

Multivariate Statistical Analysis MSA'2024

Organizing Committee / Komitet organizacyjny

Organizing Committee Chair / Przewodnicząca Komitetu Organizacyjnego
Dr hab. Alina Jędrzejczak, prof. UŁ

Scientific Secretaries / Sekretarze naukowi

Dr Piotr Szczepocki
Dr Kamila Trzcińska

University of Łódź

Uniwersytet Łódzki

Faculty of Economics and Sociology Wydział Ekonomiczno-Socjologiczny

Department of Statistical Methods Katedra Metod Statystycznych



Scientific Committee / Komitet naukowy

Scientific Committee Chair / Przewodniczący Komitetu Naukowego
Prof. dr hab. Czesław Domański, Uniwersytet Łódzki

Scientific Committee Members / Członkowie Komitetu Naukowego

Prof. Francesca Greselin,
University of Milano Bicocca

Prof. Marie Huskova,

Charles University in Prague

Prof. dr hab. Krzysztof Jajuga,
Uniwersytet Ekonomiczny we Wrocławiu

**Prof. dr hab. Mirosław
Krzyśko,**
Uniwersytet im. Adama Mickiewicza w Poznaniu

**Prof. dr hab. Włodzimierz
Okrasa,**
Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie

Prof. dr hab. Józef Pociecha,
Uniwersytet Ekonomiczny w Krakowie

Prof. dr hab. Mirosław Szreder,
Uniwersytet Gdański

Prof. dr hab. Janusz Wywiół,
Uniwersytet Ekonomiczny w Katowicach

Co-Organizers / Współorganizatorzy

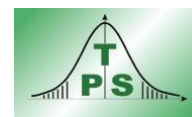
**Institute of Statistics and Demography
of the University of Łódź /
Instytut Statystyki i Demografii
Uniwersytetu Łódzkiego**



**Polish Academy of Sciences /
Polska Akademia Nauk**



**Polish Statistical Association, Branch in Łódź /
Polskie Towarzystwo Statystyczne, Oddział w Łodzi**



Honorary Patronage / Patronat honorowy

Rector of University of Łódź / Rektor Uniwersytetu Łódzkiego



Patronage of Rector
of University of Lodz

**The President of Statistics Poland /
Prezes Głównego Urzędu Statystycznego w Polsce**



Statistics Poland

Partner /Partner

Łódzka Specjalna Strefa



Łódzka
Specjalna Strefa
Ekonomiczna

LabMasters



Bank Pekao



StatSoft Polska Sp. z o.o.

StatSoft Polska

Invited Lectures / Wykłady zaproszone

Bogumił Kamiński, Szkoła Główna Handlowa w Warszawie

Tytuł: Statystyka w służbie metod podejmowania decyzji

Streszczenie: Teoria decyzji to obszar nauki zajmujący się badaniem procesu podejmowania decyzji. Klasyczne podejście do wspomagania podejmowania decyzji wykorzystywało głównie metody optymalizacyjne w celu wyboru najlepszego wariantu w sytuacji, gdy zagadnienie decyzyjne ma skomplikowaną postać funkcji celu i zbioru rozwiązań dopuszczalnych. Wykorzystanie tych metod w praktyce gospodarczej pokazało, że kluczowe jest ich uzupełnienie o uwzględnienie niepewności. Ta niepewność manifestuje się na dwa sposoby: niepewność wynikająca z natury modelowanego systemu oraz niepewność pomiaru parametrów tego systemu wykorzystywanych do specyfikacji modelu decyzyjnego. W referacie przedstawię jak te dwa aspekty, przy wykorzystaniu metod statystycznych, są uwzględniane w złożonych modelach wspomagania podejmowania decyzji. Prezentację zakończę przedstawieniem najnowszego trendu wykorzystania metod statystycznych do wspomagania podejmowania decyzji. W podejściu tym modele statystyczne, zwane metamodelami lub modelami surogatowymi, wykorzystywane są jako komponenty zastępujące tradycyjne moduły optymalizacyjne i symulacyjne.

Title: Statistics in the decision-making process

Abstract: Decision theory is an area of science that studies the decision-making process. The classic approach to decision support mainly used optimization methods to select the best variant when the decision-making problem had a complicated form of an objective function and a set of admissible solutions. The use of these methods in business practice has shown that it is crucial to supplement them by considering uncertainty. This uncertainty manifests itself in two ways: uncertainty resulting from the nature of the system being modeled and uncertainty in the measurement of the parameters of that system used to specify the decision model. In the presentation, I will present how these two aspects, using statistical methods, are taken into account in complex decision support models. I will end the presentation by discussing the latest trend in the use of statistical methods to support decision-making. In this approach, statistical models, called metamodels or surrogate models, are used as components replacing traditional optimization and simulation modules.

Francesca Greselin, Giorgia Zaccaria, University of Milano-Bicocca

Title: Studying hierarchical latent structures in heterogeneous populations with missing information

Abstract: An ultrametric Gaussian mixture model is a powerful tool for modeling hierarchical relationships among latent concepts, making it ideal for studying complex phenomena in diverse and potentially heterogeneous populations. However, in many cases, only an incomplete set of observations is available on the phenomenon under study. To address this issue, we propose MissUGMM, an ultrametric Gaussian mixture model which takes into account the missing at random mechanism for the unobserved values. Our approach is estimated using the expectation-maximization algorithm and achieves favorable results in comparison to other existing mixture models in simulations conducted with synthetic and benchmark data sets, even without a theorized ultrametric structure underlying the data. Furthermore, MissUGMM is applied to a realworld problem for exploring the sustainable development of cities across countries starting from incomplete information provided by municipalities. Overall, our results demonstrate that MissUGMM is a powerful and versatile model in dealing with missing data and is applicable to a broader range of real-world problems.

Keywords: ultrametricity, Gaussian mixture model, missing data, hierarchy of latent concepts, cities' sustainable development

Łukasz Smaga, Uniwersytet im. Adama Mickiewicza w Poznaniu

Tytuł: Przegląd metod funkcjonalnej analizy wariancji

Streszczenie: Analiza danych funkcjonalny (FDA) to gałąź statystyki skupiająca się na modelowaniu i wyciąganiu wniosków z obserwacji traktowanych jako funkcje, krzywe lub powierzchnie. Zapewnia skuteczne podejście do analizy jednej lub większej liczby zmiennych mierzonych w czasie lub przestrzeni, co jest częstym scenariuszem spotykanym w różnych dziedzinach. W prezentacji rozważymy analizę wariancji dla danych funkcjonalnych. Metody statystyczne rozwiązania tego problemu opierają się na różnych podejściach, na przykład na redukcji wymiaru (rozwińnięcie bazowe, funkcjonalne składowe główne), losowych projekcjach na dane wielowymiarowe i agregowaniu statystyk punktowych. Podczas wykładu przedstawiamy różne aspekty i strategie metod funkcjonalnej analizy wariancji. Ta ostatnia obejmuje w szczególności analizę jedno- i wielokierunkową, niezależne obserwacje i powtarzane pomiary. W prezentacji zostaną zaprezentowane rzeczywiste przykłady ilustrujące poruszane problemy.

Title: A Review of Methods for Functional Analysis of Variance

Abstract: Functional data analysis (FDA) is a branch of statistics that focuses on modeling and making inferences from observations treated as functions, curves, or surfaces. It provides a powerful approach for analyzing one or more variables measured over time or space, which is a common scenario encountered in various fields. In the talk, we consider the analysis of variance for functional data. The statistical methods for this problem are based on different approaches, for example, dimension reduction (basis expansion, functional principal components), random projections on multivariate data, and aggregating pointwise statistics. During the talk, we present the

different aspects and strategies of the functional analysis of variance methods. The latter covers, in particular, one-way, multi-way, independent observations, and repeated measurements. The presentation will feature real data examples to illustrate the problems discussed.

Papers / Referaty

Stanislav Anatolyev, Charles University, CERGE-EI

Title: AIC for many-regressor heteroskedastic regressions

Abstract: The original Akaike information criteria (AIC) and its corrected version (AICc) have been routinely used for model selection for ages. The penalty terms in these criteria are tied to the classical normal linear regression, characterized by conditional homoskedasticity and a small number of regressors relative to the sample size, which leads to very simple and computationally attractive penalty forms. We derive, from the same principles, a general version that takes account of conditional heteroskedasticity and regressor numerosity. The new AICm penalty takes a form of a ratio of certain weighted average error variances, and encompasses the classical ones: it is approximately equal to the AIC penalty when the regression is conditionally homoskedastic and regressors are few, and to the AICc penalty when the regression is conditionally homoskedastic but the number of regressors is not negligible. In contrast to those of AIC and AICc, the AICm penalty is stochastic and thus not immediately implementable, as it in addition depends on the pattern of conditional heteroskedasticity in the sample. The infeasible AICm criterion, however, can be operationalized via unbiased estimation of individual variances. The feasible AICm criterion still minimizes the expected Kullback-Leibler divergence up to an asymptotically negligible term that does not relate to regressor numerosity. In simulations, the feasible AICm does select models that deliver systematically better out-of-sample predictions than the classical criteria.

Keywords: linear regression, Akaike information criterion, conditional heteroskedasticity, many regressors

Iwona Bąk, Katarzyna Wawrzyniak, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

Tytuł: Bezpieczeństwo ekologiczne a przestrzenne zróżnicowanie województw w zakresie ochrony środowiska

Streszczenie: Bezpieczeństwo ekologiczne jest podstawą zrównoważonego rozwoju. Uznawane jest współcześnie za jeden z najważniejszych elementów niemilitarnej odmiany bezpieczeństwa. Celem badania jest zidentyfikowanie przestrzennego zróżnicowania województw w Polsce ze względu na stan realizacji przyjętych

norm środowiskowych. Wartości tych norm określono na podstawie analizy rozkładów przyjętych do badania wskaźników. Dane do obliczeń zaczerpnięto z Banku Danych Lokalnych GUS. Do oceny stopnia realizacji norm w województwach Polski w 2022 roku wykorzystano klasyczną analizę korespondencji na podstawie danych przedstawionych w złożonej macierzy znaczników oraz metodę Warda. Zastosowanie tych metod pozwoliło na wskazanie słabych i mocnych stron województw ze względu na stan ochrony środowiska. Okazało się, że najlepsza ze względu na cel badania była grupa obejmująca trzy województwa (lubelskie, pomorskie, wielkopolskie), w której 50% przyjętych do badania wskaźników okazało się być zgodnych z normą i nie zidentyfikowano w niej żadnego wskaźnika, który tej normy nie spełnia. Natomiast najgorsza sytuacja miała miejsce w województwach: opolskim i śląskim, w których nie zaobserwowano wskaźników zgodnych z normą, a cztery z nich normy nie spełniły.

Słowa kluczowe: bezpieczeństwo ekologiczne, ochrona środowiska, województwa, klasyczna analiza korespondencji, metoda Warda

Title: Ecological safety and the spatial differentiation of voivodeships in the area of environmental protection

Abstract: Ecological safety is a cornerstone of sustainable development. Nowadays it is considered as one of the most important elements of non-military security. The aim of the study is to identify the spatial differentiation of voivodeships in Poland in terms of the condition of implementation of adopted environmental standards. The values of these norms were determined on the basis of an analysis of the distributions of the indicators adopted for the study. Data for the calculations were taken from the Local Data Bank of Statistics Poland. To assess the degree of realisation of the standards in the Polish voivodeships in 2022, classical correspondence analysis was used on the basis of data presented in a binary matrix of indicators and Ward's method. The application of these methods made it possible to indicate the strengths and weaknesses of the voivodeships in terms of the state of environmental protection. It turned out that the best, considering the aim of the study, was a group containing three voivodeships (Lubelskie, Pomorskie and Wielkopolskie), in which 50% of the indicators adopted for the study turned out to be compliant with the standard and no indicator was identified that did not meet the standard. On the other hand, the worst situation was in the following voivodeships: Opolskie and Śląskie, in which no indicators compliant with the standard were observed, and four of them did not meet the standard.

Keywords: ecological safety, environmental protection, voivodeships, classical correspondence analysis, Ward's method

Tytuł: Propozycja skorygowanego jakościowo indeksu typu GEKS

Streszczenie: Szeroka gama sprzedawców detalicznych (supermarkety, sklepy z elektroniką domową, sklepy internetowe itp.) udostępniają dane ze skanera zawierające informacje o sprzedawanych produktach na poziomie kodu kreskowego, stosując np.: Globalny Numer Jednostki Handlowej (GTIN, ang. Global Trade Article Number) lub Europejski Numer Artykułu (EAN, ang. European Article Number). Ponieważ dane skanowane dostarczają pełnych informacji o transakcji, na najniższym poziomie agregacji danych możemy zastosować dowolną formułę indeksu cen, nawet formułę ważoną. Metody indeksów multilateralnych wydają się być najlepszym wyborem w przypadku tego rodzaju danych ze względu na dużą rotację zeskanowanych produktów oraz obserwowaną sezonowość produktów. Indeksy te operują na oknie czasowym i są tranzytywne, co jest kluczem do eliminacji efektu łańcuchowego dryfu. Zgodnie jednak z mocniejszym wymogiem w postaci testu identyczności możemy oczekiwać, że nawet jeśli tylko ceny powrócą do swoich pierwotnych wartości, indeks cenowy przyjmie wartość jeden. Niestety powszechnie stosowane formuły indeksów multilateralnych (np. GEKS, CCDI, GK, TPD, TDH) nie spełniają testu identyczności. W artykule omówiono autorską propozycję indeksu multilateralnego (GEKS-GAQI). Z jednej strony konstrukcja proponowanego indeksu opiera się na idei indeksu GEKS. Z drugiej strony, podobnie jak metoda Geary'ego-Khamisa, wymaga on dostosowania jakościowego (korekty jakościowej). Pokazano, że zaproponowany indeks spełnia test identyczności i większość innych pożądaných aksjomatów (testów). Zaproponowany indeks porównano z indeksami GEKS-L i GEKS-GL, które są stosunkowo nowe w literaturze i również spełniają test identyczności. Zaproponowany indeks porównujemy także z indeksami multilateralnymi wykorzystującymi korekty jakościowe (tj. z indeksami Geary-Khamisa, GEKS-AQU i GEKS-AQI). Przeprowadzone rozważania analityczne oraz badania empiryczne potwierdzają wysoką użyteczność zaproponowanego indeksu.

