

Book of Abstracts

Statistische Woche

02. – 05. September 2025 | Wiesbaden



BOOK OF ABSTRACTS

STATISTISCHE WOCHE

02. – 05. September 2025

Wiesbaden

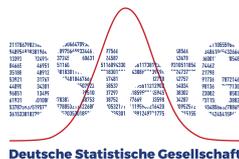


Organisers / Organisierende:

Federal Statistical Office of Germany
Statistisches Bundesamt



German Statistical Society – registered association (DStatG)
Deutsche Statistische Gesellschaft e.V. (DStatG)



Association of German Urban Statistics (VDSt)
Verband Deutscher Städtestatistik (VDSt)



German Association for Demography – registered association (DGD)
Deutsche Gesellschaft für Demographie e. V. (DGD)



STATISTISCHE WOCHE

02. – 05. September 2025

Wiesbaden

www.statistische-woche.de

Contents

Programme – Overview	p. 7
Programme	p. 9
• Session Timetable	p. 9
• Tuesday	p. 12
• Wednesday	p. 17
• Thursday	p. 22
• Friday	p. 32
Abstracts	p. 35
Social Programme	p. 152
Participants	p. 154
Building Plans	p. 163

Inhaltsverzeichnis

Programm – Übersicht	S. 7
Programm	S. 9
• Sitzungszeitplan	S. 9
• Dienstag	S. 12
• Mittwoch	S. 17
• Donnerstag	S. 22
• Freitag	S. 32
Abstracts	S. 35
Rahmenprogramm	S. 152
Teilnehmende	S. 154
Gebäudepläne	S. 163

Partners / Partner:

- KonsortSWD / RatSWD
- SAS
- Springer

Programme Committee / Programmkomitee:**German Statistical Society – registered association /
Deutsche Statistische Gesellschaft e. V.:**

Yarema Okhrin (Chair / Vorsitz DStatG, Universität Trier)
Sara Bleninger (Bayerisches Landesamt für Statistik)
Hanna Brenzel (Statistisches Bundesamt)
Christine Buchholz (Hochschule Bonn-Rhein-Sieg)
Matei Demetrescu (Technische Universität Dortmund)
Christoph Hanck (Universität Duisburg-Essen)
Solveigh Jäger (Bundesverband der Deutschen Industrie)
Carsten Jentsch (Technische Universität Dortmund)
Göran Kauermann (LMU München)
Hans Kiesel (OTH Regensburg)
Philipp Otto (University of Glasgow)
Ulrike Rockmann (Universität Oldenburg)
Gaby Schneider (Goethe-Universität Frankfurt)
Katharina Schüller (STAT-UP Statistical Consulting & Data Science GmbH)
Markus Zwick (Statistisches Bundesamt)

Association of German Urban Statistics / Verband Deutscher Städtestatistik:

Uta Thien-Seitz (Chair / Vorsitz, Direktorium - Statistisches Amt München)
Diana Andrä (Dortmunder Statistik)
Jan Dohnke (Statistik und Stadtforschung Darmstadt)
Andrea Schultz (Stadt Leipzig, Amt für Statistik und Wahlen)
Ansgar Schmitz-Veltin (Landeshauptstadt Stuttgart, Statistisches Amt)

**German Society for Demography – registered association / Deutsche Gesellschaft für
Demographie e. V.:**

Norbert Schneider (Chair / Vorsitz, Bundesinstitut für Bevölkerungsforschung)
Annelene Wengler (Robert Koch-Institut)

Conference Organiser / Tagungsbeauftragter:

Arne Johannssen (Conference Commissioner / Tagungsbeauftragter DStatG, Hochschule Harz)

Local Organisers / Lokale Organisierende:

Markus Zwick (Chair / Vorsitz, Statistisches Bundesamt)
Jannik Mühlhan (Statistisches Bundesamt)
Marco Schwickerath (Statistisches Bundesamt)
Jennifer Vallé (Statistisches Bundesamt)
Lisa Carius-Munz (Statistisches Bundesamt)
Mike Maisack (Statistisches Bundesamt)
Cornelia Koch (Statistisches Bundesamt)

Programme Overview / Programmübersicht:

Tuesday, 02. September, 2025:

08:00-09:00	Registration / Welcome Coffee
09:00-13:00	Tutorial: Tidyverse models: A modern approach to Data Science and Machine learning using R
09:00-10:40	Sessions
10:40-11:00	Coffee Break
11:00-12:40	Sessions
12:40-14:20	Lunch Break / Venue Change to Kurhaus Wiesbaden
14:20-15:10	Opening
15:10-16:00	Plenary Session (Mehdi Dagdoug): <i>Estimation of Uncertainty</i>
16:00-16:20	Coffee Break
16:20-17:10	Heinz-Grohmann Lecture (Göran Kauermann): <i>Mietspiegel in Deutschland - Ein Spannungsfeld zwischen Politik, Immobilienwirtschaft und Wissenschaft</i>
17:10-18:00	Plenary Session (Christian Brownlees): <i>Which Factors Drive Downside Risk in the U.S. Economy?</i>

Wednesday, 03. September, 2025:

09:00-10:40	Sessions
10:40-11:00	Coffee Break
11:00-12:40	Sessions
12:40-14:00	Lunch Break / Venue Change to Kurhaus Wiesbaden
13:40-14:00	Coffee at Kurhaus Wiesbaden
14:00-14:50	Plenary Session (Michael Hüther): <i>Mikro- und Makroökonomische Verteilungsfragen zu Einkommen, Vermögen und Konsum</i>
14:50-15:40	Gumbel Lecture (Yannick Hoga): <i>Forecast Comparisons via Multi-Objective Elicitability</i>

Thursday, 04. September, 2025:

09:00-10:40	Sessions
10:40-11:00	Coffee Break
11:00-12:40	Sessions
12:40-14:20	Lunch Break
14:20-16:00	Sessions
16:00-16:20	Coffee Break
16:20-18:00	Sessions
18:00-20:00	Poster Session
18:30-20:00	General Assembly Meeting DStatG

Friday, 05. September, 2025:

09:00-10:40	Sessions
10:40-11:00	Coffee Break
11:00-12:40	Sessions
13:00-13:30	Closing

Session Timetable / Sitzungszeitplan:

	09:00-10:40	11:00-12:40	09:00-10:40	11:00-12:40	09:00-10:40	11:00-12:40	14:20-16:00	16:20-18:00	09:00-10:40	11:00-12:40
Section	Tue 02/9		Wed 03/9		Thu 04/9			Fri 05/9		
Main Topic: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption									A.13.207	A.13.207
Main Topic: Uncertainty Estimation	F.04.208	F.04.208								
Main Topic: High-dimensional Time Series and Network Data							B.03.104			
Young-Academics Mini-Symposium: Modern Time Series Econometrics				A.03.212						
Computational Statistics and Data Science		A.00.Bib	A.00.Bib	A.00.Bib	A.00.Bib					
Empirical Economics and Applied Econometrics			A.03.212			A.00.Bib	A.00.Bib			
Statistics in Finance					B.03.104	B.03.104				
Competence Development: Data Literacy and Statistics	E.03.112	E.03.112								
Methodology of Statistical Surveys			E.03.112 A.03.205	E.03.112	E.03.112	E.03.112	E.03.112	E.03.112		
Regional Statistics						F.04.208			E.03.112	E.03.112
Statistics in Environmental and Engineering Sciences								F.04.208	F.04.208	F.04.208
Statistical Theory and Methods					A.03.212	A.03.212	A.03.212	A.03.212		
Nonparametric and Robust Statistics	A.03.205	A.03.205								
Economic, Social and Market Statistics	A.13.207	A.13.207	A.13.207	A.13.207 A.03.205	A.13.207	A.13.207	A.13.207	A.13.207		
(Systemic) Risk Modeling								B.03.104		
Joint DStatG and TIES Session								A.00.Bib		
DGD					A.03.205	A.03.205	A.03.205			
VDSt			F.04.208	F.04.208	F.04.208	F.04.208	F.04.208			

Plenary Sessions & Lectures**Tuesday, 02. September 2025**

Mehdi Dagdoug. „Estimation of Uncertainty“.

Tuesday, 15:10-16:00 (page 35).

Göran Kauermann. „Mietspiegel in Deutschland - Ein Spannungsfeld zwischen Politik, Immobilienwirtschaft und Wissenschaft“.

Tuesday, 16:20-17:10 (page 36).

Christian Brownlees; Carlo Pavanello; Andre Souza. „Which Factors Drive Downside Risk in the U.S. Economy?“

Tuesday, 17:10-18:00 (page 35).

Wednesday, 03. September 2025

Michael Hüther. „Mikro- und Makroökonomische Verteilungsfragen zu Einkommen, Vermögen und Konsum“.

Wednesday, 14:00-14:50 (page 36).

Tobias Fissler; **Yannick Hoga.** „Forecast Comparisons via Multi-Objective Elicitability“.

Wednesday, 14:50-15:40 (page 37).

Invited Speakers

Tuesday, 02. September 2025

Klaus Nordhausen; Christophe Croux; Mika Sipilä; Sara Taskinen. „Robust Long Short-Term Memory Models for Outlier-Contaminated Time Series“.

Tuesday, 09:00-09:50 (page 39).

Sascha Verhoeven. „Wissen. Testen. Verändern. Data & AI Literacy als Hebel für Innovation in der Verwaltung“.

Tuesday, 11:00-11:50 (page 51).

Wednesday, 03. September 2025

Sander Scholtus. „Some recent work on automatic editing of business survey data at Statistics Netherlands“.

Wednesday, 09:50-10:40 (page 60).

Peihua Qiu. „Online Monitoring of Spatio-Temporal Data Streams“.

Wednesday, 11:00-11:50 (page 69).

Thursday, 04. September 2025

Jörg Stoye. „Robust Statistical Decisions: With Application to Development Economics“.

Thursday, 11:00-11:50 (page 99).

Wanbo Lu; **Kris Boudt**; Guanglin Huang. „Estimation of Non-Gaussian Factors Using Higher-order Multi-cumulants in Weak Factor Models“.

Thursday, 11:25-12:15 (page 101).

Jean-Pierre Florens; **Anna Simoni**. „Panel data models with randomly generated groups“.

Thursday, 14:20-15:10 (page 111).

Alberto Quaini. „Nuclear Hedging Portfolios: Shrinking Complexity for Robust Diversification“.

Thursday, 17:35-18:00 (page 126).

Jethro Browell. „Forecasting the wind: From necessity to added value“.

Thursday, 17:10-18:00 (page 128).

Tommaso Proietti; Alessandro Giovannelli. „On the Estimation of Climate Normals and Anomalies“.

Thursday, 16:20-17:10 (page 128).

Tuesday, 02. September 2025, 08:00 - 09:00

Welcome Coffee

Tuesday, 02. September 2025, 09:00 - 10:40

Uncertainty Estimation (Session 1)

Chair: Timo Schmid, Room: F.04.208

Florian Schütze; Sami Diaf. „Uncovering uncertainty in narrative economics: a semantic search approach“.

Tuesday, 09:00-09:25 (page 37).

Tamar Abesadze; Christian Müller. „Methodological Issues with Measuring Economic Uncertainty“.

Tuesday, 09:25-09:50 (page 38).

Niklas Benner; Boris Blagov; Maximilian W. Dirks. „Measuring political instability at a high frequency – a text mining approach“.

Tuesday, 09:50-10:15 (page 38).

Pascal Goemans. „EU-MD, macroeconomic uncertainty and a state-dependent Phillips curve“.

Tuesday, 10:15-10:40 (page 39).

Nonparametric and Robust Statistics (Session 1)

Chair: Gabriel Frahm, Room: A.03.205

Klaus Nordhausen; Christophe Croux; Mika Sipilä; Sara Taskinen. „Robust Long Short-Term Memory Models for Outlier-Contaminated Time Series“.

Tuesday, 09:00-09:50 (page 39).

Angelika Silbernagel; Christian H. Weiß; Alexander Schnurr. „Non-parametric tests for cross-dependence based on multivariate extensions of ordinal patterns“.

Tuesday, 09:50-10:15 (page 40).

Competence Development: Data Literacy and Statistics (Session 1)

Chair: Christine Buchholz, Katharina Schüller, Room: E.03.112

Christine Buchholz. „Einstellungen zur Statistik bei Studierenden“.

Tuesday, 09:00-09:25 (page 40).

Alexander Neumann; Ramona Voshage. „Digitale Didaktik in der Hochschullehre: Technologische Entwicklungen und ihr Nutzen für datengetriebene Berufsfelder am Beispiel der Mathematik“.

Tuesday, 09:25-09:50 (page 41).

Andreas Köhler; Isabelle Kunze; **Frederik von Reumont**; **Christoph Alfken**; Alexandra Budke; Maren Franzen. „Data Literacy für den Geographieunterricht: Mit amtlicher Statistik kartenbasiert argumentieren“.

Tuesday, 09:50-10:15 (page 42).

Gaby Schneider. „Learning to apply statistics – the statistical praticum in Frankfurt“.

Tuesday, 10:15-10:40 (page 43).

Economic, Social and Market Statistics (Session 1)

Chair: Solveigh Jäger, Room: A.13.207

Andreas Kladroba; **Annette Lotz**. „Die Auswirkungen der EU-Einheitenverordnung auf die FuE-Erhebung (Teil 2)“.

Tuesday, 09:00-09:25 (page 43).

Markus Zimmermann. „Produktivität von Unternehmen. Ergebnisse der Unternehmensstrukturstatistik“.

Tuesday, 09:25-09:50 (page 44).

Johanna Jung. „Umstellung des statistischen Unternehmensregisters auf die WZ 2025“.

Tuesday, 09:50-10:15 (page 44).

Anke Rink. „Gründungen im Handwerk“.

Tuesday, 10:15-10:40 (page 45).

Coffee Break: 10:40 - 11:00

Tuesday, 02. September 2025, 11:00 - 12:40

Computational Statistics and Data Science (Session 1)

Chair: Li Chen, Room: A.00.Bibliothek

Sigbert Klinke; Kleio Chrysopoulou Tseva. „'argList': Lazy Argument Management and Propagation Toolkit“.

Tuesday, 11:00-11:25 (page 45).

Florian Schütze; Sami Diaf. „Multilingual Monetary Policy: Unfolding Language and Policy Preferences of Swiss Central Bankers“.

Tuesday, 11:25-11:50 (page 46).

Jan Anders; Göran Kauermann. „Heterogeneity in Voter Movements in Germany – A Mixture Model Approach“.

Tuesday, 11:50-12:15 (page 46).

Karolina Gliszczynska. „Estimating Heterogeneous Causal Effects with Tree-Based Methods under Imperfect Compliance and Overlap Violations“.

Tuesday, 12:15-12:40 (page 47).

Uncertainty Estimation (Session 2)

Chair: Timo Schmid, Room: F.04.208

Oliver Zeigermann; Gunnar Moys. „Schätzung von Unsicherheit in Gen AI Systemen mit einem LLM-as-a-Judge Ansatz“.

Tuesday, 11:00-11:25 (page 48).

Daniel C. Schneider; Mikko Myrskylä; Christian Dudel; Angelo Lorenti. „Statistical Inference for Discrete-Time Multistate Models: Asymptotic Covariance Matrices, Partial Age Ranges, and Group Contrasts“.

Tuesday, 11:25-11:50 (page 48).

Daniel C. Schneider; Mikko Myrskylä; Christian Dudel; Angelo Lorenti. „Statistical Inference for Discrete-Time Multistate Models: Extensions to Markov Chains with Rewards“.

Tuesday, 11:50-12:15 (page 49).

Bernhard Stefan Zins. „Inference with Non-Probability Samples: A Mixture-Modell Approach“.

Tuesday, 12:15-12:40 (page 50).

Nonparametric and Robust Statistics (Session 2)

Chair: Gabriel Frahm, Room: A.03.205

Johannes Resin; Daniel Wolfram; Johannes Bracher; Timo Dimitriadis. „Shift-Dispersion Decompositions of Wasserstein and Cramér Distances“.

Tuesday, 11:00-11:25 (page 50).

Joachim Schnurbus; Harry Haupt; Willi Semmler. „Convergence clubs in the European Union“.

Tuesday, 11:25-11:50 (page 51).

Competence Development: Data Literacy and Statistics (Session 2)

Chair: Christine Buchholz, Katharina Schüller, Room: E.03.112

Sascha Verhoeven. „Wissen. Testen. Verändern. Data & AI Literacy als Hebel für Innovation in der Verwaltung“.

Tuesday, 11:00-11:50 (page 51).

Sebastian Waßenberg; Jochen Kirstein; Anna-Lena Dr. Wagner; Leontine Dr. von Kulmiz. „SchuldenMonitor NRW: Struktur und Entwicklung kommunaler Schulden“.

Tuesday, 11:50-12:15 (page 52).

Victor Tuekam; Sebastian Wichert; Oliver Falck; Göran Kauermann. „From Mobile Phone Data to Mobility Patterns“.

Tuesday, 12:15-12:40 (page 53).

Economic, Social and Market Statistics (Session 2)

Chair: Andreas Kladroba, Room: A.13.207

Eva Gruner; Carsten Schumann. „Wie reagieren Betriebe des Bauhauptgewerbes auf die Krise im Wohnungsbau?“*Tuesday, 11:00-11:25 (page 53).***Bayram Oruc**. „Neue Konjunkturperspektiven durch Firm-Level Mikrodaten der Industrie: Innovationspotenziale für die volkswirtschaftliche Analyse“.*Tuesday, 11:25-11:50 (page 54).***Edesa Yadegar**. „t+20 - Ein Projekt zur Schnellschätzung von Konjunkturindikatoren“.*Tuesday, 11:50-12:15 (page 54).***Alexander Weigert**; Carsten Schumann; Marianne Schepers. „Handlungsbedarf in der Bautätigkeitsstatistik“.*Tuesday, 12:15-12:40 (page 55).*

Lunch Break / Venue Change to Kurhaus Wiesbaden: 12:40 - 14:20

Opening: 14:20 - 15:10Room: Kurhaus

Plenary Session 1

Chair: Markus Zwick, Room: Kurhaus

Mehdi Dagdoug. „Estimation of Uncertainty“.*Tuesday, 15:10-16:00 (page 35).*

Heinz-Grohmann Lecture

Chair: Ulrich Rendtel, Daniel Vorgrimler, Room: Kurhaus

Göran Kauermann. „Mietspiegel in Deutschland - Ein Spannungsfeld zwischen Politik, Immobilienwirtschaft und Wissenschaft“.*Tuesday, 16:20-17:10 (page 36).*

Plenary Session 2

Chair: Roxana Halbleib, Room: Kurhaus

Christian Brownlees; Carlo Pavanello; Andre Souza. „Which Factors Drive Downside Risk in the U.S. Economy?“*Tuesday, 17:10-18:00 (page 35).*

Other events on Tuesday

- 09:00 : Tutorial: Tidyverse|models}: A modern approach to Data Science and Machine learning using R, Room: B.03.104
- 19:00 : Reception by the City of Wiesbaden at the Town Hall accompanied by the Destatis Choir, see page 152
- 20:30 : Relaxed Get-Together – Rathauskeller Andechser, see page 152

Wednesday, 03. September 2025, 09:00 - 10:40

Computational Statistics and Data Science (Session 2)

Chair: Jan-Ole Koslik, Room: A.00.Bibliothek

Jens Klenke; Lennard Maßmann. „Shrinkage Bayesian Causal Forest with Instrumental Variable“. *Wednesday, 09:00-09:25 (page 55)*.

Lennard Maßmann; Eva-Maria Oeß. „Bayesian Causal Forests for Cost-Effectiveness Analysis“. *Wednesday, 09:25-09:50 (page 56)*.

Jan-Ole Koslik. „Efficient nonparametric estimation of Markov-switching models“. *Wednesday, 09:50-10:15 (page 56)*.

Li Chen; Yuanhua Feng. „Time series forecasting in SAP using a data-driven semiparametric ARMA model“. *Wednesday, 10:15-10:40 (page 57)*.

Empirical Economics and Applied Econometrics (Session 1)

Chair: Carsten Jentsch, Room: A.03.212

Florian Schütze; André Esslinger. „Explaining the difference between the excess yield on short-term German government bonds and the ECB’s deposit facility“. *Wednesday, 09:00-09:25 (page 57)*.

Julika Herzberg; Thomas Knetsch; Dilyana Popova; **Jannik Schaller**; Patrick Schwind; Sebastian Weinand. „Empirical findings on upper-level aggregation issues in the HICP“. *(page 58)*.

Linus Nüsing; Benedikt Schwab. „Order-invariant Identification in a non-linear Structural Vector Autoregression“. *Wednesday, 09:25-09:50 (page 58)*.

Methodology of Statistical Surveys: Bundesbank-Session - Measuring International Economics

Chair: Robert Kirchner, Room: A.03.205

Walter J. Radermacher. „Balance of Payments – Part of the System of Official Statistics“. *(page 61)*.

Stefan Hauf. „Balance of payments statistics – a building block for national accounts“. *(page 61)*.

Gerhard Ziebarth. „The Role of the Balance of Payments for Central Bank Policy“. *(page 62)*.

Martin Schmitz. „The Euro Area Balance of Payments and International Investment Position and their Role in ECB Policy Making“. *(page 62)*.

Matthias Ludwig. „Evidence-Based Policymaking in the European Union Using Balance of Payments Statistics“. *(page 63)*.

Ralf Münnich. „Evidence-Based Policies and Data Quality – What is Missing?“ (page 63).

Methodology of Statistical Surveys (Session 1)

Chair: Florian Dumpert, Room: E.03.112

Arijana Amina Ramic; Thorsten Konietzko; Shanna Appelhanz. „Modernisierung der PL-Prozesse in der amtlichen Statistik in Deutschland“.

Wednesday, 09:00-09:25 (page 59).

Katja-Verena Bürk; Florian Dumpert. „Plausibilisierung in der amtlichen Statistik: Vergangenheit, Gegenwart und eine mögliche Zukunft“.

Wednesday, 09:25-09:50 (page 60).

Sander Scholtus. „Some recent work on automatic editing of business survey data at Statistics Netherlands“.

Wednesday, 09:50-10:40 (page 60).

Vdst (Session 1): Umwelt-, Klima- und Gesundheitsmonitoring in Städten und Regionen

Chair: Cornelia Müller, Room: F.04.208

Daniela Schüller. „Ökonomischer Status und Pandemie-Effekte. Eine Analyse der Daten der Bürgerbefragung "Leben in Koblenz" 2023“.

Wednesday, 09:00-09:25 (page 64).

Judith Kaschowitz; **Dorothee Winkler.** „Sozialräumliche Unterschiede und individuelles Gesundheitsverhalten in deutschen Großstädten. Eine Analyse auf Basis der Innerstädtischen Raumbearbeitung und des Sozio-oekonomischen Panels“.

Wednesday, 09:25-09:50 (page 64).

Stefan Schiele; Rosanna Purrucker. „Pflegestrukturplanung in Bayern: Eine einheitliche Datengrundlage für die Sozialplanung“.

Wednesday, 09:50-10:15 (page 65).

Georg Wiegleb. „Hitzebedingte Mortalität: Möglichkeiten und Grenzen statistischer Modellierung im kommunalen Kontext“.

Wednesday, 10:15-10:40 (page 66).

Economic, Social and Market Statistics (Session 3)

Chair: Solveigh Jäger, Room: A.13.207

Simon Rommelspacher; Ralf Droßard. „Zielbild für ein neues System der Unternehmensstatistiken: SysdU“.

Wednesday, 09:00-09:25 (page 66).

Wolf-Dietmar Speich. „Berichtspflichten zu amtlichen Wirtschaftsstatistiken aus Sicht eines Statistischen Landesamtes“.

Wednesday, 09:25-09:50 (page 67).

Franziska Große; Patrizia Mödinger. „20 Jahre Verwaltungsdatenverwendung – Entwicklungsstand und Perspektiven“.

Wednesday, 09:50-10:15 (page 68).

Thorsten Gores; Hendrik Kruse. „Identifikation von VAT tradern in der Außenhandelsstatistik“.

Wednesday, 10:15-10:40 (page 68).

Coffee Break: 10:40 - 11:00

Wednesday, 03. September 2025, 11:00 - 12:40

Computational Statistics and Data Science (Session 3)

Chair: Philipp Otto, Arne Johannssen, Room: A.00.Bibliothek

Peihua Qiu. „Online Monitoring of Spatio-Temporal Data Streams“.

Wednesday, 11:00-11:50 (page 69).

Rainer Dyckerhoff; Stanislav Nagy. „Exact computation of angular halfspace depth“.

Wednesday, 11:50-12:15 (page 69).

Matei Demetrescu; **Fabian Schmidt.** „Fast Factor Extraction for Mixed Data Types“.

Wednesday, 12:15-12:40 (page 70).

Methodology of Statistical Surveys (Session 2)

Chair: Florian Dumpert, Room: E.03.112

Florian Dumpert; Markus Pauly; **Maria Thurow;** Inken Veips. „Imputing Missing Values in Official Statistics: Assessing Imputation Accuracy across Various Imputation Methods.“

Wednesday, 11:00-11:25 (page 70).

Steffen Moritz; Florian Dumpert. „Von der Praxis zur Methodik: Evaluierung von Imputationen in der amtlichen Statistik“.

Wednesday, 11:25-11:50 (page 71).

Elena Stäger; Muhammet Akman; Richard Bündgens; **Christian Borgs.** „Möglichkeiten für ML-basierte Imputationsverfahren für Item-Nonresponse am Beispiel der Statistiken des Verarbeitenden Gewerbes (StatVG)“.

Wednesday, 11:50-12:15 (page 71).

Christian Aßmann; **Ariane Würbach;** Katja-Verena Bürk; Florian Dumpert. „Handling constraints in automated statistical data editing via full conditional distributions“.

Wednesday, 12:15-12:40 (page 72).

VDSSt (Session 2): Demokratie und Wahlen aus statistischer Sicht

Chair: Andrea Schultz, Room: F.04.208

Michael Haußmann. „Die hybride MigrantInnenbeiratswahl in Freiburg“.*Wednesday, 11:00-11:25 (page 73).***Christopher Pollak; Philipp Greunke.** „Wie entwickelt sich demokratische Kultur im urbanen Raum? Ergebnisse des Leipziger Demokratie-Monitors 2024“.*Wednesday, 11:25-11:50 (page 73).***Volker Holzendorf.** „Wahlhochrechnung statt reiner Ergebnispräsentation“.*Wednesday, 11:50-12:15 (page 74).*

Economic, Social and Market Statistics (Session 4)

Chair: Solveigh Jäger, Room: A.13.207

Tanja Lenz. „Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (I)“.*Wednesday, 11:00-11:25 (page 74).***Angelika Becker.** „Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (II)“.*Wednesday, 11:25-11:50 (page 75).***René Rimpler.** „Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (III)“.*Wednesday, 11:50-12:15 (page 76).***Dr. Florian Bayer.** „Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (IV)“.*Wednesday, 12:15-12:40 (page 76).*

Economic, Social and Market Statistics (Session 5): Labour Markets and Social Security I

Chair: Bernd Hofmann, Room: A.03.205

Anton Klaus. „Arbeitsmarktintegration geflüchteter Menschen - Bewertung des Job-Turbos aus Sicht der Arbeitsmarktstatistik“.*Wednesday, 11:00-11:25 (page 77).***Ehsan Vallizadeh.** „Integrationsverläufe von Geflüchteten: Eine Kohortenanalyse“.*Wednesday, 11:25-11:50 (page 77).***Malte Reimann-Peters.** „Fachkräfteeinwanderung in der Beschäftigungsstatistik“.*Wednesday, 11:50-12:15 (page 78).***Sarah Kuhn; Barbara Schwengler; Holger Seibert; Doris Wiethölter.** „Demografischer Wandel - Ausländische Beschäftigte ersetzen in vielen Bereichen fehlendes einheimisches Personal“.*Wednesday, 12:15-12:40 (page 78).*

Young-Academics Mini-Symposium: Modern Time Series Econometrics

Chair: Karsten Reichold, Room: A.03.212

Rainer Alexander Schüssler. „Learning Signal-to-Noise Ratios from Forecast Errors: A Simulation-Based Calibration Framework“.*Wednesday, 11:00-11:25 (page 79).***Sven Pappert.** „Asymptotic Properties of MAGMAR-Copula Time Series Models“.*Wednesday, 11:25-11:50 (page 79).***Sven Otto;** Luis Winter. „Functional Factor Regression with an Application to Electricity Price Curve Modeling“.*Wednesday, 11:50-12:15 (page 80).***Karsten Reichold;** Carsten Jentsch; Christoph Hanck. „Bootstrap Inference in Panels of Cointegrating Regressions with Global Stochastic Trends“.*Wednesday, 12:15-12:40 (page 80).*

Lunch Break / Venue Change to Kurhaus Wiesbaden: 12:40 - 14:00

Plenary Session 3

Chair: Katharina Schüller, Markus Zwick, Room: Kurhaus

Michael Hüther. „Mikro- und Makroökonomische Verteilungsfragen zu Einkommen, Vermögen und Konsum“.*Wednesday, 14:00-14:50 (page 36).*

Gumbel Lecture

Chair: Christoph Hanck, Room: Kurhaus

Tobias Fissler; **Yannick Hoga.** „Forecast Comparisons via Multi-Objective Elicitability“.*Wednesday, 14:50-15:40 (page 37).*

Other events on Wednesday

18:00 : Conference Dinner, see page 152

Thursday, 04. September 2025, 09:00 - 10:40

Computational Statistics and Data Science (Session 4)

Chair: Jan Weymeirsch, Room: A.00.Bibliothek

Sara Schiesberg; Heiko Limberg. „Transformer-Modelle in der Außenhandelsstatistik: Potenziale und Herausforderungen im Vergleich mit klassischen Vektorraummodellen“.

