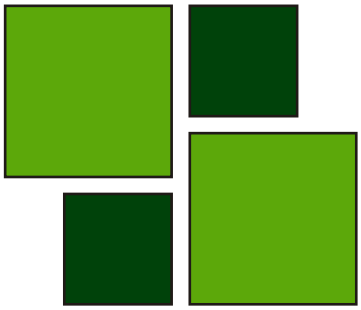


WWW.JAQM.RO

JOURNAL OF APPLIED QUANTITATIVE METHODS

Quantitative Methods Inquires

Vol. 10
No. 3
Fall
2015

ISSN 1842–4562

JAQM Editorial Board

JAQM Advisory Board

JOURNAL OF APPLIED QUANTITATIVE METHODS

Contents

JOURNAL OF APPLIED QUANTITATIVE METHODS

Quantitative Methods Inquires

TEST STATISTICS FOR COMPARING TWO PROPORTIONS WITH PARTIALLY OVERLAPPING SAMPLES

Ben DERRICK[1]
PhD Candidate
University of the West of England, Bristol

E-mail:

Anselma DOBSON-MCKITTRICK[2]
University of the West of England, Bristol

E-mail:

Deirdre TOHER[3]
PhD, Senior Lecturer
University of the West of England, Bristol

E-mail:

Paul WHITE[4]
PhD, Associate professor
University of the West of England, Bristol

E-mail: paul.white@uwe.ac.uk

Abstract

Standard tests for comparing two sample proportions of a dichotomous dependent variable where there is a combination of paired and unpaired samples are considered. Four new tests are introduced and compared against standard tests and an alternative proposal by Choi and Stablein (1982). The Type I error robustness is considered for each of the test statistics. The results show that Type I error robust tests that make use of all the available data are more powerful than Type I error robust tests that do not. The Type I error robustness and the power among tests introduced in this paper using the phi correlation coefficient is comparable to that of Choi and Stablein (1982). The use of the test statistics to form confidence intervals is considered. A general recommendation of the best test statistic for practical use is made.

Key words: Partially overlapping samples, Partially matched pairs, Partially correlated data, Equality of proportions

1. Introduction

Tests for comparing two sample proportions of a dichotomous dependent variable with either two independent or two dependent samples are long established. Let π_1 and π_2 be the proportions of interest for two populations or distributions. The hypothesis being tested is $H_0 : \pi_1 = \pi_2$ against $H_1 : \pi_1 \neq \pi_2$. However, situations arise where a data set comprises a combination of both paired and unpaired observations. In these cases, within a sample there are, say

a total of ' n_{12} ' observations from both populations, a total of ' n_1 ' observations only from population one, and a total of ' n_2 ' observations only from population two. The hypothesis being tested is the same as when either two complete independent samples or two complete dependent samples are present. This situation with respect to comparing two means has been treated poorly in the literature (Martinez-Camblor et al, 2012). This situation with respect to comparing proportions has similarly been poorly treated.

Early literature in this area with respect to comparing proportions, refers to paired samples studies in the presence of incomplete data (Choi and Stablein, 1982; Ekbohlm, 1982), or missing data (Bhoj, 1978). These definitions have connotations suggesting that observations are missing only by accident. Recent literature for this scenario refers to partially matched pairs (Samawi and Vogel, 2011), however this terminology may be construed as the pairs themselves not being directly matched. Alternatively, the situation outlined can be referred to as part of the 'partially overlapping samples framework' (Martinez-Camblor et al, 2012). This terminology is more appropriate to cover scenarios where paired and independent samples may be present by accident or design. Illustrative scenarios where partially overlapping samples may arise by design include:

i) Where the samples are taken from two groups with some common element. For example, in education, when comparing the pass rate for two optional modules, where a student may take one or both modules.

ii) Where the samples are taken at two points in time. For example, an annual survey of employee satisfaction will include new employees that were not employed at time point one, employees that left after time point one and employees that remained in employment throughout.

iii) When some natural pairing occurs. For example, a survey taken comparing views of males and females, there may be some matched pairs 'couples' and some independent samples 'single'.

Repeated measures designs can have compromised internal validity through familiarity (e.g. learning, memory or practise effects). Likewise, a matched design can have compromised internal validity through poor matching. However, if a dependent design can avoid extraneous systematic bias, then paired designs can be advantageous when contrasted with between subjects or independent designs. The advantages of paired designs arise by each pair acting as its own control helping to have a fair comparison. This allows differences or changes between the two samples to be directly examined (i.e. focusing directly on the phenomenon of interest). This has the result of removing systematic effects between pairs. This leads to increased power or a reduction in the sample size required to retain power compared with the alternative independent design. Accordingly, a method of analysis for partially overlapping samples that takes into account any pairing, but does not lose the unpaired information, would be beneficial.

Historically, when analysing partially overlapping samples, a practitioner will choose between discarding the paired observations or discarding the independent observations and proceeding to perform the corresponding 'standard' test. It is likely the decision will be based on the sample sizes of the independent and paired observations. Existing 'standard' approaches include:

Option 1: Discarding all paired observations and performing Pearson's Chi square test of association on the unpaired data.

Option 2: Discarding all unpaired observations and performing McNemar's test on the paired data.

Option 3: Combining p-values of independent tests for paired and unpaired data. This can be done by applying Fisher's inverse Chi square method or Tippett's test. These approaches make use of all of the available data. These techniques were considered by Samawi and Vogel (2011) and are shown to be more powerful than techniques that discard data. However, it should be noted that the authors did not consider Type I error rates.

Other ad-hoc approaches for using all available data include randomly pairing any unpaired observations, or treating all observations as unpaired ignoring any pairing. These ad-hoc approaches are clearly incorrect practice and further emphasise the need for research into statistically valid approaches.

Choi and Stablein (1982) performed a small simulation study to consider standard approaches and ultimately recommended an alternative test making use of all the available data as the best practical approach. This alternative proposal uses one combined test statistic weighting the variance of the paired and independent samples, see Section 3.2 for definition. The authors additionally considered an approach using maximum likelihood estimators for the proportions. This approach was found to be of little practical benefit in terms of Type I error rate or power. Others have also considered maximum likelihood approaches. For example Thomson (1995) considered a similar procedure, using maximum likelihood estimators, and found the proposed procedure to perform similarly to that of Choi and Stablein (1982). It was noted by Choi and Stablein (1982) that given the additional computation, the maximum likelihood solution would not be a practical solution.

Tang and Tang (2004) proposed a test procedure which is a direct adaption of the best practical approach proposed by Choi and Stablein (1982). This adaption is found to be not Type I error robust in scenarios considered when $n_1 + n_2 + 2n_{12} = 20$. The test proposed by Choi and Stablein (1982) is found to be Type I error robust in this scenario. The literature reviewed suggests that a solution to the partially overlapping samples case will have to outperform the best practical solution by Choi and Stablein (1982). Tang and Tang (2004, p.81) concluded that, 'there may exist other test statistics which give better asymptotic or unconditional exact performance'.

In this paper, we introduce four test statistics for comparing the difference between two proportions with partially overlapping samples. These test statistics are formed so that no observations are discarded. The statistics represent the overall difference in proportions, divided by the combined standard error for the difference.

This paper will explore test statistics for testing H_0, in the presence of partially overlapping samples. In Section 2, existing 'standard' approaches and variants of are defined. In Section 3, our alternative proposals making use of all the available data are then introduced, followed by the most practical proposal of Choi and Stablein (1982).

In Section 4, a worked example applying all of the test statistics is given, followed by the simulation design in Section 5.

In Section 6.1, for all of the test statistics, the Type I error robustness is assessed when H_0 is true. This is measured using Bradley's (1978) liberal criteria. This criteria states that the Type I error rate should be between $\alpha_{\text{nominal}} \pm 0.5\,\alpha_{\text{nominal}}$.

There is no standard criteria for quantifying when a statistical test can be deemed powerful. The objective is to maximise the power of the test subject to preserving the Type I

error rate α_{nominal}. If Type I error rates are not equal it is not possible to correctly compare the power of tests. The preferred test where Type I error rates are not equal should be the one with the Type I error rate closest to α_{nominal} (Penfield 1994). In Section 6.2, power will be considered under H_1 for the test statistics that meet Bradley's liberal criteria.

There is frequently too much focus on hypothesis testing. Confidence intervals may be of more practical interest (Gardner and Altman 1986). Confidence intervals allow insight into the estimation of a difference and the precision of the estimate. In Section 6.3, the coverage of the true difference under H_1 within 95% confidence intervals is considered. This is considered only for the most powerful test statistics that are Type I error robust.

2. Definition of standard test statistics

Assuming a dichotomous dependent variable, where a comparison in proportions between two samples is required, the layout of frequencies for the paired and the independent samples would be as per Table 1 and Table 2 respectively.

Table 1. Paired samples design for two samples and one dichotomous dependent variable

	Response Sample 2		
Response Sample 1	Yes	No	Total
Yes	*a*	*b*	*m*
No	*c*	*d*	n_{12} - *m*
Total	*k*	n_{12} -*k*	n_{12}

Table 2. Independent samples design for two samples and one dichotomous dependent variable

	Response		
	Yes	No	Total
Sample 1	*e*	*f*	n_1
Sample 2	*g*	*h*	n_2

2.1. Option 1: Discarding all paired observations

For two independent samples in terms of a dichotomous variable, as per Table 2, a Chi-square test of association is typically performed. This test will be displayed in standard textbooks in terms of χ_1^2. A chi square distribution on one degree of freedom is equivalent to the square of the z-distribution. Therefore under the null hypothesis an asymptotically N(0,1) equivalent statistic is defined as:

$$z_1 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} \quad \text{where } \hat{p}_1 = \frac{e}{n_1},\ \hat{p}_2 = \frac{g}{n_2} \text{ and } \hat{p} = \frac{e+g}{n_1+n_2}.$$

For small samples, Yates's correction is often performed to reduce the error in approximation. Yate's correction is given by:

$$z_2 = \sqrt{\frac{(n_1+n_2)((|eh-fg|-0.5(n_1+n_2))^2}{(e+g)(f+h)n_1n_2}}.$$

The statistic z_2 is referenced against the upper tail of the standard normal distribution.

An alternative to the Chi square approach is Fisher's exact test. This is computationally more difficult. Furthermore, Fisher's exact test is shown to deviate from Type I error robustness (Berkson, 1978). Fisher's exact test will not be considered for the analysis of the partially overlapping samples design in this paper.

2.2. Option 2: Discarding all unpaired observations

For two dependent samples in terms of a dichotomous variable, as per Table 1, McNemar's test is typically performed. Under the null hypothesis, the asymptotically N(0,1) equivalent to McNemar's test is:

$$z_3 = \frac{b-c}{\sqrt{b+c}}.$$

When the number of discordant pairs is small, a continuity correction is often performed. McNemar's test with continuity correction is the equivalent to:

$$z_4 = \sqrt{\frac{(|b-c|-1)^2}{b+c}}.$$

The statistic z_4 is referenced against the upper tail of the standard normal distribution.

Test statistics based on Option 1 and Option 2 are likely to have relatively low power for small samples when the number of discarded observations is large. A method of analysis for partially overlapping samples that takes into account the paired design but does not lose the unpaired information could therefore be beneficial.

2.3. Option 3: Applying an appropriate combination of the independent and paired tests using all of the available data

Given that test statistics for the paired samples and dependent samples can be calculated independently, an extension to these techniques which makes use of all of the available data would be some combination of the two tests.

In terms of power, Fisher's test and Tippett's test are comparable to a weighted approach using sample size as the weights (Samawi and Vogel, 2011). Tippett's method and Fisher's method are not as effective as Stouffer's weighted z-score test (Kim et al, 2013). Stouffer's weighted z-score, for combining z_1 and z_3 is defined as:

$$z_5 = \frac{wz_1 + (1-w)z_3}{\sqrt{w^2+(1-w)^2}} \text{ where } w = \frac{n_1+n_2}{2n_{12}+n_1+n_2}.$$

Under the null hypothesis, the test statistic z_5 is asymptotically N(0,1).

Many other procedures for combining independent p-values are available, but these are less effective than Stouffer's test (Whitlock, 2005).

The drawbacks of Stouffer's test are that it has issues in the interpretation and confidence intervals for the true difference in population proportions cannot be easily formed.

3. Definition of alternative test statistics making use of all of the available data

The following proposals are designed to overcome the drawbacks identified of the standard tests. In these proposals observations are not discarded and the test statistics may be considered for the formation of confidence intervals.

3.1. Proposals using the phi correlation or the tetrachoric correlation coefficient

It is proposed that a test statistic for comparing the difference in two proportions with two partially overlapping samples can be formed so that the overall estimated difference in proportions is divided by its combined standard error, i.e.

$$\frac{\bar{p}_1 - \bar{p}_2}{\sqrt{Var(\bar{p}_1) + Var(\bar{p}_2) - 2r_x Cov(\bar{p}_1, \bar{p}_2)}}$$

where $Var(\bar{p}_1) = \frac{\bar{p}_1(1-\bar{p}_1)}{n_{12}+n_1}$, $Var(\bar{p}_2) = \frac{\bar{p}_2(1-\bar{p}_2)}{n_{12}+n_2}$,

$$Cov(\bar{p}_1, \bar{p}_2) = \frac{\sqrt{\bar{p}_1(1-\bar{p}_1)}\sqrt{\bar{p}_2(1-\bar{p}_2)}n_{12}}{(n_{12}+n_1)(n_{12}+n_2)}$$

and r_x is a correlation coefficient.

Test statistics constructed in this manner will facilitate the construction of confidence intervals, for example a 95% confidence interval θ would be equivalent to:

$$\theta = (\bar{p}_1 - \bar{p}_2) \pm 1.96 \times \sqrt{Var(\bar{p}_1) + Var(\bar{p}_2) - 2r_x Cov(\bar{p}_1, \bar{p}_2)}\,.$$

Pearson's phi correlation coefficient or Pearson's tetrachoric correlation coefficient are often used for measuring the correlation r_x between dichotomous variables.

Pearson's phi correlation coefficient is calculated as $r_1 = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$.

The result of r_1 is numerically equivalent to Pearson's product-moment correlation coefficient and Spearman's rank correlation coefficient applied to Table 1, using binary outcomes '0' and '1' in the calculation. In this 2×2 case, r_1 is also numerically equivalent to Kendall's Tau-a and Kendall's Tau-b as well as Cramér's V and Somer's *d* (symmetrical). This suggests that r_1 would be an appropriate correlation coefficient to use.

Alternatively, assuming the underlying distribution is normal, a polychoric correlation coefficient may be considered. A special case of the polychoric correlation coefficient for two dichotomous samples is the tetrachoric correlation coefficient.

An approximation to the tetrachoric correlation coefficient as defined by Edwards and Edward (1984) is:

$$r_2 = \frac{s-1}{s+1} \text{ where } s = \left(\frac{ad}{bc}\right)^{0.7854}.$$

Other approximations are available, however there is no conclusive evidence which is the most appropriate (Digby, 1983). In any event, r_1 is likely to be more practical than r_2 because if any of the observed paired frequencies are equal to zero then the calculation of r_2 is not possible.

Constructing a test statistic using correlation coefficients r_1 and r_2 respectively, the following test statistics are proposed:

$$z_6 = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_{12}+n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_{12}+n_2} - 2r_1\left(\frac{\sqrt{\bar{p}_1(1-\bar{p}_1)}\sqrt{\bar{p}_2(1-\bar{p}_2)}n_{12}}{(n_{12}+n_1)(n_{12}+n_2)}\right)}}$$

$$z_7 = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_{12}+n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_{12}+n_2} - 2r_2\left(\frac{\sqrt{\bar{p}_1(1-\bar{p}_1)}\sqrt{\bar{p}_2(1-\bar{p}_2)}n_{12}}{(n_{12}+n_1)(n_{12}+n_2)}\right)}}$$

where: $\bar{p}_1 = \frac{a+b+e}{n_{12}+n_1}$ and $\bar{p}_2 = \frac{a+c+g}{n_{12}+n_2}$.

Under H_0, $\pi_1 = \pi_2 = \pi$, therefore two additional test statistics that may be considered are defined as:

$$z_8 = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_{12}+n_1} + \frac{\bar{p}(1-\bar{p})}{n_{12}+n_2} - 2r_1\left(\frac{\sqrt{\bar{p}(1-\bar{p})}\sqrt{\bar{p}(1-\bar{p})}n_{12}}{(n_{12}+n_1)(n_{12}+n_2)}\right)}}$$

$$z_9 = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_{12}+n_1} + \frac{\bar{p}(1-\bar{p})}{n_{12}+n_2} - 2r_2\left(\frac{\sqrt{\bar{p}(1-\bar{p})}\sqrt{\bar{p}(1-\bar{p})}n_{12}}{(n_{12}+n_1)(n_{12}+n_2)}\right)}}$$

where $\bar{p} = \frac{(n_1+n_{12})\bar{p}_1 + (n_2+n_{12})\bar{p}_2}{2n_{12}+n_1+n_2}$.

The test statistics z_6, z_7, z_8 and z_9 are referenced against the standard normal distribution.

In the extreme scenario of $n_{12} = 0$, it is quickly verified that $z_8 = z_9 = z_1$. Under H_0 in the extreme scenario of $n_1 = n_2 = 0$, as $n_{12} \to \infty$ then $z_8 \to z_3$. This property is not observed for z_9. The properties of z_8 give support from a mathematical perspective as a valid test statistic to interpolate between the two established statistical tests where overlapping samples are not present.

3.2. Test statistic proposed by Choi and Stablein (1982)

Choi and Stablein (1982) proposed the following test statistic as the best practical solution for analysing partially overlapping sample:

$$z_{10} = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left\{\frac{\psi_1^2}{n_1} + \frac{(1-\psi_1)^2}{n_{12}} + \frac{\psi_2^2}{n_2} + \frac{(1-\psi_2)^2}{n_{12}}\right\} - 2D}}$$

where $\psi_1 = \frac{n_1}{n_1 + n_{12}}$, $\psi_2 = \frac{n_2}{n_2 + n_{12}}$ and $D = \frac{(1-\psi_1)(1-\psi_2)(p_a - \bar{p}^2)}{n_{12}}$.

The test statistic z_{10} is referenced against the standard normal distribution.

The authors additionally offer an extension of how optimization of w_1 and w_2 could be achieved, but suggest that the additional complication is unnecessary and the difference in results is negligible.

In common with the other statistics presented, z_{10} is computationally tractable but it may be less easy to interpret, particularly if $\psi_1 + \psi_2 \neq 1$.

4. Worked example

The objective of a Seasonal Affective Disorder (SAD) support group was to see if there is a difference in the quality of life for sufferers at two different times of the year. A binary response, 'Yes' or 'No' was required to the question whether they were satisfied with life. Membership of the group remains fairly stable, but there is some natural turnover of membership over time. Responses were obtained for $n_{12} = 15$ paired observations and a further $n_1 = 9$ and $n_2 = 6$ independent observations. The responses are given in Table 3.

Table 3. Responses to quality of life assessment.

Response Time 1	Response Time 2 Yes	No	Total
Yes	8	1	9
No	3	3	6
Total	11	4	15
	Response Yes	No	Total
Time 1	5	4	9
Time 2	6	0	6

The elements of the test statistics (rounded to 3 decimal places for display purposes), are calculated as: $\hat{p}_1 = 0.556$, $\hat{p}_2 = 1.000$, $\hat{p} = 0.733$, $\bar{p}_1 = 0.583$, $\bar{p}_2 = 0.810$, $\bar{p} = 0.689$, $r_1 = 0.431$, $r_2 = 0.673$, $w = 0.333$, $\psi_1 = 0.375$, $\psi_2 = 0.286$, $D = 0.002$. The resulting test statistics are given in Table 4.

Table 4. Calculated value of test statistics (with corresponding p-values)

-score	1.907	.311	1.000	.500	1.747	2.023	2.295	1.937	2.202	1.809
-value	.057	.190	.317	.617	.081	.043	.022	.053	.028	.070

At the 5% significance level, whether H_0 is rejected depends on the test performed. It is of note that the significant differences arise only with tests introduced in this paper, z_6, z_7 and z_9.

Although the statistical conclusions differ for this particular example, the numeric difference between many of the tests is small. To consider further the situations where differences between the test statistics might arise, simulations are performed.

5. Simulation design

For the independent observations, a total of n_1 and n_2 unpaired standard normal deviates are generated. For the n_{12} paired observations, additional unpaired standard normal deviates X_{ij} are generated where *i* = (1,2) and *j* = (1,2,...., n_{12}). These are converted to correlated normal bivariates Y_{ij} so that:

$$Y_{1j} = \sqrt{\frac{1+\rho}{2}}X_{1j} + \sqrt{\frac{1-\rho}{2}}X_{2j} \text{ and } Y_{2j} = \sqrt{\frac{1+\rho}{2}}X_{2j} - \sqrt{\frac{1-\rho}{2}}X_{1j}$$

where ρ = correlation between population one and population two.

The normal deviates for both the unpaired and correlated paired observations are transformed into binary outcomes using critical values $C_{\pi i}$ of the normal distribution. If $X_{ij} < C_{\pi i}, Y_{ij} = 1$, otherwise $Y_{ij} = 0$

10,000 iterations of each scenario in Table 5 are performed in a 4×4×5×5×5×7=14000 factorial design.

Table 5. Values of parameters simulated for all test statistics.

Parameter	Values
π_1	0.15, 0.30, 0.45, 0.50
π_2	0.15, 0.30, 0.45, 0.50
n_1	10, 30, 50, 100, 500
n_2	10, 30, 50, 100, 500
n_{12}	10, 30, 50, 100, 500
ρ	-0.75, -0.50, -0.25, 0.00, 0.25, 0.50, 0.75

A range of values for n_1, n_2 and n_{12} likely to be encountered in practical applications are considered which offers an extension to the work done by Choi and Stablein (1982). Simulations are conducted over the range π from 0.15 to 0.5 both under H_0 and H_1. The values of π have been restricted to π <= 0.5 due to the proposed statistics being palindromic invariant with respect to π and $1-\pi$. Varying ρ is considered as it is known that ρ has an

impact on paired samples tests. Negative ρ has been considered so as to provide a comprehensive overview and for theoretical interest, although $\rho < 0$ is less likely to occur in practical applications.

Two sided tests with $\alpha_{\text{nominal}} = 0.05$ is used in this study. For each combination of 10,000 iterations, the percentage of p-values below 0.05 is calculated to give the Type I error rate α. The Type I error rate under H_0, for each combination considered in the simulation design, should be between 0.025 and 0.075 to meet Bradley's liberal criteria and to be Type I error robust.

All simulations are performed in R.

6. Simulation Results

A comprehensive set of results with varying independent and paired sample sizes, correlation, and proportions was obtained as outlined in Section 5.

6.1. Type I error rates

Under H_0, 10,000 replicates were obtained for $4\times5\times5\times5\times7=3500$ scenarios. For assessment against Bradley's (1978) liberal criteria, Figure 1 shows the Type I error rates for all scenarios where $\pi_1 = \pi_2$ using $\alpha_{\text{nominal}} = 0.05$.

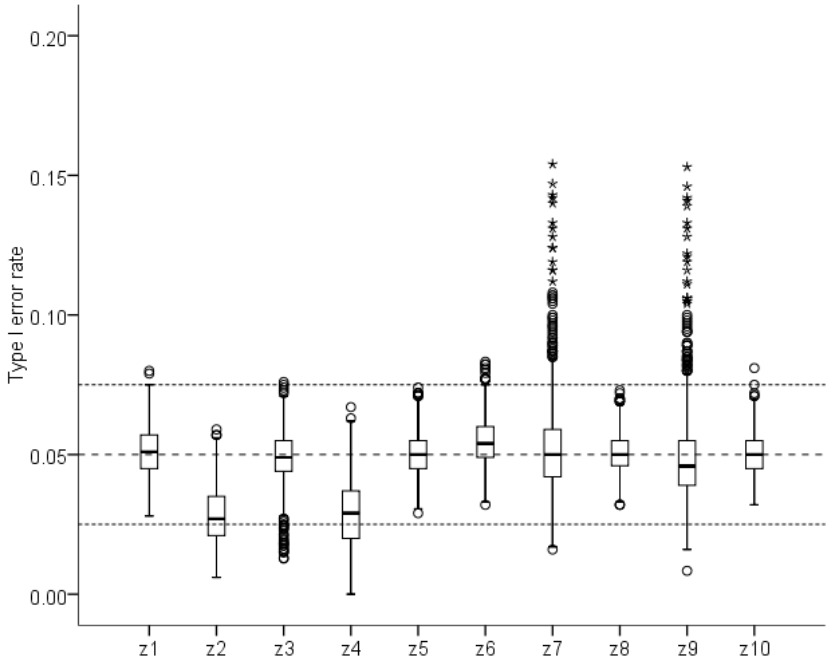


Figure 1: Type I error rates for each test statistic.

As may be anticipated, z_1 is Type I error robust because matched pairs are simply ignored. Similarly, z_3 performs as anticipated because the unpaired observations are ignored. Deviations from robustness for z_3 appear when n_{12} is small and ρ is large. Although deviations from stringent robustness are noted for z_3, this is not surprising since the cross product

ratio is likely to be small when the proportion of success is low and the sample size is low. Crucially, the deviations from Type I error robustness of z_3 are conservative and will result in less false-positives, as such the tests statistic may not be considered unacceptable.

The corrected statistics, z_2 and z_4, generally give Type I error rates below the nominal alpha, particularly with small sample sizes. Ury and Fleiss (1980) found that z_1 is Type I error robust even with small samples, however applying Yate's correction is not Type I error robust and gives Type I error rates less than the nominal alpha. It is therefore concluded that z_2 and z_4 do not provide a Type I error robust solution.

The statistics using the phi correlation coefficient, z_6 and z_8, are generally liberal robust. For z_6 there is some deviation from the nominal Type I error rate. The deviations occur when $\min\{n_1, n_2, n_{12}\}$ is small, $\max\{n_1, n_2, n_{12}\} - \min\{n_1, n_2, n_{12}\}$ is large and $\rho < 0$. In these scenarios the effect of this is that z_6 is not liberal robust and results in a high likelihood of false-positives. It is therefore concluded that z_6 does not universally provide a Type I error robust solution to the partially overlapping samples situation.

The statistics using the tetrachoric correlation coefficient, z_7 and z_9, have more variability in Type I errors than the statistics that use the phi correlation coefficient. The statistics using the tetrachoric correlation coefficient inflate the Type I error when $\rho > 0.25$ and n_{12} is large. When $\min\{n_1, n_2, n_{12}\}$ is small the test statistic is conservative. A test statistic that performs consistently would be favoured for practical use. It is therefore concluded that z_7 and z_9 do not provide a Type I error robust solution to the partially overlapping samples situation.

Three statistics making use of all of the available data, z_5, z_8 and z_{10}, demonstrate liberal robustness across all scenarios. Analysis of Type I error rates show near identical boxplots to Figure 1 when each of the parameters are considered separately. This means these statistics are Type I error robust across all combinations of sample sizes and correlation considered.

6.2. Power

The test statistics z_2, z_4, z_6, z_7 and z_9 are not Type I error robust. Therefore only z_1, z_3, z_5, z_8 and z_{10} are considered for their power properties (where H_1 is true). Table 6 summarises the power properties where $\pi_1 = 0.5$.

Table 6. Power averaged over all sample sizes.

π_1	π_2	ρ	z_1	z_3	z_5	z_8	z_{10}
0.5	0.45	> 0	0.095	0.173	0.208	0.221	0.221
		0		0.133	0.168	0.186	0.186
		< 0		0.112	0.150	0.166	0.166
0.5	0.3	> 0	0.509	0.653	0.807	0.856	0.855
		0		0.569	0.772	0.828	0.827
		< 0		0.508	0.746	0.801	0.801

		>0		0.874	0.975	0.989	0.989
0.5	0.15	0	0.843	0.834	0.970	0.985	0.986
		<0		0.795	0.966	0.980	0.982

For each of the test statistics, as the correlation increases from -0.75 through to 0.75 the power of the tests increase. Similarly, as sample sizes increase the power of the test increases.

Clearly, z_5 is more powerful than the other standard tests z_1 and z_3, but it is not as powerful as the alternative methods that make use of all the available data.

The power of z_8 and z_{10} are comparable. Separate comparisons of z_8 and z_{10} indicates that the two statistics are comparable across the factorial combinations in the simulation design. Either test statistic could reasonably be used for hypothesis testing in the partially overlapping samples case.

6.3. Confidence interval coverage

For z_8 and z_{10}, the coverage of the true difference of population proportions within 95% confidence intervals has been calculated as per the simulation design in Table 5 where $\pi_1 \neq \pi_2$. The results are summarised in Figure 2.

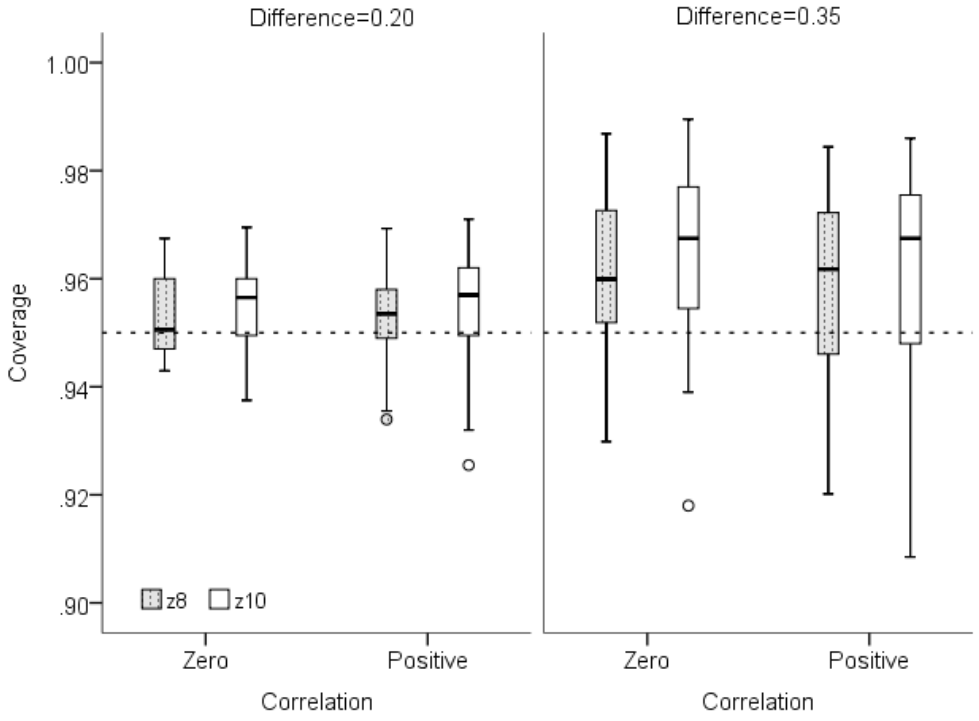


Figure 2: Percentage of iterations where the true difference is within the confidence interval.

Both z_8 and z_{10} demonstrate reasonable coverage of the true population difference $\pi_1 - \pi_2$. However, Figure 2 shows that z_8 more frequently performs closer to the desired 95%

success rate. Taking this result into account, when the objective is to form a confidence interval, z_8 is recommended as the test statistic of choice in the partially overlapping samples case.

7. Conclusion

Partially overlapping samples may occur by accident or design. Standard approaches for analysing the difference in proportions for a dichotomous variable with partially overlapping samples often discard some available data. If there is a large paired sample or a large unpaired sample, it may be reasonable in a practical environment to use the corresponding standard test. For small samples, the test statistics which discard data have inferior power properties to tests statistics that make use of all the available data. These standard approaches and other ad-hoc approaches identified in this paper are less than desirable.

Combining the paired and independent samples z-scores using Stouffer's method is a more powerful standard approach, but leads to complications in interpretation, and does not readily extend to the creation of confidence intervals for differences in proportions. The tests introduced in this paper, as well as the test outlined by Choi and Stablein (1982) are more powerful than the test statistics in 'standard' use.

The alternative tests introduced in this paper, z_6, z_7, z_8 and z_9, overcome the interpretation barrier, in addition confidence intervals can readily be formed.

Tests introduced using the phi correlation coefficient, z_6 and z_8, are more robust than the equivalent tests introduced using the tetrachoric correlation coefficient, z_7 and z_9.

The most powerful tests that are Type I error robust are z_8 and z_{10}. The empirical evidence suggests that z_8 is better suited for forming confidence intervals for the true population difference than z_{10}. Additionally, z_8 has relative simplicity in calculation, strong mathematical properties and provides ease of interpretation. In conclusion, z_8 is recommended as the best practical solution to the partially overlapping samples framework when comparing two proportions.

References

1. Berkson J. **In dispraise of the exact test,** Journal of Statistic Planning and Inference. 2, 1978, pp. 27–42.
2. Bhoj D. **Testing equality of means of correlated variates with missing observations on both responses**, Biometrika. 1978; 65:225-228.
3. Bradley JV. **Robustness?** British Journal of Mathematical and Statistical Psychology. 31(2), 1978, pp.144-152.
4. Choi SC, Stablein DM. Practical tests for comparing two proportions with incomplete data. Applied Statistics. 1982; 31:256-262.
5. Digby PG. **Approximating the tetrachoric correlation coefficient**. Biometrics. 1983; pp. 753-757.
6. Edwards JH, Edwards AWF, **Approximating the tetrachoric correlation coefficient**. Biometrics. 40(2), 1984, 563.
7. Ekbohm G., **On testing the equality of proportions in the paired case with incomplete data.** Psychometrika. 47(1), 1982, pp. 115-118.

8. Gardner MJ, Altman DG, **Confidence intervals rather than p values: estimation rather than hypothesis testing**. BMJ.292(6522), 1986, pp. 746-750.
9. Kim SC, Lee SJ, Lee WJ, Yum YN, Kim JH, Sohn S, Park JH, Jeongmi L, Johan Lim, Kwon SW., **Stouffer's test in a large scale simultaneous hypothesis testing**. Plos one. 8(5):e63290, 2013
10. Martinez-Camblor P, Corral N, De la Hera JM., **Hypothesis test for paired samples in the presence of missing data**. Journal of Applied Statistics. 40(1), 2012, pp. 76-87.
11. Penfield DA. **Choosing a two-sample location test**. Journal of Experimental Education. 62(4), 1994, 343-360.
12. R Core Team, **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org. 2014; version 3.1.2.
13. Samawi HM, Vogel R., **Tests of homogeneity for partially matched-pairs data**. Statistical Methodology. 8(3), 2011, pp. 304-313.
14. Tang ML, Tang NS., **Exact tests for comparing two paired proportions with incomplete data**. Biometrical Journal. 46(1), 2004, pp. 72-82.
15. Thomson PC., **A hybrid paired and unpaired analysis for the comparison of proportions**. Statistics in Medicine. 14, 1995, pp. 1463-1470.
16. Ury HK, Fleiss JL., **On approximate sample sizes for comparing two independent proportions with the use of Yates' correction**. Biometrics. 1980, pp. 347-351.
17. Whitlock MC., **Combining probability from independent tests: the weighted z-method is superior to Fisher's approach**. Journal of Evolutionary Biology. 18(5), 2005, pp. 1368-1373.

---

[1] Ben holds a first class honours degree in accounting and statistics and completed his masters in biometry with distinction at the University of Reading. He has previously worked in clinical research, business management information, training, and quality control. Ben is a full-time lecturer at the University of the West of England, Bristol with a teaching portfolio in statistical modelling, times series analysis, and in applications in the business sciences. He is undertaking doctoral work on the analysis of partially overlapping samples and is an active member of the consultancies arranged through the Applied Statistics group.

[2] Anselma has graduated from the University of the West of England, 2015, with first class honours in mathematics and statistics and was a recipient of the IMA (Institute of Mathematics and Its Applications) student prize. She optionally completed a full-year sandwich placement at the Office for National Statistics, where she worked collaboratively in methodological research most notably publishing papers in support of the Johnson Review of Consumer Price Indices

[3] Dr Deirdre Toher holds at PhD in applied statistics from Trinity College Dublin and is a senior lecturer in statistics at the University of the West of England. Deirdre is an active member of the Applied Statistics group working collaboratively in a wide range of diverse multi-disciplinary teams such as in disease profiling using high dimensional sensor data, providing methodological support to medical researchers, design and analysis of large scale national surveys, and the analysis of complex data such as experimental interventions in domestic water consumption. Deirdre is a Fellow of the Higher Education Academy, a fellow of the Royal Statistical Society (RSS), a Statistical Ambassador for the RSS, and trainer for the RSS Science Journalism Programme.