Słowa kluczowe: dane skanowane, wskaźnik cen towarów i usług konsumpcyjnych, indeksy multilateralne, test identyczności

Title: Proposition of the quality adjusted GEKS -type price index

Abstract: A wide variety of retailers (supermarkets, home electronics, Internet shops, etc.) provide scanner data containing information at the level of the barcode, e.g. the Global Trade Item Number (GTIN) or the European Article Number (EAN). As scanner data provide complete transaction information, we may use any price index formula, even a weighted one, at the lowest data aggregation level. Multilateral index methods seem to be the best choice in the case of scanner data due to high scanned product churn and seasonality of these products. Multilateral indices work on a whole-time window and are transitive, which is key to the elimination of the chain drift effect. Following what is called an identity test, however, it may be expected that even when only prices return to their original values, the index becomes one. Unfortunately, the commonly used multilateral index formulas (e.g. GEKS, CCDI, GK, TPD, TDH) do not meet the identity test. The paper discusses the proposal of new multilateral

index (GEKS-GAQI). On the one hand, the design of the proposed index is based on the idea of the GEKS index. On the other hand, similarly to the Geary-Khamis method, it requires quality adjusting. It is shown that the proposed index meets the identity test and most other desired axioms (tests). The proposed index is compared with the GEKS-L and GEKS-GL indices, which are relatively new and also meet the identity test. We also compare the proposed index with multilateral indices which use quality adjusting (i.e. with the Geary-Khamis, GEKS-AQU and GEKS-AQI indices). The analytical considerations as well as empirical studies confirm the high usefulness of the proposed index.

Keywords: scanner data, Consumer Price Index, inflation, price indices, multilateral indices, identity test

Jacek Bialek*, Anna Bobel**, Dagmara Oprych-Franków**, * Uniwersytet Łódzki,
** Główny Urząd Statystyczny**

Tytuł: Automatyczna detekcja downsizingu i upsizingu na danych skanowanych

Streszczenie: Pomiar zmian cen coraz częściej kojarzy się z „ukrytą” inflacją, gdy ceny produktów pozostają niezmienione, ale ze względu na zmniejszenie zawartości i wielkości opakowania konsumenci płacą więcej za jednostkę produktu. Redukcja opakowań, nie tylko w grupie towarów i produktów spożywczych, to proces zmniejszania ilości (wagi) produktów, a czasami nawet zmiany ich składu lub obniżenia jakości przy jednoczesnym utrzymaniu cen tych produktów na dotychczasowym poziomie. Czasem zdarza się, że zmniejszeniu ilości produktu w opakowaniu towarzyszy wzrost ceny, chociaż – jak pokazują badania – konsumentów bardziej zniechęca podwyżka ceny niż zmniejszenie wielkości opakowania. Praktyka kurczenia się produktów (downsizing) wpływa negatywnie na zdolność konsumentów do podejmowania świadomych wyborów zakupowych, a także negatywnie wpływa na wiarygodność i postrzeganie CPI. Innym niebezpiecznym zjawiskiem jest sytuacja, gdy producent zwiększa rozmiar produktu, ale wzrost ceny nie jest proporcjonalny (upsizing). Dużym wyzwaniem dla statystyki publicznej jest wykrywanie tego typu praktyk, a co za tym idzie wiarygodna weryfikacja próby produktów i automatyzacja doboru produktów poddawanych zmniejszeniu lub powiększeniu. W artykule przedstawiono wyniki badań autorów nad zjawiskami downsizingu i upsizingu jakie obserwujemy na dużych zbiorach danych skanowanych. W artykule zdefiniowano sześć rodzajów downsizingu i upsizingu oraz zaproponowano automatyczną procedurę detekcji tego rodzaju produktów. Na koniec zaprezentowano i omówiono wyniki analizy tych zjawisk dla wybranych grup zeskanowanych produktów spożywczych i niespożywczych.

Słowa kluczowe: dane skanowane, obniżanie rozmiaru produktu, shrinkflacja, podwyższanie rozmiaru produktu

Title: Automatic detection of downsizing and upsizing on scanner data

Abstract: Measuring price changes is increasingly associated with "hidden" inflation, when product prices remain seemingly unchanged, but due to the reduction in the content and size of packaging, consumers pay more per unit

of product. Reducing packaging, not only in the group of goods and food products, is a process of reducing the quantity (weight) of products and sometimes even changing their composition or lowering quality while maintaining the prices of these products at the current level. It sometimes happens that reducing the quantity of a product in a package is accompanied by an increase in price, even though, as research shows, consumers are more discouraged by price increases than by reductions in package size. The practice of shrinkflation negatively affects consumers' ability to make informed purchasing choices and also affects the credibility and perception of CPI. Another dangerous phenomenon is when the manufacturer increases the size of the product but the price increase is not proportional (upsizing). A big challenge for public statistics is the detection of this type of practices, and therefore reliable verification of a sample of products and automation of the selection of products being downsized or upsized. This paper presents the results of the authors' research on the phenomenon of downsizing and upsizing occurring when using scanner data. The paper defines six types of downsizing and upsizing and it proposes an automatic procedure for selecting these products affected. Finally, the results of the analysis of this phenomenon for selected food and non-food groups of scanned products are presented and discussed.

Keywords: scanner data, downsizing, shrinkflation, upsizing

Beata Bieszk-Stolorz*, Krzysztof Dmytrów*, Michał Bernard Pietrzak, * Uniwersytet Szczeciński, ** Politechnika Gdańska**

Tytuł: Jakość zużycia energii w procesie transformacji energetycznej w krajach UE – analiza wielowymiarowa

Streszczenie: Sektor energetyczny jest jedną z kluczowych gałęzi gospodarki i odgrywa istotną rolę w zrównoważonym wzroście gospodarczym. Zużycie energii w dużym stopniu przyczynia się do degradacji środowiska w wyniku emisji dwutlenku węgla i innych zanieczyszczeń ze spalania paliw kopalnych. Postępujące zmiany klimatyczne skłoniły wiele krajów do przejścia na energię odnawialną. Dzięki temu poprawia się jakość zużycia energii, co z kolei polepsza stan środowiska naturalnego. Celem badania jest wielowymiarowa ocena krajów UE pod względem jakości zużycia energii i szybkości adaptacji tego aspektu procesu transformacji energetycznej. W badaniu wykorzystano dane z bazy Eurostat za lata 2004-2021. Na ich podstawie skonstruowano rankingi krajów UE pod względem jakości zużycia energii. Wykorzystano w tym celu zmodyfikowaną metodę COPRAS, aby uwzględniła ona także aspekt dynamiczny. Następnie zbadano dynamikę krajów UE w zakresie transformacji energetycznej. Państwa zostały porównane pod względem dynamiki przy zastosowaniu metody Dynamic Time Warping. W ostatnim etapie badania państwa UE zostały pogrupowane pod względem podobieństwa dynamiki zmian jakości zużycia energii za pomocą grupowania hierarchicznego. Najlepszymi krajami pod względem jakości zużycia energii były Malta, Cypr i Austria, a najgorszymi Polska, Czechy i Słowacja. W przypadku większości krajów UE proces transformacji energetycznej był podobny i wskazywał na poprawę jakości zużycia energii. Jedynie w Rumunii jakość zużycia energii spadła w analizowanym okresie, a w Holandii pozostała na mniej więcej podobnym poziomie.

Słowa kluczowe: jakość zużycia energii, transformacja energetyczna, metoda COPRAS w wersji dynamicznej, metoda Dynamic Time Warping, analiza skupień

Title: Quality of energy consumption in the energy transition process in EU Countries – a multivariate analysis

Abstract: The energy sector is one of the key industries of the economy and plays an important role in sustainable economic growth. Energy consumption is a major contributor to environmental degradation through emissions of carbon dioxide and other pollutants from the burning of fossil fuels. Progressive climate change has prompted many countries to switch to renewable energy. This improves the quality of energy consumption, which in turn improves the environment. The aim of the study is to provide a multidimensional assessment of EU countries in terms of the quality of energy consumption and the dynamics of adaptation of this aspect of the energy transition process. The study uses data from the Eurostat database for the period 2004-2021, on the basis of which rankings of EU countries in terms of the quality of energy consumption were constructed. A modified COPRAS method was used for this purpose, in order to take into account the aspect of dynamics. The dynamics of changes of energy transition in the EU countries were then examined. The countries were compared in terms of dynamics using the Dynamic Time Warping method. In the final stage of the study, the EU countries were clustered in terms of the similarity of the dynamics of the changes in the quality of energy consumption using hierarchical grouping. The best countries in terms of energy consumption quality were Malta, Cyprus and Austria, and the worst were Poland, Czechia and Slovakia. For most EU countries, the energy transition process was similar and indicated an improvement in the quality of energy consumption. Only in Romania the quality of energy consumption declined during the analysed period, while in the Netherlands it remained at a more or less similar level.

Keywords: quality of energy consumption, energy transition, dynamic COPRAS method, Dynamic Time Warping method, cluster analysis

Stefano Bonni, Michela Borghesi, University of Ferrara

Title: Multivariate two-sample permutation test with directional alternative for categorical data

Abstract: This work concerns a distribution-free test, based on the permutation approach, on treatment effects with a multivariate categorical response variable. The motivating example is a typical case-control biomedical study, to investigate the effect of the treatment called “assisted motor activity” (AMA) on the health of comorbid patients affected by “low back pain” (LBP), “hypertension” and “diabetes”. Specifically, the goal is to test whether AMA determines an improvement in the functionality and perceived health status of patients. Two independent samples (treated and control group) are compared according to 13 different binary or ordinal outcomes. The null hypothesis of the test consists of the equality in the distribution of the multivariate responses of the two groups, whereas, under the alternative hypothesis, the health status of the treated patients is better. The approach proposed

in this work is based on the Combined Permutation Test (CPT) method, which is suitable for analyzing multivariate categorical data in the presence of confounding factors. A stratification of the groups and intra-stratum permutation univariate two-sample tests are conducted to avoid possible confounding effects. The p-values from the partial tests are combined using the CPT approach to create a suitable test statistic for the overall problem.

Keywords: nonparametric statistics, permutation test, multivariate statistics, categorical data

**Svitlana Chugaievska, Dariusz Fatuła, Uniwersytet Andrzeja Frycza Modrzewskiego
w Krakowie**

**Tytuł: Aktywność ekonomiczna członków gospodarstw domowych ukraińskich na
polskim rynku pracy w kontekście wojny w Ukrainie**

Streszczenie: Artykuł obejmuje badanie specyfiki aktywności ekonomicznej członków ukraińskich gospodarstw domowych na polskim rynku pracy w warunkach wojny na Ukrainie. Od rozpoczęcia pełnoskalowej inwazji Rosji na Ukrainę w lutym 2022 r. Polska stała się jednym z głównych kierunków dla ukraińskich uchodźców i migrantów poszukujących nie tylko bezpieczeństwa, ale także stabilności ekonomicznej. Autorzy analizują kluczowe czynniki wpływające na zaangażowanie obywateli Ukrainy na polskim rynku pracy, w szczególności prawne aspekty zatrudnienia, cechy demograficzne i poziom adaptacji do nowego środowiska. Szczególną uwagę zwrócono na strukturę zatrudnienia Ukraińców w Polsce, podział według sektorów gospodarki i różnice między regionami kraju. Rozważono również wyzwania, przed którymi stoją ukraińscy migranci, w tym barierę językową, uznawanie kwalifikacji i integrację społeczną. Opracowanie opiera się na analizie danych z badań polskich i międzynarodowych, a także na badaniu przymusowych migrantów ukraińskich w pierwszym i drugim roku wojny. Autorzy analizują również oficjalne dane statystyczne dotyczące rynku pracy w Polsce. Wnioski dotyczące wpływu wojny na procesy migracyjne oraz rekomendacje dotyczące poprawy warunków aktywności ekonomicznej migrantów ukraińskich na polskim rynku pracy mają znaczenie praktyczne.

Słowa kluczowe: Rynek pracy, migracje ukraińskie, gospodarstwa domowe, aktywność ekonomiczna, zatrudnienie.

**Title: Ukrainian households members' economic activities on the polish labor market in
the context of the war in Ukraine**

Abstract: The article covers the study of the peculiarities of the economic activity of members of Ukrainian households on the Polish labor market in the conditions of the war in Ukraine. Since the start of Russia's full-scale invasion of Ukraine in February 2022, Poland has become one of the main destinations for Ukrainian refugees and migrants seeking not only safety, but also economic stability. The authors analyze the key factors influencing the involvement of Ukrainian citizens in the Polish labor market, in particular the legal aspects of employment, demographic features and the level of adaptation to the new environment. Special attention is paid to the structure

of employment of Ukrainians in Poland, distribution by economic sectors and differences between regions of the country. Challenges faced by Ukrainian migrants, including the language barrier, recognition of qualifications and social integration, are also considered. The study is based on the analysis of data from Polish and international studies, as well as on a survey of forced Ukrainian migrants during the first and second years of the war. The authors also analyze official statistical data related to the labor market in Poland. Conclusions regarding the impact of the war on migration processes and recommendations for improving the conditions of economic activity of Ukrainian migrants on the Polish labor market are of practical importance.

Keywords: Labor market, Ukrainian migration, households, economic activity, employment

Marta Dziechciarz, Józef Dziechciarz, Uniwersytet Ekonomiczny we Wrocławiu

Tytuł: Pomiar i analiza subiektywnej percepcji dobrostanu. Wybrane aspekty dynamicznej oceny efektywności nakładów

Streszczenie: Przedmiotem badania jest analiza efektywności wydatków socjalnych w krajach europejskich na przestrzeni czasu. Metody użyte do analizy statystycznej to konwencjonalne i dynamiczne techniki analizy obwiedni danych (DEA). Główne wyniki uzyskano na podstawie danych EU-SILC i OurWorldInData.org analizowanych przy użyciu DEA. Badanie wykazało znaczące wahania subiektywnego dobrostanu, z wyraźną długoterminową tendencją wzrostową. Wśród badanych krajów są takie, w których postęp jest spowodowany czynnikami wewnętrznymi, podczas gdy w krajach śródziemnomorskich dominują czynniki zewnętrzne. Poziom szacowanych wskaźników efektywności DEA (TC i EC) wskazuje na poprawę efektywności nakładów w znacznej liczbie krajów, np. w Estonii i na Cyprze, przy stale niskich wartościach wskaźników w niektórych państwach członkowskich, we Włoszech i Luksemburgu. Podsumowując, autorzy wskazali obszary w poszczególnych krajach, w których konieczne są próby poprawy polityki społecznej w celu podniesienia i utrzymania pożądanych poziomów subiektywnego dobrostanu.