Thursday, 09:00-09:25 (page 81).

Ulrich Reincke; Frank Gramüller. „Model Governance und Compliance unter den speziellen Arbeitsbedingungen im öffentlichen Sektor.“

Thursday, 09:25-09:50 (page 81).

Lorenz Ade. „Datenschutzkonforme Methodik für den Zugang zu Mobilfunksignalen“.

Thursday, 09:50-10:15 (page 82).

Jan Weymeirsch; Ralf Münnich. „Erstellung eines flächendeckenden Wohnraumregisters für dynamische Mikrosimulationen in Deutschland“.

Thursday, 10:15-10:40 (page 83).

DGD (Session 1): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen I

Chair: Philipp Deschermeier, Room: A.03.205

Christian Schneemann; **Johanna Zenk; Gerd Zika.** „Aktuelle Ergebnisse der QuBe-Bevölkerungsprojektion für Kreise und kreisfreie Städte“.

Thursday, 09:00-09:25 (page 83).

Ines Thobe; Linus Ronsiek; Anja Sonnenburg. „Projektion der Nachfrage nach Fachkräften in Gesundheitsberufen“.

Thursday, 09:25-09:50 (page 84).

Patrizio Vanella; Christina Benita Wilke; Moritz Heß. „Long-Term Care in Germany in the Context of the Demographic Transition—An Outlook for the Expenses of Long-Term Care Insurance through 2050“.

Thursday, 09:50-10:15 (page 85).

Statistics in Finance (Session 1)

Chair: Roxana Halbleib, Room: B.03.104

Lukas Bauer. „Forecast combination with an application to financial tail risk“.

Thursday, 09:00-09:25 (page 86).

Kerstin Hötte; Anastasia Mantziou; Gesine Reinert; Mihai Cucuringu. „GDP nowcasting with large-scale inter-industry payment data in real time—A network approach“.

Thursday, 09:25-09:50 (page 86).

Yarema Okhrin; Maziar Sahamkhadam; **Andreas Stephan.** „Forecasting Bond Returns With a Copula-Based Dynamic Factor Pricing Model“.

Thursday, 09:50-10:15 (page 87).

Jing Zou; Ostap Okhrin. „Stagewise crop yield prediction with multisource functional indices“.
Thursday, 10:15-10:40 (page 87).

Methodology of Statistical Surveys (Session 3)

Chair: Hanna Brenzel, Room: E.03.112

Stefanie Hoffmann; Tobias Laick. „Von Momentaufnahmen zu Studienverläufen: Methodische Innovationen in der amtlichen Hochschulstatistik“.
Thursday, 09:00-09:25 (page 88).

Katharina Giar; Franziska Hohlstein. „Aufbau eines Bildungsverlaufsregisters in Deutschland“.
Thursday, 09:25-09:50 (page 88).

Maximilian Bach; Jonathan Kohl. „Weiterentwicklung und Evaluierung der Erfassung non-formaler Weiterbildungsaktivitäten im Mikrozensus“.
Thursday, 09:50-10:15 (page 89).

Ewa-Juditha Wójcik. „Zensus 2031: Methodentest Bevölkerung“.
Thursday, 10:15-10:40 (page 90).

Statistical Theory and Methods (Session 1)

Chair: Christian Weiß, Room: A.03.212

Ulrich Müller-Funk; Christina Ungerer. „Market Shocks: Counting Processes with Restarting Property“.
Thursday, 09:00-09:25 (page 90).

Marc-Oliver Pohle; Timo Dimitriadis; Jan-Lukas Wermuth. „Measuring Dependence between Events“.
Thursday, 09:25-09:50 (page 91).

Jan-Lukas Wermuth. „Proper Correlation Coefficients for Nominal Random Variables“.
Thursday, 09:50-10:15 (page 91).

Christian Weiß; Fukang Zhu; Hee-Young Kim. „Tobit models for count time series“.
Thursday, 10:15-10:40 (page 92).

VDSSt (Session 3): Regionale Energie- und Gebäudedaten

Chair: Uwe Meer, Room: F.04.208

Katharina Hembach-Stunden; Britta Stöver; Philip Ulrich. „Regionale Energiewende: Dashboard zur Steuerung der sozial-ökologischen Transformation“.
Thursday, 09:00-09:25 (page 92).

Jan Lunge. „Erarbeitung eines Gebäudekatasters für Wolfsburg im Kontext der Kommunalen Wärmeplanung“.
Thursday, 09:25-09:50 (page 93).

Amelie Plöger; Elena Stäger; Christoph Alfken; Marcel Noack; **Johannes Rohde**. „Validierung von Small Area-Methoden für den systematischen Einsatz in der amtlichen Statistik am Beispiel der Schätzung von Bestandsmieten auf Gemeindeebene in Nordrhein-Westfalen“.

Thursday, 09:50-10:15 (page 94).

Economic, Social and Market Statistics (Session 6): Labour Markets and Social Security II

Chair: Bernd Hofmann, Room: A.13.207

Jonas Krinitz; Johannah Zenk; Christian Schneemann; Jan-Phillip Schroer. „Arbeitsmarktanspannung: Wenn Wohnortwechsel und Pendeln nicht mehr reichen“.

Thursday, 09:00-09:25 (page 94).

Michael Hartmann. „Branchenwechsel von Beschäftigten“.

Thursday, 09:25-09:50 (page 95).

Karolin Hiesinger; **Laura Pohlan**; Franka Vetter. „The Employment Statistics of Severely Disabled People: Description and Research Potential“.

Thursday, 09:50-10:15 (page 95).

Coffee Break: 10:40 - 11:00

Thursday, 04. September 2025, 11:00 - 12:40

DGD (Session 2): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen II

Chair: Patrizio Vanella, Room: A.03.205

Christian von Malottki. „Häuser oder Wohnungen: Kann der Neubau seniorengerechter Wohnungen den Neubedarf an EFH reduzieren?“

Thursday, 11:00-11:25 (page 96).

Timon Hellwagner. „(Nicht-)Erwerbsbeteiligung im Kontext der demografischen Schrumpfung: ein Blick auf die Datenlage“.

Thursday, 11:25-11:50 (page 97).

Jan Philipp Schroer; **Alexander Schur**. „Modellierung des Arbeitsangebotes von Gesundheitsberufen für ein Fachkräftemonitoring des Bundesministeriums für Gesundheit“.

Thursday, 11:50-12:15 (page 97).

Michael Kalinowski. „QuBe - Projektion des zukünftigen Arbeitsangebots bis 2040“.

Thursday, 12:15-12:40 (page 98).

Empirical Economics and Applied Econometrics (Session 2)

Chair: Christoph Hanck, Room: A.00.Bibliothek

Jörg Stoye. „Robust Statistical Decisions: With Application to Development Economics“.
Thursday, 11:00-11:50 (page 99).

Tymon Słoczyński; **Derya Uysal**; Jeffrey Wooldridge. „Covariate Balancing and the Equivalence of Weighting and Doubly Robust Estimators of Average Treatment Effects“.
Thursday, 11:50-12:15 (page 99).

Lennard Maßmann; Eva-Maria Oeß. „Heterogeneous net treatment effects“.
Thursday, 12:15-12:40 (page 100).

Statistics in Finance (Session 2)

Chair: Roxana Halbleib, Room: B.03.104

Ilya Archakov; Roxana Halbleib; **Jasper Rennspies.** „Realized Covariance Modeling by Logarithmic Transformations and Dimensionality Reduction“.
Thursday, 11:00-11:25 (page 100).

Wanbo Lu; **Kris Boudt**; Guanglin Huang. „Estimation of Non-Gaussian Factors Using Higher-order Multi-cumulants in Weak Factor Models“.
Thursday, 11:25-12:15 (page 101).

Vasyl Golosnoy; Bastian Gribisch; Wolfgang Schmid; Miriam Isabel Seifert. „Combining Portfolio Rules to Improve Prediction of Global Minimum Variance Portfolio Weights“.
Thursday, 12:15-12:40 (page 101).

Methodology of Statistical Surveys (Session 4)

Chair: Hanna Brenzel, Room: E.03.112

Yannik Garcia Ritz; **Veronika Chakraverty.** „Auf der Suche nach dem „Sweet Spot“ zwischen Nutzen, Risiko und Komfort bei der Nutzung amtlicher Daten – Ermöglicht partielle Datensynthese-Anreicherung von Remote Access-Daten?“
Thursday, 11:00-11:25 (page 102).

Sarah Bohnensteffen; Pascal Gläske; Ralf Münnich. „Modellierung des Wohnraumbestands im Mikrosimulationsmodell MikroSim“.
Thursday, 11:25-11:50 (page 103).

Julian Ernst; Jan Weymeirsch; Ralf Münnich. „Modellrekalibrierung für räumlich-dynamische Mikrosimulationen“.
Thursday, 11:50-12:15 (page 103).

Kerstin Brückweh; **Svetlana Burmistr**; **Paul Treffenfeldt.** „Wie die Vergangenheit zählt. Zur Geschichte des Statistischen Bundesamtes (GeStat)“.
Thursday, 12:15-12:40 (page 104).

Statistical Theory and Methods (Session 2)

Chair: Matei Demetrescu, Room: A.03.212

Alexander Mayer; **Dominik Wied**; Victor Troster. „Quantile Granger Causality in the Presence of Instability“.

Thursday, 11:00-11:50 (page 105).

Uwe Hassler; Marc-Oliver Pohle; **Tanja Zahn**. „Simultaneous Inference Bands for Autocorrelations“.

Thursday, 11:50-12:15 (page 105).

Vivien Less; Philipp Sibbertsen. „A Perturbation Robust Test Against Spurious Long Memory“.

Thursday, 12:15-12:40 (page 106).

Regionalstatistik/VDSt: Regionale Verteilung von Einkommen

Chair: Diana Andrä, Room: F.04.208

Silke Joebges. „Kleinräumige Daten aus der Lohn- und Einkommensteuerstatistik – Potenziale für Analysen“.

Thursday, 11:00-11:25 (page 106).

Andrea Schultz. „Das Nettoäquivalenz-Resteinkommen als Sozialindikator für Miethaushalte“.

Thursday, 11:25-11:50 (page 106).

Economic, Social and Market Statistics (Session 7): National Accounts, Welfare Measurement

Chair: Albert Braakmann, Room: A.13.207

Albert Braakmann. „Das neue System of National Accounts 2025 - Neuerungen, Änderungen, Auswirkungen“.

Thursday, 11:00-11:25 (page 107).

Stefan Linz. „Strukturwandel in der Industrie aus unterschiedlichen Perspektiven: Produktionsindex der Konjunkturstatistik und Produktion in den Volkswirtschaftlichen Gesamtrechnungen“.

Thursday, 11:25-11:50 (page 107).

Klaus Pötzsch; **Maximilian Rotsche**. „Erkenntnisse aus 5 Jahren Arbeit der Large Cases Unit“.

Thursday, 11:50-12:15 (page 108).

Lunch Break: 12:40 - 14:20

Thursday, 04. September 2025, 14:20 - 16:00

DGD (Session 3): Regional, socioeconomic, and sex differences in morbidity and mortality

Chair: Michael Mühlichen, Enno Nowossadeck, Room: A.03.205

Michael Mühlichen; Pavel Grigoriev; Sebastian Klüsener. „Regionale Sterblichkeitsentwicklung nach Todesursachen in Deutschland vor, während und nach der COVID-19-Pandemie“.

Thursday, 14:20-14:45 (page 108).

Steffen Maretzke. „Ergebnisse unserer aktuellen Arbeit an einer Pflegebedarfsprognose. Eine weitere Komponente im System der Raumordnungsprognose des BBSR“.

Thursday, 14:45-15:10 (page 109).

Elena Rakuša; Daniela Georges; Gabriele Doblhammer. „Untersuchung des Einflusses sozialer Netzwerke und kognitiver Funktionen auf die Gesamtmortalität älterer Erwachsener unter Verwendung der Daten der NAKO Gesundheitsstudie“.

Thursday, 15:10-15:35 (page 110).

Sonja Nowossadeck; Svenja M. Spuling. „Soziodemografische und gesundheitliche Determinanten sportlicher Freizeitaktivität im mittleren und höheren Erwachsenenalter: Eine Analyse mit Daten des Deutschen Alterssurveys (DEAS) 2023“.

Thursday, 15:35-16:00 (page 111).

Empirical Economics and Applied Econometrics (Session 3)

Chair: Carsten Jentsch, Room: A.00.Bibliothek

Jean-Pierre Florens; **Anna Simoni**. „Panel data models with randomly generated groups“.

Thursday, 14:20-15:10 (page 111).

Jan Scherer. „Inference on Directionally Differentiable Functions of Nonparametric Estimators“.

Thursday, 15:10-15:35 (page 112).

Björn Höppner. „Group-Specific Heterogeneity in Short Binary Outcome Panels“.

Thursday, 15:35-16:00 (page 112).

High-dimensional Time Series and Network Data

Chair: Roxana Halbleib, Room: B.03.104

Miao Yu; Philipp Sibbertsen. „Decomposing Price-Energy in Germany and China’s Divergent Electricity Markets using STL and Multivariate Functional Data Analysis“.

Thursday, 14:20-14:45 (page 113).

Chris Fotso; Andrew Harvey; **Yeliz Özer**; Dario Palumbo; Philipp Sibbertsen. „Go with the Flow? Forecasting Regime Switches in Wind Direction“.

Thursday, 14:45-15:10 (page 114).

Daniel Dzikowski; Carsten Jentsch. „Sparsity and Fusion Penalization in Large Vector Autoregressions“.

Thursday, 15:10-15:35 (page 114).

Sven Pappert; Antonia Arsova. „Probabilistic forecasting and forecast reconciliation for wind power production“.

Thursday, 15:35-16:00 (page 115).

Methodology of Statistical Surveys (Session 5)

Chair: Michael Buchner, Room: E.03.112

Michael Buchner; Robert Garthoff. „Geheimhaltung georeferenzierter Ergebnisse der Lohn- und Einkommensteuerstatistik durch lokale Aggregation mit Hilfe des Quadtree-Verfahrens“.

Thursday, 14:20-14:45 (page 115).

Christoph Alfken; **Sören Görner**. „Veröffentlichungspotenziale von georeferenzierten Daten der Lohn- und Einkommenssteuerstatistik am Beispiel der kleinräumigen Einkommensverteilung in Nordrhein-Westfalen“.

Thursday, 14:45-15:10 (page 116).

Martin Möhler. „'Fitness for use' von demographischen Rasterdaten mit künstlichem Messfehler“.

Thursday, 15:10-15:35 (page 117).

Lorena Gril; Ulrich Rendtel. „Regionale Auswertung von anonymisierten georeferenzierten Daten“.

Thursday, 15:35-16:00 (page 117).

Statistical Theory and Methods (Session 3)

Chair: Lennart Empting, Room: A.03.212

Malte Jahn. „Stochastically interpretable distributional regression with neural networks“.

Thursday, 14:20-14:45 (page 118).

Matei Demetrescu; **Christoph Hanck**; Hajo Holzmann. „Assessing Loss Preferences on the Basis of Noisy Forecast Errors“.

Thursday, 14:45-15:10 (page 118).

Lennart Empting; **Saskia Öztürk**; Simone Maxand; Konstantin Wagner. „Inference in Panel SVARs with Cross-Sectional Dependence of Unknown Form“.

Thursday, 15:10-15:35 (page 119).

Lennart Empting. „The pvars R-Package: VAR modeling for heterogeneous panels“.

Thursday, 15:35-16:00 (page 119).

Economic, Social and Market Statistics (Session 8)

Chair: Andreas Kladroba, Room: A.13.207

Judith Niehues. „Einkommensverteilung und Armut in Deutschland: Aktuelle Verteilungsdaten auf dem Prüfstand“.

Thursday, 14:20-14:45 (page 120).

Michael Webá. „Schätzen und Testen der Übersterblichkeiten während der Corona-Pandemie im Zeitraum 2020 bis 2022“.

Thursday, 14:45-15:10 (page 121).

Christian Salwiczek. „Internationalität der Unternehmensgruppen am Wirtschaftsstandort Hamburg“.

Thursday, 15:10-15:35 (page 121).

Benjamin Fuchs; Martin Kastner. „Die Entwicklung der E-Mobilität in Deutschland - Eine Analyse regionaler Unterschiede“.

Thursday, 15:35-16:00 (page 122).

General Assembly Meeting VDSSt

Thursday, 14:20-16:00, Room: F.04.208

Coffee Break: 16:00 - 16:20

Thursday, 04. September 2025, 16:20 - 18:00

Methodology of Statistical Surveys (Session 6)

Chair: Marcel Preising, Room: E.03.112

Clara Seesing-Coelho; Christina Jaeger; Marcel Preising. „Bereitstellung von Längsschnittgewichten für das Beschäftigtenpanel“.

Thursday, 16:20-16:45 (page 123).

Charlotte Articus. „EU-SILC: Änderungen im Auswahl- und Gewichtungsverfahren für die deutsche SILC-Stichprobe ab 2026“.

Thursday, 16:45-17:10 (page 123).

Lara Krell; Ralf Münnich. „Das Einheitenproblem in Unternehmensstatistiken - Herausforderungen und Auswirkungen auf Schätzungen“.

Thursday, 17:10-17:35 (page 124).

Jacques Saliba. „Standard procedure for methodological aspects of thematic surveys at the Federal Statistical Office, Switzerland“.

Thursday, 17:35-18:00 (page 125).

(Systemic) Risk Modeling

Chair: Yannick Hoga, Room: B.03.104

Carlotta Pacifici; **Luca Trapin.** „Monitoring time-varying tail risk with the extended Generalized Pareto distribution“.

Thursday, 16:20-17:10 (page 125).

Sander Barendse. „Expected Shortfall LASSO“.

Thursday, 17:10-17:35 (page 126).

Alberto Quaini. „Nuclear Hedging Portfolios: Shrinking Complexity for Robust Diversification“.
Thursday, 17:35-18:00 (page 126).

Statistics in the Environmental Sciences, Natural Sciences and Technology (Session 1)

Chair: Gaby Schneider, Room: F.04.208

Sebastian Döhler; Iqraa Meah. „Some new null-proportion estimators with plug-in FDR control.“
Thursday, 16:20-16:45 (page 127).

Tiago Vasconcelos Afonso; **Florian Heinrichs.** „EEG-based Eye-Tracking: A Benchmark Dataset for Functional Data Analysis with Open Challenges and Baseline Results“.
Thursday, 16:45-17:10 (page 127).

Jethro Browell. „Forecasting the wind: From necessity to added value“.
Thursday, 17:10-18:00 (page 128).

Statistical Theory and Methods (Session 4)

Chair: Matei Demetrescu, Room: A.03.212

Tommaso Proietti; Alessandro Giovannelli. „On the Estimation of Climate Normals and Anomalies“.
Thursday, 16:20-17:10 (page 128).

Joint DStatG and TIES Session

Chair: Philipp Otto, Room: A.00.Bibliothek

Steffen Maletz; Konstantinos Fokianos; Roland Fried. „glmSTARMA - An R package for fitting spatio-temporal models based on generalized linear models“.
Thursday, 16:20-16:45 (page 129).

Finn Hughes; Manuel Stavridis; Manuel Marschall. „Bayesian estimation for virtual experiments in metrology featuring semi-parametric modelling using Gaussian processes“.
Thursday, 16:45-17:10 (page 129).

Merle Mendel; Roland Fried. „Statistical Modeling of Clustering and Seasonality in Return Times of Midlatitude Cyclones“.
Thursday, 17:10-17:35 (page 130).

Stefan Irrgang; Maren Köhlmann; Frederik Stellmach; Carola Stolle. „Earth Observation Data and AI for Construction Statistics“.
Thursday, 17:35-18:00 (page 130).

Economic, Social and Market Statistics (Session 9): Labour Markets and Social Security III

Chair: Christian Zemann, Room: A.13.207

Johannes Ludsteck. „Multiple Imputation zur Umschätzung der Berufsklassifikationen der Deutschen Arbeitsmarktstatistik“.

Thursday, 16:20-16:45 (page 131).

Martina Rengers. „Veränderte Arbeitszeitwünsche: Abbildung der Realität oder Effekte der Befragungsmethode? – Ergebnisse des Mikrozensus“.

Thursday, 16:45-17:10 (page 132).

Gerald Seidel. „Saisonalität am Arbeitsmarkt“.

Thursday, 17:10-17:35 (page 133).

Thursday, 04. September 2025, 18:00 - 20:00**Poster Session**

Chair: Arne Johannssen, Room: D-Foyer

Martin Popp; Michael Oberfichtner. „Arbeitsmarktkonzentration in Deutschland“. *(page 133).*

Norah Efosa; Adrian Gadiant. „Bevölkerungsprojektionen mit dem R-Paket popop“. *(page 134).*

Marlon Suárez. „Methodische Grundlagen und Herausforderungen einer bundesweiten Kulturstatistik“. *(page 134).*

Gloria Deetjen; Natalie Rosenski. „Methoden für die Integration von Mobilfunkdaten mit weiteren Datenquellen“. *(page 135).*

Lukas Olbrich; Joseph W. Sakshaug; Eric Lewandowski. „Evaluating methods to prevent and detect inattentive responding in web surveys“. *(page 135).*

Christian Mueller. „pRDM: A note on a popular collusion marker“. *(page 136).*

Ulf von Kalckreuth. „Harnessing the power of Input-Output analysis for sustainability. A simulation study based on US data“. *(page 136).*

Dominik Schulz. „Das R-Paket „deseats“ zur simultanen Schätzung von Trend und Saisonalität in Zeitreihen anhand datengesteuerter lokal gewichteter Regression“. *(page 137).*

Markus Zwick; Natalie Rosenski; **Johannes Volk.** „NeDaMo - Neue Daten für Mobilitätsanalysen“. *(page 138).*

Other events on Thursday

18:30 : General Assembly Meeting DStatG, Room: A.13.207

18:30 : Relaxed Get-Together – Das Wohnzimmer, see page 152

21:00 : Wiesbaden Casino, see page 152

Friday, 05. September 2025, 09:00 - 10:40

Regional Statistics (Session 1)

Chair: Anna Herget, Room: E.03.112

Dirk Brunken; Corinna Mundzeck; **Christian Salwiczek**; **Julian Schulz**. „Hotspotanalyse der Automobilindustrie in Deutschland“.

Friday, 09:00-09:25 (page 138).

Georg Wiegleb. „Methodische Ansätze zur Modellierung der Lage in Mietspiegeln“.

Friday, 09:25-09:50 (page 139).

Tobias Wolfanger. „HILDA - Ein Chatbot für die Landesdatenbank NRW“.

Friday, 09:50-10:15 (page 139).

Statistics in the Environmental Sciences, Natural Sciences and Technology (Session 2)

Chair: Anja Bettina Schmiedt, Room: F.04.208

Tom Hoger; Tobias Diergardt; Dennis Quandt. „Die Entwicklung der Energiebilanzierung in den Bundesländern“.

Friday, 09:00-09:25 (page 140).

Ulf von Kalckreuth. „Measuring the carbon footprints of portfolios. Emission intensities and financed emissions indicators from Multi-Regional IO and Air Emission Accounts data“.

Friday, 09:25-09:50 (page 140).

Jekaterina Zukovska. „Sensibilität für versunkene Kosten am Beispiel des Mausverhaltens in der Restaurant-Row-Aufgabe.“

Friday, 09:50-10:15 (page 141).

Anja Bettina Schmiedt; Narayanaswamy Balakrishnan; Erhard Cramer. „Goodness-of-fit Tests for Exponential and Weibull Distributions in Progressively Type-II Censored Data“.

Friday, 10:15-10:40 (page 142).

Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 1)

Chair: Katharina Schüller, Room: A.13.207

Michael Neutze; Tobias Lentz. „Visualisierung schiefer Verteilungen am Beispiel von Einkommen und Miete“.

Friday, 09:00-09:25 (page 142).

Gerit Offermann. „Einkommensverteilung in SILC“.

Friday, 09:25-09:50 (page 143).

Robert Buhles; **Roland Jeske**. „Wer trägt die Einkommensteuerlast in Deutschland?“

Friday, 09:50-10:15 (page 144).

Coffee Break: 10:40 - 11:00

Friday, 05. September 2025, 11:00 - 12:40

Panel Discussion: 20 Jahre Code of Practice: Kritische Würdigung und Reformbedarf der europäischen Qualitätskriterien

Chair: Daniel Vorgrimler, Room: A.00.Bibliothek

Susanne Hagenkort-Rieger; Walter J. Radermacher; Kerstin Schneider. „20 Jahre Code of Practice: Kritische Würdigung und Reformbedarf der europäischen Qualitätskriterien“.
Friday, 11:00-12:40 (page 144).

Regional Statistics (Session 2)

Chair: Sara Bleninger, Room: E.03.112

Andreas Gleich. „Melderegister vs. Zensus zur Ermittlung der Einwohnerzahl“.
Friday, 11:00-11:25 (page 145).

Andreas van de Loo. „Der Zensus 2022 - eine besondere Datenquelle zum Thema Wohnen und Haushalte“.
Friday, 11:25-11:50 (page 145).

Leonie Grauwinkel. „Der Zensus 2022 – ein Datenschatz“.
Friday, 11:50-12:15 (page 146).

Statistics in the Environmental Sciences, Natural Sciences and Technology (Session 3)

Chair: Sven Knoth, Room: F.04.208

Duohong Sun; Ansgar Steland. „Asymptotic studies of adapted threshold detectors based on density processes“.
Friday, 11:00-11:25 (page 146).

Sven Knoth. „EWMA control charts for exponential data — numerical methods and transformations“.
Friday, 11:25-11:50 (page 147).

Anton Imm; Ansgar Steland. „Detecting change-points of multivariate time series using the Wasserstein distance between the empirical marginal distributions“.
Friday, 11:50-12:15 (page 147).

Maik Ulmer. „EWMA control charts for the correlation coefficient“.
Friday, 12:15-12:40 (page 148).

Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 2)

Chair: Markus Zwick, Room: A.13.207

Ricarda Senger; Stefan Hauf. „Makroökonomische Einkommensverteilung: Hintergrund und Anforderungen aus dem SNA 2025 und der G20 Data Gap Initiative“.*Friday, 11:00-11:25 (page 148).***Ricarda Senger; Felix Lederer; Frederik Wijnck;** Aaron Hennig. „Makroökonomische Einkommensverteilung: Überblick des methodischen Vorgehens“.*Friday, 11:25-11:50 (page 149).***Junyi Zhu.** „Microfoundations in measurement and theory“.*Friday, 11:50-12:15 (page 150).*Julia Le Blanc; **Tobias Schmidt;** Anna Stelzer. „The Joint Distribution of Income, Wealth and Consumption in Germany“.*Friday, 12:15-12:40 (page 150).*

Closing: 13:00 - 13:30

Room: A.13.207

Other events on Friday

13:45 : City Tour, see page 152

1 Estimation of Uncertainty

Mehdi Dagdoug

Mc Gill University, Kanada
mehdi.dagdoug@mcgill.ca

SECTION: Plenary Session 1

Predictive models are increasingly used in surveys for tasks such as model-based and model-assisted estimation, as well as handling nonresponse through imputation and reweighting. The rise of statistical learning has provided survey statisticians with flexible tools, but incorporating them into survey estimation strategies poses significant challenges for valid inference. In particular, variance estimation becomes delicate when using black-box machine learning methods.

I will show that naively plugging machine learning predictions into classical estimators can lead to invalid inference and misleading uncertainty quantification. Through theoretical and empirical results, I will illustrate why many standard approaches fail with nonparametric regression. To address this, I will present an extension of the Double Machine Learning framework with cross-fitting to survey sampling that supports valid inference with arbitrary black-box methods. Under realistic conditions, the resulting estimators are shown to be square-root n consistent and asymptotically normal. I will also introduce new variance estimators based on cross-fitting that remain consistent across methods, enabling the construction of asymptotically valid confidence intervals. Issues related to model selection and aggregation will also be discussed. Simulation studies demonstrating the strong performance of the proposed methods will be presented.

2 Which Factors Drive Downside Risk in the U.S. Economy?

Christian Brownlees (1), Carlo Pavanello (2), Andre Souza (3)

1: Universitat Pompeu Fabra, Spain; 2: Universitat Pompeu Fabra, Spain; 3: ESADE Business School, Spain
christian.brownlees@upf.edu

SECTION: Plenary Session 2

We assess the role of common factors in the prediction of downside risk for a large panel of US macroeconomic time series obtained from the FRED-MD database. The panel exhibits pervasive cross-sectional tail dependence, with the joint occurrence of extreme events being power-law distributed. The pool of candidate factors we consider includes macroeconomic, financial, text-based and statistical factors. Results show that the downside risk is captured by one factor that can be associated with macroeconomic volatility and appears to be best captured by the macroeconomic uncertainty index of Jurado et al. (2015).