[4] Paul holds a PhD in pure and applied statistics. He is an associate professor in applied statistics and is the academic lead for the Applied Statistics Group at the University of the West of England. He has a wide range of undergraduate and postgraduate teaching across the University as well as providing PhD supervisory support in quantitative investigations. He has worked on over 100 projects in economics, psychology, health, and the bio- and medical sciences, as well as being a member of research ethics committees, data monitoring committees, and reviewer of grant applications.

JOURNAL OF APPLIED QUANTITATIVE METHODS

Quantitative Methods Inquires

COMPARISON OF CLASSIFICATION SUCCESS OF HUMAN DEVELOPMENT INDEX BY USING ORDERED LOGISTIC REGRESSION ANALYSIS AND ARTIFICIAL NEURAL NETWORK METHODS

Emre YAKUT
Osmaniye Korkut Ata University, Turkey, Faculty of Economics and Administrative Sciences, Department of Management Information Systems

E-mail: emreyakut@osmaniye.edu.tr

Murat GUNDUZ
UsakUniversity, Turkey , Faculty of Economics and Administrative Sciences, Department of Econometrics

E-mail: murat.gunduz@usak.edu.tr

Ayhan DEMİRCİ
Dr. Turkish Armed Forces, Turkey

E-mail: kho1993@hotmail.com.tr

Abstract

Economic development and growth are among the most important objectives for many countries. Not only economic development but also human development, which means enhancing and improving people's quality of life, plays an important role for reaching this objective. In this way, it's possible to take a more human oriented perspective by taking education, health and welfare dimensions of development into consideration by widening the perspective, which is focused narrowly on economic growth only. For this reason, human development index has become a widely preferred and recognized numerical indicator for comparison and classification of countries. Human Development Index (HDI) calculated by Human Development Report Office of United Nations Development Program (UNDP) measure people's level of welfare every year. The purpose of this research is to compare the classification success of Human Development Index by using ordered logistic regression and artificial neural network. The data of 81 countries, which has United Nations Development Program's Human Development Index, between the years of 2010-2012 were used in this study. Countries are classified for having very high, high and moderate levels of human development. The results of the ordered logistic regression model indicate that determinants including infant mortality rate, health expenses, number of internet users, import and export were observed as statistically significant. As a result of the analysis, Ordered Logistic Regression Analysis proved 88% success in classification while Elman's back propagation learning algorithm showed 92% success.

Keywords: *Human Development Index, Ordered Logistic Regression, Artificial Neural Network*

1. Introduction

Human development is the process of enhancing and improving people's life skills. This process aims to make a positive contribution to human and their living standards by equipping them with skills and capacity (UNDP 1990:1).

By its *Human Development Index* (HDI) developed in 1990, United Nations Development Program (UNDP) takes a more composite and human oriented perspective by taking education, health and welfare dimensions of development into consideration by widening the perspective which is focused narrowly on economic growth only (Lind, 1992:89).

Until 2010, GDP calculated per person based on purchasing power parity was taken into consideration for the economic dimension while life expectancy since birth was used for the health dimension and literacy and schooling were used for the education dimension. HDI calculates the arithmetic mean. Both economy and health dimension has one indicator while education dimension has two being literacy (2/3) and schooling (1/3) (Ivanova et al, 1999:159-160).

The human development approach comprises two central theses about people and development which are concerned with evaluating improvements in human lives as a distinctive development objective and what human beings can do to achieve such improvements particularly policy and political changes (Fukuda-Parr, 2003).

In 2010, index calculation was significantly changed. In this context, index calculation was based on arithmetic average instead of geometric average. With regards to education dimension, literacy among adults was excluded and the average of schooling rate and estimated schooling rate was considered (Morse 2014:249).

In the contemporary era, the concept of development has been in greater need of analysis and clarification and the word has come to be extraordinarily widely used in public discourse probably more so than ever before in its history (Payne & Phillips, 2010; Eren, et, al, 2014).

HDI is scored between 0 and 1. 1 shows the highest human development status. The human development report in 2014 stated 4 levels of human development as very high, high, moderate and low. Countries with HDI value lower than 0,550 was classified as low, 0,550–0,699 as moderate, 0,700–0,799 as high and higher than 0,800 as very high (UNDP 2014:156).

The purpose of this study is to compare the success of multiple classifications and to determine the effective factors by using logistic regression analysis and Elman ANN, multilayer ANN and LVQ network. This study is comprised of 3 parts. In the first part, logistic regression analysis is introduced while the second part focuses on Elman ANN and LVQ network. In the third part, application results are compared.

2. Ordered Logistic Regression Model

Logistic regression models are used for modelling the relation between dependent variables measured in different categories and independent variables of categorical or continuous measurement. Ordered logistic regression (OLOGREG) is used when dependent variables consists of at least three categories and measured by ordinal scale (Demirtas v.d.2009:869).

The main features of ordered logistic regression model are as follows (Chen and Hughes, 2004: 4):

- ✓ Outcome variable of categorical and ordinal measurement is a variable, which can be rearranged multiple times from an unobserved continuous latent variable, however it's not clear whether the space between the categories of this ordinal outcome variable is equal.
- ✓ Ordered logistic regression analysis, uses a correlation function to explain the effects of independent variables on ordered and categorical outcome variable. This model does not require normality and constant variance assumption.
- ✓ Since regression coefficient value is not dependent on the categories of categorical output variable, ordered logistic regression model assumes that the relation between explanatory variables and ordered categorical output variable is independent from categories.

Ordered logistic regression model is actually based on the existence of an continuous and unobserved random Y^* latent variable under a categorical dependent Y variable. The categories of this variable are estimated as sequential intervals on a continuous plane named as cut-off point or threshold value (McCullagh, 1980:109).

In $\theta_{s-1} < Y^* < \theta_s$, s=1,…, j interval and in the event of $\theta_0 = -\infty$ and $\theta_j = +\infty$ (Anderson, 1984:3), this latent Y^* variable is stated as in equity (1).

$$Y^* = \sum_{k=1}^{K} \beta_k x_k + \varepsilon \qquad (1)$$

Here θ refers to threshold value, x_k refers to independent variables vector, β_k refers to parameter vector and ε refers to error term.

The relation between observed Y variable and unobserved Y^* is shown in equity (2)(Liao, 1994:37-38):

$$y = \begin{cases} 1 & \text{if } y_i^* \le \theta_1 \\ 2 & \text{if } \theta_1 < y_i^* \le \theta_2 \\ 3 & \text{if } \theta_2 < y_i^* \le \theta_3 \\ \vdots & \\ j & \text{if } \theta_{j-1} < y_i^* \end{cases} \qquad (2)$$

θ's refer to threshold values that separate categories dependent variable. F being distribution function of error term, which is assumed to be distributed logistically, general probability of observed dependent variable's falling into k. Category is shown in equity (3) for given independent variables:

$$Prob(y = j \backslash x) = F\left[\theta_j - \sum_{k=1}^{K} \beta_k x_k\right] - F\left[\theta_{j-1} - \sum_{k=1}^{K} \beta_k x_k\right] \qquad (3)$$

There are many correlation functions, which are formations of cumulative probabilities in order to estimate ordered logistic model. These functions are shown in Table 1 (Elamir and Sadeq, 2010: 652):

Table 1. Correlation Functions and Typical Application

Function	Form	Application Area
Logit	$\log(\frac{x}{1-x})$	Categories are distributed evenly
Complementary Log-Log	$\log(-\log(1-x))$	High categories are more likely
Negative Log-Log	$-\log(-\log(x)$	Low categories are more likely
Probit	$F^{-1}(x)$	Variable is distributed normally
Couchit	$\tan(\pi(x-0.5))$	Variable has excessive values

L being logit distribution function in ordered logit model, the probability of observed variables' falling into categories of dependent variable is shown in equity (4) (Akkus v.d., 2010:323):

$$Prob(y=1) = L\left(-\sum_{k=1}^{K} x_k\beta_k\right)$$

$$Prob(y=2) = L\left(\theta_2 - \sum_{k=1}^{K} x_k\,\beta_k\right) - L\left(-\sum_{k=1}^{K} x_k\beta_k\right)$$

$$Prob(y=3) = L\left(\theta_3 - \sum_{k=1}^{K} x_k\,\beta_k\right) - L\left(\theta_2 - \sum_{k=1}^{K} x_k\beta_k\right)$$

$$Prob(y=j) = 1 - L\left(\theta_{j-1} - \sum_{k=1}^{K} x_k\beta_k\right) \qquad (4)$$

The most important assumption in ordered logistic regression model is the assumption of parallel curves. According to this assumption, regression parameters obtained in the model is the same in all categories of the dependent variable. In other words, the relation between independent variables and dependent variable does not change according to the categories of dependent variable, and parameter estimations do not change according to different threshold values. Thus, if there's a dependent variable of J category, "β_k " parameter is only one. On the other hand, there is θ_{j-1}cut-off point (threshold value) for J-1 logit comparisons (Akin and Senturk, 2012:185).

It's challenging to interpret parameters in ordered logistic regression. Methods of calculation of standardizes coefficients, calculation of estimated probabilities, calculation of factor change in estimated probabilities and percentage change in estimated probabilities are used for interpreting parameters. Odds ratio can also be used for interpreting parameters. In the event that all other variables are held constant, exp(β_k) is odd ratio for dummy variable. To standardize odds ratios, s_k: showing standard deviation, $\exp(\beta_k * s_k)$ is calculated provided that all other variables are held constant. For continuous variables; the percentage is found by $[\exp(\beta - 1) * 100]$(Ucdogruk vd., 2001).

3. Artificial Neural Networks

ANNs are cellular systems that can receive, store and use information. ANNs are parallel systems, which are formed by connecting many connecting elements with links of variable weights. Multi-layer artificial neural network is the most popular one among many artificial neural networks (Lippman 1987: 15). ANN is a system based on simple neural networks,

which can receive interconnected information as input, process them and submits to other units, and which can even use the outputs as inputs again (Pissarenko, 2001-2002: 35). ANN simulates the operation of a simple biological neural system. ANNs provide solutions to problems, which normally requires a person's natural ability to think and observe.

Artificial neural networks are computer systems which are developed to perform some characteristics of a human brain automatically without getting any help such as getting new information through learning, creating new information and discovering (Oztemel, 2003: 29). Artificial neural networks are used to achieve one or more processes including learning through using available data, associating, classification, generalization and optimization (Sen, 2004: 13).

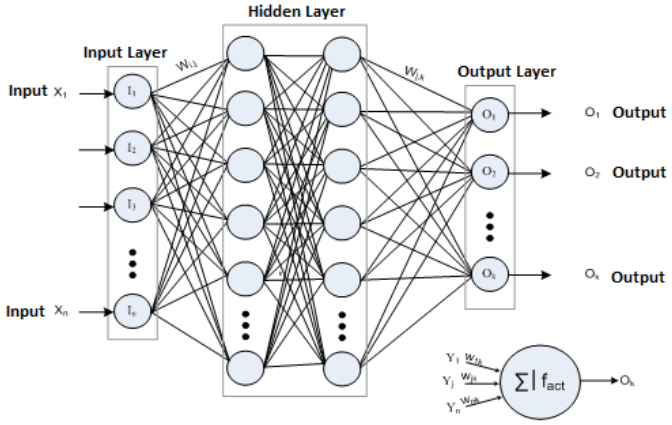


Figure 1. General Structure of an Artificial Neural Network

Use of artificial neural networks in such problems whose algorithmic solution could not be found has increased due to the fact that artificial neural networks can find solutions to new occurrences by way of examining former instances and learning the relationship between inputs and outputs of the said occurrence, regardless of whether the relationship is linear or not, from the current instances in hand. The biggest problem in artificial neural networks is that there is a need for such artificial neural networks that contain either very large neurons or multi-layered and a great amount of neurons in order to solve complicated problems (Kohonen,1987: 1-79). An artificial neural network is an intensively parallel-distributed processor which is comprised of simple processing units, has a natural tendency to collecting experiential information, and enabling them to be used (Haykin, 1999:2).

In a general artificial neural network system, neurons gather on the same direction to form layers (Yildiz, 2001: 51-67). There is parallel flow of information from the input layer to the exit in an architectural structure. Such flow is possible with parallel placed cells.

3.1. Elman Network

Elman network is an ANN type, which includes the whole multi layer ANN as well as interlayer outputs as a parallel input (Sen, 2004: 144).

Elman network delivers not only input values of a given time but also previous activity values of interlayers as an input into the network. Once the inputs are determined, the network becomes a multilayer feed-forward receptor. These inputs are used to determine network's forward outputs (Elman, 1990: 182).

Although Elman network is quite similar to Jordan network, there are significant differences. First of all, they obtain feedback activation values from the interlayer instead of output layer. Secondly, content elements are not self-connected. This network structure is shown below (Kucukonder, 2011: 78)

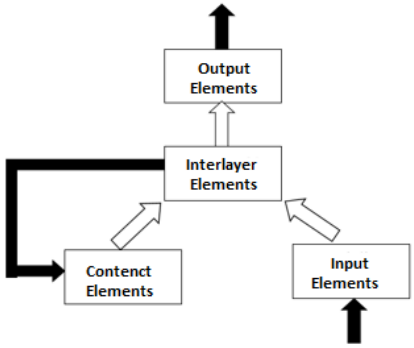


Figure2. Structure of Elman Network

In Elman network, learning is achieved in two steps according to generalized delta learning rule. Firstly, weight of net input value received by processor elements in interlayer is multiplied with and added to the weight of element values of input layer. Secondly, these connection values from content elements is multiplied with and added to previous activation values in interlayers. Elman network is shown in detail in Figure 3 below (Oztemel, 2003: 167).

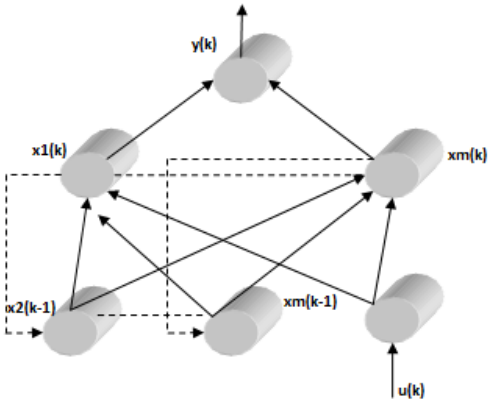


Figure 3. Detailed demonstration of Elman network

3.2. LVQ Model (Learning Vector Quantization)

Developed by Kohenon (1984), LVQ Model is a network structure using reinforcement-learning model. LVQ networks are mostly used for solving classification problems.

Learning means finding which set of vectors (reference vector) should represent input vector. LVQ network's duty is to determine the set of vectors, i.e. the ones that may be a member of input vectors, by means of learning. It learns by Kohonen's learning principle. Only one output is valued as 1 while the others are given 0 value, and if the output is valued as 1, it means that the input belongs to the class represented by the output. Since LVQ network is used as a method of statistical classification and distinction, its purpose is classifying input data (Kohonen, 2001: 245).

LVQ network consists of entrance, exit and Kohenon levels and all neurons of the entrance layer are in connection with all neurons in interlayers. The main purpose of this network is mapping a vector of n dimension in sets of vectors (Kucukonder, 2011: 70).

During training, input vector is breakdown based on the nearest neighbour rule. The model looks for the shortest distance between the input vector and reference vectors and it's assumed that the input vector belongs to the nearest vector group. The weights of the network are changed in order to determine reference vectors for an accurate breakdown of input vectors. For this purpose, reinforcement learning strategy is used. For determining the output value, "winner-takes-all" strategy is used. In training process, not the output value but whether it's correct or not is stated for each iteration. Only the values of the winner vector, which is the closest to the input vector, are changed (the weights of the network for this vector). The architectural structure of LVQ network demonstrating these details is shown in Figure 4 (Oztemel, 2003: 116).

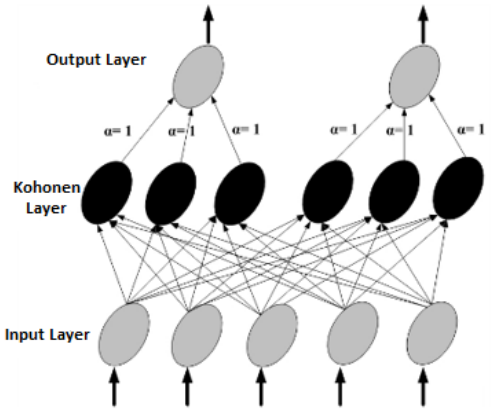


Figure 4. Structure of LVQ Network

LVQ network is composed of three layers (Adiyaman, 2007: 38):

- **Input layer:** There's no data processing in this layer, and incoming information forms the vector. Each process element in this layer is connected to each process element in Kohonen layer. Learning is achieved by changing weights of input layer and Kohonen layer.
- **Kohonen layer:** In this layer, the weight vector, which is the closest to the input set, is determined. Each element in this layer shows a reference vector and composed of weight values of links connecting input values to process elements in Kohonen layer. The number of elements in reference vector is equal to the number of elements in input layer.
- **Output layer:** In this layer, the class, which the input belongs to, is determined. Process elements in Kohonen layer is linked to only one process element in output layer. The weights between Kohonen layer and output layer is (a) and equal to 1.

4. Scope and Method of Study

This part gives information about the purpose, scope, limitations, universe and analysis method of the study.

4.1. Purpose and Significance of Study

Economic development and growth are among the most important objectives for all countries. For this reason, human development index has become a widely preferred and recognized numerical indicator for comparison and classification of countries. The purpose of

this research is to compare the classification success of Human Development Index and determine the by using ordered logistic regression as well as Elman ANN, multi layer ANN and LVQ network as artificial neural networks.

4.2. Scope, Limitations and Constraints of Study

In this study, data of highly developed, developed, moderately developed and less developed countries produced by Human Development Report Office of United Nations Development Program (UNDP), which is published annually. This data covers a period of three years and is obtained from 2010-2012 Human Development Index published at United Nations Development web site.Since it was difficult to find data about less developed countries, they are excluded from the analysis. The study classifies 81 countries based on data of three years. With the addition of one more country in 2012, the universe of the study is total of 244 countries.

Income, education and life expectancy are indispensable components in calculation of human development index. Since human development level would be subject to multiple classification process, 11 independent variables were added to these three variables, thus 14 independent variables were studied in order to enable a more detailed examination of methods. The following variables were used in the analysis respectively:

X1: BÖO "infant mortality rate",
X2: GSMH "gross national product",
X3: LKO "high school enrolment",
X4: BY "growth",
X5: DYY "direct foreign investment",
X6: ET "energy consumption",
X7: EÜ "energy production",
X8: E "inflation",
X9: IH "export",
X10: IKS "number of internet users",
X11:ISZ "unemployment",
X12: ITH "import",
X13: MTAS "number of mobile phone subscribers",
X14: SH "health expenses"

Answer variables for Human Development are coded as follows for Ordered Logistic Regression Analysis:

0: Moderately Developed,
1: Developed,
2: Highly Developed.

4.3. Data Analysis Method

Countries are classified for their Human Development Level by using Stata 11.2 package, Ordered Logistic Regression and Matlab 2012 software with Elman, Multi Layer Neural Networks and LVQ Network methods. In artificial neural network method, if a country is highly developed in human development, it is valued as 1 while the others' output values are stated as 0.

4.4. Ordered Logistic Regression Analysis

Stata 11.2 statistical analysis program is used to classify Human Development Index of countries.

Table 2. Relevance Values of the Established Model

Log-Lik Intercept Only:	**-256.865**	**Log-Lik Full Model:**	**-98.392**
D(228):	196.784	LR(12):	316.947
		Prob>LR:	0.000
McFadden's R2:	0.617	McFadden's Adj R2:	0.555
ML (Cox-Snell) R2:	0.727	Cragg-Uhler(Nagelkerke) R2:	0.828
McKelvey&Zavoina's R2:	0.886		
Variance of y*:	28.850	Variance of error:	3.290
Count R2:	0.881	Adj Count R2:	0.779
AIC:	0.938	AIC*n:	228.784
BIC:	-1.056.570	BIC':	-250.981
BIC used by Stata:	273.745	AIC used by Stata:	224.784

In order to test relevance of the model used in study, fitstat command was run in Stata 11.2 software. Table 2 shows the conformity values of the model. Akaika information criteria (AIC) was found to be 228.784 while Bayes information criteria (BIC) was found as -250.981.Prob>LR: 0.000 model with all independent variables was found to be statistically significant. McKelvey and Zavonia's R^2 value approximates to R^2 value, which is obtained by estimating linear regression model (Long andFreese, 2001: 148). McKelvey and Zavonia's R^2 value of the model was measured as 88.6%.

Table 3. ParallelCurves Assumption Test

Varia-ble	**chi2**	**p>chi2**	**df**
All	14.12	0.118	14

Table 3 shows the test results of parallel curves assumption of logistic regression model. Zero hypothesis (H_0) and alternative hypothesis (H_a) used for testing parallel curves assumption test are as follows:

"H_0: Relevant regression coefficients are the same in all categories of dependent variable

H_a: Relevant regression coefficients are different at all levels of dependent variable"

Brant, detail command was run in Stata software. In the model, as P>χ^2: 0,118, H_0 hypothesis was accepted,thus it shows that estimated regression coefficients are the same in each category of the dependent variable and parallel curves assumption is realized with P>0,05.

Table 4. Established Model's Goodness of Fit

Goodness of Fit Test of Ordered Logit Model	
Pearson χ^2 test statistic	10.24
Liberty Level	14
P>χ^2	0.332

"H_0: Model-data fit is sufficient in terms of parameter's decisiveness

H_a: Model-data fit is not sufficient in terms of parameter's decisiveness"

Omodel logit command was run in Stata software. Likelihood ratio test known as χ^2goodness of fit test, evaluates ordered logistic regression model as a whole. In the model, as P>χ^2: 0,332, H_0 hypothesis was accepted, thus it was found that ordered logistic regression model is a sufficient model for classification of human development index. It can be stated that model's goodness of fit is quite good and parameters are good at classification decisiveness.

Table 5. Results of Ordered Logistic Regression Analysis of Variables Effecting Human Development Level

						Number of obs		**244**
						LRchi2(12)		**316,95**
						Prob> chi2		**0.0000**
Log likelihood		-110,706				**Pseudo R2**		**0,617**
HDII	**Coef.**	**Std. Err.**	**Wald**	**z**	**P>z**	**Odds Ratio**	**[95% Conf. Interval]**	
BÖO	-0,132	0,035	13,985	-3,740	0,000	0,876	0,817	0,939
GSMH	0,000	0,000	1,840	1,360	0,174	1,000	1,000	1,000
LKO	0,017	0,013	1,749	1,320	0,186	1,017	0,992	1,043
BY	-0,076	0,063	1,482	-1,220	0,223	0,926	0,819	1,048
DYY	0,000	0,000	0,175	-0,420	0,674	1,000	1,000	1,000
ET	0,000	0,000	0,603	-0,780	0,438	1,000	1,000	1,000
EÜ	0,000	0,000	0,011	-0,110	0,916	1,000	1,000	1,000
E	0,051	0,037	1,837	1,360	0,175	1,052	0,978	1,131
IH	0,058	0,023	6,475	2,540	0,011	1,059	1,013	1,108
IKS	0,104	0,020	26,893	5,190	0,000	1,109	1,067	1,154
ISZ	-0,005	0,044	0,012	-0,110	0,914	0,995	0,914	1,084
ITH	-0,062	0,024	6,788	-2,610	0,009	0,940	0,897	0,985
MTAS	-0,006	0,009	0,506	-0,710	0,477	0,994	0,977	1,011
SH	0,237	0,110	4,616	2,150	0,032	1,267	1,021	1,573
/cut1	1,563	1,539					-1,453	4,580
/cut2	6,487	1,614					3,323	9,651

Results of the ordered logistic regression analysis of dependent and independent variables using mlogit command in Stata 11.2 package program are shown in Table 5. As seen in Table 5, the number of observations is 244 in the model and χ^2value is statistically significant (p <0.01). Log likelihood value of the model was found to be -110.71. The first column of Table 5 shows β coefficients of ordered logistic regression analysis. BÖO "infant mortality rate", IKS "number of internet users" and ITH "import" with 0.01 significance level and IH "export" and SH "health expenses" with 0.05 significance level were observed as statistically significant. BÖO and ITH variables above the statistically dependent variable is marked negative while estimated value of İKS, İH and SH variables are marked positive. In addition, marginal effects will be calculated for the change in probabilities of dependent variable pursuant to change in β coefficients.

In odds ratio, one unit increase in BÖO variable decreases odds of high level human development rate by 12.4% while one unit increase in ITH variable decreases it by 6% provided that all other independent variables are held constant against moderate and low level of human development rate. One unit increase in IH variable increases odds of high level human development rate by 5.9% while one unit increase in IKS variable increases it by 10.9% and one unit increase in SH variable increases it by 26.7% against moderate and low level of human development rate. Thus, the most important variable that has a positive effect on human development level is SH "health expenses" variable, the second one is IKS "number of internet users" variable and the third one is IH "export" variable while the most important variable with negative effect is BÖO "infant mortality rate".

For the number of categorical variables, M-1=2 cut-off value is obtained.

Ordered logistic regression model which is established for dependent variable with 3 categories used for calculation of probabilities is shown in Table 5.

$Z=\sum_{k=1}^{K}\beta_k X_k = -0{,}132\text{BÖO}+1{,}37*10^{-12}\text{GSMH}+0{,}017\text{LKO} - 0{,}076\text{BY}-1*10^{-11}\text{DYY} - 5{,}38*10^{-6}\text{ET}$
$- 3{,}42*10^{-7}\text{EÜ}+0{,}051\text{E}+0{,}058\text{IH}+0{,}104\text{IKS}-0{,}005\text{ISZ}-0{,}062\text{ITH}-0{,}006\text{MTAS}+0{,}237\text{SH}$

Probabilities will be calculated by ordered logistic regression model.Z values obtained from the above equation will be written in the formulas below and the class the countries belong to as per their human development levels will be determined by calculating probability values of moderate, developed and highly developed index of countries.

$P(Y=0) = 1 - \frac{\exp(Z_i - cut_1)}{1+\exp(Z_i - cut_1)}$ = probability values of moderate development index,

$P(Y=1) = \frac{\exp(Z_i - cut_2)}{1+\exp(Z_i - cut_2)} - \frac{exp(Z_i - cut_1)}{1+\exp(Z_i - cut_1)}$ = values of development index,

$P(Y=2) = \frac{\exp(Z_i - cut_2)}{1+\exp(Z_i - cut_2)}$ = probability values of high development index,

Probability Values of Human Development Index of Turkey in 2010:

$Z_1(x_i) = -0{,}132*\text{BÖO}(16)+1{,}37*10^{-12}*\text{GSMH}(7{,}31\text{E}+11)+0{,}017*\text{LKO}(56)-0{,}076*\text{BY}(9{,}29)+$
$-1*10^{11}*\text{DYY}(9{,}04\text{E}+09)-5{,}38*10^{-6}*\text{ET}(105133{,}1)-3{,}42*10^{-7}$
$*\text{EÜ}(32225)+0{,}051*\text{E}(5{,}7)+0{,}058\text{I}*\text{H}(21)+0{,}104\text{I}*\text{KS}(39{,}8)$
$-0{,}005\text{I}*\text{SZ}(11{,}9)-0{,}062*\text{ITH}(27)-0{,}006*\text{MTAS}(85)+0{,}237*\text{SH}(6{,}7)=3{,}441$

$Z_i - cut_1 = 3{,}441 - 1{,}563 = 1{,}878$

$Z_i - cut_2 = 3{,}441 - 6{,}487 = -3{,}046$

P(Y=0)= 0,133,probability of being a moderately developed country,

P(Y=1)=0,822, probability of being a developed country,

P(Y=2)=0,045, probability of being a highly developed country,

It's determined that Turkey, which was included in the category of developed countries for human development in 2010, was classified correctly in the category of developed countries for human development with the highest probability result of 0.822 by ordered logistic regression analysis.

Probability Values of Human Development Index of Turkey in 2011:

$Z_1(x_i) = -0{,}132$*BÖO(15)+1,37*10^{-12} *GSMH(7,75E+11)+0,017*LKO(61)–0,076*BY(8,8)-1*10^{11}*DYY(1,6E+10)–5,38*10^{-6} *ET(112458,7)–3,42*10^{-7} *EÜ(32064)+0,051*E(8,6)+0,058*IH(24)+0,104*IKS(43,1)-0,005*ISZ(9,8)-0,062*ITH(33) -0,006*MTAS(89)+0,237*SH(6,7)=3,916

$Z_i - cut_1 = 3{,}916 - 1{,}563 = 2{,}353$

$Z_i - cut_2 = 3{,}916 - 6{,}487 = -2{,}571$

P(Y=0)= 0,087, probability of being a moderately developed country,

P(Y=1)=0,842, probability of being a developed country,

P(Y=2)=0,071, probability of being a highly developed country,

It's determined that Turkey, which was included in the category of developed countries for human development in 2011, was classified correctly in the category of developed countries for human development with the result of 0.842 by ordered logistic regression analysis.

Probability Values of Human Development Index of Turkey in 2012:

$Z_1(x_i) = -0{,}132$*BÖO(14)+1,37*10^{-12} *GSMH(7,89E+11)+0,017*LKO(62)–0,076*BY(2,2) -1*10^{11} *DYY(1,26E+10)–5,38*10^{-6} *ET(115701,2)–3,42*10^{-7} *EÜ(31117)+0,051*E(6,8)+0,058*IH(26)+0,104*IKS(45,1) -0,005*ISZ(9,3)-0,062*ITH(32)-0,006*MTAS(91)+0,237*SH(6,7)=4,891

$Z_i - cut_1 = 4{,}891 - 1{,}563 = 3{,}328$

$Z_i - cut_2 = 4{,}891 - 6{,}487 = -1{,}596$

P(Y=0)= 0,035, probability of being a moderately developed country,

P(Y=1)=0,797, probability of being a developed country,

P(Y=2)=0,168, probability of being a highly developed country,

It's determined that Turkey, which was included in the category of developed countries for human development in 2012, was classified correctly in the category of developed countries for human development with the result of 0.797 by ordered logistic regression analysis.

In this way, Turkey's human development rate was successfully classified.

Table 6. Marginal Effects on Probability

Variable	Probability of Moderate Development	Probability of Development	Probability of High Development
BÖO	0,0017023	0,0285864	-0,0302887
GSMH	-1,76E-14	-2,96E-13	3,14E-13
LKO	-0,0002201	-0,0036953	0,0039154
BY	0,0009831	0,0165083	-0,0174914
DYY	1,29E-13	2,17E-12	-2,30E-12
ET	6,92E-08	1,16E-06	-1,23E-06
EÜ	4,40E-09	7,39E-08	-7,83E-08
E	-0,0006499	-0,0109138	0,0115637
IH	-0,0007436	-0,0124877	0,0132313
IKS	-0,0013342	-0,0224047	0,0237389
ISZ	0,0000608	0,0010203	-0,0010811
ITH	0,0007969	0,0133821	-0,014179
MTAS	0,0000785	0,0013188	-0,0013973
SH	-0,0030488	-0,0511975	0,0542463

Marginal effects show the effect of one unit change of average on the probability of dependent variable categories. In order to calculate marginal effects, mfx command was run in Stata 11.2 software. When statistically significant variables are taken into consideration in the equation of ordered logistic regression and provided that all other variables are held constant, one unit increase in BÖO variable decreases the probability of being a highly developed country with regards to human development by 3% while it increases the probability of being a developed country with regards to human development by 2.9% and being a moderately developed country with regards to human development by 0.17%.

One unit increase in IHT variable decreases the probability of being a highly developed country with regards to human development by 1% while it increases the probability of being a developed country with regards to human development by 1% and being a moderately developed country with regards to human development by 0.08%.

One unit increase in IH variable increases the probability of being a highly developed country with regards to human development by 1% while it decreases the probability of being a developed country with regards to human development by 1% and being a moderately developed country with regards to human development by 0.07%.

One unit increase in IKS variable increases the probability of being a highly developed country with regards to human development by 2% while it decreases the probability of being a developed country with regards to human development by 2% and being a moderately developed country with regards to human development by 0.1%.

One unit increase in SH variable increases the probability of being a highly developed country with regards to human development by 5% while it decreases the probability of being a developed country with regards to human development by 5% and being a moderately developed country with regards to human development by 0.3%.

Table 7. Classification Success of Ordered Logistic Regression Analysis

Ordered Logistic Regression Analysis		Estimated Group				Accuracy Rate
		Moderate	Developed	Highly Developed	Total	
Observed Group	Moderate	**44**	9	0	53	83
	Developed	7	**66**	5	78	84,6
	Highly Developed	2	6	**105**	113	92,9
	Total	53	81	110	**244**	88,1

Table 7 shows the results of 244 countries' human development level classification results by using ordered logistic regression analysis. Ordered logistic regression analysis of development classification variable showed that 44 of 53 moderately developed countries, 66 of 78 developed countries and 105 of 113 highly developed countries were estimated accurately hence achieving 83% success rate for moderately developed countries, 84.6% for developed countries and 92.9% for highly developed countries. Total classification success for all countries is 88.1%.

4.5. Artificial Neural Networks Analysis

Matlab R2012a computer software was used to establish artificial neural network models. In order to determine the appropriate artificial neural network method, trial-

and-error method is commonly used and many tests are performed. In this context, different combinations of parameters including number of hidden layers, number of nodes in hidden layers, momentum term, activation function, number of cycles were tested on both the training set and the test set to find the best performing network. Elman artificial neural network, multi layer artificial neural network and LVQ network were used as artificial neural networks in this study.

Table 8: Parameters of Elman ANN, Multi Layer ANN and LVQ Network

Network type	Elman ANN	Multi Layer ANN	LVQ ANN
Learning Algorithm	Levenberg-Marquardt Optimization (Supervised Learning)	Levenberg-Marquardt Optimization (Supervised Learning)	Learnk (Reinforcement Learning)
Learning Rule	Gradient descent rule	Gradient descent rule	Kohonen rule
Number of Nodes in Entrance Layer	14	14	14
Number of Hidden Layers	1	1	1
Number of Nodes in Hidden Layer	9	9	10
Number of Nodes in Exit Layer	3	3	3
Learning Ratio	0,01	0,01	0,01
Number of Cycles	21	28	110
Learning Time (seconds)	1	2	192
Transfer Function for Hidden Layers	Tansig	Tansig	Learnlv1 (LVQ1 weight learning function)
Transfer Function for Output Layer	Purelin	Purelin	Purelin
Training Function of BackPropagation Network	Trainlm	Trainlm	No back propagation

The network structure of the models with the best number of layers and nodes for ANN used for classification estimation is given in Table 8. Elman ANN, multi-layer ANN and LVQ models were used to classify human development levels of countries. All three models had 14 nodes in the input layer and this gives the results of 14 independent variables after normalization used in classifying human development levels. Multi-layer ANN and Elman ANN had 1 hidden layer with 9 nodes in each layer while the LVQ had 10 nodes in hidden layer. In order perform three-category classification in Elman ANN, multi-layer ANN and LVQ network, there were 3 nodes in the output layer. Figure 5 shows the architectural structure of ANN Models developed by Matlab 2012 program.

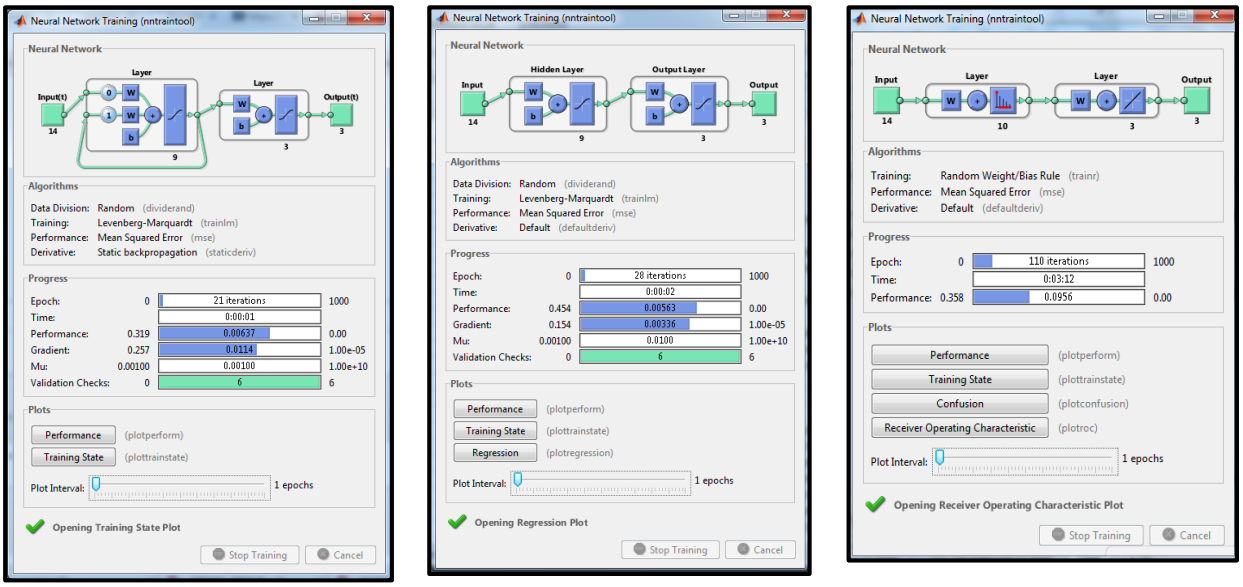


Figure 5. Models Developed by Matlab Program to Classify Human Development Level of Countries

After the training processes were completed for the established ANN models, numerous tests were performed. For Elman ANN and multi-layer ANN models, the following learning algorithm gave the best classification results for human development rate of countries: Levenberg-Marquardt Optimization learning algorithm, "tansig" as the sigmoid transfer function in hidden layers, "purelin" as the transfer function in output layer and "trainline "functions for training of back propagation network. Final ANN models, which provided the best classification results for human development levels are given in Figure 1. Reinforcement learning algorithm was used as the learning algorithm for LVQ networks, LVQ1 weight learning function was used in hidden layers and "purelin" functions were used for output layer. The most appropriate neural network models were selected as 14-1-3 for Elman neural ANN, Multi layer ANN and LVQ network.

