with the sample mean (median, mode) for the observed indicator. The limitation of mean value based imputation and its variations is its focus on a specific variable without taking into account the overall similarities between instances (Ayuyev et al., 2009). This is the easiest way of explicit modelling, but not always precise enough. Therefore, we could use other two more sophisticated techniques. Firstly, the ***regression imputation*** where missing values are imputed with the predicted values obtained by regression. Here we observe dependent variable and independent variable(s). The dependent variable is the indicator for which we miss some values, and the independent variable(s) is (are) the individual indicator(s) which show strong relationship (usually high correlation) with the dependent variable. ***Expectation maximization imputation*** focuses on the interdependence between parameters of the model and the missing values. It is an iterative process. First, the missing values are predicted based on initial estimates of the model parameters values. These predictions are then used to update the parameters values, and the process is repeated. The sequence of parameters converges to maximum-likelihood estimates, and the time to convergence depends on the proportion of missing data and the flatness of the likelihood function. For more detailed mathematical explanation on explicit modelling techniques see OECD & EC-JRC (2008, pp. 55-58).

Multiple imputation is considered to be one of the most powerful approaches to missing values estimation (Ayeyev et al., 2009). It is a general approach that does not require a specification of parameterised likelihood for all data. The missing data is imputed with a random process that reflects uncertainty. Imputation is done N times, to create N "complete" datasets. The parameters of interest are estimated on each data set, together with their standard errors. Average (mean or median) estimates are combined using the N sets and between-and within-imputation variance is calculated. Although any imputation method can be used in multiple imputation (used repeatedly to obtain N values), one of the most general models is the ***Markov Chain Monte Carlo (MCMC) method***. It is a sequence of random variables where the distribution of the observed element depends on the value of the previous one. It is assumed that data are drawn from a multivariate normal distribution and requires the following assumptions: missing at random (MAR) and missing completely at random (MCAR). For more detailed explanation see OECD & EC-JRC (2008, pp. 58-61). For example, the Environmental Sustainability Index uses the MCMC technique to substitute the missing values (Srebotnjak, 2001).

Based on the amount of operations performed, Zhang (2011) presents the following categorisation of imputation techniques: single, multiple, fractional and iterative. **Fractional imputation** represents a compromise between the single and multiple imputation methods, while **iterative imputation** techniques primarily use a generate-and-test mechanism, taking into account useful information (including incomplete cases).

Fujikawa and Ho (2002) consider a **clustering based approach for missing data imputation**, where the premise is that units could be grouped such that all the imputations in identified groups are independent from other groups. ***Distance-based clustering*** is focused mainly on development of supervised clustering methods and mean/mode based imputations in these clusters (De Mántaras, 1991). They are based on a strict separation for objects within clusters, so it is assumed that there is no influence of instances in one cluster to an imputation process in other clusters. Ayuyev et al. (2009) suggest the improved ***dynamic clustering-based imputation (DCI)*** of missing values in mixed type data. They consider the appropriate choice of a method for imputation especially important when the fraction of missing values is large and the data are of mixed type. The proposed DCI algorithm relies on similarity information from shared neighbours, where mixed type variables are considered together. Around each instance with a missing value they deterministically construct an independent cluster of similar instances with no missing values for a particular attribute. In contrast to a typical clustering method, they allow cluster intersections meaning that the same unit may be included in many clusters. It relies on a distance measure that considers both categorical and continuous variables and is applicable for estimation of missing values in high dimensional mixed type data.

Different authors propose and analyze other complex algorithms for missing data imputation. For example, Abdella & Marwala (2005) introduced a new method for imputing the missing values which uses a combination of **genetic algorithms and neural networks** for approximation of the missing data. Nelwamondo et al. (2007) compare neural network and expectation maximization techniques, while Lobato

et al. (2015) offer a **multi-objective genetic algorithm** for missing data imputation (MOGAImp) which is an iterative imputation technique. Explanation of these complex methods is beyond the scope of this paper.

3. MISSING DATA IMPUTATION: MAKING A CHOICE AMONG SIX SINGLE IMPUTATION TECHNIQUES – RESULTS AND DISCUSSION

While assessing the technological capabilities of countries, we faced the missing data problem. Firstly, we needed to exclude some of the indicators because of a large amount of data missing for the observed period of time. Secondly, we excluded some countries with incomplete datasets, where meaningful imputations could not be performed. Among the countries left, we needed to impute the missing values for some crucial indicators that could not be excluded from the analysis, such as *Researchers in R&D (per million people) – IND1* and *R&D expenditures (% of GDP) – IND2*. In this section we will try to elaborate and explain how to choose the "right" imputation technique on the example of these two indicators, selecting from the six single imputation techniques applied. Table 1 shows the results of imputation using the following techniques: CD – Cold desk imputation, SM – Simple mean imputation, EM – Expectation maximisation imputation, R – Regression imputation, C-M – Cluster mean imputation, C-R – cluster regression imputation. For two countries we apply C-R_{II}, which will be explained later in the text. For the first and the second indicator we analyse four and five countries, respectively. For each country, in the first row we present the existing dataset for the observed period. Framed positions are missing. In the next six (or seven) rows we present different results obtained by applying the mentioned techniques. We marked the chosen (most appropriate) techniques by putting an asterisk next to the technique(s) name(s) and formatting numbers in bold.

As we can see from the Table, CD imputation can be used always. We propose it to be applied this way: if the value for one year exists and we miss the value for the previous one, we impute the missing value by copying the value from the current year (e.g. Chile for both indicators). Otherwise, if we have the previous values, we always consider that to be a better situation, and copy that value for the next year missing (e.g. Greece for IND1). This technique is surely not accurate enough, but sometimes better to be applied that the others. This refers mainly to the cases where we miss the values for long time period. It is hard to predict the trend of the indicator in the past/future, especially if we miss the values for the same period for the indicator that is correlated with the one observed. This is the case with Chile, where we miss the values for the period 2000-2006 for both indicators and cannot apply the trend from one indicator to the other. For example, for Israel we also miss the values for the long time period 2000-2010, but we are in better situation because we have the whole dataset for Israel for IND2. In this case we firstly clustered the observed countries according to R&D researchers for the periods for which the values exist (2011-2013) and obtained Israel to be in a cluster with Japan, Korea, Sweden and Finland. Then we applied C-R_{II} imputation, meaning that regression analysis is performed over the obtained cluster, using R&D researchers as a dependent variable, and R&D expenditures as the independent. Of course, it is questionable whether we can accept any of these results, since a lot of values are missing. If not, the safest approach is to exclude these two countries from the further analysis. If we accept these values, than we need to be cautious when processing this data because illogical result may appear. In other cases (except Israel, Australia and Greece for IND1) C-R imputation means that we firstly performed the cluster analysis, and then the regression over the observed cluster, for one indicator observed (we did not connect two indicators because the data was missing for both).

Regarding the SM imputation, it is best to be applied in the situation when we miss one value between two which exist, and regression is not applicable because of the unstable trend which cannot be captured. This is the case with Australia and New Zealand for IND2, and for Australia (2005, 2007) and Greece (2000, 2002, 2004) for IND1. The rest of the missing values for both Australia and Greece for IND1 cannot be obtained by SM imputation. We tried to perform C-R_{II} which did not provide satisfactory results, so we use CD for Australia for 2009-2013 (copying value from 2008) and for Greece for 2008-2010 (copying from 2007).

