ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Silesian Statistical Review
2024, nr 22(28) ISSN 2449-9765

Extended Abstracts of the 26th Scientific Statistical Seminar „Marburg-Wrocław"

DOI: 10.15611/sps.2024.22.08

Identification of Significant Changes: A Case Study of the Age Structure of Polish Society (2018-2022)

Joanna Dębicka
Department of Statistics, Wroclaw University of Economics and Business, Poland
e-mail: joanna.debicka@ue.wroc.pl

Edyta Mazurek
Department of Statistics, Wroclaw University of Economics and Business, Poland
e-mail: edyta.mazurek@ue.wroc.pl

Katarzyna Ostasiewicz
Department of Statistics, Wroclaw University of Economics and Business, Poland
e-mail: katarzyna.ostasiewicz@ue.wroc.pl

Change is a continuous phenomenon encompassing various spheres of human life and the natural world. It can be perceived, experienced, observed, planned, and managed. Defined as a transformation or difference between the frequencies of classes in two structures, change can be positive, negative, or neutral. In the presentation, we focus on identifying significant changes in structural tables, distinguishing random fluctuations from meaningful shifts. Various similarity coefficients and statistical tests determine whether observed differences arise from chance or signify substantial deviations. These methods exhibit reasonably high accuracy when addressing profound/revolutionary changes but lack sensitivity when a structural transformation is mild and gradual.

Let us consider a variable for sample populations *X* and *Y* composed of N_X and N_Y units, which are divided into *n* classes and the numbers of units in each class are marked with the symbols $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ respectively. We are interested in the question of when the differences between the elements of the structures $d_i = \omega(x_i) - \omega(y_i) = \frac{x_i}{N_X} - \frac{y_i}{N_Y}$ can be considered *distinctive*. We have named them distinctive due to the concepts of outliers and atypical observations already reserved in statistics. Note that we can also treat the differences d_i as a transformation in time and then we consider the changes in class frequencies for the same population changing over time.

When it comes to statistical methods, theoretically we have a wide range of methods that can be used in comparative analysis. In the case of features measured on interval or ratio scales, we have

many different ways to compare two distributions. We have different tests, depending on the assumptions and for almost every situation, for this type of data, we will find a method. In the case of unmeasurable features, we basically have chi-square tests of independence or homogeneity with several modifications/corrections due to the sample size. In addition, to measure the conformity of structures, primarily various measures of similarity or dissimilarity based on the distance or pseudo-distance function are used. While various measures of similarity of structures are presented in the literature, statistical tests that allow determining statistical significance in assessing similarity or dissimilarity are rarely presented.

In practice, the methods mentioned are not always insufficient. Commonly known methods will detect revolutionary changes in structures, i.e. those that usually occur quickly, in a short period of time and lead to fundamental changes. However, if we are dealing with evolutionary changes, meaning those that occur gradually in a long-term process. In such situations, intuition suggests that changes exist, but statistical tests and structure similarity coefficients show that the empirical distributions of the studied variable in both populations are similar. So, to detect such changes, we need a more sensitive tool.

This type of tool was proposed by J. Dębicka and E. Mazurek in Identification of distinctive changes in structural tables (manuscript). According to this method, a relative change in the structure in the *i*-th class $r_i = \frac{d_i}{g_p} = \frac{\omega(x_i)-\omega(y_i)}{g_p}$ can be considered *distinctive* against all changes in the structure of the remaining classes if it belongs to the area $\left[\frac{\omega_p-1}{g_p}; -1\right) \cup \left(1; \frac{1-\omega_p}{g_p}\right]$, where $g_p = min\{|min\{d_1, d_2, \ldots, d_n\}|, |max\{d_1, d_2, \ldots, d_n\}|\}$ and $\omega_p = \sum_{i=1}^{n} min\{\omega(x_i), \omega(y_i)\}$ is structure similarity coefficient. The depth of distinctive changes can be different, therefore they are assessed as follows. If $|r_i|$ it belongs to the interval:

- (1, 1.10) – then the change is insignificantly distinctive,
- [1.10, 1.25) – then we observe a little distinctive change,
- [1.25, 1.40) – then we observe a moderately distinctive change,
- [1.40, 1.60) – then we observe a very large distinctive change,
- Above 1.6 – then the distinctive change is huge.

The age structure of the population is a feature for which we do not observe spectacular changes in short periods of time, and usually, statistical tests, as well as structure similarity coefficients, indicate that in two moments of time, the structures are very similar. Due to the fact that demographic changes in the years 2018-2022 included:

- a decline in the birth rate (lifestyle changes, postponing the decision to start a family, as well as economic issues and instability),
- an increase in mortality (especially due to the COVID-19 pandemic),
- an ageing population (the increase in the number of people of post-productive age and the increase in average life expectancy),
- the impact of migration, especially from Ukraine (before 2022, mainly due to economic factors, such as better working conditions. In 2022, after the outbreak of the war in Ukraine, there was a dramatic increase in the number of refugees, mainly women and children),

accumulated and can suggest potential differences in the age structure of the Polish population, we decided to choose this period and check with the authors' method in which age groups and whether we will observe any distinctive changes at all.

The analysis will be based on data from a Eurostat database.

We made comparative analyzes comparing subsequent years to 2018 and comparing age structures year to year (com. Table 1).

Table 1. Identification of Significant Changes in the Age Structure of Polish Society (2018-2022)

	Compared to 2018				Year-on-year			
Years	(2018,2019)	(2018,2020)	(2018,2021)	(2018,2022)	(2018,2019)	(2019,2020)	(2020,2021)	(2021,2022)
	Structure Similarity Coefficient							
	0.99	0.98	0.97	0.95	0.99	0.99	0.99	0.99
Age groups	Significant Relative Changes							
45-49		1.02	1.07	1.05			1.13	
70-74	1.49	1.65	1.41	1.28	1.49	1.51		
75-79								1.17

Source: own elaboration.

Note that the age structures of Poles are very similar regardless of the years of the 2018-2022 period from which we compare the age structures (see the fourth row of Table 1). Despite this, in the period under review we observe distinctive changes in the age structures of the three groups specified in the last three rows of Table 1, in which the values that are significant (greater than 1.1) are colored grey. The share of these age groups in the age structure of the Polish population increased more than in other age groups. In the 45-49 age group, there are distinctive changes, but they are insignificant (random), only the change in the share of this age group in the structure of the Polish population in the third year of the pandemic (2021) compared to the second (2020) can be considered a small but significantly distinctive change (1.13). We observe a similar situation for the 75-79 age group, but here it concerns the years 2021 and 2022 (1.17).

In the 70-74 age group, we observe changes that stand out in the age structures of 2019-2022 when we compare them to 2018. Note that these changes may have two sources of origin and may be related to the fact that the Baby Boom generation (born in 1946-1964) in 2018 accounted for about 60% of people belonging to this age group (which consisted of people born in 1944-1948), and in 2022 they accounted for 100% of people belonging to the 70-74 age group (people born in 1948-1952 in 2022 belonged to the Baby Boom generation). Comparing the structures from 2019-2021 with the age structure of the Polish population in 2018, we see that the increase in the share of this age group in the entire age structure of Poles is large (1.49, 1.41) or even huge (1.65), and then falls to the moderately in 2022 (1.28). This may also indicate that this age group has been very affected by mortality due to Covid19, because although it includes people from the Baby Boom generation, its share in the entire age structure of Poles has dropped significantly compared to 2018. This is confirmed by the distinctive changes in the studies of age structures year to year, where we do not observe distinctive changes in the structure in this age group in the years (2020, 2021) and (2021, 2022). In order to confirm this hypothesis, the age structure of deaths was examined. Figure 1 graphically presents only fragments of the structures for the years 2018-2022, i.e. for age groups from 65 years of age.

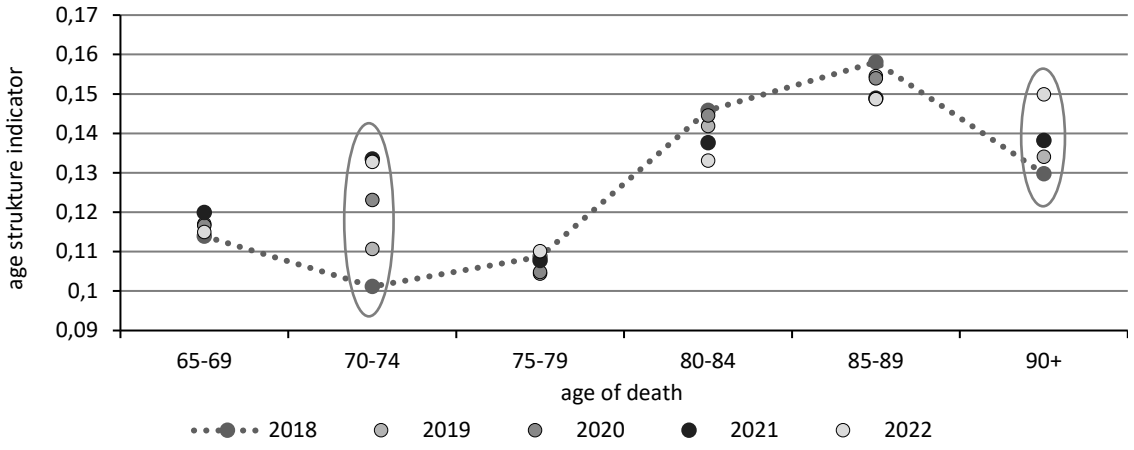

Figure 1. Age structure indicators of death for age groups 65+

Source: own elaboration.

Note that the largest increases in structure indicators of death in the entire population compared to 2018 are observed in the age groups 70-74 and 90+ (marked with ellipses in Figure 1). The significance of these changes was confirmed by indicating distinctive changes in the age structure of deaths among Poles in 2018-2022 (see Table 2).

Table 2. Identification of Significant Changes in the Age Structure of Poles' Deaths in 2018-2022

	Compared to 2018				Year-on-year			
Years	(2018,2019)	(2018,2020)	(2018,2021)	(2018,2022)	(2018,2019)	(2019,2020)	(2020,2021)	(2021,2022)
	Structure Similarity Coefficient							
	0.98	0.97	0.95	0.94	0.98	0.98	0.98	0.98
Age groups	Significant Relative Changes							
70-74	2.21	2.37	2.82	2.12	2.21	1.91	1.51	
90+	1.02			1.35				2.19

Source: own elaboration.

The results presented in Table 2 confirm that the share of structure indicators of death in the entire population increased significantly in 2019-2021 compared to 2018 in the age group 70-74. Additionally, we observe a significant increase in the share of the oldest people (90+) in the structure of deaths in 2022 (both compared to 2018 and 2021).

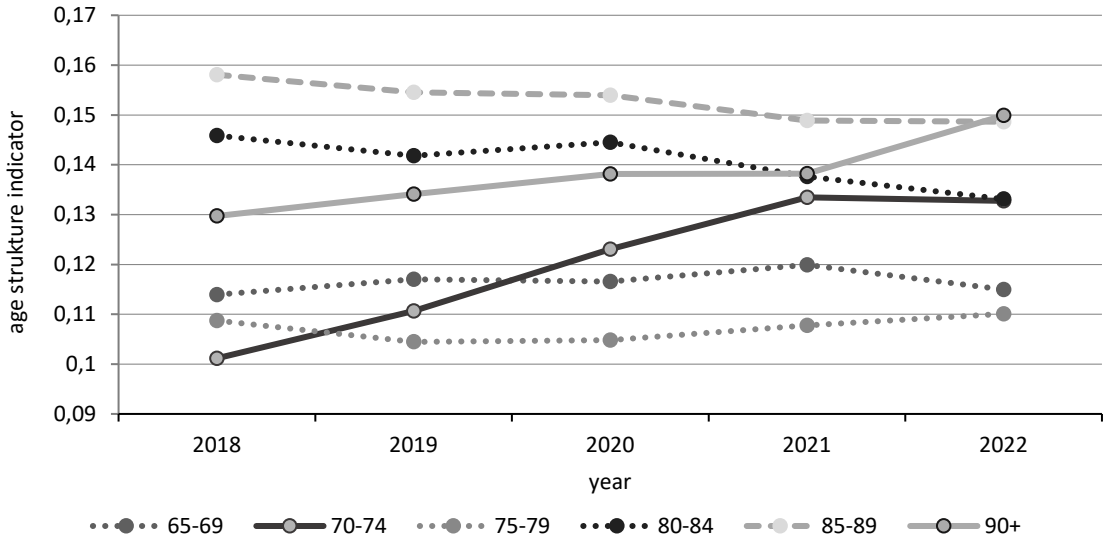

Figure 2. Age structure indicators of death for age groups in 2018-2022

Source: own elaboration.

Note that in the age groups 70-74 and 90+ the dynamics of changes in the age structure indicators of death is greater than in other age groups, which is presented in the graphs of Figure 2. Greater increases are observed for the age group 70-74, which coincides with the huge values of significant relative changes (com. Table 2).

An additional element of the analysis is to examine the impact of the Covid19 pandemic on population forecasts. We compared the actual age structure in 2020-2022 with the forecast based on data from 2013-2019 to capture distinctive changes. The relative differences between the actual values of the age structure indicators and the forecasted value of the age structure indicators for a given year are illustrated in Figure 3. Note that only in the age group 70-74 do we observe significant relative changes in all years (dots on the graph outside the grey area between -1 and 1, which indicates relative changes which as not significant). This means that the actual values are significantly higher than those forecasted based on the years 2013-2019.