3 Mikro- und Makroökonomische Verteilungsfragen zu Einkommen, Vermögen und Konsum

Michael Hüther

Institut der deutschen Wirtschaft, Deutschland
huether@iwkoeln.de

SECTION: Plenary Session 3

Die ökonomischen Auswirkungen außergewöhnlicher Krisen – insbesondere der Coronapandemie und der Energiepreiskrise – haben zu erheblichen staatlichen Ausgabenerhöhungen geführt und werfen neue verteilungspolitische Fragen auf. Der Vortrag beleuchtet zentrale mikro- und makroökonomische Verteilungsfragen mit Blick auf Einkommen, Vermögen und Konsum. Die Einkommensverteilung zeigt zwischen 2005 und 2019 eine weitgehende Seitwärtsbewegung des Ungleichheitsniveaus. Eine unklare Datenlage aufgrund methodischer Probleme in allen verfügbaren Befragungsdaten erschwert jedoch eine verlässliche Aussage zur weiteren Entwicklung ab 2020. Die Vermögensverteilung weist gegenüber der Zeit vor der globalen Finanz- und Wirtschaftskrise tendenziell eine rückläufige Ungleichheit auf. Das im internationalen Vergleich hohe Ungleichheitsniveau in Deutschland muss aber vor dem Hintergrund der umfangreichen sozialstaatlichen Absicherung eingeordnet werden.

Ergänzend werden politökonomische Implikationen im Vortrag diskutiert, etwa die ausgeprägte Inflationswahrnehmung von Anhängern der politischen Ränder in Deutschland sowie der fehlende bivariate Zusammenhang zwischen Vermögensungleichheit und rechtem Wahlverhalten auf europäischer Ebene.

Der Beitrag unterstreicht die Notwendigkeit verfügbarer, methodisch konsistenter Datenquellen, um Verteilungswirkungen aktueller Krisen belastbar analysieren und politisch fundiert adressieren zu können.

4 Mietspiegel in Deutschland - Ein Spannungsfeld zwischen Politik, Immobilienwirtschaft und Wissenschaft

Göran Kauermann

LMU München, Deutschland
goeran.kauermann@lmu.de

SECTION: Heinz-Grohmann-Lecture

Mietspiegel sind seit gut 50 Jahren ein etabliertes Instrument zur Bestimmung der „ortsüblichen Vergleichsmiete“. Seit wenigen Jahren fordert das „Mietspiegelgesetz“ die Erstellung eines Mietspiegels in allen Gemeinden mit mehr als 50.000 Einwohnern. Dabei wird zwischen einfachen und qualifizierten Mietspiegeln unterschieden; letztere müssen „nach anerkannten wissenschaftlichen Grundsätzen“ erstellt sein. Zur Erstellung von Mietspiegeln werden Daten erhoben, und diese durch Analysen zu einem Vorhersagemodell für die ortsübliche Vergleichsmiete verarbeitet. Und hier kommt die Statistik ins Spiel.

Der Vortrag betrachtet Mietspiegel von statistischer Seite her. Dabei wird auf neuere gesetzliche Änderungen eingegangen und das Spannungsfeld aufgezeigt zwischen rechtlicher und statistischer Betrachtung der Materie. Das Gesetz und die Rechtsprechung geben den Rahmen vor, oder ist es die Statistik und Wissenschaft, die den Rahmen vorgeben sollte?

Das Thema Mietspiegel berührt viele Bereiche des öffentlichen Lebens. Ein Großteil der Menschen in Deutschland lebt in Mietwohnungen, der Preisdruck in Metropolen ist groß und der Immobilienmarkt ist attraktiv. Wie kann Statistik hier für Durchblick und wissenschaftliche Sichtweise sorgen. Wir geben ein paar Anregungen im Vortrag. Ganz im Sinne von Prof. Dr. Heinz Grohmann verbinden wir statistische Wissenschaft mit Anwendung und politischer Diskussion.

5 Forecast Comparisons via Multi-Objective Elicitability

Tobias Fissler (1), Yannick Hoga (2)

1: ETH Zürich, Schweiz; 2: Universität Duisburg-Essen, Deutschland
yannick.hoga@vwl.uni-due.de

SECTION: Gumbel-Lecture

This talk summarizes recent joint work with Tobias Fissler on the use of multivariate loss functions in forecast evaluation. Statistical functionals (e.g., mean, variance, quantiles) for which such multivariate loss functions exist are termed multi-objective elicitable. We discuss two applications. First, we show that many systemic risk measures (such as CoVaR) are not elicitable in the usual sense, i.e., no suitable univariate loss functions exist. However, bivariate (multi-objective) loss functions exist that allow for meaningful comparisons of systemic risk forecasts. Second, we demonstrate that while copula forecasts can be evaluated jointly with the marginals using standard methods, this does not allow practitioners to distinguish between the quality of the copula component only. Yet, by using suitable bivariate loss functions, such an attribution becomes feasible.

6 Uncovering uncertainty in narrative economics: a semantic search approach

Florian Schütze (1), Sami Diaf (2)

1: Helmut-Schmidt-Universität, Deutschland; 2: University of Hamburg, Deutschland
schuetzf@hsu-hh.de

SECTION: Uncertainty Estimation (Session 1)

Monetary policy, as an important application of narrative economics, carries uncertainty in its decision-making process as a forward guidance tool, as well as being a strategic aspect of its communication policy. It has witnessed several attempts to construct uncertainty indices using uncertainty-related word counts, yielding questionable measurements that overlook key technical terms not included in word lists, thus rendering the computed indices biased and semantically agnostic regarding the actual jargon. This work proposes an in-depth assessment of uncertainty in a collection of international central bankers' speeches (1997-2022) and identifies its key drivers using semantic search models, namely Top2vec, to uncover nested semantic topic structures at the national and international levels as proxies for uncertainty sources. These were found to be robust in discriminating uncertainty features related to probability and risk, in line with the Keynes-Knight debate on uncertainty in macroeconomics. In addition, a robust uncertainty index could be constructed from the similarity between each document and the notion of uncertainty, either country-specific or international, which was found to follow the major global financial and banking events of the last two decades. The method of constructing an uncertainty indicator

that measures uncertainty in central bank communications has been shown to have significant predictive power for GDP growth. This ability is comparable to other measures of uncertainty, such as the Economic Policy Uncertainty Index and the Monetary Policy Uncertainty Index. Unlike the other two, the proprietary World Uncertainty Indicator cannot predict a country's economic growth.

7 Methodological Issues with Measuring Economic Uncertainty

Tamar Abesadze, Christian Müller

Constructor University, Germany
tabesadze23@gmail.com

SECTION: Uncertainty Estimation (Session 1)

This paper provides a review and critical analysis of how economic uncertainty has been defined and measured across foundational and modern literature. One primary contribution of this paper is the critical assessment of whether empirical measures are in lines with their foundational theoretical definitions. By developing a conceptual framework derived from fundamental definitions, this paper evaluates the consistency of recent approaches and challenges the bias to equate uncertainty with observable fluctuations. It argues that although many approaches apply innovative measures, they are inconsistent with the core theoretical and interpretive frameworks of uncertainty. This paper emphasises the importance for clearer linkage between theory and measurement and calls for careful consideration in treating model driven volatility as representative of uncertainty.

8 Measuring political instability at a high frequency – a text mining approach

Niklas Benner (1,2), Boris Blagov (1), Maximilian W. Dirks (1,3)

1: RWI - Leibniz Institute for Economic Research; 2: Chair of Business and Social Statistics, Department of Statistics, TU Dortmund University; 3: Chair of Empirical Macroeconomics, Faculty of Management and Economics, Ruhr-University Bochum
nbenner@rwi-essen.de

SECTION: Uncertainty Estimation (Session 1)

This study introduces a novel monthly index of political instability, the WIPI, covering four dimensions for 182 countries from 1996 to 2023. Employing text mining techniques, we extract data from the Economist Intelligence Unit's country reports. We show that our monthly index carries the same information as other frequently used indices of political instability but with a much higher frequency. Using local projections, we identify that political instability reduces economic output significantly. Decomposing the index into its dimensions reveals that instability within the political regime and mass civil protests are the key drivers of the sizable decline in prosperity. We further show that estimates of models employing annual data suffer from a sizeable temporal aggregation bias.

9 EU-MD, macroeconomic uncertainty and a state-dependent Phillips curve

Pascal Goemans

FernUniversität in Hagen, Deutschland
pascal.goemans@fernuni-hagen.de

SECTION: Uncertainty Estimation (Session 1)

We construct EU-MD, a novel database similar to FRED-MD featuring more than 100 monthly economic time series, to construct macroeconomic uncertainty indices for the set of EU countries and the euro area aggregate. For each cross-section, macroeconomic uncertainty relies on the common uncertainty estimated via stochastic volatility models in the forecast errors of factor augmented autoregressive models for all economic time series. To deal with the large impact of the COVID-19 pandemic, we use COVID-indicators to de-COVID the time series prior to factor estimation. The proposed econometric-based uncertainty indicator has several advantages over existing measures in the literature based on surveys and newspapers. In contrast to survey-based uncertainty measures, our index can be updated regularly at low costs. With regards to newspaper-based indices, our index reflects the uncertainty over a lot of domestic economic variables and is less vulnerable to being influenced by global events or ideological distortions from periodicals in different countries. We use the constructed macroeconomic uncertainty index to estimate a state-dependent Phillips curve with a monthly panel starting in 1999M1 of the euro area countries. The results imply different signs in the response of inflation to the unemployment gap for various threshold values used to distinguish across uncertainty states. In contrast, the persistence of inflation is relatively independent of the uncertainty level.

10 Robust Long Short-Term Memory Models for Outlier-Contaminated Time Series

Klaus Nordhausen (1), Christophe Croux (2), Mika Sipilä (3), Sara Taskinen (3)

1: University of Helsinki, Finland; 2: KU Leuven, Belgium; 3: University of Jyväskylä, Finland
klaus.nordhausen@helsinki.fi

SECTION: Nonparametric and Robust Statistics (Session 1)

Long Short-Term Memory (LSTM) models are a special case of recurrent neural networks and have become a standard tool in the deep learning community for time series prediction. Despite the common belief that deep neural networks can handle highly nonlinear and noisy data, standard LSTM models are not robust to outliers.

In this talk, we introduce a robust version of LSTM. Two quick fixes for improving robustness are considered: (i) scaling the time series using robust statistics such as the median and MAD, and (ii) replacing the least squares loss function with the Huber loss. A further improvement is achieved by adding a cleaning step within an iterated version of LSTM, which forms our proposed robust LSTM model.

Using simulation experiments, we show that robust LSTM can handle different types of outliers, including level shifts and patches of outliers. Finally, we explore how robust LSTM models can be used for outlier detection in time series and evaluate the accuracy of this detection approach.

11 Non-parametric tests for cross-dependence based on multivariate extensions of ordinal patterns

Angelika Silbernagel (1), Christian H. Weiß (1), Alexander Schnurr (2)

1: Helmut-Schmidt-Universität, Deutschland; 2: Universität Siegen, Deutschland
silbernagelangelika@gmail.com

SECTION: Nonparametric and Robust Statistics (Session 1)

Since their introduction, ordinal patterns have become a popular tool for data analysis. As the name may already suggest, ordinal patterns capture the ordinal structure of the underlying data. They have many desirable properties like invariance under monotone transformations, robustness with respect to small noise and simplicity in application. In particular, ordinal patterns are able to capture possibly non-linear dependence.

Recently, there has been a growing interest in extending ordinal patterns to multivariate time series in a way that takes potential correlations between the movement of the variables into account. We describe different concepts for measuring cross-dependence in sequentially observed random vectors, which are based on ordinal patterns or multivariate generalizations of them. In all cases, we derive the limiting distribution of the corresponding test statistics. In a simulation study, we compare the performance of these statistics with three competitors, namely, classical Pearson's and Spearman's correlation as well as the rank-based Chatterjee's correlation coefficient.

12 Einstellungen zur Statistik bei Studierenden

Christine Buchholz

Hochschule Bonn-Rhein-Sieg, Deutschland
christine.buchholz@h-brs.de

SECTION: Competence Development: Data Literacy and Statistics (Session 1)

Dieser Beitrag thematisiert die Bedeutung der Einstellungen zur Statistik bei Studierenden. Positive Einstellungen zur Statistik werden als förderlich für das Lernverhalten und den Lernerfolg angesehen. In diesem Kontext werden empirische Befunde aus der Befragung von Studierenden der Wirtschaftspsychologie und der Betriebswirtschaftslehre an der Hochschule Bonn-Rhein-Sieg im Semesterverlauf dargestellt.

Um den Einfluss der Einstellungen gegenüber Statistik auf das Lernverhalten und den Lernerfolg zu analysieren, werden zwei Sachverhalte untersucht: Zum einen die Zusammenhänge zwischen den individuellen Voraussetzungen der Studierenden und ihren Einstellungen zur Statistik, zum anderen aber auch die Veränderungen dieser Einstellungen im Verlauf eines Semesters.

Dabei werden die Einstellungen zur Statistik als multidimensionales Konstrukt verstanden, das sich aus affektiven, kognitiven und verhaltensbezogenen Komponenten zusammensetzt. Diese wiederum beeinflussen jeweils das Lernen und die Anwendung statistischen Wissens.

Darüber hinaus wird das multidimensionale Konstrukt in sechs Einstellungsdimensionen differenziert: Wert, Affekt, Anstrengung, kognitive Kompetenz, Schwierigkeit und Interesse. Diese Dimensionen werden hinsichtlich ihres Einflusses betrachtet.

Die Datenerhebung erfolgte im Rahmen einer Panelstudie mittels Online-Fragebogen, wobei Daten von 144 Studierenden vorliegen. Die Auswertung erfolgte mittels Korrelationstests und t-Tests für abhängige Stichproben.

Die Ergebnisse zeigen überwiegend neutrale bis leicht positive Einstellungen der Studierenden gegenüber der Statistik. Ein positiver Zusammenhang resultiert auch aus guten mathematischen Vorkenntnissen und dem Vertrauen in die eigenen Fähigkeiten. Statistisch signifikante Veränderungen im Semesterverlauf zeigten sich in den Dimensionen Schwierigkeit, kognitive Kompetenz und Anstrengung, während Affekt, Interesse und Wert zeitlich konstant blieben.

Insgesamt bilden die neutralen bis leicht positiven Einstellungen eine wichtige Grundlage für die zukünftige Anwendung statistischen Wissens und eine nachhaltig positive Einstellung zur Statistik. Die Ergebnisse verdeutlichen das Interesse an weiterer Forschung, um Studierende gezielt hinsichtlich ihrer Einstellung und Kompetenzentwicklung unterstützen zu können.

13 Digitale Didaktik in der Hochschullehre: Technologische Entwicklungen und ihr Nutzen für datengetriebene Berufsfelder am Beispiel der Mathematik

Alexander Neumann, Ramona Voshage

Hochschule für Wirtschaft und Recht Berlin, Deutschland
ramona.voshage@hwr-berlin.de

SECTION: Competence Development: Data Literacy and Statistics (Session 1)

Angesichts der wachsenden Bedeutung von Künstlicher Intelligenz (KI), Data Science und Programmierung in der Arbeitswelt ist eine intensivere Einbindung dieser Themen in die Hochschullehre unverzichtbar. Zwar bringen viele Studierende bereits Interesse an diesen Bereichen mit, doch fehlen ihnen häufig fundierte praktische Erfahrungen und theoretische Vorkenntnisse, um den steigenden beruflichen Anforderungen gerecht zu werden.

Um diesem Bedarf zu begegnen, wird in diesem Beitrag das Konzept eines neu entworfenen adaptiven digitalen Selbstlernkurses vorgestellt, der bereits erfolgreich für das selbstständige Lernen von Mathematik eingesetzt wird. Der Kurs hat das Potenzial um Inhalte aus den Bereichen KI, Data Science, Programmierung und Statistik erweitert zu werden und so Studierenden essenzielle Kompetenzen in diesen zunehmend geforderten Fachgebieten zu vermitteln. Durch automatisch generierte Übungsaufgaben und auf Fehlerquellen basierend personalisiertes Feedback schafft das System eine interaktive und flexible Lernumgebung, die Studierende gezielt in ihren individuellen Lernprozessen unterstützt. Die in den letzten Jahren entwickelte technische Infrastruktur zur dynamischen Aufgaben- und Feedbackgenerierung des Kurses orientiert sich dabei an erprobten didaktischen Konzepten der adaptiven Lernförderung sowie neuen Erkenntnissen der Neurowissenschaften und nutzt neue Softwarelösungen zum Auslesen und Verarbeiten algebraischer Ausdrücke, um diese Konzepte umzusetzen.

Ziel des Vortrags ist es, ein skalierbares Konzept für ein digitales Selbstlernangebot zu präsentieren, mit dem Hochschulen ihre Studierenden praxisnah und zukunftsorientiert auf die aktuellen Herausforderungen der datengetriebenen Arbeitswelt vorbereiten können.

14 Data Literacy für den Geographieunterricht: Mit amtlicher Statistik kartenbasiert argumentieren

Andreas Köhler (1), Isabelle Kunze (2), **Frederik von Reumont** (2), **Christoph Alfken** (1), Alexandra Budke (2), Maren Franzen (1)

1: IT.NRW, Statistisches Landesamt Nordrhein-Westfalen; 2: Institut für Geographiedidaktik, Universität zu Köln
andreas.koehler@it.nrw.de

SECTION: Competence Development: Data Literacy and Statistics (Session 1)

Data Literacy - der bewusste und kritische Umgang mit Daten - umfasst nicht nur Statistikkompetenz oder ein mathematisches Grundverständnis, sondern auch Fertigkeiten wie Digital- und Medienkompetenz. Entsprechende Kompetenzen sollten bereits im Kindes- und Jugendalter in der Schule gebildet werden. So ist es in Zeiten von Fake News und Populismus im Interesse der amtlichen Statistik, dass ihre Ergebnisse und ihre Rolle als zuverlässige Quelle in den Curricula stärker abgebildet werden.

Durch das breite Spektrum an Daten der amtlichen Statistik gibt es Ansatzpunkte zu verschiedenen Fächern. Große fachliche Überschneidungen konnten mit Curriculumsinhalten des Geographieunterrichts identifiziert werden. Beispielsweise lassen sich die Kartenmaterialien aufgrund unterschiedlichster sozioökonomischer und teils ökologischer Indikatoren den Bereichen der Bevölkerungsentwicklung, Stadt- und Raumplanung, Migration, Standortentscheidungen und weiteren, für den Geographieunterricht relevanten Themengebieten zuordnen.

In Zusammenarbeit mit dem Lehrstuhl für Geographiedidaktik der Universität zu Köln hat Information und Technik Nordrhein-Westfalen (IT.NRW) als Statistisches Landesamt Aufgaben für den Geographieunterricht verschiedener Klassenstufen konzipiert und veröffentlicht, die auf Grundlage der Kartendienste von IT.NRW zu bearbeiten sind. Geschaffen wurde so ein breites Angebot an Aufgaben für verschiedene Klassenstufen, die auf Daten zu unterschiedlichen humangeographischen Themen rekurrieren. Die Aufgaben wurden mit Lehramtsstudierenden in Seminaren entwickelt und in Gruppendiskussionen getestet und verbessert.

Den Aufgaben liegt dabei das Konzept der kartenbasierten Argumentation zugrunde. Dessen Ziel ist, zu kontroversen, geographischen Fragestellungen eine fachliche Argumentation zu entwickeln, deren Belege sich auf Karten des Statistischen Landesamts NRW stützen. Neben der Fähigkeit der kartenbasierten Argumentation wird gleichzeitig die Medienkompetenz durch den Einsatz digitaler Karten, die zunehmend an Bedeutung gewinnen, gefördert. Die digitalen Kartenangebote des Statistischen Landesamts NRW eignen sich daher besonders für den Einsatz im Unterricht.

Im Vortrag werden die Zusammenarbeit von Wissenschaft und amtlicher Statistik, das Themenspektrum, der Ansatz der kartenbasierten Argumentation, das erstellte Angebot an Aufgaben und Ergebnisse der Diskussionen mit Lehramtsstudierenden und der ersten Feedbacks vorgestellt.

15 Learning to apply statistics – the statistical practicum in Frankfurt

Gaby Schneider

Goethe University Frankfurt, Deutschland
schneider@math.uni-frankfurt.de

SECTION: Competence Development: Data Literacy and Statistics (Session 1)

The Statistical Practicum is an interdisciplinary seminar designed for mathematics students with a background in statistics. The course focuses on developing statistical literacy, data competence, and communication skills for working across disciplines. While formal proofs and analytical methods are commonly taught in standard lectures, courses like this one are essential for building practical skills that enable students to apply theoretical knowledge in real-world contexts.

Each semester, the practicum collaborates with a research group from the natural or life sciences to address a specific research question. Over the course of the seminar, students explore a range of statistical techniques of varying complexity and apply them to an empirical dataset.

In a series of carefully prepared presentations, each student examines the potential and limitations of their chosen method. These presentations emphasize (1) the rigorous formal explanation of the method, (2) an intuitive illustration of how it works, (3) a critical discussion of its suitability for the specific dataset, (4) the careful application of the method, (5) clear and accessible communication of the method to researchers from other disciplines, and (6) appropriate and effective graphical presentation of the results.

The course concludes with a final session featuring concise presentations prepared for the partnering research group. An accompanying website showcases illustrated abstracts — see www.math.uni-frankfurt.de/schneide/StatistischesPraktikum.html

We also discuss feedback, career benefits, and the teaching resources required for the course.

16 Die Auswirkungen der EU-Einheitenverordnung auf die FuE-Erhebung (Teil 2)

Andreas Kladroba (1), Annette Lotz (2)

1: Stifterverband für die Deutsche Wissenschaft, Deutschland; 2: Statistisches Bundesamt
andreas.kladroba@stifterverband.de

SECTION: Economic, Social and Market Statistics (Session 1)

Die EU-Einheitenverordnung (EWG) Nr. 696/93 definiert das Unternehmen als eine Einheit zur Erzeugung von Waren und Dienstleistungen, die insbesondere in Bezug auf die Verwendung der ihr zufließenden laufenden Mittel über eine gewisse Entscheidungsfreiheit verfügt". Dabei kann es sich um eine rechtliche Einheit oder eine Zusammenfassung mehrerer rechtlicher Einheiten handeln. Die Verordnung ist in den meisten EU-Ländern lange Jahre weitgehend ignoriert worden und rückt erst in jüngerer Zeit auf Druck von Eurostat in den Fokus der Statistischen Ämter.

Das Statistische Bundesamt hat dabei im Rahmen des sogenannten "Profiling" große Anstrengungen unternommen, um im Unternehmensregister die Unternehmen gemäß der oben genannten Verordnung zu identifizieren.

Der Stifterverband für die Deutsche Wissenschaft, der im Auftrag des BMBF die FuE-Erhebung im Unternehmensregister durchführt, ist als privatrechtliche Organisation nicht zum Zugriff auf das Unternehmensregister berechtigt. Für ein eigenes "Profiling" fehlen aber schlicht die Ressourcen.

Als Ausweg aus diesem Dilemma wurde der Arbeitsablauf der FuE-Erhebung und dabei vor allem die Zusammenarbeit zwischen Stifterverband und Statistischem Bundesamt verändert. Der Stifterverband liefert jetzt seine Mikrodaten an das Statistische Bundesamt, das dann eine Auswertung auf der Basis des europäischen Unternehmensbegriffs vornimmt. Nach einer Machbarkeitsstudie in den Jahren 2020/2021 ist der neue Ablauf für das Erhebungsjahr 2023 erstmals produktiv geworden.

Der Vortrag berichtet über die wichtigsten Erkenntnisse dieses ersten Jahres.

17 Produktivität von Unternehmen. Ergebnisse der Unternehmensstrukturstatistik

Markus Zimmermann

Statistisches Bundesamt, Deutschland
markus.zimmermann2@destatis.de

SECTION: Economic, Social and Market Statistics (Session 1)

Der Beitrag untersucht anhand der Unternehmensstrukturstatistik die Unterschiede der Arbeitsproduktivität zwischen und innerhalb von Wirtschaftsbereichen. Hierbei wird insbesondere die Rolle des neuen statistischen Einheitenbegriffs analysiert. Während vor dem Berichtsjahr 2018 die Rechtliche Einheit die relevante Darstellungseinheit in der Strukturstatistik war, werden die Ergebnisse seitdem für die Darstellungseinheit Unternehmen nach EU-Definition veröffentlicht. Ein Unternehmen kann aus einer oder mehreren Rechtlichen Einheiten bestehen. Dies hat Auswirkungen auf die Messung von Unternehmenskennzahlen wie Vorleistungsquoten, Wertschöpfung oder Arbeitsproduktivität. Ein weiterer Teil der Analyse betrifft die Streuung der Produktivität zwischen und innerhalb von Wirtschaftsbereichen sowie die Frage, ob eine Reallokation von Ressourcen hin zu produktiveren Unternehmen zu beobachten ist.

18 Umstellung des statistischen Unternehmensregisters auf die WZ 2025

Johanna Jung

Statistisches Bundesamt (Destatis), Deutschland
Johanna.Jung@destatis.de

SECTION: Economic, Social and Market Statistics (Session 1)

Das Statistische Unternehmensregister (URS) stellt nach EU-Vorgaben bis Ende 2025 auf die Wirtschaftszweigklassifikation (WZ) 2025 um. Nach der maschinellen Umstellung der Datenbank erfolgt die Nachqualifizierung der WZ 2025 in mehreren Phasen. Da direkt nach der Umstellung der Datenbank noch keine Trainingsdaten für maschinelles Lernen vorliegen, werden zunächst die Einheiten mit mehrdeutigen WZ-Positionen (1:n oder m:n) mittels einer Schlagwortliste der richtigen WZ 2025 zugeordnet. Zur Erstellung der Schlagwortliste wurden zuerst in den Texten der Rechtsträger- und Niederlassungsbezeichnung sowie die Bezeichnungen in den Meldungen der Bundesagentur für Arbeit die häufigsten Wörter auf tiefster Ebene der WZ 2008 (5-Steller) ausgewertet und (wenn möglich) einer der n möglichen Positionen im WZ 2025 zugeordnet. Die Schlagwortliste umfasst also Wörter, die auf jeden Fall vorkommen müssen aber auch Wörter, die auf keinen Fall vorkommen dürfen. Im Text der Rechtsträger- und Niederlassungsbezeichnung

sowie in den Bezeichnungen der Meldungen der Bundesagentur für Arbeit wird nun nach diesen Schlagwörtern gesucht, die eindeutig auf einen der n möglichen Wirtschaftszweige hindeuten. Nur Niederlassungen, die ein Wort oder mehrere Wörter für einen möglichen WZ 2025 beinhalten, können auf diese Weise umgeschlüsselt werden. Wird kein treffendes Wort gefunden oder mehrere Wörter, die auf verschiedene Wirtschaftszweige hindeuten, muss die Nachqualifizierung auf andere Weise erfolgen.

19 Gründungen im Handwerk

Anke Rink

Statistisches Bundesamt, Deutschland
anke.rink@destatis.de

SECTION: Economic, Social and Market Statistics (Session 1)

Die Unternehmensdemografie ermittelt echte Gründungen und Schließungen von Unternehmen. Bei der Berechnung von echten Gründungen von Unternehmen werden demografische Ereignisse wie Abspaltung, Übernahmen, Fusionen und ähnliche Ereignisse berücksichtigt.

Diese Ergebnisse dienen als Grundlage der Ermittlung von echten Gründungen von Handwerksunternehmen. Der Vortrag beleuchtet die Unterschiede zwischen der Gesamtwirtschaft und dem Handwerk und Entwicklungen über mehrere Jahre.

20 'argList': Lazy Argument Management and Propagation Toolkit

Sigbert Klinke, Kleio Chrysopoulou Tseva

Humboldt-Universität zu Berlin, Lehrstuhl für Statistik, Deutschland
sigbert@wiwi.hu-berlin.de

SECTION: Computational Statistics and Data Science (Session 1)

This work-in-progress encompasses a structured and systematic approach to handling complex and nested parameter structures in R, using the 'plot.matrix' function as an example. Internally, plot.matrix decomposes into a series of base R plotting functions ('plot', 'polygon', 'text', 'axis') and exposes a range of configurable arguments. While powerful, this interface can become verbose, rigid and difficult to manage, particularly when working with nested or partially specified parameters. To address this issue, we prototype a more modular interface based on the 'argList' package that captures function arguments as programmable values. This allows us to capture both unevaluated and evaluated function arguments in a programmable form.

Our approach is centered on three complementary strategies: parameter propagation, which expands shorthand inputs into structured lists; global parameter injection, which distributes shared parameters to nested components; and local parameter injection, which supports fine-grained control using a dot-notation syntax (e.g., 'plot1.type' → 'plot1\$type').