Regression imputation, which refers to performing the regression analysis over the whole sample, is applicable and good when the trend is captured. This is the case with Greece and Luxembourg for IND2. Besides this, both SM and CD could have been applied. This is where we easily see how our decision in this step would affect the final result (score of country). It is interesting to notice that when comparing R and C-R imputation, we obtain quite different results (see e.g. Greece for IND2, period 2008-2010). However, none of these gives good results if a lot of data is missing, and the trend cannot be captured.

C-M imputation has not proved to be a good solution in any of the cases presented. Large deviations from the existing values are noticed. For example, for Chile for IND1 we obtained the following values for the missing period 2000-2006: 543.62, 539.82, 569.74, 592.36, 640.42, 700.28, 683.89, and the existing value for 2007 is 337.18. It doesn't reach even 400 until the end of the period observed. The problem with C-M is that the instances in one cluster should be really close in order to be in a position to perform this imputation

technique. That cannot be the case with Chile which is among the countries with the lowest values in the sample. In this case average of both the cluster and the sample would be over the values for Chile. This is also the case if we observe the country among those with the highest values in the sample (the mean value, both cluster and sample, will always be lower than needed).

Table 1: Comparison of the applied techniques – results for R&D researchers and R&D expenditures

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
INDICATOR 1: R&D researchers per million people *(World Bank)*														
Australia	3454.26	3549.72	3749.7	3664.8	4062.53		4231.78		4335.39					
CD*						4062.53		4231.78		**4335.39**	**4335.39**	**4335.39**	**4335.39**	**4335.39**
SM*						**4147.15**		**4283.59**		N.A.	N.A.	N.A.	N.A.	N.A.
EM						3741.31		4305.16		3709.33	4063.27	3789.79	4182.98	3855.18
R						3843.00		4125.18		4350.37	4357.64	4108.17	3886.09	3831.81
C-M						3061.57		3753.61		3875.63	4005.01	4121.30	4261.08	4291.60
C-R						3840.89		4188.89		4098.72	4203.41	4139.55	3770.85	3900.52
C-R_{II}						4147.15		4283.59		3806.46	3802.31	3729.17	N.A.	N.A.
Chile								337.18	357.99	288.71	319.72	353.37	390.95	389.41
CD*	**337.18**	**337.18**	**337.18**	**337.18**	**337.18**	**337.18**	**337.18**							
SM	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.							
EM	349.85	359.16	372.02	318.46	341.41	296.53	376.64							
R	306.51	326.91	308.26	325.33	327.14	327.79	360.91							
C-M	543.62	539.82	569.74	592.36	640.42	700.28	683.89							
C-R	314.71	346.82	338.05	334.93	279.91	283.18	336.52							
Greece		1307.94		1417.21		1769.95	1793.7	1887.83				2212.33	2232.29	2628.16
CD*	1307.94		1307.94		1417.21				**1887.83**	**1887.83**	**1887.83**			
SM*	**1253.31**		**1362.58**		**1593.58**				N.A.	N.A.	N.A.			
EM	1348.65		1293.19		1480.10				2070.03	2096.98	1833.73			
R	1504.79		1524.36		1674.95				2009.35	2164.72	2539.52			
C-M	1333.15		1394.17		1430.71				1686.58	1765.66	1846.37			
C-R	1307.94		1769.95		1887.83				1417.21	1769.95	1307.94			
C-R_{II}	1253.31		1362.58		1593.58				1558.75	1553.60	1549.57			
Israel												7296.25	8282.31	7934.86
CD	7296.25	7296.25	7296.25	7296.25	7296.25	7296.25	7296.25	7296.25	7296.25	7296.25	7296.25			
SM	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.			
EM	7603.40	7514.73	8260.92	8207.10	8323.30	8233.91	7119.58	7363.84	8068.52	8202.68	8379.79			
R	7168.00	7777.79	7664.46	7120.56	7064.76	7459.99	7702.47	8375.10	8159.20	7615.86	7157.08			
C-M	6731.79	7109.84	7427.94	8002.61	7841.8	7544.65	7673.06	7373.24	7692.44	7648.88	7717.48			
C-R	7679.28	7667.56	7257.47	8551.63	7401.66	7614.01	6944.95	7075.53	8344.48	8205.42	8129.85			
C-R_{II}*	**7168**	**7777.79**	**7664.46**	**7120.56**	**7064.76**	**7459.99**	**7702.47**	**8375.1**	**8159.2**	**7615.86**	**7157.09**			
INDICATOR 2: R&D expenditures - % of GDP *(World Bank)*														
Australia	1.576		1.753		1.853		2.182		2.401		2.386	2.249		
CD		1.576		1.753		1.853		2.182		2.401				
SM*		**1.665**		**1.803**		**2.018**		**2.288**		**2.394**				
EM		2.144		2.091		2.122		2.345		1.881				
R		1.608		1.839		2.08		2.146		2.363				
C-M		1.848		1.795		1.789		1.812		1.904				
C-R		2.023		2.066		2.178		2.529		2.014				
Chile								0.311	0.375	0.353	0.331	0.353		
CD*	**0.311**	**0.311**	**0.311**	**0.311**	**0.311**	**0.311**	**0.311**							
SM	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.							
EM	0.361	0.338	0.354	0.346	0.356	0.338	0.341							
R	0.333	0.337	0.323	0.339	0.321	0.307	0.313							
C-M	0.517	0.520	0.504	0.508	0.502	0.519	0.526							
C-R	0.337	0.336	0.316	0.380	0.323	0.334	0.301							
Greece		0.560		0.548	0.529	0.579	0.561	0.576	0.662	0.626	0.598	0.67		
CD	0.560		0.560											
SM	N.A.		0.554											
EM	0.591		0.592											
R*	**0.593**		**0.553**											
C-M	0.517		0.504											
C-R	0.559		0.512											
Luxembourg	1.574			1.652	1.625	1.585	1.692	1.645	1.649	1.719	1.503	1.411		
CD		1.574	1.574											
SM		1.599	1.626											
EM		1.606	1.6											
R*		**1.615**	**1.633**											
C-M		1.902	1.892											
C-R		1.570	1.715											
New Zealand		1.1		1.147		1.117		1.161		1.262		1.248		
CD	1.1		1.1		1.147		1.117		1.161		1.262			
SM*	**1.077**		**1.124**		**1.132**		**1.139**		**1.211**		**1.255**			
EM	1.197		1.188		1.198		1.171		1.168		1.127			
R	1.110		1.103		1.125		1.171		1.174		1.231			
C-M	0.907		0.994		1.027		1.135		1.257		1.406			
C-R	1.113		1.108		1.129		1.177		1.183		1.228			

Performing this analysis, we can easily see how the final country score would surely depend on the imputation technique we decide to apply. Looking for example at Australia for 2005 (IND1), we obtained six pretty different values: 4062.53, 4147.15, 3741.31, 3843.00, 3061.57, and 3840.89 using six techniques. Deciding which to use is completely subjective and enables the manipulation of results. This is a huge problem, since a country could be many places up or down on the list depending on the technique accepted. Besides this, it should be highlighted that one imputation technique cannot be applied for one indicator for all the countries for which data is missing. It highly depends on the nature of the existing data, and the amount of data missing. Additionally, we could conclude that when imputing we should always look at the bigger picture, connecting (when possible) different observed indicators (in our case the correlated R&D expenditures and researchers), in order to ensure the reliability of the imputations.