Słowa kluczowe: Subiektywna percepcja dobrostanu; gospodarstwo domowe; DEA; panel DEA; dynamiczna DEA; indeks Malmquista; drabina Cantrila; Paradoks Easterlina; analiza czasoprzestrzenna; ogólna satysfakcja z życia

Title: Measurement and analysis of subjective perception of Wellbeing. Selected aspects of dynamic assessment of input effectiveness

Abstract: The task of the study is the analysis of social expenditure efficiency in European countries over time. The methods used for statistical analysis are conventional and dynamic data envelopment analysis (DEA) techniques. The EU-SILC and OurWorldInData.org data analysed with DEA provided the main results. The study identified significant fluctuations in subjective well-being, with a clear long-term upward trend. Among the surveyed countries, there are those where progress is due to internal factors, while in the Mediterranean countries,

external factors predominate. The level of the estimated DEA efficiency indicators (TC and EC) indicate the improvement in input efficiency in a substantive number of countries, e.g., Estonia and Cyprus, along with persistently low indicator values in some member states, Italy and Luxemburg. In conclusion, the authors indicated areas in individual countries where attempts to improve social policy are needed to raise and maintain desirable levels of subjective well-being.

Keywords: subjective well-being; household; DEA; panel DEA; dynamic DEA; Malmquist index; Cantril ladder; Easterlin paradox; spatiotemporal analysis; overall life satisfaction

Piotr Fiszeder*, **Witold Orzeszko***, **Radosław Pietrzyk****, **Grzegorz Dudek*****,
* **Uniwersytet Mikołaja Kopernika w Toruniu**, ** **Uniwersytet Ekonomiczny we Wrocławiu** oraz **Empirica Sp. z o.o.**, *** **Politechnika Częstochowska**

Tytuł: Identyfikacja determinantów zmienności bitcoina przy wykorzystaniu modeli statystycznych oraz uczenia maszynowego

Streszczenie: Niniejsze badanie rozszerza prognozowanie zmienności Bitcoina (BTC) poprzez analizę obszernego zestawu 62 zmiennych objaśniających, w tym zachowania rynku kryptowalut, trendów wyszukiwania Google, indeksów finansowych i wskaźników ekonomicznych. W badaniu stosujemy metody Bayesian Model Averaging (BMA), Least Absolute Shrinkage and Selection Operator (LASSO) i Random Forest (RF) do oceny znaczenia zmiennych i dokładności prognoz. Nasze badania pokazują, że modele LASSO i RF uwzględniające zmienne egzogeniczne znacznie poprawiają zarówno dzienne, jak i tygodniowe prognozy zmienności BTC w porównaniu z modelami wykorzystującymi jedynie opóźnione zmienności BTC. Kluczowe czynniki wpływające na zmienność BTC obejmują opóźnione zrealizowane wariacje, wolumen obrotu i intensywność wyszukiwania w Google. Badanie pokazuje, że wpływ tych zmiennych na zmienność BTC jest zmienny w czasie, odzwierciedlając ewoluujący związek Bitcoina z szerszymi wskaźnikami ekonomicznymi i nastrojami rynkowymi. Uzyskane wnoszą wkład do literatury, zapewniając kompleksową analizę czynników wpływających na zmienność BTC, oceniając skuteczność transformacji zmiennych i porównując skuteczność zaawansowanych metod prognozowania w radzeniu sobie z ekstremalną zmiennością kryptowaluty. Spostrzeżenia te mogą być cenne dla badaczy, inwestorów, zarządzających portfelami i decydentów na dynamicznym rynku kryptowalut.

Słowa kluczowe: Zmienność kryptowalut; random forest; LASSO; Bayesian model averaging; HAR

Title: Identification of Bitcoin Volatility Drivers Using Statistical and Machine Learning Methods

Abstract: This study advances the understanding of Bitcoin (BTC) volatility forecasting by analyzing an extensive set of 62 explanatory variables, including cryptocurrency market behaviour, Google search trends, financial indices, and economic indicators. We employ Bayesian Model Averaging (BMA), Least Absolute Shrinkage and

Selection Operator (LASSO), and Random Forest (RF) methods to assess variable importance and forecast accuracy. Our research demonstrates that LASSO and RF models incorporating exogenous variables significantly improve both daily and weekly BTC variance forecasts compared to models using only lagged BTC volatilities. Key factors influencing BTC volatility include lagged realized variances, trading volume, and Google search intensity. The study reveals that the impact of these variables on BTC volatility is time-varying, reflecting Bitcoin's evolving relationship with broader economic indicators and market sentiment. Our findings contribute to the literature by providing a comprehensive analysis of BTC volatility drivers, evaluating the effectiveness of variable transformations, and comparing the performance of advanced forecasting methods in handling the cryptocurrency's extreme volatility. These insights are valuable for researchers, investors, portfolio managers, and policymakers navigating the dynamic cryptocurrency market.

Keywords: Cryptocurrency volatility; market and policy uncertainty; random forest; LASSO; Bayesian model averaging; HAR

Elżbieta Gołata, Uniwersytet Ekonomiczny w Poznaniu

Tytuł: Projekcje demograficzne – czy jesteśmy w stanie przewidzieć przyszłość?

Streszczenie: W 2023 roku GUS opublikował dwie projekcje demograficzne według krajowej definicji ludności oraz ludności rezydującej. W 2023 roku opublikowana została także projekcja Eurostatu: EuroPop2023, a w bieżącym roku ukazała się kolejna rewizja projekcji ONZ: World Population Prospects WPP 2024. W perspektywie roku 2060 projekcje te przewidują spadek liczby ludności Polski. Jednak w każdym przypadku szacunki są inne. Największy ubytek liczby ludności w Polsce przewiduje WPP 2024 i jest to ponad 7,1 mln osób. Według oficjalnej prognozy GUS (dla ludności krajowej) liczba ludności zmniejszy się o 6,8 mln. Natomiast Eurostat szacuje spadek ludności Polski o 4,3 mln. Zgodnie z prognozą GUS najmniej, bo o 3,8 mln osób zmniejszy się populacja ludności rezydującej. Uwzględniając oficjalną prognozę GUS i ONZ w 2060 r. ludność Polski zmniejszy się o 17-18% stanu w 2022 roku. W przypadku EuroPop2023 i projekcji GUS dla rezydentów zmniejszenie populacji stanowi 10-11% stanu wyjściowego. Wszystkie te projekcje zgodnie przewidują jeszcze większy, zarówno w wymiarze absolutnym (4,6-8,4 mln) jak i względnym (21-35%) ubytek ludności w wieku produkcyjnym. W referacie podjęto dwa wzajemnie powiązane wątki: wyjaśnienia przyczyn rozbieżnych szacunków oraz oceny wiarygodności projekcji. Próba odpowiedzi na pytanie o przyczyny rozbieżności obejmuje kilka zagadnień. Jest to w pierwszej kolejności ocena jakości danych wyjściowych, a więc definicji i szacunku liczby ludności. Następnie założenia dotyczące płodności i umieralności oraz migracji. Powodem rozbieżności jest także zastosowana metoda prognozowania. Wszystkie te czynniki różnicując wyniki projekcji wpływają na ocenę ich wiarygodności. Stosując metody analizy demograficznej oraz porównawczej analizy statystycznej przedstawiona zostanie ocena wiarygodności założeń odnośnie do płodności i umieralności. Wskazane zostaną także ograniczenia dotyczące szacunku ruchu wędrownego oraz powiązania migracji z procesami ruchu naturalnego. Rozbieżności stanów wyjściowych oraz dalszych szacunków i założeń mają swoje odzwierciedlenie w strukturze wieku ludności, która kształtuje ostateczny wynik projekcji. Złożony proces konstrukcji projekcji

oraz wzajemne powiązania poszczególnych jego elementów i etapów utrudniają jednoznaczną ocenę wiarygodności. Jednak poznanie procesów demograficznych umożliwia ocenę realności poszczególnych scenariuszy co ma strategiczne znaczenie dla rozwoju społecznego i gospodarczego.

Słowa kluczowe: projekcje demograficzne, płodność, umieralność, migracje

Title: Demographic projections – can we predict the future?

Abstract: In 2023 Statistics Poland (GUS) published two demographic projections: one based on the national population definition and the other on the resident population. Additionally, in 2023, Eurostat released its projection: EuroPop2023, and this year another revision of the United Nations' projection, World Population Prospects (WPP) 2024, was published. Looking ahead to 2060, all these projections predict a decline in Poland's population, though the estimates differ. The WPP 2024 projects the largest population loss for Poland, over 7.1 million people. According to GUS's official projection (for the national population definition), the population will decrease by 6.8 million. Meanwhile, Eurostat estimates a population decrease of 4.3 million. According to GUS, the resident population will decrease the least, by 3.8 million people. Considering GUS's official projection and the UN's estimates, by 2060, Poland's population will decrease by 17-18% from its 2022 level. In the case of EuroPop2023 and GUS's projection for residents, the population decline is expected to be 10-11% of the initial value. All these projections consistently forecast an even greater loss in the working-age population, both in absolute terms (4.6-8.4 million) and relative terms (21-35%). The paper addresses two interrelated issues: explaining the reasons for these differing estimates and assessing the credibility of the projections. The attempt to explain the discrepancies involves several considerations. First is the evaluation of the quality of the data, including the definition and estimation of the population. Next are the assumptions regarding fertility, mortality, and migration. The forecasting method used also contributes to the divergence. All these factors, by influencing the projection results, affect the assessment of their reliability. Using demographic analysis methods and comparative statistical analysis, the paper will present an evaluation of the credibility of assumptions about fertility and mortality. It will also highlight the limitations in estimating migration and the relation between migration and natural population processes. Discrepancies in the initial data, as well as assumptions and further estimates, are reflected in the age structure of the population, which shapes the final outcome of the projections. The complex process of constructing projections and the interconnections between its various elements make it difficult to definitively assess their reliability. However, understanding demographic processes allows for evaluating the plausibility of different scenarios, which is of strategic importance for social and economic development.

Keywords: demographic projections, fertility, mortality, migration

Wioletta Grzenda*, Agnieszka Marszałek, * Szkoła Główna Handlowa w Warszawie,
** badaczka niezależna**

**Tytuł: Rola wykształcenia i płci w kształtowaniu ścieżek zawodowych polskich
millenialsów: analiza z wykorzystaniem shared frailty survival models**

Streszczenie: Celem badania było zbadanie wpływu płci i wykształcenia na mobilność zawodową młodych pracowników na przykładzie polskiego rynku pracy. Skupiamy się na Polsce jako kraju, gdzie bezrobocie wśród młodych ludzi jest szczególnie niskie, a tradycyjne role społeczne związane z płcią nadal są uważane za ważne. Analizując zmiany pracy, wychodzimy poza dotychczasowe badania, biorąc pod uwagę zarówno czas trwania poszczególnych epizodów pracy, a także zmienność w czasie niektórych cech młodych ludzi, takich jak wykształcenie czy stan cywilny. Analizę przeprowadzono z wykorzystaniem metod analizy przeżycia, w tym shared frailty models. Korzystając z danych z badania Generation and Gender Survey, ujawniamy, że wpływ badanych czynników na mobilność zawodową zależy od płci. Otrzymaliśmy, że wpływ posiadania dziecka na mobilność zawodową był istotny jedynie w przypadku kobiet. Matki miały większe ryzyko zmian pracy w porównaniu z kobietami bezdzietnymi. Stabilizacja karier zawodowych mężczyzn następuje wraz z wiekiem i związana jest z opuszczeniem domu rodzinnego oraz zawarciem związku małżeńskiego. Natomiast, posiadanie wyższego wykształcenia w większym stopniu wpływało na ryzyko zmian pracy w przypadku mężczyzn niż w przypadku kobiet.

Słowa kluczowe: mobilność zawodowa, zmiany pracy, wykształcenie, płeć, shared frailty models, analiza przeżycia

**Title: The role of education and gender in shaping the career paths of polish millennials:
a shared frailty survival model analysis**

Abstract: Our study aims to examine the influence of gender and education on job mobility among young employees, using the example of the Polish labour market. We focus on Poland as a country where unemployment among young people is particularly low, but traditional gender-based social roles are still considered important. When analysing job changes, we go beyond previous studies by considering the duration of individual job episodes and the time-varying nature of some characteristics in young people, such as education or marital status. Our analysis was based on survival analysis methods including shared frailty survival models. Using data from the Generation and Gender Survey, we reveal that the impact of the examined factors on job mobility depends on gender. We received that the impact of having a child on job mobility was significant only for women. Mothers had a lower risk of job changes compared to childless women. The stabilisation of men's careers occurs with age and is associated with leaving the family home and marriage. However, in the case of women, these factors had no impact on professional mobility. Moreover, having higher education more significantly impacts the risk of job changes for men than women.

Keywords: job mobility, job changes, education, gender, shared frailty models, survival analysis

Adam Idczak, Jerzy Korzeniewski, Uniwersytet Łódzki

Tytuł: Analiza sentymentu w procesie podejmowania decyzji inwestycyjnych - akcje, kryptowaluty i towary

Streszczenie: Wykorzystanie informacji wyodrębnionych z nieustrukturyzowanych danych, takich jak dokumenty tekstowe, do podejmowania decyzji inwestycyjnych może być alternatywą dla tradycyjnych metod inwestycyjnych wykorzystujących ustrukturyzowane dane. Badanie ma na celu próbę weryfikacji powyższej hipotezy, w której informacje wyodrębnione z dokumentów tekstowych są sentymentem. Wydźwięk dokumentów tekstowych został określony przy użyciu dwuetapowego algorytmu. W pierwszym kroku algorytmu niektóre dokumenty są przypisywane do klasy dokumentów pozytywnych lub negatywnych przy użyciu zestawu reguł leksykalnych i gramatycznych oraz zestawu terminów kluczowych. Kluczowe terminy nie muszą być wprowadzane przez użytkownika, algorytm znajduje je samodzielnie. W drugim kroku pozostałe dokumenty są dołączane do jednej z klas przy użyciu reguł opartych na słownictwie znalezionym w dokumentach pogrupowanych w pierwszym kroku. W badaniu wykorzystano notowania i dane z Twittera z lat 2019-2023 dla spółek notowanych na rynku amerykańskim, kryptowalut i towarów.