These mechanisms are implemented via utility functions ('propagate_param', 'global_param', and 'local_param') that facilitate clear and priority-aware argument resolution. The current implementation is demonstrated through basic examples, which are intentionally minimal to highlight the internal logic rather than practical use cases.

Although still in early stages, this framework is designed for generalization and reuse, particularly in scenarios where function interfaces require flexibility, nesting, or extensibility. The long-term goal is to enable more concise, maintainable, and user-friendly APIs in R, applicable not only in

graphical workflows but in any domain involving complex parameter management. Future work will focus on practical applications and interface validation.

21 Multilingual Monetary Policy: Unfolding Language and Policy Preferences of Swiss Central Bankers

Florian Schütze (1), Sami Diaf (2)

1: Helmut-Schmidt-Universität, Deutschland; 2: University of Hamburg, Deutschland
schuetzf@hsu-hh.de

SECTION: Computational Statistics and Data Science (Session 1)

Understanding monetary policy has always been of paramount economic and political importance. However, it remains a difficult task, despite transparency efforts and the regular flow of information to the public, which becomes even more complex when communication channels are multilingual. This paper examines the policy narratives of the Swiss National Bank (SNB) in terms of language and policy preferences, using the corpus of speeches delivered by its members over the period 1997-2022. Using a dynamic semantic search strategy based on top2vec, the framework analysis was able to identify interlingual similarities and differences with the help of pre-trained multilingual models. The results show that the SNB's communication strategy is strongly oriented towards the objectives assigned to the central bank, with attention being paid to systemic risks, banking regulation and financial markets, which emerge as second but no less important objectives, closely linked to the international environment, in particular the Eurosystem as a strategic aspect of the stability of the Swiss franc. The results suggest that English is used exclusively to address core central banking issues (monetary policy, inflation and interest rates), while uncertainty concerns seem to be reported more in German or French. The resulting dual semantic space, consisting of embedded words and documents, yielded relevant topics with respect to the size and scope of the corpus. Furthermore, informative indices could be constructed for policy measurement, as a crisis index was found to be consistent with the business cycle fluctuations and technical recessions experienced in Switzerland over the last 25 years.

22 Heterogeneity in Voter Movements in Germany – A Mixture Model Approach

Jan Anders, Göran Kauermann

Ludwig-Maximilians-Universität München, Germany
jan.anders@stat.uni-muenchen.de

SECTION: Computational Statistics and Data Science (Session 1)

Understanding how voters transition between parties is central to post-election analysis, particularly amid recent rightward shifts in electoral outcomes. Beyond traditional polling-based methods, voter transition matrices can be estimated from aggregate election results using ecological inference methods. Previous approaches primarily employed hierarchical Bayesian models, while recent work has shifted to constrained optimization techniques. Both approaches perform well but face challenges, notably the underlying assumption that voter transition behaviour is homogeneous across electoral districts. We regard this assumption as critical and propose a more flexible framework for modelling voter transitions.

To relax the homogeneity assumption, we introduce a mixture-model approach, whereby electoral districts are grouped into clusters, that is mixture components, exhibiting similar voter movement patterns. We treat the number of mixture components as a hyperparameter to be determined separately.

Estimation is carried out using a stochastic Expectation-Maximization algorithm, which proves to be numerically flexible. As result we obtain the cluster specific transition matrices and for each voting district a (posterior) distribution for the mixture components. The stochastic nature of the EM algorithm allows to assess the estimation variability of the mixture components. In particular, we can assess the variability of the membership probabilities, which can be used to quantify our confidence in each assignment.

We illustrate our method with an analysis of the 299 electoral districts from the 2025 German Federal Election and present the resulting voter movement estimates and clusters.

23 Estimating Heterogeneous Causal Effects with Tree-Based Methods under Imperfect Compliance and Overlap Violations

Karolina Gliszczynska

Universität Duisburg-Essen, Deutschland
karolina.gliszczynska@uni-due.de

SECTION: Computational Statistics and Data Science (Session 1)

Estimating the Complier Average Causal Effect (CACE) in instrumental variable (IV) settings is critical for uncovering causal relationships, especially when treatment compliance varies across subpopulations. In practice, policy interventions often suffer from imperfect compliance and regions of covariate space where the overlap assumption is violated—challenges that undermine the reliability of standard IV estimation methods. This work extends the Bayesian Additive Regression Trees with Instrumental Variables (BART-IV) framework to a (Transformed) Random Forest-IV setting to improve both flexibility and computational efficiency in CACE estimation.

We introduce a novel methodology that incorporates kernel-based weighting to balance observable covariates between instrument-defined groups, thereby addressing overlap violations and mitigating problems from near-deterministic instrument assignment probabilities. Our contributions are threefold:

1. We reinterpret the BART-IV framework as a general two-step procedure, enabling the use of alternative machine learning models, such as Random Forests, in place of BART for CACE estimation.
2. We develop a Random Forest-IV approach that offers competitive or superior performance relative to BART-IV, particularly when the binary covariate assumption is relaxed.
3. We integrate kernel-based weights into the transformed Random Forest-IV and GRF-IV framework, improving robustness in settings with extreme propensity scores.

Through simulation studies, we demonstrate that our approach maintains high estimation accuracy across varying degrees of treatment effect heterogeneity. The kernel-weighted extension is especially effective in stabilizing estimates when propensity scores approach 0 or 1, conditions that often lead to extreme inverse probability weights.

24 Schätzung von Unsicherheit in Gen AI Systemen mit einem LLM-as-a-Judge Ansatz

Oliver Zeigermann, Gunnar Moys

Techniker Krankenkasse, Deutschland
oliver.zeigermann@tk.de

SECTION: Uncertainty Estimation (Session 2)

Generative KI-Systeme, insbesondere Large Language Models (LLMs), bieten vielfältige Möglichkeiten zur Automatisierung komplexer textbasierter Aufgaben. Ihre inhärente Unsicherheit stellt jedoch eine Herausforderung für den produktiven Einsatz dar – insbesondere in sensiblen Anwendungsfeldern wie dem Gesundheitswesen. In diesem Vortrag zeigen wir am Beispiel der Techniker Krankenkasse, wie mithilfe eines sogenannten LLM-as-a-Judge-Ansatzes die Unsicherheit von KI-generierten Ausgaben abgeschätzt und für Menschen verständlich gemacht werden kann. Dabei beurteilt ein zweites LLM kritisch die Ausgabe eines ersten Modells im Kontext von Eingabe und Aufgabe und weist auf potenzielle Probleme, Unstimmigkeiten oder Unsicherheiten hin. Wir teilen unsere praktischen Erfahrungen und beantworten unter anderem folgende Fragen: - Wie entwickelt man ein verlässliches LLM-as-a-Judge-System? - Wie lässt sich Unsicherheit für Anwender:innen verständlich ausdrücken? - Wann ist der Einsatz solcher Systeme sinnvoll – in der Entwicklungsphase oder im Produktivbetrieb? - Wie kann man die metakognitive Bewertungsaufgabe so gestalten, dass das LLM sie möglichst effizient und treffsicher lösen kann? Darüber hinaus diskutieren wir methodische und ethische Implikationen sowie mögliche Erweiterungen für andere Einsatzgebiete.

25 Statistical Inference for Discrete-Time Multistate Models: Asymptotic Covariance Matrices, Partial Age Ranges, and Group Contrasts

Daniel C. Schneider (1), Mikko Myrskylä (1,2,3), Christian Dudel (1,3,4), Angelo Lorenti (1)

1: Max Planck Institute for Demographic Research, Germany; 2: Helsinki Institute for Demography and Population Health, University of Helsinki; 3: Max Planck-University of Helsinki Center for Social Inequalities in Population Health; 4: Federal Institute for Population Research, Germany
schneider@demogr.mpg.de

SECTION: Uncertainty Estimation (Session 2)

Multistate life tables (MSLTs), or multistate survival models, have become a widely used analytical framework among medical researchers, epidemiologists, social scientists, and demographers. MSLTs can be cast in continuous time or discrete time. While the choice between the two approaches depends on the concrete research question and available data, discrete-time models have a number of appealing features: They are easy to apply; the computational cost for point estimates is typically low; and today's empirical studies are frequently based on regularly spaced longitudinal data, which naturally suggests modelling in discrete time. Up to now, explicit formulas for the covariance matrices of the outcome statistics of discrete-time multistate models (DTMS) have only been developed to a limited extent, which is why many research papers have to resort to costly bootstrap procedures.

This presentation lays out several new asymptotic inference results for DTMS, which substantially cut the computational burden and open new possibilities for the combination and presentation of model outcome statistics. First, we derive asymptotic covariance matrices for the outcome statistics of conditional and/or state expectancies, mean age at first entry, and lifetime risk. We then discuss group comparisons of these outcome measures, which require the calculation of a joint covariance matrix of two or more results. Finally, new procedures are presented for the estimation of multistate models over a partial age range, and how these subrange calculations relate to the result that is obtained from the full age range of the model. All newly derived expressions are checked against bootstrap results in order to verify correctness of results and to assess performance.

26 Statistical Inference for Discrete-Time Multistate Models: Extensions to Markov Chains with Rewards

Daniel C. Schneider (1), Mikko Myrskylä (1,2,3), Christian Dudel (1,3,4), Angelo Lorenti (1)

1: Max Planck Institute for Demographic Research, Germany; 2: Helsinki Institute for Demography and Population Health, University of Helsinki; 3: Max Planck-University of Helsinki Center for Social Inequalities in Population Health; 4: Federal Institute for Population Research, Germany
schneider@demogr.mpg.de

SECTION: Uncertainty Estimation (Session 2)

Discrete-time multistate models (DTMS) have become a widely used analytical framework among epidemiologists, social scientists, and demographers. Markov Chains with rewards (MCWR) have been shown to be a useful modelling extension to discrete-time multistate models. In this paper, we substantially improve and extend the possibilities that MCWR holds for DTMS. We make several contributions. First, we develop a system of creating and naming different rewards schemes, so-called "standard rewards". While some of these schemes are of interest in their own right, several new possibilities emerge when dividing one rewards result by another, the result of which we call "composite rewards". In total, we can define at least ten new useful outcome statistics based on MCWR that have not yet been used in the literature. Secondly, we derive expressions for asymptotic covariance matrices that are applicable for any standard rewards definition. Thirdly, we show how joint covariance matrices of any number of rewards results can be obtained, which leads to expressions for the joint covariance matrices of (any number of) composite rewards. Lastly, expressions for point estimates and covariance matrices of partial age ranges are derived. We confirm correctness of results by comparisons to simulation-based results (point estimates) and by comparisons to bootstrap-based results (covariance matrices).

27 Inference with Non-Probability Samples: A Mixture-Modell Approach

Bernhard Stefan Zins

Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur für Arbeit (BA), Germany
stefan.zins@iab.de

SECTION: Uncertainty Estimation (Session 2)

Non-probability sampling remains a popular method for collecting survey data at low cost. However, statistical inference based on non-probability relies heavily on assumptions about the unknown sampling design. A common way to address this problem is to pair the non-probability sample with a probability sample, which allows for unbiased point and variance estimation.

We propose to improve the precision of estimates based on a probability sample by borrowing strength from a larger non-probability sample.

The goal is to estimate a regression model that holds for the population of interest, which can be tested using the probability sample.

The basic assumption is that the observations in the non-probability sample come from a finite mixture, some of which are drawn from the distribution in the population of interest, that is, for them the regression model holds, for others a regression model with possibly quite different parameters holds. One can borrow strength from the non-probability sample by identifying which observations can be considered as observations from the regression model that is being investigated. The essential requirement is that a probability sample is available. Only then is it possible to evaluate whether or not observations from the non-probability sample can contribute.

An EM algorithm is used to estimate the propensities that an element in the non-probability sample is part of the target population. The method is applied to the pooled sample, where we know with certainty that the elements from the probability sample are part of the target population.

28 Shift-Dispersion Decompositions of Wasserstein and Cramér Distances

Johannes Resin (1), Daniel Wolffram (2), Johannes Bracher (2), Timo Dimitriadis (3)

1: Universität Heidelberg; 2: Karlsruher Institut für Technologie; 3: Goethe-Universität Frankfurt
resin@econ.uni-frankfurt.de

SECTION: Nonparametric and Robust Statistics (Session 2)

Divergence functions are measures of distance or dissimilarity between probability distributions that serve various purposes in statistics and applications. We propose decompositions of Wasserstein and Cramér distances—which compare two distributions by integrating over their differences in distribution or quantile functions—into directed shift and dispersion components. These components are obtained by dividing the differences between the quantile functions into contributions arising from shift and dispersion, respectively. Our decompositions add information on how the distributions differ in a condensed form and consequently enhance the interpretability of the underlying divergences. We show that our decompositions satisfy a number of natural properties and are unique in doing so in location-scale families. The decompositions allow us to derive sensitivities of the divergence measures to changes in location and dispersion, and they give rise to weak stochastic order relations that are linked to the usual stochastic and the dispersive order. Our theoretical developments are

illustrated in two applications, where we focus on forecast evaluation of temperature extremes and on the design of probabilistic surveys in economics.

29 Convergence clubs in the European Union

Joachim Schnurbus (1), Harry Haupt (1), Willi Semmler (2)

1: Universität Passau, Deutschland; 2: New School for Social Research, New York
joachim.schnurbus@uni-passau.de

SECTION: Nonparametric and Robust Statistics (Session 2)

The convergence of economic growth is analyzed for regions of the European Union and the Eurozone. Using recent approaches to data-driven identification of countries' club membership, considerable variation is found in club composition across countries and time, and a nonparametric approach is used to analyze and explain this variation.

30 Wissen. Testen. Verändern. Data & AI Literacy als Hebel für Innovation in der Verwaltung

Sascha Verhoeven

Stadt Heidelberg, Deutschland
sascha.verhoeven@heidelberg.de

SECTION: Competence Development: Data Literacy and Statistics (Session 2)

Verwaltungen stehen zunehmend unter Druck: Mehr Aufgaben, weniger Personal, enge Budgets und hohe Erwartungen an digitale Lösungen. Als Antwort darauf hat die Stadt Heidelberg mit dem Team „Innovation und Wissenschaftskooperationen“ eine zentrale Einheit geschaffen, die Innovationspotenziale frühzeitig identifiziert, Akteure aus Verwaltung, Wissenschaft und Wirtschaft zusammenbringt und Ideen gezielt in konkrete Projekte überführt. Ziel des Teams ist es, den praktischen Einsatz von Daten und KI zu ermöglichen, Freiräume für offenes Experimentieren zu schaffen und eine Kultur zu fördern, die auf Kooperation, Prototyping und gemeinsames Lernen setzt.

Innovation wird dabei nicht abstrakt gedacht, sondern konkret umgesetzt. Programme wie „Test in Heidelberg“ bieten reale Testumgebungen, in denen Daten gesammelt, Ideen getestet und KI-gestützte Lösungen in enger Zusammenarbeit mit Start-ups, Unternehmen und Forschungseinrichtungen entwickelt werden. Parallel entstehen innerhalb der Verwaltung Pilotprojekte, etwa zur Nutzung von Sprachmodellen, zur Automatisierung von Routineprozessen oder zur datengestützten Priorisierung von Aufgaben.

Ein zentrales Element bildet dabei eine gelebte interne Lernkultur. Regelmäßige Skill-Sharings, ein intelligenter Matching-Bot und der jährlich wachsende Innovationstag machen Wissen und Kompetenzen sichtbar und fördern den Austausch. Zuletzt nahmen über 500 Mitarbeitende teil und zeigten damit eindrucksvoll, welches Innovationspotenzial bereits in der Verwaltung vorhanden ist.

Ein weiteres Alleinstellungsmerkmal ist die enge Vernetzung mit der Wissenschaft. Gemeinsam mit Hochschulen entstehen praxisnahe Fallstudien auf Basis städtischer Daten, offene Daten werden in Lehrformate integriert. Die Verwaltung wird so aktiv zum Forschungsgegenstand und zum Forschungspartner und gewinnt wertvolle Erkenntnisse für ihre eigene Weiterentwicklung.

Der Vortrag zeigt, wie eine daten- und KI-gestützte Transformation erfolgreich umgesetzt werden kann: mit klarer Zielorientierung, Offenheit für Experimente und einer Kultur, die konsequent auf Lernen und Zusammenarbeit setzt. Ein zentraler Hebel auf diesem Weg ist das kontinuierliche Erschließen und Verstehen von Daten. Ihre systematische Nutzung wird weiter ausgebaut, um datenbasierte Entscheidungen in der Verwaltung gezielt zu verbessern.

31 SchuldenMonitor NRW: Struktur und Entwicklung kommunaler Schulden

Sebastian Waßenberg, Jochen Kirstein, Anna-Lena Dr. Wagner, Leontine Dr. von Kulmiz

Information und Technik Nordrhein-Westfalen - Statistisches Landesamt
sebastian.wassenberg@it.nrw.de

SECTION: Competence Development: Data Literacy and Statistics (Session 2)

Die Verschuldungssituation einer Kommune lässt sich nicht mit einer einzelnen Kennzahl erfassen – sie erschließt sich erst durch das Verstehen, Einordnen und Vergleichen verschiedener Daten. Genau hier setzt der SchuldenMonitor NRW an: Die interaktive Anwendung unterstützt Nutzende dabei, zentrale Informationen zur kommunalen Verschuldung einzusehen und einzuordnen. Sie soll das Verständnis der Daten fördern, indem sie finanzpolitische Zusammenhänge verständlich visualisiert, Orientierung bietet und den Zugang zu kommunalen Finanzdaten erleichtert.

Die Höhe, Struktur und Entwicklung kommunaler Schulden sind wichtige Indikatoren zur Beurteilung der kommunalen Haushaltslage und Handlungsfähigkeit. Sie ermöglichen sowohl eine detaillierte Analyse der Situation einer einzelnen Kommune als auch fundierte Vergleiche zwischen mehreren Kommunen. So kann es beispielsweise vorkommen, dass zwei Kommunen eine gleich hohe Pro-Kopf-Verschuldung von 1.500 Euro aufweisen, jedoch Kredite zu unterschiedlichen Zwecken aufgenommen haben. Während bei Kommune A 90 % der Schulden auf Liquiditätskredite zur kurzfristigen Überbrückung finanzieller Engpässe entfallen, bestehen bei Kommune B 90 % aus Investitionskrediten zur Finanzierung langfristiger Projekte wie den Infrastrukturausbau. Trotz gleicher Pro-Kopf-Verschuldung liegt es nahe, dass sich die finanziellen Spielräume und folglich auch die Handlungsfähigkeit der beiden Kommunen deutlich unterscheiden.

Der SchuldenMonitor NRW macht solche Unterschiede sichtbar. Karten bieten einen Überblick über die Höhe der Gesamtverschuldung sowie über die Höhe und den Anteil von Liquiditäts- und Investitionskrediten. Ergänzende Informationen helfen bei der Einordnung, etwa durch Vergleiche mit Einheiten der gleichen Verwaltungsform oder Größenklasse. Auch die zeitliche Entwicklung der Verschuldung wird stets zusammen mit entsprechenden Referenzwerten dargestellt. Neben der Analyse einzelner Kommunen oder Kreise sind darüber hinaus auch individuelle Vergleiche mehrerer Gebietseinheiten möglich.

Der SchuldenMonitor NRW macht zentrale Kennzahlen der kommunalen Verschuldung intuitiv erfassbar, schafft Transparenz und ermöglicht einen detaillierten Blick auf langfristige finanzielle Entwicklungen. Die bereitgestellten Visualisierungen und Einordnungen sollen politische Entscheidungsprozesse unterstützen und das öffentliche Verständnis über kommunale Finanzen fördern.

32 From Mobile Phone Data to Mobility Patterns

Victor Tuekam (1,2), Sebastian Wichert (2), Oliver Falck (2), Göran Kauermann (1)

1: LMU Munich, Germany; 2: ifo Institute
tuekam@lmu.de

SECTION: Competence Development: Data Literacy and Statistics (Session 2)

This study demonstrates how large high-frequency mobile phone data can be transformed into flow networks to analyze human mobility patterns. Using anonymized data provided by a large mobile network provider in Germany, we utilize digital traces generated through user interactions with mobile networks, such as calls, messages, and automatic connectivity updates, to infer movements between locations. The spatial resolution is thereby recorded at the level of cell tower locations, while mobility patterns are of interest at the level of administrative districts (e.g., postcode areas). We demonstrate a large-scale data analysis that is still privacy-preserving, examining population mobility without relying on GPS or user-consented tracking.

Our approach focuses on two main tasks: preprocessing the raw mobile phone records and conceptualizing them as network representations of spatial flows. Since the data reflect connections to cell towers rather than precise locations, the spatial resolution is defined by the geographic positions of the towers and resulting individual trajectories, which themselves are imprecise. We address challenges related to trajectory reconstruction and data sparsity by applying a scalable, spatio-temporal linear interpolation technique to estimate user movement between towers. These estimated flows are then aggregated to the postcode level to create interpretable mobility networks. This work contributes a practical method for transforming infrastructure-based mobile phone data into meaningful representations of urban and regional mobility. By emphasizing data preprocessing and network construction—while avoiding user-level tracking—the study supports the development of scalable, privacy-conscious methods for analyzing large-scale human movement patterns. For example, we demonstrate movement patterns in Munich and apply a dynamic gravity model to explain them. We also tackle the question of obtaining information on the mode of transportation.

33 Wie reagieren Betriebe des Bauhauptgewerbes auf die Krise im Wohnungsbau?

Eva Gruner, Carsten Schumann

Statistisches Bundesamt, Deutschland
eva.gruner@destatis.de

SECTION: Economic, Social and Market Statistics (Session 2)

Seit 2022 ist der deutsche Wohnungsbau zunehmend in eine Krise geraten. Das Team der Konjunkturstatistik im Baugewerbe hat eine Sonderauswertung der Einzeldaten aus dem Monatsbericht im Bauhauptgewerbe erstellt, um herauszufinden, ob die Betriebe mit einem langjährigen Schwerpunkt im Wohnungsbau in der Lage waren, ihre Produktionskapazitäten zu erhalten und in anderen Bauarten (gewerblicher Bau oder Tiefbau) einzusetzen. Untersucht wurden also eventuelle Verlagerungseffekte auf Ebene der einzelnen Betriebe aus dem Wohnungsbau in die anderen Bauarten.

34 Neue Konjunkturperspektiven durch Firm-Level Mikrodaten der Industrie: Innovationspotenziale für die volkswirtschaftliche Analyse

Bayram Oruc

Statistisches Bundesamt, Deutschland
Bayram.Oruc@destatis.de

SECTION: Economic, Social and Market Statistics (Session 2)

Die Konjunkturanalyse stützt sich traditionell auf aggregierte Makroindikatoren, die wesentliche wirtschaftliche Entwicklungen abbilden. Durch den Einsatz amtlicher Mikrodaten auf Unternehmensebene aus dem Verarbeitenden Gewerbe Deutschlands eröffnen sich neue Möglichkeiten, die Konjunkturberichterstattung differenzierter zu gestalten. Diese Arbeit beschäftigt sich mit der Entwicklung innovativer Indizes, die unter anderem Unternehmensgrößen, Performance-Merkmale und Marktdynamiken berücksichtigen. Zu den Beispielen gehören der Mittelstandsindex, der Big Player Index und der Top-Performance Index, die eine detaillierte Betrachtung wirtschaftlicher Entwicklungen jenseits aggregierter Kennzahlen ermöglichen. Darüber hinaus wird untersucht, wie auf Basis von Mikrodaten ein Industrie-Geschäftsklima-Indikator entwickelt werden kann, der Umsatz-, Produktions- und Auftragsdaten kombiniert, um frühzeitige Signale zur wirtschaftlichen Entwicklung zu liefern. Ein weiterer Schwerpunkt liegt auf Verteilungs- und Strukturanalysen, die tiefere Einblicke in die Entwicklung von Entgelten und Produktivität sowie in die Marktkonzentration, differenziert nach Wirtschaftszweigen, ermöglichen.

35 t+20 - Ein Projekt zur Schnellschätzung von Konjunkturindikatoren

Edesa Yadegar

Statistisches Bundesamt, Deutschland
edesa.yadegarshirabad@destatis.de

SECTION: Economic, Social and Market Statistics (Session 2)

Das Projekt t+20 des Statistischen Bundesamtes entwickelt Methoden zur beschleunigten Bereitstellung von Konjunkturindikatoren für das Verarbeitende Gewerbe. Ziel ist es, bereits 15 bis 20 Tage nach Monatsende verlässliche Schätzungen zu ermöglichen. Im Rahmen des Projekts wurden verschiedene Verfahren getestet, darunter auch maschinelle Lernmethoden wie Regressionsbäume und Nearest-Neighbour-Ansätze. Aktuell basiert das Verfahren auf einem Imputationsverfahren mittels linearer Regression, da dieses sich als besonders robust erwiesen hat. Nach rund anderthalb Jahren „Echtzeitbetrieb“ zeigen die Ergebnisse für den Umsatz im Verarbeitenden Gewerbe eine sehr gute und robuste Qualität. Die mittlere absolute prozentuale Abweichung liegt dabei aktuell bei 1,1%. Anfang des Jahres wurden die Daten daher erstmals als experimentelle Statistik veröffentlicht. Angesichts der steigenden Nachfrage nach schnellen und präzisen Wirtschaftsdaten – verstärkt durch die Corona-Pandemie und aktuelle wirtschaftliche Herausforderungen – bietet das Projekt eine wertvolle Entscheidungsgrundlage für Politik und Wirtschaft

36 Handlungsbedarf in der Bautätigkeitsstatistik

Alexander Weigert, Carsten Schumann, Marianne Schepers

Statistisches Bundesamt, Deutschland
alexander.weigert@destatis.de

SECTION: Economic, Social and Market Statistics (Session 2)

Die Bundesregierung der 20. Wahlperiode aus SPD, Grünen und FDP (Ampel-Koalition) hatte im Herbst 2023 einen Gesetzesentwurf zur Novellierung des Hochbaustatistikgesetzes (HBauStatG) vorgelegt. Die Novelle sollte Datenlücken schließen, insbesondere bei Baubeginnen und bei Programmen zur sozialen Wohnraumförderung. Im weiteren Verlauf wurde das Gesetz vom Deutschen Bundestag um ein Merkmal zur Barrierefreiheit von Wohnungen sowie präzisere Vorgaben zur digitalen Datenübermittlung ergänzt und in zweiter und dritter Lesung im Bundestag beschlossen. Allerdings rief der Bundesrat in seiner zweiten Lesung im Juli 2024 den Vermittlungsausschuss an. Bis zum vorzeitigen Ende der Legislaturperiode wurde eben jener nicht mehr einberufen. Damit war die Gesetzgebung zur Novellierung des HBauStatG gescheitert. In diesem Vortrag wird erörtert, was dieses Scheitern für die Bautätigkeitsstatistiken bedeutet und welche Schritte nun notwendig sind. Die Bautätigkeitsstatistik steht unter akutem Handlungsdruck. Qualitätsanalysen verdeutlichen den Innovationsstau im Datenfluss und legen hohes Verbesserungspotenzial bei der Aktualität und Aussagekraft der veröffentlichten Statistiken offen. Während langfristig die Digitalisierung der Datenflüsse anzustreben ist, muss mittelfristig das Veröffentlichungskonzept modernisiert werden, um weiterhin qualitativ hochwertige Ergebnisse über die konjunkturelle der Baubranche und die Entwicklung des Wohnungsbaus in Deutschland zu erhalten.

37 Shrinkage Bayesian Causal Forest with Instrumental Variable

Jens Klenke (1), Lennard Maßmann (1,2)

1: Universität Duisburg-Essen, Deutschland; 2: Ruhr Graduate School in Economics, Essen, Deutschland
jens.klenke@vwl.uni-due.de

SECTION: Computational Statistics and Data Science (Session 2)

This paper proposes a novel framework for estimating heterogeneous treatment effects using Instrumental Variables (IV) in observational studies with sparse data and imperfect compliance. To address these limitations, we build upon the Bayesian Instrumental Variable Causal Forest (BCF-IV) framework that has been developed to estimate the conditional Complier Average Causal Effect (CACE) non-parametrically while retaining interpretability. BCF-IV uses Bayesian Additive Regression Trees (BART) to identify treatment effect heterogeneity and to estimate the conditional CACE based on the conditional Intention-To-Treat (ITT) effects and the proportion of compliers. Our approach extends BCF-IV by proposing a Shrinkage Bayesian Instrumental Variable Causal Forest (SBCF-IV) algorithm. SBCF-IV adopts the SoftBART algorithm and makes two major contributions. First, SBCF-IV implicitly discriminates between relevant and irrelevant covariates when estimating conditional ITT effects and proportions of compliers. Secondly, our approach implements varying posterior splitting probabilities from SoftBART into the discovery of heterogeneous subgroups. These modifications enhance SBCF-IV's ability to handle sparse data and to detect variables that drive the heterogeneity of treatment effects. A simulation study suggests that a more precise estimation of conditional CACE can be achieved while maintaining interpretability,

particularly in scenarios with sparsity, confounding, and nonlinearity. In an empirical application, we revisit the Oregon Health Insurance Experiment to demonstrate the use of SBCF-IV in comparison to BCF-IV and discuss the differences in the estimates for the conditional CACE.