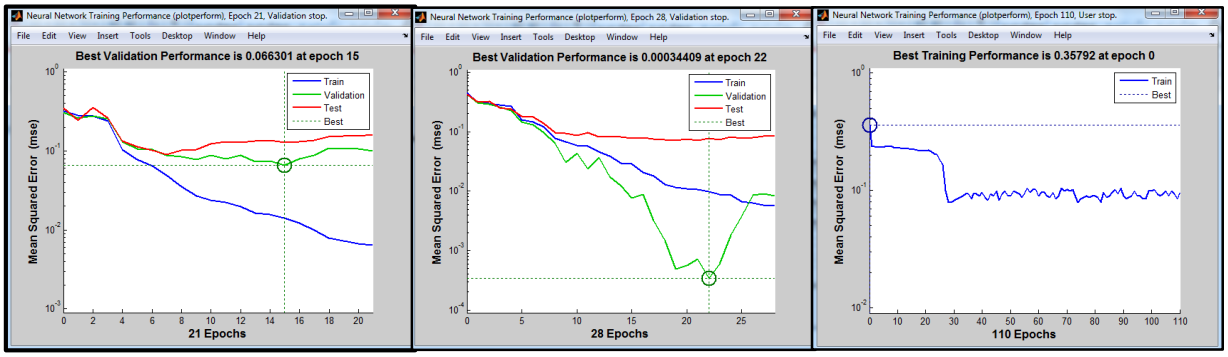


Figure 6. Cycling Performance of ANN Models

Figure 6 shows the cycling performance of ANN models. At the beginning, 1000 iterations were given for the training of the established artificial neural networks and Elman

ANN spent 1 second for 21 iterations, multi layer ANN spent 2 seconds for 28 iterations and LVQ network spent 192 seconds for 110 iterations and completed the learning process.

Table 9. Classification Success of Elman ANN Analysis

Elman ANN		Estimated Group				Accuracy Percentage
		Moderate	Developed	Highly Developed	Total	
Observed Group	Moderate	**48**	5	0	53	90,6
	Developed	0	**74**	4	78	94,9
	Highly Developed	2	3	**108**	113	95,6
	Total	50	82	112	**244**	94,3

Table 9 shows the results of 244 countries' human development level classification results by using Elman ANN. Elman ANN analysis of development classification variable showed that 48 of 53 moderately developed countries, 74 of 78 developed countries and 108 of 113 highly developed countries were estimated accurately hence achieving 90.6% success rate for moderately developed countries, 94.9% for developed countries and 95.6% for highly developed countries. Total classification success for all countries is 94.3%.

Table 10. Classification Success of Multi Layer ANN Analysis

Multi Layer ANN		Estimated Group				Accuracy Percentage
		Moderate	Developed	Highly Developed	Total	
Observed Group	Moderate	**53**	0	0	53	100
	Developed	2	**75**	1	78	96,2
	Highly Developed	2	2	**109**	113	96,5
	Total	57	77	110	**244**	97,1

Table 10 shows the results of 244 countries' human development level classification results by using Multi Layer ANN. Multi Layer ANN analysis of development classification variable showed that 53 of 53 moderately developed countries (all of them), 75 of 78 developed countries and 109 of 113 highly developed countries were estimated accurately hence achieving 100% success rate for moderately developed countries, 96.2% for developed countries and 96.5% for highly developed countries. Total classification success for all countries is 97.1%.

Table 11. Classification Success of LVQ Network Analysis

LVQ Network	Estimated Group				Accuracy Percentage
	Moderate	Developed	Highly Developed	Total	

Observed Group	**Moderate**	**46**	7	0	53	86,8
	Developed	13	**62**	3	78	79,5
	Highly Developed	4	8	**101**	113	89,4
	Total	63	77	104	**244**	85,7

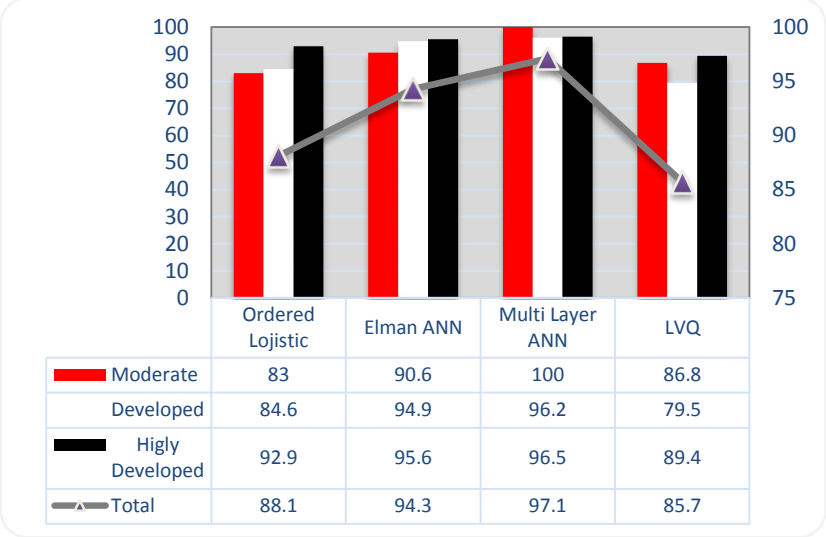
Table 11 shows the results of 244 countries' human development level classification results by using LVQ network analysis. LVQ network analysis of development classification variable showed that 46 of 53 moderately developed countries, 62 of 78 developed countries and 101 of 113 highly developed countries were estimated accurately hence achieving 86.8% success rate for moderately developed countries, 79.5% for developed countries and 89.4% for highly developed countries. Total classification success for all countries is 85.7%.

5. Conclusion

In this study, human development level of countries were classified by using ordered logistic regression, Elman ANN, multi layer ANN and LVQ network. Multiple classification methods were used to determine the relation between moderately developed, developed and highly developed countries as the dependent variable with ordered variable of 3 categories and 14 independent variables. The data of 81 countries, which has United Nations Development Program's Human Development Index, between the years of 2010-2012 were used in this study.

In ordered logistic regression analysis among 14 independent variables, it was observed that BÖO "infant mortality rate" and ITH "import" had a statistically significant negative effect while IKS "number of internet users", IH "export" and SH "health expenses" a statistically significant positive effect. Thus, the most important variable that has a positive effect on human development level was SH "health expenses" variable and the second one was IKS "number of internet users" varible while the most important variable with negative effect was BÖO "infant mortality rate". Marginal effects for independent variables were calculated. It was observed that one unit increase in the most important variable SH increases the odds of being a highly developed country with regards to human development by 5% while the second most important variable IH increases this odd by 1%. The classification of Turkey, which was classified as a developed country with regards to human development in 2010, 2011 and 2012, was successfully estimated by using ordered logistic regression analysis.

For ANN, which is another classification technique for human development level; Elman ANN, multi layer ANN and LVQ network were used. For classification estimation, there was 1 hidden layer for Elman, Multi layer ANN and LVQ. For Elman ANN and Multi layer ANN, 9 nodes were used in hidden layers, and 10 nodes were used in LVQ network.



	Ordered Lojistic	Elman ANN	Multi Layer ANN	LVQ
Moderate	83	90.6	100	86.8
Developed	84.6	94.9	96.2	79.5
Higly Developed	92.9	95.6	96.5	89.4
Total	88.1	94.3	97.1	85.7

Chart 1. Performance Criterion for All Four Methods of Analysis

Chart 1 shows percentage values of all four methods of analysis for moderate, development and high development levels and the overall classification accuracy. Multi Layer ANN analysis had the best performance among all. Its classification success was 100% for moderately developed countries, 96.2% for developed countries and 96.5% for highly developed countries. As a result of comparison of analyses, it's seen that Muli Layer ANN provides results with a higher accuracy percentage compared to Elman ANN while ordered logistic regression analysis provides results with a higher accuracy compared to LVQ network. In all four methods of analysis, it was proved that Multi Layer ANN had a better performance compared to the other three methods with regards to total classification results of moderately, developed and highly developed countries.

References

1. Adıyaman, F., **Talep Tahmininde Yapay Sinir Ağlarının Kullanılması, (Yayımlanmamış Yüksek Lisans Tezi)**, İstanbul Teknik Üniversitesi Fen Bilimler Enstitüsü, İstanbul, 2007.
2. Akın, H.B., ve Şentürk, E., "**Bireylerin Mutluluk Düzeylerinin Ordinal Lojistik Regresyon Analizi İle İncelenmesi"**, Öneri, C.10, S.37, 2012, ss.183-193.
3 .Akkuş, Ö, Türkan, S., Tatlıdil, H., **"Sıralı Bağımlı Değişken Modeli Ve Diskriminant Analizi'nin Ticari Bankaların Mali Performanslarına Göre Sınıflandırılmasında Kullanımı"**, Suleyman Demirel University The Journal of Faculty of Economics and Administrative Sciences, Vol.15, No.2, 2010, pp:319-332.
4. Chen, C-K., and Hughes, J. Jr., **"Using Ordinal Regression Model to Analyze Student Satisfaction Questionnaires",** IR Applications, Association for Institutional Research, Volume 1, 2004, pp:1-13.
5. Demirtas E. A., Anagun A.S., Koksal G., **"Determination Of Optimal Product Styles By Ordinal Logistic Regression Versus Conjoint Analysis For Kitchen Faucets"**. International Journal of Industrial Ergonomics 39, 2009, pp 866–875.
6. Elamir, E., ve Sadeq, H., **"Ordınal Regression to Analyze Employees Attitudes Towards The Application of Total Quality Management"**, Journal of Applied Quantitative, Vol.5, No.4, 2010, pp. 647-658

7. Elman, J.L., **"Finding structure in time"**, Cognitive Science, 14(2), 1990, pp. 179-211.
8. Eren, M., Çelik, A.K., and Kubat, A., **"Determinants of the Levels of Development Based on the Human Development Index: A Comparison of Regression Models for Limited Dependent Variables"**, Review of European Studies, Vol. 6, No. 1, 2014, pp.10-22.
9. Fukuda-Parr, S., **"The human development paradigm: Operationalizing Sen's ideas on capabilities. Feminist Economics"**, 9(2-3), 2003, pp. 301-317.
10. Haykin, S., **"Neural Networks: A Comprehensive Foundation"**, Prentice Hall Inc, New Jersey, 1999.
11. Ivanova, I., Arcelus F.J. and Srinivasan G., **"An assessment of The Measurement Properties of The Human Development Index"**, Social Indicators Research 46, 1999, pp. 157-179.
12. Kohonen, T., **"State of the Art in Neural Computing"**, "Int. Conf. on AI", 1987, pp.1-79.
13. Kohonen, T., **Self-Organizing Maps**, Springer Series in Information Sciences, (3rd Edition), New York, 2001.
14. Küçükönder, H., **"Yapay Sinir Ağları ve Tarımda Bir Uygulama", (Yayımlanmamış Doktora Tezi)**, Kahranmanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü, Kahramanmaraş, 2011.
15. Liao, T. F., **Interpreting Probability Models, Logit, Probit, and Other Generalized Linear Model, Sage Publications**, In Sage University Papers Series. Thousand Oaks, Calif: Sage, 1994.
16. Lind Niels C., "**Some Thoughts on The Human Development Index**", Social Indicators Research 27, 1992, pp. 89-101.
17. Lippman, R. **"An introduction to computing with neural nets"**, IEEE ASSP Mag., 4, 1987, pp. 4-22.
18. Long, J. Scott, ve Freese, J., **Regressıon Models For Categorıcal Dependent Varıables Usıng Stata**, Texas, A Stata Press Publication, 2001.
19. McCullagh, P., **"Regression Models for Ordinal Data"**, Journal of the Royal Statistical Society. Series B (Methodological), Vol. 42, No. 2, 1980, pp.109-142.
20. Morse, S., **"Stirring The Pot. Influence of Changes in Methodology of The Human Development Index on Reporting By The Pres"**, Ecological Indicators 45, 2014, pp.245–254.
21. Öztemel, E., **"Yapay Sinir Ağları"**, Papatya Yayıncılık, İstanbul, 2003.
22. Payne, A., & Phillips, N., **Development. Cambridge**, UK: Polity Press, 2010.
23. Pissarenko, D., "Neural Networks For Financial Time Series Prediction: Overwiev Over Recent Research", 2002, http://members.inode.at/d.pissarenko/fyp/Pissarenko2002.pdf
24. Şen, Z., **Yapay Sinir Ağlarının İlkeleri**, Özener Matbacılık, İstanbul, 2004.
25. UNDP, **Human Development Report**, 1990.
26. UNDP, **Human Development Report**, 2010.
27. UNDP, **Human Development Report**, 2011.
28. UNDP, **Human Development Report**, 2012.
29. UNDP, **Human Development Report**, 2014.

30. Üçdoğruk, Ş., Akın, F., ve Emeç, H., **"Hanehalkı Harcamaların Olasılıklarını Sıralı Regresyon Modeli ile Tahmin Etme"**, V. Ulusal Ekonometri ve İstatistik Sempozyumu, Adana, 2001. (http://idari.cu.edu.tr/sempozyum/bil13.htm).

31. Yıldız, B., **"Finansal Başarısızlığın Öngörülmesinde Yapay Sinir Ağı Kullanımı ve Halka Açık Şirketlerde Ampirik Bir Uygulama"**, İMKB Dergisi, Sayı:17, 2001, s.51-67.

ANNEX-1 Human Development Levels by Countries

Country	Development Level	Country	Development Level
Germany	Highly Developed	**Guatemala**	Moderately Developed
USA	Highly Developed	**Georgia**	Developed
Australia	Highly Developed	**Croatia**	Developed
Austria	Highly Developed	**The Netherlands**	Highly Developed
Azerbaijan	Developed	**Honduras**	Moderately Developed
Belgium	Highly Developed	**England**	Highly Developed
Bolivia	Moderately Developed	**Ireland**	Highly Developed
Bosnia and Herzegovina	Developed	**Spain**	Highly Developed
Brazil	Developed	**Israel**	Highly Developed
Bulgaria	Developed	**Sweden**	Highly Developed
Algeria	Developed	**Switzerland**	Highly Developed
Czech Republic	Highly Developed	**Italy**	Highly Developed
China	Moderately Developed	**Iceland**	Highly Developed
Denmark	Highly Developed	**Jamaica**	Developed
Dominic Republic	Moderately Developed	**Japan**	Highly Developed
Ecuador	Moderately Developed	**Canada**	Highly Developed
El Salvador	Moderately Developed	**Kazakhstan**	Developed
Estonia	Highly Developed	**Cyprus**	Highly Developed
Morocco	Moderately Developed	**Kyrgyzstan**	Moderately Developed
Philippines	Moderately Developed	**Colombia**	Developed
Finland	Highly Developed	**Korea**	Highly Developed
France	Highly Developed	**Costa Rica**	Highly Developed

Country	Development Level	Country	Development Level
Latvia	Developed	**Serbia**	Developed
Lithuania	Developed	**Slovenia**	Highly Developed
Luxemburg	Highly Developed	**Slovakia**	Highly Developed
Hungary	Highly Developed	**Sri Lanka**	Moderately Developed
Macedonia	Developed	**Syria**	Moderately Developed
Malaysia	Developed	**Saudi Arabia**	Developed
Malta	Highly Developed	**Chile**	Developed
Mexico	Developed	**Thailand**	Moderately Developed
Egypt	Moderately Developed	**Trinidad and Tobago**	Developed
Moldova	Moderately Developed	**Tunisia**	Developed
Nicaragua	Moderately Developed	**Turkey**	Developed
Norway	Highly Developed	**Ukraine**	Developed
Pakistan	Moderately Developed	**Uruguay**	Developed
Panama	Developed	**Jordan**	Developed
Paraguay	Moderately Developed	**Venezuela**	Developed
Peru	Developed		
Poland	Highly Developed		
Portugal	Highly Developed		
Romania	Developed		
Russia	Developed		
New Zealand	Highly Developed		
Greece	Highly Developed		

JOURNAL OF APPLIED QUANTITATIVE METHODS

Quantitative Methods Inquires

EVALUATING THE RISK OF PROJECTS IMPLEMENTATION IN VARIOUS SITUATIONS USING GENERALIZED TOPSIS MODEL AND BUSINESS PLAN

Gholamreza JANDAGHI
Professor, University of Tehran, Farabi College, Iran

E-mail: jandaghi@ut.ac.ir

Mahbube HOSSEINI
MBA Student, University of Tehran, Farabi College, Iran

E-mail: m_hosseini88@ut.ac.ir

Abstract

Due to the growing acceptance of project management, the application of appropriate knowledge, processes, skills, tools and updated and new techniques can affect the success of project. Risk of project implementation has always been discussed in prosperity or not reaching the desired goal, and due to complex world, the updated techniques and tools must be used for measuring of this kind of risk. So the purpose of this paper is to evaluate this risk in various situations with the new models that its name is generalized TOPSIS and also used BP that it is one of the important tools before the implementation of the project. In this method the new concept is used; as peril, hazard and risk. A case study of the proposed model will be introduced in the MATLAB that the end method is modeled in it. The sensitivity analysis of the proposed method are discussed in the next and will be measured by using the Microsoft Excel.

Keywords: *Project management, Peril, Hazard, Implementation risk, Business plan, Generalized TOPSIS model, Sensitivity analysis*

1. Introduction

Innovation is the origin and driving force of economic growth in the post-industrial era. The competitiveness of the national economy is stably provided Only by new products and improving existing products. In today's dynamic world innovation is accelerating and goods quickly are copied.[15] But should we act hastily? Business executives in years ago have tried using techniques such as total quality management to restrain implications of changes in their business processes. Approach such as total quality management allows administrators to minimize the risk of changes with the scientific methods and run their business management step-

by-step. Bur developments in the economies that are take place recently, need the new requirements in its business scope. In the new economic situation, different firms and organizations need more agility and flexibility to accept the changes that have been inevitable.[13] Therefore many managers want to know how they can beget greatest change in the business process in the shortest time and at the same time minimize the risk of changes. Sometimes the change is necessary because without it business is highly declined and can not compete with other actors in the economy. However, if changes are not made intelligently, the risks will be created for the business that they may undermine the philosophy of all actions and efforts. [7]

2. What is Business Plan?

James Short and Davenport in 1990 in an article define the idea of designing business plan in this way: " Analysis and design of cycles and new business processes." Design of business processes means analyzed delicately business and designed the existing processes to achieve significant improvements in business efficiency. In the TQM improvement process and more modification are based on statistical control of the business situation, while the process of innovation and change in the design of business processes is based more on the role of information technology. So from this point to discuss IT is come in the change management and business process. In the methodology of designing business process the issue is to make rapid changes that the tools of these changes are IT. Sometimes these changes are to create environmental sustainability that beget IT-based tools and systems and sometimes IT is used to implement changes in business processes. [17]

2.1 Why do you need BP?

According to the definition business plan is a plan that determines the future and develops a business and and often takes a period of several years. [24] Indeed it is a document that expresses company actions is in a specified period and typically includes a detailed list of the risks and uncertainty and analyzes them. [25] Consequently, it is a tool for predicting and managing better of existing businesses or new ones. These works can be accomplished through focus on priorities, monitoring and evaluating progress and contributing to achieve the predetermined objectives. Such a plan would enable individuals to gain necessary preparation for promotion programs in various stages of businesses. [4] In fact, the plan is to respond such question as; why? What thing? How to? Who? What time? How much? And One of the stages of preparation of it is the assessment of the business risk. [25]

3. Risk and its management

Risk in general is probability of incurring losses, and project risk is an uncertain and plausible event and if it occurs, it affects on the positive impact of the project results and objectives (opportunities) or negative impact (threats). [18] Consequently, it is essential in today's world, the risks should be identified and also controlled. So discussion of risk management is important too. Risk management is the systematic process that develops and implements to increase the positive risks (opportunities) and reduce the negative risks (threats).[23] For management policy the Risk Management is one of the good tools in any organization

that he can reduce them by checking and evaluating of existing risks in the system. This tool is widely used in investment, project management, etc. [21]

4. The position of risk management in project management

Project management is a process that plans, organizes, directs and controls the context of project implementation by the possible ways to achieve the desired objectives. In other words, it is a process that coordinates all components of a project. [14] The risk management process is continuously carried out in the project life cycle. [8]

In figure 1 the project management has been modeling and the position of RM is showed in it. [22]

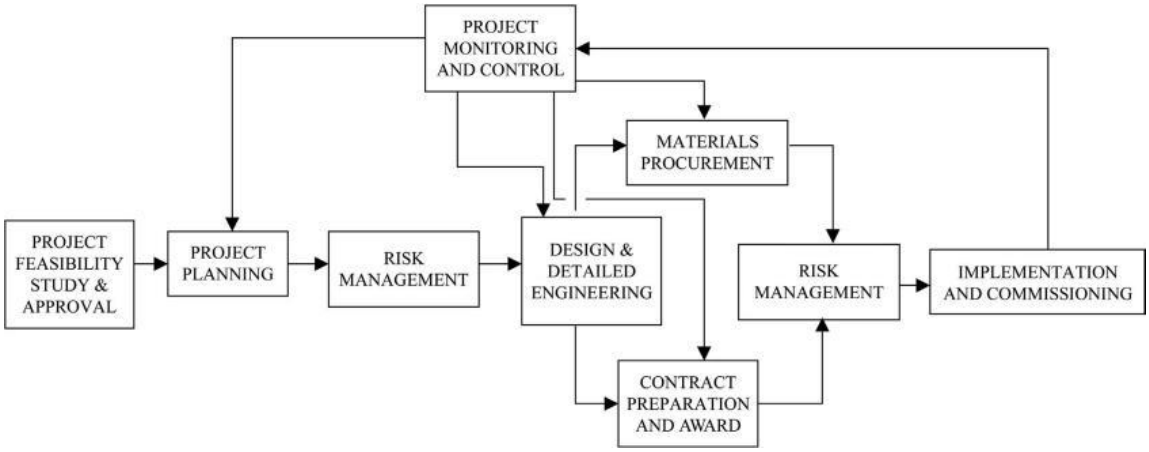


Figure 1. Project management model

5. The main steps in the implementation of risk management [6]

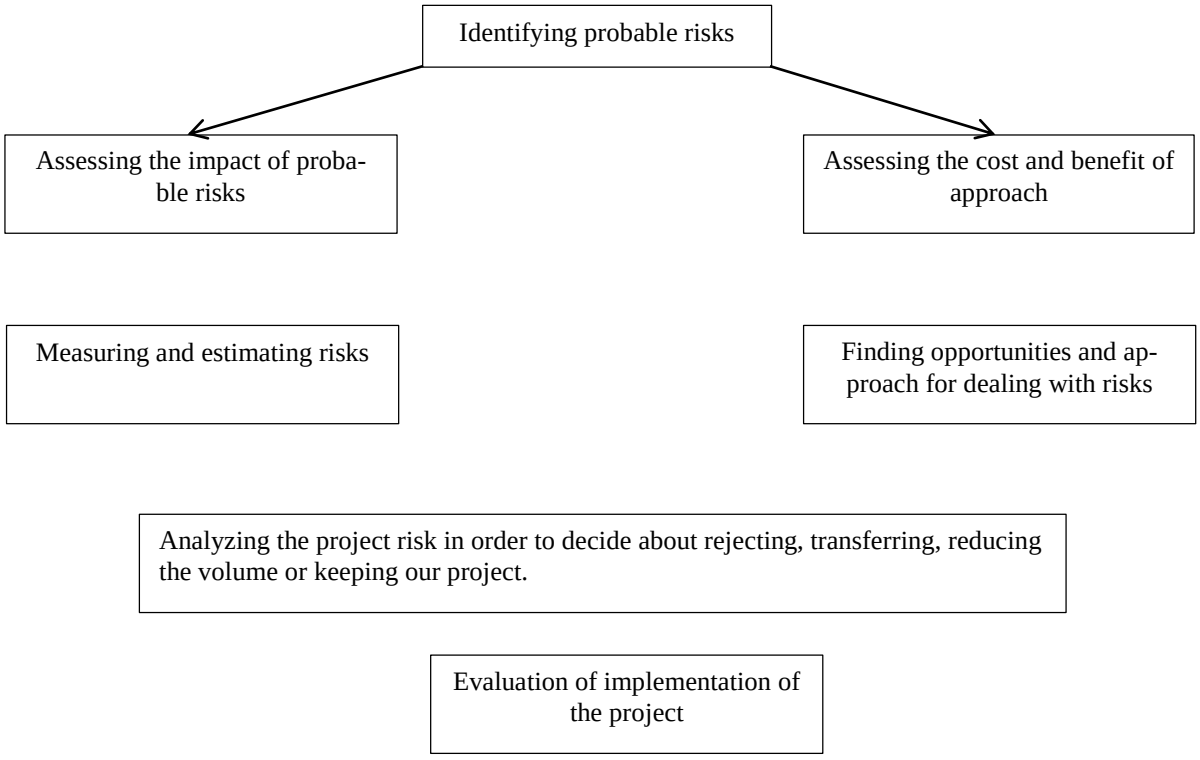


Figure 2. Implementation of risk management

6. Benefit of RM in the project

RM has produced criteria and procedures that persons, financial institutions (commercial and industrial) and non-profit organizations and governments can use it in evaluating prospective of job and controlling and also financing hazard. Accordingly risk management manage the risks by systematic approach. Therefore always in working it is important to respond to two basic questions about probable future events: The first question is "what will happen?" The second one is that "what should we do?" Risk management planes to deal with possible future events. [11] Also it provides the ability for project manager to reduce the risk of failing in achieving the specified benefits by creating the efficient rating system in the process of project implementation, resourcing allocation and activities implementation and phases of project. [2]

7. The relationship between risk management and business plan:

By definition that has been presented can see that the risk management and business plan are closely related. To be able to face the peril before it individuals must be prepared to deal with it. As a result, the business plan can be identified risks.

But it should be noted that the business plan is based on a number of assumptions stated. Regardless of whether you have considered assumptions carefully, The probability that everything goes exactly according to your plan is too low. So if you plan for probable outcome, you're ready to run. It is very important that know areas of your business plan as fragile and vulnerable. [16]

Due to fallow the business plan can be used to manage and control project risks:

1- It can help manager or entrepreneur to specify, focus and examine the aspects and development of their project.
2- Creates logical and accepted framework in which a work can be developed and measures related to the profession will follow in the next few years.
3- Offers criteria for the assessing the situation of a real business to what it should be. [24]

8. Research issue

The purpose of this article is to measure the risk of project implementation in different situations. In fallowing mathematical model offers for reaching to the aim.

Some assumption in this model must be considered:

1- The main objective of the model is to optimize the locating of project and measure the risk of project implementation in different places.
2- It is assumed that by changing the location of the project the ratio of cost and income is constant. [10]

$$\frac{C_1}{I_1} = \frac{C_2}{I_2} = \cdots \qquad (1)$$

C_i= the cost made by project in the i location

I_i= income of project in the i location

3- In this model risk is created only as a threat and do not consider the positive aspects of it. Indeed ,for example, the rate is 0.9 implies that if the probable or unlikely event occurs , probability of threat in that location is equal to 0.9. And it does not mean that 90% can be converted to an opportunity or a threat. [19]

9. Methodology

1- Obtain an indicator that during the project create risk if the amount of them change.

2- Rank the locations of project with the TOPSIS model. This method is based on CL_i criterion and projects will be ranked by this criterion. CL_i of alternative that is higher that ones is in priority. [12] But then you will see that what is important to choose the project is probability of peril.

Before continuing some definition need to determine:

Peril: It is a situation that can lead to an accident or incident.

Risk: It is a chance of peril detonating to an accident or incident.

Hazard: Personal injury arising out of an accident, direct or indirect financial loss caused by incident or accident is known hazard.

Consequences: It is a condition that occurs as a result of actual peril.

Peril= Financial hazard * Risk [8]

Probability (peril)= probability (financial hazard) * probability (risk)

Probability of peril= probability of financial hazard causing * probability of risk

$$\boldsymbol{P_{p_i} = P_{fh_i} * P_{r_i}} \qquad (2)$$

So for choosing project not only the risk is important, but also financial hazard is important, too. In fact, what it is important to choose the project is probability of peril.

3- Calculate the probability of the risk

$$1 - CL_i = P_{r_i} \qquad (3)$$

CL_i: index ranking of TOPSIS method

P_{r_i} = risk probability in i-th location

4- Calculate the probability of financial hazard

For calculating this parameter at first it is required to calculate the financial hazard for each of the locations in each of the indicators.

4.1 Method of calculating the financial hazard for each of the indicators

Financial hazard for any indicator per unit loss should be estimated for each location.

$$fh_{ij} = P_{ij}(create\ loss) * COI_{ij} \qquad (4)$$

i=1,2,….,n

j= 1,2,….,m

fh_{ij}= financial hazard for i location in the j indicator

COI_j (cost or income)= For j indicator when threat is created in the project the amount of cost become high or the amount of income become low

4.1.1 Calculate COI_j

COI amount depends on the type of index that any of its damage to which part of the plan. It may be changed the cost or income. Cost may increase or income may decrease. Depending on the nature of index by the business plan the amount of costs or incomes that the index are influences them consider as amount of COI for that index. For example, consider the supply of raw materials indicator for the construction of the sugar production factory. If the cost of material become high, it's influence the cost. So the cost change. As a result, this index COI is equal to the amount of cost.

4.1.2 Method of calculating the P_{ij}(create loss)

Depending of the type of index that is quantitative or qualitative, calculating the probability of losses is different.

4.1.2.1 Probability of losses by the qualitative index

Weighted index for the j-th criterion (w_j) in the TOPSIS model indicates that the this criterion affects the project with the amount of w_i%. As a result, when the level of criterions become low probability of the creating loss for the project is w_i.

$$P_{ij}(create\ loss) = w_j \quad (5)$$

4.1.2.2 Probability of losses by the quantitative index

$$P_{ij}(create\ loss\ per\ unit\ reduction) = \binom{x_{ij}}{1} w_i (1 - w_j)^{(x_{ij}-1)} \quad (6)$$

If the probability of creating failure or loss is w_i, the probability that x_{ij} (amount of j-th index for the i-th location in the TOPSIS model) reduces per unit, defines by the binomial distribution with (x_{ij},w_j) parameters.

Exception: positive and negative indicators that are at their lowest levels they will not cause harm for the project. Therefore, the probability of loss for these index will be zero.

4.2. Calculate the probability of financial hazard to the entire project

To calculate the peril probability it should be considered that all the parameters are not gone up or down at the same time to cause harm for the project.

$$P_i(hazard\ from\ the\ i-th\ project) = \frac{fh_i}{fh_T} \quad (7)$$

$$fh_i = \sum_{j=1}^{m} fh_{ij} \quad (8)$$

$$fh_T = \sum_{i=1}^{n} fh_i \quad (9)$$

FH_i= financial hazard for the i-th location

FH_{ij}= financial hazard for i location in the j indicator

FH_T= all project financial hazard for total locations

It must estimate the probability that at least one of the m index create financial hazard to the project. This probability is equal to:

$$p_{fh_i} = \sum_{j=1}^{m} \binom{m}{j} p_i^j (1 - p_i)^{m-j} \quad (10)$$

p_{fh_i} = probability of financial hazard creation

m= amount of indicators

p_i= probability of fh from the i-th location

5- Calculate the peril probability

$$\boldsymbol{P_{p_i} = P_{fh_i} * P_{r_i}} \quad (11)$$

P_{p_i}= peril probability of i-th location

p_{fh_i} = probability of financial hazard creation

P_{r_i} = risk probability in i-th location

P_{p_i} represents the parameter that in this article known as risk of the project implementation in different situations. The more peril probability (P_{p_i}) is less, the more project is ideal.

10. Case Study: appropriate location for Municipal solid waste landfills using geographic information system (GIS) and Analytical Hierarchy Process (AHP) (Case study of Iran (city: Ilam))

In this paper we introduce 12 indexes that their weights of each index were determined according to the AHP method. Based on geographic information systems and indicators top priority (location 1) have been selected from 3 locations.

Some assumptions are considered for comparison:

1- Impact indicators and weights of them is intended according to the thesis.
2- Decision matrix according to the definitions in the thesis, is prepared by an expert.
3- COI matrix is expressed with respect to the business plan of the project.

10.1. Decision matrix and indicators weights

Table 1. Decision matrix and indicators weights

Indicators	Type of indicators	Weights	Location 1	Location 2	Location 3
Far from city	Quantitative	0.212	6	7	4
Far from airport	Qualitative	0.167	6	7	6
Stay away from the blue zone	Qualitative	0.13	8	6	8
Stay away from the main roads	Quantitative	0.104	6	4	8
Slope	Quantitative	0.09	8	6	4
Stay away from protected areas	Qualitative	0.068	7	7	7
Geology	Quantitative	0.06	5	8	7
Land use	Qualitative	0.048	8	6	8
Stay away from the village	Qualitative	0.041	7	7	7
Fault	Qualitative	0.033	7	7	7
Height	Quantitative	0.026	8	6	4
Precipitation	Qualitative	0.021	6	6	6

10.2. Calculate the relative proximity index and risk probability

Table 2. Relative proximity index and risk probability

	Location 1	Location 2	Location 3
CL	0.633	0.592	0.381
P_r	0.367	0.408	0.619

10.3. COI matrix

Table 3. COI index

Indicators	COI (dollars per ton)
Far from city	0.51
Far from airport	0.51
Stay away from the blue zone	0.51
Stay away from the main roads	0.51
Slope	93.75
Stay away from protected areas	0.51
Geology	93.75
Land use	93.75
Stay away from the village	0.51
Fault	973.75
Height	973.75
Precipitation	973.75

10.4. Calculate the probability of financial hazard

Table 4. Probability of financial hazard

	Location 1	Location 2	Location 3
P_{fh}	0.99998997436	0.999997555	0.99999980368

10.5. Measuring risk of the project implementations

Table 5. Risk of the project implementation

	Location 1	Location 2	Location 3
P_p	0.366561994	0.407779	0.6194821

Location 1 with the lowest risk (36.66%) is selected as the preferred location. Risk of the project implementation in location 2 is 40.78% and in location 3 is 61.94%. This ranking is consistent with the end priorities of the thesis. But this method shows risk and also ranking.

In any decision method risk has not been investigated and values obtained from the models are used to rank the options. For this reason, this method is distinct from other methods. Furthermore in the next section another distinctive feature of this method will be specified.

11. Sensitivity analysis of the proposed method (as used in this case) using a stability index

Sensitivity analysis (SA) is the study of how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input. The SA is hence considered by some as a prerequisite for model building in any setting, be it diagnostic or prognostic, and in any field where models are used. Kolb (quoted in Rabitz, 1989) noted that theoretical methods are sufficiently advanced, so that it is intellectually dishonest to perform modeling without SA. [3]

In fact, it checks the relationship between input and output of information and discuss the modifications that affect the output of the model. [20] There are different ways for sensitivity analysis that Campolongo (2001) and Saltelli and Chan (2000) fully reviewed these methods.

One of these methods is changing the index weights. But it should be noted that the output of this proposed model are two different things:

1- Ranking the project locations
2- Calculating the risk of project implementation in the site chosen

Therefore, it is necessary to do sensitivity analysis on both output per changing the index weights. To change the weights of indicators it is used an index that called instability index and its sensitive will be discussed in those two output. [1]

11.1. sensitivity analysis of first section

This part is focused on changing of the final ranking of project per changing of the indicator weights.