Słowa kluczowe: eksploracja tekstu, sentyment dokumentów, grupowanie dokumentów, decyzje inwestycyjne, rynek akcji, kryptowaluty, towary

Title: Sentiment analysis in investment decisions - stock, cryptocurrencies and commodities

Abstract: Using information extracted from unstructured data such as text documents to make investment decisions can be an alternative to traditional investment methods using structured data. The study aims to try to verify the above hypothesis, in which the information extracted from text documents is sentiment. The output of text documents was determined using a two-step algorithm. In the first step of the algorithm, some documents are assigned to the class of positive or negative documents using a set of lexical and grammatical rules and a set of key terms. The key terms do not have to be entered by the user, the algorithm finds them on its own. In the second step, the remaining documents are attached to one of the classes using rules based on the vocabulary found in the documents grouped in the first step. The study used quotes and Twitter data from 2019-2023 for companies listed on the US market, cryptocurrencies and commodities.

Keywords: text mining, document sentiment, document clustering, investment decisions, stock market, cryptocurrencies, commodities

**Stanisław Jaworski, Wojciech Zieliński, Szkoła Główna Gospodarstwa Wiejskiego
w Warszawie**

**Tytuł: Optymalny dobór rozmiaru próby w estymacji przedziałowej w modelu Item
Count Technique dla pytań drażliwych**

Streszczenie: Jednym z problemów w badaniach ekonomiczno-społecznych jest oszacowanie odsetka pozytywnych odpowiedzi na pytanie drażliwe. Pytanie drażliwe jest takim pytaniem, na które respondent może nie chcieć udzielić rzetelnej odpowiedzi. W pracy podano optymalny rozmiar próby dla konstrukcji przedziału ufności dla tego odsetka.

Słowa kluczowe: pytania drażliwe, model Item Count Technique, przedział ufności, optymalny rozmiar próby

**Title: Optimal sample size in interval estimation for the fraction of sensitive questions in
a Item Count Technique model**

Abstract: In socioeconomic research estimation of the fraction of positive answers to a sensitive question. The sensitive question is such that the respondent would not give a honest answer. In the paper the problem of optimal sample size in construction of a confidence interval for such fraction is discussed.

Keywords: sensitive questions, Item Count Technique model, confidence interval, optimal sample size

Magdalena Kawecka, Zuzanna Krysiak, Uniwersytet Ekonomiczny w Katowicach

Tytuł: Migracje pracowników w kontekście starzenia się społeczeństwa w Polsce

Streszczenie: W obliczu dynamicznych zmian demograficznych, w tym starzenia się społeczeństwa, problematyka migracji pracowników nabiera szczególnego znaczenia dla rynku pracy w Polsce. Niniejszy artykuł analizuje wpływ migracji, zarówno napływu pracowników z zagranicy, jak i emigracji młodych Polaków, na sytuację demograficzną i strukturalne zmiany w zatrudnieniu. Celem pracy jest zrozumienie, w jaki sposób migracje mogą modyfikować tendencje związane ze starzeniem się społeczeństwa oraz jak wpływają na kształtowanie się rynku pracy w Polsce. Dotychczasowe badania skupiały się przede wszystkim na wpływie starzenia się populacji na rynek pracy, natomiast w mniejszym stopniu analizowały rolę migracji w tym kontekście. W ramach badań podjęto analizę danych statystycznych dotyczących przepływów migracyjnych oraz ich wpływu na strukturę wiekową populacji i poziom aktywności zawodowej. Wyniki badań mogą dostarczyć cennych informacji dla polityków i decydentów w zakresie opracowywania strategii na rzecz zrównoważonego rozwoju demograficznego i ekonomicznego kraju.

Słowa kluczowe: Migracje; Rynek pracy; Demografia

Title: Worker migration in the context of an ageing society in Poland

Abstract: In the face of dynamic demographic changes, including the ageing of society, the issue of employee migration is becoming particularly important for the Polish labor market. This article analyzes the impact of migration, both the inflow of workers from abroad and the emigration of young Poles, on the demographic situation and structural changes in employment. The aim of the work is to understand how migration can modify the trends related to the ageing of society and how it affects the formation of the labor market in Poland. Previous studies have focused primarily on the impact of population ageing on the labor market, while to a lesser extent they have analyzed the role of migration in this context. The research involved the analysis of statistical data on migration flows and their impact on the age structure of the population and the level of professional activity. The research results can provide valuable information for politicians and decision-makers in the development of strategies for the sustainable demographic and economic development of the country.

Keywords: Migration; Labor Market; Demographics

Grzegorz Kończak, Uniwersytet Ekonomiczny w Katowicach

Tytuł: Test zgodności dla wielowymiarowej normalności

Streszczenie: W referacie przedstawiona zostanie propozycja wielowymiarowego testu normalności z ustalonym wektorem wartości oczekiwanych i macierzą wariancji-kowariancji. Idea testu została oparta na znanym w literaturze teście pustych cel. Na podstawie próby n -elementowej wyznaczana jest liczba pustych cel, tzn. cel, w których nie znalazł się żaden element z próby. Zaletą testu jest możliwość stosowania go dla stosunkowo nielicznych prób. W referacie przedstawione zostaną wyniki symulacyjnej analizy mocy proponowanego testu.

Słowa kluczowe: test zgodności, wielowymiarowa normalność, wnioskowanie statystyczne, test pustych cel, symulacje komputerowe

Title: On a new goodness of fit test for multivariate normality

Abstract: The paper proposes a multivariate normality test with a fixed vector of expected values and a variance-covariance matrix. The idea of the test is based on the well-known empty cells test in the literature. Based on the n -element sample, the number of empty cells is determined, i.e., cells in which no element from the sample was found. The advantage of the test is that it can be applied to relatively small samples. The paper will present the results of a simulation power analysis of the proposed test.

Keywords: goodness of fit test, multivariate normality, inferential statistics, empty cells test, Monte Carlo study

Tytuł: Struktura miksu energetycznego na rynkach o podobnej zmienności cen energii elektrycznej

Streszczenie: Ze względu na trwającą transformację rynków energii w miksie energetycznym zwiększa się udział energii wytwarzanej ze źródeł odnawialnych. Z jednej strony daje to nadzieję na niższe ceny energii elektrycznej, z drugiej – ze względu na niestabilność źródeł energii takich jak słońce i wiatr – może przyczyniać się do wzrostu zmienności cen. W badaniu analizujemy średnie dzienne ceny energii elektrycznej na wybranych europejskich rynkach dnia następnego oraz dzienne udziały energii wytwarzanej z różnych źródeł. Za pomocą modeli dyskryminacyjnych mieszanek dla funkcjonalnych szeregów czasowych wyodrębniliśmy grupy rynków podobnych pod względem zmienności cen energii elektrycznej. Następnie przeprowadziliśmy analizę otrzymanych grup pod względem struktury miksu energetycznego. Celem badania było sprawdzenie czy rynki o podobnej zmienności cen posiadają podobną strukturę energii wytwarzanej z różnych źródeł. W szczególności interesowało nas, czy większej zmienności cen towarzyszy wyższy udział energii z niestabilnych źródeł. Uzyskane wyniki mogą być interesujące w kontekście badań nad optymalnym składem miksu energetycznego.

Słowa kluczowe: energia elektryczna, zmienność cen, miks energetyczny, funkcjonalne szeregi czasowe, grupowanie

Title: Energy mix structure in markets with similar electricity price volatility

Abstract: Due to the transformation of energy markets, the share of energy generated from renewable sources is increasing in the energy mix. On the one hand, this offers hope for lower electricity prices, but, on the other hand, the vulnerability of energy sources such as solar and wind may contribute to increased price volatility. In the study, we analyze average daily electricity prices in selected European day-ahead markets and the daily shares of energy generated from different sources. Using discriminant mixture models for functional time series, we identified groups of markets similar in terms of electricity price volatility. We then analyzed the obtained groups in terms of the structure of the energy mix. This study aimed to examine whether markets with similar price volatility have a similar structure of energy generated from different sources. In particular, we were interested in whether higher price volatility is accompanied by a higher share of energy from unstable sources. The results obtained may be of interest in the context of research on the optimal composition of the energy mix.

Keywords: electricity, price volatility, energy mix, functional time series, clustering

Tytuł: Obciążenie estymacji parametrów rozkładu dochodów na podstawie prób nielosowych: Symulacja na bazie danych jednostkowych z badania EU-SILC

Streszczenie: Celem badania jest ocena obciążenia hipotetycznej strategii badania społecznego, w której podstawą wnioskowania jest próba nielosowa, a estymacja jest wspomagana informacjami dodatkowymi z niezależnej, losowej próby referencyjnej oraz całej populacji w zakresie zmiennych wspomagających, a przedmiotem badania są parametry rozkładu dochodów gospodarstw domowych. Badanie opiera się na symulacji Monte Carlo na zbiorze danych jednostkowych z badania EU-SILC zrealizowanego w Polsce w 2022 r. Badanie autorskie polega na odtworzeniu prawdopodobnej populacji, pobieraniu prób nielosowych oraz losowych prób odniesienia, estymacji parametrów populacji i ocenie obciążenia wyników. Czynniki kontrolowanymi w symulacji są: stopień wadliwości próby nielosowej, typ parametru populacji, rodzaj estymatora, liczba zmiennych pomocniczych i wielkość próby. Wyniki wskazują, że wadliwość próby nielosowej wpływa negatywnie na obciążenie szacunków. Wykorzystanie zmiennych społeczno-demograficznych do estymacji parametrów rozkładu dochodów z próby nielosowej zwykle zmniejsza obciążenie szacunków, ale rezultat może być niezadowolający, jeżeli wadliwość próby jest duża.

Słowa kluczowe: próby nielosowe, estymacja podwójnie odporna, kalibracja, EU-SILC

Title: The bias of the estimation of income distribution parameters in non-probability surveys: an experiment based on EU-SILC microdata

Abstract: The aim of the study is to assess the bias of a hypothetical sampling strategy, in which the basis for inference is a non-probability sample, and the estimation is strengthened by additional information on auxiliary variables from an independent probability sample and the entire population, and the subject of the study is the parameters of household income distribution. The study is based on a Monte Carlo simulation on a set of unit data from the EU-SILC survey carried out in Poland in 2022. The author's study consists of reconstructing the probable population, drawing non-probability samples and probability reference samples, estimating population parameters, and assessing the bias of the results. The factors controlled in the simulation are the non-probability sample defectiveness, the type of population parameter, the type of estimator, the number of auxiliary variables, and the sample size. The results indicated that the defectiveness of a non-probability sample has a negative impact on the bias of estimates. The use of socio-demographic variables to estimate the parameters of the income distribution from a non-probability sample usually reduces the bias of the estimates, but the result may be unsatisfactory if the data defect index is large.

Keywords: non-probability samples, doubly robust estimation, calibration, EU-SILC

Tytuł: O własnościach predyktorów typu plug-in w przypadku pewnej błędnej specyfikacji modelu

Streszczenie: Rozważamy problem predykcji wybranych charakterystyk w domenach z użyciem predyktorów typu plug-in. Analizy przeprowadzone zostały przy założeniu liniowego modelu mieszanego z dwoma skorelowanymi wektorami efektów losowych. W badaniu uwzględniono także problem błędnej specyfikacji modelu wynikającej z heteroscedastyczności efektów losowych. Celem referatu jest symulacyjna analiza własności rozważanych predyktorów. Szczególna uwaga została zwrócona na dokładność oraz precyzję oszacowań. Analizy Monte Carlo przeprowadzono z wykorzystaniem języka R oraz rzeczywistych danych o charakterze ekonomicznym.

Słowa kluczowe: plug-in, predyktor, statystyka małych obszarów, liniowe modele mieszane

Title: On the properties of plug-in predictors for some model misspecification

Abstract: We consider the problem of predicting selected characteristics in domains using plug-in predictors. The analyses were carried out assuming a linear mixed model with two correlated random effects vectors. The study also considered the problem of misspecification of the model resulting from the heteroscedasticity of random effects. The purpose of the paper is a simulation analysis of the properties of the considered predictors. Particular attention was paid to the accuracy and precision of the estimates. Monte Carlo analyses were carried out using the R language and real economic data.

Keywords: plug-in, predictor, small area estimation, linear mixed models

Paweł Kufel*, Marcin Błażejowski*, Jacek Kwiatkowski, * Uniwersytet WSB Merito w Toruniu, ** Uniwersytet Mikołaja Kopernika**

Tytuł: Baysowskie uśrednianie modeli wektorowej autoregresji - automatyczna procedura w gretlu

Streszczenie: W referacie zostanie przedstawiona bayesowska procedura uśredniania modeli wektorowej autoregresji (VAR). W tej koncepcji, do wnioskowania dopuszczana jest większa liczba modeli, a następnie dokonywane jest uśrednianie wyników na podstawie nie jednego, a wielu modeli. Na podstawie bayesowskich oszacowań przestrzeni modeli wyliczane są uśredniane wartości ocen parametrów strukturalnych oraz odchylenia standardowe tych ocen. Wyznaczane są również dodatkowe miary jak prawdopodobieństwa modeli (posterior model probability) oraz prawdopodobieństwo włączenia zmiennej do modelu (posterior inclusion probability).

Implementacja tego rozwiązania będzie zastosowana w oprogramowaniu gretl w formie skryptów oraz pakietów. Prezentacja działania procedury oraz interpretacja uzyskanych

Słowa kluczowe: BMA, VAR, baysowskie uśrednianie modeli, gretl

Title: Bayesian averaging of vector autoregressive models - an automatic procedure in gretl

Abstract: We present a Bayesian procedure for averaging (BMA) vector autoregressive (VAR) models. In this concept, a larger number of models are allowed for estimation, and then the results are averaged on the basis of not one but many models. Our procedure is based on Bayesian estimates of the many models. We calculate the averaged values of estimates based on not only a single model but the entire model space. Additional measures like posterior model probability and posterior inclusion probability of a variable are also determined. The implementation of this solution will be used in gretl software in the form of scripts and packages. The procedure and interpretation of the obtained results will be presented using an empirical example.