38 Bayesian Causal Forests for Cost-Effectiveness Analysis

Lennard Maßmann (1,3), Eva-Maria Oeß (2)

1: Universität Duisburg-Essen, Deutschland; 2: Universität zu Köln, Deutschland; 3: Ruhr Graduate School in Economics, Essen, Deutschland
lennard.massmann@uni-due.de

SECTION: Computational Statistics and Data Science (Session 2)

We introduce a novel approach that combines Bayesian Causal Forests (BCF) with Cost-Effectiveness Analysis (CEA) to assess effect heterogeneity of a binary treatment under unit-varying costs. Recently, CEA Forests have been proposed to estimate heterogeneous effects in a frequentist setting using Generalized Random Forests (GRF). This approach requires, first, estimating the uplift of the cost and the outcome effect, and, second, differencing the unit-level effects. We translate CEA Forests into a Bayesian framework, which holds promise for improving accuracy, facilitating seamless uncertainty quantification, and more effectively capturing sparsity within the underlying data generating process. A simulation study illustrates guidelines for visual and metric-based CEA using our approach. In comparison to the GRF-based method, we find that our method outperforms the CEA Forests, especially under smoothness in the covariate signal and in small samples.

39 Efficient nonparametric estimation of Markov-switching models

Jan-Ole Koslik

Universität Bielefeld, Deutschland
jan-ole.koslik@uni-bielefeld.de

SECTION: Computational Statistics and Data Science (Session 2)

Markov-switching models are powerful tools that allow capturing complex patterns from time series data driven by latent states. Recent work has highlighted the benefits of estimating components of these models nonparametrically, enhancing their flexibility and reducing biases, which in turn can improve state decoding, forecasting, and overall inference. Formulating such models using penalised splines is straightforward, but practically feasible methods for a data-driven smoothness selection in these models are still lacking. Traditional techniques, such as cross-validation and information criteria-based selection suffer from major drawbacks, most importantly, their reliance on computationally expensive grid search methods, hampering practical usability for Markov-switching models. An alternative approach treats spline coefficients as random effects and employs marginal likelihood maximisation via the TMB R package, avoiding grid search but introducing a computationally demanding nested optimisation problem and potential numerical instability. As an alternative, we propose using the so-called extended Fellner-Schall method for smoothness selection, which leverages the relatively simple structure of penalised splines treated as random effects. This method provides an efficient and general mechanism for smoothness selection, avoiding the need for nested optimisation and higher-order derivatives, improving numerical stability, and

significantly reducing computational costs. Our approach enables the practical estimation of flexible Markov-switching models, even in complex settings.

40 Time series forecasting in SAP using a data-driven semiparametric ARMA model

Li Chen (1,2), Yuanhua Feng (1)

1: Universität Paderborn, Deutschland; 2: MHP Management und IT Beratung GmbH
lichen@campus.uni-paderborn.de

SECTION: Computational Statistics and Data Science (Session 2)

Motivated by more and more semi- or nonparametric models applied in time series forecasting and their demonstrated superior performance in many empirical researches, this paper explores the adoption and integration of a semiparametric ARMA model in an enterprise system landscape. We begin by reviewing basic construction of the semiparametric ARMA model, the iterative plug-in algorithm for estimating the trend component of trend stationary times series, forecast techniques and quality measurements, which were well researched and published with the R package *smoots*. Subsequently, we showcase a novel approach to adopt the semiparametric ARMA model in a forecast application based on SAP Analytics Cloud (SAC), which leverages the platform's strengths in system integrity, state-of-the-art user interface (UI) design as well as seamless connection to a R engine with *smoots* package embedded. The forecast application addresses key challenges in terms of cost efficiency, user experience, and the requirement for in-house statistical or machine learning expertise while adopting such statistical algorithms in enterprise context. Finally, we empirically evaluate the forecast quality of the integrated semiparametric ARMA model using real-world data, demonstrating promising results overall.

41 Explaining the difference between the excess yield on short-term German government bonds and the ECB's deposit facility

Florian Schütze, André Esslinger

Helmut-Schmidt-Universität, Deutschland
schuetzf@hsu-hh.de

SECTION: Empirical Economics and Applied Econometrics (Session 1)

The present study employs nonlinear regression models to analyse the relationship between the difference in yield on short-term German government bonds and the ECB's deposit rate, as well as the deposit rate itself. The aforementioned variables are positively correlated, and the difference can be described by a concave function depending on the deposit rate. To achieve a more comprehensive modelling, a methodology for determining thresholds is introduced, and additional economic indicators are considered as explanatory variables. In conclusion, it has been determined that the incorporation of an inflation limit, situated in close proximity to the target value as stipulated by the ECB, in conjunction with the volatility exhibited by the EURIBOR, is the most significant factor contributing to the enhancement of explanatory power. This approach yielded an R-squared of 0.6, without using lagged variables. The exclusion of lagged variables is predicated on the objective of the research, which is to elucidate the factors contributing to excess

returns, as opposed to making predictions about future returns. The utilization of only same-day data was imperative to ensure optimal interpretability.

42 Empirical findings on upper-level aggregation issues in the HICP

Julika Herzberg, Thomas Knetsch, Dilyana Popova, **Jannik Schaller**, Patrick Schwind, Sebastian Weinand

Deutsche Bundesbank, Deutschland
jannik.schaller@bundesbank.de

SECTION: Empirical Economics and Applied Econometrics (Session 1)

We analyse potential mismeasurement of the Harmonised Index of Consumer Prices (HICP) at the upper level of aggregation, focusing on two sources of measurement error: the choice of index formula (representativity component) and the reliability of weights (data vintage component). HICP weights are annually updated based on national accounts data, which at the time of use have preliminary status. The use of final data is expected to yield more reliable weights and, thus, a better estimate of inflation. With national accounts vintage data, we calculate bias and inaccuracy metrics in order to analyse mismeasurement at the upper level of aggregation in the HICPs for Germany, France, Italy, Spain and the Netherlands, as well as for the country group, over the period from 2012 to 2021. For the representativity component, the data availability allows an additional analysis of the period until 2024. Measured in terms of annual HICP rates, the total upper-level aggregation bias falls short of one-tenth of a percentage point. Further, the representativity component, which captures the fact that a Laspeyres-type index such as the HICP suffers from a systematic overestimation of inflation due to the disregard of changes in consumption patterns, and the data vintage component are both found to contribute to overall bias. The contribution of both the representativity and the data vintage components amounts to quite similar shares. However, during the recent high inflation period around 2022 and 2023, the representativity component reveals a significantly higher bias compared to previous years.

43 Order-invariant Identification in a non-linear Structural Vector Autoregression

Linus Nüsing, Benedikt Schwab

Universität Konstanz, Deutschland
linus.nuesing@uni-konstanz.de

SECTION: Empirical Economics and Applied Econometrics (Session 1)

Bayesian Additive Regression Trees (BART) have been shown to be a flexible, non-parametric regression approach that captures non-linear interactions between covariates and response variables. Building on a multivariate extension of the BART framework, we propose a non-linear vector autoregressive model, which we refer to as the Seemingly Unrelated Bayesian Additive Vector Autoregressive Tree (SUBAVART).

Through a comprehensive Monte Carlo study, we demonstrate that the estimated generalized impulse responses converge to their underlying true values for linear as well as non-linear data generating processes. This indicates that the SUBAVART correctly recovers both the dynamic structure and the error covariance matrix of the true model, highlighting its flexibility with respect to the complexity of the data.

To induce sparsity in the model, we incorporate a Dirichlet prior over the splitting variables, which effectively shrinks the predictor space by selecting only the most relevant (lagged) variables for building the trees.

The multivariate extension enables us to incorporate various (structural) identification methods beyond the order-dependent recursive Cholesky decomposition to identify macroeconomic shocks, including for instance identification based on external instruments.

We illustrate the usefulness of the model in an empirical application on monetary policy shocks.

44 Modernisierung der PL-Prozesse in der amtlichen Statistik in Deutschland

Arijana Amina Ramic, Thorsten Konietzko, Shanna Appelhanz

Statistisches Bundesamt, Deutschland
thorsten.konietzko@destatis.de

SECTION: Methodology of Statistical Surveys (Session 1)

Die Plausibilisierung (PL) von Erhebungsdaten spielt eine zentrale Rolle im Statistikproduktionsprozess der amtlichen Statistik in Deutschland. Zur Verbesserung der Ergebnisqualität wie auch zur Reduzierung des bestehenden Konfliktes zwischen großem Ressourcen- und Zeiteinsatz einerseits und Aktualität der Ergebnisse andererseits arbeitet die amtliche Statistik in Deutschland an der Modernisierung, Optimierung und Standardisierung der bestehenden Prozesse und verwendeten IT-Produkte für die Durchführung der PL.

Die Modernisierung des PL-Prozesses umfasst daher drei Stränge:

(1) Anhand der in der „Landkarte der Standardaufbereitungsprozesse“ dokumentierten typischen statistikübergreifenden Funktionalitäten, fachlichen Anforderungen, Arbeitsschritten und Anwendungsszenarien der Datenaufbereitungsphase des Statistikerstellungsprozesses (inkl. der PL) wird eine auf State-of-the-Art-Technologien und geltenden Standards basierende IT-Unterstützung mit dem Fokus auf Effizienz von Backendprozessen konzipiert und umgesetzt. Dabei wird ein hoher Automatisierungsgrad bei der PL-Durchführung angestrebt und auch ein ML/KI-Einsatz soll ermöglicht werden.

(2) Um die Anzahl fehlerhafter Datensätze zu reduzieren, sollen möglichst viele PL-Prüfungen bereits beim Empfang der Erhebungsdaten in den angeschlossenen IT-Meldesystemen und entsprechende Korrekturen (automatisch oder manuell durch die Auskunftspflichtigen) durchgeführt werden.

(3) Mittels der Geschäfts- und Unternehmensarchitektur und den dort genutzten Methoden und BPMN 2.0-, ArchiMate- oder Entity-Relationship-Modellen sowie der „Landkarte der Standardaufbereitungsprozesse“ sollen die bestehenden PL-Prozesse und Zusammenhänge analysiert und – darauf basierend – optimierte Soll-Prozesse und zugehörige Bebauungspläne für die IT-Landschaft erstellt werden.

Die amtliche Statistik in Deutschland hat sich mit dem Modernisierungsvorhaben zum Ziel gesetzt, den Prozess der Plausibilisierung mit seinen vier Grundfunktionen – Fehlererkennung, Fehlersteuerung, Fehlerbeseitigung und Überprüfung der Fehlerbeseitigung – zukunftsorientiert und zukunftsfähig zu gestalten.

Schlüsselwörter: Plausibilisierungsprozesse, Optimierung und Standardisierung, Geschäftsprozessmodellierung, IT-Landschaft

45 Plausibilisierung in der amtlichen Statistik: Vergangenheit, Gegenwart und eine mögliche Zukunft

Katja-Verena Bürk, Florian Dumpert

Statistisches Bundesamt, Deutschland
irina.simankova@destatis.de

SECTION: Methodology of Statistical Surveys (Session 1)

Die Prüfung der Datenplausibilität ist ein wichtiger Bestandteil der Datenaufbereitung in der amtlichen Statistik und spielt eine entscheidende Rolle, um Qualität und Verlässlichkeit statistischer Veröffentlichungen sicherzustellen. Wesentliches Ziel der Plausibilisierung ist es, im Zuge der Datenerhebung entstandene fehlerhafte oder unplausible Werte zu identifizieren und zu lokalisieren, um diese anschließend korrigieren zu können.

In der bisherigen Praxis ist die Plausibilisierung häufig von aufwändigen, manuell durchgeführten Arbeitsschritten abhängig – ein zeit- und ressourcenintensives Vorgehen. Um Effizienz und Skalierbarkeit zu erhöhen, erproben statistische Ämter zunehmend stärker automatisierte Vorgehensweisen – etwa auf Basis deterministischer Regeln, statistischer Modelle sowie Methoden der Künstlichen Intelligenz und des maschinellen Lernens. Diese datengetriebenen Ansätze bieten großes Potenzial zur Verbesserung und Optimierung der Plausibilisierungsprozesse, sind jedoch häufig auch mit nicht zu unterschätzenden technischen und organisatorischen Herausforderungen verbunden.

Der Vortrag bietet einen Überblick über den Entwicklungsstand der Plausibilisierung in der amtlichen Statistik und skizziert Perspektiven für ihre zukünftige Weiterentwicklung.

46 Some recent work on automatic editing of business survey data at Statistics Netherlands

Sander Scholtus

Statistics Netherlands, Netherlands
s.scholtus@cbs.nl

SECTION: Methodology of Statistical Surveys (Session 1)

Traditionally, national statistical institutes and other producers of official statistics have spent much time and effort on manual editing of business survey data to correct measurement errors. Over the past decades, this editing process has been made more efficient by, first, focusing manual editing efforts on the most influential errors and, second, developing automatic editing methods to handle the remaining errors.

A common approach for automatic editing is based on the paradigm of Fellegi and Holt. Given a set of rules that should be satisfied by the data (e.g., "profit equals total revenues minus total costs"), this paradigm states that the observed data should be made consistent with these rules by changing the fewest (weighted) values. Each value can be given a so-called reliability weight, to indicate that some observed values are more likely to be erroneous than others. The paradigm can be formulated as a mixed-integer linear optimization problem, known as the error localization problem. In practice, error localization under the Fellegi-Holt paradigm often has only limited success in mimicking the strategies of manual editors.

A few years ago, I proposed a generalization of the Fellegi-Holt paradigm, which allows edit operations to be used that correct errors that affect more than one variable at the same time. (For

instance: an error where two values on a survey form are interchanged.) Again, edit operations can be given different reliability weights, to indicate that the error corresponding to an edit operation is more or less likely to occur. Two practical questions when applying this error localization approach are: how to define suitable edit operations and how to choose suitable reliability weights. In this contribution, some recent research at Statistics Netherlands into automatic error localization will be presented, including applications to data of the Dutch structural business statistics.

47 Balance of Payments – Part of the System of Official Statistics

Walter J. Radermacher

LMU München, Deutschland
wjr@outlook.de

SECTION: Methodology of Statistical Surveys: Bundesbank-Session - Measuring International Economics

One result of the Enlightenment was the marriage of the nation state and (official) statistics. The establishment of nation states required reliable statistics, as well as necessitating political decisions within said states. Similarly, the emergence of international (e.g. UN) and supra-national (EU) institutions has been accompanied by the development of statistical infrastructure. Systems of official statistics were designed to measure national economies and economic activities, with the exception of international trade in goods. With the increasing importance of globalisation, these conditions and the requirements for official statistics systems have changed fundamentally. These changes affect all levels (basic statistics, accounts, indicators) and all statistical processes (design, production, communication). They require new methods of cooperation between national statistical institutions and new conventions between users and producers concerning the priorities and quality requirements set for statistical components. Last but not least, these changes also require close and trust-based cooperation between the actors at the national and interior supranational statistical levels. Meeting these challenges requires more than just a technical and methodological approach. Rather, it is necessary to understand and manage the effects of further developments and modernisation as part of the overall system of official statistics.

48 Balance of payments statistics – a building block for national accounts

Stefan Hauf

Statistisches Bundesamt, Deutschland
stefan.hauf@destatis.de

SECTION: Methodology of Statistical Surveys: Bundesbank-Session - Measuring International Economics

The national accounts and balance of payments are both standardised in terms of their methodology in coordinated sets of rules. These are currently the System of National Accounts 2008 for the national accounts, and the Balance of Payments and International Investment Position Manual 2009 for the balance of payments.

In keeping with these methodological guidelines (which, in the case of the national accounts, are fleshed out further in the European System of Accounts 2010 for the European Union), the Federal

Statistical Office and the Deutsche Bundesbank cooperate closely in producing their results based on mutually supplied data.

These concern imports and exports, which are made up of cross-border flows of goods and services. In an open economy, what is known as the external balance (that is to say, exports less imports) is needed to calculate gross domestic product (GDP), the key indicator of an economy's growth. Similarly, gross national income is likewise an important variable for the national accounts, and can only be computed by cooperating with the Bundesbank in the field of the balance of payments. Gross national income is obtained from gross domestic product by adding primary incomes received by resident units from the rest of the world and subtracting primary incomes paid to the rest of the world.

To obtain gross national disposable income from gross national income, secondary incomes received from the rest of the world (taxes, social contributions and benefits, and other transfers) are added and secondary incomes paid to the rest of the world are deducted. Balance of payments statistics provided by the Bundesbank are important inputs for these cross-border transactions as well.

49 The Role of the Balance of Payments for Central Bank Policy

Gerhard Ziebarth

Aktionskreis Stabiles Geld, Germany
gerhard.ziebarth@gmail.com

SECTION: Methodology of Statistical Surveys: Bundesbank-Session - Measuring International Economics

This contribution delves into the intricate relationship between monetary policy and the Balance of Payments in an open market economy. It underscores the indispensable value of BOP as the flagship and conceptual anchor for an entire family of external statistics and as an essential pillar in an integrated system of macroeconomic statistics and international accounting standards.

To be most effective, the monetary objectives and the analytical background have to be clearly defined and comprehensive data sets fit for purpose should be at hand.

Normative issues concerning BOP are put to the test.

A glance into the "engine room of external statistics" reveals important necessary conditions in order to meet the expectation of high-quality and up to date services of the various user groups in a rapidly changing "global village".

50 The Euro Area Balance of Payments and International Investment Position and their Role in ECB Policy Making

Martin Schmitz

European Central Bank, Germany
martin.schmitz@ecb.europa.eu

SECTION: Methodology of Statistical Surveys: Bundesbank-Session - Measuring International Economics

This contribution outlines the methods used by the European Central Bank (ECB), in cooperation with the European System of Central Banks (ESCB) to compile statistics on the euro area balance of payments and international investment position (BOP/IIP). It describes the key principles for

compiling these statistics, including the treatment of the euro area as a single economic territory and the focus on transactions and positions with non-euro area residents. It also presents the main data sources, primarily national contributions from euro area countries, emphasizing the importance of data quality and harmonization for accurate euro area aggregates. Additionally, the chapter discusses euro area specific compilation practices. It also highlights the analytical value of BOP/IIP statistics for ECB policymaking, particularly in assessing the euro area's exposure to external shocks and the international impact of ECB policies. The contribution also presents recent advancements in the development of euro area BOP/IIP statistics, including more detailed breakdowns for special-purpose entities and enhanced sectoral and geo-graphic information.

51 Evidence-Based Policymaking in the European Union Using Balance of Payments Statistics

Matthias Ludwig

European Commission - Eurostat, Deutschland
matthias.ludwig@ec.europa.eu

SECTION: Methodology of Statistical Surveys: Bundesbank-Session - Measuring International Economics

This contribution covers the multifaceted role of balance of payments (BOP) statistics in providing a basis for shaping European Union (EU) policies and monitoring their impact. It starts by describing the process of preparing and collecting BOP data, explaining the importance of sound methodology and the efforts taken by Eurostat, the statistical office of the EU, to ensure the high quality of the data. Subsequently, the contribution outlines various uses of BOP statistics, from detecting macroeconomic imbalances to analyzing the potential economic impacts of specific trade and customs policies, from enhancing transparency to strengthening accountability, from monitoring and evaluating, to informing the public. The contribution also details the specific case of the accounts of the EU institutions in BOP.

52 Evidence-Based Policies and Data Quality – What is Missing?

Ralf Münnich

Universität Trier, Deutschland
muennich@uni-trier.de

SECTION: Methodology of Statistical Surveys: Bundesbank-Session - Measuring International Economics

Official statistics institutions aim to collect data of high quality. In general, guidelines are available that help in achieving a high quality of data and, as a result, a high quality of statistics. However, in times of rising rates of non-response, alternative data collection methods come into play. This raises the question of whether statistics drawn from data collected via alternative methods still fulfil sufficient standards of quality. A related question concerns quality standards that are to be measured with regard to specific content. The aim of this paper is to provide an overview of modern challenges in the use of data while taking consideration of high quality standards. This is especially important in the context of data used as evidence for policy purposes or legislative processes. Special emphasis is placed on raising awareness of quality considerations as well as data and statistical literacy and its integration into education and life-long learning.

53 Ökonomischer Status und Pandemie-Effekte. Eine Analyse der Daten der Bürgerbefragung "Leben in Koblenz" 2023

Daniela Schüller

Stadtverwaltung Koblenz, Deutschland
Daniela.Schueller@stadt.koblenz.de

SECTION: VDSSt (Session 1): Umwelt-, Klima- und Gesundheitsmonitoring in Städten und Regionen

Nach dem offiziellen Ende der Corona-Pandemie haben manche Menschen weiterhin mit deren Folgen zu kämpfen. Obwohl es keinen Einfluss haben sollte, ob jemand reich oder arm ist, kann vermutet werden, dass wohlhabendere Personen besser durch die Pandemie gekommen sind. Dazu werden die Folgen „gesundheitliche Beeinträchtigungen“, „wirtschaftliche/finanzielle Einbußen“ und „soziale Isoliertheit“ für verschieden ökonomisch ausgestattete Gruppen und Lebensstiltypen näher betrachtet.

Der Artikel zeigt anhand ausgewählter Items der Koblenzer Mehrthemen-Bürgerbefragung selbstberichtete Auswirkungen der Corona-Pandemie auf die unterschiedlichen ökonomisch-Statusgruppen.

54 Sozialräumliche Unterschiede und individuelles Gesundheitsverhalten in deutschen Großstädten. Eine Analyse auf Basis der Innerstädtischen Raubeobachtung und des Sozio-oekonomischen Panels

Judith Kaschowitz, Dorothee Winkler

BBSR, Deutschland
judith.kaschowitz@bbr.bund.de

SECTION: VDSSt (Session 1): Umwelt-, Klima- und Gesundheitsmonitoring in Städten und Regionen

Bewohnerinnen und Bewohner in deutschen Großstädten finden unterschiedliche Lebensbedingungen vor in Bezug auf städtebauliche Strukturen, soziale Infrastruktur oder Belastungen durch Lärm und Hitze. Weiter gibt es bestimmte Stadtteile in denen armutsgefährdete Gruppen eher segregiert und unter vergleichsweise schwierigen Lebensbedingungen in Bezug auf Wohnen, öffentlicher Raum und Umweltbelastungen leben. Solche Unterschiede in den Lebensbedingungen könnten sich auf das Gesundheitsverhalten von Bewohnerinnen und Bewohnern auswirken, da individuelle Gesundheit und Gesundheitsverhalten das Resultat einer Wechselbeziehung von Individuum und (sich verändernden) Lebens- und Umweltbedingungen sein können. Dieser Beitrag knüpft an eine Studie an, in der die Gesundheit von Bewohnerinnen und Bewohner in Abhängigkeit des „sozialen Status“ (SGB-II Quote) des Stadtteils untersucht wurde. Die Ergebnisse unserer Vorläuferstudie haben gezeigt, dass neben individuellen Einflussfaktoren auch die kleinräumige Ebene eine moderierende Rolle für die individuelle Gesundheit spielt. Daran anknüpfend fokussieren wir uns in diesem Beitrag auf den Einfluss kleinräumiger Disparitäten auf das Gesundheitsverhalten. Die sozialräumlichen Unterschiede erfassen wir mit den Daten der Innerstädtischen Raubeobachtung (IRB) über die SGB-II Quote auf Stadtteilebene. Zur Untersuchung des Gesundheitsverhaltens wie körperliche Aktivität, Schlafdauer und Rauchverhalten, nutzen wir die Befragungsdaten des Sozio-oekonomischen Panels (SOEP). Über die Verschneidung der beiden Datenquellen soll der Einfluss des Stadtteils auf gesundheitsbezogene Verhaltensweisen, unter Berücksichtigung individueller sozioökonomischer

und demographischer Merkmale, untersucht werden. Erste deskriptive Ergebnisse zeigen, dass sich das Gesundheitsverhalten von Personen, die in Stadtteilen mit einer hohen SGB-II Quote leben, von dem Gesundheitsverhalten von Personen aus Stadtteilen mit niedriger SGB-II Quote unterscheidet. So sind Personen in Gebieten mit hoher SGB-II Quote seltener körperlich aktiv, rauchen mehr und schlafen weniger als Personen in Gebieten mit niedriger SGB-II Quote. Aufbauend darauf soll multivariat untersucht werden, welchen Einfluss sozialräumliche Unterschiede neben individuellen Merkmalen auf das Gesundheitsverhalten wie z.B. die körperliche Aktivität haben.

55 Pflegestrukturplanung in Bayern: Eine einheitliche Datengrundlage für die Sozialplanung

Stefan Schiele, Rosanna Purrucker

Bayerisches Landesamt für Statistik, Deutschland
Rosanna.Purrucker@statistik.bayern.de

SECTION: VDSt (Session 1): Umwelt-, Klima- und Gesundheitsmonitoring in Städten und Regionen

Eine regionalisierte Pflegestrukturplanung ermöglicht es den Landkreisen und kreisfreien Städten, den Bedarf an pflegerischer Versorgung, der sich durch demografische Entwicklungen ergibt, frühzeitig abschätzen und entsprechend zukunftsorientiert handeln zu können. Angesichts des erwarteten fortschreitenden Anstiegs der Anzahl an Pflegebedürftigen gewinnt dies zunehmend an Bedeutung. Relevante Einflussfaktoren wie die Altersstruktur der Bevölkerung unterscheiden sich regional und müssen individuell berücksichtigt werden. Das Bayerische Staatsministerium für Gesundheit, Pflege und Prävention (StMGP) hat hierfür im Jahr 2021 das Projekt Pflegestrukturplanung initiiert. Dabei unterstützen das Landesamt für Statistik (LfStat) und das Landesamt für Pflege (LfP) die Kommunen durch eine regelmäßig aktualisierte und regionalisierte Datengrundlage.

Das Ziel des Projektteams ist es, für die Sozialplanenden in Bayern die Basis für eine einheitliche Pflegestrukturplanung zu schaffen, indem verlässliche und grundlegende Daten auf regionaler Ebene sowie eine Handlungsleitlinie zur Verfügung gestellt werden. Den Kommunen soll der Einstieg in die eigenständige Durchführung der Pflegebedarfsermittlung erleichtert werden. Durch eine abgestimmte Methodik sollen ein fachlicher Austausch und ein Vergleich der regionalen Planungen ermöglicht werden. Die Bayerische Handlungsleitlinie wurde von Expertinnen- und Experten aus den Kommunen und Bezirken erarbeitet und umfasst neben methodischen Aspekten ein für die Berechnung entwickeltes Basismodell.

Die Berechnung basiert auf altersspezifischen Quoten aus der Pflegestatistik sowie Angaben zur Bevölkerung aus der Bevölkerungsfortschreibung und den Ergebnissen der Bevölkerungsvorausberechnung. Ab 2027 wird das LfStat eine regionalisierte Vorausberechnung des Pflegebedarfs auf Kreisebene zur Verfügung stellen, die verschiedene Aspekte der Pflege berücksichtigt. Dabei wird unter anderem die vorausberechnete Zahl der Pflegebedürftigen nach Pflegegrad und Versorgungsart ausgewiesen. Zusätzlich wird auch die Anzahl an verfügbaren Pflegeplätzen sowie der Bedarf und das Angebot an Pflegekräften betrachtet.

Durch die Bereitstellung der Daten auf regionaler Ebene können die Kommunen frühzeitig auf die zukünftige Entwicklung reagieren. Die regionalisierte Pflegestrukturplanung leistet damit einen wichtigen Beitrag für die Sozialplanung in den Kommunen.

56 Hitzebedingte Mortalität: Möglichkeiten und Grenzen statistischer Modellierung im kommunalen Kontext

Georg Wiegleb

Landeshauptstadt Magdeburg, Deutschland
georg.wiegleb@stat.magdeburg.de

SECTION: VDSSt (Session 1): Umwelt-, Klima- und Gesundheitsmonitoring in Städten und Regionen

Die Untersuchung des Einflusses von Hitze auf Gesundheit und Mortalität gewinnt zunehmend auch im kommunalen Kontext an Bedeutung – etwa im Rahmen stadtbezogener Hitzeaktionspläne. Statistische Auswertungen und modellhafte Analysen rücken damit stärker in den Fokus kommunaler Statistikstellen und städtischer Entscheidungsprozesse. In diesem Vortrag sollen Möglichkeiten zur Untersuchung des Konzepts hitzebedingter Mortalität aufgezeigt werden – unter Verwendung vorhandener Datenquellen wie Temperaturreihen des Deutschen Wetterdienstes, kommunaler Sterbe- und Bevölkerungsstatistiken sowie Statistiken zu Todesursachen. Im Vordergrund stehen Fragen der konzeptionellen Abgrenzung, Datenqualität und methodischen Anforderungen: Wie lässt sich „Hitze“ definieren? Welche Schwellenwerte, Referenzgrößen, Maßzahlen oder Abstraktionen können geeignet sein, um das Konzept zu erfassen? Welche statistischen Verfahren sind grundsätzlich denkbar – und welche Anforderungen stellen sie an Verfügbarkeit, Auflösung und Struktur der Daten? Ziel ist es, die Machbarkeit und Grenzen solcher Analysen im Rahmen kommunaler Statistik realistisch einzuordnen.