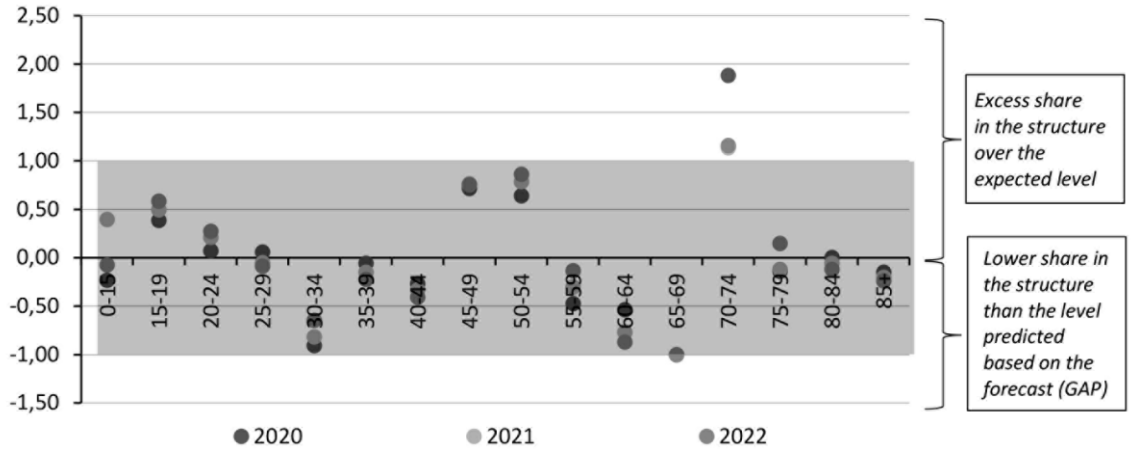

Figure 3. Relative differences between the actual age structure elements in 2020-2022 and those forecasted based on data from 2013-2019

Source: own elaboration.

In summary, the applied method for detecting distinctive changes in the structure indicates that the 70-74 age group is particularly sensitive to the changes that occurred between 2018 and 2022. Although this method does not allow for identifying the reasons why changes in the structural indicators of this group are more significant than those in other groups, it strongly suggests that changes in this age group warrant closer examination. Demographic processes can be analyzed from different perspectives (e.g., the age structure of the population is correlated with the age structure of deaths, and both of these factors affect the accuracy of demographic forecasts). However, the fact that this method consistently highlights the same age group, regardless of the specific issue analyzed, suggests that this may be an important area for further research.

Generally detecting significant structural changes enables a better understanding of the world around us and the processes that influence it. This facilitates adaptation to changing conditions and informed decision-making. Changes in the age structure of society can impact various areas of social life, including healthcare, retirement, and economic systems. Further research is needed to comprehend better these changes and their potential consequences for society and the economy.

Keywords: Structural tables, identification of changes in structure, demographic analysis

References

Andrec, M., Snyder, D., Zhou, Z., Young, J., Montelione, G., & Levy, R. (2007). A Large Data Set Comparison of Protein Structures Determined by Crystallography and NMR: Statistical Test for Structural Differences and the Effect of Crystal Packing. *Proteins Structure Function and Bioinformatics*, *69*(3), 449-465.

Beauchamp, J., Johnson, R., & Wichern, D. (1983). Applied Multivariate Statistical Analysis. *Technometrics*, *25*(4), 385.

Blanchin, M., Guilleux, A., Hardouin, J., & Sébille, V. (2019). Comparison of Structural Equation Modelling, Item Response Theory and Rasch Measurement Theory-based Methods for Response Shift Detection at Item Level: A Simulation Study. *Statistical Methods in Medical Research*, *29*(4), 1015-1029.

Bongaarts, J. (2004). Population Aging and the Rising Cost of Public Pensions. *Population and Development Review*, *30*(1), 1-23.

Dosselmann, R., & Yang, X. (2009). A Comprehensive Assessment of the Structural Similarity Index. *Signal Image and Video Processing*, *5*(1), 81-91.

Johnston, C., Crosnoe, R., Mernitz, S., & Pollitt, A. (2019). Two Methods for Studying the Developmental Significance of Family Structure Trajectories. *Journal of Marriage and Family*, *82*(3), 1110-1123.

Kończak, G., Kosińska, M. (2023). O testowaniu istotności różnic w strukturach populacji na podstawie prób o małych liczebnościach. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, (3), 145-160.

Manton, K. G., & Gu, X. (2001). Changes in the Prevalence of Chronic Disability in the United States: 1982-1999. *Public Health Reports*, *116*(1), 1-12.

National Institute on Aging. (2017). *Aging and Health: A Global Perspective*.

Sei, Y., & Ohsuga, A. (2021). Privacy-preserving Chi-squared Test of Independence for Small Samples. *BioData Mining*, *14*(1).

Sokołowski, A. (1993). Propozycja testu podobieństwa struktur. *Przegląd Statystyczny*, *40*(3-4), 295-301.

Sheskin, D. (2004). *Handbook of Parametric and Nonparametric Statistical Procedures* (3rd ed). Chapman & Hall/CRC.

Smith, J. P., & Kington, R. S. (1997). Demographic and Economic Correlates of Health in Old Age. *Demography*, *34*(1), 1-20.

Wan, H., Sengupta, M., Velkoff, V. A., & DeBarros, K. A. (2005). *Current Population Reports, P23-209,65+ in the United States: 2005*. U.S. Government Printing Office.

World Health Organization. (2021). *World Report on Ageing and Health*.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 22(28)

Some Remarks on the Correlation of Statistical Tests

Karlheinz Fleischer

Research group statistics, Philipps-University of Marburg, Germany

e-mail: k.fleischer@wiwi.uni-marburg.de

Summary

In empirical analyses we usually have just one sample for all analyses, we want to do. Usually, several statistical tests are carried out with this sample. Hence, we expect that the sample results are not independent of each other. We are interested in the correlation coefficient between the test statistics and between p-values of two tests, a test of the mean of a normal population with unknown variances and a test of the variance of the population and we will investigate some properties.

Problem

Let us assume that we draw a sample of length n from a normal distribution with mean μ and variance σ². First the sample is used to test the hypothesis $H_0^{(1)}$: $\mu = \mu_0$, where the variance σ^2 is unknown, and secondly to test the hypothesis $H_0^{(2)}$: $\sigma^2 = \sigma_0^2$.

The test statistics for these tests are

$$T_1 = \frac{\overline{X}-\mu_0}{S}\sqrt{n} \qquad \text{and} \quad T_2 = \frac{(n-1)S^2}{\sigma_0^2}$$

with sample mean $\overline{X}$ and sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$.

We are interested in the strength of the connection between these tests. Especially, we are interested in the correlation coefficient between the test statistics and between the p-values of both tests.

Some simulation results concerning test statistics

Starting with 1 million simulations we recognize first, that the correlation depends on the sample size n (see figures 2 and 3), but seems to be the same if the ratio $\frac{\mu_0-\mu}{\sigma}$ is unchanged (see figure 1). Furthermore the correlation seems to increase with the sample size and tends to a limit very quickly (figures 2 and 3).

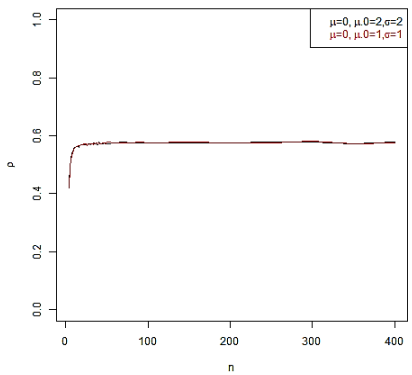

Figure 1. Correlation of test statistics for (μ.0-μ)/σ=1

Source: own elaboration.

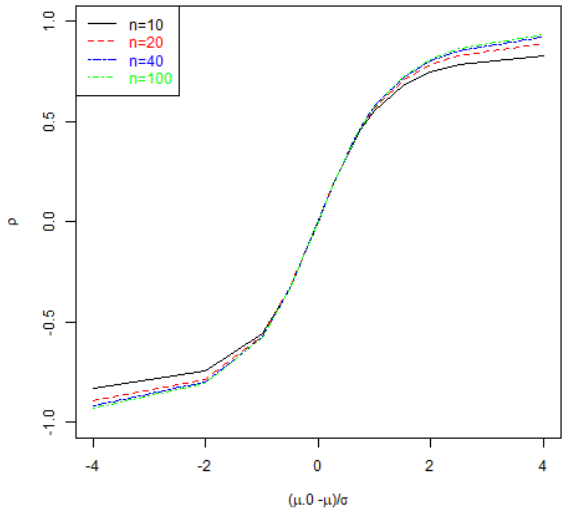

Figure 2. Correlation of test statistics for different n

Source: own elaboration.

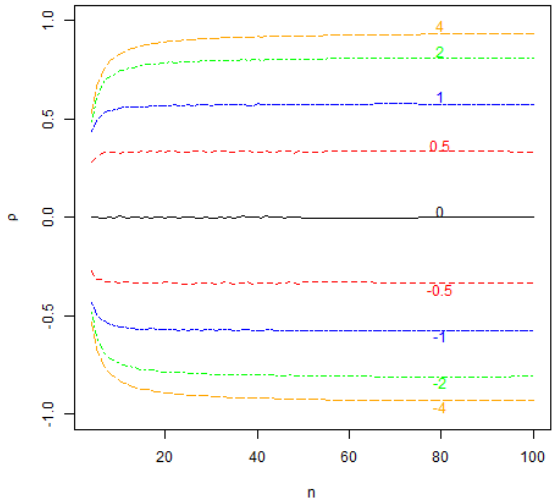

Figure 3. Correlation of test statistics for different (μ.0-μ)/σ

Source: own elaboration.

Figure 3 shows that the limit seems to depend on the ratio $\frac{\mu_0-\mu}{\sigma}$ only, where the sign of the correlation coefficient depends on the sign of $\mu_0 - \mu$ and it is symmetrically. The larger the absolute value of the ratio the larger is the (absolute value of the) correlation coefficient. If the first hypothesis is true the correlation coefficient seems to be 0 for all sample sizes.

Additionally, further simulations show that the correlation coefficient of the test statistics does not depend on whether the second hypothesis is true or false.

Theoretical result

Now, we try to calculate the exact correlation coefficient of the test statistics.

Using the facts that $\overline{X}$ and S^2 are independent, when sampling from a normal distribution, and that S is χ −distributed with n-1 degrees of freedom, we can calculate the correlation coefficient between the test statistics T_1 and T_2. We find

$$\rho_{T_1,T_2} = \frac{\frac{\mu_0-\mu}{\sigma}}{\sqrt{2}\cdot\sqrt{\frac{(n-1)(n-2)^2}{n(n-3)}\left(\frac{c_n}{c_{n-1}}\right)^2 + \frac{(\mu_0-\mu)^2}{\sigma^2}\left(\frac{(n-1)(n-2)^2}{n-3}\left(\frac{c_n}{c_{n-1}}\right)^2 - (n-1)\right)}},$$

where $c_n = \frac{1}{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)}$ with gamma function $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ (for $x > 0$).

This formula shows some of the observed results: The correlation coefficient does not depend on the real values of μ, μ_0, σ^2 or σ_0^2, but only on the ratio $\frac{\mu_0-\mu}{\sigma}$. It is symmetrically and the values of the correlation coefficient can be calculated for fixed ratios $\frac{\mu_0-\mu}{\sigma}$ and fixed sample sizes n, e.g., for $\frac{\mu_0-\mu}{\sigma} = 1.5$ we find $\rho_{T_1,T_2} \approx 0.6762$ for $n = 10$, $\rho_{T_1,T_2} \approx 0.7065$ for $n = 20$ and for $n = 50$ it holds $\rho_{T_1,T_2} \approx 0.7201$. The exact correlation coefficient of the test statistics for different values of $\frac{\mu_0-\mu}{\sigma}$ between -4 and 4 and for $n = 5, 10, 15, \ldots, 100$ is visualized in figure 4.

From figures 2, 3 and 4 we can observe that the correlation coefficient seems to increase with the ratio $\frac{\mu_0-\mu}{\sigma}$, and seems to increase only slightly with the sample size for a given ratio.

The question arises, if it is possible to prove the convergence of the correlation coefficient of the test statistics for increasing sample sizes and if we are able to determine the limit. The behavior of the correlation coefficient depends on the behavior of $\left(\frac{c_n}{c_{n-1}}\right)^2$. It is possible to prove the following limits for n →∞:

$$\left(\frac{c_n}{c_{n-1}}\right)^2 \longrightarrow 0\,, n\left(\frac{c_n}{c_{n-1}}\right)^2 \longrightarrow 1, \frac{(n-1)(n-2)^2}{n(n-3)}\left(\frac{c_n}{c_{n-1}}\right)^2 \longrightarrow 1.$$

For the proof of these statements some properties connected with the Wallis product $\prod_{i=1}^{\infty}\frac{(2i)^2}{4i^2-1} = \frac{\pi}{2}$ are helpful.

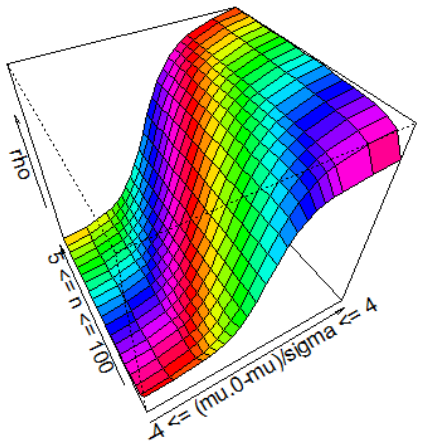

Figure 4. Correlation for different (mu.0-mu)/sigma and n

Source: own elaboration.

Furthermore, simulations suggest that $\frac{(n-1)(n-2)^2}{n-3}\left(\frac{c_n}{c_{n-1}}\right)^2 - (n-1) \longrightarrow \frac{1}{2}$, which would lead to the following limit for the correlation coefficient

$$\rho^{\lim}_{T_1,T_2} = \frac{\frac{\mu_0-\mu}{\sigma}}{\sqrt{2}\cdot\sqrt{1+\frac{1}{2}\cdot\frac{(\mu_0-\mu)^2}{\sigma^2}}} = \frac{\frac{\mu_0-\mu}{\sigma}}{\sqrt{2+\frac{(\mu_0-\mu)^2}{\sigma^2}}} = \frac{sgn(\mu_0-\mu)}{\sqrt{1+\frac{2}{\frac{(\mu_0-\mu)^2}{\sigma^2}}}}$$

(last equation only if $\mu_0 \neq \mu$). Here sgn is the sign-function $sgn(x) = -1$, if $x < 0$, $sgn(0) = 0$, $sgn(x) = 1$, if $x > 0$.

Figure 5 shows the potential limits for ratios $\frac{\mu_0-\mu}{\sigma}$ between -4 and 4. E.g., for $\frac{\mu_0-\mu}{\sigma} = 1.5$ the limit would be 0.7276, hence, only about 2.90% higher than the real correlation coefficient of 0.7065 for $n = 20$, and only about 1.03% higher then the real value 0.7201 for $n = 50$.