For the purpose of testing accuracy of imputation methods, we performed another analysis that includes more datasets for two other indicators used for assessing countries' technological capabilities, for which complete data exists across the observed domain of countries: *ICT goods imports (% of total goods imports) – IND3* and *Electric power consumption (kWh per capita) – IND4*. The idea of this analysis is to delete the values for the same years that were missing for IND1 (in Table 1), and calculate the root mean square error (RMSE) for each technique applied, for each country. The results for IND3 are presented in the Table 2. For each country, in the first row we present the existing (complete) dataset for the observed period. In the second row we deleted the values; framed positions are now "missing". Next six rows present different results obtained by applying the same techniques used in the previous analysis. For each technique we calculated the RMSE (presented in the last column) by applying the following formula (Chai & Draxler, 2014):

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \quad (1)$$

where y_t is the observed, actual value and $\hat{y}_t$ is the predicted, imputed value.

The aim of this analysis is to examine whether we would intuitively choose the technique with the smallest error as the most appropriate, following the principle we used for the first analysis. For each country, the techniques with the smallest RMSE are marked by putting an asterisk next to the technique(s) name(s).

Table 2: Testing the accuracy of the applied techniques – Calculating the root mean square error (RMSE)

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	RMSE
INDICATOR 3: ICT goods imports - % of total goods imports *(World Bank)*															
Australia	15.53	13.49	12.38	12.12	12.77	11.76	11.66	10.70	9.13	10.51	10.64	9.83	8.89	8.86	
Deleted values	15.53	13.49	12.38	12.12	12.77		11.66		9.13						
CD						12.77		11.66		9.13	9.13	9.13	9.13	9.13	0.982
SM*						12.22		10.40		N.A.	N.A.	N.A.	N.A.	N.A.	**0.388**
EM*						12.21		9.94		9.79	10.33	8.98	8.42	8.51	**0.593**
R						12.59		11.66		12.28	12.91	10.30	9.89	10.39	1.387
C-M						10.36		8.79		8.68	9.33	8.73	8.27	8.16	1.350
C-R						12.34		11.24		11.14	12.00	10.96	10.08	10.63	1.115
Chile	10.19	9.10	8.57	8.07	8.73	8.24	8.84	7.50	6.18	7.35	7.86	7.08	7.20	7.64	
Deleted values								7.50	6.18	7.35	7.86	7.08	7.20	7.64	.
CD	7.50	7.50	7.50	7.50	7.50	7.50	7.50								1.470
SM	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.								N.A.
EM*	10.07	8.97	7.58	8.33	8.47	8.90	8.01								**0.570**
R	7.27	7.33	7.62	7.14	7.13	7.26	7.53								1.633
C-M	11.41	10.64	9.26	9.16	9.62	10.00	9.57								1.192
C-R	8.11	6.88	7.11	6.64	6.78	6.55	7.87								1.733
Greece	7.43	6.70	5.81	5.84	6.20	5.74	5.67	5.85	5.17	5.45	4.98	4.99	5.10	3.95	
Deleted values		6.70		5.84		5.74	5.67	5.85				4.99	5.10	3.95	
CD	6.70		6.70		5.84				5.85	5.85	5.85				0.687
SM*	6.91		6.27		5.79				N.A.	N.A.	N.A.				**0.329**
EM	5.42		6.80		5.33				4.95	5.45	5.37				0.998
R*	6.70		6.16		5.71				5.37	5.16	4.70				**0.428**
C-M	10.00		8.35		8.32				6.84	7.36	7.61				2.270
C-R	6.49		6.20		5.98				6.88	5.85	5.22				0.839
Israel	14.13	11.48	9.85	9.25	10.40	9.75	9.57	8.48	7.85	9.72	9.20	8.73	8.86	8.76	
Deleted values												8.73	8.86	8.76	
CD	8.73	8.73	8.73	8.73	8.73	8.73	8.73	8.73	8.73	8.73	8.73				2.019
SM	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.				N.A.
EM	11.92	11.97	11.38	10.88	11.00	11.32	11.27	10.71	9.68	10.27	10.11				1.516
R	8.73	8.86	8.90	8.76	8.76	8.75	8.79	8.81	8.75	8.79	8.63				1.990
C-M*	13.49	12.66	11.34	10.43	10.85	11.05	10.91	9.31	8.67	8.58	8.81				**1.041**
C-R	8.71	8.76	8.83	8.79	8.80	8.82	8.74	8.67	8.86	8.82	8.74				2.005

Regarding the values for Australia for IND3 (Table 2), we would certainly choose SM imputation for years 2005 and 2007 (which also has the lowest RMSE), but would definitely make a mistake by choosing CD imputation (RMSE=0.982) for the years 2009-2013, since EM has lower RMSE value (0.593). But not having

the values with which to connect and compare, it would be brave and uncertain to choose EM values as the most appropriate. Here we can easily see how the results are questionable when the data is imputed. Secondly, regarding Chile, the only method we could use is CD since a large amount of data is missing and we cannot predict the trend by connecting to another correlated value (performing C-R_{II}). CD imputation has RMSE value of 1.47, which is much higher than the EM which again proves to be better, with RMSE of 0.57. However, again it would be hard to justify such a decision on choosing EM for imputation. Regarding Greece, we would again be fine with the values for 2000, 2002 and 2004, performing the SM imputation (RMSE=0.329). This shows that it is not a problem when one value is missing between two which exist. However, for the years 2008-2010, a mistake would be made by using CD (RMSE=0.687) instead of R (RMSE=0.428). But it is encouraging that in this case the mistake wouldn't be large, since the values of RMSE are close (unlike for Australia). Israel is the most challenging case, since a large amount of data is missing. As we can see from the RMSE values for all techniques for Israel, the errors are higher. This indicates that inaccurate decision would surely be made when a lot of data is missing. This result is logical and expected, and supports the claim that entities with a large number of missing data should be excluded from the analysis. For the opposite reason SM is always the best technique (but can be applied only in a few cases). For Israel, C-M imputation stands as the best one but still with a large error of 1.041.

In order to obtain a general value of accuracy of the applied techniques, total RMSE is calculated using the aforementioned formula, but gathering the data for all countries. The results are presented for IND3 and IND4 (Table 3). Because of the space limitations, we did not provide the detailed analysis for RMSE values for IND4 for each country, but only the total. SM imputation is applied for 5 missing values, and other techniques are used for 31 imputations for IND3, i.e. 30 for IND4 (because IND4 is not measured for 2013).

Table 3: Total RMSE values

IND 3 - ICT goods imports						IND 4 – Electric power consumption					
CD	**SM**	**EM**	R	C-M	C-R	CD	**SM**	**EM**	R	C-M	C-R
1.498	**0.436**	**1.078**	1.574	1.454	1.588	358.18	**161.72**	**270.67**	450.33	1447.79	459.45

From the Table 3 we can see that SM expectedly proves to be the most accurate technique, since it is applicable only in a few cases and is measured as the average value between the two which exist. More importantly, EM proves to be the best technique among those which can be applied always. It is an interesting result, since we did not choose EM in none of the cases presented in the first analysis. Looking at Tables 1 and 2, it is obvious why we did not choose EM, since there is no explanation on such a decision. We can explain why we choose simple methods as CD, believing that we make smaller mistake by arguing that the trend has not changed over time. This analysis proves this to be wrong. Other techniques have higher errors, ranked differently for these two indicators. It is not possible to claim which one is the most inaccurate; it depends on the sample.