Since the thesis is used AHP method and paired comparisons for weighting the indicators, for analyzing the sensitivity the instability index (L') will be used. This index indicates that in which extent decision maker judging can be such that the final ranking does not change. Therefore in the AHP matrix elements above the main diagonal should be multiplied on L' and the elements below the main diagonal should be divided on L' ($0<L'<1$). The L 'gradually

reduced from the unit until the final ranking does not change. As soon as ranking changing causes, the decreasing of L' must be stopped and the instability index is obtained. Stability index (L) is defined as fallowed:

$$L = \frac{1}{L'} \qquad (12)$$

The lower the instability index is, the less sensitive model is to changes in weights. [5]

Figure 1 shows the trend of changes in the instability index and the final result for the case study. According to the entries expressed first part of curves that intersect will be presented as an instability indicator.

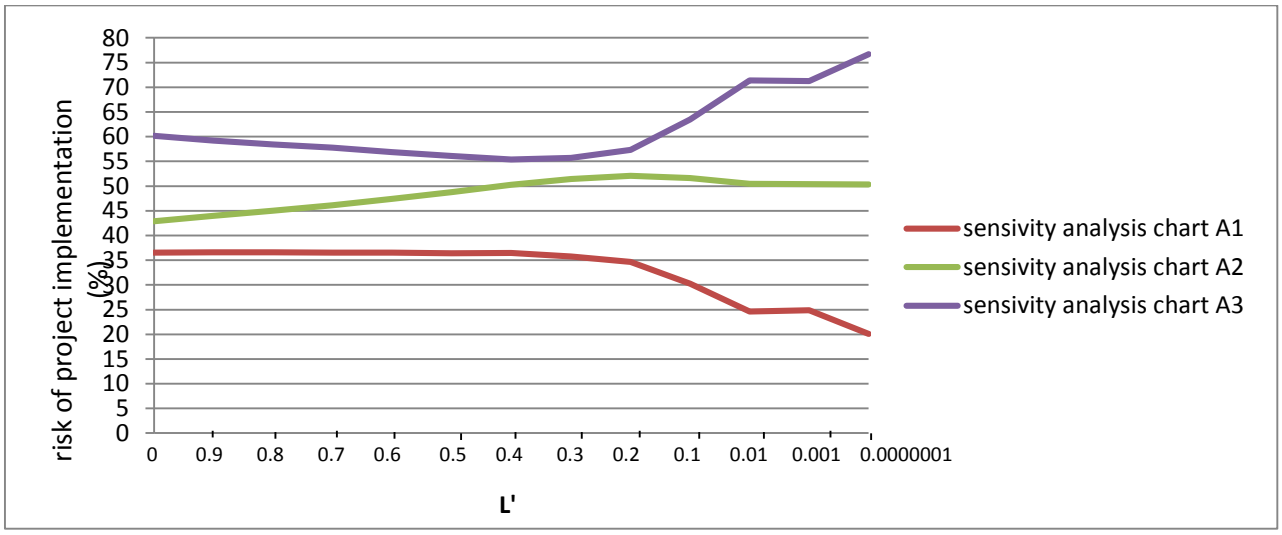


Chart 1: sensitivity analysis based on the instability index for the first section

Decreasing L' has continued until 10^{-7} but as the chart shows none of the graphs intersect. So we can say that the final ranking model is not sensitive of weights changing or it's almost zero. This case seems reasonable. Because the ultimate goal of the model is measuring the locations risk and final ranking should not be changed by changing the weights. For accurate investigation, this claim will be examined in future research.

11.2. Sensitivity analysis for second section

The previous section outlined that model ranking sensitivity is very few by changing in the weights. In this part the sensitivity analysis will be done based on the risk of implementing the project in location 1 in case study by changing the weights of indicators.

In this section we will use the instability index, too. But the difference is that amount of L' will decrease until the amount of deviation in risk has arrived to 5%. Chart 2 shows decreasing of L' and percentage of changing in the amount of risk. [5.]

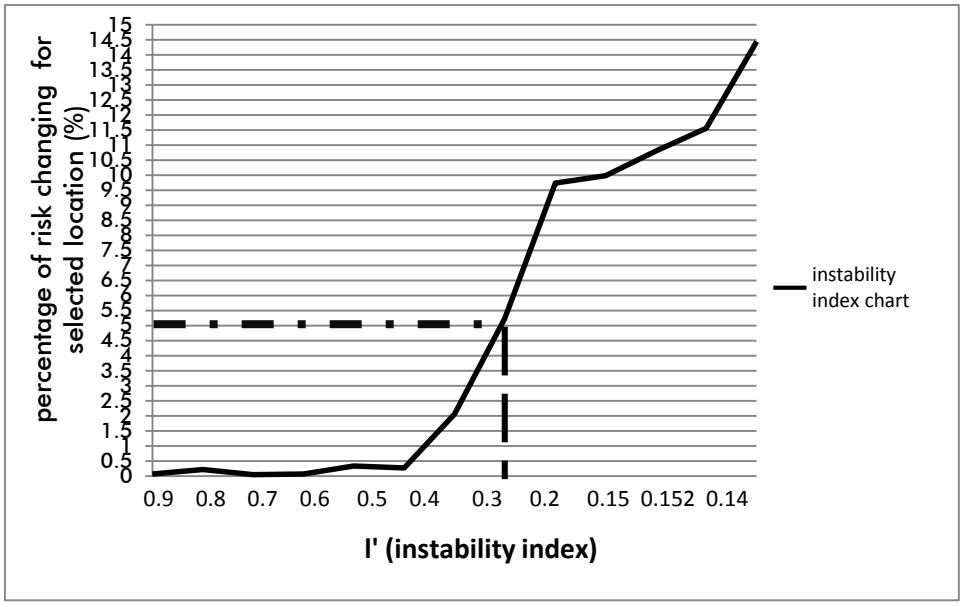


Chart 2: sensitivity analysis based on the instability index for the second section

According to the figure 1 when the L'=0.2, Percentage of risk changing for location 1 is equal to 5%. As a result L=5. Since the instability index is low, so the amount of risk is less sensitive to changes in weight.

12. Discussion and conclusion

At the management level, a clear vision that is created by the use of risk management system is to facilitate and improve the quality of decision making in the project. To create a clear vision should always combine qualitative and quantitative approaches to decide better. [9] In this paper, new model called generalized TOPSIS method has been presented by integrating the business plan as a basis for the start of a project and TOPSIS method as a quantitative method for ranking decision problems. In this method three elements have defined: peril, financial hazard and risk and with these elements risk of project implementation measures and the best options is selected in the terms of the lowest risk.

One of the attributes of this model that differentiate it from other forms of multi-criteria decision making is that in addition to ranking options this model measures the peril probability and Final numbers obtained will be not applications only for ranking.

Sensitivity analysis is also examined in this article in a particular case and has been identified that final ranking of the model is not sensitive relative to the weights changing and the method for analyzing the sensitivity is stability index and model sensitivity from this index is 5 relative to the 5% changing in the amount of the project implementation risk in the best location. Since the stability index for the peril probability of the project implementation is very high, the model has very few sensitivity.

13. Suggestion

To develop more and better the model the fallowing items will be studied in future research, included:

1- According to the first assumption in this model risk is considered as its negative aspect. In the future research this one can be developed and the positive aspect of it can be considered, too. One the way that can remove this problem is that after measuring the risk of project implementation this factor can again come to the model as a indicator and with the one of the multi-criteria decision making model such as TOPSIS resolved the issue. The name of this indicator can be "unforeseen options".
2- One of the factors influencing the decision-making model and also this model is to explain the impact indicators. On the other hand the PMBOK standard is a complete guide for Project Management. So with Combining this standard by business plan and fuzzy expert system can explain effective indicators on these issues and define the scientific basis for the indicators. [19]
3- About the sensitivity analysis in future article can speak better and more. Also can analyze sensitivity of the model with the others one.

References

1. Alfredo del Can˜o, P.E., M.ASCE, and M. Pilar de la Cruz, P.E , "**Integrated Methodology for Project Risk Management**", Journal of construction engineering and management, Nov. Dec 2002.
2. Alison Mobey, David Parker, "**Risk evaluation and its importance to project implementation**", International Journal of Productivity and Performance Management, 2002.
3. Andrea Saltelli, "**Sensitivity Analysis for Importance Assessment**", Risk Analysis, Vol. 22, No. 3, 2002.
4. Benson Honig, Tomas Karlsson, "**Institutional forces and the written business plan**", Journal of Management vol. 30 no. 1, February 2004, 29-48,.
5. Che-Wei Chang, Cheng-Ru Wu, Chin-Tsai Lin, Huang-Chu Chen, "**An application of AHP and sensitivity analysis for selecting the best slicing machine**", journal of Computers & Industrial Engineering Volume 52, Issue 2, March 2007, Pages 296–307.
6. Crouhy, M., Dangalai, R, "**The essentials of risk management**", chapter 1, Risk Management, 2000.
7. David L. Cleland, "**Project Manager's Handbook**", University Of Pittsburgh-Pittsburgh, and Lewis R. Ireland, 2008.
8. FAA, "**The system safety handbook",** Dec. 2000.
9. Fiona D. Patterson, Kevin Neailey, "**A Risk Register Database System to aid the management of project risk**", International Journal of Project Management 20 (2002), 365–374.
10. Frank T. Anbari, "**Earned value project management method and extensions**", Project Management Journal; Dec 2003.
12. Gary Stoneburner, Alice Goguen, and Alexis Feringa, "**Risk Management Guide for Information Technology Systems**", Recommendations of the National Institute of Standards and Technology, July 2002.

13. G.R. Jahanshahloo, F. Hosseinzadeh Lotfi, M. Izadikhah, "**An algorithmic method to extend TOPSIS For decision-making problems with interval data**", Applied Mathematics and Computation 175, 2006.
14. J. Rodney Turner, "**The handbook of project-based management**", 3rd Edition, 2009.
15. Jack R. Meredith, Samuel J. Mantel. JR, "**Project management, a managerial approach**", 2011.
16. James M. Lyneis, David N. Ford, "**System dynamics applied to project management: a survey, assessment, and directions for future research**", System Dynamics Review, Summer/Fall 2007.
17. John Rafter, "**Risk analysis in project management**", 1994.
18. Linda Pinson, "**Anatomy of a business plan: the step-by-step guide to building a business and securing your company's future**" , 2008.
19. Lionel Galway, "**Working paper: Quantitative Risk Analysis for Project Management ; A Critical Review**", WR-112-RC, February 2004.
20. Kalle Kähkönen, "**Integration Of Risk And Opportunity Thinking In Projects**", Presented at the Fourth European Project Management Conference, PMI Europe 2001, London UK, 6-7/ June 2001
21. M. Saisana, A. Saltelli and S. Tarantola, "**Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators**" J. R. Statist. Soc.A, 2005.
22. Paul Hopkin, "**Fundamental of risk management**", third edition, 2014.
23. Prasanta Kumar Dey, "**Decision support system for risk management: a case study**", Management Decision, 2001.
24. Project Management Institute, Inc, "**A Guide to the Project Management Body of Knowledge**", chapter 11, Risk Management, 2004.
25. Tomas Karlsson, Benson Honig, "**Judging a business by its cover: An institutional perspective on new ventures and the business plan**", Journal of Business Venturing, Pages 27–45, January 2009.
26. http://www.entrepreneur.com/article/38290

SHA FAMILY FUNCTIONS

Stelian DUMITRA
PdD, Department of Economic Informatics Doctoral School
The Bucharest University of Economic Studies
Bucharest, Romania

E-mail: stelian.dumitra@endava.com; stelu20d@yahoo.com

Abstract

This paper presents a general overview regarding SHA family functions. A lot of hash functions were proposed in the last three decades and most of them are based upon the MD construction, especially the MD4 family. The most popular hash functions that belong to the MD4 family are: MD5, the SHA family (SHA-0, SHA-1, SHA-224, SHA-256, SHA-384, SHA-512) and the RIPEMD family (RIPEMD, RIPEMD-128, RIPEMD-160, RIPEMD-256, RIPEMD-320). In 2004, collision attacks against MD5 and SHA-0 were demonstrated by Xiaoyun Wang and one year later she extended a theoretical attack against SHA-1. NIST took the Wang attack into serious consideration and decided to open a competition to develop the next secure hash algorithm, named SHA-3. After five years of competition NIST selected Keccak as winner. Keccak is a family of sponge functions. Also, this paper describes the sponge construction and its security. Various comparisons between SHA-3, SHA-2 and SHA-1 regarding the security strengths, performances and construction details are summarized. The conclusions are presented at the end of the paper.

Keywords: *Keccak, SHA-3, hash function, collision resistance, preimage attack, sponge construction*

1. Introduction

Cryptographic hash functions play an important role in the current cryptographic protocols and they are used to ensure data integrity, data origin authentication, password protection, random number generation and the likes of them. They produce a fixed-length output, *digest*, that can be also treated as a *fingerprint* of the input data [7]. A lot of security protocols and applications use hash functions: digital signature scheme for authentication data such as DSS (Digital Signature Scheme), XML signature, computing MAC (Message Authentication Code)/HMAC (Keyed-Hash Message Authentication Code), secure communication protocols such as SSH (Secure Shell Host), SFTP (Secure File Transfer Protocol), SSL (Secure Socket Layer), IPSec (Internet Protocol Security) etc, Kerberos protocol for authentication and data integrity, PGP (Pretty Good Privacy), S/MIME (Secure/Multipurpose Internet Mail Extensions) for integrity of e-mail messages.

A hash function ($h: M \rightarrow M_h$) is a function that transforms a variable-length input into a fixed-length output – *hash value* (for example, 128, 160, 224, 256, 384, 512 bits). In practice there are two classes of hash functions:

1. *One Way Hash Functions – OWHF*
2. *Collisions Resistant Hash Functions – CRHF*

An OWHF function satisfies the following properties

- Given a digest value $d \in M_h$ it is computationally infeasible to find a message $m \in M$ so that $d = h(m)$.
- Given a message m_1 it is computationally infeasible to find another message $m_2 \neq m_1$ so that $h(m_1) \neq h(m_2)$.

For a CRHF function it is computationally infeasible to find two different messages m_1 and m_2 so that $h(m_1) = h(m_2)$. This means that the digests are almost unique for each given message. The OWHF functions are also known as *one-way weak collision resistance* and CRHF functions are also known as *strong collision resistance*.

Another class of hash functions is MAC (Message Authentication Codes) that is a function of the symmetric key *k* and the message *m*, $m = MAC_k(x)$. A lot of hash functions were proposed in the last three decades and most of them are based upon the MD (Merkle-Damgård) construction [8], especially the MD4 family [1]. The most popular hash functions that belongs to the MD4 family are: MD5 [2], the SHA family (SHA-0, SHA-1, SHA-224, SHA-256, SHA-384, SHA-512) [3] and the RIPEMD family (RIPEMD, RIPEMD-128, RIPEMD-160, RIPEMD-256, RIPEMD-320) [4, 5, 6]. The MD4 digest algorithm was developed by Ronald Rivest and the digest length is 128 bits. All operations are bitwise Boolean functions: AND, OR, XOR and negation. Boer, Bosselaers and Hans Dobbertin demonstrated weaknesses in MD4 and this algorithm is not recommended for secure hashing [9, 10]. A new strengthened version of MD4 was introduced by Rivest in 1991 and computes a 128-bit output, called MD5 and possess a collision resistance of about 2^{64}. As MD4, this presented potential weaknesses because pseudo-collisions were found on its compression function. In 1993, the US National Institute of Standard and Technology (NIST) published a new message digest standard, SHA (Secure Hash Algorithm). The first version was SHA-0 and, in 1994, a new version was published, SHA-1, derived from SHA-0 with some changes. SHA-0 and SHA-1 produce an output length of 160 bit. In the absence of analytical attacks, the maximum collision resistance of SHA-0 and SHA-1 is of about 2^{80}. The known attack on SHA-0 was developed by Joux and Chabaud [11], a differential attack that finds two messages hashing with the same value in about 2^{61}evaluations. In 2000 NIST introduced three more variants of SHA-1: SHA-256, SHA-384 and SHA-512, functions that produce a message digest with a length of: 256, 384 and 512 bits. These functions were adopted as standard, SHA-2, by FIPS in 2002. A new modification of SHA-1 was introduced in 2004, SHA-224, to fit the security level of 3DES. This function was also included in SHA-2 standard.

In 2004, collision attacks against MD5 and SHA-0 were demonstrated by Xiaoyun Wang [12]. One year later Wang extended a theoretical attack against SHA-1 and it was claimed that a collision search would take 2^{69} steps [13]. An improved version of this attack was presented by Wang in August 2005 with the time complexity of 2^{63}(a brute-force search would require 2^{80} operations) [14]. Other cryptographic attacks on SHA-1 were proposed by Christophe De Cannière and Christian Rechberger [15], Grechnikov [16], Stéphane Manuel [17], Cameron McDonald, Philip Hawkes and Josef Pieprzyk [18], Marc Stevens [19].

Although no serious flaws were disclosed against SHA-2, NIST took the Wang attack into serious consideration and decided to open a competition to develop the next secure hash

algorithm, named SHA-3. NIST did not plan to replace SHA-2 with SHA-3, as it considered that both functions should co-exist. Below is the timeline of the SHA-3 selection process [20]:

- November 2, 2007: NIST announces a request for a new cryptographic hash function – SHA-3 [21]
- October 31, 2008: Submission deadline. 64 submissions were received from the international cryptography community.
- December 10, 2008: The first round began. NIST selected 51 algorithms for Round 1.
- July 24, 2009: The second round was announced. NIST selected 14 algorithms for Round 2.
- December 9, 2010: The third round was announced and 5 algorithms were selected:
  - *BLAKE* by Jean-Philippe Aumasson, Luca Henzen, Willi Meier and Raphael C.-W.Phan
  - *Grøstl* by Lars Ramkilde Knudsen, Praveen Gauravaram, Krystian Matusiewicz, Florian Mendel, Christian Rechberger, Martin Schläffer and Søren S. Thomsen
  - *JH* by Hongjun Wu
  - *Keccak* by Joan Daemen, Guido Bertoni, Michaël Peeters and Gilles Van Assche
  - *Skein* by Bruce Schneier, Niels Ferguson, Stefan Lucks, Doug Whiting, Mihir Bellare, Tadayoshi Kohno, Jesse Walker, Jon Callas
- October 2, 2012: NIST selected Keccak as winner.

Table 1. General information about SHA-3 finalists and SHA-2 in bits

Algorithm	Domain Extender	Underlying Primitive	Primitive size	Hash size
BLAKE	HAIFA	Block cipher	k=512 b=512	224 256
			k=1024 b=1024	384 512
Grøstl	Grøstl	A pair of permutations	512 512	224 256
			1024 1024	384 512
JH	JH	Permutation	1024	224 256 384 512
Keccak	Sponge	Permutation	1600	224 256 384 512
Skein	UBI	Tweakable block cipher	k=512 b=512 t=128	224 256 384 512
SHA-2	MD	Block cipher	k=512 b=256	224 256
			k=1024 b=512	384 512

2. The SHA-2 Hash Family

The SHA-2 functions: SHA-224, SHA-256, SHA-384, SHA-512 are the next generation of SHA-1 function standardized by US NIST in 2002 [3]. Initially there were two functions in this standard: SHA-256 and SHA-512. Later, in addition, another two truncated versions were standardized: SHA-224 and SHA-384. The SHA-2 functions are described in detail below.

Padding the Message. The message *M* of length *l* bits is right-padded with a binary *"1"* followed by *k* zero bits, followed by *s-bits* suffix containing the binary length of the original message.

$$\begin{aligned} l+1+k &\equiv 448\, mod\ 512,\ s=64, && \text{for SHA-224 and SHA-256} \\ l+1+k &\equiv 896\, mod\ 1024,\ s=128, && \text{for SHA-384 and SHA-512} \end{aligned} \tag{1}$$

The length of the padded message should be a multiple of 512 bits. The padded message must be parsed into N-512 bit blocks, for SHA-224/256, respectively N-1024 bit blocks, for SHA-384/512, $M^{(1)}, M^{(2)}, \dots, M^{(N)}$.

Computing the Message Digest. The algorithm uses 64 (resp. 80) round functions for processing a single message block. An input block $M^{(i)}$ is expressed as $M_0^{(i)} M_1^{(i)} .. M_{15}^{(i)}$, where $M_t^{(i)}$ are 32-bit (resp. 64-bit) words. The words of the message schedule are labeled $W_0, W_1, \dots, W_t$, where t denotes the number of the rounds – 1, 63 for SHA-224/256, respectively 79 for SHA-384/512.

After the padding phase, eight working state registers *a, b, c, d, e, f, g, h* are initialized with the $(i-1)^{th}$ hash value $H_0^{(i-1)}, H_1^{(i-1)}, \dots, H_7^{(i-1)}$. The initial hash $H^{(0)}$ must be initialized with 32-bit constants, for SHA-224/256, respectively 64-bit constants for SHA-384/512. The steps for computing the message digest are described below:

Step 1. Preparing the message schedule, $\{W_t\}$***:***

$$W_t = \begin{cases} M_t^{(i)} & 0 \le t \le 15 \\ \sigma_1(W_{t-2}) + W_{t-7} + \sigma_0(W_{t-15}) + W_{t-16} & 16 \le t \le 63\ (or\ 79) \end{cases} \tag{2}$$

For SHA-224/256, the functions σ_0 and σ_1 are defined as:

$$\begin{aligned} \sigma_0(x) &= ROTR^7(x) \oplus ROTR^{18}(x) \oplus SHR^3(x) \\ \sigma_1(x) &= ROTR^{17}(x) \oplus ROTR^{19}(x) \oplus SHR^{10}(x) \end{aligned} \tag{3}$$

For SHA-384/512, the functions σ_0 and σ_1 are defined as:

$$\begin{aligned} \sigma_0(x) &= ROTR^1(x) \oplus ROTR^8(x) \oplus SHR^7(x) \\ \sigma_1(x) &= ROTR^{19}(x) \oplus ROTR^{61}(x) \oplus SHR^6(x) \end{aligned} \tag{4}$$

Step 2. Initializing the variables a, b, c, d, e, f, g, h with the $(i-1)^{th}$ ***hash value:***

$$a = H_0^{(i-1)}, \dots, h = H_7^{(i-1)} \tag{5}$$

Step 3. For t = 0 to 63, the following values are calculated:

$$
\begin{aligned}
T_1 &= h + \sum_1(e) + Ch(e,f,g) + K_t + W_t \\
T_2 &= \sum_0(a) + Maj(a,b,c) \\
h &= g \\
g &= f \\
f &= e \\
e &= d + T_1 \\
d &= c \\
c &= b \\
b &= a \\
a &= T_1 + T_2
\end{aligned}
\tag{6}
$$

For SHA-224/256, the functions $\sum_0$ and $\sum_1$ are defined as:

$$
\begin{aligned}
\sum_0(x) &= ROTR^2(x) \oplus ROTR^{13}(x) \oplus ROTR^{22}(x) \\
\sum_1(x) &= ROTR^6(x) \oplus ROTR^{11}(x) \oplus ROTR^{25}(x)
\end{aligned}
\tag{7}
$$

For SHA-384/512, the functions $\sum_0$ and $\sum_1$ are defined as:

$$
\begin{aligned}
\sum_0(x) &= ROTR^{28}(x) \oplus ROTR^{34}(x) \oplus ROTR^{39}(x) \\
\sum_1(x) &= ROTR^{14}(x) \oplus ROTR^{18}(x) \oplus ROTR^{41}(x)
\end{aligned}
\tag{8}
$$

Ch and *Maj* are two logical functions that operate on 32-bit (64-bit) words and three variables, *x, y, z*.

$$
\begin{aligned}
Ch(x,y,z) &= (x \wedge y) \oplus (\neg x \wedge z) \\
Maj(x,y,z) &= (x \wedge y) \oplus (x \wedge z) \oplus (y \wedge z)
\end{aligned}
\tag{9}
$$

K_t are round constants on 32-bit (64-bit) words, where $0 \leq t \leq 63$ or $0 \leq t \leq 79$.

Step 4. Computing the i^{th} intermediate hash value $H^{(i)}$:

$$
H_0^{(i)} = a + H_0^{(i-1)}, \ldots, H_7^{(i)} = h + H_7^{(i-1)} \tag{10}
$$

After processing $M^{(N)}$ the resulting message digest of the message, M, is:

$$
H_0^{(N)} \parallel H_1^{(N)} \parallel H_2^{(N)} \parallel H_3^{(N)} \parallel H_4^{(N)} \parallel H_5^{(N)} \parallel H_6^{(N)} \parallel H_7^{(N)} \tag{11}
$$

SHA-224/SHA-384 is defined in the same manner as SHA-256/SHA-512 with the following differences: uses different constants initialization, $H^{(0)}$ and the message digest is truncated at 224/384 bits as: $H_0^{(N)} \parallel H_1^{(N)} \parallel H_2^{(N)} \parallel H_3^{(N)} \parallel H_4^{(N)} \parallel H_5^{(N)} \parallel H_6^{(N)}$, respectively

$H_0^{(N)} \parallel H_1^{(N)} \parallel H_2^{(N)} \parallel H_3^{(N)} \parallel H_4^{(N)} \parallel H_5^{(N)}$.

More details about implementing SHA-2 are described in [3].

3. Merkle-Damgård construction

Merkle-Damgård construction is a hash construction method which was designed by R. Merkle [22] and I. Damgård [23] in 1989.

It transforms a compression function, $h: \{0,1\}^{m_c} \times \{0,1\}^{n} \rightarrow \{0,1\}^{m_c}$, into a hash function. The m_c denotes the size of the chaining value and n denotes the block size for the compression function. Most of the hash functions are built upon MD construction (Figure 1). It begins with a padding step where the message M is padded so that the message length becomes a multiple of message block length, n. The most used procedure is: *the message M of length l bits is right-padded with a binary "1" followed by k zero bits, followed by s-bit suffix containing the binary length of the original message, so that*

$$l + 1 + k + s \equiv 0 (mod\ n). \tag{12}$$

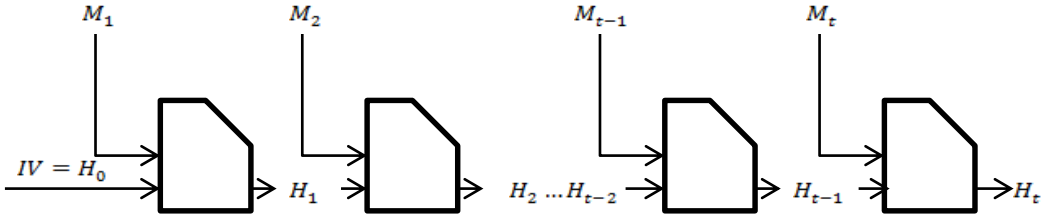


Figure 1. The Merkle-Damgård costruction

The message is divided into block of n bits each, $M = M_1 M_2 \ldots M_t$. An initial chaining value is set for the hash function, publicly known initialization vector, *IV*, and a process is repeated *t* times:

$$\begin{aligned} H_0 &= IV \in \{0,1\}^{m_c} \\ H_i &= h(H_{i-1}, M_i), i = 1, 2, .., t \end{aligned} \tag{13}$$

The final H_t is outputted as the hash value, i.e. $H(M) = H_t$.

If the compression function is collision resistant then the hash function itself is collision resistant, so the collision resistance is preserved. Also, the pre-image resistance and second pre-image resistance of the compression function are preserved [22, 23].

4. Sponge construction

The sponge construction is a mode of operation, based on a fixed-length permutation (or transformation) *f*, a padding rule and a parameter bitrate *r*, which builds a function mapping variable-length input to variable-length output [24]. The permutation *f* operates on a fixed number of bits, the width *b*. The value $c = b - r$ is called the *capacity* [25]. This construction is used for building hash functions and stream ciphers. When it is used as a hash function, called SHA-2 Replacement Mode, the sponge function receives a variable-length input and produces a fixed-length output (SHA-3 224/256/384/512). If a sponge function is used as a stream cipher, called Variable-length Output Mode, it receives a fixed-length input and produces a variable-length output [26]. Also, the sponge functions are used for generating pseudo-random bits. A sponge construction can be expressed as a random permutation,

where the construction is called a P-sponge (random sponge), or random function, where the construction is called a T-sponge [27]. Before the sponge construction phase is performed a pre-processing phase where all the bits of the state are initialized to zero, the message is padded to a multiple of *r* and cut into block of *r* bits.

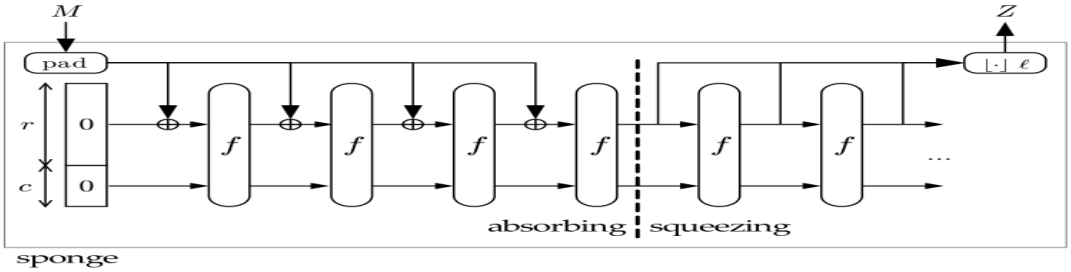


Figure 2. The sponge construction (Source: [24], p.13)

The construction consists of two phases: the absorbing phase and the squeezing phase.

- In the absorbing phase the *r*-bit input message blocks are XOR-ed and overwritten to the state, the *f* is applied to this state. After all blocks are processed, second phase is applied.
- In the squeezing phase a part of the state is returned as output blocks and *f* is applied to the state. The process is repeated in the same manner until the number of the chosen blocks by the user is achieved.

The difference between the compression functions of MD construction and the functions from sponge construction is that a function used in sponge construction maps 1 bit input into 1 bit output. When *f* is expressed as a random permutation the lower bound for the complexity of a collision is $min(2^{n/2}, 2^{c/2})$ and of a pre-image and second pre-image is $min(2^{n}, 2^{c/2})$, where *n* is the hash size. If c ≥ 2n and *f* is a random function, then the sponge construction is differentiable from a random oracle, the strength against signature forgery is increased from $2^{n/2}$ to 2^{n}. More about theory of sponge constructions and their security properties are provided in [24].

5. Keccak hash function

Keccak, became the new SHA-3 standard, is a family of sponge functions. It can be used in two principle modes:

- **SHA-2 Replacement Mode** – SHA-3 produces a fixed-length output of 224, 256, 384 or 512 bits
- **Variable-length Output Mode** – SHA-3 can generate arbitrarily many output bits, so it can be used as a stream cipher or pseudorandom bit generator.

In the pre-processing phase the message *m* is padded as follows:

$pad(m) = m \parallel P10^{*}1$, where P is the bit string representation of the message m, followed by a 1, then by a smallest number of 0s and then again a 1, so that $len(P10^{*}1)\ mod\ r \equiv 0$.

In the case of SHA-3, the width of the state, *b*, is:

$b = r + c = 5 \cdot 5 \cdot 2^{l},\ l = 0,1,\ldots,6$, so $b \in \{25, 50, 100, 200, 400, 800, 1600\}$

The values $b = 25$ and $b = 50$ are not used in practice. For SHA-3 a state of $b = 1600$ bits is used and $r \in \{1088, 1344\}$. The parameters of SHA-3 are represented in bits, in Table 2:

Table 2. The parameters of SHA-3

b (state)	r	c	security level	hash output
1600	1344	256	128	224
1600	1344	256	128	256
1600	1088	512	256	384
1600	1088	512	256	512

5.1 The Keccak-f permutation

Keccak-f[b] is a permutation over Z_2^b. The state (figure-3) consists of a 5×5 array of 64-bit words, a three-dimensional array of elements of *GF(2)*, *a[5][5][w]*, where $w = 2^l$. An element is denoted as *a[x][y][z]*, $x, y \in Z_5,\ z \in Z_w$. The string representation of the state is denoted as *s* and its bits are indexed from *0* to *b-1*.

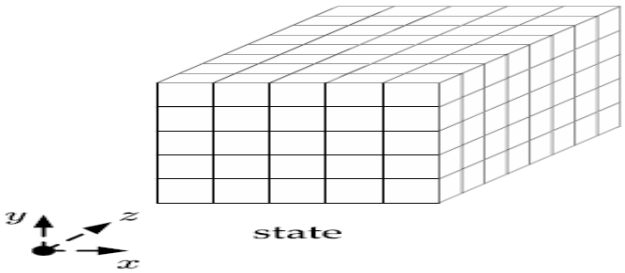


Figure 3. The state of Keccak where each small cube represents one bit (Source: [25], p.11)

The mapping between the bits of *a* and those *s* is:

$$s[w(5y + x) + z] = a[x][y][z] \tag{14}$$

The *Keccak-f[b]* function consists of $n_r = 12 + 2l$ rounds, where each round consists of *b* bits.

Table 3. Number of rounds within Keccak-f (for SHA-3: b=1600, $n_r = 24$)

state width b	number of rounds n_r
25	12
50	14
100	16
200	18
400	20
800	22
1600	24

Each round consists of five sub-rounds, denoted by Greek letters: $\theta(theta), \rho(rho), \pi(pi), \chi(chi)\ and\ \iota(iota)$ [25], [28].

$\theta(a)$:

$$c[x] = a[x,0] \oplus a[x,1] \oplus a[x,2] \oplus a[x,3] \oplus a[x,4], x = 0,1,2,3,4$$
$$d[x] = c[x-1] \oplus rot(c[x+1],1), \quad x = 0,1,2,3,4$$
$$a[x,y] = a[x,y] \oplus d[x], \quad x,y = 0,1,2,3,4$$

$\rho(a)$:

$$a[x][y][z] = a[x][y][z-(t+1)(t+2)/2], where\ 0 \le t < 24, \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}^t \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

$\pi(a)$:

$$a[x][y] = a[x'][y'],\ where \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (15)$$

$\chi(a)$:

$$a[x] = a[x] \oplus (a[x+1] \oplus 1) \wedge a[x+2]$$

$\iota(a)$:

$$a[0,0] = a[0,0] \oplus RC[i_r],$$ where $RC[i_r]$ - round constants, $0 \le i_r \le n_r - 1$

5.2 Keccak vs. SHA-2

After collision attacks against MD5 and SHA-0 were demonstrated by Wang in 2004 and theoretical attacks against SHA-1 were proved to exist serious doubts that these will pose a practical threat against SHA-2 arose. NIST decided to simultaneously use two standards: SHA-2 and SHA-3. The SHA-3 functions are alternatives to the SHA-2 functions. These two standards use different design principles: SHA-2 uses MD principle and Keccak uses Sponge construction. The security strengths, performances and a various details regarding construction of the functions SHA-1, SHA-2 and SHA-3 are summarized in Table 4.

Table 4. Comparison of SHA functions (Source: [28, 29])

Function	Output Size (bits)	Internal State size (bits)	Block size (bits)	Max message size(bits)	Rounds	Example Performance (MiB/s)[29]	Security Strengths in Bits		
							Collision	Preimage	2nd Preimage
SHA-1	160	160 (5x32)	512	$2^{64}-1$	80	192	<80	162	160-L(M)
SHA-224	224	256 (8x32)	512	$2^{64}-1$	64	139	112	224	min(224, 256-L(M))
SHA-256	256						128	256	256-L(M)
SHA-384	384	512 (8x64)	1024	$2^{128}-1$	80	154	192	384	384
SHA-512	512						256	512	512-L(M)
SHA-512/224	224						112	224	224
SHA-512/256	256						128	256	256
SHA3-224	224	1600 (5x5x64)	1152	Unlimited	24	-	112	224	224
SHA3-256	256		1088				128	256	256
SHA3-384	384		832				192	384	384
SHA3-512	512		576				256	512	512
SHAKE128	d		1344				min(d/2, 128)	≥min(d, 128)	min(d, 128)
SHAKE256	d		1088				min(d/2, 256)	≥min(d, 256)	min(d, 256)

Keccak, as hash function, provides 224, 256, 384 and 512 bit output sizes as well as SHA-2. Also, Keccak can be used as a stream cipher or pseudorandom bit generator, so it

supports variable output length and plays well with HMAC and KDFs. Both, SHA-2 and SHA-3, support N/2 bit collision resistance, N preimage resistance. SHA-3, SHA-384, SHA-512/224 and SHA-512/256 support N bit second preimage resistance, while SHA-256 and SHA-512 support N-L(M) second preimage resistance, where N is the output size in bits and *L(M)* is a function defined as $[log_2(len(M) / B)]$, with B - the block length of the function. Keccak is very hardware friendly and is better suited for embedded applications that are power or cost constrained, but is slower than SHA-2 in software area, it overall has a good performance, fairly high quality, in-depth analysis [31].