Keywords: BMA, VAR, bayesian model averaging, gretl

Joanna Landmesser-Rusek, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Tytuł: Porównanie dwóch rodzajów sieci topologicznych dla rynku walutowego: bazujących na współczynnikach korelacji oraz na koncepcji przyczynowości

Streszczenie: Sieci topologiczne pozwalają na rozpoznanie własności strukturalnych rynku walutowego. Sieci takie mogą być konstruowane na podstawie wartości współczynników korelacji między parami walut, a popularny algorytm minimalnego drzewa rozpinającego (minimum spanning tree, MST) pozwala zrozumieć istotne relacje zachodzące na rynku. Alternatywna miara odległości, bazująca na koncepcji przyczynowości (np. Grangera) dla szeregów czasowych, umożliwia pomiar nie tylko siły powiązań pomiędzy parami walut, ale także kierunkowości tych relacji. Odpowiednikiem MST na grafie skierowanym jest drzewostan o minimalnym koszcie (minimum-cost arborescence, MCA). Zaletą sieci przyczynowości jest fakt, iż mogą być one pomocne w badaniu dynamicznej propagacji szoków w systemie walutowym. W pracy budowane będą w sposób kroczący sieci powiązań dla najważniejszych walut światowych w okresie od 1.01.2019 do 31.09.2024. Porównanie dwóch typów konstruowanych sieci – korelacyjnych oraz przyczynowości – dokonane zostanie na podstawie wyznaczonych charakterystyk topologicznych sieci, takich jak gęstość, średnia odległość, wskaźnik centralizacji, stopnie wierzchołków.

Słowa kluczowe: sieci topologiczne, rynek walutowy, korelacja, przyczynowość

Title: Comparison of two types of topological networks for the foreign exchange market: based on correlation coefficients and on the concept of causality

Abstract: Topological networks make it possible to recognize the structural properties of the currency market. Such networks can be constructed based on the value of correlation coefficients between currency pairs, and the popular minimum spanning tree (MST) algorithm makes it possible to understand the relevant relationships occurring in the market. An alternative measure of distance, based on the concept of causality for time series (e.g. Granger causality), allows measuring not only the strength of relationships between currency pairs but also the directionality of these relationships. The equivalent of MST on a directed graph is minimum-cost arborescence (MCA). The advantage of causality networks is that they can help study the dynamic propagation of shocks in the currency system. In this work, correlation networks will be constructed in a stepwise manner for the most important world currencies in the period from 1.01.2019 to 31.09.2024. The comparison of two types of constructed networks - correlation and causality networks - will be made based on the determined topological characteristics of the network, such as density, average distance, centralization index, and degrees of vertices.

Keywords: topological networks, foreign exchange market, correlation, causality

Paweł Lula, Uniwersytet Ekonomiczny w Krakowie

Tytuł: Wykorzystanie modeli typu BERTopic w analizie polskiego dorobku publikacyjnego w obszarze ekonomii, finansów i zarządzania

Streszczenie: Modele klasy BERTopic należą do grupy najnowszych i najbardziej zaawansowanych metod identyfikacji tematów występujących w zbiorze dokumentów tekstowych. Wykorzystują reprezentację tekstu opartą na osadzeniach i składają się z wielu połączonych ze sobą modeli neuronowych. Są w stanie wygenerować łatwe w interpretacji opisy zidentyfikowanych tematów. Stanowią alternatywę dla znacznie prostszych modeli klasy LDA (Latent Dirichlet Allocation Model). W prezentacji zostaną porównane oba podejścia i na tej podstawie sformułowane będą wnioski wskazujące, czy zastosowanie bardzo złożonych modeli klasy BERTopic jest uzasadnione istotnym polepszeniem jakości wyników w porównaniu z wynikami dostarczonymi przez modele klasy LDA. W części empirycznej pracy zostaną pokazane wyniki analizy polskiego dorobku publikacyjnego w obszarze ekonomii, finansów i zarządzania zrealizowane za pomocą modeli klasy BERTopic.

Słowa kluczowe: eksploracyjna analiza tekstu, identyfikacja tematów, modele BERTopic, modele LDA

Title: The application of BERTopic models in the analysis of Polish research publication in the fields of economics, finance and management

Abstract: BERTopic models belong to a group of the latest and most advanced methods of topic identification existing in a collection of text documents. They use an embedding-based text representation and consist of multiple

interconnected neural networks models. They are able to generate easy-to-interpret descriptions of identified topics. They are an alternative to the much simpler Latent Dirichlet Allocation Model (LDA) class of models. The presentation will compare these two approaches and, on this basis, draw conclusions indicating whether the use of highly complex BERTopic class models is justified by a significant improvement in the quality of the results compared to those provided by LDA models. The empirical part of the paper will show the results of the analysis of the Polish research publication in the field of economics, finance and management realised with the use of BERTopic models.

Keywords: topic identification, BERTopic models, Latent Dirichlet Allocation models

Wojciech Łukaszonek*, Marcin Szymkowiak, Waldemar Wołyński**, * Uniwersytet Kaliski, ** Uniwersytet Ekonomiczny w Poznaniu**

Tytuł: Przestrzenno-czasowa analiza składowych głównych na potrzeby klasyfikacji lokalnych rynków pracy dla powiatów województwa wielkopolskiego

Streszczenie: Dynamika oraz kierunki obserwowanych w kontekście rynku pracy zmian stanowią odzwierciedlenie głębokich transformacji cywilizacyjnych, zarówno w obszarze gospodarki, jak i społecznym. Ograniczenia elastyczności tych rynków, szczególnie w zakresie podaży, wynikają m.in. z długotrwałych procesów kształcenia i rozwoju kompetencji zawodowych. W związku z tym analiza rynków pracy, nawet na poziomie lokalnym, następuje poważnych wyzwań metodologicznych. Niezwykle istotnym aspektem wspomnianych badań są występujące na rynku pracy powiązania przestrzenne, szczególnie wyraźne w bardziej rozwiniętych regionach oraz ich częściach, takich jak metropolie czy aglomeracje, przyciągające siłę roboczą z mniej rozwiniętych obszarów. Niniejszy referat demonstruje możliwości przestrzenno-czasowej analizy składowych głównych (STPCA) w ocenie zróżnicowania lokalnych rynków pracy (na przykładzie powiatów województwa Wielkopolskiego). Metoda ta pozwala uwzględnić zarówno zmienność w czasie, jak i wzajemne zależności przestrzenne badanych jednostek. W ramach prezentacji skoncentrowano się na klasyfikacji powiatów województwa wielkopolskiego, opierając się na 12 parametrach związanych z rynkiem pracy, rejestrowanych w latach 2004-2022, a jako ich źródło wykorzystano Bank Danych Lokalnych. Uwzględniono przy tym różne macierze odległości oraz miary autokorelacji przestrzennej. Uzyskane wyniki poddano grupowaniu metodą Warda prowadzącą ostatecznie do klasyfikacji powiatów województwa wielkopolskiego w kontekście ich sytuacji na rynku pracy, przy uwzględnieniu zarówno relacji czasowych, jak i przestrzennych. Wyników zostanie przedstawiona na przykładzie empirycznym.

Słowa kluczowe: STPCA, rynek pracy, klasyfikacja powiatów

Title: Spatio-temporal principal component analysis for the classification of local labour markets for the counties of the Wielkopolska voivodeship

Abstract: The dynamics and directions of the changes observed in the context of the labour market reflect profound civilizational transformations, both in the economic and social spheres. The limitations on the flexibility of these markets, especially in terms of supply, are due, among other things, to the long-standing processes of training and development of professional competence. Consequently, the analysis of labour markets, even at the local level, poses serious methodological challenges. An extremely important aspect of the aforementioned research is the spatial linkages that exist in the labour market, which are particularly pronounced in more developed regions and their parts, such as metropolises or agglomerations, attracting labour from less developed areas. This paper demonstrates the possibilities of spatio-temporal principal component analysis (STPCA) in assessing the differentiation of local labour markets (using the example of the counties of the Wielkopolska voivodeship). This method makes it possible to take into account both temporal variation and spatial interdependence of the studied units. The presentation focuses on the classification of the counties of the Wielkopolska voivodeship, based on 12 parameters related to the labour market, recorded in 2004-2022, with the Local Data Bank used as their source. Various distance matrices and measures of spatial autocorrelation were taken into account. The obtained results were subjected to clustering using Ward's method, ultimately leading to the classification of the counties of Wielkopolska voivodeship in the context of their labour market situation, taking into account both temporal and spatial relationships.

Keywords: STPCA, labour market, classification of the counties

Stanisław Maciej Kot, Politechnika Gdańska

Tytuł: Wyznaczanie skali ekwiwalentności, awersji do nierówności dochodowych i do nierówności rangowych z pojedynczego rozkładu dochodów

Streszczenie: W ekonomii dobrobytu stosowanej dobrobyt społeczny ucieleśniony w danym rozkładzie dochodów ocenia abstrakcyjny planista społeczny, SP, za pomocą indywidualnej funkcji dobrobytu społecznego (SWF). Ponieważ każdy członek społeczeństwa może pełnić rolę SP, to może istnieć tyle odmiennych ocen dobrobytu ilu jest owych członków. Kot (2024) zaproponował metodę wyznaczania pojedynczej reprezentacji owych wielu ocen dobrobytu społecznego. Autor założył dwie generalne rodziny planistów społecznych. Pierwsza rodzina $\{SP_\varepsilon\}$, indeksowana parametrem $\varepsilon > 0$ awersji do nierówności dochodowych, odpowiada SWF $_\varepsilon$ Atkinsona. Druga rodzina $\{SP_\nu\}$, indeksowana parametrem $\nu > 1$ awersji do nierówności rangowej, odpowiada SEF $_\nu$ indukowanej przez uogólniony indeks Giniego $G(\nu)$. Autor wykazał istnienie par SP_ν SP_ε , splecanych dobrobytowo takich, że $SWF_\nu = SWF_\varepsilon$. Dla empirycznego ciągu (v_i, ε_i) , $i=1, 2, \dots, n$, istnieje pojedyncza para (ν^*, ε^*) reprezentująca wszystkie n par. W owym artykule dochody gospodarstw domowych korygowano za pomocą potęgowej skali ekwiwalentności z parametrem θ przyjmującym arbitralną wartość 0.5. Obecny artykuł oferuje ogólną metodę wyznaczania zarówno ν , ε jak i θ na podstawie danego rozkładu dochodów gospodarstw domowych. Metodę tę

zastosowano do oceniania dobrobytu społecznego i nierówności dla 27 krajów-członków Unii Europejskiej w 2021 roku.

Słowa kluczowe: Awersja do nierówności, skala ekwiwalentności, dobrobyt społeczny

Title: Eliciting an equivalence scale, aversion to income inequality, and to rank inequality from a single income distribution

Abstract: In Applied Welfare Economics, an abstractive social planner, SP, assesses social welfare embodied in a given income distribution using an individual social welfare function (SWF). As every member of society may play the role of SP, there could be as many distinct welfare assessments of the same income distribution as society members. Kot(2024) proposed a method of finding a single representation of such many social welfare assessments. He assumed two general families of SPs. The first family, $\{SP_\varepsilon\}$, indexed by the parameter $\varepsilon > 0$ of aversion to income inequality, corresponds to Atkinson's SWF_ε . The second family, $\{SP_\nu\}$, indexed by the parameter $\nu > 1$ of aversion to rank inequality, corresponds to SEF_ν induced by the generalised Gini index $G(\nu)$. The author showed that there are pairs of welfare-entangled SP_ν and SP_ε , such as $SWF_\nu = SWF_\varepsilon$. For an empirical sequence of the pairs (ν_i, ε_i) , $i=1, 2, \dots, n$, a single pair (ν^*, ε^*) exists, representing all n pairs. In that paper, household incomes were adjusted by the power equivalence scale with the parameter θ arbitrarily set to 0.5. The current paper offers a general method of eliciting ν , ε , as well as θ from a given distribution of household incomes. This method is applied to assess social welfare and inequality for 27 European Union member countries in 2021.

Keywords: inequality aversion, equivalence scale, social welfare

Anna Majdzińska, Uniwersytet Łódzki

Tytuł: Wzorce struktur demograficznych jednoosobowych gospodarstw domowych w Polsce w świetle wyników NSP 2002, 2011 i 2021

Streszczenie: Opracowanie przedstawia analizy porównawcze struktur wieku ludności w jednoosobowych gospodarstwach domowych w Polsce w świetle wyników NSP 2002, 2011 i 2021 (ogółem i według płci). W tym celu przeprowadzona została dekompozycja zmian w strukturze wieku populacji, a także wykorzystano wybrane narzędzia wielowymiarowej analizy porównawczej. Źródło danych stanowiły publikacje Głównego Urzędu Statystycznego.

Słowa kluczowe: 1-osobowe gospodarstwa domowe, struktura ludności wg wieku, NSP 2002, 2011 i 2021, Polska

Title: Demographic structures of one-person households in Poland based on the National Census results in 2002, 2011, and 2021

Abstract: The study presents comparative analyzes of population age structures in one-person households in Poland in the light of the results of the National Census 2002, 2011 and 2021 (total and by gender). For this purpose, a decomposition of changes in the age structure of the population was carried out, and selected multivariate statistical methods were used. The source of the data were publications of Statistics Poland.

Keywords: one-person households, population age structure, National Censuses 2002, 2011 and 2021, Poland

Agnieszka Orwat-Acedańska, Uniwersytet Ekonomiczny w Katowicach

Tytuł: Analiza przestrzenna współczynnika dzietności w powiatach Polski

Streszczenie: Współczynnik dzietności w Polsce od lat przyjmuje wartości poniżej poziomu prostego zastępowania pokoleń, co jest szczególnie niepokojące w obliczu starzejącego się społeczeństwa. W połowie 2024 roku wskaźnik ten wynosił 1,16, i był najniższy w historii Polski i jednym z najniższych w Europie. Przeprowadzanie badań statystycznej analizy przestrzennej współczynnika dzietności na poziomie powiatów ma kluczowe znaczenie dla zrozumienia demograficznych trendów, prognozowania wartości tego współczynnika na poziomie powiatów oraz formułowania skutecznych polityk społecznych. Głównym celem artykułu jest identyfikacja zależności przestrzennych zachodzących pomiędzy powiatami Polski pod względem poziomu współczynnika dzietności ogólnej. Realizacji celu służy zastosowanie miar i testów globalnej oraz lokalnej autokorelacji przestrzennej Morana i Geary'ego. Okresem badawczym są lata 2002-2023. Dane wykorzystane w badaniu pochodzą z Banku Danych Lokalnych GUS. Identyfikację zależności przestrzennych poprzedzono analizą szeregów czasowych współczynnika dzietności na poziomie powiatów. Przeprowadzone badanie wskazuje na istotne zróżnicowanie przestrzenne oraz dużą zmienność pozycji powiatów w hierarchii wartości WDO w ostatnich dekadach w Polsce. Z przeprowadzonych analiz statystyk globalnych wynika, że w badanym okresie pomiędzy powiatami występowała dodatnia autokorelacja przestrzenna współczynnika dzietności. Natomiast analiza otrzymanych statystyk lokalnych umożliwiła wyodrębnienie klastrów podobnych powiatów oraz powiatów odstających.