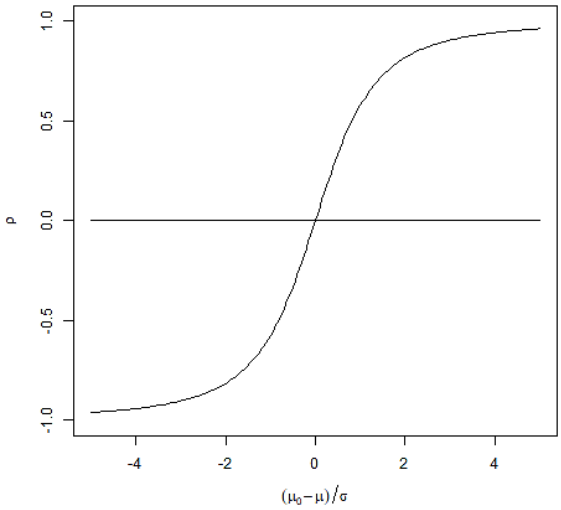

Figure 5. Limit of Correlation dependent of (μ0-μ)/σ

Source: own elaboration.

Unfortunately, the proof of the limit function is still missing, only simulations support that result until now.

Simulations concerning the p values and test results

Now, let's have a look on the p values of both tests. Exact calculations are difficult, hence, we rely on simulations, again.

The p values are mainly negatively correlated, but only of moderate size and nearly 0 if the first hypothesis is only slightly false and the second one is true (see figures 6 and 7). For increasing sample sizes the correlation coefficients seem to tend to zero (figure 7) and, furthermore, the p values seem to be identical if only the sign of the ratio $\frac{\mu_0-\mu}{\sigma}$ changes (figure 7).

The bivariate dependence of the correlation coefficient from the ratio $\frac{\mu_0-\mu}{\sigma}$ and the sample size n is illustrated in figure 8. Even if 1 million simulations are carried out, the correlation coefficient is estimated very inaccurately. One problem of these simulations is that, especially for larger sample sizes, the standard deviations of the p values are very small, such that very often they are estimated as 0, which leads to missing values during the calculations.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 22(28)

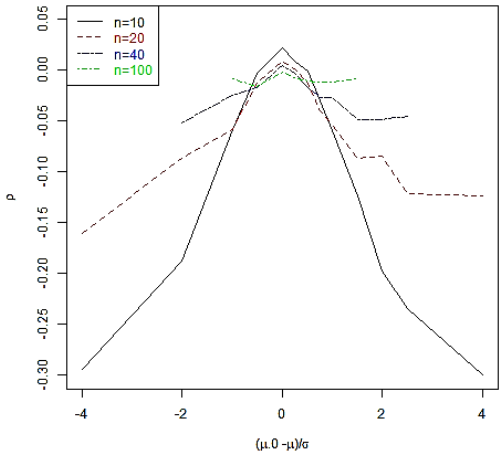

Figure 6. Correlation of p values for different n

Source: own elaboration.

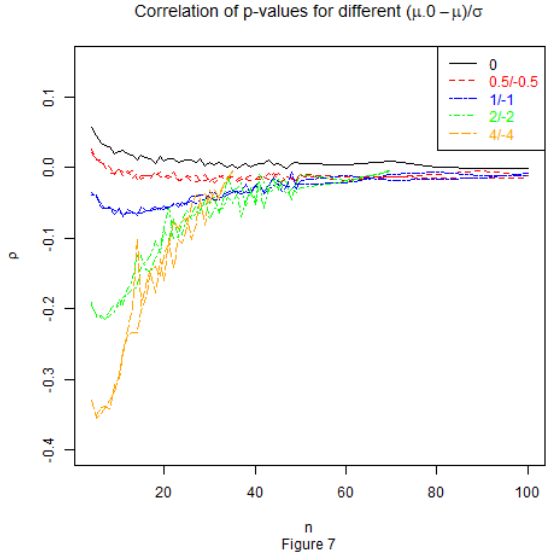

Figure 7. Correlation of p values for different (μ.0-μ)/σ

Source: own elaboration.

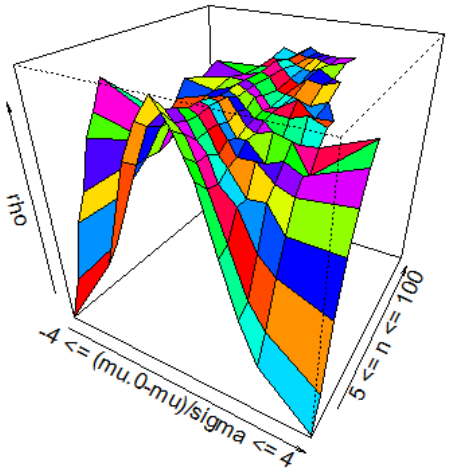

Figure 8. Correlation for different (mu.0-mu)/sigma and n

Source: own elaboration.

Furthermore, we can simulate the connection of the test decisions of both tests (see Table 1).

Table 1. Test decision

$\mu = \mu_0 = 0,$ $\sigma = \sigma_0 = 1,$ $n = 20, \alpha = 0.05$	$H_0^{(2)}$ rejected	$H_0^{(2)}$ not rejected
$H_0^{(1)}$ rejected	0.0048	0.0451
$H_0^{(1)}$ not rejected	0.0451	0.9051
$\mu = 0,\ \mu_0 = 1,$ $\sigma = \sigma_0 = 1,$ $n = 20, \alpha = 0.05$	$H_0^{(2)}$ rejected	$H_0^{(2)}$ not rejected
$H_0^{(1)}$ rejected	0.0488	0.9399
$H_0^{(1)}$ not rejected	0.0015	0.0099
$\mu = 0,\ \mu_0 = 1,$ $\sigma = \sigma_0 = 1,$ $n = 100, \alpha = 0.05$	$H_0^{(2)}$ rejected	$H_0^{(2)}$ not rejected
$H_0^{(1)}$ rejected	0.0503	0.9497
$H_0^{(1)}$ not rejected	0.0000	0.0000

Source: own elaboration.

E.g., for $\mu = \mu_0 = 0, \sigma = \sigma_0 = 1, n = 20, \alpha = 0.05$, the probability that, e.g., $H_0^{(1)}$ is rejected and $H_0^{(2)}$ is not rejected is estimated as 4.51%. The test statistics are uncorrelated in this case, both hypotheses are true, hence, for both hypotheses the probability of an error of the first kind is $\alpha = 0.05$. These probabilities are estimated as $0.0048 + 0.0451 = 0.0499$ quite well for both hypotheses. The probability that both hypotheses are rejected simultaneously is about 0.5%. If the test statistics would be independent, this probability should be 0.25%.

The last two tables contain the probabilities, if the first hypothesis is false and second one is true. Other situations can be simulated in the same manner.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 22(28)

Freight Rate Options in Container Shipping

Albert Gardoń

Department of Statistics, Wroclaw University of Economics and Business, Poland

e-mail: albert.gardon@ue.wroc.pl

The container shipping market is highly competitive. In such markets the psychology of many independent competitors plays a crucial role and hence, random models become adequate for describing the behavior of the market. Models assuming the normality of real data are widely used in many disciplines, first and foremost in finances, where the price processes are represented by the geometrical Brownian motion (the famous Black-Scholes model – see Karatzas, Shreve, 1998). Unfortunately, in last years this purely gaussian approach has been criticized because of heavy tails of empirical data distributions which contradict the normality (see Andersen, Bollerslev, Diebold, 2007; Cont, Tankov, 2004; Das, 2002; Johannes, 2004). However, the researchers have tried to overcome this problem by creating models based on the general jump-diffusion processes where large price changes have been extracted from the empirical distribution and modeled separately as jumps driven by the Poisson process. Differential equations describing such models must be usually solved by means of numerical methods (see Gardoń, 2004 and 2006). But it helps neither because the empirical distributions of a financial data, even after extraction of jumps, are asymmetrical or too slender, that means they have got significantly different skewness or higher kurtosis than in the normal case (see Gardoń, 2011; Peiró, 1999).

In the case of the container shipping industry the problem appears as well but fortunately there are parts of this market where the jump-diffusion model may be relevant (see Gardoń, 2014). This model is defined by the following stochastic differential equation in the integral form:

$$X_t = X_{t_0} + \int_{t_0}^{t} aX_s ds + \int_{t_0^+}^{t} b\bar{X}_s dW_s + \int_{t_0^+}^{t} \bar{C}_s \bar{X}_s dN_s \quad , \mathrm{t} > \mathrm{t}_0, \text{ a.s.} \tag{1}$$

where the modeled process *X* denotes the weekly average net freight (or rate, which is a transportation price consisting of basic ocean freight and different surcharges like e.g. fuel surcharge) per transported unit (TEU or FFE — the volume of a 20 or 40 feet long standard container), $\bar{X}_t = X_{t^-} = \lim_{s \to t} X_s$, *W* is a standardized Wiener process, *N* is a homogeneous. Poisson process with intensity λ and the both driving processes are said to be independent. Coefficients *a*, *b* and *C* (overlined C_s means the left-hand side limit at s, as for X) are called here the drift, the volatility and the relative jump size, respectively. Further, *C* denotes a process with almost all paths partially constant, changing its values at the jump points of the Poisson process *N,* but with values independent from the both driving processes *W* and *N* (see Mancini, 2009). This means the jump sizes are realizations of an independent identically distributed (iid) sequence of random variables (C_{Γ_i}) which may be treated as independent copies of a certain random variable *C*, whereas the jump times of the process *N* are denoted by (Γ_i). Additionally, the sequence (τ_i) will represent times when the process is observed.

We would like to introduce the idea of derivatives into the container shipping market (see Karatzas, Shreve, 1998). The idea is strongly considered in recent years (see Koekebakker, Adland, Sødal, 2007; Nomikos at al, 2013). Precisely speaking, a simple bilateral EuroCall options. European options are a basic example of derivatives. They give the owner the right for buying (Call) an underlying good at the fixed time (expiration date *T*) in the future for a fixed price (strike

price *K*). This underlying good is usually an asset or a market index but it may be also the shipping service cost.

Another type of derivatives is already widely used in the shipping industry – forward/future contracts. They are rather informal and consist in long term agreements between shippers (commodity owners) and carriers (vessel owners). Unfortunately, such contracts are unreliable and therefore cannot be an effective hedging tool for the risk reduction. If the average market freight rate changes significantly, especially if it drops, the shippers strongly insist the carriers for renegotiation and reducing the previously contracted prices. Since the market is competitive the carriers have no other choice as to decline rates in order not to lose customers.

The idea of the bilateral EuroCalls should resolve the problem of price renegotiations in the case of freight drop. The cost of an option is called the premium and is calculated as the discounted expected value of the future profit with respect to the martingale probability measure (see Karatzas, Shreve, 1998; Nomikos et al, 2013). The premium is paid by an option buyer at the time of the enter into the agreement and is usually a tiny part of the strike price. In the instance of a freight decrease in the future a shipper which buys the EuroCall will have the possibility for shipping the commodities for the market price lower than the strike price and resigning from the option. A carrier will earn in this case less for the transportation but it will have received an additional income for issuing the EuroCall compensating the freight rate drop. On the other hand, if the rates will increase, a shipper will have the guarantee that the cargo will be shipped for a previously agreed lower strike price. A carrier will gain in this instance less than it could have regarding actual market rates, though, it will still receive the satisfactory strike price from the option plus the premium already gained when the EuroCall was issued. This operation should reduce the freight change risk for both contractors, hence, should be advantageous for both parties.

As it was mentioned above the net premium of a EuroCall is generally calculated as the discounted expected value of the future profit with respect to the objective (risk free) martingale probability measure Q, i.e.:

$$O_t = E_t^Q\left(e^{-\rho(T-t)}\, max(X_t - K,\ 0)\right)\ ,\ t > t_0, \tag{2}$$

where *X* is the net freight process, *T* is the expiration time, *K* is the strike price, *ρ* is the riskless rate and EURIBOR is usually taken as such a rate. *Q* is called a pricing probability measure such that the net freight process *X* is a local martingale with respect to *Q*. The measure *Q* realizes the so-called no arbitrage assumption (see Karatzas, Shreve, 1998) which says that there cannot be an opportunity on the market for a riskless relative profit greater than *ρ*. In practice, this assumption requires that in Eqn. (1) the drift coefficient $a = \rho - \lambda\, EC$ (see Kou, 2002).

If in the Eqn.(1) the jump component vanishes, i.e. if *C=0* a.s., then the jump-diffusive equation reduces to the well-known Black-Scholes model (see Karatzas, Shreve, 1998) which is widely used for pricing the derivatives on financial markets. In this instance the explicit formula for the EuroCall net premium is known:

$$O_t = \Phi(d_+)X_t - \Phi(d_-)Ke^{\rho(T-t)}, \text{ where } d_\pm = \frac{log\frac{X_t}{K} + \left(\rho \pm \frac{b^2}{2}\right)(T-t)}{b\sqrt{T-t}} \tag{3}$$

and *Φ* is the standard normal distribution function. Since this purely gaussian approach has been criticized in last years because of heavy tails of empirical distributions in our article we will compare the premium evaluation (3) to the premium calculated based on the jump-diffusive model (1). In this latter case it cannot be calculated explicitly and it will be simulated by means of the Monte Carlo method. Hence, the jumps have to be identified and the model needs to be calibrated.