Many performance comparisons of the SHA-3 finalists and SHA-2 can be found at [20].

5.3 Security of Keccak

The SHA-3 hash functions were designed to resist collision, pre-image, second preimage or length-extension attacks, resistance which should be equal or exceeds the resistance that the corresponding SHA-2 functions provide [28]. More about the theory of sponge construction and their security properties can be found at [24], [30]. The security of Keccak has been thoroughly researched by a number of cryptanalysts [20].

Aumasson and Khovratovich provided two possible distinguishers on reduced-round Keccak-f[1600]. First, they detected non-ideal behavior in the algebraic description of the permutation applying cube-testers. Second, the authors tried to solve the constrained-input constrained-output (CICO) problem using automated algebraic techniques [32, 33].

Aumasson and Meier presented *zero-sum distinguishers*. This distinguisher was applied to the inner permutation of the hash function of Keccak and succeeded up to 16 rounds [33, 34].

Boura and Canteaut extended the *zero-sum distinguisher* of Aumasson and Meier to 18 rounds by analyzing the Walsh spectrum of the non-linear part and bounding the degree of the rounds more tightly [33], [35].

Boura and Canteaut extended their zero-sum distinguishers to 20 rounds [33], [36].

Morawiecki and Srebrny used SAT-solver techniques to find preimages for three rounds of Keccak, with 40 unknown message bits [33], [37].

Various research papers regarding the security analysis of Keccak can be found at [33].

6 Conclusions

A general overview regarding SHA family functions was presented. A brief design of SHA-2 and Keccak algorithms was described. Many cryptographic attacks against MD5 and SHA-0 were demonstrated by cryptanalysts and these functions were finally broken. Wang extended a theoretical attack against SHA-1. After these attacks, NIST opened a new competition for the next secure hash algorithm, named SHA-3. Also, NIST decided to simultaneously use two standards: SHA-2 and SHA-3. Five finalists in this competition are BLAKE, Grøstl, JH, Keccak and Skein. Keccak was announced as the winner. SHA-2 and Keccak are designed completely differently: SHA-2 uses MD construction and Davies-Meyer compression function, while Keccak is based on the Sponge construction. If an attack could work on SHA-2, the same attack would not work on SHA-3. Both functions support the same hash lengths and Keccak can be used as a stream cipher or pseudorandom bit generator because it supports variable output length. Also, Keccak has higher performance in hardware implementations than SHA-2. Various comparisons regarding construction, performance and security strength were summarized.

References

1. Announcing Request for Candidate Algorithm Nominations for a New Cryptographic Hash Algorithm (SHA-3) Family, Federal Register, Vol. 72, No. 212, 62212, http://csrc.nist.gov/groups/ST/hash/documents/FR_Notice_Nov07.pdf, Friday, November 2, 2007
2. Aumasson, J.P., Khovratovich, D., **First Analysis of Keccak**, comment on the NIST Hash Competition, https://131002.net/data/papers/AK09.pdf, 2009
3. Aumasson, J.P., Meier, W., **Zero-sum distinguishers for reduced Keccak-f and for the core functions of Luffa and Hamsi**, CHES 2009 rump session, https://131002.net/data/papers/AM09.pdf, 2009
4. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G., **Cryptographic sponge functions**, version 0.1, http://sponge.noekeon.org/CSF-0.1.pdf, January 2011
5. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G., **The KECCAK reference**, version 3.0, http://keccak.noekeon.org/Keccak-reference-3.0.pdf, January 2011
6. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G., **On the security of the keyed sponge construction**, Symmetric Key Encryption Workshop, http://sponge.noekeon.org/SpongeKeyed.pdf, February 2011
7. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G., **The Keccak sponge function family, Third-party cryptanalysis**, http://keccak.noekeon.org/third_party.html, 2012
8. Boer, B., Bosselaers, A., **An attack on the last two rounds of MD4**, Advances in Cryptology – Crypto'91, LNCS, vol. 576, Springer, Berlin Heidelberg, 1992, pp 194-203
9. Boura, C., Canteaut, A., **A zero-sum property for the Keccak-f permutation with 18 rounds**, comment on the NIST Hash Competition, https://www.rocq.inria.fr/secret/Anne.Canteaut/Publications/zero_sum.pdf, 2010
10. Boura, C., Canteaut, A., **Zero-Sum Distinguishers for Iterated Permutations and Application to Keccak-f and Hamsi-256**, Selected Areas in Cryptography, LNCS, vol. 6544, Springer, Berlin Heidelberg, 2011, pp 1-17
11. Burr, B.: National Institute of Standards and Technology (NIST), **SHA3 WHERE WE'VE BEEN WHERE WE'RE GOING**, http://csrc.nist.gov/groups/ST/hash/sha-3/documents/burr_dimacs2013_presentation.pdf, 2013
12. Chabaud, F., Joux, A., **Differential collisions in SHA-0**, Advances in Cryptology – Crypto'98, LNCS, vol. 1462, Springer, Berlin Heidelberg, 1998, pp 56-71
13. Damgård, I., **A Design Principle for Hash Functions**, Advances in Cryptology — CRYPTO' 89 Proceedings, LNCS, vol. 435, Springer, Berlin Heidelberg, 1990, pp 416-427.
14. De Cannière, C., Rechberger, C., **Finding SHA-1 Characteristics: General Results and Applications**, Advances in Cryptology – ASIACRYPT 2006, LNCS, vol. 4284, Springer, Berlin Heidelberg, 2006, pp 1-20.
15. Dobbertin, H., **Cryptanalysis of MD4**, Journal of Cryptology, LNCS, vol. 11, issue 4, Springer, Berlin Heidelberg, 1998, pp 253-271
16. Dobbertin, H., Bosselaers, A. and Preneel, B., **The hash function RIPEMD-160**, http://homes.esat.kuleuven.be/~bosselae/ripemd160.html, February 2012
17. Gauravaram, P., Millan, W., Dawson, E., Viswanathan, K., **Constructing Secure Hash Functions by Enhancing Merkle-Damgård Construction**, LNCS, vol. 4058, Springer, Berlin Heidelberg, 2006, pp 407-420

18. Grechnikov, E.A., **Collisions for 72-step and 73-step SHA-1: Improvements in the Method of Characteristics**, Cryptology ePrint Archive: Report 2010/413, 2010
19. Madhuravani, B., Murthy, D.S.R., **Cryptographic Hash Functions: SHA Family**, International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 2, issue 4, Springer, Berlin Heidelberg, 2013
20. Manuel, S., **Classification and generation of disturbance vectors for collision attacks against SHA-1**, Designs, Codes and Cryptography, vol. 59, issue 1-3, Springer, US, 2011, pp 247-263.
21. Matusiewicz, K., **Analysis of Modern Dedicated Cryptographic Hash Functions, PhD thesis,** Macquarie University, 2007.
22. McDonald, C., Hawkes, P., Pieprzyk, J., **Constructing Nonlinear Differentials in SHA-1**, Cryptology ePrint Archive: Report 2009/259, 2009
23. Mendel, F., Nad, T., Scherz, S., Schläffer, M., **Differential Attacks on Reduced RIPEMD-160**, LNCS, vol. 7483, Springer, Berlin Heidelberg, 2012, pp 23-38
24. Merkle, R., **One Way Hash Functions and DES**, Advances in Cryptology - CRYPTO' 89 Proceedings, LNCS, vol. 435, Springer, Berlin Heidelberg, 1990, pp 428-446.
25. Morawiecki, P., Srebrny, M., **A SAT-based preimage analysis of reduced Keccak hash functions**, Cryptology ePrint Archive: Report 2010/285, (2010)
26. National Institute of Standards and Technology (NIST), **FIPS 180-4**, http://csrc.nist.gov/publications/fips/fips180-4/fips-180-4.pdf, 2012.
27. National Institute of Standards and Technology (NIST), **Third Round Report of the SHA-3 Cryptographic Hash Algorithm Competition, NISTIR 7896**, http://nvlpubs.nist.gov/nistpubs/ir/2012/NIST.IR.7896.pdf, 2012
28. National Institute of Standards and Technology (NIST), **SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions, DRAFT FIPS PUB 202**, http://csrc.nist.gov/publications/drafts/fips-202/fips_202_draft.pdf, 2014
29. Paar, C., Pelzl, J., **SHA-3 and The Hash Function Keccak**, Springer, http://professor.unisinos.br/linds/teoinfo/Keccak.pdf, 2010
30. Rivest, R.L., **The MD4 message digest algorithm**, Advances in Cryptology - Crypto'90, Springer-Verlag, 1991, pp 303-311
31. Rivest, R.L., **The MD5 message-digest algorithm**, Request for Comments (RFC 1320), Internet Activities Board, Internet Privacy Task Force, 1992.
32. Stevens, M., **Counter-Cryptanalysis**, Advances in Cryptology – CRYPTO 2013, LNCS, vol. 8042, Springer, Berlin Heidelberg, 2013, pp 129-146.
33. Wang, G., Wang, S., **Preimage Attack on Hash Function RIPEMD**, LNCS, vol. 5451, Springer, Berlin Heidelberg, 2009, pp 274-284
34. Wang, X., Yu, H., **How to Break MD5 and Other Hash Functions**, Advances in Cryptology – EUROCRYPT 2005, LNCS, vol. 3494, Springer, Berlin Heidelberg, 2005, pp 19-35
35. Wang, X., Yin, Y.L., Yu, H., **Finding Collisions in the Full SHA-1**, Shoup, V.(ed.) CRYPTO 2005, LNCS, vol. 3621, Springer, Berlin Heidelberg, 2005, pp 17-36.
36. Wang, X., Yao, A., Yao, F., **Cryptanalysis of SHA-1**, Cryptographic Hash Workshop hosted, NIST, October 2005
37. http://en.wikipedia.org/wiki/SHA-3

AN OVERVIEW OF CAUSAL DIRECTED ACYCLIC GRAPHS FOR SOCIAL WORK RESEARCHERS[1]

Michael A. LEWIS[2,3]
Hunter College School of Social Work, New York, USA

E-mail: mle0050@hunter.cuny.edu

Abstract:

Given the mission of social work to improve people's lives by intervening in ways to enhance their well-being, social work researchers are interested in the causal effects of various types of social interventions (de Anda, 2007; Ohmer, and Jorr, 2006; Hawkins, 2006). Statisticians, econometricians, and other experts in quantitative methods tend to view randomized controlled trials (RCTs) as the "gold standard" when it comes to estimating causal effects, and there's a key reason for this--RCTs are based on randomly assigning units to at least two different intervention groups. Random assignment tends to result in the groups being balanced on variables that may have a causal relationship with the outcome of interest other than the intervention which is the focus of a given study. At least one of these "balanced variables" may also be causally related to the intervention itself. "Balanced" means that the average values of these other variables are equal across the different groups.

Key words: directed acyclic graphs, DAG, social work researchers, randomized controlled trials, RCT

Introduction

Variables that affect both the outcome and the intervention under study are called *confounding variables*. The fact that randomization tends to generate balance on potentially confounding variables means that if a researcher observes a difference (or differences) across intervention groups on the outcome of interest, they can be confident that this difference is due to the intervention being evaluated.

The problem many social work researchers face is that RCTs are often neither ethical nor feasible, yet these researchers are still interested in estimating causal effects. Perhaps unbeknownst to many social work researchers, computer scientists have been concerned with causality as well. This concern has come largely out of the area of *artificial intelligence* (AI). Some of the leaders in this field are UCLA electrical engineer/computer scientist Judea Pearl and Carnegie Mellon faculty members Peter Sprites, Clark Glymour, and Richard Scheines (SGS).

AI has focused on trying to program thinking in computers. One feature of thinking is thinking about causal relationships. Over the years Pearl (2000) and SGS (2000) have been working on the problem of how to represent thinking about such relationships in computers,

and they've developed mathematical tools for modeling such thinking. These mathematical tools are called causal Directed Acyclic Graphs (DAGS) and are the subject of this paper.

Causal DAGs aren't only useful to researchers in AI but are to quantitative researchers more generally who're interested in estimating causal effects using non-experimental or observational data. This is so for two reasons. First, causal DAGs provide a way of precisely, yet intuitively, specifying a researcher's causal assumptions. That is, they provide a language researchers can use to clearly state their assumptions about what is causing what. Such clarity allows other researchers to critically review those assumptions. Precise statements of assumptions and constructive critiques of them is a big part of the enterprise of science. Thus, causal DAGs can serve as an additional resource in a scientific approach to quantitative social work research.

Second, causal DAGs provide rules for addressing the problem of confounding when a researcher is faced with non-experimental or observational data. These rules can be drawn upon to guide specification of statistical models for use in social work research. For example, causal DAGs can provide guidance regarding which variables ought and ought not to be included in a *regression or propensity score matching model*, assuming that the causal assumptions encoded in a given DAG are true.

In order to make the fairly abstract ideas discussed in this paper more concrete, I'll refer to an example related to social work, more specifically social policy. I chose a social policy example because this is my substantive area of expertise. Hopefully, it'll be clear how causal DAGs can be used in other areas of social work research as well.

The Basic Income and Economic Security

As any social worker will know from either professional experience or coursework in social policy, the U.S., like all "advanced industrialized societies," is considered a welfare state. One of the hallmarks of welfare states is that, under certain conditions, they provide income to some of their residents that these residents don't have to work for. By "work" I mean sell one's labor in the formal labor market in return for a wage. For the purposes of this paper, I'll refer to such support as government provided non-wage (GNW) income.

Examples of GNW income programs in the U.S. are Temporary Assistance for Needy Families (TANF), Supplemental Security Income (SSI), Social Security, and Unemployment Insurance (UI). In social work we tend to focus on the degree to which these programs promote social justice or meet people's needs. Following this concern, I'll focus in this paper on what I'll call Economic Security, a variable meant to capture how economically or financially secure one feels.

A particular type of GNW a number of social scientists have written about (Widerquist, Lewis, and Pressman, 2005) is called a Basic Income (BI). BI is essentially an **unconditional** minimum income provided by government. That is, some level of government would grant all citizens or residents a minimum income without requiring recipients to engage in certain types of behavior in return for the benefit (e.g., enrolling in a workfare program, forming "proper" family structures, voting in elections, etc.). The policy appears to be based on the idea that people have a right to at least a subsistence level income, regardless of whatever else they might or might not be doing.

In this paper, I'll use a hypothetical study of the causal effect of BI on Economic Security to provide an overview of causal DAGs. Imagine that some political jurisdictions in

the U.S. chose to enact a law granting their residents a BI while other jurisdictions did not. Those who ended up receiving or not receiving a BI weren't randomly assigned to their "treatment" groups. Instead, normal processes operating in day to day politics led to these differences in treatment. Imagine that we end up with non-experimental data for a sample of residents from various jurisdictions. For each resident we have data on whether they received BI, their level on the Economic Security variable, and their levels on other variables to be defined below. We think BI is a cause of Economic Security and we'd like to estimate this causal effect.

What is meant by Cause?

Perhaps the most appropriate place to begin is with the concept of *cause*. In a paper co-authored with one of his students (Chen and Pearl, 2013), Pearl discusses cause in terms of probability distributions as well as expected values of outcomes. As those familiar with mathematical statistics know, the term "expected value" can be thought of as synonymous with "average". Since social work researchers often use regression models in their work and since such models typically focus on expected values of outcomes, I'll rely on the expected value based conception of cause.

According to Chen and Pearl (2013), x is a cause of y if intervening to change the value of x results in a change in the expected value of y. For example, suppose we take BI as x and Economic Security as y. Then BI would be a cause of Economic Security if intervening to change the value of BI resulted in a change in the expected value of Economic Security.

To be a bit more concrete, let BI take two possible values—$15,000 per year and $0 per year. Since no political jurisdiction in the U.S. currently has a BI program, let's take this to mean that the BI variable currently takes on the value $0 per year for everyone in the country.[4] Suppose currently the average level of Economic Security is 5 units. Now let's say some jurisdictions enact a BI grant of $15,000 per year to all their residents and, as a result, these residents' average Economic Security increases from 5 to 25 units. This would indicate that BI is a cause of Economic Security.

Note that the idea of intervening or doing something to change the value of some variable is crucial to the notion of causality. Also note that an intervention isn't necessarily tied up with an RCT, although it can accommodate RCTs. We could intervene by randomly assigning some people to receive a BI of $15,000 per year and others to receive one of $0 per year. Or, as in our hypothetical "natural experiment," laws could be passed stipulating that residents receive a BI of $15,000 per year. As long as the BI intervention changes the expected value of Economic Security, **by definition** we have a causal effect.

Identification versus Estimation

In order to understand the role causal DAGs might play in social work research, one must understand the distinction between *identification* and *estimation*. In Elwert (2013, p. 247 http://www.ssc.wisc.edu/soc/faculty/pages/docs/elwert/Elwert%202013.pdf) we find the following passage: "identification…determines whether and under what conditions, it is possible to strip an observed association of all its spurious components."

Social work researchers are no doubt familiar with the concept of correlation and its relationship to causality. Two variables, let's call them x and y, are correlated if x causes y, y causes x, or they share a common cause. What Elwert is getting at in the passage above is the

idea that identification is related to conditions under which one can isolate "causal correlation" from "non-causal correlation".

Estimation has to do with statistical methods, such as OLS regression, Two-Stage Least Squares regression, and others used to obtain estimates of causal effects. Identification "comes before" estimation in the sense that one needs to determine if causal association can be isolated from non-causal association before trying to estimate the magnitude of causal association.

As will become clear later in this paper, part of the utility of causal DAGs to social work researchers is that they provide clear rules for isolating causal from non-causal correlation, **assuming the causal relationships encoded in a given DAG actually hold true**. That is, causal DAGs tell us how to identify the causal effects encoded in a given DAG and, in that way, they provide guidance for the process of model specification and estimation.

Elements of Causal DAGs

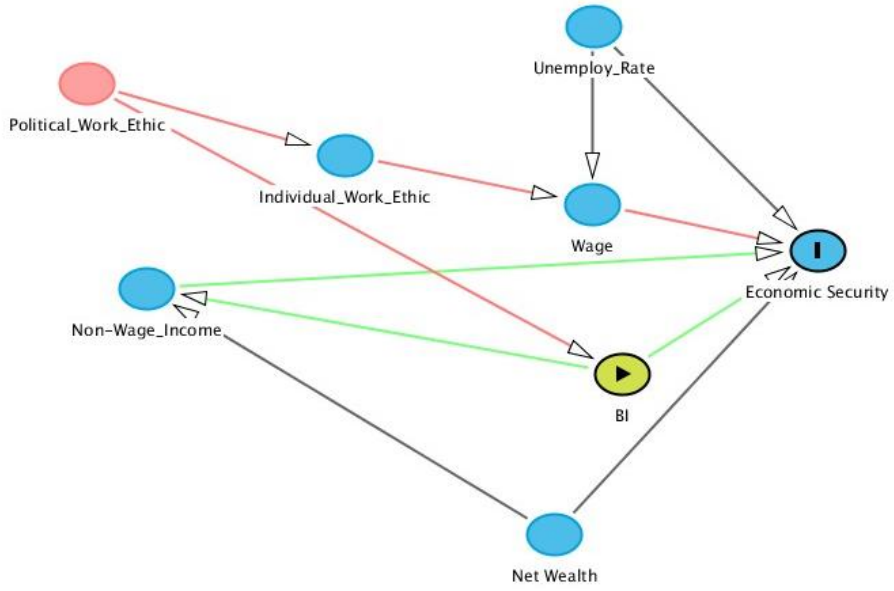


Figure 1. Causal DAG Representing Causes of Economic Security

Figure 1 will be used to explain the basic elements of causal DAGs. It was created using an open source online program for drawing causal graphs called Dagitty (Textor, Hardt, and Knüppel, 2011). Dagitty was explicitly written to implement the ideas of Pearl and others regarding DAGs. The color scheme in the diagram is done automatically in Dagitty as a way of labelling the role of certain variables in the diagram. Since knowing the details of that color scheme isn't necessary for our purposes, I'll have nothing to say about them.

I used Dagitty to draw Figure 1 on the basis of assumptions I made about the causal relationship between BI and Economic Security as well as how those two variables are causally related to a set of other variables. To do so I drew largely on economic theory and intuition. For example, on the basis of economic theory I'm assuming that the unemployment rate in an area where a person resides can influence the wage they receive for two reasons: 1) the unemployment rate can influence whether or not they receive a wage at all and 2) if they do receive a wage, their wage level can be affected by the unemployment rate because the more

people there are searching for jobs the less employers might be able to pay current workers (because higher unemployment might give employers an advantage over employees when it comes to wage bargaining).

Drawing a causal DAG is the first step involved in using such tools to aid social work research. That is, a researcher must draw a DAG which encodes the assumptions they're[5] making about what causes what. However, causal DAGs must be drawn according to certain rules.

The diagram in Figure 1 is an example of what mathematicians call a *graph*. In fact, the "G" in DAG stands for the word "graph". This is a different use of the term from what many readers might be accustomed to. That is, the term "graph" doesn't refer to a coordinate system along with a curve in that system. Instead it is a set of vertices or nodes along with edges or arrows connecting those vertices. The set of vertices in Figure 1 is {Unemploy Rate, Political Work Ethic, Individual Work Ethic, Wage, Economic Security, Non-Wage Income, Net Wealth, BI}. The vertices in a causal DAG represent variables. The edges or arrows in a causal DAG represent causal relationships.

Notice that there is an arrow "coming out of" BI and "going into" Economic Security. This encodes the assumption that BI causes or has a causal effect on Economic Security. In general, an arrow coming out of x and going into y means that x is assumed to cause y. The use of arrows to represent causal relationships is what makes these kinds of graphs causal graphs.

"Directed Acyclic" means that there can be no cycles or loops in a DAG. Perhaps the best way to understand this is to see a graph where there is a cycle:

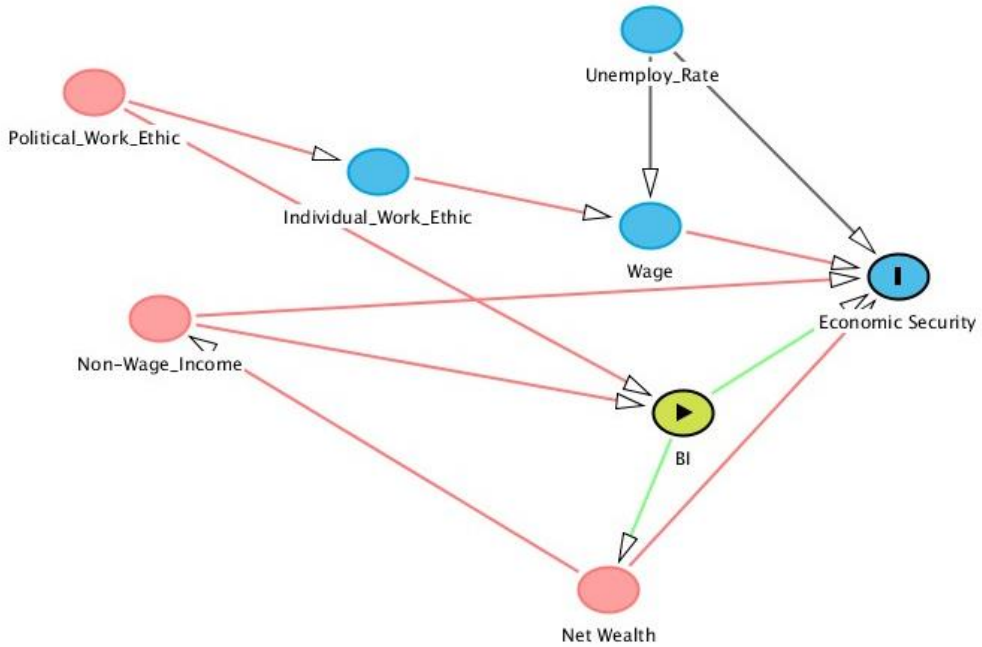


Figure 2. Graph Representing Causes of Economic Security with a Cycle

Notice that there is an arrow going from Non-Wage Income to BI from BI to Net Wealth and from Net Wealth back to Non-Wage Income. This is an example of a cycle because a "train" of causal connections circles back to where it started. Another example of a cycle

would be if there were an arrow going from BI to Economic Security and one going from Economic Security back to BI. Again, causal DAGs don't allow these cycles.

Readers familiar with structural equation models might be thinking that causal DAGs are nothing but structural equation models without latent variables. This is true. Structural equation models (as well as path models) can be thought of as special cases of causal DAGs.[6] However, structural equation models are based on specific assumptions regarding functional form, such as linearity, while causal DAGs are not (Greenland, Pearl, and Robins, 1999). For example, the assumption that BI causes Economic Security encoded in Figure 1 makes no assumption at all about the functional form of that causal relationship.

Here are some of the details regarding the variables in Figure 1:

VARIABLE	VALUES
BI	1 = $15,000/year 0 = 0$/year
Economic Security	Scale which measures how economically/financially secure one feels
Non-Wage Income	Weekly income in U.S. dollars from sources other than work
Individual Work Ethic	scale which measures the degree of effort expended at work[7]
Political Work Ethic	scale which measures elected officials'(within the study participant's jurisdiction) commitment to idea that able bodied should work for their subsistence[8]
Net Wealth	Total Assets – Total Liabilities measured in U.S. dollars
Unemployment Rate	Jurisdictional unemployment rate measured as a percentage
Wage	Weekly wage measured in U.S. dollars

Table 1. Values of Variables from Figure 1.

As I stated earlier, whenever an arrow comes out of one variable and goes into another, the variable the arrow comes out of is assumed to be a cause and the variable it goes into is assumed to be an effect. More specifically, this is a case of a *direct causal effect*.

Now let's take three variables x, y, and z, and suppose that an arrow comes out of x and goes into y. Also suppose that an arrow comes out of y and goes into z. In this case the causal DAG would be encoding the assumption that x is a direct cause of y, y is a direct cause of z, and x is an *indirect cause* of z. Readers familiar with the use of path analysis and structural equation modeling might recognize this as y being a "mediator" of the causal relationship between x and z. I should also add that the indirect causal effect of one variable on another can "work through" more than one intervening or mediating variable. The total causal effect of one variable on another is the sum of its direct and indirect causal effects on that variable.

Figure 1 also includes examples of indirect causal effects. Notice that Political Work Ethic is assumed to be an indirect cause of Economic Security. The idea is that if we could intervene to change the degree to which a set of elected officials in a given jurisdiction are committed to the idea that all able bodied people should work for their subsistence this would result in a change in the expected value of work effort expended by residents in that jurisdiction. This change in the average level of work effort would result in a change in the average wage, and this, in turn, would result in a change in the expected value of Economic Security.

Readers may not think much of the assumptions I've spelled out regarding the indirect causal effect of Political Work Ethic on Economic Security. But that's beside the point. The

point is that a causal DAG should encode whatever assumptions a researcher is making about causal relationships. These types of graphs provide a precise yet intuitive way of doing so with the benefit that others can easily see what a researcher is thinking regarding what causes what.

The next key concept relevant to drawing causal DAGs is *path*. A path in a causal DAG is a sequence of variables connected to each other by arrows.[9] A *directed path* between two variables is one where "travel" along arrows between the variables takes place so that travel is always from the tails to the heads of these arrows. Consider Figure 1 again. The path from Political Work Ethic to Individual Work Ethic to Wage to Economic Security is a directed path between Political Work Ethic and Economic Security.

An *undirected path* between two variables is one where travel along arrows takes place ignoring the direction of the arrows along the path. In Figure 1 the path from BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security is an example of an undirected path. The path from Political Work Ethic to Individual Work Ethic to Wage to Economic Security is an undirected path between Political Work Ethic and Economic Security. That is, "ignoring the direction of arrows" means that it doesn't matter which way arrows along the path are pointing. So a directed path, such as Political Work Ethic to Individual Work Ethic to Wage to Economic Security, is a special case of an undirected path.

Having defined directed and undirected path, I can state the no cycles constraint, discussed earlier, a little differently. Causal DAGs aren't allowed to have a directed path that ends with the variable it started with. The path from Non-Wage Income to BI to Net Wealth back to Non-Wage Income is an example of such a directed path. So it isn't allowed in a causal DAG.

Collider and *descendant* are two other concepts crucial to understanding causal DAGs. A collider is a variable which has two arrows coming from two different variables going into it. In Figure 1 Wage is a collider because arrows from both Individual Work Ethic and Unemploy Rate go into it. To understand the notion of descendant, assume that there is a directed path from x to y. There may or may not be variables between x and y. Since the path starts at x and ends at y, y is called a descendant of x. In Figure 1 Individual Work Ethic is a descendant of Political Work ethic and so is Wage.

Next I need to define the notion of *conditioning*. Other terms for conditioning are *subgroup analysis* and *stratification* (Morgan and Winship, 2007). Within the context of causal DAGs, conditioning occurs when an analyst examines the causal relationship between two variables **for given values** of at least one other variable, the variable which is being conditioned on.

For example, suppose in Figure 1 that Individual Work Ethic can only assume 10 possible values, the values 1-10 inclusive. We could estimate the causal relationship between Wage and Economic Security **only** for those with a level of 1 on the Individual Work Ethic variable, only for those with a level of 2 on it, only for those with a level of three on this variable, etc. This would be an example of conditioning on Individual Work Ethic.

It's also possible for two or more variables to be simultaneously conditioned on. For example, in the present case we might examine the causal relationship between BI and Economic Security for given values of both Individual Work Ethic and Political Work Ethic simultaneously.

Confounding, Regression Models, Propensity Score Matching, and Causal DAGS

In this section I'll spend a good deal of time discussing the identification of causal effects. I'll assume that we have a random sample of some population of interest and that each member of the sample has complete data on all the variables in Figure 1. This is to assume away all issues having to do with lack of a random or probability sample as well as issues having to do with missing data. I assume away these matters not because they are unimportant but because the purpose of this paper is to introduce readers to causal DAGs. Bringing in these other issues would needlessly complicate matters, given this purpose.

Regression models are ubiquitous in social work research. A typical example of the use of such models can be found in Studts, Stone, and Barber's paper *Predictors of Access to Health Care Services among Groups of TANF Recipients in Kentucky* (2014). They report findings from a set of logistic regression models involving predictors of access to health care. This use of the term "predictor" is ambiguous because it can be used when one is concerned with correlations between variables or causal relationships between them.

From about the end of November to the beginning/mid-December in the U.S., I suspect there's a strong correlation between the number of Christmas trees bought per day and the number of times the song Rudolf the Red Nosed Reindeer is played per day. So number of Christmas trees bought per day would be a good predictor of number of times Rudolf is played per day, but that wouldn't necessarily mean that the buying of Christmas trees causes the playing of Rudolf the Red Nosed Reindeer.

Social work researchers, including Studts, Stone, and Barber, are no doubt aware of this issue. That, in fact, is why they so often turn to regression models. Such models are thought to address the "correlation does not imply causation problem" by including independent variables other than the one of primary interest as "control variables." These control variables are thought to potentially confound the causal relationship between the independent variable and outcome of interest, and including them in a regression model is believed to take care of this problem. What causal DAGs can provide is guidance regarding the conditions under which including such control variables does take care of it, assuming the causal relationships encoded in a given causal DAG actually hold true.

In addition to regression models, social work researchers have made increasing use of a set of tools from statistics called propensity score matching (PSM) (Guo, Barth, and Gibbons, 2006). In fact, two of the field's most highly regarded quantitative methodologists, Shenyang Guo and Mark W. Fraser recently wrote a book (2010) on this topic. PSM is often used to deal with what econometricians call the *selection problem*.

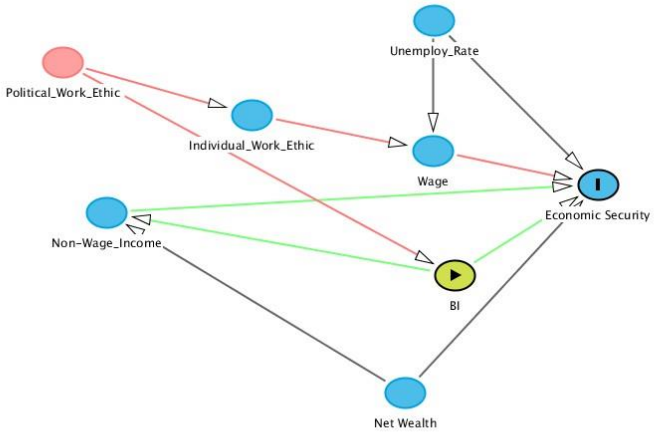
One of the problems researchers have when it comes to drawing causal conclusions on the basis of non-experimental data is that when the research participants are human beings they often self-select into treatment. This means that there could be at least one factor which causes both selection into treatment and a certain outcome. The result is that treatment and outcome variables will be correlated, and a researcher may mistake this correlation for a causal relationship.

As an example of the selection problem, suppose we're interested in the causal effect of Diet (magnitude of calories one consumes over a period) on Exercise (magnitude of exercise one engages in over a period). We might run into a selection problem because Depression (how depressed one is) may influence both Diet and Exercise.

PSM addresses the selection problem by using observed variables to model selection. That is, values on observed variables are used to estimate the probability that a case ends up "receiving" the intervention or treatment of interest. These estimated probabilities are called *propensity scores*. Cases are then matched on propensity scores and an analysis, often a type of regression model, is run in an effort to estimate the causal effect of some variable of interest. There're a number of ways propensity scores can be used for matching, and Guo and Fraser (2010) have a nice discussion of some of the different methods involved. A detailed discussion of these methods is beyond the scope of this paper. What is relevant is the fact that causal DAGs can provide guidance regarding what variables ought to be included in a propensity score model. In order to understand how causal DAGs can aid in the specification of regression and propensity score models, one first needs to understand the concepts *backdoor path*, *unblocked* or *open backdoor path*, *blocked* or *closed backdoor path*, *intercepting a path*, *confounding path*, and *confounders*.

A backdoor path from x to y is a path which has an arrow going into x. (Morgan and Winship, 2007; Greenland, Pearl, and Robins, 1999). A path is intercepted by a variable if that variable is on the path but isn't one of the variables on either end of the path (Greenland and Pearl, 2009). For example, the BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security path is intercepted by Political Work Ethic because Political Work Ethic is on the path but isn't on either end of the path. A path is intercepted by a set of variables if the members of that set intercept the path. So the BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security path is intercepted by the set {Political Work Ethic, Individual Work Ethic, Wage}.

Now suppose we have a backdoor path between x and y. This path would be blocked or closed if it has at least one collider and unblocked or open if it doesn't (Greenland, Pearl, and Robins, 2000). Take a look at Figure 1 again:



The path BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security is a backdoor path because BI and Economic Security are connected by a path that begins with an arrow going into BI. The path BI to Political Work Ethic to Individual Work Ethic to Wage to Unemploy Rate to Economic Security is also a backdoor path for the same reason.

The BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security path is unblocked because it doesn't contain a collider along it. The BI to Political Work Ethic to Individual Work Ethic to Wage to Unemploy Rate to Economic Security path is blocked because it contains a collider, namely Wage.

A confounding path is an unblocked or open backdoor path, and the variables which intercept that path are confounders (Greenland and Pearl, 2009). In the BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security path, members of the set {Political Work Ethic, Individual Work Ethic, Wage} are confounders.

Given the causal assumptions encoded in Figure 1 and the rules of causal DAGs, the existence of the confounding path referred to in the previous paragraph means that the relationship between BI and Economic Security is confounded by Political Work Ethic, Individual Work Ethic and Wage. So in order to identify the causal effect of BI on Economic Security, we need to do something about these confounders. This brings us to the notion of conditioning on a set of variables to block a confounding path.

Suppose **Z** is a set of confounders along a confounding path. Conditioning on the variables in Z blocks this path if:

1. any node along the path which has an arrow coming out of it is a member of Z or
2. the path has at least one collider which is not a member of Z and no descendant of any collider is a member of Z (Greenland and Pearl, 2009).

Conditioning on a set of variables to block an otherwise confounding path is the causal DAG version of controlling for confounding variables (or confounders) to identify a causal effect if interest. Thus, if the assumptions encoded in a particular causal DAG hold true, there's a set of variables Z which intercepts **all** confounding paths between x and y, and conditioning on Z blocks all backdoor paths between x and y, then the causal effect of x on y can be identified by conditioning on the elements of **Z.** (Pearl, 2000 and Greenland and Pearl, 2009).

Going back to Figure 1, notice again that the backdoor path BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security is intercepted by Political Work Ethic Individual Work Ethic and Wage. The backdoor path BI to Political Work Ethic to Individual Work Ethic to Wage to Unemploy Rate, to Economic Security is intercepted by Political Work Ethic, Individual Work Ethic, Wage, and Unemploy Rate. These are the only backdoor paths in the DAG in Figure 1.