57 Zielbild für ein neues System der Unternehmensstatistiken: SysdU

Simon Rommelspacher, Ralf Droßard

Statistisches Bundesamt, Deutschland
simon.rommelspacher@destatis.de

SECTION: Economic, Social and Market Statistics (Session 3)

Das bestehende System der Unternehmensstatistiken ist in seiner heutigen Form, historisch ausgehend von der Beobachtung der Industrieproduktion, stetig gewachsen. Neue Anforderungen und Datenbedarfe wurden erfüllt, indem der Erfassungsbereich der Unternehmensstatistiken immer wieder ergänzt und erweitert wurde, um eine nahezu vollständige Abbildung der Wirtschaft zu ermöglichen. Dies führte teilweise zu einem unterschiedlichen Vorgehen bei der Statistikerstellung in den einzelnen Wirtschaftsbereichen.

Mit dem Zielbild für ein neues System der Unternehmensstatistiken (SysdU) soll ein methodisch ganzheitliches System der Unternehmensstatistiken für den Statistischen Verbund entwickelt und implementiert werden. Das neue System soll Kohärenz gewährleisten und eine Flexibilität sowohl in der Statistikproduktion als auch in der Datennutzung zulassen, da auf neue Datenbedarfe schnell reagiert werden muss. Die Ergebnisse müssen belastbar und regional darstellbar sein.

Das Zielbild und die Potentiale, die sich daraus für eine Entlastung der meldenden Unternehmen und der Verwaltung durch eine effiziente Statistikproduktion ergeben, sollen vorgestellt und diskutiert werden. Abschließend wird ein Ausblick auf einen möglichen Umsetzungsplan gegeben.

58 Berichtspflichten zu amtlichen Wirtschaftsstatistiken aus Sicht eines Statistischen Landesamtes

Wolf-Dietmar Speich

Statistisches Landesamt des Freistaates Sachsen, Deutschland
wolf-dietmar.speich@statistik.sachsen.de

SECTION: Economic, Social and Market Statistics (Session 3)

Das Thema „Berichtspflichten zu amtlichen Wirtschaftsstatistiken“ ist Teil einer intensiv geführten Diskussion bezüglich Bürokratiebelastung der Wirtschaft. Diese statistischen Berichtspflichten fußen auf bundesgesetzlichen bzw. EU-Regelungen. Die dabei ermittelten Daten sind unverzichtbar für die Entscheidungsfindung in Politik, Verwaltung und Wirtschaft.

Obwohl die Gesamtbelastung durch amtliche Statistiken deutlich weniger als 1 % der gesamten Bürokratiekosten der deutschen Wirtschaft in Höhe von 65.000 Mill. EUR (Sachsen [BIP-Anteil von 3,8 %]: 12,3 Mill. EUR bezogen auf 2.470 Mill. EUR) beträgt, sind die Statistikämter kontinuierlich mit diesem Thema beschäftigt. Letztlich ist in Sachsen zwar lediglich ein Fünftel der im Unternehmensregister geführten 155.000 rechtlichen Einheiten berichtspflichtig, doch bindet jede Online-Meldung in den Unternehmen finanzielle und personelle Ressourcen. Bei kleineren Unternehmen ist die damit verbundene relative Belastung hoch.

Das Statistische Landesamt ist daher bestrebt, Wege zu finden, damit Unternehmen weitgehend belastungsarm den in gesetzlichen Regelungen auf EU-Ebene bzw. in Deutschland festgeschriebenen Datenbedarf bedienen können. Nach den bisherigen Erfahrungen ist dies insbesondere durch wertschätzende Kommunikation und über Unterstützungsangebote zu den Statistikmeldungen gegeben. Das betrifft u.a. den Hinweis auf eine automatisierte Datenerhebung aus dem betrieblichen Rechnungswesen mittels des Meldeverfahrens eSTATISTIK.Core. Beispielsweise gehen bei der Verdienststatistik bereits über 80% aller Meldungen so ein. In Zusammenarbeit mit den sächsischen Kammern und weiteren Partnern wird dieses Meldeverfahren gegenwärtig verstärkt beworben und soll daran interessierten Unternehmen in einem Webinar vorgestellt werden.

Unter den heutigen gesetzlichen Rahmenbedingungen stellt diese Initiative einen wichtigen Schritt zur Minimierung der Belastung der Unternehmen bezüglich amtlicher Wirtschaftsstatistiken dar. Perspektivisch sind weitere Entlastungen der Wirtschaft insbesondere durch Schaffung bundesgesetzlicher Regelungen zur Mehrfachnutzung einmal an Verwaltungen gelieferter Angaben (Once-Only-Prinzip) oder durch Nachnutzung von in Unternehmen vorliegenden Datenbeständen (z. B. Geschäftsberichte in elektronischer Form) sowie eigener Datenaufbereitung in den Statistikämtern bei adäquater Ressourcenausstattung vorstellbar.

59 20 Jahre Verwaltungsdatenverwendung – Entwicklungsstand und Perspektiven

Franziska Große (1), Patrizia Mödinger (2)

1: Landesamt für Statistik Niedersachsen, Deutschland; 2: Statistisches Bundesamt, Deutschland
franziska.grosse@statistik.niedersachsen.de

SECTION: Economic, Social and Market Statistics (Session 3)

Die Verwendung von Verwaltungsdaten in den Wirtschaftsstatistiken erfolgt mittlerweile seit über 20 Jahren. Die Verwaltungsdaten bilden nicht nur eine wesentliche Grundlage zum Aufbau des Statistischen Unternehmensregisters, sie führten auch zum vollständigen Ersatz von Großzählungen und zur Reduzierung von Auskunftspflichten im Rahmen der Konjunkturstatistiken.

Die Verwendung von Verwaltungsdaten hat eine Reihe von Vorteilen. Die Daten liegen bereits für bestimmte Zwecke vor und können durch die amtliche Statistik „zweitverwertet“ werden. Oftmals handelt es sich um einen vollständigen Datenbestand, der fortlaufend aktualisiert wird. Dies ermöglicht es der amtlichen Statistik beispielsweise, die Handwerkszählung komplett auf Basis von Verwaltungsdaten zu berechnen und damit zur Entlastung von Berichtspflichten beizutragen. Die Verwaltungsdatennutzung bringt jedoch auch Nachteile mit sich, da diese Daten nicht für statistische Zwecke gemacht wurden. Beispielsweise fehlen Plausibilisierungen oder die Wertmerkmale entsprechen nicht den statistischen Definitionen. Erschwert bzw. verlangsamt wird eine Nutzung potentieller Verwaltungsdaten auch durch den zeitlichen Vorlauf zur Schaffung der rechtlichen Rahmenbedingungen.

Die Erstellung der Wirtschaftsstatistiken ist heutzutage ohne die Nutzung von Verwaltungsdaten gar nicht mehr denkbar. Um das Potenzial der Verwaltungsdaten bestmöglich auszuschöpfen, ist es zwingend erforderlich, die amtliche Statistik von Beginn an beim Aufbau von Registern oder Datenbeständen mit einzubeziehen. Erst dann ist gewährleistet, dass die Verwaltungsdaten die notwendigen Qualitätskriterien (zu Aktualität, Genauigkeit, Konsistenz, Kohärenz) der amtlichen Statistik erfüllen können.

Der Vortrag soll beleuchten, was durch die Verwaltungsdatenverwendung in den Wirtschaftsstatistiken schon erreicht wurde, wo die Vorteile, aber auch die Grenzen der Verwaltungsdaten liegen und was bei der Ausgestaltung neuer Verwaltungsdaten beachtet werden sollte.

60 Identifikation von VAT Tradern in der Außenhandelsstatistik

Thorsten Gores, Hendrik Kruse

Statistisches Bundesamt, Deutschland
thorsten.gores@destatis.de

SECTION: Economic, Social and Market Statistics (Session 3)

VAT-Trader (auch Non-Resident VAT-Trader) sind ausländische Unternehmen, die in einem Land für die Mehrwertsteuer (VAT, Value Added Tax) registriert sind, ohne dort eine physische Präsenz zu haben. Sie verfügen weder über Mitarbeiter noch Betriebsstätten oder Produktionsaktivitäten im Land der steuerlichen Registrierung. Diese Unternehmen führen umsatzsteuerpflichtige Transaktionen durch, wie etwa innergemeinschaftliche Lieferungen und Erwerbe, und benötigen daher eine Umsatzsteuer-Identifikationsnummer (USt-IdNr.) im jeweiligen Land.

Die physischen grenzüberschreitenden Warenbewegungen dieser Unternehmen werden in der Außenhandelsstatistik erfasst. Da jedoch kein wirtschaftlicher Eigentumswechsel zwischen gebietsansässigen und nicht gebietsansässigen Einheiten stattfindet, bleiben diese Warenströme in den Volkswirtschaftlichen Gesamtrechnungen und der Zahlungsbilanzstatistik unberücksichtigt.

Im Jahr 2023 wurde erstmals eine systematische Identifikation von VAT-Tradern in der Außenhandelsstatistik durchgeführt. Grundlage hierfür waren Unternehmensdaten aus dem Außenhandelsregister, dem Unternehmensregister, dem Gemeindeverzeichnis, den ATLAS-Stammdaten sowie Steuerdaten der Oberfinanzdirektionen. Seitdem wird die Liste der VAT-Trader jährlich aktualisiert. Der Vortrag stellt die angewandte Identifikationsmethode vor und gibt Einblicke in die neuesten Erkenntnisse zu VAT-Tradern in Deutschland.

61 Online Monitoring of Spatio-Temporal Data Streams

Peihua Qiu

University of Florida, USA
pqiu@ufl.edu

SECTION: Computational Statistics and Data Science (Session 3)

In applications such as environmental monitoring and disease surveillance, spatial data are often collected sequentially over time, forming spatio-temporal data streams. Effectively monitoring these streams in real time is critical for detecting anomalies in the underlying spatio-temporal processes. Due to the complex nature of spatio-temporal data—including intricate correlation structures, evolving mean patterns, and nonstandard distributions—this remains a challenging research problem. In this talk, we will present several recent methodologies developed by our research team, using spatio-temporal local linear kernel smoothing, exponentially weighted spatial LASSO, and other advanced data smoothing techniques. We will also showcase applications of these methods in analyzing real infectious disease datasets.

62 Exact computation of angular halfspace depth

Rainer Dyckerhoff (1), Stanislav Nagy (2)

1: University of Cologne, Cologne, Germany; 2: Charles University, Prague, Czech Republic
rainer.dyckerhoff@statistik.uni-koeln.de

SECTION: Computational Statistics and Data Science (Session 3)

Much recent research has focused on directional data, i.e., data on the unit sphere. The angular halfspace depth is a tool for nonparametric analysis of directional data. This depth was proposed as early as 1987, but its widespread use has been hampered by significant computational problems. We present an efficient algorithm for the exact computation of the angular halfspace depth in arbitrary dimensions, which does not require the data to be in general position. The algorithm is based on a two-step projection scheme. In the first step, the data are repeatedly projected onto a lower-dimensional sphere. Then, the data are projected from this low-dimensional sphere onto a linear space in which the usual halfspace depth is computed with respect to a signed measure. Compared to known algorithms, this new algorithm is significantly faster. However, the main advantage of the proposed algorithm is that it is able to compute the depth of all data points in a sample (with respect to that sample) with the same time complexity as the depth of a single point. Another important advantage of our algorithm is its good parallelizability.

63 Fast Factor Extraction for Mixed Data Types

Matei Demetrescu, **Fabian Schmidt**

TU Dortmund, Deutschland
f.schmidt@statistik.tu-dortmund.de

SECTION: Computational Statistics and Data Science (Session 3)

Empirical research has access to ever larger data sets as technological advances make it easier and less costly to collect large amounts of data. However, the amount of available data often exceeds the capabilities of the methods ultimately used to answer the research question, meaning that either not all of the data can be used or techniques such as shrinkage or dimensionality reduction have to be used. In the latter case, assuming a latent factor structure in the data is common. One of the most widely used methods for factor extraction is solving a principal component estimation problem for which efficient implementations exist. Although mixed data types may be considered via generalized linear models driven by latent factors, PCA is not available for those. To enable the use of such data, we propose a maximum likelihood-based iterative alternating least squares procedure capable of accommodating mixed data types, and we empirically demonstrate its practical applicability.

64 Imputing Missing Values in Official Statistics: Assessing Imputation Accuracy across Various Imputation Methods.

Florian Dumpert (1), Markus Pauly (2,3), **Maria Thurow** (2,3), Inken Veips (2)

1: Statistisches Bundesamt (Destatis), Germany; 2: TU Dortmund University, Germany; 3: UA Ruhr, Research Center Trustworthy Data Science and Security, Germany
maria.thurow@tu-dortmund.de

SECTION: Methodology of Statistical Surveys (Session 2)

In the process of data preparation, handling missing values is an important element that may affect later research outcomes. In this study, we conduct a comparative simulation study of imputation methods, including widely used methods such as missRanger (Random Forest based imputation) and MICE (Multiple Imputation by Chained Equations) as well as two approaches that were especially established for the use in Official Statistics: CANCEIS (Canadian Census Edit and Imputation System) from Statistics Canada, which includes the option of adding plausibility rules within the imputation and VIM (Visualization and Imputation of Missing Value) from Statistics Austria.

Using the German Structure of Earnings Survey from 2010 to 2018, we show how to assess imputation methods based on their (multivariate) imputation accuracy. Since the term “imputation accuracy” is not uniquely defined in theory and practice, we use different measures in our analyses: Besides often used accuracy measures like the normalized root mean squared error (NRMSE) and the proportion of false classification (PFC), we also focus on distribution distance measures.

Since in Official Statistics data plausibility is another crucial aspect in data preparation, we place an additional focus on this aspect comparing the different imputation methods regarding their ability to impute data fulfilling predefined so-called edit rules.

65 Von der Praxis zur Methodik: Evaluierung von Imputationen in der amtlichen Statistik

Steffen Moritz, Florian Dumpert

Statistisches Bundesamt, Deutschland
steffen.moritz@destatis.de

SECTION: Methodology of Statistical Surveys (Session 2)

Eine hohe Datenqualität ist eine Grundvoraussetzung für verlässliche amtliche Statistiken. Da in Erhebungen regelmäßig fehlerhafte oder fehlende Angaben auftreten, kommt der Imputation eine zentrale Bedeutung zu. Durch geeignete Imputationsverfahren können potenzielle Verzerrungen reduziert und die Aussagekraft statistischer Ergebnisse gesichert werden.

Die Auswahl eines geeigneten Verfahrens gestaltet sich in der Praxis jedoch häufig als äußerst komplex. Der wahre, „fehlende“ Wert ist in der Regel unbekannt und kann somit nicht zur Qualitätsbeurteilung herangezogen werden. Diese fehlende Ground Truth erschwert eine objektive Bewertung der Imputationsgenauigkeit und den methodischen Vergleich verschiedener Verfahren. Es besteht die Gefahr, dass Verfahren bevorzugt werden, die für bestimmte Analyseziele geeignet erscheinen, dabei aber in anderen Kontexten systematische Verzerrungen verursachen.

Derzeit in der Praxis eingesetzte Evaluationsmethoden umfassen u.a. erklärende Grafiken zum Vergleich von Verteilungen, Modellgütemaße sowie Simulationsstudien. Die Auswahl an Metriken reicht von RMSE, MAE und MAPE bis hin zu verschiedensten verteilungsbezogenen Kennzahlen. Jede dieser Methoden und Metriken bringt spezifische Stärken und Schwächen mit sich – ein universell geeignetes Verfahren existiert allerdings nicht. Diese Vielfalt erschwert eine konsistente Anwendung in der Praxis, da ein übergreifender Bewertungsrahmen in der amtlichen Statistik bisher fehlt. Entsprechend erfolgt die Auswahl von Verfahren häufig ad hoc, und verwendete Ansätze variieren sowohl im internationalen Kontext als auch innerhalb einzelner Statistikämter stark.

Der Vortrag diskutiert die Notwendigkeit eines strukturierten und robusten Bewertungsrahmens für Imputationen, der bestehende Methoden systematisiert, konsistente Anwendung unterstützt und die Vergleichbarkeit erhöht. Ziel ist es, Wege aufzuzeigen, wie Transparenz und Reproduzierbarkeit in der Qualitätssicherung fehlender Daten gestärkt werden können – national wie international.

66 Möglichkeiten für ML-basierte Imputationsverfahren für Item-Nonresponse am Beispiel der Statistiken des Verarbeitenden Gewerbes (StatVG)

Elena Stäger, Muhammet Akman, Richard Bündgens, **Christian Borgs**

IT.NRW, Statistisches Landesamt Nordrhein-Westfalen, Deutschland
christian.borgs@it.nrw.de

SECTION: Methodology of Statistical Surveys (Session 2)

Im Rahmen der Statistiken des Verarbeitenden Gewerbes (StatVG) wird der Inland-Umsatz als wichtige Kennzahl momentan bei fehlenden Werten durch den letzten bekannten Wert des entsprechenden Betriebs ersetzt (LOCF (Last Observation Carried Forward) -Imputation), wenn ein Unternehmen keine Meldung vornimmt. Auch bereits erprobt wurde die Imputation von fehlenden Werten mit Canceis, welches auf einem k-NN- Algorithmus basiert.

Ziel des Projektes ist es, sowohl LOCF, wie auch Canceis bei fehlenden Einzelwerten durch eine bessere Methode zu schlagen. Dabei sollen auch maschinelle Lernverfahren getestet werden: Es werden Zeitreihen-Imputationen (imputeTS) mit baumbasierten Verfahren (missRanger missForest) und neuronalen Netzen (LSTM NN) verglichen.

Dabei werden Umsatz-Werte aus dem StatVG-Datensatz zufällig als fehlend gesetzt. So können die Imputationsverfahren mit den bekannten Werten verglichen werden. Das zentrale Ziel ist es, die Abweichung der imputierten Werte und der eigentlichen Werte gegenüber den bisherigen Verfahren (über den RMSE ausgewertet) deutlich zu verringern.

Es zeigt sich, dass ein ML-basiertes Imputationsverfahren (missRanger) gegenüber den Referenzverfahren bei einzelnen fehlenden Werten zu einer deutlichen Reduktion der Abweichungen zwischen wahren und imputierten Werten führt. Damit ist diese Methode für eine Weiterentwicklung der Imputationsmethodik für fehlende Einzelwerte ein vielversprechender Anwendungsfall für die StatVG und für den Statistischen Verbund.

67 Handling constraints in automated statistical data editing via full conditional distributions

Christian Aßmann (1,2), **Ariane Würbach** (1), Katja-Verena Bürk (3), Florian Dumpert (3)

1: Leibniz Institute for Educational Trajectories Bamberg, Germany; 2: Chair of Survey Statistics and Data Analysis, Otto-Friedrich-University Bamberg, Germany; 3: German Federal Statistical Office
ariane.wuerbach@lifbi.de

SECTION: Methodology of Statistical Surveys (Session 2)

Reported survey data are prone to inaccuracies due to respondent error as reported values may be missing or implausible, i.e., they do not satisfy logical constraints. When such logical constraints are due to the interaction of multiple variables, it is also unclear which variable or variables are actually erroneous. A standard method used by Statistical Offices to correct data and ensure data consistency are edit-imputation routines following the Fellegi-Holt paradigm. Using such an easily computable heuristic does not necessarily exploit all the information available in the observed data. Another way that incorporates all available information is to apply Bayesian methods in the form of full conditional distributions of missing values to properly account for the uncertainty that arises in the process of replacing erroneous values. While Bayesian approaches based on parametric models are available in the literature for categorical and continuous data, this paper presents a method for specifying full conditional distributions using classification and regression trees instead, while taking into account nested balance constraints, i.e., linked constraints involving multiple variables. The CART algorithm was chosen, because it provides flexible univariate approximations to the full conditional distributions of the variables while reducing the computational intensity of the overall Bayesian approach. Results from simulation suggest that, compared to complete case analysis, the average root mean squared error of moment estimates can typically be reduced by 20 to 30 percent when using the nonparametric Bayesian approach and the corresponding specification of full conditional distributions using the CART algorithm.

68 Die hybride MigrantInnenbeiratswahl in Freiburg

Michael Haußmann

Stadt Freiburg i.Br., Deutschland

michael.haussmann@freiburg.de

SECTION: VDSSt (Session 2): Demokratie und Wahlen aus statistischer Sicht

In Freiburg wird der städtische MigrantInnenbeirat, der den Gemeinderat berät, seit den 1980er Jahren gewählt. Eine besondere Herausforderung ist dabei die geringe Beteiligung aufgrund zahlreicher Faktoren, unter anderem aufgrund mangelnder Deutschkenntnisse. Das Freiburger Wahlamt hat deshalb in diesem Jahr erstmals neben der Wahl im Wahllokal die Online-Wahl in 12 Sprachen angeboten. Die Bilanz fiel trotz eines versuchten Wahlbetrugs ausgesprochen positiv aus.

69 Wie entwickelt sich demokratische Kultur im urbanen Raum? Ergebnisse des Leipziger Demokratie-Monitors 2024

Christopher Pollak (1), Philipp Greunke (2)

1: Universität Leipzig, Deutschland; 2: Stadt Leipzig, Deutschland

c.pollak@uni-leipzig.de

SECTION: VDSSt (Session 2): Demokratie und Wahlen aus statistischer Sicht

Wie steht es aktuell um die demokratische Verfasstheit der bevölkerungsreichsten Stadt in Sachsen? Der Leipziger Demokratie-Monitor 2024 liefert auf diese Frage empirisch fundierte Antworten. Aufbauend auf jährlich erhobenen Daten der Kommunalen Bürgerumfrage bietet der Monitor einen Überblick über die Verbreitung gruppenbezogener Menschenfeindlichkeit, Vertrauen in etablierte Institutionen, Einstellungen zu demokratischer Teilhabe und weiteren demokratiebezogenen Einstellungen in der Leipziger Stadtgesellschaft.

Dabei zeigen die Daten neben der Zustimmung zu demokratischen Werten auch besorgniserregende Entwicklungen, etwa zunehmenden Autoritarismus und ein weit verbreitetes Gefühl mangelnder politischer Selbstwirksamkeit. Die Einstellungen werden sowohl mit objektiven Faktoren wie Alter, Bildung, Einkommen oder Wohnlage als auch mit subjektiven Faktoren (Zukunftserwartung, Lebenszufriedenheit, Wahrnehmung der eigenen wirtschaftlichen Situation) in Beziehung gesetzt. Dabei zeigt sich: Wer sozial eingebunden, wirtschaftlich abgesichert und optimistisch ist, weist häufiger demokratiebefürwortende Haltungen auf. Der Vortrag stellt neben der Methodik und ausgewählten Befunden auch deren Relevanz für die lokale Demokratieförderung vor.

70 Wahlhochrechnung statt reiner Ergebnispräsentation

Volker Holzendorf

Stadtverwaltung Jena, Deutschland
volker.holzendorf@jena.de

SECTION: VDSSt (Session 2): Demokratie und Wahlen aus statistischer Sicht

In Jena fanden seit Mai 2024 Wahlen auf allen Ebenen (Oberbürgermeister-, Stadtrats-, Landtags-, Bundestags und Europawahl) statt. Dies nahm die Statistikstelle Jena zum Anlass, ein Wahl-dash-board zu entwickeln. (www.shiny.jena.de/Wahlen/). Mit diesem Dashboard können sich die Bürgerinnen und Bürger intensiv über verschiedene Auswahlfelder mit den vergangenen Wahlen für Analysezwecke beschäftigen. Das Dashboard zeigt Wahlergebnisse (Wahlgebietsgewinner, Hochburgenkarte) und die Metadaten zum Wahlgebiet in verschiedenen Formen kleinräumig an. In einer Ausbaustufe des Dashboard, das auch den Mitgliedsstädten der KO.R-Gemeinschaft des KOSIS Verbundes zu Verfügung steht, soll die Ergebnispräsentation am Wahlabend realisiert werden. Dazu wird auch eine Prognose des gesamtstädtischen Ergebnisses gehören. Das gesamt-städtische Ergebnis und das Ergebnis der zwei Landtagswahlkreise in Jena soll auf Grund der bereits eingegangenen Wahlbezirke vorhersagt werden. Derzeit werden am Wahlabend in Kommunen wie auch in Jena lediglich die eingegangenen Meldungen von Wahlbezirken dargestellt. Dabei wird das Ergebnis einzelner Wahlbezirke lediglich zusammengezählt ohne zu berücksichtigen, dass Wahlkreise unterschiedlich groß sind und wegen ihrer Sozialstruktur unterschiedlich wählen. Die Entwicklungsschritte des Prognosemodells werden im Vortrag vorgestellt.

71 Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (I)

Tanja Lenz

Bundesverband Baustoffe - Steine und Erden e.V. (bbs)
s.jaeger@bdi.eu

SECTION: Economic, Social and Market Statistics (Session 4)

Die Wirtschaftsverbände gehören zu wichtigen Akteuren im deutschen Gesellschafts- und Politiksystem. Sie vertreten die Interessen ihrer Mitglieder, vom Einzelunternehmer bis zum Weltkonzern, gegenüber Politik und Verwaltung, anderen Bereichen der Wirtschaft, der Wissenschaft und den Medien. Sie bündeln differenzierte Meinungen und Positionen innerhalb ihrer Mitgliedschaft und bieten der Politik abgestimmte, repräsentative Positionen der Wirtschaft. Sie fungieren damit als ein wichtiger Player im demokratischen Diskurs und schärfen das Bewusstsein für spezifische Themen auch gesellschaftspolitisch.

Für ihre Aufgaben benötigen die Wirtschaftsverbände solide und aussagekräftige statistische Informationen. Wirtschaftsverbände nutzen daher sehr intensiv ein breites Spektrum an amtlichen und nichtamtlichen Daten, von den Statistischen Ämtern von Bund und Ländern und anderen Behörden, von Wirtschaftsforschungsinstituten sowie privaten Datenanbietern. Auch eigene Verbandserhebungen ergänzen das breite Spektrum an Datenquellen. Viele Verbände bieten auf verschiedensten Kommunikationswegen ihren Mitgliedern einen umfassenden Datenservice an, beispielsweise über eigene regelmäßige Publikationen oder verbandspezifische Datenbanken.

Die Session möchte aufzeigen, wie und für welche Zwecke die Wirtschaftsverbände ihre Daten nutzen, erheben, verarbeiten und kommunizieren. Verbandsvertreterinnen und Vertreter aus verschiedenen Wirtschaftsbereichen stellen das breite Spektrum ihrer verbandsinternen statistischen Arbeit vor und möchten für die vielfältigen Bedarfe in diesem Aufgabenbereich sensibilisieren.

I) Tanja Lenz vom Bundesverband Baustoffe - Steine und Erden (bbs) wird im ersten Vortrag der Session über die statistische Arbeit des Verbandes für die Wirtschaftsbereiche Bau- und Rohstoffindustrie berichten.

72 Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (II)

Angelika Becker

Verband der Chemischen Industrie e.V. (VCI)
s.jaeger@bdi.eu

SECTION: Economic, Social and Market Statistics (Session 4)

Die Wirtschaftsverbände gehören zu wichtigen Akteuren im deutschen Gesellschafts- und Politiksystem. Sie vertreten die Interessen ihrer Mitglieder, vom Einzelunternehmer bis zum Weltkonzern, gegenüber Politik und Verwaltung, anderen Bereichen der Wirtschaft, der Wissenschaft und den Medien. Sie bündeln differenzierte Meinungen und Positionen innerhalb ihrer Mitgliedschaft und bieten der Politik abgestimmte, repräsentative Positionen der Wirtschaft. Sie fungieren damit als ein wichtiger Player im demokratischen Diskurs und schärfen das Bewusstsein für spezifische Themen auch gesellschaftspolitisch.

Für ihre Aufgaben benötigen die Wirtschaftsverbände solide und aussagekräftige statistische Informationen. Wirtschaftsverbände nutzen daher sehr intensiv ein breites Spektrum an amtlichen und nichtamtlichen Daten, von den Statistischen Ämtern von Bund und Ländern und anderen Behörden, von Wirtschaftsforschungsinstituten sowie privaten Datenanbietern. Auch eigene Verbandserhebungen ergänzen das breite Spektrum an Datenquellen. Viele Verbände bieten auf verschiedensten Kommunikationswegen ihren Mitgliedern einen umfassenden Datenservice an, beispielsweise über eigene regelmäßige Publikationen oder verbandspezifische Datenbanken.

Die Session möchte aufzeigen, wie und für welche Zwecke die Wirtschaftsverbände ihre Daten nutzen, erheben, verarbeiten und kommunizieren. Verbandsvertreterinnen und Vertreter aus verschiedenen Wirtschaftsbereichen stellen das breite Spektrum ihrer verbandsinternen statistischen Arbeit vor und möchten für die vielfältigen Bedarfe in diesem Aufgabenbereich sensibilisieren.

Angelika Becker vom Verband der Chemischen Industrie (VCI) wird im zweiten Vortrag der Session über die statistische Arbeit des Verbandes für den Wirtschaftsbereich der chemisch-pharmazeutischen Industrie als Teil des Verarbeitenden Gewerbes berichten.