We decided to conduct the calculations based on the real data from one of the leading carriers. The data set consists of 656 weekly net freight returns from the over 12 years long time period from January 2 2000 to August 5 2012 concerning the most important and competitive South-East Asia – Europe trade, head haul (more profit-yielding) direction (see Fig. 1). The data seems to be normally distributed apart from a set of outliers and the jump-diffusive model seems to be relevant in this instance (see Gardoń, 2014). The Black-Scholes model requires that the discretized relative price changes:

$$Z_n = \frac{X_{\tau_n} - X_{\tau_{n-1}}}{X_{\tau_{n-1}}} = \frac{\Delta X_{\tau_n}}{X_{\tau_{n-1}}} \approx ln \frac{X_{\tau_n}}{X_{\tau_{n-1}}} \ , \qquad n = 1, \dots, L \text{ , a.s.}$$

where L is the number of observations, are realizations of a normally distributed random variables sequence. But since (τ_n) is not said to be equidistant the data must be firstly standardized:

$$\forall n = 1, \dots, L \qquad Z_n^* = \frac{Z_n - a(\tau_n - \tau_{n-1})}{b\sqrt{\tau_n - \tau_{n-1}}} \overset{A}{\approx} N(0,1), \text{ a.s.}$$

The symbol $\overset{A}{\approx}$ means the asymptotical distribution. Hence, the drift and volatility parameters *a* and *b* must be firstly estimated. However, as it is already mentioned the so-called no arbitrage insists on taking $a = \rho - \lambda\, EC$. On November 19 2012 corresponding to the data set this riskless rate was equal to 0.86% p.a. for 1 year USD contracts which implies for the time unit 1 week: $\rho = 0{,}0165\%$. On the contrary, the volatility may be estimated in the maximal likelihood sense by the standard deviation of normalized returns:

$$\hat{b} = \sqrt{\frac{1}{L}\sum_{n=1}^{L} \frac{Z_n^2}{\Delta\tau_n} - \left(\frac{1}{L}\sum_{n=1}^{L} \frac{Z_n}{\sqrt{\Delta\tau_n}}\right)^2} \ , \tag{4}$$

which follows immediately from the fact, that the relative price changes (Z_n) have to be normally distributed and *b* is in such an instance the infinitesimal variance of the normalized return.

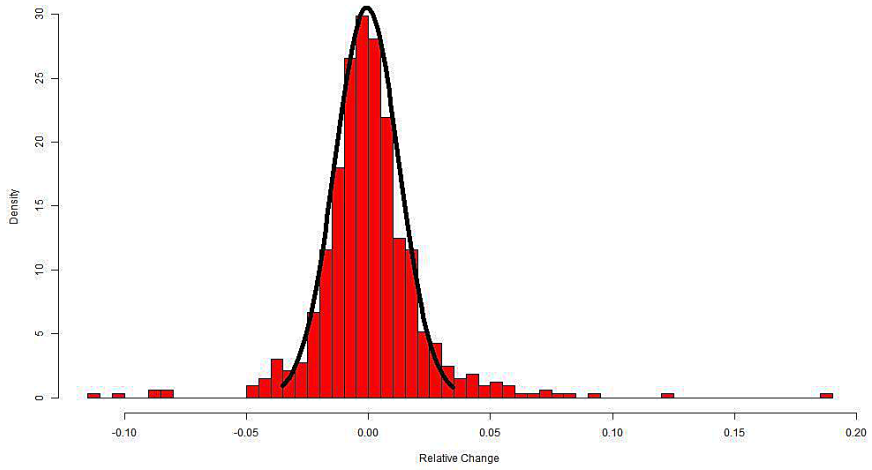

Figure 1. The normal vs empirical density of the relative rate changes

Source: own elaboration.

In the Figure 2 it is shown how the net premium based on the jump-diffusion model depends on the strike price and the time to expiration, when the actual freight rate is 3000 USD/FFE. Each case was evaluated by means of 1000 artificially generated paths of the net freight process *X*. Even with the strike price about 10% lower than the actual net freight the net cost of the EuroCall has been 10 – 15% of the actual rate depending on the time to expiration. More precisely it is presented in Figure 3 where the premium of the 6 months ahead EuroCall depends only on the strike price.

In the instance of the jump-diffusive model the jumps have to be recognized firstly and extracted from the remaining "continuous" data. One of possible ways is the threshold method (see Mancini, 2009; Gardoń, 2011) with the threshold condition $\frac{Z_n^2}{\hat{b}^2} > r(\Delta\tau_n)$, where the threshold function $r(t) = \beta t^{1-\varepsilon} = 7{,}3576t^{0{,}9}$ and $\hat{b}$ are evaluated by means of an iterative procedure based on Eqn.(4) regarding to the remaining "continuous" (still not recognized as jumps) returns at each step. Eventually, the estimation of the Poissonian intensity λ has been evaluated for $\hat{\lambda} = 0{,}08$ and the relative jump sizes have been simulated directly from the empirical distribution. The results following from the Black-Scholes model are not presented in figures because this part of investigation was conducted just for comparison. We only sum up that the premium in this instance has been usually 50 – 100 USD lower than in the jump-diffusive case, especially for strike prices strongly differing from the initial rate. This is not surprising since the gaussian models underestimate the probability for outliers.

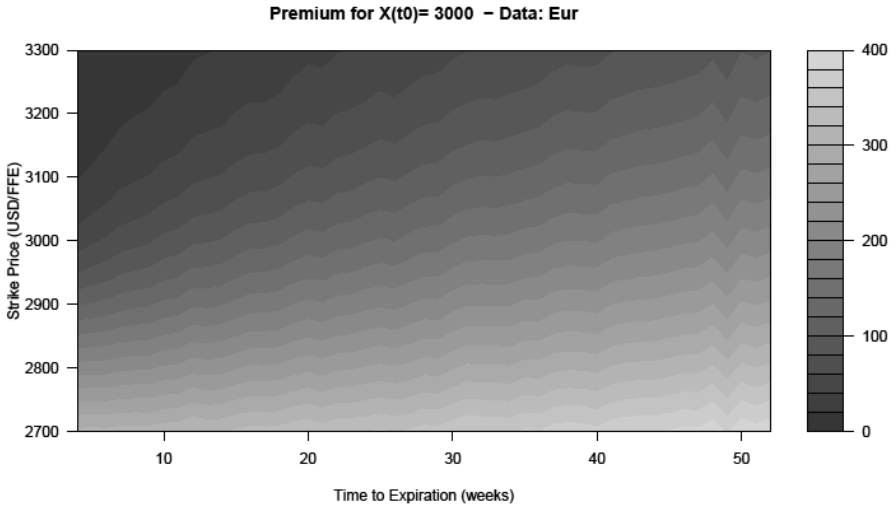

Figure 2. The net premium of the EuroCall for the net freight from South-East Asia to Europe with the initial freight rate 3000 USD/FFE, depending on the strike price and the time to expiration

Source: own elaboration.

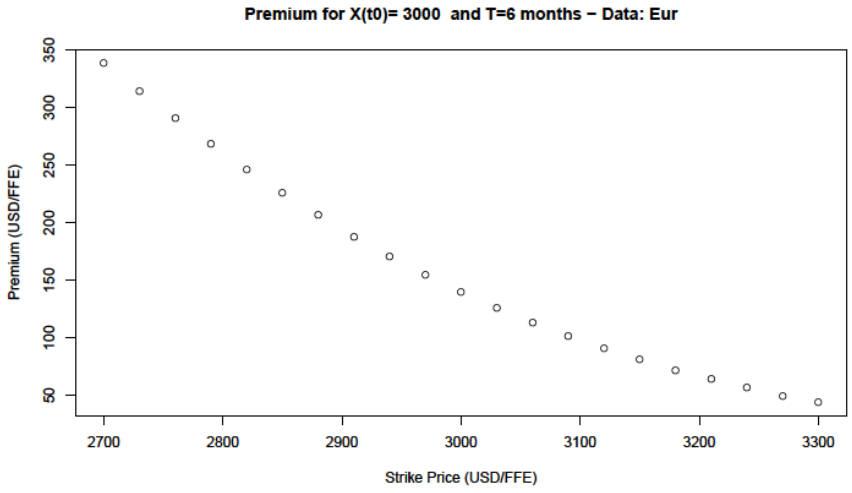

Figure 3. The net premium of the European Call for the 6 months ahead net freight from South-East Asia to Europe with the initial freight rate 3000 USD/FFE depending on the strike price

Source: own elaboration.

Eventually, we conducted a special simulation for checking how much a company would lose or gain additionally if it issued EuroCalls for the entire data period, i.e. from the beginning of the current millennium. The scenario consists of several assumptions:

- every week the EuroCalls were issued for 10% of shipped containers;
- all EuroCalls were issued for the time horizon of 3 months=91days=13 weeks;
- all EuroCalls were issued with the strike price $K = X_{t_0}$;
- EuroCalls were issued only if the sum of the strike price and the premium exceeded the minimal profit-yielding level;
- to each premium the 10% commission is added but not less than 10 USD.

Further, we have not included the costs of the option issuing. We have believed they are negligible, especially in a longer time horizon, because the way an option contract would be agreed does not differ from the way the parties usually enter into a forward/future contract. The week by week additional profit or loss is drawn in Figure 4. Already at the first look it is visible that this profit/loss behaves more stable in the first half of the time period investigated. There is a simple explanation of this fact, namely the volume of transported cargo which has been increasing systematically by 5 – 10% yearly within 13 years. The increasing absolute amount of possibly issued Calls enlarged the volatility of the profit in the second half of observations.

Generally, the issuing of EuroCalls would have brought the carrier a slight additional profit in tough years of a price decline. On the other hand it would lead to quite dramatic losses exceeding even 1 million USD weekly during good times when the freights exploded. In those weeks the cargo would have had to be shipped for a much lower option strike price than the actual market net freight. The crucial question is what is the total impact of EuroCalls for the company profit during the entire investigated time period. The accumulated profit/loss is shown in Figure 5 and looks extremely optimistic. It is not monotone increasing, though, it has never appeared negative. Moreover, after almost 13 years of application of the EuroCalls it would deliver the additional profit equal to about 100% of the total actual profit in the year 2011 which has been the last whole calendar year during the period of investigation. It is equivalent to about 8% of an extra income yearly.

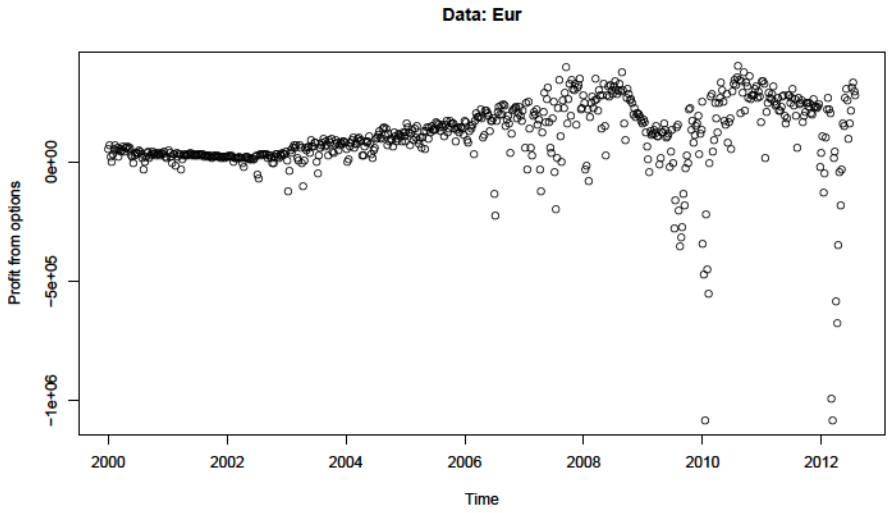

Figure 4. The simulated weekly additional company profit or loss corresponding to the European Call options issued for the net freight

Source: own elaboration.

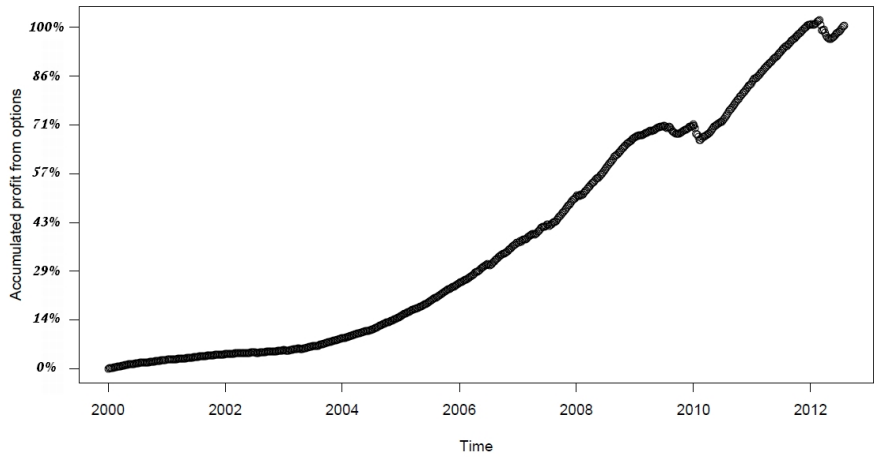

Figure 5. The simulated accumulated additional company profit corresponding to the European Call options issued for the net freight expressed by the percentage of the total profit in the year 2011 (the last whole year in the period investigated)

Source: own elaboration.

Although the scenario discussed in the previous section is rather conservative, it would give the carrier issuing the EuroCalls for the net freight a significant additional profit. What is more, it is distributed in such a way that the EuroCall premiums generate an additional profit in the time of price decline (when the EuroCalls are not executed by shippers) whereas in the time of price increase the options executed by shippers imply the lowering of the profit for a carrier. Hence, the options decrease the volatility of the weekly net freight making the income more stable and limiting the risk of a carrier. On the other hand, the EuroCalls assure the upper bound for the transportation price paid by a shipper limiting its risk as well. Summing it up, this type of contract between a shipper and a carrier should be attractive for both parties.

The investigation is conducted based on the data from a chosen company only, though, since this part of the market is significantly competitive, the competitors' freight rates should be quite similar one to another. Thus, the results described may be useful for each company operating in this industry.

Acknowledgement

This research was funded by the Danish Maritime Fund, through grant no. 2011-58.

References

Andersen, T., Bollerslev, T., & Diebold, F. (2007). Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility. *Review of Economics and Statistics*, *89*(4), 701-720.

Cont, R., & Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC.

Gardoń, A. (2004). The Order of Approximations for Solutions of Itô -type Stochastic Differential Equations with Jumps. *Stochastic Analysis and Applications*, *3*(22), 679-699.

Gardoń, A. (2006). The Order 1.5 Approximations for Solutions of Jump-Diffusion Equations. *Stochastic Analysis and Applications*, *6*(24), 1147-1168.

Gardoń, A. (2011). The Normality of the Financial Data after an Extraction of Jumps in the Jump-Diffusion Model. *Mathematical Economics*, *7*(14), 79-92.

Gardoń, A. (2014). The Normality of Weekly Relative Changes of the Freight Rate in Container Shipping. *Conference Proceedings of the 17-th International Scientific Conference Applications of Mathematics and Statistics in Economics*, (95-103).