Słowa kluczowe: autokorelacja przestrzenna, współczynnik dzietności ogólnej, powiaty

Title: Spatial analysis of the fertility rate in Polish counties

Abstract: The total fertility rate in Poland has been below the level of simple generation replacement for years, which is particularly worrying in the face of an aging society. In mid-2024, this indicator was 1.16, and was the lowest in the history of Poland and one of the lowest in Europe. Conducting statistical spatial analysis of the fertility rate at the county level is crucial for understanding demographic trends, forecasting the value of this rate at the county level, and formulating effective social policies. The main aim of the article is to identify spatial

dependencies between Polish counties in terms of the total fertility rate. The objective is achieved by using Moran and Geary's measures and tests of global and local spatial autocorrelation. The research period is 2002-2023. The data used in the study come from the Local Data Bank of the Central Statistical Office. The identification of spatial dependencies was preceded by the analysis of time series of the fertility rate at the counties level. The conducted research indicates significant spatial differentiation and high variability of the position of counties in the hierarchy of total fertility rate values in recent decades in Poland. The analyzes of global statistics show that in the examined period there was a positive spatial autocorrelation of the fertility rate between counties. However, the analysis of the obtained local statistics made it possible to distinguish clusters of similar counties and outlier counties.

Keywords: spatial autocorrelation, total fertility rate, counties

Michał Bernard Pietrzak, Politechnika Gdańska

Tytuł: Wykorzystanie przestrzennej macierzy wag w procedurze wyznaczania modelu ilorazu potencjałów

Streszczenie: W przestrzennych badaniach ekonomicznych stosowany jest model potencjału, który zastosowany do wybranych zjawisk ekonomicznych pozwala na określenie potencjału wszystkich analizowanych regionów. W wykonywanych analizach regionalnych stosowany jest także ilorazu potencjału, który stanowi iloraz potencjału dochodu oraz potencjału ludności. Wielkości potencjału dochodu lub potencjału ludności dla wybranych regionów wyznacza się uwzględniając wartości zjawiska ekonomicznego dla wszystkich regionów oraz odległości fizycznej między regionami. Istotną determinantą wielkości potencjału regionów, oprócz odległości fizycznej, stanowi ich podobieństwo ekonomiczne, wyrażone za pomocą odległości ekonomicznej. Oznacza to, że zastosowanie odległości ekonomicznej w procedurze wyznaczania ilorazu potencjału powinno pozwolić na uwzględnienie zależności przestrzennych między regionami, które wynikają z ich wzajemnych powiązań społeczno-ekonomicznych. W związku z tym w artykule przedstawiona zostanie propozycja wykorzystania przestrzennej macierzy wag w procesie wyznaczania odległości ekonomicznej oraz modelu modelu ilorazu potencjałów.

Słowa kluczowe: przestrzenna analiza ekonomiczna, model potencjału, iloraz potencjału, odległość ekonomiczna, przestrzenna macierz wag

Title: Use of the spatial weighting matrix in the procedure for determining the potential quotient model

Abstract: In spatial economic studies, a potential model is used, which, when applied to selected economic phenomena, makes it possible to determine the potential of all the regions analyzed. In the regional analyses performed, the potential quotient is also used, which is the quotient of income potential and population potential. The magnitudes of income potential or population potential for selected regions are determined taking into account the values of the economic phenomenon for all regions and the physical distance between regions. An important determinant of the size of the potential of regions, in addition to physical distance, is their economic similarity, expressed by economic distance. This means that the use of economic distance in the procedure for determining

the potential quotient should allow consideration of the spatial dependencies between regions that result from their socio-economic interrelationships. Accordingly, the article will propose the use of a spatial weight matrix in the process of determining economic distance and the model of the potential quotient model.

Keywords: spatial economic analysis, potential model, potential quotient, economic distance, spatial weight matrix

Józef Pociecha, Barbara Pawełek, Uniwersytet Ekonomiczny w Krakowie

Tytuł: Uczenie maszynowe a uczenie statystyczne - wielowymiarowa analiza bibliometryczna na podstawie bazy Scopus

Streszczenie: We współczesnej analizie danych powszechnie posługujemy się terminami: uczenie maszynowe oraz uczenie statystyczne. Celem referatu jest przedstawienie definicji tych pojęć, określenie stopnia podobieństwa ich rozumienia oraz wzajemnego powiązania i rozwoju. W artykule przedstawiono wyniki wielowymiarowej analizy bibliometrycznej rozwoju liczby publikacji z zakresu uczenia maszynowego (ML) oraz uczenia statycznego (SL) wraz z siecią powiązań między występującymi pomiędzy nimi. W tym celu przeanalizowano rozkłady liczby publikacji na temat ML oraz SL według obszaru tematycznego, typu dokumentu, typu źródła, słowa kluczowego, kraju ukazania się publikacji, nazwiska autora i jego afiliacji. Zbudowano sieci współwystępowania słów kluczowych w tych publikacjach, mapy tematyczne i mapy ewolucji tematów w oparciu o słowa kluczowe, sieci współpracy między autorami, instytucjami i krajami. Przeanalizowano również ich podział na grupy w rozpatrywanych podokresach, częstość występowania autorów i ich powiązanie oraz grupowanie ze względu na kraje, powiązanie i grupowanie autorów ze względu na nazwisko. Źródłem danych do wykonanej analizy bibliometrycznej jest baza danych Scopus, dostępna w ramach Wirtualnej Biblioteki Nauki. Wyniki przeprowadzonej analizy bibliometrycznej pozwalają na sformułowanie wielu ciekawych wniosków dotyczących rozwoju i powiązania ML oraz SL, jako nowych i gwałtownie rozwijających się subdyscyplin wiedzy informatycznej i statystycznej.

Słowa kluczowe: analiza bibliometryczna, analiza danych, uczenie maszynowe, uczenie statystyczne

Title: Machine learning and statistical learning - multivariate bibliometric analysis based on the Scopus database

Abstract: In modern data analysis, we commonly use the terms: machine learning and statistical learning. The aim of the paper is to present the definition of these concepts, determination the degree of similarity of their understanding, mutual connection and development. The paper presents the results of a multidimensional bibliometric analysis of the development the number of publications in the field of machine learning (ML) and statistical learning (SL) along with the network of relationships between them. Distributions of the number of publications on ML and SL by subject area, document type, source type, keyword, country of publication, author's

name and affiliation were analyzed for this purpose. Networks co-occurrence of keywords in these publications, thematic maps and keyword-based maps of topic evolution, networks of cooperation between authors, institutions and countries were built. Their division into groups in the considered sub-periods, the frequency of occurrence of authors and their association, as well as grouping by country, association and grouping of authors by surname were also analyzed. The source of data for the bibliometric analysis is the Scopus database, available as part of the Virtual Library of Science. The results of the bibliometric analysis allow us to formulate many interesting conclusions regarding the development and connection of ML and SL as new and rapidly developing sub-disciplines of computer science and statistical knowledge.

Keywords: bibliometric analysis, data analysis, machine learning, statistical learning

Aneta Ptak-Chmielewska*, Paweł Kopciuszewski, * Szkoła Główna Handlowa
w Warszawie, ** ING Hubs Poland**

Tytuł: Scenariusze ryzyka klimatycznego w stress testach modeli PD dla portfeli detailed

Streszczenie: Brak jest ogólnodostępnych wiarygodnych i reprezentacyjnych danych historycznych zawierających informacje na temat ESG i opracowanych modeli dla ryzyka klimatycznego. W artykule analizie poddano portfele kredytów hipotecznych i niezabezpieczonych pożyczek z banku z południowej Europy. Wiarygodne szacunki PD i rzeczywisty wpływ ryzyka przejścia na portfel jest ważnym celem dla banków aby zapewnić strategię zgodną z Net Zero celem do 2050 oraz pomocną w spełnieniu kryteriów wejścia do przejścia na bardziej ekonomiczną/niskoemisyjną gospodarkę. Przyszły wpływ zarówno procesu przejścia klimatycznego and zmian klimatycznych są bardzo ważne dla szacunku kapitałów w banku. Scenariusze ryzyka przejścia koncentrują się na przejściu z paliw kopalnych, wysokoemisyjnej produkcji i konsumpcji do emisyjnie neutralnych alternatyw. Scenariusze są zdefiniowane wariantowo i zawierają czynniki ryzyka. Głównym celem artykułu jest sprawdzenie tendencji zmian w prognozach PD w relacji do scenariuszy ryzyka klimatycznego z wykorzystaniem bazy NGFS. Uwzględniono wewnętrzną bankową informację o klientach banku oraz makroekonomiczne zmienne i scenariusze przejścia. Badanie oparto o zachowanie dwóch portfeli banku w latach 2008-2022 (pierwszy portfel) oraz 2013-2023 (drugi portfel). Dla każdego klienta połączono wewnątrzbankowe dane z danymi makroekonomicznymi. Dane z bazy scenariuszy NGFS zostały wykorzystane do prognozy ryzyka portfela dla perspektywy kilkudziesięciu lat.

Słowa kluczowe: ryzyko klimatyczne, ryzyko przejścia, stress testy, NGFS

Title: Climate risk scenarios in stress testing PD models for retail portfolios

Abstract: There are no reliable and representative historical data including ESG information and developed models for the climate risk estimation. In our paper, we analyze retail mortgage and unsecured loans bank portfolios from the South of Europe. The reliable PD prediction and the actual impact of the transition risk on the

portfolio risk is a very important goal of the banks in order to ensure that the strategy is aligned with Net Zero targets by 2050 as well as helping in the admission criteria to transition to a greener/low-carbon economy. The future impact of both the climate transition processes and climate change are very important for the estimation of banks' capital. Transition risk scenarios mainly focus on transition from fossil fuels, carbon-intensive production and consumption towards emission-neutral alternatives. The scenarios are defined for different variants and include transition risk drivers. The main goal of the paper is to check the tendency in the PD's prediction in relation to climate risk scenarios using NGFS database. Ultimately, we took into account an internal information for bank customers along with macroeconomic variables and the transition scenarios. The research study is based on the behavior of two separate portfolios in years 2008-03-31 to 2022-11-30 (1st portfolio) and 2013-01-31 to 2023-05-31 (2nd portfolio). For each customer we merged customer related information with macroeconomic data. We then used the NGFS climate risk data to predict the portfolio risk over multiple years ahead.

Keywords: climate risk, transition risk, stress tests, NGFS

Marek Rocki, Szkoła Główna Handlowa w Warszawie

Tytuł: Ranking uczelni wyższych jako predyktor zatrudnialności absolwentów

Streszczenie: Corocznie ukazujące się rankingi uczelni wyższych porządkują w pewnym sensie informacje o systemie szkolnictwa wyższego, ale tworzone są z myślą o bieżących analizach, a w tym do podejmowaniu decyzji przez kandydatów na studia. Współistnienie ogólnopolskiego systemu monitorowania ekonomicznych losów absolwentów skłania do zestawienia wyników wybranego rankingu z losami absolwentów. Wykorzystano dane o wynikach rankingu w podziale na uczelnie publiczne i niepubliczne oraz dane ze wspomnianego systemu monitorowania. Celem prac było sprawdzenie, czy wybranie przez kandydatów na studia uczelni zajmujących czołowe miejsca w rankingu daje podstawy do przewidywania lepszej – niż w przypadku wyboru innej uczelni – sytuacji na rynku pracy w pięć lat po uzyskaniu dyplomu. Wskazano, że w przypadku uczelni publicznych ranking jest dobrym predyktorem w tym zakresie. W przypadku uczelni niepublicznych analizy nie wskazały na istotnie lepszą sytuację absolwentów czołówki rankingu od absolwentów pozostałych uczelni.

Słowa kluczowe: ranking, szkolnictwo wyższe, absolwenci szkół wyższych, zatrudnialność, rynek pracy

Title: University rankings as a predictor of graduate employability

Abstract: The annual rankings of universities, in a sense, organize information about the higher education system, but they are created with current analyzes in mind, including decision-making by candidates for studies. The coexistence of a nationwide system for monitoring the economic fate of graduates prompts us to compare the results of the selected ranking with the fate of graduates. The article uses data on the ranking results divided into public and private universities and data from the above-mentioned monitoring system. The aim of the work was to check whether the candidates' choice of universities with top rankings provides grounds for predicting a better

situation on the labor market than in the case of choosing another university five years after obtaining a diploma. It was indicated that in the case of public universities, the ranking is a good predictor in this respect. In the case of non-public universities, the analyzes did not indicate a significantly better situation for graduates at the top of the ranking than graduates of other universities.