4. CONCLUSION

Using composite indices for assessing large phenomena and ranking the countries accordingly has been widely accepted. Even being aware of the shortages of composite indices, the race for a better country rank on different lists has taken its toll. Understanding the importance of composite indices today, many scholars try to deal with different methodological problems mentioned in this paper, all with the aim to create something maybe unreachable, and that is the objective and precise methodology for creating indices, in order to ensure the reliability of the results. The problem we address in this paper is the missing data. It is almost impossible to have complete datasets for all the countries observed according to different indicators that an index includes. To avoid exclusion of indicators or the countries observed, we need to opt for an imputation technique that will make the data complete, and countries performance possible to assess. In this paper we showed an overview of different imputation techniques, both single and multiple. Having different techniques among which we can choose is good, but we stay with a big problem of deciding which method to use to fill in empty data spaces. Luengo et al. (2012) considered this problem, choosing the best imputation method considering three groups of classification methods. We also try to provide an answer to this question in the Section 3 where we analysed two basic indicators needed for assessing technological capabilities of countries, which cannot be excluded from the analysis, and for which a lot of data is missing. On the examples of different countries we show how to select one appropriate out of six different single imputation techniques applied. Additionally, in order to strengthen the results and principles used in the first analysis, we performed an analysis on measuring accuracy of the applied techniques on another two indicators for the same countries for which the whole datasets exist, calculating the RMSE. This analysis enhances the claim that imputation of the missing data leads to doubtful results, enabling the manipulation.

It is important to emphasize that one imputation technique cannot always be applied for all the countries observed according to one indicator (sometimes even not for one country for the whole time period), but depends on the amount of the missing values and the nature of the existing data (e.g. whether the observed country has higher, lower or average values comparing to the others in the sample; whether we can catch the trend in the existing values which can be used for predicting the missing once), and depends on the sample (if the imputation technique is not separated from the sample). Besides these, it is important to notice that when we speak about the values for indicators for a period of time, we deal with time-series and not random values, and therefore need to be couscous when imputing. As mentioned in OECD-JRC (2008) there is no "a" method that should be used for imputing the missing values, but the method should be fitted to the characteristics of the missing information.

As Dempster and Rubin (1983, pp.3-10) remark, imputation “is seductive because it can lull the user into the pleasurable state of believing that the data are complete after all, and it is dangerous because it lumps together situations where the problem is sufficiently minor that it can legitimately be handled in this way and situations where standard estimators applied to real and imputed data have substantial bias.” As Foa and Turner (2012) highlight, there is a serious problem of legitimacy when nations are rated based on estimated, rather than actual values. In such cases, it is very difficult to guard against challenges by critics from countries which are rated poorly, arguing that the scores are inaccurate. This is why the use of imputation is more common among academicians than among international organizations such as the United Nations, the European Commission, or the World Bank.

REFERENCES

Abdella, M., & Marwala, T. (2005, April). The use of genetic algorithms and neural networks to approximate missing data in database. *In Computational Cybernetics, 2005. ICCC 2005. IEEE 3rd International Conference on* 207-212. IEEE.

Archibugi, D., & Coco, A. (2004). A new indicator of technological capabilities for developed and developing countries (ArCo). *World development*, 32(4), 629-654.

Ayuyev, V.V., Jupin, J., Harris, P.W., & Obradovic, Z. (2009). Dynamic clustering-based estimation of missing values in mixed type data. Proceedings of the *11th International Conference on Data Warehousing and Knowledge Discovery - DaWaK '09* (pp. 366-377). Springer-Verlag Berlin.

Bandura, R. (2008). A survey of composite indices measuring country performance: 2008 update. *New York: United Nations Development Programme, Office of Development Studies (Working Paper).*

Chai, T., & Draxler, R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247-1250.

Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., & Tarantola, S. (2006). Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index.*Available at SSRN 888461.*

De Mántaras, R. L. (1991). A distance-based attribute selection measure for decision tree induction. *Machine learning,* 6(1), 81-92.

Dempster, A.P, & Rubin, D.B. (1983). *Incomplete Data in Sample Surveys*, in Madow, W.G., Olkin, I., Rubin, D. (Eds),Sample Surveys Vol. II: Theory and Annotated Bibliography, Academic Press, New York.

Foa, R, & Tanner, J.C. (2012). *Methodology of the Indices of Social Development* (No. 2012-04). ISD Working Paper Series. Retrieved from http://hdl.handle.net/1765/50510

Filippetti, A., & Peyrache, A. (2011). The patterns of technological capabilities of countries: a dual approach using composite indicators and data envelopment analysis. *World Development*, *39*(7), 1108-1121.

Fujikawa, Y., & Ho, T.B. (2002). Cluster-based algorithms for dealing with missing values. *Knowledge Discovery and Data Mining Conference* (pp. 549-554). Springer, Berlin.

Grupp, H., & Mogee, M. E. (2004). Indicators for national science and technology policy: how robust are composite indicators?. *Research Policy*, 33(9), 1373-1384.

Hudrlikova, L. (2013). Composite indicators as a useful tool for international comparison: the Europe 2020 example. *Prague Economic Papers*, 22(4). doi:10.18267/j.pep.462

Little, R.J.A., & Rubin, D.B.(1987). *Statistical analysis with missing data*. John Wiley & Sons, New York.

Luengo, J., García, S., & Herrera, F. (2012). On the choice of the best imputation methods for missing values considering three groups of classification methods. *Knowledge and Information Systems*, 32(1), 77-108.

Lobato, F., Sales, C., Araujo, I., Tadaiesky, V., Dias, L., Ramos, L., & Santana, A. (2015). Multi-objective genetic algorithm for missing data imputation. *Pattern Recognition Letters*, 68, 126-131.

Nelwamondo, F.V., Mohamed, S., Marwala, T. (2007). Missing Data: A comparison of neural network and expectation maximization techniques. *Current Science* 3(11), 1514-1521.

OECD, & European Commission-Joint Research Centre. (2008). *Handbook on constructing composite indicators: Methodology and User guide*. OECD publishing.

Srebotnjak, T. (20 November 2001). *Multiple Imputations of Missing Data in the Environmental Sustainability Index - Pain or Gain?* United Nations Statistics Division, New York, USA in cooperation with the Center for International Earth Science Information Network (CIESIN), Columbia University, USA Retrieved from http://unstats.un.org/unsd/methods/wgci/ESIpubl.pdf

Zhang, S. (2011). Shell-neighbor method and its application in missing data imputation. *Applied Intelligence*, 35(1), 123-133.