The second backdoor path in Figure 1 can be blocked without conditioning on anything since Wage is a variable along that path and Wage is a collider. In this case we'd say that we're conditioning on the empty set Z = {} because the empty set is the one with no members. The second backdoor path can also be blocked by conditioning on Z = {Political Work Ethic, Individual Work Ethic, Unemploy Rate} since all the variables in this set have arrows coming out of them along that path.

The first backdoor path in Figure 1 can be blocked by conditioning on Z = {Political Work Ethic, Individual Work Ethic} since all the variables in this set have arrows coming out of them along that path. Notice that Political Work Ethic and Individual Work Ethic block both (all) paths between BI and Economic Security. Hence, the causal effect of BI on Economic Security can be identified by conditioning on Political Work Ethic and Individual Work Ethic.

To see how conditioning on a collider variable enters into all this, take a look at Figure 1 again. The path BI to Political Work Ethic to Individual Work Ethic to Wage to Unemploy Rate to Economic Security contains a collider variable, namely Wage. Suppose in an effort to identify the causal effect of BI on Economic Security, we conditioned **only** on Wage. That is suppose Z = {Wage}. Now take a look again at the second criterion for blocking a confounding path by conditioning on a set of variables[10]: the path has at least one collider which is **not** a member of Z and no descendant of any collider is a member of Z. Since the only member of Z is Wage and Wage is a collider, the BI to Political Work Ethic to Individual Work Ethic to Wage to Unemploy Rate to Economic Security is **unblocked** or **opened** by conditioning on Wage. That is, by conditioning on Wage we create a confounding path which is exactly what we don't want to do. Hence, if a researcher is interested in identifying a causal effect while addressing the problem of confounding, one shouldn't condition only on a collider variable.

I said earlier that conditioning on the variables Political Work Ethic and Individual Work Ethic would allow us to identify the causal effect of BI on Economic Security because these variables block all backdoor paths between BI and Economic Security. The set of two variables Z = {Political Work Ethic, Individual Work Ethic} is an example of what's called a *sufficient set*. This set of variables is sufficient in the sense that conditioning on them would be enough to identify the causal effect of BI on Economic Security. In general a set of variables is sufficient for identifying the causal effect of x on y if conditioning on those variables blocks all backdoor paths between x and y. A set of variables is *minimally sufficient* for identifying the causal effect of x on y if no proper subset of the set is sufficient (Greenland and Pearl, 2009 and Pearl, 2000).[11]

The set of variables {Political Work Ethic, Individual Work Ethic} is a sufficient set but it isn't minimally sufficient. To see this look again at the first criterion for blocking a backdoor path by conditioning on a set of variables along that path: any node along the path which has an arrow coming out of it. In Figure 1 the variable in set Z = {Political Work Ethic} has an arrow coming out of it along both backdoor paths. So conditioning on it identifies the causal effect of BI on Economic Security. The same would be true by conditioning on the variable in Z = {Individual Work Ethic}. The sets Z = {Political Work Ethic} and Z = {Individual Work Ethic} are both subsets of {Political Work Ethic, Individual Work Ethic} so {Political Work Ethic, Individual Work Ethic} isn't a minimally sufficient set.

The relevance of all this talk about conditioning on confounders to block backdoor paths is this: if we were willing to make the causal assumptions encoded in Figure 1 and willing to assume a linear (in parameters) functional form, then conditioning on a set of variables would amount to including them in an OLS regression model, along with the causal variable of interest (BI in this case), or including them in a model of propensity scores and subsequently running an appropriate statistical model on the basis of those scores. Let's take the OLS regression example first. Given the causal assumptions encoded in Figure 1, the effect of BI on Economic Security can be estimated by the following Ordinary Least Squares (OLS) regression model:

$$\text{Economic Security} = a + b_1\text{BI} + b_2\text{Political Work Ethic} + b_3\text{Individual work ethic}$$

The coefficient b_1 would be an estimate of the total causal effect of BI on Economic Security, controlling for the effects of Political Work Ethic and Individual Work Ethic on Economic Security. That is, b_1 would estimate how much the average level of Economic Security

would change **as a result of** an intervention to change BI by one unit (recall the "units" of BI are 1 = $15,000 per year and 0 = $0 per year). This is the sense in which causal DAGs can provide guidance regarding how to take care of the "correlation does not imply causation problem" through regression modeling.

Given what I said about sufficiency and minimal sufficiency, the following models could also be run to estimate the effect of BI on Economic Security:

$$\text{Economic Security} = a + b_1\text{BI} + b_2\text{Political Work Ethic}$$
$$\text{Economic Security} = a + b_1\text{BI} + b_2\text{Individual Work Ethic}$$

Earlier I discussed propensity score matching as a way of addressing the selection problem. What causal DAGs add to propensity score methodology is guidance on what variables should be included in a propensity score model. Those variables are the ones which appear in a given set **Z**—the ones which block all backdoor paths between a cause and effect variable of interest (Morgan and Winship, 2007). Thus, if one were interested in using propensity score matching to estimate the causal effect of BI on Economic Security, one could do the following.

First, use Political Work Ethic and/or Individual Work Ethic to model selection into treatment, BI in the present case. Second, use the saved propensity scores along with an appropriate method of matching, such as *nearest-neighbor*, *kernel*, etc., in a model to estimate the effect of BI on Economic Security. Both Guo and Fraser (2010) and Morgan and Winship (2007) discuss various matching methods for use with propensity scores.

OLS regression models and propensity score matching are often discussed as two different methodologies and in some respects they are. The insight that causal DAGs provide, however, is that in a sense they are equivalent ways of adjusting for or controlling for confounding variables.

Much of this section of the paper has been fairly theoretical and abstract. What I want to do now is discuss two things researchers can do if they're interested in applying these ideas. One relates to determining if confounding is present in a given causal DAG. The other has to do with finding a sufficient set of variables to condition on in order to control for confounding.

It's relatively easy to visually determine if confounding is present in a causal DAG if one recognizes the following fact: backdoor paths between x (cause of interest) and y (effect of interest) which contain a variable along them which is a common cause of both x and y are the "candidates" for confounding paths. Thus, one can use the following procedure to determine if confounding is present (Greenland, Pearl, and Robbins, 1999):

1. delete all arrows coming out of x
2. check to see whether the remaining graph contains variables which cause both x and y whether directly or indirectly
3. if there are common causes of x and y, then the backdoor paths going through those common causes are confounding paths (unless such a backdoor path goes through a collider or descendant of one) and so confounding is present; if there aren't such common causes of x and y, then confounding is absent

If this algorithm were applied to the DAG of Figure 1, we'd have to delete the arrow going from BI to Economic Security (step 1). We'd then find that Political Work Ethic is the only

common cause of BI and Economic Security (step 2). Thus, all back door paths going through Political Work Ethic would be candidates for confounding paths (step 3). The BI to Political Work Ethic to Individual Work Ethic to Wage to Economic Security is unblocked so it's a confounding path. The BI to Political Work Ethic to Individual Work Ethic to Wage to Unemploy Rate, to Economic Security is blocked by Wage since Wage is a collider. In order to address the problem of confounding we have to block all backdoor paths between BI and Economic Security by conditioning on a sufficient set.

To find a sufficient set of variables to condition on we could use the following procedure:

1. For each backdoor path collect the variables which intercept that path into a set.
2. Any such set which contains a collider or descendant of a collider blocks the relevant backdoor path.
3. Any variables in any such set which are non-colliders or non-descendants of colliders can be conditioned on the block the relevant path.
4. The sufficient set of variables is that set which blocks all backdoor paths between the variables of interest. A minimally sufficient set can be obtained by deleting variables from a sufficient set one at a time, thereby ending up with a series of subsets of the sufficient set until no other variables can be dropped without ending up with a set of variables which no longer blocks all backdoor paths between the cause and effect variables of interest.

If we followed the two procedures or algorithms spelled out above in regard to Figure 1, we end up with Z = {Political Work Ethic, Individual Work Ethic} as a sufficient set and with Z = {Political Work Ethic} and Z = {Individual Work Ethic} as minimally sufficient sets. These algorithms can be tedious to implement, however, without a computer. Dagitty, the program I used to draw the DAGs in this paper, can also be used to implement these two procedures. Thus, social work researchers interested in applying the ideas of causal DAGS would do well to acquaint themselves with Dagitty or some similar program.

Causal DAGs and RCTs

For most of this paper, I've used a policy example to illustrate the core ideas involved in causal DAGs. I've also focused on the use of DAGs in research with observational or non-experimental data. It should be said, however, that causal DAGs can be used by researchers working at a more "micro" level, and they can be used by those conducting RCTs. Here's an example to give readers a feel for what might be involved in such uses.

Suppose a social work researcher is interested in the effectiveness of a treatment designed to decrease trauma. More specifically, the researcher believes that traumatized persons exposed to this treatment will see a bigger reduction in their symptoms than such persons not so exposed. This researcher also believes that level of violent crime in a traumatized person's neighborhood and level of stress in a traumatized person's life also causally affect magnitude of traumatic symptoms. Further assumptions this researcher makes are 1) violent crime in a traumatized person's neighborhood causes treatment and 2) violent crime causes stress. By violent crime causing treatment I have in mind the idea that if we intervened to increase

the level of violent crime in traumatized persons' neighborhoods this would cause an increase in the average level of treatment for trauma that they chose to undergo.[12]

The causal assumptions spelled out in the previous paragraph are encoded in the following graph.

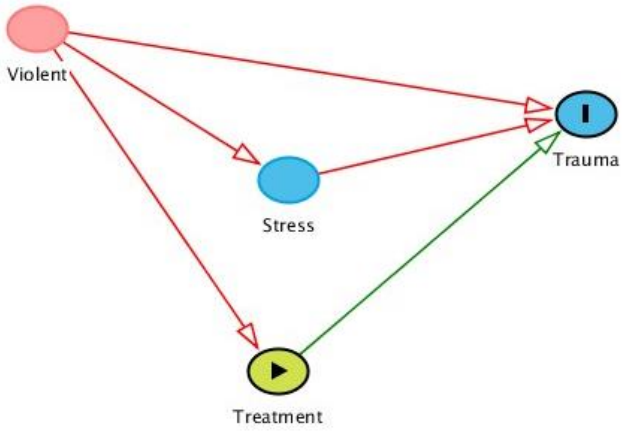


Figure 3 Causal Effect of Treatment on Trauma

Notice that there are two backdoor paths in this graph. One is Treatment to Violent to Trauma and the other is Treatment to Violent to Stress to Trauma.

We can see if there's confounding present in the graph by applying the algorithm discussed earlier and which is repeated here for convenience:

1. delete all arrows coming out of x
2. check to see whether the remaining graph contains variables which cause both x and y whether directly or indirectly
3. if there are common causes of x and y, then the backdoor paths going through those common causes are confounding paths (unless such a backdoor path goes through a collider or descendant of one) and so confounding is present; if there aren't such common causes of x and y, then confounding is absent

Here x is Treatment and y is Trauma, and the only arrow coming out of Treatment is the one going into Trauma. So, applying step 1, that arrow would be deleted. Applying step 2, we'd find that Violent is the only common cause of Treatment and Trauma. Since there are no colliders or descendants of colliders along the two backdoor paths and Violent is a variable on each of these paths, applying step 3 we find that both of those paths are confounding paths—the confounder, of course, is Violent.

If the researcher conducted an RCT and things went as planned, the connection between Violent and Treatment would be broken by randomization. That is, by randomly assigning people to either receiving the treatment or not, the researcher would end up with a graph like that in Figure 4:

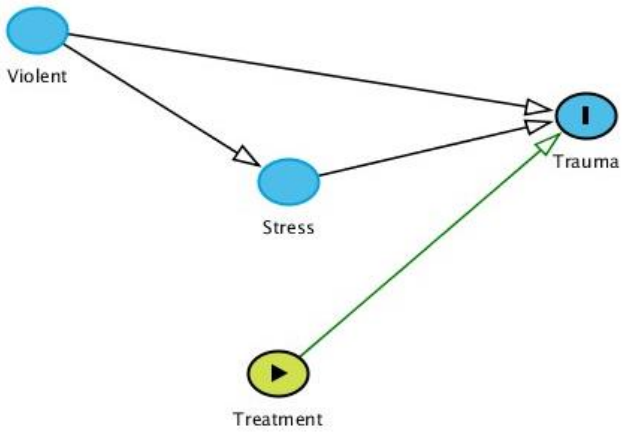


Figure 4

That is by way of random assignment Violent is no longer a cause of treatment. What this example shows is that when it comes to conducting RCTs, causal DAGs can help researchers think through the confounders they need to "watch out for". This, of course, is only true if the causal assumptions encoded in the DAG hold.

Another use of causal DAGs for RCT researchers can be seen by considering what can go wrong. Humans have a way of "not behaving themselves" as RCT participants. This misbehavior can occur in different ways but for the purposes of this paper let's just consider attrition. Attrition is when participants drop out of a study or are loss to follow up.

Suppose in an RCT to test the effectiveness of the treatment for trauma, "drop outs" from the group who received the treatment tend to be those who live in relatively safer neighborhoods while drop outs from the control group tend to be those from less safe neighborhoods. Thus, the group who received the treatment would tend to come from less safe neighborhoods while those in the control group would tend to come from more safe ones. This is another example of the selection problem, which was discussed earlier in the paper and we've already discussed how causal DAGs can help address it. Since Violent is a confounder, the treatment effect could be estimated by either a propensity score model which represented the probability of receiving treatment as a function of Violent before estimating this effect or a regression model of the effect of Treatment on Trauma, controlling for the effect of Violent.

RCTs are often talked about as if there's a sharp divide between experimental and non-experimental or observational data. In theory there is, but once humans start doing all they can to "mess up" pristinely designed RCTs, the line between experimental and observational data starts to look finer. Causal DAGs can help researchers think through how to address confounding in such situations.

Conclusion

I've said nothing in this paper so far about the relationship between hypothesis testing and causal DAGs. This is because causal DAGs have entered the social sciences more as a way of guiding how statistical models are set up in an effort to estimate causal effects than how to test hypotheses regarding such effects. At this point that is, causal DAGs have entered the social sciences more as guides to model specification. But in this closing section I'll make

some brief comments regarding what I see as the relationship between causal DAGs and hypothesis testing.

Social work researchers will often have ideas about which variables are causally related to others. These may come from theory, previous research findings, "common sense," or some combination of these. What causal DAGs provide is a precise language for visually representing these ideas. Yet they provide more than that—they also provide rules which stipulate the conditions under which causal effects can be identified, assuming the causal relationships encoded in a given graph are true. As we've seen from earlier sections of the paper, these rules about what needs to be done can be "translated" into ways of specifying statistical models to estimate such causal effects.

The role that hypothesis testing plays in all this is that it allows us to bring data to bear to determine if the assumed causal relationships encoded in a DAG mesh with available data. Rejection of the null hypothesis in this context suggests that a causal DAG is consistent with available data because causal relationships encoded in DAGs constitute alternative hypotheses (from the hypothesis testing perspective). Looked at this way what causal DAGs do for us is offer a precise way of determining what hypotheses we ought to test. And in that role they can be a potent addition to the tool set of quantitative social work researchers.

References

1. Chen, Bryant, Pearl, Judea, **"Regression and Causation: a Critical Examination of Six Econometrics Textbooks"** Real World Economics Review (65), 2013, p. 2-20.
2. de Anda, Diane, **"Intervention Research and Program Evaluation in the School Setting",** Children and Schools, vol. 29 (2), 2007, p. 107-121.
3. Elwert, F. **"Graphical Causal Models"** http://www.ssc.wisc.edu/soc/faculty/pages/docs/elwert/Elwert%202013.pdf, 2015
4. Greenland, Sander and Pearl, Judea, **"Causal Diagrams",** International Encyclopedia of Statistical Science, Lovic, M. (editor), 2011, p. 208-216.
5. Greenland, Sander, Pearl, Judea, Robins, James, **"Causal Diagrams for Epidemiologic Research"** http://www.hsph.harvard.edu/wp-content/uploads/sites/343/2013/03/causaldia.pdf, 1999
6. Guo, Shenyang and Fraser, Mark, **Propensity Score Analysis: Statistical Methods and Applications.** California: Sage Publications, Inc, 2010.
7. Guo, S., Barth R. P., Gibbons, C., **"Propensity Score Matching Strategies for Evaluating Substance Abuse Services for Child Welfare Clients",** Children and Youth Services Review, vol. 28 (4), 2006, p. 357-383.
8. Hawkins, J.D, **"Science, Social Work, Prevention: Finding the Intersections",** Social Work Research, vol. 30 (3), 2006, p. 137-152.
9. Morgan, Stephen L., Winship, Christopher. **Counterfactuals and Causal Inference: Methods and Principles for Social Research,** New York: Cambridge University Press, 2007.
10. Ohmer, M.L. and Korr, W.S, **"The Effectiveness of Community Practice Interventions: a Review of the Literature",** Research on Social Work Practice, vol. 16 (2), 2006, p. 132-145.
11. Pearl, Judea, **Causality: Models, Reasoning, and Inference,** New York: Cambridge University Press, 2000.

12. Spirtes, Peter, Glymour, Clark N., Scheines, Richard, **Causation, Prediction, and Search.** Massachusetts: MIT Press, 2000.
13. Studts, Christina, Stone, Romona, Barber, Gerald, **"Predictors of Access to health Care Services among Groups of TANF Recipients in Kentucky",** Social Service Review, 2006, p. 527-548.
14. Textor, J., Hardt, J. Knuppell, S., **"DAGitty: a Graphical Tool for Analyzing Causal Graphs",** Epidemiology, vol. 22 (5), 2011.
15. Widerquist, Karl, Lewis, Michael, Pressman Steve, **The Ethics and Economics of the Basic Income Guarantee**. England: Ashgate Press, 2005.

---

[1] **Acknowledgements**
I'd like to thank Sharon Schwartz, Dan Herman, Kristin Ferguson-Colvin, and Mark Fraser for their helpful comments. Judea Pearl was also helpful at clarifying issues in a series of email exchanges I had with him. I suspect the paper is better than it otherwise would have been as a result of those exchanges. Any errors or omissions are my responsibility alone.

[2] **Correspondence:** Michael A. Lewis, 129 East 79th Street, New York, NY 10075

[3]Michael Lewis holds a masters degree in social work From Columbia University and a doctorate in Sociology from the City University of New York Graduate Center. He teaches courses in social welfare policy and political economy at the Hunter College School of Social Work. Lewis' research interests are in poverty, social welfare policy, and quantitative methods. His work, some co-authored with Eri Noguchi, has appeared in a number of peer-review journals.

[4] Actually, Alaska does have a policy called the Permanent Fund Dividend which is arguably a basic income but, for the sake of discussion, let's ignore this reality.

[5] I'll use "they" and similar constructions, instead of "he," "she," etc., to be gender neutral.

[6] As long as we're only considering recursive path or structural equation models, since these are the ones consistent with the no cycles or loops constraint regarding causal DAGs.

[7] A 0 on the scale simply means that a person doesn't work at all (outside the home that is).

[8] The idea here is that the degree to which elected officials are committed to this view affects the behavior of individuals in those officials' jurisdiction. More concretely, if members of the legislature in a given locality have made public statements indicative of a high level of commitment to the view that able bodied persons should work for their subsistence this is assumed to cause individual residents of that jurisdiction to expend more effort at work.

[9] When I used the term "train" earlier I really had in mind the idea of a path.

[10] Even though I said "variables," the set in question, as is the case here, can have only one member.

[11] Suppose set $Z_1 = \{a, b, c\}$ and set $Z_2 = \{a, b\}$. Then Z_2 is a proper subset of Z_1 because every member of Z_2 is a member of Z_1 but Z_2 and Z_1 are not equal to or the same as one another.

[12] Intervening to increase crime in someone's neighborhood isn't likely to get support from a Human Subjects Review Board but that doesn't matter for present purposes. I'm really just going back to the definition of cause referred to earlier in the paper.

JOURNAL OF APPLIED QUANTITATIVE METHODS

Quantitative Methods Inquires

A SPREADSHEET SOLUTION TO H. E. DUDENEY'S PUZZLE "VISITING THE TOWNS"

Mike C. PATTERSON
Dillard College of Business, Midwestern State University, USA

E-mail: mike.patterson@mwsu.edu

Daniel FRIESEN
University of North Texas at Dallas, USA

E-mail:

Bob HARMEL
Dillard College of Business, Midwestern State University, USA

E-mail:

Abstract:

The purpose of this paper is to present an optimization model approach to one of Dudeney's puzzles, "Visiting the Towns," which appears as puzzle number 243 in Amusements in Mathematics (Dudeney, 1917). Henry Ernest Dudeney (1857-1930) was one of the most prolific mathematical puzzle creators of the past two hundred years. His first collection of puzzles, The Canterbury Puzzles and Other Curious Problems, was published in 1907. This popular book is still in print today (Dudeney, 1907). Beginning in 1917, he created five other popular puzzle collections, three of which were published posthumously. In addition to prodigious quantities of puzzles, his range of puzzle-types was exceedingly large, including geometry-based, logic puzzles , combinatorics" (Bremner, 2011), "cryptarithmetic" (Kilpelainen, 2012), cross-number puzzles (Sit, 1991), and chess problems (Nowlan, n.d.). During his time, there were no calculators; his mathematical puzzles were approached through pencil and paper calculations.

Key words: spreadsheet solution, H. E. Dudeney, puzzle, Visiting the Towns

Introduction

Dudeney had little in the way of formal education, a fact that launches Bremner (2011) into an analysis of Dudeney's methods for finding exact solutions. For instance, the puzzle of the Silver Cubes has as its solution (x,y) = (104940/40831, 11663/40831). Although his earliest efforts appeared in periodicals, including *The Weekly Dispatch, The Queen, Blighty, Cassell's Magazine* and *The Strand* (Henry Ernest Dudeney, n.d.), his column in *The Strand,* named "Perplexities," ran for over thirty years (Nowlan, n.d.). Many of these puzzles were published under the pseudonym "Sphinx."

A brief literature review follows this section. Next, the puzzle is given in its original format. A section describing the development of the model, along with a detailed description of the model proper is next. We conclude with a discussion and summary.

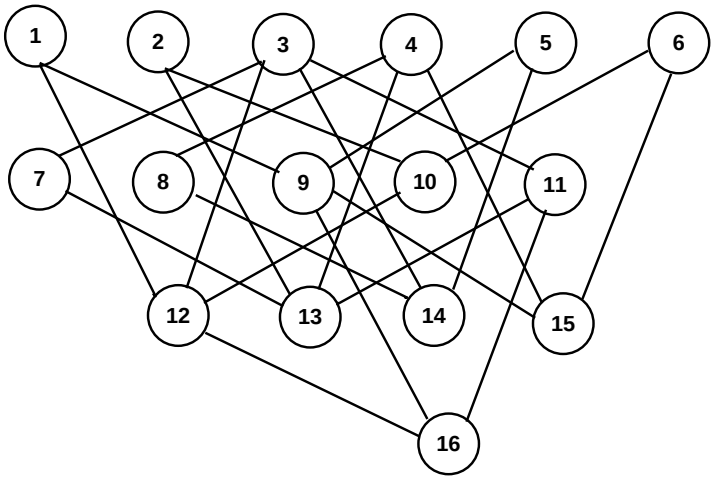
Literature Review

In addition to Dudeney, some of the better known writers in the field of recreational mathematics include Sam Loyd, Raymond Smullyan, Martin Gardner and Charles Lutwidge Dodgson, better known as Lewis Carrol, author of *Alice's Adventures in Wonderland*. "Recreational mathematics" is the term frequently used to describe mathematical games, puzzles and riddles. The best known peer-reviewed journal devoted to recreational mathematics is the *Journal of Recreational Mathematics* (Baywood Publishing, Inc., n.d.). Many journals give some passing attention to the subject, often through dedicated columns. For example, *Communications of the ACM* regularly publishes a column named "last byte" (Winkler, 2012). Alexander Dewdney (not to be confused with H. E. Dudeney!) wrote a "famous section" in *Scientific American* during the 1980s, as did Gardner for over twenty-four years prior. Dewdney's column was named "Computer Recreations" while Gardener's column was named "Mathematical Games" (Jimenez & Munoz, 2011). The column "Classroom Capsules" appears in *The College Mathematics Journal* and a brief survey shows that problems and puzzles often appear in the column for the purpose of providing "effective teaching strategies for college mathematics instruction (Alfaro, 2008)." As early as 1923, Carver suggested using Dudeney's puzzles as "stimulus" to undertake investigation, for both students and teachers!

In defining "recreational programming," Jimenez and Munoz (2011) refer to the practice as one of studying computer programming by solving problems of a playful nature. They describe the discipline as "similar to recreational mathematics." Demain (2010) uses the term "recreational computer science." Kino and Uno (2012) briefly discuss the incorporation of computers into the study of games and puzzles, and the reasons therefore. These authors modelled the game Tantrix using an interger programming formulation and solved it with an IBM software product. Kilpelainen (2012) uses problems from recreational mathematics, including some of Dudeney's other puzzles, to assess the features and utility of a new programming language. He notes that the process of formulating puzzle solution algorithms for computation invites programmers to consider problem generalization.

The Problem: Visiting the Towns (Dudeney, 1917)

243—VISITING THE TOWNS



"A traveller (sic), starting from town No. 1, wishes to visit every one of the towns once, and once only, going only by roads indicated by straight lines. How many different routes are there from which he can select? Of course, he must end his journey at No. 1, from which he started, and must take no notice of cross roads, but go straight from town to town. This is an absurdly easy puzzle, if you go the right way to work."

A Model of Visiting the Towns

Operations researchers will recognize the similarity between Dudeney's puzzle and the much researched Traveling Salesman Problem (TSP). Lawler, Lenstra, Kan, & Shmoys (1985) trace the origins of the TSP to the German publication *Der Handlungsreisende, wie er sein soll und was er zu thun hat, urn Aufträge zu erhalren und eines gluckichen Erfolgs in seinen Geschäften gewiss zu sein Von einem alten Comis-Voyageur* (Voight, 1831; Muller-Merbach, 1983). The Visiting the Towns puzzle is similar to the TSP in two ways. First, the tour must end in the same city where it begins. In this case the origin is the town labelled number "1." Second, each city must be visited only once.

The major difference between this problem and the TSP deals with the objective of the TSP, which is to minimize the total length of the journey (called a "tour"). Dudeney's problem does not identify the distance between the towns which are connected by roads. Thus, there is no objective to minimize the tour length. Instead, the objective is to designate a tour that begins in town "1," ends in town "1," and enters each town once. The general TSP does not constrain the solution space by specifying allowable roads as Dudeney's puzzle does (Press, Flannery, Teukolsky, & Vetterling, 1988).

Model Formulation of the Visiting the Towns Puzzle

The initial development of this spreadsheet model required the use of the Solver add-in tool for Excel. Solver is an optimization software tool developed by Frontline Systems and is available with Excel. During initial development, two characteristics of this problem presented particularly challenging issues. First, the model requires a constraint known as "alldifferent," which is particularly powerful in combinatorial problems, such as the TSP and, in this case, the Visiting the Towns puzzle. (Frontline Systems, n.d.). Secondly, Solver provides a genetic algorithm approach to optimization for non-smooth models and proved to be essential in the successful formulation of this model in order to arrive at an optimal solution.

The initial Solver formulation of the problem is shown in Table 1. Row 2 and column B, both highlighted, are used to identify each of the towns, labeled 1 to16. Cells C3:R18 identify those towns that are connected by a road, thus making a connection between those towns possible. A "1" indicates the presence of a connecting road. Likewise, a "0" indicates the absence of a road. For example cell D3, holding a "0," indicates that the towns 1 and 2 are not connected. Cell K1, with a value of "1" indicates a road connection between towns 1 and 9. Cells S2:S17, initially set to values 1-16, will hold the Excel-developed sequence when optimization is complete. Column U, labeled PATH, will be utilized to assure that connections are made between towns which have direct road connections. Cells C3:R18, which as previously discussed define the feasible routes between cities, is named MATRIX in Excel. While the

model does not deal with distances between towns, as would be the case in a TSP, model formulation requires an objective. The defined objective in the problem is to maximize the distance of the tour. In reality, what is accomplished is the assurance that only values of 1 are included in the tour, meaning that only feasible connections between towns are used. The requirement that each city be visited only once is addressed with the "alldifferent" constraint. The Solver parameters are displayed in Figure 2.

Table 1. Initial Spreadsheet Model Formulation Visiting the Towns Puzzle

1\A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
2		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	1	Path
3	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	2	0
4	2	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	3	0
5	3	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	4	0
6	4	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	5	0
7	5	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	6	0
8	6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	7	0
9	7	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	8	0
10	8	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	9	0
11	9	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	10	0
12	10	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	11	0
13	11	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	12	0
14	12	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	13	0
15	13	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	0	14	0
16	14	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	15	0
17	15	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	16	0
18	16	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	0
19																		Total	0
20																		Test	1

The formula view of the model is displayed in Table 2. The INDEX function returns the value in the cell at the intersection of a particular row and column. The genetic algorithm /evolutionary solver proposes a tour in column S. Since allowable connections (i.e. the roads) are worth 1 while unallowable connections are worth 0, Solver is programmed to increase the value of the resulting tour where the value is the sum of the connection values. This forces the solution to contain allowable connections since the objective is maximization. The "alldifferent" constraint forces the solution to contain only one visit to each town. The solution output from Solver is displayed in Table 3. As previously discussed, column S defines the suggested tour to travel from location 1 to 16, entering each once and only once. The sequence is 1-12-16-11-3-7-13-2-10-6-15-4-8-14-5-9-1. Figure 3 displays a flow diagram of the optimal tour. The astute observer will identify that the inverse tour (1-9-5-14-8-4-15-6-10-2-13-7-3-11-16-12-1) is also optimal. As Dudeney (1917) points out, this path (and its inverse) represent the only optimal solution to this puzzle. Solver required approximately 40 seconds to formulate the optimal answer with no adjustment to the default settings.

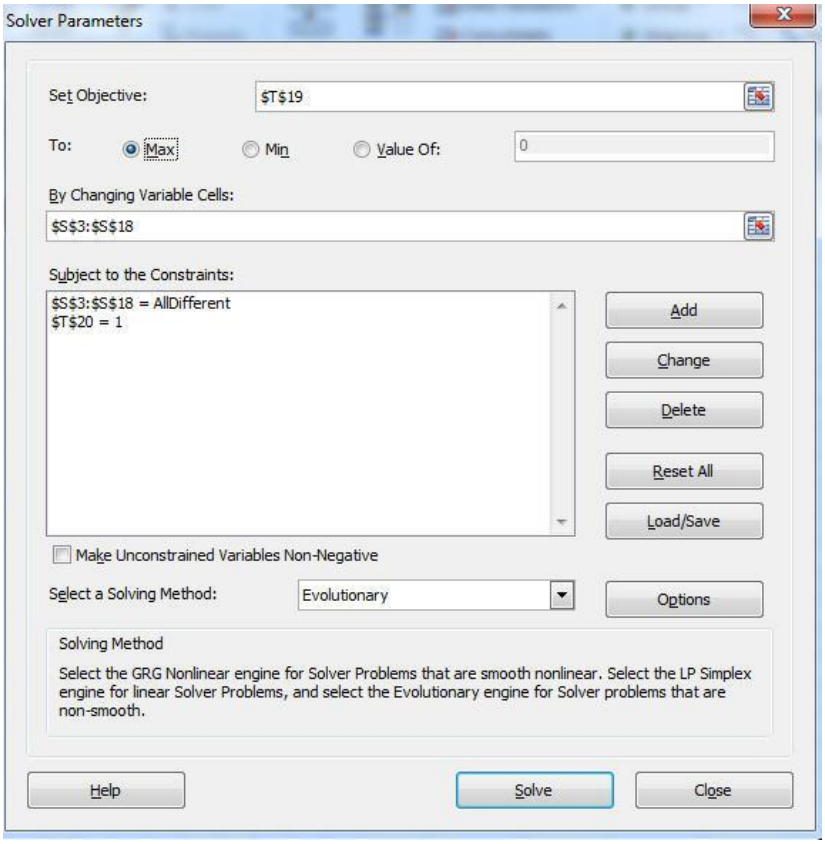


Figure 2. Solver Parameters Visiting the Towns Puzzle

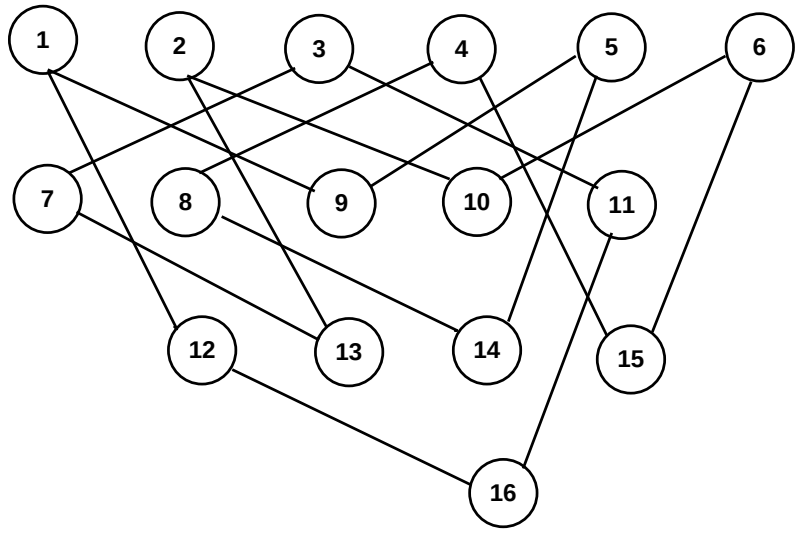
Table 2. Formula View Visiting the Towns Puzzle

1\A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
2		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	1	Path
3	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	2	=INDEX(MATRIX,S2,S3)
4	2	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	3	=INDEX(MATRIX,S3,S4)
5	3	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	4	=INDEX(MATRIX,S4,S5)
6	4	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	5	=INDEX(MATRIX,S5,S6)
7	5	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	6	=INDEX(MATRIX,S6,S7)
8	6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	7	=INDEX(MATRIX,S7,S8)
9	7	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	8	=INDEX(MATRIX,S8,S9)
10	8	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	9	=INDEX(MATRIX,S9,S10)
11	9	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	10	=INDEX(MATRIX,S10,S11)
12	10	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	11	=INDEX(MATRIX,S11,S12)
13	11	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	12	=INDEX(MATRIX,S12,S13)
14	12	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	13	=INDEX(MATRIX,S13,S14)
15	13	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	0	14	=INDEX(MATRIX,S14,S15)
16	14	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	15	=INDEX(MATRIX,S15,S16)
17	15	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	16	=INDEX(MATRIX,S16,S17)
18	16	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	=INDEX(MATRIX,S17,S18)
19																		Total	=SUM(T3:T18)
20																		Test	=S18

Table 3. Model Solution Visiting the Towns Puzzle

1\A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
2		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	1	Path
3	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	12	1
4	2	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	16	1
5	3	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	11	1
6	4	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	3	1
7	5	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	7	1
8	6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	13	1
9	7	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	2	1
10	8	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	10	1
11	9	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	6	1
12	10	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	15	1
13	11	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	4	1
14	12	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	8	1
15	13	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	0	14	1
16	14	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	5	1
17	15	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	9	1
18	16	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1
19																		Total	16
20																		Test	1

Figure 3. Visiting the Towns Solution



Discussion and Summary

Puzzles utilizing mathematics enjoy immense popularity. The phenomenon Sudoku is perhaps the most recent example, see for example (Friesen, Patterson, Harmel, 2010). H. E. Dudeney is one of best-known mathematical riddle developers of all time. In 1917 he published *Amusements in Mathematics*, which is a collection of 430 mathematical and logic puzzles. One of the puzzles titled "Visiting the Towns" present a challenging combinatorial problem. In this paper, we presented an optimization model approach to solving this puzzle.

Kilpelainen (2012) noted that researchers do not apply computing power to puzzles to take the fun out of them; rather, a possible purpose is to "develop instructive methods" for general classes of problems. This is certainly true here.

Finally, Dudeney (1917) makes two interesting remarks in his puzzle description. First he teases "How many different routes are there from which he can select?" TSP problems do not restrict the roads between towns; a TSP problem composed of 16 towns would have (16-1)!/2 which is approximately equal to 6.538E+11 (Malkevitch, 2005). Indeed, TSPs are known to be NP-complete (Press, Flannery, Teukolsky, & Vetterling, 1988). Of course, with the roads restricted, the number of possible tours is much reduced.