Keywords: ranking, higher education, university graduates, employability, labor market

Agnieszka Rossa, Uniwersytet Łódzki

Tytuł: Wykorzystanie metody predykcji w pomiarze dzietności

Streszczenie: Zagadnienie płodności i dzietności leży obecnie w centrum uwagi demografów i polityków społecznych, w związku z obserwowanym wyraźnym spadkiem wskaźników płodności w wielu krajach rozwiniętych, w tym także w Polsce. Jest to zjawisko niepokojące, prowadzi bowiem m.in. do zakłóceń w systemie zabezpieczeń społecznych, w podaży na rynku pracy czy też w dostępie do usług zdrowotnych. Istnieją dwa zasadniczo różne podejścia, które można zastosować w analizie płodności i dzietności, tj. podejście przekrojowe i podejście wzdłużne (kohortowe). Pierwsze z nich daje spojrzenie na zachowania reprodukcyjne żyjącej w danym okresie populacji kobiet, podczas gdy drugie ujęcie daje spojrzenie na zachowania reprodukcyjne konkretnych generacji kobiet, uwzględniające ich specyfikę i przeszłą historię reprodukcyjną. Współczynnik dzietności co do zasady jest kohortową miarą płodności, mimo to jest rutynowo obliczany w oparciu o dane przekrojowe. Przekrojowy współczynnik dzietności TFR (total fertility rate) jest związany z kohortą syntetyczną, tj. cząstkowe współczynniki płodności wg wieku obserwowane w populacji kobiet w danym roku są traktowane tak, jakby charakteryzowały doświadczenie życiowe pewnej hipotetycznej generacji kobiet. Taka perspektywa jest często stosowanym sposobem analizy dzietności, ponieważ dane statystyczne zbierane są zazwyczaj w ujęciu rocznym. Z kolei pełny (kohortowy) współczynnik dzietności CFR (Complete Fertility Rate) nie może być dokładnie obliczony, dopóki dana kohorta kobiet nie zakończy swojej aktywności, co wymaga wieloletniej obserwacji. Innymi słowy, dane kohortowe związane z płodnością kohort są zwykle niepełne, tj. są znane do bieżącego roku, a to, co wydarzy się w przyszłości, pozostaje nieznane. W związku z tym zamiast tego obliczany jest przekrojowy TFR. Niestety, na podstawie TFR nie jesteśmy w stanie odróżnić efektu kalendarza płodności od efektu poziomu płodności. Innymi słowy, zmiany w TFR mogą być spowodowane zmianami w poziomie płodności (quantum effect) oraz w kalendarzu płodności (tempo effect). Na przykład, może się zdarzyć, że w pewnym roku niektóre kobiety odłożą w czasie decyzję dotyczącą posiadania lub powiększenia liczby potomstwa, co przekłada się na obniżone cząstkowe współczynniki płodności w tym roku, ale może to nie wpłynąć na pełną (kohortową) płodność kobiet. W takim przypadku TFR w danym roku zmniejszy się, ale CFR nie. Bongaarts i Feeney (1998) zaproponowali metodę korekty TFR w celu wyeliminowania efektu tempa i uzyskania miary czystego efektu quantum. Skorygowana miara TFR reprezentuje całkowity współczynnik dzietności, który zostałby zaobserwowany w danym roku, gdyby nie nastąpiło przesunięcie w kalendarzu urodzeń. Procedura korekty zaproponowana przez Bongaarts i Feeneya działa na poziomie specyficznym dla wieku, opiera się na

dostosowaniu z osobna każdego współczynnika płodności specyficznego dla wieku. Wzór Bongaarta-Feeneya odtwarza poprawnie współczynnik CFR przy dwóch założeniach, tj. (1) rozkład współczynników płodności wg wieku ma stały kształt w czasie oraz (2) przesuwa się do wyższego lub niższego wieku o ustaloną wartość, tj. wzrost lub spadek średniego wieku rodzenia dziecka obserwowany w ciągu jednego roku utrzymuje się w każdym kolejnym roku przez cały okres życia hipotetycznej kohorty. Zeng i Land (1991) wykazali, że skorygowany TFR jest generalnie odporny na naruszenie pierwszego założenia, tj. założenia stałego kształtu rozkładu płodności. Jednakże gdy naruszony jest warunek drugi, wówczas równość między skorygowanym TFR a CFR nie zachodzi. Kim i Schoen (2000) wykazali, że formuła Bongaarta-Feeneya daje niestabilne i niewłaściwe wartości, gdy rozkład płodności zmienia się cyklicznie w czasie. Tak więc, ogólnie rzecz biorąc, korekta Bongaarta-Feeneya nie zapewnia eliminacji efektu tempa. Ze względu na krytykę, z jaką spotkało się podejście Bongaarta-Feeneya, proponuję użycie pomysłu Butza i Warda (1979) do oszacowania średniego poziomu dzietności w danym roku. Idea opiera się na obliczeniu tzw. indeksu tempa (timing index), który odzwierciedla związek między tempem i quantum. Indeks ten został zdefiniowany jako odsetek płodności kohortowej wniesiony w danym roku przez wszystkie aktywne kohorty w tym roku. Ponieważ jednak dane kohortowe są zwykle niepełne, proponuję zastosowanie dodatkowo metody predykcji cząstkowych kohortowych współczynników płodności dla przyszłych okresów. Rozwiązanie to adaptuje metodę Lee i Cartera (1992) prognozowania umieralności.

Cytowana literatura:

Bongaarts, J., Feeney G. (1998), On the quantum and tempo of fertility, *Population and Development Review*, 24(2), 271-291.

Butz, W.P., Ward, M.P., (1979), Will U.S. fertility remain low? A new economic interpretation, *Population and Development Review*, 5, 663-688.

Kim, Y. J., Schoen, R. (2000), On the quantum and tempo fertility: limits to the Bongaarts-Feeney adjustment, *Population and Development Review*, 26, 554-559.

Lee, R.D., Carter, L. R., (1992), Modeling and Forecasting U. S. Mortality, *Journal of the American Statistical Association*, 87(419), 659-671.

Zeng, Y., Land, K.C., (2001), A sensitivity analysis of the Bongaarts-Feeney method of adjusting bias in observed period total fertility rates, *Demography*, 38, 17-28.

Andrzej Sokółowski*, Katarzyna Budny, Małgorzata Markowska***, * Uniwersytet
Andrzeja Frycza Modrzewskiego w Krakowie, ** Uniwersytet Ekonomiczny w
Krakowie, *** Uniwersytet Ekonomiczny we Wrocławiu**

Tytuł: Mediana wielowymiarowa – nowe podejście

Streszczenie: Przedstawiona praca jest trzecią z serii poświęconej zagadnieniu mediany wielowymiarowej. Nowatorskie podejście polega przede wszystkim na innej koncepcji mediany wielowymiarowej. Jest ona definiowana jako punkt tylko w przestrzeni jednowymiarowej. Dla zmiennej dwuwymiarowej jest prostą, dla zmiennej trójwymiarowej płaszczyzną, a dla zmiennej wielowymiarowej – hiperpłaszczyzną. Referat koncentruje

się na metodach estymacji mediany-prostej dla zmiennej dwuwymiarowej. Wykorzystywane jest pojęcie środka zbioru szacowanego poprzez rzutowanie punktów empirycznych na okrąg lub hipersferę, a następnie zastosowanie jednej z pięciu metod estymacji punktu środkowego. Zakłada się, że prosta lub hiperpłaszczyzna mediany powinna zawierać ten punkt. Do znalezienia kolejnych punktów definiujących medianę wykorzystywany jest wskaźnik agregatowy. Dla zmiennej dwuwymiarowej kolejnym krokiem jest poszukiwanie drugiego punktu definiującego prostą. Przy parzystej liczbie obserwacji przeszukuje się wszystkie punkty empiryczne. Przy nieparzystej liczbie obserwacji przeszukuje się wszystkie punkty leżące w połowie odległości pomiędzy punktami sąsiednimi. To sąsiedztwo jest specyficzne i do jego określenia zaproponowano nową miarę odległości nazwaną odległością śmigłową. Mediana wielowymiarowa powinna po pierwsze dzielić próbę na połowy równe co do liczebności, a po drugie powinna rozdzielać obserwacje, których średnie wartości wskaźnika agregatowego różnią się maksymalnie.

Słowa kluczowe: mediana wielowymiarowa, środek zbioru, estymacja

Title: Multidimensional median – new approach

Abstract: The paper is the third one in the series on the multidimensional median. The new idea of the multidimensional median is in heart of this innovative approach. The median is a point only for one-dimensional variable, while it is a straight line for two-dimensional variable, and a hyper-plane for the multidimensional one. The current presentation is concentrated on the median estimation for the two-dimensional variable. The idea of the middle of the set of points found through five estimation procedures performed on the projection onto circle or hypercube is used. We assume that the multidimensional median should include this point. Also the composite indicator is used. For two-dimensional variable we should find the second point defining the line. For the even number of objects all of them are being searched. For the odd number of observations we search the half points between neighbours. This neighbourhood is defined in a special way, by the newly proposed, point based propeller distance measure. Multidimensional median should divide the sample into equal halves, and their means of composite indicator should be maximally different.

Keywords: multidimensional median, the middle, estimation

Tomasz Stachurski, Uniwersytet Ekonomiczny w Katowicach

Tytuł: Dokładność szacowania kwantyli metodami wspomaganymi modelowo w badaniach reprezentacyjnych

Streszczenie: Jednym z kluczowych wyzwań związanych z wnioskowaniem na podstawie badań reprezentacyjnych jest uzyskanie wystarczającej dokładności otrzymanych oszacowań. Zagadnienie to jest szczególnie istotne w przypadku wnioskowania o parametrach podpopulacji, których liczebność w próbie bywa niewielka lub nawet zerowa. Z taką sytuacją badacze mają do czynienia w przypadku potrzeby wnioskowania o

parametrach na niskim poziomie agregacji, a także w przypadku analiz jednocześnie w wielu wymiarach. W przypadku badań ekonomicznych, rozkłady analizowanych zmiennych zazwyczaj cechują się asymetrią, a także występowaniem wartości oddalonych. Wówczas bardziej odpowiednią miarą położenia są kwantyle aniżeli średnia. Dlatego też w przeprowadzonym badaniu rozważano metody wspomaganie szacowania kwantyli w populacji oraz małym obszarze poprzez wykorzystanie zmiennych pomocniczych. Do analizy zastosowano badanie symulacyjne, w którym porównano własności klasycznych metod estymacji kwantyli niewykorzystujących zmiennych pomocniczych z metodami wspomaganymi modelowo. Wśród metod wspomaganych modelowo rozważano estymatory regresyjne, estymatory kalibrowane oraz metody oparte na modelu. Dokonano oceny obciążenia oraz błędu średniokwadratowego rozważanych metod przy zastosowaniu warstwowego planu losowania. Przeprowadzone badanie symulacyjne wskazuje w wielu przypadkach na przewagę metod wspomaganych modelowo, w tym także autorskich propozycji w stosunku do metod istniejących już w literaturze.

Słowa kluczowe: estymacja kwantyli, odwracanie dystrybuanty, statystyka małych obszarów, estymatory regresyjne, liniowe modele mieszane

Title: Accuracy of quantile estimation using model-assisted methods in survey sampling

Abstract: One of the key challenges in the inference based on survey data is achieving sufficient accuracy. This issue becomes particularly demanding in the case of estimation parameters for subpopulations, especially those with a small or even zero sample size. Such situations arise when inferences need to be made for parameters at a low level of aggregation or in analyses across multiple dimensions simultaneously. In economic research, the distributions of analysed characteristics are often asymmetric and contain outliers. In such cases, quantiles are a more appropriate measure of location than the mean. Therefore, in this study, methods of quantile estimation for the population and small areas using auxiliary variables were considered. A simulation study was conducted to compare the properties of classical quantile estimation methods, which do not utilize auxiliary variables, with model-assisted methods. Among the model-assisted methods, regression estimators, calibrated estimators, and model-based methods were examined. The bias and mean square error of the considered methods were evaluated under stratified sampling. The results of the simulation study indicate that, in many cases, model-assisted methods, including some original proposals, outperform methods already existing in the literature.

Keywords: quantile estimation, inversion of the distribution function, small area estimation, regression estimators, linear mixed models

Tytuł: Nieznane właściwości rozkładu Double Weibull lub bimodal power Laplace

Streszczenie: Balakrishnan i Kocherlakota (1985), wykorzystując transformację Fernandez i Steela (1998) do rozkładu Weibulla z jednostkowym parametrem skali, zdefiniowali rozkład double Weibull (DW). Autor artykułu, wykorzystując standardowy rozkład Laplace'a, definiuje rozkład bimodal power Laplace (BPL). Okazuje się, że funkcje gęstości wspomnianych rozkładów są takie same, mimo że zostały utworzone w różny sposób. Artykuł przedstawia chronologiczny przegląd dużej rodziny rozkładów Laplace'a, przypomina znane właściwości BPL i wprowadza nowe właściwości, takie jak punkty przegięcia, miara Moorsa, generator liczb pseudolosowych, entropia Shannona-Tsallisa-Renyiego i macierz informacji Fishera. Autor podaje przykłady stosowalności BPL. Główne kody R i dowody niektórych twierdzeń podano w załączniku.

Słowa kluczowe: Rozkład Laplace'a; odchodzenie od rozkładu Laplace'a; symulacje Monte Carlo

Title: Unknown properties of Double Weibull or bimodal power Laplace distribution

Abstract: Balakrishnan and Kocherlakota (1985), using the Fernandez and Steel (1998) transformation to the Weibull distribution with a unity scale parameter, introduced the double Weibull distribution (DW). The author of the article, using the standard Laplace distribution, defines the bimodal power Laplace (BPL) distribution. It turns out that the density functions of the mentioned distributions are the same, even though they were created in different ways. The article presents a chronological overview of the large family of Laplace distributions, recalls known properties of the BPL, and introduces new properties such as inflection points, Moors measure, pseudorandom number generator, Shannon-Tsallis-Renyi entropy and Fisher information matrix. Illustrative examples of the applicability of the BPL are provided. The main R codes and proofs of some theorems are given in the Appendix.

Keywords: Laplace distribution; Departure from Laplacity; Monte Carlo simulation.

Tytuł: Rozkład normalny ze składnikiem uplastyczniającym I

Streszczenie: Pierwszym (głównym) celem jest wprowadzenie nowego elastycznego rozkładu zdefiniowanego w nieskończonej domenie. Ten rozkład nazywa się rozkładem normalnym skośnym ze składnikiem uplastyczniającym. Drugim (dodatkowym) celem jest chronologiczny przegląd rozkładów należących do dużej rodziny rozkładów uplastycznionych normalnych. Przedstawiono niektóre właściwości propozycji, takie jak PDF, CDF, kwantyle, generator, momenty, skośność, kurtoza i momenty. Nieznane parametry nowego rozkładu oszacowano metodą największego prawdopodobieństwa. Przedstawiono również entropię Shannona i macierz informacji Fishera. Podano ilustrujące przykłady stosowalności i elastyczności wprowadzonego rozkładu.