THE CONJOINT ANALYSIS OF PARENTS' PREFERENCES TOWARDS CHILDREN'S PHYSICAL ACTIVITIES

Jelena Šormaz*, Marija Kuzmanović, Veljko Jeremić
University of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: jelenasormaz91@gmail.com

Abstract: *Knowing parents' attitudes and preferences, as well as the challenges they face when choosing and organising their children's activities gives us the opportunity to perceive key problems in this field and to offer possible solutions. The right choice of the activities influences the correct development of children, in the present, as well as in the future. In this study it is shown how Conjoint analysis, which has rarely been used in this field, could help in obtaining detailed insight into the current situation, parents' priorities and wishes, but also potential problems that could arise. The study may help to draw attention to the importance of this field, to point out the possibility of its research through different disciplines and to initiate further efforts towards its comprehension and development.*

Keywords**:** *parents' preferences, physical activities, Conjoint analysis, children's development, health risks*

1. INTRODUCTION

Childhood presents the key period for creating a basis for the development of all aspects of personality, primarily through different activities and experiences. This is the period during which children are gradually forming their attitudes, interests, preferences, habits which usually stay present later in the future, with more or fewer alterations.

Naturally, parents have the pivotal role in the choice of children's activities in the early childhood. Later on in life, the importance of other factors prevails, such as friends, school, and media. With time, the latter becomes more and more influential.

We can conclude that there is a combination of different factors, pressures, and limitations that influences children as well as their parents when it comes to deciding which activities the child should choose. Lack of time and financial resources, peer and media pressure represent just a few among the challenges they face on a daily basis.

The aim of this paper is to examine how they cope with these challenges, which criteria they take into account when choosing children's activities and whether the parents are aware of the importance that each activity has on their child's development. Also, we would like to find out what potential problems could arise when it comes to choosing and organising those activities. If we could understand parents' preferences, as well as the challenges they face, we would be able to point out some major issues and suggest potential solutions.

First, we will talk about the importance of the physical activities in childhood and their impact on developing permanent healthy habits. Moreover, we will discuss the possible consequences of inactivity, as well as certain limiting factors that impede parents from organising their children's activities. Also, we are going to introduce Conjoint methodology, used for obtaining the results of our survey, that will also be presented in more details within this part. Afterwards, we need to analyse the obtained results. First of all, we will get to know our sample and its characteristics. Later on, we will talk about attitudes and experiences parents shared with us in the second part of the survey. Lastly, we will present the results of the Conjoint analysis, discussing the most important conclusions about respondents' preferences. The last part of the paper is dedicated to the discussion of the new findings. We have tried to give a comprehensive overview of the key problems, their possible solutions and potential recommendations for the future research.

2. THE IMPORTANCE OF PHYSICAL ACTIVITY

The way children spend their leisure time has changed over time. Electronic entertainment represents one of the indoors activities, which takes up to 7.5 hours a day of children's free time (Hofferth & Curtin, 2003). In the last decade, the use of electronic devices increased dramatically. Even though the time dedicated to watching television is still significant, we are now witnessing the increase of usage of other media (Hofferth &

Sandberg, 2001). It is alarming that this way of life leads to the increase of chronic diseases. It is significant to emphasise that this happens despite the high level of children's participation in organised physical activities. A lot of children practise different sports, but it might be possible they are not aware of the relevance of physical activities. Therefore, they can't form healthy habits in the right way.
In our research, we decided to consider parents whose children attend primary school as our target group because the period of the primary school is vital for children's personality development. Furthermore, it enabled us to examine how physical activities are combined with school responsibilities since parents often have the tendency to believe that the school results and the level of physical activities are conflicting. In other words, they think the more time children dedicate to physical activities, the poorer results they will achieve at school (Inspiring Children's Physical Activity: Exploratory Research with Parents, 2003). This shows us why it is important to explain to parents that this is a mistake and physical activity, in fact, has a positive impact on different aspects of a child's life, even on the development of their intellectual skills and on school results presently, as well as in the future.

The beginning of primary school (between six and eight years) represents the crucial period for the development of physical capabilities (Dauer & Pangrazi, 1989; Fox, 2009) during which children form habits related to physical activity and nutrition (Jago et al., 2005). It has been proven that the habits formed during childhood stay present during adulthood as well (Malina, 1996; Twisk et al., 2000). Consequently, the acquisition of healthy habits in the early childhood has had a tremendous significance and impact on children's health in the present, as well as in the future. Adequate participation in physical activities during childhood and adolescent period represents the key factor for keeping good health in the long term (Trost, 2010).

Having in mind broad prevalence of physical inactivity and constant spreading of the obesity problem (Ogden et al., 2008), the promotion of regular activity among the youth has become the top priority of health organisations worldwide. Establishment of healthy behaviour in the early childhood constitutes a pivotal role for future habits and activities. Encouraging physical activity, especially outdoor, could be the key factor in the struggle against health risks caused by modern lifestyle.

Parents have the possibility to influence the children's behaviour, and they have a pivotal role in encouraging their physical activity (Brockman et al., 2009). The research have shown that this encouragements, together with parents' logistical support (providing resources and equipment for the activities, driving or accompanying their children to the venue where the activity is taking place), has the major impact on the children's level of activity (Davison et al., 2003; Hennessy et al., 2010; Jago et al., 2010). Bearing in mind the increase of problems with overweight and obesity, the extent to which parents influence their children's habits concerning nutrition, physical activity and passive ways of spending free time (watching television, using computers, etc.) became the issue of special interest (Pocock et al., 2010).

However, we have to be aware of the factors that limit the possibility to take part in different activities. Some of the factors parents identified as limiting were lack of time (Hesketh et al., 2005; Mitchell et al., 2011), lack of resources and excessive costs (Mitchell et al., 2011; Smith, 2005), security, violent peers, as well as the peers' pressure to sedentary activities (such as watching television and computer usage), lack of adequate program opportunities and offers (Smith, 2005) and a great distance to the venue of the training (Hesketh et al., 2005). Parents' experiences show that these obstacles make it difficult or even impossible for them to involve their children in different activities.

Since the sedentary way of life has become more and more prominent in the modern life, we can notice the increase of its consequences on health and other aspects of children's personality. It is important to point out these consequences continue later on in life because the habits acquired in childhood continue in the mature age as well. That is one of the main reasons why we decided to focus on the physical activity in our research.

Understanding the social and environmental factors that influence physical activity play vital role in developing effective measures and initiatives for increasing the level of children's and adolescent's activity. On the one hand, our results could encourage parents to pay more attention to this topic and help them to get a broader insight into its importance. At the same time, these results can also be beneficial to schools, relevant ministries, and organisations when it comes to starting different initiatives. They could help them create innovative programs and focus on the key matters revealed during the research.

3. METHODOLOGY AND RESEARCH

3.1. Conjoint analysis

Our goal was to determine which factor is the most important between different types of training, as well as which combination of characteristics is the most appealing one for parents. To be able to answer these questions, we used Conjoint analysis, which reveals respondents' preferences through the simulation of the real problem of choice.

The previously mentioned term was introduced in 1978 by Green and Srinivasan. However, there isn't a clear-cut definition of this term, but several are provided. Hair (Hair et al., 1998.) defined it in his paper saying that 'Conjoint analysis is a multivariate technique used specifically to understand how respondents develop preferences for products or services. It is based on the simple premise that consumers evaluate the value of a product or service by combining the separate amounts of value provided by each attribute.' Sudman and Blair (1998) pointed out that it must be regarded as a type of 'thought experiment' designed to show how various elements of products or services (price, brand, style) predict customer preferences for a product or service.