Dudeney's second interesting, possibly cryptic, remark likely speaks to the puzzle's complexity: "This is an absurdly easy puzzle, if you go the right way to work" (Dudeney 1917). Logically, if a town has only two roads, one must be incoming and one must be exiting. This directly removes the choice of road from towns 1, 2, 5, 6, 7, 8 and indirectly removes any choice of road from towns 9, 10, 13, and 14. Starting from town 1 and proceeding *right* to town 12, there is a choice. Taking the *right-most* path, we arrive at town 16. From town 16, there are no further choices. Armed with this solution algorithm, the exact solution is easily verified. This is a good quality for recreational computing exercises to possess.

References

1. Bremner, A. **Positively prodigious powers or how Dudeney done it?** Mathematics Magazine, Vol. 84, No. 2, 2011, pp. 120-125
2. Carver, W. B. **The mathematical puzzle as a stimulus to investigation,** The American Mathematical Monthly, Vol. 30, No. 3, 1923, pp. 132-135
3. Demaine, E. D. **Recreational computing,** American Scientist, Vol. 98, No. 6, 2010, pp. 452-456
4. Dudeney, H. E. **Amusements in mathematics,** London: Nelson, 1917
5. Dudeney, H. E. **The Canterbury puzzles and other curious problems,** London: Nelson, 1907
6. Friesen, D. D., Paterson, M. C. and Harmel, B. **A spreadsheet optimization model for solving sudoku problems,** Business Management Dynamics, Vol. 2, No. 9, 2013, pp. 15-22
7. Jimenez, B. C. R. and Munoz, R. R. **From recreational mathematics to recreational programming and back,** International Journal of Mathematical Education in Science and Technology, Vol. 42, No. 6, 2011, pp. 775-787
8. Kilpelainen, P. **Using XQuery for problem solving,** Software: Practice and Experience, Vol. 42, 2012, pp. 1433-1465
9. Kino, F. and Uno, Y. **An integer programming approach to solving Tantrix on fixed boards,** Algorithms, Vol. 5, 2012, pp. 158-175
10. Lawler, E., Lenstra, J., Kan, A. and Shmoys, D. **The traveling salesman problem,** Chichester: John Wiley & Sons, 1985
11. Malkevitch, J. **Sales and chips,** American Mathematical Society: Feature Column, 2005 September. Retrieved on January 12, 2013 from http://www.ams.org/samplings/ feature-column/fcarc-tsp
12. Nowlan, R. A. **Henry Ernest Dudeney,** In A chronicle of mathematical people. (n.d.). Retrieved on January 15, 2013 from Robertnowlan.com
13. Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. **Numerical recipes in C: The art of scientific computing,** Cambridge: Cambridge University Press, 1988

14. Voight, B. F. **Der Handlungsreisende, wie er sein soll und was er zu thun hat, urn Aufträge zu erhalren und eines gluckichen Erfolgs in seinen Geschäften gewiss zu sein Von einem alten Comis,** Voyageur. Ilmenau, Germany, 1981
15. Winkler, P. **Puzzled: Find the magic set,** Communications of the ACM, Vol. 55, No. 8, 2012, pp. 120
16. * * * Baywood Publishing Company. (n.d.) Journal of Recreational Mathematics. Retrieved on June 2, 2013 from http://www.baywood.com/journals/previewjournals.asp?id=0022-412X
17. * * * Frontline Systems. (n.d.). Retrieved December 21, 2012 from http://www.solver.com/ xlsplatform4.htm.
18. * * * **Henry Ernest Dudeney,** (n.d.). Downloaded on January 3, 2013 from https://en. wikipedia.org/wiki/Henry_Dudeney

INTELLECTUAL CAPITAL COMPONENTS AS DRIVERS OF ROMANIAN SMES PERFORMANCE

Bogdan-Vasile ILEANU[1]
PhD Assistant Professor
Bucharest University of Economic Studies, Department of Statistics and Econometrics

E-mail: bogdan.ileanu@csie.ase.ro

Alexandru ISAIC-MANIU[2]
PhD Professor
Bucharest University of Economic Studies

E-mail: alex.maniu@gmail.com

Claudiu HERTELIU[3]
PhD Professor
Bucharest University of Economic Studies

E-mail: claudiu.herteliu@gmail.com

Tudorel ANDREI[4]
PhD Professor
Bucharest University of Economic Studies

E-mail: andreitudorel@yahoo.com

Abstract:

The paper studies the possibility of evaluation of the intellectual capital in Science, Information and Technology domain from Romania. The analysis is made starting from a base sample of 1400 SMEs. Quantitative methods such as Structural Equation Modelling method are used to determine the group of latent factors used to evaluate the IC components of SME. Due to the particularities of Romanian SMEs and of the indicators used, some findings such as reference age of manager, the role of strong internal organization and structural capital are revealing for the first time some aspects relating the intangible assets and their potential resource for companies' performance. The results are similar to some particular studies made on SMEs from countries with economies in transition.

Key words: Structural Equation Modelling, SMEs, Intellectual Capital, Romania, Science, Technology, Human Capital

Introduction

The whole world is changing in all its forms: nature, human, economy, etc. As a result of the resource reduction there should be find other resources and to use more efficient than those we already have.

The notions of production have had to be revised. Each company has now a valuable but not entire known resource, the intellectual capital. This resource must be well managed

and exploited in order to succeed. In order to develop and benefit from Intellectual capital of a company, information must be allowed to circulate. Nowadays it is important that enterprises make access to knowledge easier for their employees, and they enrich the structure of the knowledge and share with partners.

One of the most important long-term goals of European Union is to become the most dynamic and competitive knowledge-based economy goals based also on Small and medium-sized enterprises which are well-known considered the engine of the economies.

It is also important to point that national (Feleaga, L. et al, 2012) and international studies (Storey, J. et al, 2009) claim that 75 to 90% of international market capitalization is attributable to intangible assets. Since the intellectual capital is already a major component of macro and micro economic growth and the economy based on knowledge gives rise to a new type of business, new workers and new professions why this model which have the core based on Intellectual capital(IC) and which function in many countries, shouldn't be followed by Romanian's companies too. To enhance the role of IC in company's management strategies, a brief description of the role of IC and its development is presented in the next paragraph.

I. Literature review

I.1. About intellectual capital

There are different intellectual capital definitions; most of them describe it as an integrated term for intangible factors of organisation work. For example, Annie Brooking defines intellectual capital as the "combined intangible assets which enable the company to function" (Brooking, 1996). A wider determination is given by N. Bontis— "the hidden values of individuals, enterprises, institutions, communities and regions that are the current and potential sources for wealth creation. These hidden values are the roots for nourishment and the cultivation of future wellbeing" (Bontis, 2004).

The intellectual capital can also be analysed as a mixture of human capital concept and knowledge management ideas: it integrates together knowledge and their parent — individual. Such vision underlines the intellectual capital model presented by Edvinsson and Malone (1997): according to them the intellectual capital is subdivided into the human capital ("the capabilities of the company's employees necessary to provide solutions to customers, to innovate and to renew") and structural capital ("includes the quality and reach of information technology systems, company images, databases, organizational concept and documentation"), and later includes the relational and organizational capital as specificity of external and internal interaction of the company employees.

I.2. Why measure intellectual capital?

Through systematic literature review we were able to identify the main reasons for measuring intellectual capital. These are: to help organizations to formulate their strategy; (2) to assess strategy execution; (3) to assist in diversification and expansion of decisions; (4) to use these as a complementary information and or measure for compensation; and (5) to communicate this key measures to external stakeholders. (Marr, B, 2003). Intellectual capital measure is also useful when the evaluation of company is made. The stakeholders are interested to invest or rather to buy the entire company if they know the values of the most important intangible asset. In many cases a hierarchy between departments, geographical subsidiaries or even between companies or countries is needed. As many stakeholders want to

evaluate the IC, a lot of methods where developed to achieve this goal. At macro-level we think that the achievement of a National Intellectual Capital Index (NICI), as Bontis (2004) named it, will be also helpful to complete relevance of Human Development Index and to establish and maybe control the corruption and shadow economy.

Due to its major importance a lot of methods were developed during years. In the following paragraph we mention the most cited methods in the related literature.

I.3. Brief review of Principal Methods used to measure Intellectual Capital

The methods of measuring intellectual capital started from the micro level analysis for more than 30 years. The most important methods found in literature in the last 25 years, considered also as reference methods, are: The invisible Balance Sheet by Sveiby developpend in 1989, Balance Score Card developped by Kaplan and Norton in 1992, Intangible Asset Monitor, by Sveiby in 1994, Skandia Navigator, by Edvinsson and Malone, 1997, IC-Index, Roos, Roos, Dragonetti & Edvinsson in 1997, Intellectus model, by Sanchez in 2002, EVVICAE by McCutcheon, 2008. Starting from these basic methods many case studies for different regions and countries or for different domains of economic activities can be found.

During the time there were developed also different methods of regional or national intellectual capital measurement. Among them the most well-known are the National Intellectual Capital Index by Nick Bontis, based on Edvinsson's model, Intellectual Capital of Nations by Amidon D.M., and the Intellectual Capital Monitor by Andriessen and Stam which integrates the classic intellectual capital model from "The Intangible Assets Monitor" by Sveiby (1997).

An important regional approach was developed by Schiuma, Lerro and Carlucci in 2008. This method, named RICI, takes into account four perspectives: hardware, Netware, wetware, software.

In the particular case of enterprises, few important research initiatives rather methodologies of measure than methods of measure in the true sense of the word are: 'Wissensbilanz", made in Germany (Edvinsson, L., Kivikas, M., 2007), Danish Guidelines or MERITUM project. These are the pioneers in establishing a set of steps and indicators which should be followed to measure IC of a firm.

Starting of these ideas we try to propose an adapted model for measuring the IC of Romanian SMEs. Before presenting our studies we are briefing few case-studies of SMEs-IC measure met in the international literature.

II. Methodology and Research Study

II.1. Proposed model of study

As it is normal, in every study there are some particularities regarding the evidence or relevance of a particular component of the intellectual capital asset, but during the years many researchers agreed (Stewart, 1999, Sveiby & Buck, 2001, Bontis, 2004, Mertins, K. & all, 2009, Martin, C. et all, 2011, etc.) that the Intellectual Capital of a company could be expressed as a sum of multiple components such as: Human Capital, Organizational Capital (Internal Component) and Relational Capital (External component). These components combined with the financial power of the company are the drivers of the company's performance and of the economy in general.

In Romania, in the case of Science, Technology and Informatics domains, we consider that the hard core of Intellectual Capital is given by human and organizational components, as we previous showed (Ileanu & all, 2011). Similar results, where relational capital has a lower influence on the IC formation than other components, were obtained by Halim, S. (2010) in a study made on the German and few other European countries' SMEs, or by Martin de Castro and Lopez Saez (2008) in a study made on 49 Spanish companies from IT domain, etc.

Since the relational component has a lower influence as some studies showed and we were not able to measure indicators for this component of IC we tested for now the relevance of some indicators and their effects on the main components of IC which could be indirect measured. Before the presentation of the model we briefly present in the next paragraph the data source of the research

II.2. Sample and data preparation

The applied research is done on a representative sample of 1400 small and medium enterprises (SME). Data collection was made during 2010, using face-to-face method by fieldwork interviewers. All the respondents were specialists in HR. The sample is cross stratified by number of employees, NCAE (domain activity) and regions. Economic theory and practice recommend the analysis within a single domain because of the large heterogeneity of the factors and results. As a result, we took from the base sample only the companies from the Scientific, Information and Technical Activities (SCIT) domains, considering that these are the most appropriate/relevant domains in measuring the intellectual capital. After the data collection, we have cleaned the database by eliminating the cases with many multiple missing answers. In this case the final sample is formed by 162 SMEs from the SCIT domains. The structure of the sample after the data cleaning is the following: 70% micro-enterprises(less than 10 employees) 20% small enterprises (10-49 employees), 10% medium enterprises (over 49 employees).

Taking into account that methodology we need a large number of observations for the SEM, we used two methods for data imputations instead of deleting cases in those cases where there were not multiple variables with missing values on the same case. For the continuous variables we used the mean value method of imputation and the mode value for imputation method for the categorical variables.

II.3. Description of the variables

The variables used in our study are lying on different ranges, are becoming from different patterns, have distinct units of measure and are measured on different scales. For these reasons they were standardized. The classical method of standardization was applied in this case: $z = \frac{X - Mean(X)}{Standard\ deviation\ (X)} \sim N(0,1)$, where X= the initial variable and Z the transformed(standardized) variable.

In order to make the analysis clearer we kept the letter "z" in front of the variable to reveal that the used variable is standardized.

zangtr represent the percentage of employees in the company which benefit of training during the last year. Its values are lying between zero and one hundred percent. **zwage** is a categorical variable and represent the changes on the employees' salary during the last year comparative to the previous one. It is measured on a scale of nine points. **zperf** represent the performance of the enterprise in the last year compared to the previous year, measured on a scale of six points describing the performances from very low to very high; **zinvinov** is measuring the percentage of the total investment of the enterprise for developing new products or services, this variable

is measured on a scale with six points and has values from zero to over seventy five percent; z**angexp,** defines the percentage of employees which have more than fifteen years of experience in the principal domain of activity of the enterprise. It has values from 0 to 100% unconstrained by predefined classes; **ztrday**, counts the average number of days of training of employees during the last year. This variable is measured on a scale of four points, it cannot be assumed as continuous; **zangsex2**, represents the square of the variable percentage of men employees; z**angedu**, represent the percentage of employees which have high level of education; **zcanewp** is a continuous variable which measures the percentage from turnover gained from new products developed in the last year; **zvrstman2** measures the square of age of the manager, it is assumed as continuous variable and it has a mean equal with 40. **Zcatr** represents a continuous variable calculated as a report between the amount of money allocated for employees training and the turnover of the company. As endogenous result variable we took a six-scaled variable named **zperf**, which measures the performance of an analysed company.

II.4 Description of the quantitative method used in analysis.

Many quantitative methods were used to evaluate the IC of a company. The most frequent methods used in this domain are: Data Envelopment Analysis (Campisi, D, 2008, Matei, M, 2010, Yang, C.,2010) Principal Component Analysis (Ileanu,B., 2010, Yang, C, 2010), Structural Equation Models (SEM) with latent variables (Bontis, 2004, Martinez-Torres,M.R., 2006, Carrington, D. and Tayles, M, 2011, etc). We decided to use SEM because is the most appropriate method in our case. Some of the reasons are:

SEM is a better method because it can show the relation between variables as they are in the reality.

SEM approach doesn't alter the practical significance of the variable as PCA does it. DEA require in general relevant inputs or outputs and searches an efficient frontier and efficient companies. In our case it is possible that the companies are not efficient even if they invest in Intellectual Capital.

SEM is defined as multiple equations model in which the response variable in one regression equation can appear as an explanatory variable in other equation SEM has the ability to test the specified models by economic theory including unobservable (latent) variables. The latent variables can be measured indirectly through their effects (indicators), or sometimes through their observable causes.

SEM can be modelled using two components: structural equation model and the measurement model.

The structural Equation model:

This model describes the relations between the exogenous and endogenous latent variables. A formal representation is the following:

$\eta_y = a + B\eta_y + \Gamma\eta_x + \varepsilon$, where η_y is vector of endogenous latent variables (given by IC-variable) in our case), η_x is the vector of exogenous latent variables and ε the vector of residual components, assumed as white noises.

The measurement model describes the relations between the exogenous and endogenous observed or latent variables. A formal representation could be:

$$y = \alpha + L_y\eta_y + w$$
$$x = \beta + L_x\eta_x + u$$

, where y is a vector of endogenous observed variables, and x vector of exogenous observed variables, *u* and *w* are error vectors having the same properties like ε

Estimation.

According to Schumacker R.E., and Lomax G.R (2010, p.217) when the observed variables are from small or medium samples, normal distributed, with no many missing values, Maximum Likelihood(ML) and Generalized Least Square(GLS) are recommended. In case of large samples with non-normality Weighted Least Square (WLS) is recommended. According to Kline (2011, p.176) in cases of severely non-normality the GLS as part of WLS family methods can be used.

In order to apply the most suitable method for parameters estimation we have verified the normality distribution of the variables using Jarque–Bera test (Jarque, C.M, Bera, A.K, 1980). Principal descriptive characteristics of the variables are present below in the Table 1.

Table 1. Descriptive characteristics of the sample variables

	ZINVINOV	ZVRSTMAN	ZANGTR	ZCATR	ZTRDAY	ZWAGE	ZANGEDU	ZANGEXP	ZCANEWP
Mean	0	0	0	0	0	0	0	0	0
Median	-0.27011	-0.06818	-0.64238	-0.45085	-0.81549	-0.70711	0.56017	0.01194	-0.32673
Maximum	2.19508	2.25217	1.62594	5.61669	2.07741	2.47487	0.828100	3.15430	1.99988
Minimum	-0.88640	-1.83164	-0.89442	-0.52247	-0.81549	-0.70710	-1.851200	-0.87437	-0.90838
Std. Dev.	1	1	1	1	1	1	1	1	1
Skewness	0.82844	0.27045	0.75887	3.23927	0.94836	1.06229	-0.722280	1.71194	0.67474
Kurtosis	2.42056	2.07226	1.83010	14.3990	2.62066	2.56200	1.868188	5.84658	2.07261
Jarque-Bera (JB)	20.7970	7.78470	24.7876	1160.38	25.2547	31.7636	22.73233	133.825	18.09771
Probability (JB)	**0.00003**	**0.02039**	**0.00000**	**0.00000**	**0.00000**	**0.00000**	**0.000012**	**0.000000**	**0.000118**
Observations	162	162	162	162	162	162	162	162	162

In the present case, the most appropriate estimation method for this system is the GLS method, taking into account that we have a medium sized sample, some continuous and ordinal but assumed as continuous variables and non-normal distributed (Probability of JB-Test is less than critical level of significance).

The estimation using GLS implies the minimization of the discrepancy function $G(\theta) = G(S;\Sigma)$ where S =the sample covariance of the observed variables and Σ is the theoretical covariance; The general form of the minimization function is:

$$G(\theta) = \left(s - f(\theta)^t\right) W (s - f(\theta))$$

where, **s** = vector containing the variances and co-variances of the observed variables**;**

$f(\theta)$ = a specified function of an unknown parameter, in most of the cases this function is linear

W = weight matrix which corresponds in the above function to the estimation method chosen. **W** is chosen to minimize **G**, and G(N-1) gives the fitting function, in most cases a χ^2 distributed statistic. The performance of the χ^2 is affected by sample size, error distribution,

factor distribution, and the assumption that factors and errors are independent (Ullman, 2003).

The function G in GLS case is equivalent with $G_{GLS} = \frac{1}{2} tr\{[S - \sum(\theta)]W^{-1}\}^2$. According to Bollen (1989) in most of the cases, $W = S^{-1}$ is a matrix which determines consistent estimators.

In order to achieve significant parameters and a best fit we imposed some constraints taking into account some findings met in the international literature.

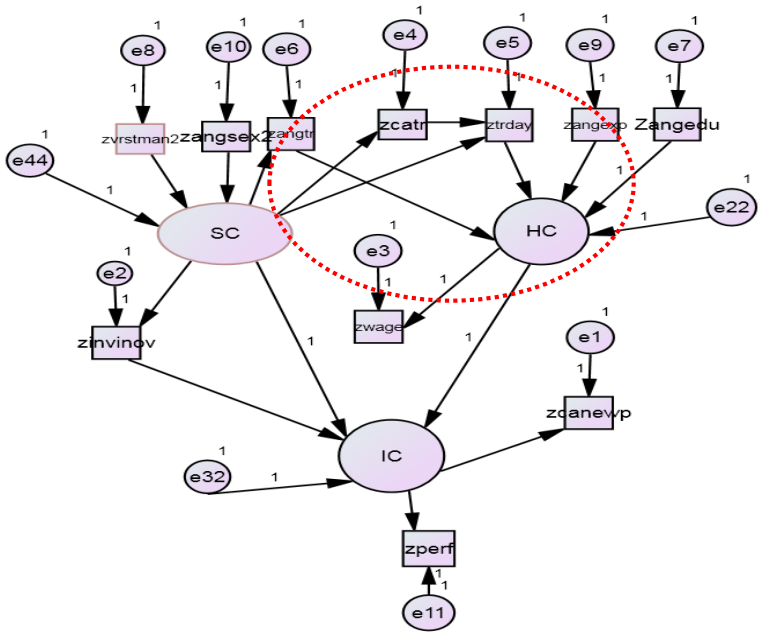
For example we imposed that IC is the simple sum of HC and SC. We also imposed that education strongly influence human capital and finally we imposed the fact that human capital increase is seen in the increase in wages, hypothesis empirical tested before by Ileanu and Tanasoiu (2008).

The results of estimation, made with specific software (AMOS) are presented in the next paragraph.

III. Results

Using the SPSS module AMOS we have made the graphical representation of the relations between the variables as the management theory and previous empirical studies sustain its. This representation could be seen in the figure presented below. The variables e_1, e_2, …, e_{44} represent standardized normal white noises.

Figure 1 The Model of IC analysis



After the estimation using SEM in AMOS we achieved the values of the standardized coefficients of the models included in the system.

Most of the coefficients are significant at levels around 0.05 as it can be seen in the table below.

Table 2. Estimated values and significance tests of the SEM model

			Estimate	α(risk)	Standardized coefficients
SC	<---	zvrstman2	-,570	,014	-,487
SC	<---	zangsex2	-,217	,242	-,185
zcatr	<---	SC	,191	,050	,218
zangtr	<---	SC	,311	,003	,343
ztrday	<---	SC	,298	,003	,316
ztrday	<---	zcatr	,236	***	,219
HC	<---	zangexp	,138	,042	,097
HC	<---	zangtr	,031	,589	,023
HC	<---	ztrday	,150	,004	,115
HC	<---	Zangedu	1,000		,699
zinvinov	<---	SC	,283	,004	,315
IC	<---	SC	1,000		,441
IC	<---	HC	1,000		,539
IC	<---	zinvinov	1,161	***	,461
zperf	<---	IC	,016	,042	,042
zcanewp	<---	IC	,207	***	,481
zwage	<---	HC	1,000		

*Note: *** estimated coefficient is significant at level lower than 1%*

IV. Findings

Most of the coefficients are positive and are showing the direct contribution of the factors on the growth of the enterprise's Intellectual Capital. The human component is based on training investment and the effects of investment in trainings are seen by the increase of training days number.

An increase by 1 % of the percentage allocated for employees training will have positive effects (increase by 0,236 %) in the average number of days of training. An increase of the average number of days of training will generate an increase of the human capital. Here can also be observed the positive and indirect effect given by the percentage of turnover (zcatr) allocated for training on the human component.

Human capital in the case of Romanian SMEs from SCIT domain is based on education, training allocation funds and effort of training. In these evidences there are some hidden factors which contribute to the increase of HC even if they are not represented in the current model. Some examples could be: the capacity of trainers, the abilities of trainers, and the structure by age of trained persons, and job satisfaction for about some studies are showing that motivator factors are influencing employee level of satisfaction, and then the employee

job satisfaction contributes to employee engagement (Akinbobola, 2011, Brown,1996) and finally SC and HC development.

The red circle marks the connection between SC component and factors of HC. In the vision of Bontis (2001) SC "deals with the mechanisms and structures of the organization that can help support employees in their quest for optimum intellectual performance and therefore overall business performance. An individual can have a high level of intellect, but if the organization has poor systems and procedures by which to track his or her actions, the overall intellectual capital will not reach its fullest potential". In conclusion, good strategies and good abilities of the decisional factors will determine efficient allocation of funds and the effects will be seen in the increase of HC, IC and finally on the increase of company performances, outputs.

The share of trained employees is not significant (the risk=0,589>0,05). This thing takes into account that the SCIT activities in Romania are relatively recent formed. Most of employees are young and above average skilled due to the nature of activity and its requirements. Also a larger offer than demand of labor force in this domain is influencing the structure and qualification of the employees in the current domains. Most of them are working in direct production so the percentage of trained person should not count. In order to have a significant influence, this problem should be treated using a binary variable for a larger group. For example could be obtained different results if this variable is categorized as follows: "trained most of them or all vs. and only few or non-of them are trained".

The structural capital is influenced by employees distribution by gender and by the age=experience of the manager. Since the model is estimated using standardized variables and the relationships between SC and manager's age and between SC and the structure by gender are parabolic with negative quotients, the maximum of SC is reached when the initial variables are equal with their sample means. In this case the sample mean is given by the 56.6% of male proportion. Gender structure, used as human capital component was also found as significant in other studies, eg. Li J. et al (2008).

According to descriptive statistics presented in the **Table** 1 combined with the regression results from Table 2, the efficiency of the SC is achieved when the mean age is around 40, and when the distribution by gender is equilibrated (around 50%-50%). These results are normal are proving that the stability and efficiency is given by equilibrium and experience. Since we deal with SMEs, most of them being close to micro-enterprises, and where the highest importance on the company future is given by the manager' decision, the age of 40 is the most appropriate to increase the organizational degree and to improve in the end the HC and Intellectual capital.

The results of this study sustain the hypothesis regarding the theory that the performance of enterprises is directly influenced by IC. In the Table 2, the coefficient which relates IC and companies performances is positive and its standardized value shows that for an increase of IC by 1% the enterprise performance will rise by 0,04%. This weak association reveals also that the increase of IC is not time synchronized with performance. The increase of company's performance could be seen in few years. Another element which should be noted is that allocated funds for innovation have a strong influence on the intellectual capital development. As the results show it an increase of the funds for innovation and research by 1% will determine an increase of the IC by 0, 46%.

Conclusions and future researches

Providing the importance and contributions of intangible assets on the companies' performances, using local empirical studies and many international similar examples, this study gives strengths and trust among the SME leaders in understanding, discovering and driving the intangible assets that are important value for their own businesses. Since active SMEs represent a driver in sustainable regional development and in particular for development poles (Davidescu & Strat, 2014), perhaps the key factor of these results is represented by the way of how we measure and exploit the SMEs IC.

Furthermore discovering and correctly driving the SMEs IC Such awareness is essential in particular in all countries, where SMEs make up a significant proportion of business and employment.

After the experiment regarding the evaluation of IC, at least in the case of SMEs from Romania it is necessarily to verify methodology on larger samples. Due to high heterogeneity degree it is recommended to take into account the particularities of economic domain.

In the future, this study should be improved by including more indicators especially those that could reveal the relational factor. This study could be also applied on all major domains from Romanian economy to analyse the differences and particularities found in each branch

It is necessary to improve the model on a panel data set of companies in order to facilitate the asynchronous analysis due to distance in time between efforts and effects. However, even if the model has enough possibilities of development, it can be considered a pilot study, or a start point on the complex process of IC measurement and evaluation, being also the first detailed quantitative study applied on SMEs.

References

1. Akinbobola, O.I., **Conflict in Human Capital Relationships: the Impact of Job Satisfaction on Job Involvement in a Workplace**, International Journal of Social Science and Humanity, Vol. 1, No. 2, 2011
2. Alexandru A.A., Casuneanu C., A, **Two-Factor Theory Model For The Romanian Employees Motivation**, Revista Economica No.4 (51), 2010
3. Bontis, N., Cap. XVI, **Managing Organizational Knowledge by Diagnosing Intellectual Capital: Framing and Advancing the State of the Field**, 281, in Malhorta Y. Knowledge Management and Business Model Innovation, Idea Group Publishing, 2001
4. Bontis, N., **National intellectual capital index: a United Nations initiative for the Arab region**, Journal of Intellectual Capital, 5(1), 2004, pp.13–39.
5. Bollen, K., **Structural Equation with Latent Variables**, Wyley series in Probability and Matematical Statistics, John Wiley & Sons Inc., 1989
6. Brookings, A., **Intellectual Capital: Core Assets of the Third Millenium Enterprise**, International Thompson Business Press, London, 1996
7. Campisi D.- Costa, R., **A DEA-Based method to Enhance Intellctual Capital Management**, Knowledge and Process Management, 15(3), 2008, pp. 170-183
8. Carrington, D. -Tayles, M., **The Mediating Effects of Sensemaking and Measurement on the Intellectual Capital and Performance Linkage**, Electronic Journal of

Knowledge Management, 9 (3), 2011, 284-295, available online at www.ejkm.com

9. Chen, H., **Measuring Intellectual Capital using fuzzy analytic process**, International Journal of Innovation and Learning 6(1), 2009, pp. 51-61
10. Davidescu, A., Strat V.A., **Coordinates of a New Romanian Regional Policy - Identifying the Development Poles. A Case Study**, Informatica Economică vol. 18, no. 2, 2014
11. Edvinsson, L., Malone, M., **Intellectual Capital: Realizing Your Company's True Value by Finding Its Hidden Roots**, Harper Collins, New York, NY, 1997
12. Edvinsson, L., Kivikas, L., **Intellectual capital (IC) or Wissensbilanz process: some German experiences**, Journal of Intellectual Capital, 8(3), 2007, pp. 376 – 385 available on line at http://www.emeraldinsight.com/journals.htm?articleid=1621068&show=html
13. Feleaga, L. Et al., **European Evidence On Intellectual Capital: Linking Methodologies With Firm Disclosures**, Acta Oeconomica, 2012, available at http://voicudragomir.ase.ro/LF,NF,VDD,LR_AO2012.pdf
14. Halim, S., **Statistical Analysis on the intellectual capital statement**, Journal of Intellectual Capital, 11 (1), 2010, pp. 61-72;
15. Ileanu, B., **Econometric methods used for the evaluation of Intelectual Capital case study on the Romanian SMEs**, The fifth edition of the International conference "Business Excellence", October 2010, Brasov, Romania
16. Ileanu B.V., Tanasoiu O., **Factors of The Earning Functions and their Influence on the Intellectual Capital of an Organization**, Journal of Applied Quantitative Methods, Vol. 3, No. 4, 2008
17. Ileanu et. Al, **Measuring Intellectual Capital for Romanian SMEs. A comparative analysis by field of activity**, The 8th International Conference on Intellctual Capital, Knowledge Management & Organisational Learning-ICICKM, 27-28 October 2011, Bangkok, Thailand
18. Jarque, C.M., Bera, A. K., **Efficient tests for normality, homoscedasticity and serial independence of regression residuals**, Economics Letters 6: 255-259 , North-Holland Publishing Company, 1980, available online at https://net-files.uiuc.edu/abera/www/ecoletters80.pdf
19. Kline, R.B., **Principles and Practice of Structural Equation Modelling**, 3rd Edition, The Guilford Press, New York, 2011
20. Li, J. et al., **Intellectual capital disclosure and corporate governance structure in UK firms**, Acounting and Business Research, 38(2), 2008, pp. 137-159, available online at http://www.tandfonline.com/doi/abs/10.1080/00014788.2008.9663326
21. Marr, B. et al., **Why do firms measure their intellectual capital?,** Journal of intellectual capital, 4(4), 2003, pp. 441-464
22. Martinez-Torres, M.R., **A procedure to design a structural and measurement model of Intellectual Capital: An exploratory study**, Information & Management, 43(5), 2006, pp. 617-626
23. Martín de Castro, G., López Sáez, P, **Intellectual capital in high-tech firms: The case of Spain**, Journal of Intellectual Capital, 9(1), 2008, pp. 25 - 36

24. Matei, M, **Inovation efficiency analysis for Romania**, Economic Computation and Economic Cybernetics Studies and Research, 44(3), 2010, pp. 93-204;
25. Mouritsen, J., Larsden, H.T., Bukh, P.N.D., **Intellectual capital and the capable firm: narrating, visualizing and numbering for managing knowledge**. Accounting Organizations and Society, 26(7/8), 2001, pp. 735-762.
26. Roos, J., Roos, G., Dragonetti, N. C., Edvinsson, L. **Intellectual capital: Navigating new business landscape**. Houndmills: Macmillan Press, 1997
27. Schumacker, R.E., Lomax, R.G, **A Beginner's Guide to Structural Equation Modelling**, 3rd Edition, Routlege, Taylor & Francis Group, 2010.
28.Stewart, T., **Intellectual Capital**, The New Wealth of Organizations. London, Nicholas Brealey Publishing House, 1999
29. Storey, J et. Al, **The Routlege Companion to Strategic Human Resource Management**, Routlege, Taylor & Francis Group, 2009
30. Sveiby, K.E., **The New Organizational Wealth: Managing & Measuring Knowledge-Based Assets**, San Francisco, USA, Berret-Koehler, 1997
31. Ullman, J. B., Bentler, P. M., **Structural Equation Modeling**. Handbook of Psychology, 2003, pp. 607–634.
32. Yang, C.-Chen, T. **Evaluating the efficiency of intellectual capital management for Taiwan IC design industry**, African Journal of Business Management 4(15), 2010, pp. 3366-3377

[1] Bogdan-Vasile Ileanu is Assistant Professor of Statistics and Econometrics in The Academy of Economic Studies Bucharest, Department of Statistics and Econometrics. He earned his PhD degree in September 2012. His research objectives are connected with human and intellectual capital measure, regional analysis, statistical analysis, labour economics, etc. Bogdan is author or co-author of more than 20 studies presented at important international conferences or published in journals.

[2] Alexandru Isaic-Maniu is Professor in the Academy of Economic Studies Bucharest, Department of Statistics and Econometrics. In a long career of more than 35 years he published more than 300 articles or studies with national and international relevance. Also Professor Isaic-Maniu is author or co-author of more than 10 books and member of more than 20 research projects. His main domains of interest are: Statistics, Econometrics, Human Capital and Financial Evaluation etc.

[3] Claudiu Herteliu currently holds the position of professor within the Chair of Statistics and Econometrics, at the Bucharest Academy of Economic Studies. In 2007, he earned his PhD degree at the same university. His main areas of academic interest are: theoretical and macroeconomic statistics, religion statistics, education statistics, and econometrics. He attended many scientific conferences both in Romania and abroad (USA, Germany, Holland, France, Italy, Hungary, etc.) He published (as co-author and single author) over 20 articles in scientific journals, and contributed to a number of four books published in Romania.

[4] Tudorel Andrei is Professor of Econometrics at the Department of Statistics and Econometrics of the Bucharest Academy of Economic Studies, Romania. He is specialized in economic statistics and structural modelling. During 2008 he was invited Professor at the Dauphine University of Paris, France, where he presented his results on the quantitative analysis of the corruption in the public sector in Romania. He has published more than 40 scientific papers in national an international reviews like Higher Education, The International Journal of Economic Development, Economic Computation and Economic Cybernetics Studies and Research, etc. He was Fellowship Researcher to NATO, Brussels, during 2000–2002 working in the field of quantitative models for the cost of integration in NATO for aspiring countries. He participated with scientific papers in more than 20 international conferences in U.S.A., France, Denmark, Italy, Romania, etc. He has been involved in Public Administration Reform projects in Romania and acted as Director of the Central Unit for Public Administration Reform in the Romanian Government, during 2001

SUSTAINABILITY REPORTING AND CONSUMERS' PROTECTION

Ionel JIANU
Associate Professor, PhD
The Bucharest University of Economic Studies, Romania

E-mail: ionel.jianu@cig.ase.ro

Abstract:
The present article analyses the corporative practices and the consumers' protection from the perspective of sustainability reporting, offering a fundamental theory for the sustainability reporting domain, which explains why a company is responsible for consumers' protection. Based on the domain literature and on qualitative data from reports recommended by Global Reporting Initiative, within the article there is developed a conceptualization of corporate practice regarding consumers' protection, the research having implications for the practice in the domain and for further researches.

Key words: consumer, integrated reporting, sustainability reporting

1. Introduction

Businesses of large companies are carried out around two major objectives: profit and sustainability or sustainability and profit. Their order is not necessarily relevant, conditioning and ensuring one another, both contributing to diversification of the information provided by economic actors that intend to be competitive and to secure the market and the reliability of stakeholders.

Companies, especially those listed, are required to publish yearly both the financial statements and a report on corporate governance, so that users be able to substantiate their analysis and decisions. In addition, companies enlarge the information offered by highlighting the aspects of business sustainability in order to convince the whole society, in general, and the users, in particular, on responsible corporative behavior. Regarding financial reporting, literature abounds with studies and positioning of the various (Minnis, 2011; Bartkus, Glassman and McAfee, 2006; Healy and Palepu, 2001; Baker and Wallagey, 2000; Ward, 1998; Herry and Waring, 1995), the nonfinancial reporting is at the beginning, and what exists, at the level of publications, focuses primarily on social and environmental issues (Brown, de Jong and Levy, 2009; Botescu, Nicodim. and Condrea, 2008; Udayasankar, 2008; Falck and Heblich, 2007; Perrini, Russo and Tencati, 2007; Jenkins, 2006). Social responsibility follows the overall dynamics of international management trends being standardized in packages that can be implemented such as: UN Global Compact, The Global Reporting Initiative (GRI) and ISO standards (Sahlin-Andersson, 2006).