Johannes, M. (2004). The Statistical and Economic Role of Jumps in Continuous-time Interest Rate Models. *Journal of Finance, 59*(1), 227-260.

Karatzas, I., & Shreve, S. E. (1998). *Methods of Mathematical Finance*. Springer.

Koekebakker, S., Adland, R., & Sødal, S. (2007). Pricing Freight Rate Options. *Transportation Research Part E, 43*(5), 535-548.

Kou, S. G. (2002). A Jump-Diffusion Model for Option Pricing. *Management Science*, *48*(8), 1086-1101.

Mancini, C. (2004). Estimation of the Parameters of Jump of a General Poisson-Diffusion Model. *Scandinavian Actuarial Journal, 2004*(1), 42-52.

Mancini, C. (2009). Non Parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics*, *36*(2), 270-296.

Nomikos, N. K., Kyriakou, I., Papostolou, N. C., & Pouliasis, P. K. (2013). Freight Options: Price Modelling and Empirical Analysis. *Transportation Research Part E*, *51*(C), 82-94.

Peiró, A. (1999). Skewness in Financial Returns. *Journal of Banking and Finance, 23*(6), 847-862.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 22(28)

Economic Discomfort in the FRG 1951 to 2021 – a Critical Analysis of the Misery Index

Ullrich Heilemann
Roland Schuhr
University of Leipzig, Germany
e-mail: schuhr@wifa.uni-leipzig.de>

To provide a summary of the state of the US economy, in 1971 Arthur Okun defined an index of "economic discomfort" as the sum of the rates of unemployment and inflation. The exact purpose and basis of the indicator was not specified, nor was its structure. However, this did not hinder the increasing popularity of the Misery Index (MI), as it soon became known: With inflation and employment, it is based on statutory targets, it uses currently available official data and it is very easy to calculate. Since its introduction, MI has been used not only to assess the state of the economy, but also to assess the performance of US presidents and to predict election results.

Economic discomfort and its determinants are subjects of research in many fields of the social sciences, which led to numerous modifications and extensions of Okun's formula. For the US, for example, these include house and share prices, the long-term interest rate, economic growth or real per capita income. Okun himself was open to replacing the unemployment rate with income in his formula and he also referred to Gallup polls showing the importance of the two components fluctuate according to levels (Okun, 1973, 1976). However, varying substitution rates in the preference function, as MIs are also interpreted, were not given much attention until the 2000s. To check MI's selection and weighting of variables, MI's were confronted with similar other indicators. For the US, these were the University of Michigan's Consumer Sentiment Index (Lovell & Tien, 2000) and for European countries, the Eurobarometer (Di Tella et al. 2001, 2003; Blanchflower et al. 2014; Welsch, 2007).

For the FRG only very few MI results and government assessments are currently available. This paper fills this gap. We calculate the MI and two variants and test them against the "Policy Barometer Index" (PBI), a survey-based indicator of citizens' satisfaction with government performance. We examine the descriptive power of MIs and their components, their weighting and their stability over time. Then MIs are employed to evaluate governments and finally we ask about their implications for economic policy.

Misery Indices and the Policy Barometer Index

Starting point is Okun's index formular and two extensions

$$\mathrm{MI\ I}_t = a_1 \cdot u_t + a_2 \cdot \mathrm{infl}_t\,, \tag{1}$$

$$\mathrm{MI\ II}_t = a_1 \cdot u_t + a_2 \cdot \mathrm{infl}_t + a_3 \cdot \mathrm{gdp}_t\,, \tag{2}$$

$$\mathrm{MI\ III}_t = a_1 \cdot u_t + a_2 \cdot \mathrm{infl}_t + a_3 \cdot \mathrm{gdp}_t + a_4 \cdot \mathrm{defq}_t\,, \tag{3}$$

where u_t denotes the unemployment rate, infl_t the inflation rate, gdp_t the year-on-year rate of change in real GDP and defq_t the share of government net lending in nominal GDP. The weights are equal to $a_1 = a_2 = 1$ and $a_3 = a_4 = -1$. We calculated annual index values for the period

1951 to 2021 and monthly index values for the period 1977-3 to 2021-9 on the basis of official data from the Federal Employment Agency and the Federal Statistical Office. The annual data (Figure 1) was used for government evaluation, while the monthly data was tested against the PBI by regression analysis. The PBI is based on the "Politbarometer" of the Forschungsgruppe Wahlen. It has been commissioned by Zweites Deutsches Fernsehen (ZDF) since March 1977, making it the longest and most comprehensive non-electoral survey in the FRG. A representative sample is asked at irregular intervals throughout the year to rate the current government's performance on a scale from –5 (strongly disagree) to +5 (strongly agree). The PBI values used are own calculations of monthly averages of the survey results.

Test results

Three sets of regression equations were tested. The first set explores the direct relationships between MIs and PBI, the second set the influence of MI components on PBI. Following Berlemann and Enkelmann (2014) the third set examines the joint influence of MI components and specific extensions in form of event and government dummy variables. One equation in sets two and three also includes the one-period-lag of the dependent variable.

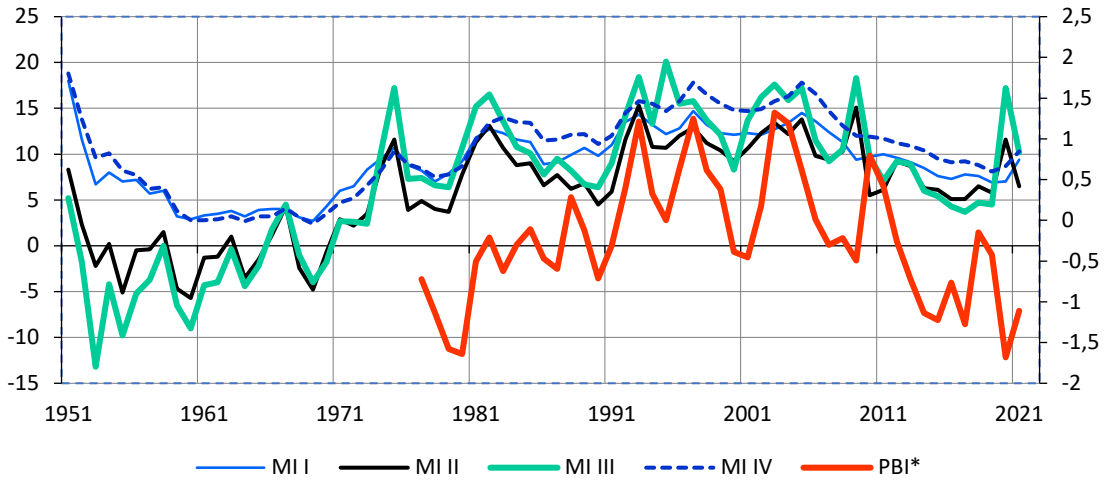

Figure 1. Misery Indices and Policy Barometer Index, 1951–2021. MI I to MI IV: left scale. PBI* = - PBI: right scale

Source: own calculations.

The explanatory power of equations varies widely. All MIs are significant, but fit is best for MI I by considerable margins. In terms of weights, the results are broadly similar. In contrast to the coefficients of unemployment and inflation the coefficients of growth and of deficit ratio are very small and not significant. Unemployment is weighted twice as high as inflation. Event and government variables improve the fit, unemployment and inflation coefficients converge slightly. While the parameters of the two equations with Koyck lags are stable, parameters of most of the other equations vary widely, as OLS CUSUM test results show. Fluctuation in the variables' influences is corroborated by 20-year moving window estimates. However, inflation is only twice and only briefly more important than unemployment.

For the whole period covered, MI I seems to provide a plausible picture of the economic discomfort of FRG citizens. Our results support Okun's choice of variables. However, they do not

support his weighting. PBI reacts almost twice as much to unemployment as to inflation. Moreover, the coefficients are changing over time. To account for the different weights of unemployment and inflation, we calculated MI IV using regression results and the restriction $a_1^* + a_2^* = 2$:

$$\text{MI IV}_t = a_1^* \cdot u_t + a_2^* \cdot \text{infl}_t, \quad (4)$$

where a_1^* = 1.291 and a_2^* = 0.709.

Government evaluation

The performance of German governments in the period from 1951 to 2021 are evaluated by MIs and PBI. We calculated and assigned indicator averages for legislatures (19 periods) and chancellors (8 periods). As the scores are not comparable, we look at rankings. In addition, a variant of the Barro Misery Index (BMI) was calculated. BMI does not focus on the mean values of the components during the term of government, but on their difference from their predecessors (Barro 1999):

$$\text{BMI}_\tau^* = a_1 \cdot \Delta u_\tau + a_2 \cdot \Delta \text{infl}_\tau, \quad (5)$$

where Δu_τ and Δinfl_τ denote the changes during the respective legislature / term-in-office (τ) and $a_1 = a_2 = 1$. All indicators include 2 or 3 of the 4 most important legal objectives of FRG macroeconomic policy since 1967, when the German Stability Law came into force. Results for chancellors are summarised in Table 1.

Table 1. Ranks of chancellors, 1951-2021 (1977-2021)

Chancellors	MI I		MI II		MI III		MI IV		PBI		BMI*	
	Ø	Rank	Ø	Rank	Ø	Rank	Ø	Rank	Ø	Rank	Index	Rank
Adenauer	6.5	4	–0.6	3	–4.4	1	7.6	4		—	–1.6[1]	2
Erhard	3.7	2	–1.3	1	–1.6	2	3	1		—	–0.1	6
Kiesinger	3.3	1	–1	2	–0.2	3	3.2	2		—	–0.7	7
Brandt	6.3	3	2	4	1.5	4	5	3		—	3.5	8
Schmidt	9.5 (9.5)	6 (2)	7.7 (7.5)	5 (1)	10.8 (10.5)	6 (2)	9.4 (9.6)	5 (1)	1.0	— (1)	1.2	7 (4)
Kohl	11.8	7 (3)	9.4	7 (3)	12.1	7 (3)	13.9	7 (3)	–0.1	— (3)	–0.9	3 (2)
Schröder	12.8	8 (4)	11.7	8 (4)	14.4	8 (4)	15.7	8 (4)	–0.4	— (4)	–0.4	5 (3)
Merkel	9.2	5 (1)	7.9	6 (2)	8.8	5 (1)	11.0	6 (2)	0.5	— (2)	–5.3	1 (1)
Ø	9.1		5.9		6.7		10.3		0.2		—	

Source: own calculations. Assignment of terms-in-of office to years by majority of months. Results for 1977-2021 in (...).
[1] Excluding the legislature Adenauer I.

Transitions from MI I to MI II and MI III only lead to minor differences in the ranking – and almost exclusively in the years before 1974. The greatest number of differences with MI I are to be found with MI III, in which Adenauer replaces Kiesinger at the top of the list. The BMI* ranking is more or less the opposite of the MIs. Chancellors Merkel and Adenauer are at the top of the list, while Kiesinger and Brandt are at the bottom. The rankings of MI I and MI IV are very strongly positively correlated with those of the PBI in the period 1977-2021 – perfectly in the case of MI IV. BMI* and PBI, on the other hand, yield contrasting results. It should be noted that the rankings of the legislative periods due to the indices show a noticeably higher variance compared to the chancellor results.

Okun's misery index provides a plausible picture of the economic discomfort of the population of the FRG, at least since the 1970s. While the tests rule out the influences of both growth and the deficit ratio, they also suggest that the unemployment rate is twice as influential as the inflation rate, the latter being broadly in line with international evidence. Both influences have changed over time. Only for short periods, however, inflation has been (slightly) more dominant. Regarding the evaluation of government performance, MIs and PBI lead to different results, as well. They differ particularly with the first and last positions. But again, since the 1970s, the rankings differ little from Okun's index. This is not the case for the BMI*, which gives an opposite picture to both MIs and PBI. Indicator choice is hence a question of user preference for specification and empirical backing. The implications of our findings may be sobering for policy makers. We demonstrated through simulations that economic discomfort appears difficult to reduce noticeably in the short term.

Note

Earlier versions of this paper have been presented at sessions of the Sonderforschungsbereich 475 der Deutschen Forschungsgemeinschaft, Complexity Reduction in Multivariate Data Structures, at the Annual meeting of the Economic Policy Committee of the Verein für Socialpolitik on 9 March 2023 in Düsseldorf, and at the 37th CIRET Conference from 10 to 14 September 2024 in Vienna.

References

Barro, R. J. (1999). Reagan vs. Clinton: Who's the Economic Champ? *Business Week*, *22*.

Berlemann, M., & Enkelmann, S. (2014). The Economic Determinants of U.S. Presidential Approval: A Survey. *European Journal of Political Economy*, *36*, 41-54.

Blanchflower, D. G., Bell, D. N. F., Montagnoli, A., & Moro, M. (2014). The Happiness Trade-Off between Unemployment and Inflation. *Journal of Money, Credit and Banking*, *46*, 117-141.

Di Tella, R., MacCulloch, R. J., & Oswald, A. J. (2001). Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review*, *91*, 335-341.

Di Tella, R., MacCulloch, R. J., & Oswald, A. J. (2003). The Macroeconomics of Happiness. *The Review of Economics and Statistics*, *85*, 809-827.

Lovell, M. C., & Tien, P.-L. (2000). Economic Discomfort and Consumer Sentiment. *Eastern Economic Journal*, *26*, 1-8.

Okun, A. M. (1973). Comments on Stigler's Paper. *American Economic Review*, *63*, 172-177.

Okun, A. M. (1976). Conflicting National Goals. In: Eli Ginzberg (ed.), *Jobs for Americans*. Englewood Cliffs, NJ: Prentice Hall, reprinted in: Pechman, J. A. (1983), (pp. 221-249).