The central assumption of the decomposition approach is that consumers evaluate overall utility of product/services combining part-worth utilities of various levels for each attribute which describe a certain product. Those partial utilities are defined as numeric values that show the level of appeal of different characteristics of product/services. When a consumer chooses between different products, they don't observe only one isolated characteristic, but they make their decision by observing (perceiving) and analysing its characteristics all at once. The name of the method comes from that fact – CONsidered JOINTly (Kuzmanovic, 2006).

The researcher defines some key attributes of the analysed entity. Then he needs to choose several different levels, variation for each attribute. Each combination of the different levels of attributes forms one profile or concept (Kuzmanovic, 2006).

Two main goals of the Conjoint Analysis are:

- To evaluate the contribution (influence) of each attribute on consumers' overall preferences
- To create a valid model of consumers' judgment which would be helpful in forecasting his reactions to the changes in the characteristics of the existing product or introduction of a new product to the market.

This method represents very well real conditions and the consumers' behaviour on the market. It simulates the whole process of choice in a certain way. Upon completing the survey, the respondent has to put themselves in the position of a consumer and to truly analyse and evaluate all of the presented offers according to their preferences, needs, and priorities.

The process of implementation of the Conjoint analysis study consists of the following key steps:

1. Attribute List Formulation - A business problem is defined, and an attribute (features) list, as well as their performance levels, is developed to study the problem.
2. Data collection - Respondents are asked to express the trade-offs they are willing to make among product features by rating, sorting or choosing among hypothetical product concepts.
3. Utility calculation - A set of preference values (also called part worth utilities or part-worths) is derived from the interview data; they reflect the trade-offs each respondent has made.
4. Market Simulation.

Owing to the Conjoint Analysis, we can understand which characteristics are the most important to them when choosing a product/service. Furthermore, we are able to see which level of each attribute is the most appealing to them. In this way, we are able to adjust the offer to their needs and preferences. We can identify the key factors which they pay attention to the most and which affect their decisions and choices. In addition, we can modify them in order to create the optimal mix of characteristics. In that way, we maximise consumers' satisfaction, as well as we maximise our opportunities/income/success. We get a realistic image of the needs and wishes on the market.

This method has rarely been applied in the field of children's development. However, considering the fact that our main goal is to get an insight into parents' preferences, it is undoubtedly suitable for this research.

What we are expecting to achieve with this choice of method is to shed the light on the new possibilities a contribution that similar multidisciplinary research could bring, broadening the knowledge and understanding of this field.

3.2. Research

During October 2015, 109 parents participated in the survey. The surveys were distributed online, using various social networks, as well as in the paper version in several primary schools in Serbia.

We decided to focus on the parents whose children attended primary schools since we have learned from the referenced literature that this was the critical period for children's development. The children included in this study were between 6 and 14 years old.

The survey consisted of two parts. In the first part, demographic questions and issues that concern parents' attitudes and experiences towards their children's physical activity were included. The second part comprised 16 scenarios analysed by the Conjoint analysis. Each scenario represented the combination of various levels of seven different attributes. All attributes with all their levels are the following:

- Type of sport:
  - Individual
  - Team
- Contact:
  - With contact
  - Without contact
- Training intensity:
  - High
  - Medium
  - Low
- Frequency of training:
  - Once a week
  - 2 – 3 times a week
  - More than 3 times a week
- Parents involvement:
  - Children need to be accompanied by their parents to the venue
  - Children don't need to be accompanied by their parents to the venue
- Possibility of competition:
  - Existing
  - Non existing
- Monthly price:
  - 10% above the average
  - Average
  - 10% below the average

Parents were asked to evaluate each offer of training (5-I like it very much, ..., 1-I do not like it at all) based on their preferences and priorities.

4. RESULTS

As it was expected, mothers represented the majority of the sample. 88 mothers and 21 fathers have completed our surveys. 88 girls and 70 boys have been included in the research. The average age of children was 10.97 years. 74.3% of parents were from the territory of Belgrade, and the remaining 25.7% came from all over Serbia, mostly from the province of Vojvodina, 16% of them. An average household income was shown to be on a rather high level, as 25.7% of parents said their average monthly income was 120.000 RSD.

In the following part of the survey, we focused on children's physical activity. The results have shown that 75.6% of children practice some kind of sport, and the team sports stand out as the most popular ones. Figure 1 illustrates the distribution of children's participation in different types of sports.

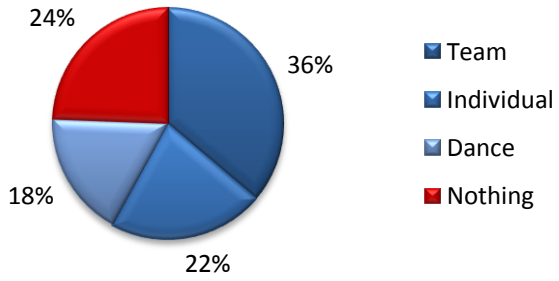

Figure 1: Children's organised physical activities

It has also been shown that there is statistically significant difference ($p<0.05$) when it comes to comparing the level of activity between boys and girls, girls being slightly less active than boys.

It was interesting to examine whether there has been some statistically significant difference between the level of activity of parents and their children. It has been found that such difference exists ($p<0.01$) and that there is a direct correlation between these two variables (parents' level of activity and child's level of activity). In other words, if parents practise, or had practiced some kind of sport, there is much greater possibility for their children to practice some sport, as well. Numerous research has confirmed this pattern – the more parents are active, the more their children will also be (Moore et al., 1991; Anderson et al., 2009).

Some of the questions concerned the parents' satisfaction with the level of their children's physical activity. The obtained results indicate that the parents are to a great extent satisfied with this level, since more than 78% of them declared as satisfied, as opposed to 5% of highly unsatisfied respondents. Moreover, it has been shown that there is no statistically significant difference between the extent of satisfaction between the boys' parents and girls' parents. It is interesting to mention that some research have shown that there has been noticed a lower level of activities among girls, but their parents did not express any discontent, since they considered it natural for girls to lose their interest in sport at some point, while the boys were more forced to be physically active (Anderson et al., 2010). Besides, it has been noted that parents who are or were actively practicing some kind of sport were somewhat more satisfied with the level of their children's activity, which is another proof that the physical activities of parents and their children are directly correlated.

The parents who were not entirely satisfied were asked to evaluate the influence of certain factors that could be responsible for children's insufficient activity. Each factor was evaluated using the 1 – 4 scale, where 1 represented 'completely without influence', and 4 'highly influential'. The results show that the lack of financial resources was identified as the greatest problem, since almost 45% of parents considered this factor as highly influential. The lack of child's free time was recognised as another important obstacle. It has been shown that the lack of adequate programs was considered as one of the least influential factors, which is consistent with the results of other similar research around the world. In these research respondents also considered the offer of programs quite satisfying, while they saw the price as the greatest obstacle (Inspiring Children's Physical Activity: Exploratory Research with Parents, 2003). The lack of child's interest was shown to be the least influential, but it is important to point out that generally there are no significant differences in the extent of influence. What we found interesting was the fact that there was statistically significant difference ($p<0.05$) between the mothers' and fathers' responses. Fathers considered the lack of parents' time to be far more influential than mothers did while for mothers the lack of financial resources represented a more important obstacle than it was for fathers.