Najważniejsze kody R podano w załączniku. Podano PDF-y dużej rodziny rozkładów uplastycznionych normalnych.

Słowa kluczowe: Odejście od normalności, składnik uplastyczniający

Title: Skew plasticizing component normal I distribution

Abstract: The first (main) goal is to introduce a new flexible distribution defined in an infinite domain. This distribution is named the skew plasticizing component normal distribution. The second (additional) goal is a chronological overview of distributions belonging to the large family of normal plasticizing distributions. Some properties of the proposal such as PDF, CDF, quantiles, generator, moments, skewness, kurtosis and moments of order statistics are presented. The unknown parameters of the new distribution are estimated by the maximum likelihood method. The Shannon entropy and Fisher Information Matrix are also presented. Illustrative examples of applicability and flexibility of the introduced distribution are given. The most important R codes are given in the Appendix. The PDFs of the large family of normal plasticizing distributions are given.

Keywords: Departure from normality, plasticizing component

Piotr Szczepocki*, Ewa Feder-Sempach*, Stanislav Uryasev, * Uniwersytet Łódzki,
** Stony Brook University**

Tytuł: Europejskie akcje o charakterystyce bezpiecznej przystani: zastosowanie drawdown betas dla spółek z indeksu STOXX EUROPE 600

Streszczenie: Głównym celem artykułu jest porównanie *drawdown betas*: Conditional Drawdown-at-Risk Beta (CdaR Beta) i Expected Regret of Drawdown Beta (EroD Beta) w okresach 2004-2013 oraz 2014-2023. Stosujemy metodologię Conditional Drawdown-at-Risk Beta (CdaR Beta) i Expected Regret of Drawdown Beta (EroD Beta). Te dwie miary, podobnie jak standardowa beta, odnoszą zwroty z aktywów do zwrotów z rynku, ale opierają się na koncepcji *drawdown*: spadku wartości aktywów od szczytu do następującego po nim dołka. Wyniki pokazują, że *drawdown betas* są bardziej wrażliwe na niepokój rynku podczas nieoczekiwanych wydarzeń i dzięki temu mogą wskazywać bezpieczne aktywa dla inwestorów zorientowanych na Europę.

Słowa kluczowe: drawdown beta, STOXX Europe 600, akcje bezpiecznej przystani

Title: European safe haven stocks: using drawdown beta for STOXX EUROPE 600 companies

Abstract: The main objective of this article is to compare drawdown betas, Conditional Drawdown-at-Risk Beta (CdaR Beta) and Expected Regret of Drawdown Beta (EroD Beta) for the periods 2004-2013 and 2014-2023. We

use the Conditional Drawdown-at-Risk Beta (CdaR Beta) and Expected Regret of Drawdown Beta (EroD Beta) methodology. These two measures, similarly to standard beta, relate asset returns to market returns, but are based on the concept of drawdowns: a decline in asset values from a peak to a subsequent trough. The results show that drawdown betas are more sensitive to market turmoil during unexpected events and may indicate safe haven assets for European-oriented investors.

Keywords: drawdown beta, STOXX Europe 600, safe-haven stocks

Tomasz Szkutnik, Uniwersytet Ekonomiczny w Katowicach

Tytuł: Zastosowanie metody bootstrapowej w ocenie nieobciążoności predykcji

Streszczenie: Badanie koncentruje się na ocenie nieobciążoności predykcji opartej na modelu regresji liniowej. Dokładność rozważanych predyktorów mierzona jest za pomocą testu statystycznego opartego na wariancji predykcji ex-post. Duża wartość stosunku wariancji predykcji ex-post do wariancji resztowej prowadzi do odrzucenia hipotezy o nieobciążoności predykcji. Rozkład ten oceniany jest za pomocą parametrycznej procedury bootstrap. Ponadto zaproponowano testy typu bootstrap, próbujące zweryfikować bezstronność trzech predyktorów regresji liniowej. Przyjmuje się, że rozkład reszt regresji liniowej jest rozkładem normalnym lub t-Studenta. Moc rozważanych testów statystycznych badana jest na podstawie obszernej analizy symulacyjnej.

Słowa kluczowe: bootstrap, nieobciążoności predykcji, moc testu

Title: Application of the bootstrap method in the assessment of the unbiasedness of predictions

Abstract: The analysis is focused on the unbiasedness of prediction based on a linear regression model. The accuracy of the considered predictors is measured by means of a statistical test based on the ex-post prediction variance. A significantly large value of the ratio of ex-post prediction variances and residual variance leads to the rejection of the hypothesis on unbiasedness of prediction. The distribution of the ratio is assessed by means of the parametric bootstrap procedure. Moreover, bootstrap-type tests, attempting to verify the unbiasedness of three predictors of the linear regression, are proposed. The distribution of the linear regression error term is assumed to be normal or t-Student. The power of considered statistical tests is examined based on an extensive simulation analysis.

Keywords: bootstrap, unbiasedness of prediction, power of the tests

Tytuł: Modelowanie efektów polityki fiskalnej w Polsce

Streszczenie: W artykule oszacowano wpływ szoków wydatkowych i szoków podatkowych na poziom aktywności gospodarczej w Polsce. W badaniu zastosowano słynny trzy równaniowy strukturalny model VAR, ale wzbogacony o informacje dotyczące wartości krótkookresowych elastyczności. Model oparty jest o nową metodę zaproponowaną przez Baumeister i Hamiltona w latach 2015, 2018 i 2019. Model pozwala na uwzględnienie niepewności dotyczącej krótkookresowych elastyczności oraz nałożenie restrykcji na ich znaki. Model oszacowano na danych od 1999 kwartał I do 2023 kwartał IV. Wstępne wyniki wskazują na mniejsze wartości krótkookresowych elastyczności PKB na zmienne fiskalne niż wartości opisywane w literaturze. Dodatkowo krótkookresowa elastyczność wpływów podatkowych na PKB wydaje się większa niż założono a priori. Uzyskano dodatni mnożnik wydatkowy, który przyjmuje maksymalną wartość 0,7 w siódmym kwartale. Natomiast mnożnik podatkowy jest ujemny, a jego maksymalna wartość -2,7 pojawia w okresie zerowym.

Słowa kluczowe: polityka fiskalna, strukturalne modele VAR, metody bayesowskie, funkcje reakcji na impuls

Title: On Modelling the Effects of Fiscal Policy Shocks in Poland

Abstract: I estimate the impact of government expenditure and tax shocks on economic activity levels in Poland. Using the well-known three-equation structural VAR model, the study enhances its analysis by incorporating researcher's knowledge about short-term elasticity values. The model draws on a novel approach introduced by Baumeister and Hamilton in their works from 2015, 2018, and 2019. The model allows to account for researcher's uncertainty regarding short-term elasticities and for imposing restrictions on their signs. It is estimated using data spanning from the first quarter of 1999 to the fourth quarter of 2023. Preliminary findings suggest that the short-term GDP elasticities related to fiscal variables are lower than those reported in literature. Interestingly, the short-term elasticity of tax revenues in relation to GDP appears to be higher than initially anticipated. The analysis reveals a positive spending multiplier, peaking at 0.7 in the seventh quarter, while the tax multiplier shows a negative value, reaching a maximum of -2.7 at the initial period.

Keywords: fiscal policy, structural VAR, Bayesian inference, impulse-response functions

Grażyna Trzpiot, Jacek Szoltysek, Uniwersytet Ekonomiczny w Katowicach

Tytuł: Wykorzystanie podejścia taksonomicznego do oceny atrakcyjności miast

Streszczenie: Autorzy postanowili podjąć próbę zbudowania ogólnej miary taksonomicznej oceny jakości życia w mieście, uwzględniającej cztery uwarunkowania dobrego funkcjonowania w warunkach miejskich: zdrowie i bezpieczeństwo, akceptację (inkluzywność), mobilność (możliwości docierania do miejsc i do ludzi) oraz

perspektywy materialne i duchowe życia. By uniknąć dyskusji na temat tego, co jest właściwe – uzyskiwanie opinii subiektywnych, czy też korzystanie z obiektywnie określonych i zmierzonych wielkości, nie gromadzonych na potrzeby tworzenia konkretnego rankingu, lecz w głównej mierze w związku z obowiązkami sprawozdawczymi wobec instytucji statystycznych. Autorzy postanowili dobrać szereg zmiennych, których interpretacja pozwala najpierw na wnioskowaniu w wybranych kategoriach zmiennych, a następnie, uwzględniając przekrojowy charakter uwarunkowań dobrego funkcjonowania w warunkach miejskich, przeprowadzić ostateczne wnioskowanie.

Słowa kluczowe: miary oceny jakości życia, demografia miast, środowisko

Title: Application of a taxonomic approach to assess the attractiveness of cities

Abstract: The authors decided to try to build a general taxonomic measure for assessing the quality of life in a city, taking into account four determinants of good urban living: health and safety, acceptance (inclusiveness), mobility (opportunities to reach places and people), and material and spiritual perspectives on life. In order to avoid debates about what is appropriate - obtaining subjective opinions, or using objectively defined and measured quantities, not collected for the purpose of creating a specific ranking, but mainly in connection with reporting obligations to statistical institutions. The authors decided to select a number of variables, the interpretation of which allows first to make inferences in selected categories of variables, and then, taking into account the cross-sectional nature of the determinants of good performance in urban conditions, to make a final inference.

Keywords: measures of quality of life assessment, urban demographics, environment

Alicja Wolny-Dominiak, Tomasz Żądło, Uniwersytet Ekonomiczny w Katowicach

Tytuł: Nieobserwowane realizacje zmiennej a wybór metody predykcji

Streszczenie: Wybór metody predykcji uwzględniający zarówno wybór modelu (parametrycznego lub nieparametrycznego), jak i klasy predyktora zazwyczaj obejmuje ocenę jakości dopasowania modelu do danych empirycznych i ocenę dokładności ex post prognoz. Natomiast ocena dokładności ex ante predykcji jest prowadzona zazwyczaj przy założeniu ostatecznej postaci wybranego modelu. Ponadto, jeśli do oceny dokładności ex ante predykcji różnych predyktorów przyjmujemy różne modele, uzyskane wyniki nie są porównywalne. W artykule proponujemy symulacyjną metodę wyboru predyktora opartą na ocenie ex ante dokładności predykcji przy jednoczesnym uwzględnieniu różnych modeli (różnych scenariuszy), a możliwą do zastosowania zarówno w przypadku modeli parametrycznych jak i nieparametrycznych (w tym uczenia maszynowego). Modele te mogą też uwzględniać zjawiska, które nie są obserwowane dla danych próbkowych. Proponowana metoda może być wykorzystywana w przypadku dowolnych danych (przekrojowych, przekrojowo-czasowych i szeregów czasowych) w przypadku problemu predykcji wektora dowolnych funkcji rozważanych zmiennych losowych

(dowolnych charakterystyk populacji lub podpopulacji). Oprócz rozważań teoretycznych zostanie przedstawiony przykład z wykorzystaniem rzeczywistych danych przekrojowo-czasowych.

Słowa kluczowe: wybór predyktora, modele parametryczne i nieparametryczne, dokładność predykcji

Title: Unobserved realisations of a variable and choice of the prediction method

Abstract: The choice of prediction method taking into account both the choice of model (parametric or non-parametric) and the class of predictor usually involves an assessment of the goodness-of-fit of the model and an assessment of the ex post prediction accuracy. In contrast, the estimation of ex ante prediction accuracy is usually conducted assuming the final form of the chosen model. Furthermore, if different models are adopted to estimate the ex ante prediction accuracy of different predictors, the results obtained are not comparable. In this paper, we propose a simulation-based predictor selection method based on the ex ante assessment of prediction accuracy taking into account different models (different scenarios), and which can be applied to both parametric and nonparametric models (including machine learning). These models can also take into account phenomena that are not observed in sample data. The proposed method can be used with arbitrary data (cross-sectional, longitudinal and time series) for the problem of predicting a vector of any functions of the considered random variables (arbitrary population or sub-population characteristics). In addition to the theoretical considerations, an example using real longitudinal data will be presented.

Keywords: selection of a predictor, parametric and non-parametric models, prediction accuracy

Artur Zaborski, Marcin Pełka, Uniwersytet Ekonomiczny we Wrocławiu

Tytuł: Analiza PROFIT obiektów symbolicznych na przykładzie oceny reklam samochodów

Streszczenie: PROFIT jest przykładem „zewnątrznej” wektorowej metody map preferencji. Jest ona połączeniem skalowania wielowymiarowego i analizy regresji wielorakiej. Danymi wejściowymi w analizie PROFIT są zarówno współrzędne punktów reprezentujących obiekty na mapie percepcyjnej jak również oceny preferencji obiektów ze względu na wybrane zmienne. Dla konfiguracji punktów reprezentujących obiekty otrzymanej za pomocą skalowania wielowymiarowego przeprowadza się analizę regresji wielorakiej, w której zmiennymi objaśniającymi są współrzędne obiektów na mapie percepcyjnej, a zmiennymi zależnymi oceny marek ze względu na poszczególne cechy. Program dokonuje rozmieszczenia na mapie percepcyjnej zmiennych w postaci wektorów wskazujących kierunek maksymalnej preferencji ze względu na daną zmienną. Celem artykułu jest przedstawienie propozycji przeprowadzenia analizy PROFIT dla obiektów symbolicznych. W części empirycznej wykorzystano metodę środków ciężkości dla obiektów symbolicznych do oceny wybranych reklam samochodów.

Słowa kluczowe: PROFIT, skalowanie wielowymiarowe, analiza regresji, analiza danych symbolicznych, mapy preferencji

Title: The PROFIT analysis for symbolic objects on the example of the car advertisement evaluation

Abstract: PROFIT is a kind of external vector analysis of preference mapping. It is a combination of multidimensional scaling and multiple regression analysis. PROFIT takes as input both a configuration of stimulus points and a set of preference rankings of the different properties of the stimuli. For stimulus space obtained by multidimensional scaling multiple regression is performed using the coordinates as independent variables and attribute as the dependent variable. The program locates each property as a vector through the configuration of points, so that it indicates the direction over the space in which the property is increasing. The aim of the article is to present suggestions how to perform PROFIT analysis for symbolic objects. In the empirical part, the center method for symbolic objects was used to evaluate selected car advertisements.

Keywords: PROFIT, multidimensional scaling, regression analysis, symbolic data analysis, preference maps