Welsch, H. (2007). Macroeconomics and Life Satisfaction: Revisiting the Misery Index. *Journal of Applied Economics*, *10*, 237-251.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 22(28)

Dependent Binomial Distribution, Generalization

Stanisław Heilpern

Department of Statistics, Wroclaw University of Economics and Business, Poland

e-mail: stanislaw.heilpern@ue.wroc.pl

We will study the following model. Let $X_1, \ldots, X_n$ be the continuous random variables with the same distribution and the threshold $d > 0$. We define the following Bernoulli random variables $I_1, \ldots, I_n$, where

$$I_j = \begin{cases} 0 & X_j \leq d \\ 1 & X_j > d' \end{cases}$$

the status of our system, and the probabilities $p = \Pr(I_j = 1)$ and $q = 1 - p_j$ (Heilpern, 2007, 2024). We will investigate the random variable

$$K = \sum_{j=1}^{n} I_j.$$

This is the number of objects meeting the condition $X_j > d_j$. Let $f_{\mathbf{I}}(i_1, \ldots, i_n) = \Pr(I_1 = i_1, \ldots, I_n = i_n)$ be the probability mass function (p.m.f), $F_{\mathbf{I}}(i_1, \ldots, i_n) = \Pr(I_1 \leq i_1, \ldots, I_n \leq i_n)$ the cumulative distribution function (c. d. f.) and the marginal c. d. f., where $i_j \in \{0, 1\}$ and $\mathbf{I} = (I_1, \ldots, I_n)$.

This model can be used in reinsurance and credit issues. In reinsurance n is the number of claims, X_j is the value of jth claim and d is the retension. The indicator I_i takes the value 1 where the reinsurer covers the claim and 0 otherwise. The random variable K is the number of claims covered by the reinsurer. In credit issues n is the number of obligors, X_j is the jth obligor's financial situation, d is the value of the credit. The indicator I_i takes the value 1 where the obligor repays the credit and 0 otherwise. The random variable K is the number of solvent obligors (Heilpern, 2007, 2024).

We assume that the random variables $X_1, \ldots, X_n$ may be dependent. We describe the dependent structure of $\mathbf{X} = (X_1, \ldots, X_n)$ using copula $C_{\mathbf{X}}$. This is the link (Nelsen, 1999)

$$F_{\mathrm{X}}(x_1, \ldots, x_n) = C_{\mathrm{X}}(F_{X_1}(x_1), \ldots, F_n(x_n)).$$

If X_j is continuous, then $C_{\mathbf{X}}$ is univocally determined, but the copula $C_{\mathbf{I}}$ is univocally determined on the points of jumps i_j only. The copulas $C_{\mathbf{X}}$ and $C_{\mathbf{I}}$ are equal at i_j. We also assume that the copula C_{X} is exchangeable, i.e. $C_{\boldsymbol{X}}(u_1, \ldots, u_n) = C_{\boldsymbol{X}}(u_{\pi(1)}, \ldots, u_{\pi(n)})$ for any permutation of set $\{1, \ldots, n\}$. Then $C_{\mathbf{I}}$ is exchangeable, too. The distribution of K is calculated using the following formula

$$\Pr(K = k) = \binom{n}{k} f_{k,n} = \sum_{j=0}^{k} (-1)^j \frac{n!}{(n-k)!\,j!\,(k-j)!} F_{k-j,n},$$

where $f_{k,n} = Pr(I_1 = 1, \ldots, I_k = 1, I_{k+1} = 0, \ldots, I_n = 0)$ and $F_{k,n} = Pr(I_{k+1} = 0, \ldots, I_n = 0)$ (Cossette et al. 2002) .We obtain that $F_{k-j,n} = C(\underbrace{1, \ldots, 1}_{k}, \underbrace{q, \ldots, q}_{n-k})$, $E(K) = np$ and $V(K) = npq + (n^2 -$

$n)(C(q, q) - p^2)$. We may say that the random variable K has a **dependent binomial distribution** and denote it as $K \sim DB(n, p, C_I)$ (Heilpern, 2007).

Now we present the selected copulas. When the random variables $X_1, \dots, X_n$ are independent, then $\Pi(u_1, \dots, u_n) = u_1 \cdot \dots \cdot u_n$. In this case $F_{k,n} = q^{n-k}$, $f_{k,n} = p^k q^{n-k}$, $\Pr(K = k) = \binom{n}{k} p^k q^{n-k}$ and $V(K) = npq$. This is the classical binomial distribution. The comonotonicity, strict positive dependence is done using the copula $M(u_1, \dots, u_n) = \min(u_1, \dots, u_n)$. We have

$$F_{k,n} = \begin{cases} q & k < n \\ 1 & k = n \end{cases}, f_{k,n} = Pr(K = k) = \begin{cases} q & k = 0 \\ 0 & 0 < k < n \\ p & k = n \end{cases}$$

and $V(K) = n^2 pq$. The copula, which is the convex combination of independence and comonotonicity $(1 - \rho)\Pi + \rho M$, where $0 \leq \rho \leq 1$, is called the Spearman copula. The parameter ρ is the Spearman correlation coefficient. We obtain

$$F_{k,n} = \begin{cases} (1-\rho)q^{n-k} + \rho q & k < n \\ 1 & k = n \end{cases}, f_{k,n} = \begin{cases} (1-\rho)q^n + \rho q & k = 0 \\ (1-\rho)p^k q^{n-k} & 0 < k < n \\ (1-\rho)p^n + \rho p & k = n \end{cases}$$

and $V(K) = npq(1 + \rho(n - 1))$.

The Archimedean copulas take a simple, quasi-additive form $\varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_n))$, where the generator φ is the decreasing and convex function satisfying conditions: $\varphi(0) = \infty$ and $\varphi(1) = 0$. We have $F_{k,n} = \varphi^{-1}((n - k)\varphi(q))$ in this case. The Archimedean copulas form families characterized by some parameters, which reflect the degree of dependence. The generator $\varphi(u) = u^{-a} - 1$, $a > 0$, induces the Clayton family with $F_{k,n} = ((n - k)(q^{-a} - 1) + 1)^{-1/a}$ and $V(K) = npq + (n^2 - n)((2q^{-a} - 1)^{-1/a} - q^2)$. The limit value of parameter $a = 0$ corresponds to independence, and $a = \infty$ implies comonotonicity. When $\varphi(u) = (-\ln u)^t$, $t \geq 1$, we get the Gumbel family with $F_{k,n} = F_{k,n} = q^{(n-k)^{1/t}}$ and $V(K) = npq + (n^2 - n)$. In case $t = 1$ we obtain independence and when $t = \infty$ we get comonotonicity.

Example. Let $n = 50$, $p = 0.3$ and the dependent structure of $X_1, \dots, X_n$ is described by the Clayton, Gumbel and Spearman copulas. The distribution of K for these copulas and different values of the Kendal τ coefficient of correlation: 0, 0.2, 0.4, 0.6, 0.8, 1 is presented in Figure 1.

The variance of the random variable K grows significantly with increasing dependency for these copulas. The variance for the Gumbel and Spearman families is greater than for the Clayton family and the variance for the Gumbel family is comparable to the variance for Spearman copulas.

Now, we can treat the number of objects as the random variable N. Then, the number of claims covered by the reinsurer or the number of solved obligors is a random sum (Heilpern, 2024)

$$K = \sum_{j=1}^{N} I_j.$$

Then we obtain, that

$$Pr(K = k) = \sum_{n=k}^{\infty} Pr(N = n)\, Pr(K = k|N = n) = \sum_{n=k}^{\infty} Pr(N = n)\, Pr(K_n = k),$$

where $K_n = I_1 + \dots + I_n$ and $K_0 = 0$.

The distribution of the number of objects K depends on the copulas and for small dependencies, the differences are greater.

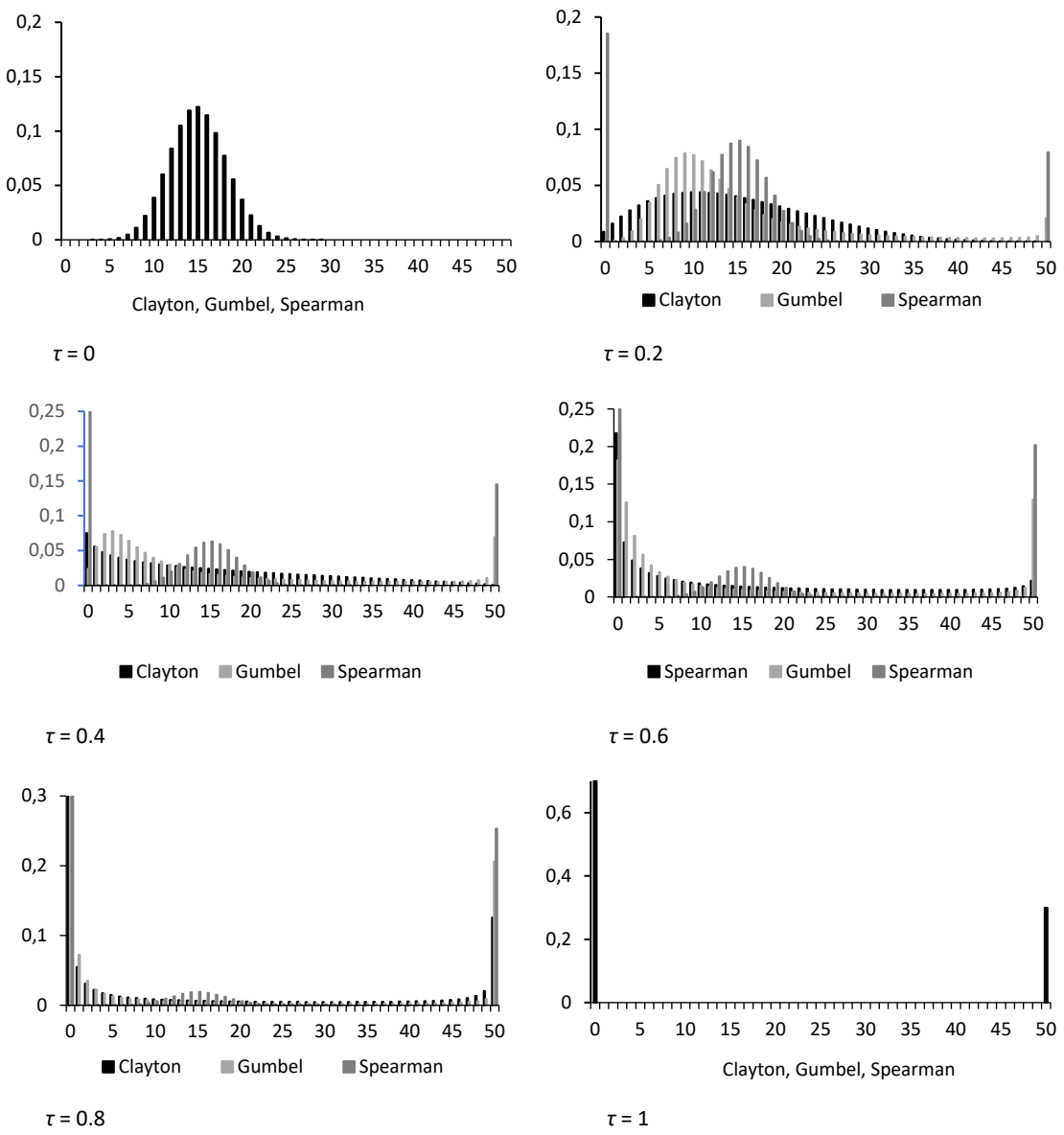

Figure 1. The distribution of *K*

Source: own elaboration.

So far we have dealt with a number of objects *K*. Now we will study the total value *T*, e.g. the total value of claims covered by the reinsurer. Let $Z_j = \max(X_j - d, 0)$, the value of the *j*th claim covered by the reinsurer. If the number of claims is random, then the total value *T* is a random variable $T = \sum_{j=1}^{N} I_j Z_j$. Then

$$F_T(x) = \sum_{n=0}^{\infty} F_{T_n}(x)\, Pr(N = n),$$

where $T_n = \sum_{j=1}^{n} I_j Z_j$ .We obtain

$$F_{T_n}(x) = Pr(K = 0) + \sum_{k=1}^{n} Pr(S_k \le x) \binom{n}{k} f_{k,n},$$

where $S_k = \sum_{j=1}^{k} Z_j = \sum_{j=1}^{k} X_j - kd$.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 22(28)

Now we will investigate a case when the probability of success p is imprecisely determined. For instance, we obtain information that it is equal to "about 0.3". For this purpose, we use fuzzy sets (Zadeh, 1965; Dubois & Prade, 1980). The fuzzy subset A of space Z is described by its membership function $\mu_A: Z \rightarrow [0, 1]$. The crisp set $A_\alpha = \{z \in Z: \mu_A(z) \geq \alpha\}$, where $0 < \alpha \leq 1$, is called α-cut. The cut A_1 is the core of A, and A_0, the closure of set $\{z \in Z: \mu_A(z) > 0\}$, is the support of fuzzy set A.

A fuzzy number is a fuzzy subset of the real line R. Its every α-cut A_α is the compact interval $[A_\alpha^L, A_\alpha^U]$ (Dubois & Prade, 1980). The trapezoidal fuzzy number $A = (a, b, c, d)$ has a linear membership function at the intervals $[a, b]$ and $[c, d]$. If $b = c$ then we obtain the triangular fuzzy number $A = (a, b, d)$. We define the mean value (Campos and Gonzalez, 1987; Heilpern, 1992) $MV(A) = \frac{1}{2}\int_0^1 (A_\alpha^L + A_\alpha^U)d\alpha$ and spread (Heilpern, 2024) $S(A) = \int_0^1 (A_\alpha^U - A_\alpha^L)d\alpha$ of fuzzy number. The image $f(A)$ of a fuzzy subset of Z, where $f: Z \rightarrow$ R, has the following membership function (Zadeh, 1965)

$$\mu_{f(A)}(x) = \sup_{f(x)=x} \mu_A(z).$$

This formula allows us to define the arithmetic operations * on fuzzy numbers. The membership function of the fuzzy number $A * B$ is equal to

$$\mu_{A*B}(z) = \sup_{x*y=z} \{min\{\mu_A(x), \mu_B(x)\}\}.$$

Therefore, we can observe that the borders of α-cuts of arithmetic operation $A * B$ are defined by the borders of A and B, e.g. $(A+B)_\alpha^L = A_\alpha^L + B_\alpha^L$, $(A+B)_\alpha^U = A_\alpha^U + B_\alpha^U$ (Dubois & Prade, 1980).