Since the finance was overall seen as the most influential factor, we were curious to see how much money are parents willing to devote to children's training on a monthly basis. The range of the amount of money was between 0 and 15 000 RSD while the average amount was 3945.4 RSD. It has been demonstrated that there was a statistically significant difference between the sums. The respondents from Belgrade were willing to invest, as opposed to the respondents from the rest of the country. The results of t-test have shown that the value of statistic was $t = 2.403$, while the p-value was $p = 0.018$ ($F = 5.416$, $p = 0.022$). The residents of Belgrade were on average willing to invest 4258.2 RSD while this amount was 2928.7 RSD when it comes to the respondents outside the capital.

Taking into account the challenges parents are facing in everyday life, such as stress, a multitude of obligations and responsibilities, etc., we wanted to examine to which extent these conditions influence the choice of children's activities. Foreign research has shown that parents have less time for organising and participating in those activities. Furthermore, some parents have admitted that sometimes they prefer that their children are less active. When they return home from work, they do not need any additional activities,

and consequently, it is easier for them that their children spend time using a computer or watching television (Inspiring Children's Physical Activity: Exploratory Research with Parents, 2003). However, although sometimes it seems easier to turn to unhealthy habits, parents have the responsibility to guide their children towards a healthy lifestyle (Sallis et al., 2000).

Parents completing our survey were asked to say whether they agreed with the given statements considering the difficulties they could have in organising and arranging children's activities. 52.3% of parents admitted they find it difficult to cope with this task and that they found it highly challenging while the rest of 47.7% of the respondents stated that they were not perceiving it as a difficulty.

After gaining insight into parents' experiences and opinions, we turned to Conjoint analysis in order to comprehend the way they choose activities for their children. We wanted to learn which criteria were the most important to them and how their priorities and preferences had reflected to their final choice.

Using Conjoint method, we came to a conclusion that the most desirable combination of characteristics which forms the optimal training type is the following:

- Team sport
- Without contact
- High intensity
- 2-3 times a week
- Children do not need parents' help to get to the venue
- Possible competitions
- Monthly price 10% above average

Among the list of the variables, the most significant ones that stand out and influence parents' choices the most are the frequency of training, price per month and training intensity. The least important one for parents is the question of a contact or non-contact sport. In the following graph (Figure 2) we can see the average importance of each variable.

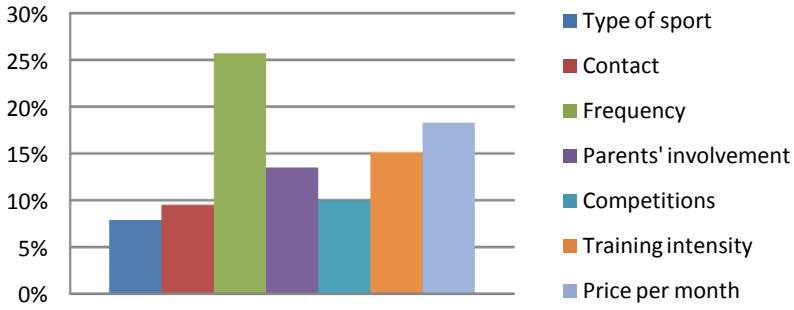

Figure 2: Average importance of attributes on aggregated level

As we can see in the following graph showing the partial utilities for each level of attributes (Figure 3), parents prefer training 2-3 times per week, and they are very sensitive to changes of this level. Furthermore, it is very important for them that children can go alone to the venue where the training is held and that the price is below the average. If these levels change, the attractiveness of the offer reduces.

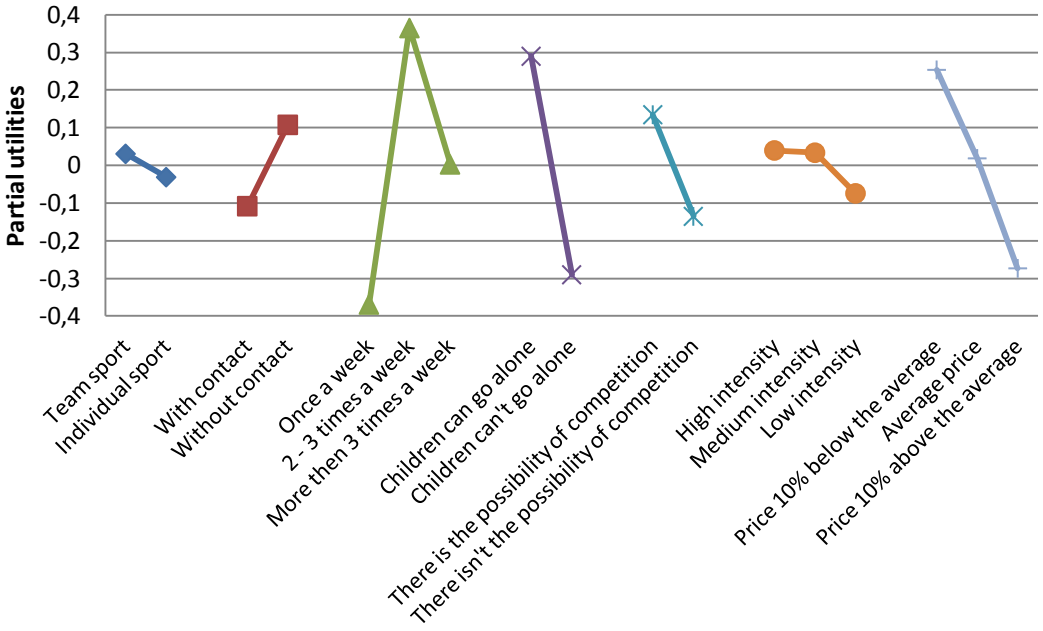

Figure 3: Partial utilities of levels of attributes

Values of Pearson's and Kendall's coefficients confirm the high level of significance and credibility of results. ($R=0.987$, $p<0.01$; $\tau=0.929$, $p<0.01$).

4. CONCLUSION

This paper deals with a very current topic in order to indicate the importance of physical activity during childhood, family support and encouragement in the process of child's development and the problems arising in this field.
Although in Serbia is not paid much attention to this issue, numerous research with a different focus (health, family relations, academic achievements, etc.) has been conducted around the world. What makes our research stand out among other similar ones is the fact we used Conjoint analysis to determine parents' preferences, which enabled us to reveal some attitudes and opinions that were not obvious instantaneously.
We found out that parents were mostly satisfied with the level of their children's activity, although they were worried about the sedentary way of life. Furthermore, the research has demonstrated that most parents find the lack of financial resources as the most limiting factor when choosing activities for their children.

Bearing in mind the conditions brought by the modern way of life – fast technological changes influencing all aspects of life, more stress and pressures from the environment, it is extremely important to draw the attention of parents to the significance of dedicating their time, resources and effort to their children's activities. Moreover, it is essential to help them organise and use their free time, as well as their children's free time, in the most effective way and, since the focus of the research was directed towards physical activity as one of the most important and simultaneously the most endangered nowadays, it is also vital to make an influence on parents and children so that they could develop and maintain their healthy habits. When it comes to parents, one of the possible strategies could be using websites, specialised magazines, and social networks to promote not only the positive effect of sport on health but also the risks of constant inactivity. Furthermore, school programs should encompass informing and educative initiatives dedicated to physical activity. The sport should be presented as a healthy habit, an essential part of everyday life, and not only as a secondary competitive activity.