73 Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (III)

René Rimpler

Zentralverband des Deutschen Handwerks (ZDH)
s.jaeger@bdi.eu

SECTION: Economic, Social and Market Statistics (Session 4)

Die Wirtschaftsverbände gehören zu wichtigen Akteuren im deutschen Gesellschafts- und Politiksystem. Sie vertreten die Interessen ihrer Mitglieder, vom Einzelunternehmer bis zum Weltkonzern, gegenüber Politik und Verwaltung, anderen Bereichen der Wirtschaft, der Wissenschaft und den Medien. Sie bündeln differenzierte Meinungen und Positionen innerhalb ihrer Mitgliedschaft und bieten der Politik abgestimmte, repräsentative Positionen der Wirtschaft. Sie fungieren damit als ein wichtiger Player im demokratischen Diskurs und schärfen das Bewusstsein für spezifische Themen auch gesellschaftspolitisch.

Für ihre Aufgaben benötigen die Wirtschaftsverbände solide und aussagekräftige statistische Informationen. Wirtschaftsverbände nutzen daher sehr intensiv ein breites Spektrum an amtlichen und nichtamtlichen Daten, von den Statistischen Ämtern von Bund und Ländern und anderen Behörden, von Wirtschaftsforschungsinstituten sowie privaten Datenanbietern. Auch eigene Verbandserhebungen ergänzen das breite Spektrum an Datenquellen. Viele Verbände bieten auf verschiedensten Kommunikationswegen ihren Mitgliedern einen umfassenden Datenservice an, beispielsweise über eigene regelmäßige Publikationen oder verbandspezifische Datenbanken.

Die Session möchte aufzeigen, wie und für welche Zwecke die Wirtschaftsverbände ihre Daten nutzen, erheben, verarbeiten und kommunizieren. Verbandsvertreterinnen und Vertreter aus verschiedenen Wirtschaftsbereichen stellen das breite Spektrum ihrer verbandsinternen statistischen Arbeit vor und möchten für die vielfältigen Bedarfe in diesem Aufgabenbereich sensibilisieren.

René Rimpler vom Zentralverband des Deutschen Handwerks (ZDH) wird im dritten Vortrag der Session über die statistische Arbeit des Verbandes für den Wirtschaftsbereich des Handwerks berichten.

74 Wirtschaftsverbände als wichtige Datennutzer: Das breite Spektrum der verbandsinternen statistischen Arbeit an vier Beispielen (IV)

Dr. Florian Bayer

Bitkom e.V.
s.jaeger@bdi.eu

SECTION: Economic, Social and Market Statistics (Session 4)

Die Wirtschaftsverbände gehören zu wichtigen Akteuren im deutschen Gesellschafts- und Politiksystem. Sie vertreten die Interessen ihrer Mitglieder, vom Einzelunternehmer bis zum Weltkonzern, gegenüber Politik und Verwaltung, anderen Bereichen der Wirtschaft, der Wissenschaft und den Medien. Sie bündeln differenzierte Meinungen und Positionen innerhalb ihrer Mitgliedschaft und bieten der Politik abgestimmte, repräsentative Positionen der Wirtschaft. Sie fungieren damit

als ein wichtiger Player im demokratischen Diskurs und schärfen das Bewusstsein für spezifische Themen auch gesellschaftspolitisch.

Für ihre Aufgaben benötigen die Wirtschaftsverbände solide und aussagekräftige statistische Informationen. Wirtschaftsverbände nutzen daher sehr intensiv ein breites Spektrum an amtlichen und nichtamtlichen Daten, von den Statistischen Ämtern von Bund und Ländern und anderen Behörden, von Wirtschaftsforschungsinstituten sowie privaten Datenanbietern. Auch eigene Verbandserhebungen ergänzen das breite Spektrum an Datenquellen. Viele Verbände bieten auf verschiedensten Kommunikationswegen ihren Mitgliedern einen umfassenden Datenservice an, beispielsweise über eigene regelmäßige Publikationen oder verbandspezifische Datenbanken.

Die Session möchte aufzeigen, wie und für welche Zwecke die Wirtschaftsverbände ihre Daten nutzen, erheben, verarbeiten und kommunizieren. Verbandsvertreterinnen und Vertreter aus verschiedenen Wirtschaftsbereichen stellen das breite Spektrum ihrer verbandsinternen statistischen Arbeit vor und möchten für die vielfältigen Bedarfe in diesem Aufgabenbereich sensibilisieren.

Dr. Florian Bayer von Bitkom wird im vierten Vortrag der Session über die statistische Arbeit des Verbandes für den Wirtschaftsbereich der ITK-Branche berichten.

75 Arbeitsmarktintegration geflüchteter Menschen - Bewertung des Job-Turbos aus Sicht der Arbeitsmarktstatistik

Anton Klaus

Statistik der Bundesagentur für Arbeit, Deutschland
anton.klaus@arbeitsagentur.de

SECTION: Economic, Social and Market Statistics (Session 5): Labour Markets and Social Security I

Im Oktober 2023 hat die Bundesregierung den Auftakt zum "Jobturbo" bekannt gegeben. Ziel war es, die Prozesse zur Arbeitsmarktintegration von Geflüchteten aus der Ukraine und anderen Herkunftsstaaten beschleunigen. Der Vortrag beleuchtet den Stand der Arbeitsmarktintegration, zeigt auf, wo die amtliche Statistik bei der Bewertung (noch) an ihre Grenzen stößt und welches Potenzial in Längsschnittdaten steckt.

76 Integrationsverläufe von Geflüchteten: Eine Kohortenanalyse

Ehsan Vallizadeh

Statistik der Bundesagentur für Arbeit, Deutschland
ehsan.vallizadeh@arbeitsagentur.de

SECTION: Economic, Social and Market Statistics (Session 5): Labour Markets and Social Security I

Um den Integrationserfolg von Zugewanderten detaillierter untersuchen zu können, entwickelt die Statistik der BA weitere Messungen und Darstellungen, die über die Standardindikatoren zur Arbeitsmarktintegration hinausgehen. Ein häufiges Problem der bisherigen Indikatoren, wie die Beschäftigungsquote, ist, dass diese verschiedene Personengruppen umfassen, die zu unterschiedlichen Zeiträumen nach Deutschland eingereist sind. Dadurch werden Entwicklung und Dynamik der Arbeitsmarktintegration aufgrund der unterschiedlichen Aufenthaltsdauer deutlich verzerrt. Um diese Lücke zu schließen, nutzt die Statistik der BA verstärkter Kohortenanalyse, bei

denen homogene Ausgangskohorten von Geflüchteten durch den ersten Eintrag in den Statistiken der BA definieren werden und in Längsschnittbetrachtungen ihr weiterer Verbleib analysiert wird. Im Rahmen des Vortrags soll dieses neue Auswertungskonzept diskutiert und die Analysepotenziale vorgestellt werden.

77 Fachkräfteeinwanderung in der Beschäftigungsstatistik

Malte Reimann-Peters

Statistik der Bundesagentur für Arbeit, Deutschland
malte.reimann-peters@arbeitsagentur.de

SECTION: Economic, Social and Market Statistics (Session 5): Labour Markets and Social Security I

Das Beschäftigungswachstum wurde 2023 und 2024 ausschließlich von Ausländern getragen. Neben einer reinen Analyse nach Staatsangehörigkeiten stehen der Beschäftigungsstatistik der BA über eine Schnittstelle mit dem Ausländerzentralregister auch Daten zum Aufenthaltsstatus der Beschäftigten zur Verfügung. Der Vortrag soll erläutern, welche Rolle die Erwerbsmigration im Kontext Beschäftigungswachstum der Ausländer spielt. Neben der grundsätzlichen Entwicklung seit Inkrafttreten des Fachkräfteeinwanderungsgesetzes im März 2020 sollen auch tiefere Analysemöglichkeiten beleuchtet werden.

78 Demografischer Wandel - Ausländische Beschäftigte ersetzen in vielen Bereichen fehlendes einheimisches Personal

Sarah Kuhn (1), Barbara Schwengler (2), Holger Seibert (1), Doris Wiethölter (1)

1: IAB Berlin-Brandenburg, Berlin; 2: Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg
Barbara.Schwengler@iab.de

SECTION: Economic, Social and Market Statistics (Session 5): Labour Markets and Social Security I

Deutschland steht vor einer enormen demografischen Herausforderung, da die geburtenstarken Jahrgänge in den kommenden Jahren den Arbeitsmarkt verlassen. Sie können durch die viel kleineren Berufseinstiegskohorten nur zum Teil ersetzt werden – insbesondere in ländlichen Regionen und in Ostdeutschland. Ausländische Beschäftigte können jedoch helfen, den Arbeitskräftemangel zu kompensieren und die demografische Situation abzufedern.

Wir untersuchen die Beschäftigungsentwicklung von 2015 bis 2024 nach Altersstruktur, Anforderungsniveau der ausgeübten Tätigkeiten (Helfer, Fachkraft, Spezialist, Experte), Staatsangehörigkeit in Ost- und Westdeutschland und auf der Ebene der siedlungsstrukturellen Kreistypen des Bundesinstituts für Bau-, Stadt- und Raumforschung (BBSR). Wir zeigen, dass es einen deutlichen Anstieg der ausländischen Beschäftigten auf allen Anforderungsniveaus gab. Insbesondere auf dem Helfer- und Fachkraftniveau ging der Beschäftigungszuwachs allein auf ausländische Beschäftigte zurück.

Auf der regionalen Ebene lässt sich feststellen, dass in fast allen Regionen die Anteile von Fachkräften in dem Untersuchungszeitraum gesunken sind, während die Beschäftigung auf Spezialisten- und Expertenniveau vor allem in den kreisfreien Großstädten deutlich gestiegen ist. Gleichzeitig haben die Helferanteile der Beschäftigten insbesondere in den städtischen und ländlichen Kreisen zugenommen, und dies deutlich stärker in den östlichen Regionen.

Die demografischen Auswirkungen auf den Arbeitsmarkt zeigen sich beim Vergleich der Altersgruppe der 58 bis 62-Jährigen, die noch im Erwerbsleben steht, mit der Altersklasse der 28 bis 32-Jährigen. Den stärksten Deckungsbedarf gibt es in Ostdeutschland, wo die älteren Kohorten deutlich stärker sind und nicht von den nachrückenden Kohorten kompensiert werden können. In den städtischen und ländlichen Regionen Westdeutschlands erhöhen die ausländischen Beschäftigten hingegen die Gesamtbeschäftigung vor allem in den jungen und mittleren Altersklassen. Es können jedoch nur in den kreisfreien Großstädten die Altersabgänge in Rente durch junge Beschäftigte kompensiert werden. Eine ungünstige Relation der gewählten jüngeren zu der älteren Altersgruppe weisen Tätigkeiten auf Fachkraft- und Helferniveau auf, während diese für Beschäftigte auf Spezialisten- und Expertenniveau günstiger ausfällt.

79 Learning Signal-to-Noise Ratios from Forecast Errors: A Simulation-Based Calibration Framework

Rainer Alexander Schüssler

University of Münster, Deutschland
raineralexanderschuessler@gmail.com

SECTION: Young-Academics Mini-Symposium: Modern Time Series Econometrics

We propose a data-driven framework to calibrate signal-to-noise ratios (SNRs) in predictive modeling based on empirical forecast errors from time series data. By matching observed forecast error magnitudes with simulated outcomes, we infer realistic noise levels that reflect actual predictive difficulty—without requiring knowledge of the data-generating process (DGP). This enables model evaluations under empirically plausible SNRs. In a macroeconomic case study on support recovery, we demonstrate that variable selection may fail even under sparsity—not due to model misspecification, but because the signal is too weak to recover. Our results emphasize a crucial but underexplored distinction between structural failure and informational limits.

80 Asymptotic Properties of MAGMAR-Copula Time Series Models

Sven Pappert

TU Dortmund, Deutschland
pappert@statistik.tu-dortmund.de

SECTION: Young-Academics Mini-Symposium: Modern Time Series Econometrics

Copula-based time series models implicitly assume a finite Markov order. In reality, a time series may not follow the Markov property. We modify copula-based time series models by introducing a moving aggregate (MAG) part into the model updating equation. The functional form of the MAG-part is given as the inverse of a conditional copula. The resulting MAG-modified autoregressive copula-based time series model (MAGMAR-Copula) is discussed in detail, and distributional properties are derived in a D-vine framework. First, we investigate stationarity and mixing of the time series, and then we explore the asymptotic properties of maximum-likelihood estimators.

81 Functional Factor Regression with an Application to Electricity Price Curve Modeling

Sven Otto, Luis Winter

University of Cologne, Deutschland
sven.otto@uni-koeln.de

SECTION: Young-Academics Mini-Symposium: Modern Time Series Econometrics

We propose a function-on-function linear regression model for time-dependent curve data that is consistently estimated by imposing factor structures on the regressors. An integral operator based on cross-covariances identifies two components for each functional regressor: a predictive low-dimensional component, along with associated factors that are guaranteed to be correlated with the dependent variable, and an infinite-dimensional component that has no predictive power. In order to consistently estimate the correct number of factors for each regressor, we introduce a functional eigenvalue difference test. Our setting allows us to construct a novel central limit theorem for the regression parameters in a fully functional model, making it possible to construct confidence bands and conduct statistical inference. The model is applied to forecast electricity price curves in three different energy markets. Its prediction accuracy is found to be comparable to popular machine learning approaches, while providing statistically valid inference and interpretable insights into the conditional correlation structures of electricity prices.

82 Bootstrap Inference in Panels of Cointegrating Regressions with Global Stochastic Trends

Karsten Reichold (1), Carsten Jentsch (2), Christoph Hanck (3)

1: TU Wien, Austria; 2: TU Dortmund University; 3: University of Duisburg-Essen
karsten.reichold@tuwien.ac.at

SECTION: Young-Academics Mini-Symposium: Modern Time Series Econometrics

Bai, Kao, and Ng (2009, *Journal of Econometrics* 149, 82–99) propose continuously updated (CUP) estimators for panel cointegrating regression models with cross-sectional dependence generated by unobserved global stochastic trends. While the CUP estimation approach generally performs well, test statistics based upon the CUP estimators suffer from enormous size distortions in finite samples. To address this problem, we propose a block-diagonal VAR sieve bootstrap to capture the second-order time series and cross-section dependence structure in the data and prove bootstrap consistency for the test statistics based upon the CUP estimators under sequential limits with $N \rightarrow \infty$ after $T \rightarrow \infty$. Simulation results reveal that using bootstrap critical values reduces size distortions of the test considerably, with negligible power losses under the alternative. An empirical application demonstrates the importance of the block-diagonal VAR sieve bootstrap in practice by analyzing the Fisher effect in 19 OECD countries.

83 Transformer-Modelle in der Außenhandelsstatistik: Potenziale und Herausforderungen im Vergleich mit klassischen Vektorraummodellen

Sara Schiesberg, Heiko Limberg

Statistisches Bundesamt, Deutschland
sara.schiesberg@destatis.de

SECTION: Computational Statistics and Data Science (Session 4)

In diesem Vortrag untersuchen wir den Einsatz von Transformer-Modellen in der Außenhandelsstatistik und vergleichen sie mit klassischen Vektorraummodellen. Transformer-Modelle haben sich in vielen Bereichen der Sprachverarbeitung als leistungsstark erwiesen. Wir diskutieren sowohl die Potenziale als auch die Herausforderungen dieser Methoden für unseren Fachbereich, insbesondere hinsichtlich ihrer Fähigkeit, komplexe Zusammenhänge in Daten zu erkennen und übertragbar auf verschiedene Anwendungsfälle zu bleiben. Dabei betrachten wir sowohl theoretische Überlegungen als auch erste praktische Erfahrungen mit diesen Ansätzen. Zudem skizzieren wir erste Ergebnisse aus unseren Analysen, die Hinweise auf die Leistungsfähigkeit und Grenzen dieser Modelle in der Außenhandelsstatistik liefern.

84 Model Governance und Compliance unter den speziellen Arbeitsbedingungen im öffentlichen Sektor.

Ulrich Reincke, Frank Gramüller

SAS Institute GmbH, Deutschland
Ulrich.Reincke@sas.com

SECTION: Computational Statistics and Data Science (Session 4)

Die Arbeitsbedingungen für Data Scientists und Statistiker im öffentlichen Sektor sind oft mit strengen regulatorischen Anforderungen und begrenzten Ressourcen verbunden. Berücksichtigt man zusätzlich die langwierigen Beschaffungswege im öffentlichen Bereich über Ausschreibungen, so scheint der Einsatz von Open Source Software wie Python und R oft im wahrsten Sinne des Wortes die „erste Wahl“ und alternativlos zu sein.

Mit dem Inkrafttreten der Europäischen Datenschutzgrundverordnung 2018 und dem AI Akt 2024 ist die Einhaltung von Governance- und Compliance-Richtlinien bei der Modellierung personenbezogener Daten verschärft worden. Diese Verschärfungen betreffen auch den öffentlichen Sektor, der im Bereich der Digitalisierung von Prozessen anderen Branchen noch hinterherhinkt.

Diese Präsentation beleuchtet diese Herausforderungen und zeigt mögliche Lösungsansätze im Bereich von Model Governance und Compliance mit der SAS-Viya Plattform. Im Vortrag wird deutlich gemacht, dass man mit der prozessübergreifenden SAS-Viya Plattform auf die Verwendung verschiedener Open Source Software wie z.B. Python und R nicht verzichten muss und bei der Einhaltung der regulativen Anforderungen unterstützt wird.

Durch praxisnahe Beispiele und Fallstudien wird gezeigt, welche Risiken und Anforderungen es gibt und wie die Integration von Open Source in der SAS Plattform eine robuste aber auch hinreichend flexible Umgebung für die Modellierung und Analyse schafft, in der die Compliance Regeln abgebildet werden können.

Besondere Aufmerksamkeit wird dabei auf die Komponenten SAS Model Manager und SAS Model Risk Management gelegt, die sicherstellen, dass Modelle genau, zuverlässig, transparent, fair, unverzerrt, nachvollziehbar sind, und nach Daten- und Softwareversionsstand auch rückwirkend auditierbar.

Die Präsentation bietet wertvolle Einblicke und praktische Empfehlungen für Data Scientists und Statistiker, die in einem komplexen regulatorischen Umfeld arbeiten und gleichzeitig auf innovative Analysetechniken nicht verzichten möchten.

85 Datenschutzkonforme Methodik für den Zugang zu Mobilfunkdaten

Lorenz Ade

Statistisches Bundesamt, Deutschland
lorenz.ade@destatis.de

SECTION: Computational Statistics and Data Science (Session 4)

Seit 2017 erstellt das Statistische Bundesamt experimentelle Statistiken auf Basis von Mobilfunkdaten. Die in diesen Projekten verwendeten Rohdaten werden bisher in einer „methodischen Blackbox“ aufbereitet und verarbeitet, da die Mobilfunkbetreiber nur aggregierte Daten auf der Grundlage ihrer vertraulichen Algorithmen bereitstellen.

Wenn Mobilfunkdaten in die amtliche Statistikproduktion einfließen sollen, müssen die Daten die notwendigen Qualitätskriterien der Transparenz, Zugänglichkeit und Vergleichbarkeit erfüllen. Außerdem müssen die Datenschutzstandards jederzeit eingehalten werden, um das Vertrauen der Öffentlichkeit in die amtliche Statistik zu erhalten. Um diese Themen zu bearbeiten kooperieren die Mobilfunkanbieter im Rahmen des Projektes „Anonymität bei integrierten und georeferenzierten Daten (AnigeD)“ mit dem Statistischen Bundesamt.

Das Projekt zielt darauf ab, die „methodische Blackbox“ der Mobilfunkdatenverarbeitung mit der Implementierung eines Verarbeitungs- und Anonymisierungsverfahrens zu öffnen. Hierzu gehören auch Anonymisierungsverfahren für die Nutzung von anonymisierten georeferenzierten Mobilfunkdaten, die Konzeption und der Aufbau der erforderlichen Entwicklungsumgebung beim Datenprovider, die Entwicklung eines Modellprozesses für die zukünftige Zusammenarbeit zwischen privaten (Mobilfunk-)Datenanbietern und dem Statistischen Bundesamt und ggf. weiteren staatlichen Institutionen.

Vorgestellt werden die aktuelle Projektergebnisse. Dies sind insbesondere die besonderen Datenschutzaspekte der Mobilfunkdaten und mögliche Anonymisierungsverfahren, als auch die technischen Herausforderungen bei der Verarbeitung sehr großer Datenmengen.

86 Erstellung eines flächendeckenden Wohnraumregisters für dynamische Mikrosimulationen in Deutschland

Jan Weymeirsch, Ralf Münnich

Universität Trier, Deutschland
weymeirsch@uni-trier.de

SECTION: Computational Statistics and Data Science (Session 4)

Räumlich-dynamische Mikrosimulationen haben ein besonderes Potenzial, Bevölkerungen auf einer sehr detaillierten geografischen Ebene, wie Stadtteilen, Häuserblöcke oder Adressen zu simulieren. Hierfür wird typischerweise ein detaillierter Gebäude- und Wohnungsdatensatz benötigt, um feingliedrige Wanderungsströme modellieren zu können, insbesondere in Hinsicht auf den sich sehr dynamisch entwickelnden Wohnungsmarkt und lokale Wohnraumkapazitäten. Es gibt jedoch zum aktuellen Zeitpunkt kein umfassendes Gebäuderegister für Deutschland, welches den Ansprüchen für die geplante Nutzung genügt, insbesondere keines, welches für die Forschungsgemeinschaft offen zugänglich ist und Gebäude hinsichtlich ihrer Nutzung als Wohnraum oder als potenzielle Arbeitsstätte unterscheidet.

In einer ersten Pilotstudie haben wir bereits Möglichkeiten evaluiert, öffentlich zugängliche Daten, insbesondere OpenStreetMap (OSM) und lokal verfügbare amtliche Daten als Grundlage für den Aufbau eines solchen Datensatzes in einer deutschen Großstadt zu nutzen. Aufbauend aus den daraus gezogenen Schlussfolgerungen, möchten wir unseren Ansatz nun flächendeckend auf ganz Deutschland ausweiten. Hierbei nutzen wir modellbasierte Ansätze sowie Machine Learning Methoden, um amtliche Daten mit öffentlichen Datenquellen wie beispielsweise OSM zu verknüpfen und Gebäude nach Ihrer Nutzung und Ihrem Wohnraum zu klassifizieren. Der sich hieraus ergebende Datensatz soll der wissenschaftlichen Gemeinschaft frei zur Verfügung gestellt werden und auch für feingliedrige räumliche Mikrosimulationen verwendet werden können.

Dieser Beitrag stellt die von uns verfolgten Strategien, Methoden sowie erste Ergebnisse in Teilbereichen von Deutschland vor.

87 Aktuelle Ergebnisse der QuBe-Bevölkerungsprojektion für Kreise und kreisfreie Städte

Christian Schneemann, **Johanna Zenk**, **Gerd Zika**

Institut für Arbeitsmarkt- und Berufsforschung (IAB), Deutschland
christian.schneemann2@iab.de

SECTION: DGD (Session 1): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen I

Seit dem Jahr 2016 basieren die BIBB-IAB-Qualifikations- und Berufsprojektionen auf einer eigens erstellten QuBe-Bevölkerungsprojektion. Sie hebt sich insbesondere durch die Unterscheidung von Personen mit und ohne deutsche Staatsbürgerschaft von bereits bestehenden Bevölkerungsvorausberechnungen ab. Durch die Unterscheidung nach Staatsbürgerschaft können die erheblichen Unterschiede bei den Geburtenziffern und dem Wanderungsverhalten berücksichtigt werden. Aufbauend auf die QuBe-Bevölkerungsprojektion für die Bundesebene und basierend auf der Methodik von Studtrucker u.a. (2022), wird seit 2022 zusätzlich eine jährliche QuBe-Bevölkerungsprojektion für die 400 Kreise und kreisfreien Städte in Deutschland bis zum Jahr 2060 erstellt.

In die Modellierung der QuBe-Bevölkerungsprojektion für Kreise und kreisfreie Städte aus dem Jahr 2024 floss unter anderem die Reform des Staatsbürgerschaftsrechts über höhere Einbürgerungsquoten ein (Schneemann u.a. 2025). Dies wirkte sich indirekt auf die Geburtenziffern und das Wanderungsverhalten aus. Dennoch stieg der Anteil von Personen ohne deutsche Staatsangehörigkeit an der Gesamtbevölkerung von rund 15,4 Prozent im Jahr 2023 auf rund 17,6 Prozent im Jahr 2040. Die Ergebnisse der regionalen Bevölkerungsprojektion bestätigten bereits bekannte demographische Entwicklungen.

Die Ergebnisse der QuBe-Bevölkerungsprojektion für Kreise und kreisfreie Städte dienen als Datengrundlage für die regionalspezifischen Arbeitsmarktprojektionen, die im Rahmen des QuBe-Projektes für die Bundesländer und 34 Arbeitsmarktregionen erstellt werden. Eine Herausforderung bei der Erstellung der regionalen Bevölkerungsprojektion stellt der Zugang zu Daten über die Zu- und Fortzüge in den einzelnen Kreisen dar. Im Rahmen der Statistischen Woche sollen die aktuellen Ergebnisse der QuBe-Bevölkerungsprojektion für Kreise und kreisfreie Städte aus dem Jahr 2025 vorgestellt werden, die auch die Zensuskorrektur aus dem Jahr 2023 berücksichtigen. Durch die Zensuskorrektur und den neuen Datenstand haben sich bereits starke Veränderungen bei der Bevölkerungsprojektion auf der Bundesebene ergeben. Der neue Datenstand wird sich auch in unterschiedlicher Stärke auf die Kreisebene auswirken und zu teilweise, im Vergleich zu den vorangegangenen Projektionen, anderen demografischen Entwicklungen in den Kreisen führen.

88 Projektion der Nachfrage nach Fachkräften in Gesundheitsberufen

Ines Thobe, Linus Ronsiek, Anja Sonnenburg

GWS mbH, Deutschland
thobe@gws-os.com

SECTION: DGD (Session 1): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen I

Im Zuge des demografischen Wandels, veränderter Verhaltensweise und rechtlicher Gestaltungen steigt hierzulande der Bedarf an Fachkräften in den Gesundheitsberufen. Gleichzeitig ist der Arbeitsmarkt durch einen Fachkräftewettbewerb geprägt, sodass sich die Frage nach einer flächendeckenden Sicherstellung einer qualitativ hochwertigen Gesundheitsversorgung stellt. Konkret heißt das: Das Gesundheitssystem muss mit Fachkräften in der richtigen Anzahl, mit der richtigen Qualifikation, am richtigen Ort und in der richtigen Zusammensetzung ausgestattet werden.

Um frühzeitig berufliche Passungsprobleme in den Berufen des Gesundheitswesens zu erkennen und geeignete Stellschrauben zur Behebung von Diskrepanzen zu identifizieren, hat das Bundesministerium für Gesundheit den Aufbau eines detaillierten langfristigen Monitorings von Angebot und Bedarf in den Gesundheitsberufen beauftragt. Gemeinsam beabsichtigt das Projektteam aus dem IAB, dem BIBB und der GWS, die Gesundheitsberufe in einer möglichst kleinteiligen Differenzierung zu erfassen und Ansatzpunkte politischer Handlungsmöglichkeiten zu operationalisieren, um Folgewirkungen von Maßnahmen zur adäquaten Behebung von Passungsproblemen abzuschätzen. Insgesamt werden 55 Berufe in drei verschiedenen Sektoren – ambulant, stationär und sonstige (Rettungsdienst, Verwaltung, etc.) – betrachtet.

Zur Erstellung dieses Monitorings auf der Nachfrageseite nimmt die GWS verschiedene Datensätze aus dem Gesundheitsbereich in den bestehenden Modellkontext des QUBE-Modells auf und schreibt diese mittels eines in einer vorangegangenen Machbarkeitsstudie entwickelten Projektionskonzepts fort. Im Zentrum steht dabei die Gesundheitspersonalrechnung des Statistischen Bundesamtes. Mithilfe weiterer Basisstatistiken wie beispielsweise der Pflege-, der Krankenhausstatistik oder auch

ärztlichen Abrechnungsdaten können Inanspruchnahmen von Gesundheits- und Pflegeleistungen sowie Personalschlüssel berechnet werden. Über die demografische Entwicklung – abgebildet durch die IAB-Bevölkerungsprojektion – kann dann die zukünftige Arbeitsnachfrage im Gesundheitsbereich sowohl im Volumen als auch in Erwerbstätigenzahlen projiziert werden.

Im Ergebnis entsteht ein Monitoring der Arbeitsmarktsituation im Gesundheitswesen, welches in seinen Ergebnissen kompatibel zu den Projektionen ist, die für das Fachkräftemonitoring des Bundesministeriums für Arbeit und Soziales verwendet werden.