GRI is nowadays the best known frame for voluntary reporting regarding environmental and social issues, being world-wide used (in more than 65 countries). The GRI aim is to harmonize the various existing reporting systems and provide a platform for active dialogue about what is sustainable performance. „Its model was the U.S. financial reporting system FASB, which GRI has sought to expand in depth (global), goal (indicators of social, economic and environmental performance), flexibility (descriptive and quantitative indicators), as well as what regards the interested public (industry, financial sector, accounting profession, civil society, NGOs working on human and environmental rights and other stakeholders)" (Brown, de Jong and Levy, 2009).

Referring to developed markets, communities influenced or affected by the activities of companies enjoy their attention and occupy a central place in the development and implementation of socially-responsible policies and measures. Thus, companies rely on the involvement and participation of communities in the design and implementation of companies' processes and activities that directly affect them (Corus and Ozanne, 2012). At the level of undeveloped markets, negative impact of corporations' activities on communities is not a very important issue for them (Garvey and Newell, 2005).

In all these aspects presented so far, the consumer role is clear and decisive as regards the business development approach. The consumer through its manifestation in the market can have a decisive contribution to the companies' prosperity and affirmation, becomes the actor around which gravitates the entire businesses system. The natural question to ask is about how companies manage to attract, retain and loyal consumers. If we consider the first of the two strategic objectives mentioned above, respectively the profit, the answer to such a question can be given by the pricing policy on products that companies sell. The consumers face two problems: less money in the pocket and higher prices on the market. Along with the understanding that seller aim to increase the price (Bolton & White, 2006), it must be also understand the consumer that wants a reasonable price. The perception of this attribute of price of being reasonable is related to the consumer's assessment ability and to its emotional condition when comparison is made with similar products in terms of acceptability, respectively utility (Xia, Monroe and Cox, 2004). Consumers who perceive price as unacceptable, have a negative attitude on that who applies it (Gebhardt, 2008) and show a self-protective behavior by complaining on that commercial relationship, and even cancel it (Monroe & Xia, 2006). If we consider the second of the two strategic objectives mentioned above, respectively sustainability, the answer lies in promoting by company of specific policies that lead to obtaining and providing products that, in addition to usefulness, should be friendly to both the purchaser and the environment in which he operates.

The proactive behavior of consumers is encouraged; their role is there is redefin by the development of a relationship between them that provides them more than just the opportunity to ensure the product they need. On this background, innovation occurs as a way to push things forward, a way to meet long-term demands of consumers, especially those relating to value. Basically, the traditional way of doing business is given up, the whole business system is reconsidered, collaboration with the consumer and placing him at the forefront of all company processes being crucial. Companies reassess their policies, reestablishes their values, target sustainable development and are forced to make known their approach through communication, by presenting financial and nonfinancial information, practical by sustainable reporting. Is the way to be more expressive, more directly, more understandable to all who aim

to obtain value and, why not, much more marketable. Integrated reporting is the path to financial stability and sustainability of any type of business in the current society.

The present study, based on stakeholder theory, examines the influence of sustainability reporting on consumer's protection, prepared on the basis of GRI practice. For this purpose, we consider a sample of 26 reports from GRI site. The research methodology used in this article is the qualitative analysis that is based on the fundamental theory. The results obtained show the interest of large companies for development of sustainability reporting with help of various categories of stakeholders (consumers in our case).

The paper is organized as follows: the second section is the specialty literature review regarding sustainability reporting, sustainability and consumer care, in the end being presented a short resume on the stakeholder theory; the third section describes the methodological approach and forth section presents the findings. Finally, the fifth section provides the conclusions and the implications of the present research.

2. Literature review

2.1. Sustainability reporting

The sustainability reporting supposes two ways of discussion and understanding: theoretical and practical. Theoretically, till now, in the literature, has not yet been developed a widely accepted definition of "sustainability reporting". Academic articles and business press use the term, but in fact they talk about "sustainability" or "corporate social responsibility". Practically, the companies, being concerned over the social and environmental impact of business and the impact of social and environmental issues on business, manage their sustainability footprint. Recent emphasis has been on the integration of ethical, social, environmental and economic, or sustainability issues within corporate reports. This has been referred to as 'triple bottom line' (Elkington, 1997), or 'sustainability' reporting (Global Reporting Initiative, 2000).

Reporting represents for companies an indirect way dialogue with stakeholders. Reporting is both the vector that provides information on the current situation of the company, its strategy and future direction, as well as the key and basis of any deep discussion about the sustainability of its products and its approaches to competitiveness, development, for supremacy in the market. The demand for information market seems to disagree with the honesty and sometimes coldness and objectiveness of financial reporting figures, focusing more on reports which present in details, on a non-financial manner, data and information about products, suppliers, consumers, company politics and actions regarding the assumed role as an organism responsible and accountable for its actions.

The companies managers are awareness that through a sustainability report they can successfully accomplish their managerial plans, so that they included the sustainability report in the companies organizational culture. Now, regarding the sustainability reporting, there are concerns about the completeness and credibility of these reports (Adams, 2004) and the motives of managers preparing them (O'Dwyer, 2003). But, taking into account that many researchers in the field of sustainability reporting are motivated by a desire to see improvement in the sustainability performance of organizations (Adams and Gonzalez, 2007), we consider that there will be enough room for improvements, by creating a Conceptual Framework for Sustainability Reporting and for a universal application of the sustainability reporting.

2.2. Sustainability and consumers care

In the domain's literature there are several theories which attempt to explain motivations for embracing sustainability path (Garriga and Mele, 2004). According to a Financial Times article by Maitland (2003) "many companies are concluding that they cannot afford not to invest in being seen as responsible." Although this study is based on stakeholders' theory, it is necessary to point out two other theories which could explain the company's consumer care, such as integrative and ethical theories. These suggest that firms engage in socially responsible activities because they have a normative (moral) commitment to serve multiple stakeholders (Berman, et.al., 1999; Garriga and Mele, 2004; Matten and Crane, 2005). Companies, such as Microsoft and Google, are among the strongest proponents of corporate social responsibility (Delaney, 2008; Guth, 2008). Consumers, the most important stakeholder in this context, are different, today we meet with 'greener' and 'ethical' consumers asserting themselves more and more often. Changing consumer preferences are also going to prove to be a tipping point for companies to be aware and act consequently.

The new consumers are asking more and more difficult questions, consumers asking brands how they deal with other products, and so on. A major change in the business world is made by introducing the consumer community to some information about the producer community. This is a very efficient way of changing the old state of things, to bring the new consumer up at the level of partnership with the producer. This relation affects the rules of the game. We are witnessing at the attempt to fill up the gap between producers and consumers by making things much more local, punctual, internal.

The business development is related to the consumer attitudes on business. Each category of stakeholders (consumers, shareholders, employees, managers and so on) is making different choices. These are the things that are bridging the divide between business and development, and it might be a very positive trend. Building this new concept of stakeholder, by introducing inside the shareholders, is based on the new type of company, a company in which the primary stakeholders have a majority share in the structure of the company and are directly linked to the consumer community. Consumer perception of companies on the importance of consumers is an approach that is taking shape not only from business environment, the governments also being actively involved in supporting and promoting measures by which the consumer to be respected. Legislation in the United Kingdom, for example, requires public listed companies to disclose the risks of ethical, social and environmental concerns in their annual reports (Porter and Kramer, 2006).

The consumers have become during the last decades almost a natural resource for the company development. The company has to promote its best policies to attract and to maintain the consumers inside the business. There are a lot of companies which understood this and provide models for "conserving" this natural resources base (the companies which are provider for the featured GRI reports). These companies, which are realizing that they depend on consumers, maintain a good balance between financial and social policies, between shareholders and stakeholders treatments, between interior and exterior process of the company. Even if one of these things moves out of balance, then everybody suffers.

2.3. Stakeholder theory

A sustainability report based on production and presentation of complex information, usually voluntary, extends the information contained in traditional financial statements. Its use may be justified by reference to stakeholder theory, according to which companies should

create wealth for all stakeholders, in contrast to the traditional financial reporting based on creating value, in principal, for the actual and potential investors and for the lenders (IASB Framework, 2010). The basic proposition of the stakeholder theory is that the firm's survival depends on its successful management of relationships with stakeholders. Considering that the information provided in financial statements is usually insufficient, the corporate transparency must be expanded to stakeholders' area and must be presented in an integrated form.

Stakeholder theory is, in our opinion, the most appropriate theory to explain the consumer's role and importance for the company development. Also integrative theory and the theory of ethics participate fully in support of sustainability reporting, in general, and consumer protection, in particular. Because the terms general and particular were used, it is necessary to bring in discussion the broader perspective of the stakeholder theory – an individual can be part of more than one stakeholder group (Freeman, 1984) and the narrow view of stakeholder theory – stakeholders are those groups that are necessary for company survival (Mitchell, Agle, and Wood, 1997). This paper adheres to the narrow perspective of stakeholder theory. Although the company interacts with many groups of stakeholders, the company's survival depends on the consumers.

3. Method

The sustainability and the consumer care is a relatively new direction in the modern economic research, and, most important, the companies' perceptions of sustainability reporting represent a complex area where more work is needed. In this context we consider that a research strategy that usually emphasizes words rather than quantification in the collection and analyses the data – the qualitative research, represents an appropriate approach to knowledge generation. Qualitative content analysis used in this paper is the most used approach in qualitative analysis, although in current research approaches the quantitative content analysis is preffered (Insch, Moore and Murphy, 1997). So, we have decided to concentrate our attention on the company's behavior in relationship to consumer, analyzing a particular aspect of organizational culture, namely how the company implements sustainability reporting (particularly for consumer protection). Given these reasons, qualitative content analysis, semiotics and hermeneutics are used based on the grounded theory to identify the dimensions of the company policies in consumers' protection.

3.1. Data collection

For the study 26 reports recommended by GRI were analyzed. All the reports are organizational documents which that due to the relevance of the information presented are recommended by GRI to be used as a model. So, we have a heterogeneous group of sources which are of particular interest for both the business world and research, due to the vastness of information relating to documentation. These documents were produced for the business development, they are non-reactive, and so, they do not affect the validity of the research results. Also, the criteria for assessing the quality of the documents (Scott, 1990): authenticity, credibility, representativeness and meaning have been met, taking into consideration that these documents were certified by the independent audit firms on the one hand and by the GRI group of experts on the other hand.

GRI report presents, as general view, the necessary information belonging to an entity so that it can improve the aspects of the elements of economic, social and environmental

performance. The indicators recommended by GRI to be included in the social responsibility report have the mission to guide the content of the report and not offer solutions on how data can be collected for presentation or calculation of indicators. The importance and recognition that GRI has gained worldwide has led us to choose this referential as a basis for study social responsibility reports. The performance indicators recommended by GRI are of three types: economic performance indicators (9 indicators), natural environmental performance indicators (30 indicators) and social performance indicators (40 indicators). Social performance indicators recommended by the GRI are grouped into four categories: labor and decent work practices (14 indicators), human rights (9 indicators), society (8 indicators) and product liability (9 indicators). The analysis was focused on product liability area, being analyzed indicators reflecting company policy on consumer protection: Customer Health and Safety, Product and Service Labeling, Marketing Communications, Customer Privacy and Compliance.

3.2. Sample

Purposive sampling was employed in the selection of the GRI reports. This means that reports were chosen that could inform us about the research problem addressed in this study, similar to the working procedures used in other works that deals with the issue of sustainability reporting (Creswell, 2007). The sampling strategy was thus driven more by theory than by representativeness. Our selection was directed to those reports that included companies with high levels of sustainability reporting behavior regarding the company's attitude towards consumers needs. At the time of data collection (October 2013), on GRI website were presented and recommended as a model 42 sustainability reports produced by large multinational corporations. From these, only 26 were analyzed (Table 1), remaining 16 reports were removed from the study because: 1 of them could not be downloaded, 9 of them were not written in English but in Spanish, Portuguese or Arabic, and six reports did not provide information about the company policy on consumer protection.

3.3. Data analysis

Grounded theory has become by far the most widely used framework for analyzing qualitative data (Glaser and Strauss, 1967). According to this theory is required to code data, to combine the codes into broader categories and themes, and to interpret the results. In the first step, categories related to sustainability reporting (open coding) were formed. In the second step, a category was identified as the central phenomenon: consumer protection. Third, axial coding dealing with the central phenomenon was undertaken. Finally, all of the categories were drawn upon to develop a typology. For data interpretation was used, through inductive reasoning, qualitative content analysis. A semiotic analysis which suggests that the symbolic order of a culture is constructed and interpreted through a system of signs was used to highlight the new meanings of the producer – consumer relations (Barley, 1983). To obtain objective results, the entirely analyses of this research was realized under umbrella of the hermeneutics approach.

Table1. The sample composition

Company Name	Activity Sector	The number of sustainability reports	The name of the present report	GRI checked	The length of the present report (pages number)	How many time the word consumer / customer appears in report
Larsen & Toubro	Construction, Industry	6	L&T's Corporate Sustainability Report 2013	A+	136	6
ÇİMSA	Industry	3	Sustainaility Report 2012	A+	92	3
Commerzbank	Banking	2	Corporate Responsibility Report 2013	A	88	6/319
Bank Asia Limited	Banking	1	Sustainability Report 2012	B	84	8/43
Liberty Global	Media Cable	2	Corporate Responsibility Report 2012	B	60	23/169
BHP Billiton	Industry	15	BHP Billiton Sustainability Report 2013	A+	54	1/15
Petrochemical Industries Company	Industry	5	PIC Sustainability Report	B	49	7/25
Family Develoment Foundation	Government Organization	2	Clear vision Towards Sustainability	A	87	0/10
Dell	Industry	3	FY13 Corporate Responsibility Summary Report	A	39	5/36
Queiroz Galvão Exploração e Produção	Industry	2	Sustainabiliy Annual Report 2012	B	39	10/34
Q8	Industry	1	Q8 Corporate Sustainability Report	C	42	1/11
Northrop Grumman	Services and Industry	5	2012 CORPORATE RESPONSIBILITY REPORT	A+	42	0/62
Votorantim	Industry	3	Integrated Report Votorantim	B+	160	8/82
United Parcel Service (UPS)	Services	10	UPS Corporate Sustainability Report 2012 - More of What Matters	A+	143	6/188
İşbank	Banking	1	Sustainability Report 2012	B	120	4/237
CapitaLand Limited	Construction Real Estate	4	CapitaLand Limited Global Sustainability Report 2012	B+	73	1/13
Alcatel-Lucent	Telecommunications services	3	2012 Sustainability Report	A+	242	11/186
Export Development Canada	Financial Services	9	CSR in the New Reality	B+	43	0/53

Company Name	Activity Sector	The number of sustainability reports	The name of the present report	GRI checked	The length of the present report (pages number)	How many time the word consumer / customer appears in report
Dubai Customs	Government Organization	5	Dubai Customs Sustainability Report - 2012	A	30	0/45
Pacific Rubiales Energy Corp	Construction Real estate	5	Pacific Rubiales Sustainability Report 2012	A+	240	2/23
Hamburger Hafen und Logistik AG	Transporturi	2	Annual Report 2012	B+	225	3/60
Hydro-Québec	Electric Utilities	5	Sustainability report 2012	B	48	8/67
Vancity	Banking	10	2012 Annual Report: The Vancity Effect	A+	72	5/6
Energy Development Corporation	Energy	6	It's Possible: EDC 2012 Performance Report	A+	194	5/64
SNAM	Industry	6	Sustainability Report 2012	A+	116	5/29
Ecopetrol		4	Reporte Integrado de gestión Sostenible 2012	A+	385	7/76

Table 2: Customer care through Product Responsibility

Dsiclosures	Customer health and safety	Products and service	Marketing Communica-	Customer privacy	Compliance (PR)
1.	Partially reported	Fully reported	Fully reported	Fully reported	Fully reported
2.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
3.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
4.	Not reported	Fully reported	Fully reported	Fully reported	Fully reported
5.	Not reported	Fully reported	Fully reported	Fully reported	Fully reported
6.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
7.	Not reported	Fully reported	Not reported	Not reported	Not reported
8.	Fully reported	Not reported	Fully reported	Fully reported	Fully reported
9.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
10.	Not reported	Not reported	Not reported	Not reported	Not reported
11.	Not reported	Not reported	Not reported	Not reported	Not reported
12.	Not reported	Not reported	Not reported	Not reported	Not reported
13.	Partially reported	Fully reported	Partially re-	Not reported	Fully reported
14.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
15.	Not reported	Not reported	Fully reported	Fully reported	Fully reported
16.	Partially reported	Not reported	Partially re-	Partially re-	Not reported
17.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
18.	Not reported	Not reported	Not reported	Not reported	Not reported
19.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported

Dsi-closures	Customer health and safety	Products and service	Marketing Communica-	Customer privacy	Compliance (PR)
20.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
21.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
22.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported
23.	Not reported	Fully reported	Fully reported	Fully reported	Fully reported
24.	Not reported	Not reported	Not reported	Not reported	Not reported
25.	Fully reported	Not reported	Not reported	Not reported	Fully reported
26.	Fully reported	Fully reported	Fully reported	Fully reported	Fully reported

4. Results

On the basis of the findings of the qualitative study, we highlighted the place of consumer in the center of the strategic objectives of corporate social responsibility which perceive the implementation and publication of the consumer protection practices as a corporate perspective of sustainability reporting.

4.1. Sustainability domains and company's perception

The term social sustainability of the company is used to cover a broad area of concepts in scientific literature, making its exact definition to be ambiguous. The broadening of its coverage area has been present in contemporary literature by including the term stakeholders and implicitly many other aspects related directly or indirectly to the company life. Inaccurate terminology is generated also by his dynamic character and changing resultat of its alignment with current social problems. Social sustainability orientation does not eliminate the problem of company profitability, which remains the main reason for the existence of the company, but add on the management agenda social and environmental problems resulting from the activity of the enterprise and how they affect communities (Nasrullah and Rahim, 2014).

We started the construction of this approach from the consideration that reporting at the company level, is undergoing a period of search and retrieval, in the idea of being close to what the free market and modern society requires. As the way to penetrate into the market is satisfying consumer demands, and the consumer, as an individual, is the basic element of modern society, awareness and moving toward offering a varied and transparent information constitues strong attributes for performance of companies. Companies are turning to its stakeholders, entering and feeding relationships without which their existence would not be possible (Freeman, et al., 2010). Ignoring by the company of non-financial aspects of performance cause loss of market share and value for the company and how it is perceived in the market and generate additional costs with informations and change negativelly the way company's products are perceived by clients (Pintea, 2011).

The positioning of the top management of the companies analyzed in the sample provide certainty on this new guidance on reporting that is required to be made. Thus, at their level has become apparent, in accordance with the principles of corporate governance, strategic planning cycles to incorporate new dimensions of thinking and proactive action. Companies is willing to focus on those issues that matter to all stakeholders, basically they assume responsibility to respond to all messages received from all interested parties.

"Liberty Global's corporate responsibility strategy focuses on the issues that are most important to our stakeholders, as well as those issues that have a material impact on the future competitiveness of the business. These topics were identified and prioritized in a materiality assessment process which we conducted in early 2012. Our top seven most material issues include: customer privacy and security; digital inclusion; e-waste; transparency and disclosure; energy use and GHG emissions; and protecting children on the internet and TV. We have grouped these topics into four key issue areas: 1.promoting a digital society; 2.building trust with our customers; 3.managing our environmental impacts; and 4.being a responsible business. Our CR Framework summarizes our approach to managing these opportunities and challenges". (Corporate Responsibility Report 2012)

New goals appear, we need the personnel specialized in the relationship with each of these stakeholders categories, as each may be decisive for the existence of the company. If until now we have shown that happens in relation to outside the company, it must be said that these efforts are considerable for stakeholders inside the company. These companies have chosen the path of transparency, chose to communicate with stakeholders bringing confidence from the inside out and from the outside taking involvement.

4.2. Corporate stages of sustainability reporting development

Not all the things we bring into discussion in this article, referring to the issue of sustainable reporting recommended by GRI reports were perceived, developed and presented so from the beginning. It therefore seems essential the need to provide guidelines for the creation of a true and fair view of sustainable reporting.

It must be said that both the theory and practice of sustainable reporting occupies a central place within the concerns and approaches of all social actors responsible in all spheres of activity, from business to the society and from the society to business. States through specialized bodies have become actors involved in the issue of sustainable reporting, these being among the most energetic vectors in social responsibility activities.

Sustainability has evolved in recent years from a concept on the company's impact on the environment in which it operates, in an overall management concept based on three pillars, namely environment, social and economic sustainability, known as triple bottom line – TBL or three pillars - profit, people, planet. TBL concept is addressed to stakeholders that participate directly or indirectly impacted by the company's actions. Principles of sustainability have as main objective the development of globalized companies, through the participation of all major economic actors (governments, companies, society) with respect to the environment and people.

The business of large companies can not and should not be seen only through the information that shows performance and financial stability because the companies are living organisms. Basically, this is achieved by sustainable reporting, informing the public about socially responsible actions, the steps for a green planet and a clean environment, to position the company as local actors, zonal, regional and even global involved in community life.

Things began timidly, given the traditional culture of the business environment - business for profits. But businesses progresses well and are about to become sustainable, without reducing profit. The best example through qualitative research that we propose, is the name given of the reports submitted by companies (Table 1, Column 5). It is said that a good title

provides more than 60% chances of selling a book, so you must agree that those who promote the culture of our corporations know how to do, they do very well and convey a lot, including the idea of evolution. Since the current Sustainability Report or Corporate Social Responsability to the Annual Report (with a retro twist) and even to the futuristic Integrated Raporting, all bring about the story and charm periods they have spent and immortality unless they devote at least place collective memory.

4.3. Sustainability reporting and consumers care

We bring into question the sustainability of the product and hence the sustainability of the business. We go a little under which we are accustomed, advocating for change in perception and claim that a sustainable product is the key to business success. No matter how involved we were in environmental protection campaigns, how many social projects we support and run (without minimizing human value), we can not change the perception of the products that do not meet consumers needs. Consumer validates corporate actions on social and environmental responsibility, while not all companies have the same level of involvement in such activities that generate additional costs. In a competitive market, achieving social responsibility actions that are not appreciated by the consumer disadvantage companies by the high costs involved (Steger, Ionescu-Somers and Salzmann, 2006).

If we talk about sustainable actions undertaken by companies, they seem to get past environmental and social labels and objectified, meaning approaching actual consumer, to meet all its requirements. Such an approach is the one proposed by a major player the German financial world, Commerzbank:

> *"We have defined a wide range of sustainability issues and reviewed the relevance of these issues for our stakeholders and the bank. The results were aggregated to form a materiality matrix. The issues that were of equally high importance to us and our stakeholders in 2012 mainly relate to the sphere of action "market and customers", for example customer satisfaction and consumer protection. By contrast, environmental issues were less significant because our banking operations have a comparatively low impact on the environment"* (Corporate Responsibility Report, 2013).

This position is certain evidence that managers of large corporations realize that its product policy is best valued from the consumer perspective. No matter how many "green" or human initiatives you might have, your products will not be marketable. However, the more consumer is satisfied, the product will record higher sales.

Another example that things are going to prioritize the relationship that the company cultivates consumers/customers is the content of the column No. 8 of Table 1. The use of the two terms of sustainability reports is quasi-permanent, with no area of interest with no reference or use of consumer concepts, namely customer.

The company's involvement with stakeholder leads to increased company value and the recent evolutions show that companies environment and mutual support stakeholders depict resolving common problems and support each other. Social sustainability practiced by the company determines a natural process of selecting good managers who can adopt such policies, and makes it imperative specialized entities such as social or ethical audit. (Cespa and Cestone, 2007). Together with participation of stakeholders and their information about the policy adopted by the company ensure sustainable development of the company, the gain

from the synergy between the company and stakeholders stimulating the creation and alignment of the company's social policies to the current requirements of the society.

The constant policies in relation to stakeholders ensure the company's prosperity. Thus, it is not irrelevant the argument of placing consumers or customers in the forefront of stakeholders, interacting with them and partnership:

> *"At Dell, we start by listening and that process is critical to our reporting. Stakeholders for us include our customers, activists, industry partners, nongovernmental organizations, government agencies and regulators, suppliers, investors and team members. We regularly engage them, soliciting feedback, reviewing goals and responding to inquiries throughout the year"* (Corporate Responsibility Summary Report).

Thus. the extremely vast landscape of responsibility regarding the company, aquire concreteness through the role of the receiver of the company, which must hear the voice of consumers, to reach a compromise between what the company wants and the signal of the consumer so that the most consumer needs are met.

> *"We consider the viewpoints of a wide range of stakeholders, including customers, employees, investors, community leaders, and NGOs through both formal and informal channels. This is extensively discussed in our report specifically as part of our materiality matrix program where we surveyed and cataloged their concerns and have a plan to prioritize and address the top-tier issues. We also list our stakeholder engagement program and collaboratively work with a variety of external initiatives to make greater impact"* (UPS Corporate Sustainability Report 2012).

We have a speech that makes part of an agreement (formal aspect of producer-consumer relationship), in a deal (informal look the same relationship), a fairly clear and convincing evidence that each party must assume and comply with the terms of the collaboration.

> *"Dialogue helps shape İşbank's business strategy and the approach to sustainability. Accountability to stakeholders –customers, employees, shareholders and others affected by the Bank's operations – involves sharing insights and addressing concerns. Understanding consumers is the basis for the İşbank brand promise. Consumer insight is decisive for both the business strategy and product development. Comprehensive interviews and visits to customers throughout the network enable the Bank to identify trends in and respond to them in the product and service offering. İşbank's stakeholders are those institutions and individuals; that are potentially influenced by İşbank's decisions, activities, products and services; that have a potential influence on İşbank with their decisions, activities, products and services".* (Sustainability Report 2012).

Regarding the relationship company (producer) - consumer, we believe it is necessary to build a fundamental theory of the sustainability reporting which emanates managerial transparency (from the company) and transparent decision-making (from the consumer). Such an approach might run as follows:

- ✓ Assuming and promoting partnership with consumers: declaring the consumer as a resource of strategic importance; harmonizations of the internal company documents with the contents of all regulations (laws, decrees, resolutions) on consumer protection; permanent consultation of the consumers by using all possibilities of communication (direct contact, using social networks, e-mail addresses); attracting the customers to the company's shareholders; involvement in joint projects and initiatives.
- ✓ Product policy: consumer health and safety; products and service labeling; marketing communications; customer privacy; compliance.
- ✓ Reporting focused on consumer: presentation of financial position statement of the company, by recognizing Asset (resource) "consumers" (traditional consumers, current consumers, occasional consumers) and Equity (source) "consumers"; statement of financial performance by classifying revenues and expenses by the embodiment (revenues and expenses in relation to traditional consumers - commercial relationships older than 5 years, revenues and expenses in relation to current customers - commercial relations more than 1 year but less than 5 years and revenues and expenses in relation to occasional consumers - commercial relations in the current year); the presentation of the cash flow statement on the consumer (traditional cash flows, current cash flows, occasional cash flows);

In all three positioning we find through appeal to the procedure semiotics, very clear signs about the role and importance in the lives of consumers and companies approaches. Thus placement in top of the list of stakeholders is an extremely strong signal ; continuous consultation with them also, and the fact that the two major players - the producer and the consumer depend on each other, completes the picture.

5. Conclusions

Based on the results of this study we can highlight the importance given by the company to its consumers, which are placed in most of analyzed reports in the forefront of the company's trusted partner. Stakeholder theory is useful in analyzing data on corporate sustainability. Group affiliation and ownership structure as well as corporate governance application, determine sustainable involvement of companies and correct information to stakeholders.

First, the results of the study highlighted the importance of stakeholders to develop business and related to this, the company management orientation for their stakeholders. Although so far we have not made reference other than the management of companies, to be responsible for providing a comprehensive image about the business, we must say that the financial accounting system is the part aimed to perform the entire procedure for financial support and responsible attitude. Purely financial reporting, still satisfactory at the time of this article, should be seasoned with everything we have shown that courses of action are moving into consumer requirements.

Secondly, the results of the study have provided, by recourse to semiotics, clear signs of positioning in the top list of consumer stakeholders. The consumer is the key to any sustainable approach developed at company level. Product sustainability translates now through performance and corporate ethics, leading to the strengthening of headquarters of a successful business: sustainability and profit. Another important aspect reflected in the study is the one that gives the required notice of balance, meaning it promotes the idea that a company can

not get from stakeholders more that it can give, if we refer only to the consumer, and that the relationship must be balanced between all categories of stakeholders, without favoring one of them over others.

Third, the results of the study reflected the importance of consumer for the company, the acceptance of the consumer in everything that means strategy, action plan and activities at the company level. Consumer becomes an agent of welfare, a cornerstone of any building policies and product strategies. We can speak now of agency theory in producer-consumer relationship, the consumer being the agent who, through his actions bring welfare to the producer. Thus, the relationship between producer and consumer is based on a permanent feedback, companies focusing on this aspect.

Acknowledgements

This paper was co-financed from the European Social Fund, through the Sectorial Operational Programme Human Resources Development 2007-2013, project number POSDRU/159/1.5/S/138907 "Excellence in scientific interdisciplinary research, doctoral and postdoctoral, in the economic, social and medical fields -EXCELIS", coordinator The Bucharest University of Economic Studies.

References

1. Adams, C. A., Larrinaga Gonz´alez, C., **Engaging with organisations in pursuit of improved social and environmental accountability and performance.** Accounting, Auditing and Accountability Journal, 20(3), 2007, pp.333–355.
2. Baker, C. R., Wallagey, P, **The Future of Financial Reporting in Europe: Its Role in Corporate Governance,** The International Journal of Accounting, 35(2), 2000, pp.173-187.
3. Bartkus, B., Glassman, M, McAfee, B, **Mission Statement Quality and Financial Performance,** European Management Journal, 24(1), 2006, pp.86–94.
4. Barley, S., **Semiotics and the study of occupational and organizational cultures**, Administrative Science Quarterly, 28, 1983, pp.393-413.
6. Berman, S. L., Wicks, A. C., Kotha, S., Jones, T. M., **Does stakeholder orientation matter? The relationship between stakeholder management models and firm financial performance,** Academy of Management Journal, 42(5), 1999, pp.488–506.
7. Bolton, L. E., Alba, J. W., **Price fairness: Good-service differences and the role of vender costs**. Journal of Consumer Research, 33, 2006, pp. 258–265.
8. Botescu, I., Nicodim, L., Condrea, E., **Business Ethics and the Social Responsability of the Company**. Amfiteatru Economic. 10(23), 2008, pp.131-135.
9. Brown, H.S., de Jong, M. and Levy, D.L, **Building institutions based on information disclosure: lessons from GRI's sustainability reporting**. Journal of Cleaner Production, 17(6), 2009, pp. 571-580.
10. Cespa, G., Cestone, G., **Corporate Social Responsibility and Managerial Entrenchment**. Journal of Economics & Management Strategy, 16, 2007, pp. 741–771
11. Corus, C. and Ozanne, J. L., **Stakeholder engagement: Building participatory and deliberative spaces in subsistence markets**, Journal of Business Research, (65), 2012, pp.1728-1735.

12. Creswell, J. W., **Qualitative inquiry and research design — Choosing among five approaches**. US: Sage, 2007
13. Delaney, K.J., **Google: From don't be evil to how to do good,** The Wall Street Journal, January 18, 2008, p:B1
14. Elkington, J., **Cannibals with Forks: The Triple Bottom Line of 21st Century Business**, Oxford: Capstone, 1997.
15. Falck, O., Heblich, S, **Corporate Social Responsibility: Doing Well by Doing Good,** Business Horizons, 50(3), 2007, pp.247-254.
16. Freeman, R. E, **Strategic management: A stakeholder approach,** Boston, MA: Pitman, 1984.
17. Freeman, R. E., Harrison, J. S.,Wicks, A. C., Parmar, B. L. and De Colle, S, ***Stakeholder theory — The state of the art,*** Cambridge University Press, 2010.
18. Glaser, B. G. and Strauss A. L, ***The Discovery of Grounded Theory: Strategies for Qualitative Research***, Chicago: Aldine, 1967
19. Garriga, E. and Mele, D., **Corporate social responsibility theories: Mapping the territory,** Journal of Business Ethics, 53(1–2), 2004, pp.51–71.
20. Garvey, N. and Newell, P., **Corporate accountability to the poor: Assessing the effectiveness of community-based strategies,** Development in Practice, 15(3), 2005, p.389–404.
21. Gebhardt, G. F., **Social justice in marketing: Fairness, satisfaction and customer lifetime value,** Marketing Science Institute, Special Report, January, 2008.
22. Grewal, R., Chandrashekaran, M., Citrin, A. V., **Customer satisfaction heterogeneity and shareholder value**, Journal of Marketing Research, 47(4), 2010, pp.612–626.
23. Guth, R.A., **Bill Gates issues call for kinder capitalism — Famously competitive, billionaire now urges business to aid the poor.** The Wall Street Journal, January 28, 2008, p.A1.
24. Healy, P. and Palepu, K., **Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature,** Journal of Accounting and Economics, 31, 2001, pp.405-441.
25. Herry, R. H., Waring, A, **A user perspective on 'making Corporate reports valuable**, British Accounting Review, 27, 1995, pp.139–152.
26. Homburg, C., Hoyer, W. D., Koschate, N., 2005. **Customers' reactions to price increases: Do customer satisfaction and perceived motive fairness matter?** Journal of the Academy of Marketing Science, 33(1), 1995, pp.36–49.
27. IASB, **The Conceptual Framework for Financial Reporting**, 2010
28. Insch, G., Moore, J., Murphy, L., **Content analysis in leadership research: exemples, procedures and suggestions for future use,** *Leadership Quarterly*, 8(1), 1997, pp. 1-25.
29. Jenkins, H., **Small Business Champions for Corporate Social Responsibility,** Journal of Business Ethics, 67(3), 2006, pp.241–256.
30. Maitland, A., **No hiding place for the irresponsible business: Companies are trying to pre-empt trouble over social, environmental and ethical issues,** Financial Times, September 29, 2003, p.2.

31. Matten, D. and Crane, A., **Corporate citizenship: Toward an extended theoretical conceptualization.** Academy of Management Review, 30(1), 2005, pp.166–179.
32. Minnis, M., **The Value of Financial Statement Verification in Debt Financing: Evidence from Private U.S. Firms.** Journal of Accounting Research, May, 2011, pp.457–506.
33. Mitchell, R. K., Agle, B. R., Wood, D. J., **Toward a theory of stakeholder identification 34. and salience: Defining the principle of who and what really counts,** Academy of Management Review, 22(4), 1997, pp.853–886.
35. Monroe, K. B. and Xia, L., **The price is unfair!** In J. N. Sheth, and R. S. Sisodia (Eds.), Does marketing need reform? (pp. 158–165). Armonk, NY and London, England: M.E. Sharpe, 2006
36. Nasrullah, N. M. and Rahim M. M., **CSR in Private Enterprises in Developing Countries**, Springer: Switzerland, 2014.
37. O'Dwyer, B., **Conceptions of corporate social responsibility: The nature of managerial capture**, Accounting, Auditing & Accountability Journal, 16, 2003, pp.523–557.
38. Perrini, F., Russo, A., Tencati, A., **CSR Strategies of SMEs and Large Firms. Evidence from Italy**, Journal of Business Ethics, 74(3), 2007, pp.285–300.
39. Pintea, M. O., **Financial and non-financial aproaches regarding the growth of the economic entities performances**, PhD thesis, 2011
40. Preston, L. E., & Donaldson, T., **Stakeholder management and organizational wealth. Academy of Management Review**, 24(4), 1999, 619–620.
41. Sahlin-Andersson, K., **Corporate social responsibility: A trend and a movement, but of what and for what?** Corporate Governance, 6(5), 2006, pag.595-608
42. Steger, U., Ionescu-Somers, A., Salzmann, O., ***Inside the mind of stakeholder, Perspectives for managers***, 2006, disponibil la http://www.imd.org/research/publications/upload/PFM139_LR_Steger_Ionescu_Salzmann.pdf (July 2013)
43. Udayasankar, K, **Corporate Social Responsibility and Firm Size**, Journal of Business Ethics, 83(2), 2008, pp.167–175.
44. Xia, L., Monroe, K. B., Cox, J. L., **The price is unfair! A conceptual framework of price fairness perceptions**. Journal of Marketing, 68(3), 2004, pp.1–15.
45. Ward, M., **Summary financial statements: an analysis of the adoption decision. A research note**. British Accounting Review, 30, 1998, pp.249–260.
46. https://www.globalreporting.org/reporting/report-services/featured-reports/Pages/default.aspx