Let K_p has dependent distribution DB(n, p, C_l). But we do not know the exact value of probability p. We only know the imprecision value of p "about p_0" (Heilpern 2020, Dębicka et al. 2022). We can treat it as a triangular fuzzy number P (Heilpern, 2024). This fuzzy number induces the fuzzy subset $\mathbf{K}$ on the family of the dependent binomial random variables with the membership function $\mu_{\mathbf{K}}(K_p) = \mu_P(p)$. The α-cat of it is equal to $\mathbf{K}_\alpha = \{K_p: p \in P_\alpha\}$. Using functions $f(K_p) = E(K_p) = np = m$ and $g(p) = V(K_p)$ we can define the expected value $E(\mathbf{K})$ and variance $V(\mathbf{K})$ of fuzzy set $\mathbf{K}$. They have the following membership functions $\mu_{E(\mathbf{K})}(m) = \mu_P\left(\frac{m}{n}\right)$, $\mu_{V(\boldsymbol{K})}(s) = \sup_{\{p:\, g(p)=s\}} \mu_P(p)$ and α-cuts $E(\mathbf{K})_\alpha = \{m = np: p \in P_\alpha\}$, $V(\mathbf{K})_\alpha = \{s = g(p): p \in P_\alpha\}$. We compute $E(\mathbf{K})$, $V(\mathbf{K})$, the mean and spread of $\mathbf{K}$ for Spearman, Gumbel and Spearman copulas for different values of Kendal coefficient of correlation τ. These values depend on the copulas and for small dependencies, the differences are greater.

Now we define the Pr($\mathbf{K} \in B$), where B is a crisp subset of R using function $f(p) = \Pr(K_p \in B)$, where $K_p \sim$ DB(n, p, C_l). It has the α-cut $\Pr(\mathbf{K} \in B)_\alpha = \{\Pr(K_p \in B): K_p \sim \text{DB}(n, p, C_l), p \in P_\alpha\}$. We compute Pr($\mathbf{K}$ = 10), where $n = 50$ and the dependent structure is described by Spearman copula. We obtain Pr($\mathbf{K}$ = 10)$_\alpha$ = $[f_\rho(0.4 - 0.1\alpha), f_\rho(0.05 + 0.25\alpha)]$, where $f_\rho(p) = \binom{50}{10}(1-\rho)p^{10}(1-p)^{40}$ (Heilpern, 2024).

We assume that the copula parameter a is imprecision (Heilpern, 2024). Let $K_a \sim$ DB(n, p, C_a) and $a = l(\tau)$, where τ is a Kendall coefficient of correlation. We can treat this coefficient as a fuzzy number T and the parameter a as a fuzzy number $A = l(T)$ in this case. The fuzzy number A induces the fuzzy copula $\mathbf{C}_A$ with membership function $\mu_{\boldsymbol{C}_A}(C_a) = \mu_A(a)$ and the fuzzy number of objects $\mathbf{K}_A$ with $\mu_{\boldsymbol{K}_A}(K_a) = \mu_A(a)$, where $K_a \sim$ DB(n, p, C_a). The expected value of $\mathbf{K}_A$ is a crisp number $E(\mathbf{K}_A) = np$, but the variance $V(\mathbf{K}_A)$ is fuzzy with $\mu_{V(\boldsymbol{K}_A)}(s) = \sup_{\{a:\, v(a)=s\}} \mu_A(a)$, where $v(a) = npq + (n^2 - n)(C_a(q, q) - q^2)$. The fuzzy probability $\Pr(\mathbf{K}_A \in B)$ has the membership function $\mu_{Pr(\boldsymbol{K}_A \in B)}(q) = \sup_{\{a:\, f(a)=q\}} \mu_A(a)$, where $f(a) = \Pr(K_a \in B)$ and $K_a \sim$ DB(n, p, C_a). It has α-cut $\Pr(\mathbf{K}_A \in B)_\alpha = \{\Pr(K_a \in B): a \in A_\alpha\}$.

References

Campos, L. M., & Gonzalez, A. (1989). A Subjective Approach for Ranking Fuzzy Numbers. *Fuzzy Sets and Systems*, *29*, 145-153.

Cossette, H., Gaillardetz, P., Marceau, E., & Rioux, J. (2002). On Two Dependent Individual Risk Models. *Insurance: Mathematics and Economics*, *30*(1), 153-166.

Dębicka, J., Heilpern, S., & Marciniuk, A. (2022). Modelling Marital Reverse Annuity Contract in a Stochastic Economic Environment. *Statistika*, *102*(3), 261-281.

Dubois, D., & Prade, H. (1980). *Fuzzy Sets and Systems: Theory and Applications.* Academic Press.

Heilpern, S. (1992). The Expected Value of a Fuzzy Number. *Fuzzy Sets and Systems*, *47*(1), 81-86.

Heilpern, S. (2007). Zależny rozkład dwumianowy i jego zastosowanie w reasekuracji i kredytach. *Badania Operacyjne i Decyzje*, *71*(1), 45-61.

Heilpern, S. (2020). *Selected Credit Risk Models*, In: B. Ciałowicz (Ed.), *Quantitative methods in the contemporary issues of economics* (pp. 48-57). edu-Libri.

Heilpern, S. (2024). Selected Reinsurance Models. *Central European Journal of Economic Modelling and Econometrics*. *16*(2), 95-124.

Nelsen, R. B. (1999). *An Introduction to copulas*, Springer.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, *8*(3), 338-353.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 22(28)

Cash Flows and Probabilistic Structure of the Model of Combined Reverse Annuity Contract and Critical Illness Insurance with the Effect of Gender

Agnieszka Marciniuk

Department of Statistics, Wroclaw University of Economics and Business, Poland

e-mail: agnieszka.marciniuk@ue.wroc.pl

Beata Zmyślona

Department of Statistics, Wroclaw University of Economics and Business, Poland

e-mail: beata.zmyslona@ue.wroc.pl

Introduction

Financial security is an important aspect of life nowadays. This issue is of particular importance in view of the phenomenon of ageing populations in many countries around the world because it causes a significant increase in the proportion of pensioners in most European countries. Pensioners are among the groups most at risk of poverty. In Poland, due to the different retirement ages, pensioners are exposed to living in extreme material conditions in retirement. The economic status of married couples is among the factors most determining their financial situation. In addition, pensioners are exposed to particularly high costs of medical treatment or palliative care, as these groups are most often affected by chronic and severe illnesses.

Thus, on the one hand, pensioners are exposed to a reduced financial standard of living, while on the other hand, they most often own real estate. In this situation, the use of equity release contracts, which allow access to capital locked up in the value of real estate, seems reasonable. The abstract addresses pensioners' financial protection issues through various financial and insurance contracts. We provide a hybrid of equity release contracts with critical illness insurance. The analysis of cash flows associated with this hybrid requires introducing the multistate Markov model. The probabilistic structure and the estimation procedures of the model are described in (Dębicka & Zmyślona, 2016, 2019). The theoretical background and details connected with the form of the contracts are presented in (Marciniuk & Zmyślona, 2022). As an example of severe illness, lung cancer is chosen. We consider the morbidity, mortality and fatality rates based on historical data from the Lower Silesian population. The examples of analysis and comparisons of cash flows under different contract variants (married and individual) for women, men and the general population are presented. The benefits derived from this contract can help improve the living conditions and the quality of life for the elderly and provide additional financial resources in case of a critical illness.

Contract

The contract consists of two elements, namely, the reverse annuity contract and the dread disease insurance. One of the equity release forms similar to the home reversion scheme is the reverse annuity contract (Marciniuk et al., 2020). In exchange for a monthly annuity, the senior citizen surrenders ownership rights to the property, guaranteed in Chapter IV of the Land Registry, to live in the apartment until death. The construction of this kind of contract is based on a model

that describes an extended lifetime considering the risk of morbidity, severe disease, and life expectancy in the critical stage.

The model connected with health insurance considers the course of a dread disease. Thus, the following states are distinguished (Dębicka & Zmyślona, 2019; Zmyślona & Marciniuk, 2020):

1. the insured is alive and healthy,
2. the insured became mildly ill during the last year,
3. the insured has been mildly ill for at least one year,
4. the insured became critically ill during the last year,
5. the insured has been critically ill for at least one year,
6. the insured is dead (D—dead).

As critical illness mortality is very dependent on the duration of the illness, the fifth state is expanded to include the survival time in a critical stage. The multistate model related to dread disease insurance is presented in Figure 1.

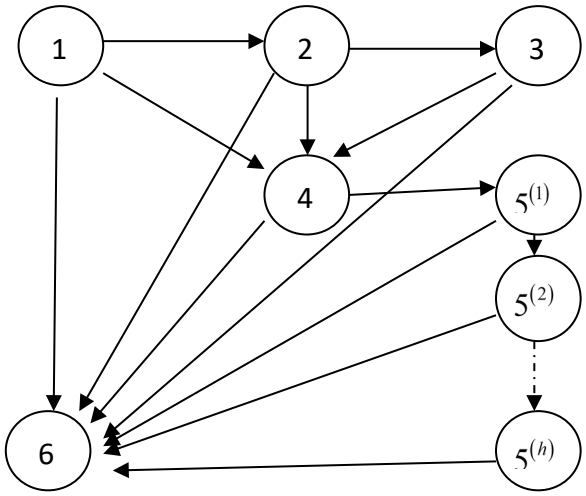

Figure 1. The multistate model for one spouse

Source: (Zmyślona and Marciniuk, 2020).

The circles present the states, and the arcs describe the direct transitions between the states. Such an extended model makes it possible to vary the probability of death according to health status. The transition probabilities matrix connected with the multistate model related to critical health insurance is described in (Dębicka and Zmyślona 2019) and is given as follows

$$\mathbf{Q}_h^{[x]}(k) = \begin{pmatrix} q_{11} & q_{12} & 0 & q_{14} & 0 & 0 & \dots & 0 & q_{16} \\ 0 & 0 & q_{23} & q_{24} & 0 & 0 & \dots & 0 & q_{26} \\ 0 & 0 & q_{33} & q_{34} & 0 & 0 & \dots & 0 & q_{36} \\ 0 & 0 & 0 & 0 & q_{45^{(1)}} & 0 & \dots & 0 & q_{46} \\ 0 & 0 & 0 & 0 & 0 & q_{5^{(1)}5^{(2)}} & \dots & 0 & q_{5^{(1)}6} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

where $q_{11} = 1 - (q_{x+k} - \varphi_{x+k}) - \chi_{x+k}$ $q_{12} = \chi_{x+k}(1 - \psi_{x+k})$, $q_{16} = q_{x+k} - \varphi_{x+k}$, $q_{ij} = 1 - q_{x+k} - \xi_{x+k}$, for i=2,3, j=3, $q_{ij} = \xi_{x+k}$, for i=2,3, j=4, $q_{i6} = q_{x+k}$, for i=2,3, $q_{45^{(1)}} = 1 - d_{x+k}^{(4,5^{(1)})}$, $q_{46} = d_{x+k}^{(4,5^{(1)})}$, $q_{5^{(i)}5^{(j)}} = 1 - d_{x+k}^{(5^{(i)},5^{(j)})}$ for i=1,2,…h-1, j=i+1, $q_{5^{(l)}6} = d_{x+k}^{(5^{(l)}6)}$, for l=1,2,…,h.

The transition probabilities depend on the rates for the whole population

q_{x+k}– the probability of death, φ_{x+k}– the dread disease mortality rate, χ_{x+k} – the dread disease incidence rate (the morbidity rate),

and the rates specified for the critical illness population:

ψ_{x+k} – the percentage of patients diagnosed in the critical stage, ξ_{x+k} – the probability of health deterioration to the critical state, $d_{x+k}^{(i,j)}$ – the fatality rate in the population of the critically ill.

We assume that *x* denotes the age of an insured.

In this part, we consider the influence of gender on the cash flows connected with this model. Gender is one of the most significant factors in the incidence of health inequalities, which are reflected in the phenomenon of excess mortality in the male population. It is described by the lower life expectancy of the male population compared to that of the female population. It occurs worldwide, although in Poland, it is at a relatively high level of around eight years. Male over-mortality is due not only to biological and genetic factors but also to socio-cultural factors consisting of lifestyle, physical activity, addictions, and diet. Health behaviours are also important, which become apparent in different attitudes towards prevention, health care and anti-health behaviours. Male over-mortality also manifests itself in increased rates of morbidity and mortality from a number of serious and chronic diseases.

The impact of gender on cash flows will be considered by comparing the size of benefits and premiums with and without a gender distinction. We consider two variants of the contract the first – marriage and the second – individual. In the first variant, the spouses have a property valued for the amount W, and they decide on a marriage reverse annuity contract, allocating a percentage 1-β of the received annuity benefit $\ddot{a}_{(x,y)}$ as a health insurance premium p for the husband and wife. The model assumes the independence of the spouses' future lifetime. In this case, the R morbidity rate is also taken into account. The second variant means that the spouses have a property valued at the amount W. Each decides on an individual reverse annuity contract at 0,5W, allocating a percentage 1-β of the received annuity benefit $\ddot{a}_x$ as a health insurance premium p for the husband and wife. In addition, both variants consider the case when the gender division is not considered, which means the values are calculated for the general population.

Calculations and Conclusions

The following assumptions are made for the calculations:

- the value of property $W = 100000$ euros,
- the percentage of property $\alpha = 50\%$,
- the age of person $x, y \in \{65,70,75,75,80,85\}$,
- $\beta = 0.99$,
- the morbidity rate $R = 50\%$ (for the general population) and $R = 69.1\%$ (for the variant, when the gender's division is considered) is included only in the marriage contract,
- the fixed long-term interest rate i = 5.79 % (was estimated based on actual Polish market data related to the yield to maturity on fixed interest bonds and Treasury bills from 2008 in the Svensson model),
- Life Tables for Poland from 2008 were used.

In addition, the following notations were adopted:

- $c_x, c_y, c_{x/y}$ - critical illness insurance benefit for wife, husband and general, respectively,
- BG - by gender,
- WG - without gender breakdown (general).

Table 1 presents cash flows in both variants.

In the individual variant, we can observe that the benefit of a reverse annuity contract is highest for men and lowest for women. The exact relationship is for the dread disease insurance premium. However, the benefit in this case is highest for women and lowest for men. When gender division is not considered, the benefit of reverse annuity contract is higher than for women, but the differences are insignificant. The exception is when women are older than 80. Then, there is an inverse relationship.

In the case of no gender split, women gain in the case of the reverse annuity but lose significantly in the case of the insurance.