Analysing the results, we came to the conclusion that if we want to make our children more active, what we need to do is to start sports programs that would not be financially demanding, making them affordable to the larger number of children.

Although we came to some important findings, there is still a wide range of opportunities when it comes to investigating this topic. For instance, it would be recommendable to include younger children in the future research. One of the reason is the fact that the processes of personality development begin long before primary school, and so we would be able to examine the effects of different factors present in the earlier period, which would give us the opportunity to react on time according to those potential new findings. Another important benefit would be the possibility to compare preschool and school children and to discover

potential differences and its causes. Moreover, focusing only on school children, we missed out the opportunity to investigate the influence of younger or older children on their parent's opinions and family dynamic in general.

Also, it would be desirable that the future research cover greater and more heterogeneous sample, not only to include the children of different age but also to reach more children out of Belgrade, so that we could have more comprehensive insight into the current state in this field.

Apart from that, having in mind the modern way of life, it would be truly prudent to thoroughly examine children's relation to the new media, especially the time they dedicate to the new technologies and its influence on different life aspects.

In conclusion, we want to emphasise once more that this is the topic that demands attention due to its importance and influence in present, as well as in the future, when we would be able to perceive the consequences of the present (in)activity. This is why it is crucial to deal with this problem now, to dedicate our attention to it and resources if we want to ensure the proper development of the new generations, in spite of the increasing challenges.

REFERENCES

Anderson, C., B., Hughes, S., O., Fuemmeler, B., F. (2009). Parent–child attitude congruence on type and intensity of physical activity: Testing multiple mediators of sedentary behavior in older children, *Health Psychology*, Vol 28(4), Jul, 428-438. Retrieved from http://dx.doi.org/10.1037/a0014522 (25.09.2015.)

Brockman, R., Jago, R., Fox, K., R., Thompson, J., L., Cartwright, K., Page, A., S. (2009). *Get off the sofa and go and play: family and socioeconomic influences on the physical activity of 10–11 year old children*, BMC Publ Health, 21:253.

Dauer, V., P., Pangrazi, R., P. (1989*). Dynamic Physical Education for Elementary School Children*, 9th edition. New York: Macmillan

Davison, K., K., Cutting, T., M., Birch, L., L. (2003). Parents' activity-related parenting practices predict girls' physical activity, *Med Sci Sports Exerc*, US National Library of Medicine, National Institutes of Health, 35 (9):1589–1595.

Fox, K., R.(2009). How to help your children become more active. In Gonzalez-Gross M. (Ed.), *Active healthy living: A guide for parents*, Brussels: Coca Cola Europe, 52–67.

Hair J.F., Anderson, R.E., Tatham, R.L., Black, W.C., (1998). *Multivariate Data Analysis*, 5th Edition, Upper Saddle River, New Jersey, Prentice Hall International

Hennessy, E., Hughes, S., O., Goldberg, J., P., Hyatt, R., R., Economos, C., D. (2010). *Parent–child interactions and objectively measured child physical activity: a cross-sectional study*, IJBNPA, 7:71.

Hesketh, K., Waters, E., Green, J., Salmon, L., Williams, J. (2005). Healthy eating, activity, and obesity prevention: a qualitative study of parent and child perceptions in Australia, *Health Promot Int*, US National Library of Medicine, National Institutes of Health, 20:19–26.

Hofferth, S., and Sandberg, J. (2001). Changes in American Children's Time, 1981-1997. In S. Hofferth & T. Owens (Eds.), *Children at the Millennium: Where did we come from, where are we going?*. New York: Elsevier Science.

Hofferth, S., L. and Curtin, S., C. (2003). Leisure Time Activities in Middle Childhood, *Positive Outcomes Conference*, Washington, DC, March 12-13

Inspiring Children's Physical Activity: Exploratory Research with Parents Prepared for: U.S. Department of Health and Human Services Centers for Disease Control and Prevention VERB™ Campaign, Prepared by: SpencerHall, May 2003

Jago, R., Baranowski, T., Baranowski, J., C., Thompson, D., Greaves, K., A. (2005). BMI from 3–6 years of age is predicted by TV viewing and physical activity, not diet., *Int J Obes Relat Metab Disord*, US National Library of Medicine, National Institutes of Health, 29(6):557–564.

Jago, R., Davison, K., K., Brockman, R., Page, A., S., Thompson, J., L., Fox, K., R. (2010). Parenting styles, parenting practices, and physical activity in 10- to 11-year olds, *Prev Med*, US National Library of Medicine, National Institutes of Health, 52:44–47.

Kuzmanović, M. (2006). *Kvantitativne metode u marketingu: Primena Conjoint analize*, Društvo operacionih istraživača, Beograd

Malina, R., M. (1996). Tracking of physical activity and physical fitness across the lifespan. *Res Q Exerc Sport*,US National Library of Medicine, National Institutes of Health, 67(3):48–57.

Mitchell, J., Skouteris, H., McCabe, M., Ricciardelli, L., A., Milgrom, J., Baur, L., A., Fuller- Tyszkiewicz, M., Dwyer, G. (2011). Physical activity in young children: a systematic review of parental influences, *Early Child Dev Care*, US National Library of Medicine, National Institutes of Health, 82(11):1–27.

Moore, L.L., Lombardi, D.A., White, M.J., Campbell, J.L., Oliveria, S.A., Ellison, R.C. (1991). Influence of parents' physical activity levels on activity levels of young children., *The Journal of Pediatrics*, Feb;118(2):215-9.

Ogden, C., L., Carroll, M., D., Flegal, K., M. (2008). High body mass index for age among US children and adolescents, 2003-2006, *JAMA*, US National Library of Medicine, National Institutes of Health, 299:2401-2405.

Pocock, M., Trivedi, D., Wills, W., Bunn, F., Magnusson, J. (2010). Parental perceptions regarding healthy behaviours for preventing overweight and obesity in young children: a systematic review of qualitative studies, *Obes Rev*, US National Library of Medicine, National Institutes of Health;11(5):338-353.

Sallis, J., F., Prochaska, J., J., Taylor, W., C. (2000). A review of correlates of physical activity of children and adolescents, *Med Sci Sports Exerc*, US National Library of Medicine, National Institutes of Health, 32(5):963–975.

Smith, M., (2005). Parental influences on the physical activity behaviour of children of various ethnic backgrounds, *Res Q Exerc Sport*, US National Library of Medicine, National Institutes of Health, 76(1):50–51.

Sudman, S., Blair, E. 1998. *Marketing research: A problem-solving approach*. New York: McGraw–Hill.

Trost, S., G. (2010). *Discussion paper for the development of recommendations for children's and youth's participation in health promoting physical activity*. Retrieved from http://espace.library.uq.edu.au/view/UQ:190465. (25.09.2015.)

Twisk, J., W., Kemper, H., C., G., van Mechelen, W. (2000). Tracking of activity and fitness and the relationship with cardiovascular disease risk factors. *Med Sci Sports Exerc*, US National Library of Medicine, National Institutes of Health, 32(8):1455–1461.