The higher total annuity in both variants is when the gender split is considered. The annuity differences increase with the rise of the spouses' age (0.6% when x = y = 65 and 11.55% when x = y = 85 for the marriage contract 2.18% when x = y = 65 and 12.8% when x = y = 85 for the individual contract). The total reverse annuity is the highest in the second variant, but we have to remember that in this case, when one of the spouses dies, the second will obtain a much lower pension. The insurance benefit for men is lower than the unisex benefit. In the individual variant, the differences range from 32% to 42%. In the marriage variant, the differences are lower, from 8% to 17%. The benefit for women is significantly higher than those without a gender split. For the individual variant, the differences decrease with the beneficiary's age from 94% to 66%, and for the marriage one from 57.3% to 18%.

Table 1. Cash flows of the described contract in two variants

First variant – marriage contract								
x=y	annuity $\ddot{b}^{BG}_{(x,y)}$	annuity $\ddot{b}^{WG}_{(x,y)}$	benefit for man c_y	benefit for woman c_x	benefit $c^{WG}_{x/y}$	premium for man p_y	premium for woman p_x	premium $p^{WG}_{x/y}$
65	4039.69	4015.53	4886.41	8316.32	5285.99	27.92	12.48	16.64
70	4568.69	4501.41	5346.56	9427.91	6137.50	31.58	14.12	19.05
75	5374.72	5217.88	6709.78	11710.01	8075.86	37.15	16.61	22.69
80	6601.92	6244.41	10233.44	15295.94	12215.22	45.64	20.41	28.12
85	8450.05	7575.06	15755.93	20230.30	17147.11	58.42	26.12	35.21

Second variant – individual contract											
x=y	$\ddot{b}_y$	$\ddot{b}_x$	$\ddot{b}_x + \ddot{b}_y$	$\ddot{b}^{WG}_{x/y}$	$2\ddot{b}^{WG}_{x/y}$	c_y	c_x	$c^{WG}_{x/y}$	p_y	p_x	$p^{WG}_{x/y}$
65	2730.05	2256.79	4986.84	2440.26	4880.52	4778.30	15033.35	7749.62	27.30	22.57	24.40
70	3187.24	2619.13	5806.37	2815.99	5631.98	5396.69	17487.65	9074.49	31.87	26.19	28.16
75	3866.40	3196.56	7062.96	3383.81	6767.62	6983.12	22531.70	12043.24	38.66	31.97	33.84
80	4882.28	4109.54	8991.82	4216.20	8432.40	10947.91	30801.80	18311.96	48.82	41.10	42.16
85	6382.61	5526.00	11908.61	5278.69	10557.38	17214.48	42794.17	25707.59	63.83	55.26	52.79

Source: own elaboration.

References

Dębicka, J., & Zmyślona, B. (2016). Construction of Multi-State Life Tables for Critical Illness Insurance – Influence of Age and Sex on the Incidence of Health Inequalities. *Śląski Przegląd Statystyczny,* (14), 41-63. https://doi.org/10.15611/sps.2016.14.03

Dębicka, J., & Zmyślona, B. (2019). Modelling of Lung Cancer Survival Data for Critical Illness Insurances. *Statistical Methods and Applications, 28*(4), 723-747. https://doi.org/10.1007/s10260-019-00449-x.

Marciniuk, A., Zimková, E., Farkašovský, V., & Lawson, C. W. (2020). Valuation of Equity Release Contracts in Czech Republic, Republic of Poland and Slovak Republic. *Prague Economic Papers*, *29*(5), 505-521. https://doi.org/10.18267/j.pep.743.

Marciniuk, A., & Zmyślona, B. (2022). Marriage and Individual Equity Release Contracts with Dread Disease Insurance as a Tool for Managing the Pensioners' Budget. *Risks*, *10*(7), 140, 1-15. https://doi.org/10.3390/risks10070140

Zmyślona, B., & Marciniuk, A. (2020). Financial Protection for the Elderly – Contracts Based on Equity Release and Critical Health Insurance. *European Research Studies Journal* XXIII*,* (1), 867-882. https://doi.org/10.35808/ersj/1798

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 22(28)

A Review on Sample Selection Bias in Corporate Governance Research

Vladlena Prysyazhna

University Marburg, Germany

e-mail: prysyazv@staff.uni-marburg.de

As an institutional framework, Corporate Governance deals with companies' internal processes to ensure company success and the satisfaction of shareholders' claims through effective decision-making, incentive systems, monitoring mechanisms, transparency and accountability. Thus, decisions made by the companies' management are determined by their ownership structure and corporate board composition and shape firms' characteristics concerning compliance, financing and strategy. From a methodological standpoint, such decisions, or 'choices', inevitably lead to selection bias when analyzing their outcomes because only the outcomes of choices made are observable, while the potential outcomes of alternative decisions cannot be assessed for the same company. Similarly, specific corporate governance characteristics, such as CEO duality, presence of female directors or institutional shareholders shape firms' strategy and affect their decision-making, but are themselves endogenously determined. Indeed, such characteristics are not "randomly assigned" to firms, but depend on their institutional environment, industry and other factors. Further, it is usually not possible to select a random sample of companies for an empirical investigation, which can also result in biased results.

Several methods exist to handle sample selection bias and address a situation when either (1.) the inability to observe the outcome of a choice not made results in an inappropriate "control group" that makes an empirical evaluation biased or (2.) a specific outcome can be observed only for a non-random sub-population of firms. Such methods primarily encompass (a.) propensity score matching that provides a basis for an estimation of the average treatment effect to address the sample selection on observables and (b.) Heckman selection model to explicitly account for the selection process which results in a non-random sub-population and can be affected by unobservable factors.

In this presentation, the methodological problem of sample selection bias in Corporate Governance research outlined above is considered. First, important internal Corporate Governance attributes are discussed to define the research questions that are confronted with the aforementioned problems following an introduction of appropriate methods to handle them. Then, a review of the approaches used in existing studies will be presented. Finally, the ability of the aforementioned methods to uncover unbiased relationships is demonstrated based on an exemplary empirical investigation of the impact of CEO duality on firm performance represented by Tobin's Q for an international sample of large listed firms. It can be shown that estimating an average treatment effect based on propensity score matching can reveal the effects that otherwise appear insignificant and discuss the need to additionally address the selection bias on unobservables in the given context.

References

Börsch-Supan, A., & Köke, J. (2002). An Applied Econometricians' View of Empirical Corporate Governance Studies. *German Economic Review*, *3*(3), 295-326.

Chen, Y., & Vann, C. E. (2017). Clawback Provision Adoption, Corporate Governance, and Investment Decisions. *Journal of Business Finance & Accounting, 44*(9-10), 1370-1397.

Hoechle, D., Schmid, M., Walter, I., & Yermack, D. (2012). How Much of the Diversification Discount Can Be Explained by Poor Corporate Governance? *Journal of Financial Economics, 103*(1), 41-60.

Iyengar, R. J., & Zampelli, E. M. (2009). Self-selection, Endogeneity, and the Relationship between CEO Duality and Firm Performance. *Strategic Management Journal, 30*(10), 1092-1112.

Jo, H., & Harjoto, M. A. (2012). The Causal Effect of Corporate Governance on Corporate Social Responsibility. *Journal of Business Ethics, 106*, 53-72.

Shen, C. H., & Chang, Y. (2012). Corporate Social Responsibility, Financial Performance and Selection Bias: Evidence from Taiwan's TWSE-listed banks. In J. R. Barth, C. Lin, & C. Wihlborg (Eds.), *Research Handbook on International Banking and Governance.* Edward Elgar Publishing Ltd.

Switzer, L. N., & Tang, M. (2009). The Impact of Corporate Governance on the Performance of US Small-cap Firms. *International Journal of Business, 14*(4), 341-355.

Tucker, J. W. (2010). Selection Bias and Econometric Remedies in Accounting and Finance Research. *Journal of Accounting Literature, 29,* 31-57.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 22(28)

Model Selection Before and After Multiple Imputation

Robert Scherf

Philipps University of Marburg, Germany

e-mail: scherfro@staff.uni-marburg.de

The Missing Data Model

Consider a linear regression model $Y = X\beta + Z\gamma + \epsilon$, where $X \in \mathbb{R}^{n\times k}, Z \in \mathbb{R}^{n\times m}$ are nonrandom matrices and $\epsilon \sim N(0, I_n\sigma^2)$ are independent unobservable error terms. In this setting X contains focus regressors, i.e. regressors that must be included, and Z contains auxiliary or doubtful regressors. We assume that X and Z are observed on the complete sample, but only for the first $r < n$ units $Y_r = (Y_1, \dots, Y_r)$ is observed and $Y_{n-r} = (Y_{r+1}, Y_{r+2}, \dots, Y_n)$ is missing at random. Multiple Imputation (MI) replaces missing values by repeated draws from a posterior probability distribution under a specified imputation model. This leads to $M \geq 2$ vectors:

$$Y^{(l)} = \left(Y_1, Y_2, \dots, Y_r, Y_{r+1}^{(l)}, \dots, Y_n^{(l)}\right) \ (l = 1, \dots, M).$$

Each imputed data set $\left(Y^{(1)}, X, Z\right), \dots, \left(Y^{(M)}, X, Z\right)$ can then be analyzed by standard complete data procedures, such as estimation of regression coefficients or standard errors. Note that the imputation and subsequent analyses may be done by separate entities.

Model Selection

When imputation and subsequent analyses are performed by different individuals, discrepancies may arise between the underlying models used. This issue, known as uncongeniality, represents a significant critique of Multiple Imputation (MI). Also, in linear regression theory it is well known that the inclusion of irrelevant or exclusion of relevant variables may introduce bias. Therefore, uncongeniality adds an additional layer of complexity. The observation that identical selection methods applied before and after MI can yield different model rankings offers a stochastic interpretation of this phenomenon.

We assume both parties only choose between the fully restricted $\hat{\beta} = (X'X)^{-1}X'Y$ or unrestricted model $\left(\hat{\beta}; \hat{\gamma}\right) = ([X;Z]'[X;Z])^{-1}[X;Z]'Y$, i.e. whether or not to include Z. The selection processes for imputation and analysis are based on the observed data (Y_r, X_r, Z_r) and the imputed data sets, respectively. The models are ranked according to their Akaike information criterion (AIC) value, which is calculated for the ith model as

$$AIC(i) = p \cdot \log\left(\hat{\sigma}_i^2\right) + 2(q_i + 1),$$

where $p \in \{r, n\}$ is the number of units included, $\hat{\sigma}_i^2$ is the MLE of σ^2 and q_i is the total number of regressors under the ith model.

In the analysis phase, the selection criterion is applied to each of the imputed data sets. Literature suggests several approaches to combine these M selected models. Two examples are the following rules:

1) choose variables that are selected in at least 50% of all data sets or
2) choose variables that are selected at least once and weigh them accordingly.

For the first rule, an uncongenial model for the analysis is selected if the initial imputation model is selected in fewer than 50% of all imputed data sets. As for the second rule, the analysis model is uncongenial if at least once a model different from the imputation is chosen. In this case, however, the degree of uncongeniality depends on its selection frequency.

The choice of AIC and later Bayesian information criterion (BIC) is due to analytical convenience as shown in section 3. However, instead of these likelihood-based approaches one could also think of model selection as in the classical context of pre-testing, where we could for instance select according to the results of t- or F-tests.

Exponential Inequalities

First, we extend the notation from section 1. Let

$$M = I - X(X'X)^{-1}X' \text{ and } \theta = (Z'MZ)^{0.5}\gamma.$$

Viewing MI as the opportunity to gain access to the complete instead of just the observed data justifies the following approach: Different models, i.e. the restricted and unrestricted estimators, before and after MI are selected by the AIC if and only if the inequality

$$rn \cdot log(1+a)\, log(1+b) - 2m(r\, log(1+a) + n\, log(1+b)) < -4m^2$$

holds for $a = \hat{\theta}_r'\hat{\theta}_r/(\hat{\sigma}_r^2 \cdot (r-k-m))$ and $b = \hat{\theta}_n'\hat{\theta}_n/(\hat{\sigma}_n^2 \cdot (n-k-m))$, where r and n as subscripts refer to an evaluation on the first r or n units, respectively. For example, selecting the unrestricted estimator on just the observed and the restricted on the complete sample corresponds to the following solution of the inequality above:

$$a > exp\left(\frac{2m}{r}\right) - 1 \text{ and } b < exp\left(\frac{2m}{n}\right) - 1,$$

where *a* and *b*, if interpreted as random variables on the whole sample space, follow non-central F-Distributions after proper rescaling. Analogous results hold for the BIC, which enables us to compare both methods with regards to the probability of changes, i.e. which method is more or less likely to select different models in both phases.

This can also be viewed independently of a missing data problem, since it detects when a change occurs after adding $(n-r)$ units. For MI, however, this is applied to all $l = 1, \dots, M$ data sets, where in each one the first $Y_r = (Y_1, \dots, Y_r)$ are equal and $Y_{n-r}^{(l)} = (Y_{r+1}^{(l)}, \dots, Y_n^{(l)})$ are random. Therefore, we are interested in the impact of these random vectors, because they are drawn from a conditional probability distribution depending on the outcome of a selection process.

Reference

Bartlett, J. W., & Hughes, R. A. (2020). Bootstrap Inference for Multiple Imputation under Uncongeniality and Misspecification. *Statistical Methods in Medical Research,* (29), 3533-3546.

Kim, J. K. (2004). Finite Sample Properties of Multiple Imputation Estimators. *The Annals of Statistics,* (32), 766-783.

Liang, H., Zou, G., Wan, A. T., & Zhang, X. (2011). Optimal Weight Choice for Frequentist Model Average Estimator. *Journal of the American Statistical Association,* (106), 1053-1066.

Rao, P. (1971). Some Notes on Misspecification in Multiple Regression. *The American Statistician,* (25), 37-39.

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika,* (63), 581-592.

Schomaker, M., & Heumann, C. (2014). Model Selection and Model Averaging After Multiple Imputation. *Computational Statistics and Data Analysis,* (71), 758-770.

Xie, X., & Meng, X.-L. (2017). Dissecting Multiple Imputation from a Multi-phase Inference Perspective: What Happens When God's, Imputer's and Analyst's Models are Uncongenial? *Statistica Sinica,* (27), 1485-1594.