89 Long-Term Care in Germany in the Context of the Demographic Transition—An Outlook for the Expenses of Long-Term Care Insurance through 2050

Patrizio Vanella (1), Christina Benita Wilke (2), Moritz Heß (3)

1: aQua-Institut; 2: FOM Hochschulzentrum Bremen; 3: Hochschule Niederrhein
patrizio.vanella@aqua-institut.de

SECTION: DGD (Session 1): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen I

Demographic aging results in a growing number of older people in need of care in many regions all over the world. Germany has witnessed steady population aging for decades, prompting policymakers and other stakeholders to discuss how to fulfill the rapidly growing demand for care workers and finance the rising costs of long-term care. Informed decisions on this matter to ensure the sustainability of the statutory long-term care insurance system require reliable knowledge of the associated future costs. These need to be simulated based on well-designed forecast models that holistically include the complexity of the forecast problem, namely the demographic transition, epidemiological trends, concrete demand for and supply of specific care services, and the respective costs. Care risks heavily depend on demographics, both in absolute terms and according to severity. The number of persons in need of care, disaggregated by severity of disability, in turn, is the main driver of the remuneration that is paid by long-term care insurance. Therefore, detailed forecasts of the population and care rates are important ingredients for forecasts of long-term care insurance expenditures. We present a novel approach based on a stochastic demographic cohort-component approach that includes trends in age- and sex-specific care rates and the demand for specific care services, given changing preferences over the life course. The model is executed for Germany until the year 2050 as a case study.

90 Forecast combination with an application to financial tail risk

Lukas Bauer

University of Freiburg, Deutschland
lukas.bauer@vwl.uni-freiburg.de

SECTION: Statistics in Finance (Session 1)

We propose a novel performance based forecast combination scheme. The scheme uses the standardized loss difference relative to the average model, thus accounting for the statistical magnitude of the losses. The theoretical goal of the scheme is to perform as well as the best candidate model, and we characterize the risk of the combination scheme relative to the best individual model.

Our scheme builds an intuitive bridge between model selection and forecast combination while showing robust performance in finite samples.

We apply the scheme to combine forecasts of financial tail risk, i.e., Value-at-Risk and expectiles. The data are large cap stocks from the NYSE Trade and Quote database. We find that the scheme is competitive, both with respect to alternative combination schemes and the individual models.

91 GDP nowcasting with large-scale inter-industry payment data in real time—A network approach

Kerstin Hötte, Anastasia Mantziou, Gesine Reinert, Mihai Cucuringu

KEDGE Business School, France
kerstin.hoette@posteo.de

SECTION: Statistics in Finance (Session 1)

Real-time economic information is essential for policy-making but difficult to obtain. We introduce a granular nowcasting method for macro- and industry-level GDP using a network approach and data on real-time monthly inter-industry payments in the UK. To this purpose we devise a model which we call an extended generalised network autoregressive (GNAR-ex) model, tailored for networks with time-varying edge weights and nodal time series, that exploits the notion of neighbouring nodes and neighbouring edges. The performance of the model is illustrated on a range of synthetic data experiments. We implement the GNAR-ex model on the payments network including time series information of GDP and payment amounts. To obtain robustness against statistical revisions, we optimise the model over 9 quarterly releases of GDP data from the UK Office for National Statistics. Our GNAR-ex model can outperform baseline autoregressive benchmark models, leading to a reduced forecasting error. This work helps to obtain timely GDP estimates at the aggregate and industry level derived from alternative data sources compared to existing, mostly survey-based, methods. Thus, this paper contributes both, a novel model for networks with nodal time series and time-varying edge weights, and the first network-based approach for GDP nowcasting based on payments data.

92 Forecasting Bond Returns With a Copula-Based Dynamic Factor Pricing Model

Yarema Okhrin (2), Maziar Sahamkhadam (1), **Andreas Stephan** (1)

1: Linnaeus University, School of Business and Economics, Sweden; 2: Augsburg University
andreas.stephan@lnu.se

SECTION: Statistics in Finance (Session 1)

This paper presents a novel copula-based no-arbitrage pricing framework for forecasting corporate bond returns and optimizing bond portfolios. Utilizing a copula-based dynamic factor model, we generate step-ahead forecasts for zero-coupon bond yields, which are subsequently applied to obtain and simulate the no-arbitrage prices for both callable and non-callable fixed-coupon bonds. These simulated bond prices serve as inputs for a novel convex multiobjective portfolio optimization, incorporating key criteria such as average returns, Conditional Value-at-Risk (CVaR), distance-to-default, transaction costs, and option-adjusted duration and convexity. Applying our methodology to a dataset of 879 corporate bonds denominated in Euros from January 2016 to July 2024, we demonstrate that the suggested copula-based no-arbitrage pricing framework takes advantage of the yield curve non-linear dependence structure and offers bond portfolios that consistently outperform those portfolios based on the classical dynamic Nelson-Siegel approach and an equally weighted (EQW) benchmark in terms of higher returns and Sharpe ratios while effectively reducing tail risk.

93 Stagewise crop yield prediction with multisource functional indices

Jing Zou (1,2), Ostap Okhrin (1,2)

1: Technische Universität Dresden, Germany; 2: Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI) Dresden/Leipzig
jing.zou@tu-dresden.de

SECTION: Statistics in Finance (Session 1)

Index insurance design involves integrating weather data, soil moisture, phenology information, and satellite imagery, which presents challenges in data fusion. This article addresses the modeling of multisource functional indices of varying lengths stagewise by boosting an ensemble of sequential models. The implemented methods, including nonparametric regression and deep learning models, aim to improve crop yield prediction by accounting for spatiotemporal dependence. Results from an applied case study demonstrate the feasibility of stagewise modeling and the hedging effectiveness of the proposed index insurance contracts.

94 Von Momentaufnahmen zu Studienverläufen: Methodische Innovationen in der amtlichen Hochschulstatistik

Stefanie Hoffmann, Tobias Laick

Statistisches Bundesamt
stefanie.hoffmann@destatis.de

SECTION: Methodology of Statistical Surveys (Session 3)

Die amtliche Hochschulstatistik liefert seit Jahrzehnten wertvolle Daten zu Studierenden, Studienanfänger/-innen sowie Absolventen/-innen. Bis zur Novelle des Hochschulstatistikgesetzes (HStatG) im Jahr 2016 war sie jedoch auf regelmäßige Querschnitterhebungen beschränkt. Dies erlaubte lediglich Momentaufnahmen, jedoch keine Analyse individueller Studienverläufe. Mit der Einführung der Studienverlaufsstatistik wurde ein paradigmatischer Wandel vollzogen: Erstmals können longitudinale Daten über die gesamte Hochschulkarriere in Deutschland hinweg erfasst und verknüpft werden. Dies eröffnet neue Möglichkeiten für die Hochschulforschung, -politik und -verwaltung und ermöglicht eine präzisere Analyse von Studienverläufen.

Dabei offenbart die Einführung der Studienverlaufsstatistik auch erhebliche methodische Herausforderungen, insbesondere im Bereich der Datenverknüpfung und des Datenschutzes. Denn eine zentrale Herausforderung im Bereich der Record Linkage besteht darin, individuelle Daten zuverlässig und konsistent zu verknüpfen. In Bezug auf die Studienverlaufsstatistik ergibt sich die zusätzliche Herausforderung, sicherzustellen, dass keine Rückschlüsse auf einzelne Personen möglich sind. Die Verknüpfung im Rahmen der Studienverlaufsstatistik erfolgt über ein datenschutzkonformes Pseudonymisierungsverfahren, das auf einem nicht-reversiblen Hash-Verfahren basiert. Dabei werden verschiedene unveränderliche Erhebungsmerkmale sowie zwei Hilfsmerkmale, nämlich die Angaben zum Tag im Geburtsdatum sowie die ersten vier Buchstaben des Vornamens verwendet, um einen pseudonymisierten Identifikator zu erstellen. Das Verfahren gewährleistet eine konsistente Verknüpfung der Datensätze über mehrere Semester hinweg und ermöglicht eine belastbare empirische Grundlage, die eine genauere Analyse von Übergängen vom Bachelor- zum Masterstudium, Studienabbrüchen und Studienerfolgen ermöglicht.

Der Beitrag beleuchtet die methodischen Herausforderungen der Studienverlaufsstatistik und bewertet die Datenqualität anhand von Zusammenführungs- und Wiederfindungsquoten. Darüber hinaus werden Ergebnisse zu den bereits entwickelten Quoten präsentiert. Diese zeigen unter anderem die fachspezifischen Unterschiede in den Übergangsraten vom Bachelor- zum Masterstudium. Abschließend werden Perspektiven für zukünftige Analysen diskutiert, die zur Weiterentwicklung der Hochschulforschung und -politik beitragen können.

95 Aufbau eines Bildungsverlaufsregisters in Deutschland

Katharina Giar, Franziska Hohlstein

Statistisches Bundesamt, Deutschland
Katharina.Giar@destatis.de

SECTION: Methodology of Statistical Surveys (Session 3)

Mit dem Aufbau eines Bildungsverlaufsregisters in Deutschland soll eine bildungsbereichsübergreifende statistische Datenbasis geschaffen werden, die Bildungsverlaufsdaten für die Wissenschaft, Politik und Wirtschaft bereitstellen kann. Die Statistischen Ämter des Bundes und der Länder

konzipieren hierzu gemeinsam mit den Kultusministerien und dem BMBF den Aufbau eines auf amtlichen Bildungsstatistiken basierenden „Bildungsverlaufsregisters im Verbund“ (BVR-V). Der föderale Aufbau des BVR-V entspricht der Kompetenzteilung zwischen Bund und Ländern: Bundesrechtlich geregelte Statistiken wie die Hochschul- und Berufsbildungsstatistik werden in einem Bundesbildungsverlaufsregister (BVR-B), länderrechtlich geregelte Statistiken wie die amtliche Schulstatistik in Bildungsverlaufsregistern der Länder (BVR-L) verarbeitet. Eine Zusammenführung der getrennten Statistikregister soll über das BVR-V stattfinden. Durch das Zuspielden einer Bildungs-ID, die mithilfe personenbezogener Merkmale in einer separaten „Vertrauensstelle“ zugeordnet oder durch die Ableitung eines eindeutigen Personenkennzeichens erstellt wird, können pseudonymisierte Bildungsverläufe erzeugt und ausgewertet werden.

Das Bildungsverlaufsregister zählt zu den Projekten, die die künftige Regierung nach aktueller Planung im Bereich Bildung umsetzen möchte. Hierfür ist die Schaffung rechtlicher Grundlagen auf Bundes- und Landesebene sowie ein Bund-Länder-Staatsvertrag notwendig. Mit dem Aufbau eines BVR wird ein wichtiger Schritt hin zu der lange geäußerten Forderung aus Wissenschaft und Forschung nach der Verknüpfbarkeit von registerbasierten Bildungsverlaufsdaten geschaffen, welches zu zahlreichen bildungsrelevanten Fragen eine qualitativ hochwertige empirische Grundlage bereitstellen kann. Zudem ist ein Vorteil der Nutzung amtlicher Daten, neben dem Wegfall typischer methodischer Probleme von Panelerhebungen, der direkte Bezug der Informationen aus der Verwaltung. Die intensive Plausibilisierung im Rahmen der statistischen Aufbereitung gewährleistet die Validität, Objektivität und die Reliabilität der gelieferten Angaben. Die Zusammenführung bisher separat erfasster Querschnittsdaten bringt neben dem umfassenden Nutzen auch methodische Herausforderungen mit sich. Hierunter fallen im Hinblick auf die Datenaufbereitung u.a. die Harmonisierung, Verknüpfung und Plausibilisierung der Bildungsstatistiken. Um die Bildungsverlaufsdaten im Anschluss analysieren zu können, müssen zudem neue Auswertungsroutinen entwickelt werden, sowie ein geeignetes Geheimhaltungsverfahren und Wege der Datenbereitstellung für die Wissenschaft abgestimmt werden.

96 Weiterentwicklung und Evaluierung der Erfassung non-formaler Weiterbildungsaktivitäten im Mikrozensus

Maximilian Bach (1), Jonathan Kohl (2)

1: Statistisches Bundesamt, Deutschland; 2: Deutsches Institut für Erwachsenenbildung
maximilian.bach@destatis.de

SECTION: Methodology of Statistical Surveys (Session 3)

Die Erfassung non-formaler Weiterbildungsaktivitäten im Rahmen großangelegter Mehrthemenbefragungen wie dem Mikrozensus ist mit erheblichen Herausforderungen verbunden. Aufgrund des begrenzten Frageumfangs haben Befragte oft einen großen Ermessensspielraum, welche Aktivitäten sie non-formaler Weiterbildung zuordnen. Insbesondere bei kürzeren Weiterbildungsaktivitäten, wie einstündigen Online-Kursen, führen längere Referenzzeiträume zudem häufig zu Erinnerungslücken. Beide Faktoren können die systematische Untererfassung der tatsächlichen Weiterbildungsbeteiligung zur Folge haben.

Vor diesem Hintergrund wurde 2023 das bestehende Fragekonzept im Mikrozensus grundlegend überarbeitet, um die Erfassung der non-formalen Weiterbildungsteilnahme zu optimieren und damit die Datenqualität für die Bildungsberichterstattung zu verbessern. Die wesentlichen Neuerungen umfassen eine neu strukturierte Abfolge der Referenzzeiträume – zunächst vier Wochen, anschließend zwölf Monate – sowie eine präzisere und eindeutige Frageformulierung, die in enger

Zusammenarbeit mit nationalen Experten des Deutschen Instituts für Erwachsenenbildung (DIE) entwickelt wurde.

Zur Evaluation der Wirksamkeit des neuen Fragekonzepts kam erstmals im Mikrozensus im dritten Quartal 2023 ein Split-Ballot-Verfahren zum Einsatz. Dabei wurde eine Unterstichprobe der CAWI-Befragten randomisiert in eine Kontrollgruppe (altes Frageformat) und eine Experimentalgruppe (neues Frageformat) unterteilt. Die Ergebnisse dieses Verfahrens zeigen einen deutlichen Anstieg der gemessenen Weiterbildungsbeteiligung von 16 % auf 21 % im Vergleich zum alten Frageformat. Auf Basis dieser positiven Ergebnisse wurde das neue Fragekonzept 2024 flächendeckend implementiert. Erste Auswertungen des Berichtsjahres 2024 bestätigen die Verbesserungen im CAWI-Modus, während sich in den assistierten Erhebungsmodi (CATI und CAPI) keine vergleichbaren Effekte zeigen. Diese Ergebnisse deuten darauf hin, dass das Split-Ballot-Verfahren ein vielversprechendes Instrument zur Evaluierung von Messinstrumenten darstellt, wenngleich die Übertragbarkeit der im CAWI-Modus gewonnenen Erkenntnisse auf andere Erhebungsmodi begrenzt zu sein scheint.

97 Zensus 2031: Methodentest Bevölkerung

Ewa-Juditha Wójcik

Statistisches Bundesamt, Deutschland
ewa-juditha.wojcik@destatis.de

SECTION: Methodology of Statistical Surveys (Session 3)

In diesem Beitrag stellen wir die Ziele, Methoden und erste Erkenntnisse zum Methodentest im Bevölkerungsmodul gemäß Registerzensuserprobungsgesetz als wichtigen Schritt für einen registerbasierten Zensus ab 2031 vor. Zentrales Element zur Qualitätssicherung ist der sogenannte Lebenszeichenansatz. Die hohen Ansprüche an die Ergebnisqualität bei der Einwohnerzahlermittlung stellen eine besondere Herausforderung dar. Um den Lebenszeichenansatz zu überprüfen, erfolgt ein Abgleich mit den Ergebnissen der bereits durchgeführten Personenerhebungen des Zensus 2022. Die Bewertung der Qualität der Vergleichsregister erfolgt u.a. durch Kennzahlen zur Bewertung der Datenqualität und einen Abgleich mit Referenzdaten aus dem Zensus 2022. Ziel des Methodentests ist es, Aussagen zur grundsätzlichen Eignung der Methoden zu treffen, um im Zensus 2031 realitätsgerechte Bevölkerungszahlen auf den Ebenen Bund, Länder und Gemeinden ermitteln zu können.

98 Market Shocks: Counting Processes with Restarting Property

Ulrich Müller-Funk (1), Christina Ungerer (2)

1: Universität Münster, Deutschland; 2: HTWG Konstanz
funk@wiwi.uni-muenster.de

SECTION: Statistical Theory and Methods (Session 1)

Sales figures and other business characteristics do not always develop smoothly over time but are also subject to disruptive influences from individual market-changing events. These include the entry or exit of market participants as well as market-regulating political measures. In the analysis, such shocks must be reflected in the underlying stochastic model. A comparable situation occurs, for example, in reliability theory. From there, the idea of a process with the "restarting property" will be adopted here. If the process is restarted at the time of such an event, the type of process is preserved, and the shock then results in a jump in the parameter space. This avoids a

complete reboot, which would neglect the previous history. The counting process proposed here is motivated by the Bass model from product diffusion theory. However, the approach should not only be viewed in this context; it can also be framed economically in other ways. It can be used, for example, to forecast various opportunities and threats for a company. For the presentation we choose the formulation in discrete time.

99 Measuring Dependence between Events

Marc-Oliver Pohle (1), Timo Dimitriadis (2,1), Jan-Lukas Wermuth (2)

1: Heidelberg Institute for Theoretical Studies, Germany; 2: Goethe University Frankfurt, Germany
pohle@kit.de

SECTION: Statistical Theory and Methods (Session 1)

Measuring dependence between two events, or equivalently between two binary random variables, amounts to expressing the dependence structure inherent in a 2×2 contingency table in a real number between -1 and 1 . Countless such dependence measures exist, but there is little theoretical guidance on how they compare and on their advantages and shortcomings. Thus, practitioners might be overwhelmed by the problem of choosing a suitable measure. We provide a set of natural desirable properties that a proper dependence measure should fulfill. We show that Yule's Q and the little-known Cole coefficient are proper, while the most widely-used measures, the phi coefficient and all contingency coefficients, are improper. They have a severe attainability problem, that is, even under perfect dependence they can be very far away from -1 and 1 , and often differ substantially from the proper measures in that they understate strength of dependence. The structural reason is that these are measures for equality of events rather than of dependence. We derive the (in some instances non-standard) limiting distributions of the measures and illustrate how asymptotically valid confidence intervals can be constructed. In a case study on drug consumption we demonstrate how misleading conclusions may arise from the use of improper dependence measures.

100 Proper Correlation Coefficients for Nominal Random Variables

Jan-Lukas Wermuth

Goethe Universität Frankfurt, Deutschland
wermuth@econ.uni-frankfurt.de

SECTION: Statistical Theory and Methods (Session 1)

I develop an intuitive concept of perfect dependence between two variables of which at least one has a nominal scale that is attainable for all marginal distributions and propose a set of dependence measures that are 1 if and only if this perfect dependence is satisfied. The advantages of these dependence measures relative to classical dependence measures like contingency coefficients, Goodman-Kruskal's lambda and tau and the so-called uncertainty coefficient are twofold. Firstly, they are defined if one of the variables is real-valued and exhibits continuities. Secondly, they satisfy the property of attainability. That is, they can take all values in the interval $[0,1]$ irrespective of the marginals involved. Both properties are not shared by the classical dependence measures which need two discrete marginal distributions and can in some situations yield values close to 0 even though the dependence is strong or even perfect.

Additionally, I provide a consistent estimator for one of the new dependence measures together with its asymptotic distribution under independence as well as in the general case. This allows to construct confidence intervals and an independence test, whose finite sample performance I subsequently examine in a simulation study. Finally, I illustrate the use of the new dependence measure in two applications on the dependence between the variables country and income or country and religion, respectively.

101 Tobit models for count time series

Christian Weiß (1), Fukang Zhu (2), Hee-Young Kim (3)

1: Helmut Schmidt University, Deutschland; 2: Jilin University, Changchun, China; 3: Korea University, Sejong, South Korea
weissc@hsu-hh.de

SECTION: Statistical Theory and Methods (Session 1)

Several models for count time series have been developed during the last decades, often inspired by traditional autoregressive moving average (ARMA) models for real-valued time series, including integer-valued ARMA (INARMA) and integer-valued generalized autoregressive conditional heteroscedasticity (INGARCH) models. Both INARMA and INGARCH models exhibit an ARMA-like autocorrelation function (ACF). To achieve negative ACF values within the class of INGARCH models, log and softplus link functions are suggested in the literature, where the softplus approach leads to conditional linearity in good approximation (DOI: 10.5705/ss.202020.0353). However, the softplus approach is limited to the INGARCH family for unbounded counts, that is, it can neither be used for bounded counts, nor for count processes from the INARMA family.

In this talk, we present an alternative solution, named the Tobit approach, for achieving approximate linearity together with negative ACF values, which is more generally applicable than the softplus approach. The main part of the talk studies a Skellam–Tobit INGARCH model for unbounded counts in detail (DOI: 10.1111/sjos.12751), including stationarity, approximate computation of moments, maximum likelihood and censored least absolute deviations estimation for unknown parameters and corresponding simulations. The practical relevance of the Skellam–Tobit INGARCH model is illustrated by real-world data examples on lottery winners, yields from a chemical process, and air quality in Beijing. The talk concludes with a presentation of some work in progress, where the Tobit approach is used to define (unbounded or bounded) INARMA-type models. By contrast to existing INARMA models for count time series, the proposed Tobit INARMA models allow for negative ACF values while showing the typical ARMA properties in close approximation.

102 Regionale Energiewende: Dashboard zur Steuerung der sozial-ökologischen Transformation

Katharina Hembach-Stunden, Britta Stöver, Philip Ulrich

Gesellschaft für Wirtschaftliche Strukturforchung (GWS), Osnabrück
hembach-stunden@gws-os.com

SECTION: VDSSt (Session 3): Regionale Energie- und Gebäudedaten

Die sozial-ökologische Transformation hin zu Klimaneutralität wird auf regionaler Ebene umgesetzt. Kommunen, Kreise und Städte stehen vor der Herausforderung, den Anforderungen an die

Energie- und Wärmewende gerecht zu werden und eine für ihre Bevölkerung und Unternehmen nachvollziehbare und sozial tragbare Ausgestaltung zu entwickeln. Ein Überblick über mögliche sozioökonomische Folgen unterschiedlicher Maßnahmen-Optionen mit regionalem Fokus ist für die Gestaltung dieses Transformationsprozesses zentral.

Das interdisziplinäre Projekt Info-EW (www.info-ew.de) unterstützt die regionalen Akteur:innen bei der sozial-ökologischen Transformation durch die Entwicklung eines datengestützten Informationssystems zur Bewertung der sozialen und wirtschaftlichen Auswirkungen der Energiewende und des Strukturwandels vor Ort. Das im Rahmen des Förderprogramms „Energiewende und Gesellschaft“ geförderte Projekt kombiniert das neue makroökonomische multiregionale Input-Output-Modell RIMES mit Haushaltsmikrosimulationen auf Kreisebene und ermöglicht so regionalspezifische Szenarioanalysen der Energiewende.

Am Beispiel des Ausbaus von Wärmepumpen und Fernwärme, zwei zentralen Elementen der Wärmewende, demonstrieren wir, wie im Dashboard integrierte Szenarien genutzt werden können, um Effekte unterschiedlicher Strategien auf Kreisebene zu analysieren. Das Wärmewende-Szenario quantifiziert die Auswirkungen des Austauschs fossiler Heizsysteme sowie des Ausbaus grüner Fernwärme unter Berücksichtigung investitionsbezogener Kenngrößen wie der Anzahl neu zu installierender Wärmepumpen, dem Fernwärme-Ausbaupfad und den damit verbundenen Kosten. Die Analyse erfolgt unter Einbeziehung regionalstatistischer Merkmale wie Wohngebäudestruktur, Heizsystemverteilung und Eigentumsverhältnissen, um Umsetzungshemmnisse und -potenziale systematisch zu erfassen. Dadurch wird eine fundierte Einschätzung regional differenzierter Transformationspfade ermöglicht.

Im Ergebnis unterstützt das Dashboard bei der Bewertung unterschiedlicher Maßnahmen auf Kreisebene. Es erfasst die Herausforderungen vor Ort und zeigt regionale Hemmnisse und Chancen der Transformation auf. Dies ermöglicht die Einordnung von Handlungsoptionen, und bietet darüberhinaus eine Argumentationsbasis, um regionale Herausforderungen gezielt anzugehen, von Best Practices anderer Regionen zu lernen und maßgeschneiderte Lösungsansätze vor Ort zu entwickeln.

103 Erarbeitung eines Gebäudekatasters für Wolfsburg im Kontext der Kommunalen Wärmeplanung

Jan Lunge

Stadt Wolfsburg, Deutschland
jan.lunge@stadt.wolfsburg.de

SECTION: VDSSt (Session 3): Regionale Energie- und Gebäudedaten

Ein Baustein des Projekts Kommunale Wärmeplanung ist die Bestandsanalyse innerhalb einer Kommune. Dabei sollen die aktuellen Energieverbräuche und -bedarfe gebäudescharf erfasst und analysiert werden. Die Statistikstelle der Stadt Wolfsburg hat diesen Projektbaustein umgesetzt und berichtet von ihren Erfahrungen bei der Erstellung eines Gebäudekatasters und der Bestandsanalyse.

104 Validierung von Small Area-Methoden für den systematischen Einsatz in der amtlichen Statistik am Beispiel der Schätzung von Bestandsmieten auf Gemeindeebene in Nordrhein-Westfalen

Amelie Plöger, Elena Stäger, Christoph Alfken, Marcel Noack, Johannes Rohde

IT.NRW - Statistisches Landesamt, Deutschland
johannes.rohde@it.nrw.de

SECTION: VDSSt (Session 3): Regionale Energie- und Gebäudedaten

Das Thema „Wohnen und Mieten“ wurde in den vergangenen Jahren häufig kontrovers diskutiert. Zuverlässige Daten stehen für die oft stark von lokalen Faktoren geprägten Mietmärkte allerdings selten flächendeckend bzw. in ausreichender Qualität zur Verfügung. Insbesondere fehlt es an einheitlichen und flächendeckenden Daten zu Bestandsmieten. Hier setzen Small Area-Methoden an: Diese bieten das Potenzial, verlässliche Ergebnisse – trotz geringerer Stichprobenumfänge – auch auf kleinräumiger Ebene veröffentlichen zu können, indem die Datenbasis durch externe Hilfsmerkmale angereichert und die Schätzungen auf diese Weise stabilisiert werden.

Erste Projekte zur Schätzung von durchschnittlichen Bestandsmieten auf Gemeindeebene mittels dieser Methodenklasse wurden anhand der Daten des Mikrozensus 2018 in Kooperation mit dem Statistischen Bundesamt und der Universität Trier sowie Statistik Nord und dem Landesamt für Statistik Niedersachsen durchgeführt. Hierbei hat sich gezeigt, dass Small Area-Verfahren valide Ergebnisse auf kleinräumiger Ebene ermöglichen und somit das Veröffentlichungsportfolio der amtlichen Statistik bereichern können.

Mit den nun vorliegenden Daten des Mikrozensus 2022 wird eine erneute Schätzung der Brutto- und Nettobestandsmieten auf Ebene der Gemeinden in Nordrhein-Westfalen durchgeführt. Die Ergebnisse werden genutzt, um einerseits die Entwicklung der durchschnittlichen Mietpreise zur letzten Schätzung zu beurteilen. Zum anderen erfolgt ein Vergleich der Small Area-Schätzungen mit den Ergebnissen aus der Gebäude- und Wohnungszählung des Zensus 2022. Somit ist erstmals eine Validierung der Ergebnisse anhand von vergleichbaren Daten möglich. Darüber hinaus werden aus den bisherigen Erfahrungen mit Small Area-Methoden erste Empfehlungen und Best Practices für einen künftig möglichen systematischen Einsatz in bestimmten Themenfeldern der amtlichen Statistik abgeleitet.

105 Arbeitsmarktanspannung: Wenn Wohnortwechsel und Pendeln nicht mehr reichen

Jonas Krinitz (1), Johannah Zenk (2), Christian Schneemann (2), Jan-Phillip Schroer (3)

1: Gesellschaft Wirtschaftliche Strukturforchung (GWS); 2: Institut für Arbeitsmarkt- und Berufsforschung (IAB);
3: Bundesinstitut für Berufsbildung (BIBB)
krinitz@gws-os.com

SECTION: Economic, Social and Market Statistics (Session 6): Labour Markets and Social Security II

Ein akuter Arbeitskräfteengpass tritt, sowie Arbeitslosigkeit immer lokal auf und kann sich regional unterschiedlich, mitunter strukturell, entfalten oder ballen. Beeinflusst wird dieses durch unterschiedliche Formen der Bewegung des Arbeitskräfteangebots. Zur Bewertung des Prozesses

dieser Koordinationsbewegung des Arbeitsmarktes, bzw. zur Identifikation von Engpässen ist es wichtig zu verstehen welche Formen von Arbeitsmobilität es gibt, wie sich diese über die Jahre entwickelt haben und was die individuellen Formen von Mobilität beeinflusst.

Mithilfe unterschiedlicher Arbeitsnachfrage- und Angebotsindikatoren entwickeln wir ein detailliertes Bild über die aktuellen Ströme und Hemmnisse der Arbeitskräftemobilität. Zudem wird ein kontrafaktisches Modell ohne Binnenwanderung entwickelt, um sowohl die regionale Bedeutung von räumlicher Mobilität heute als nach 25 Jahren ohne Binnenwanderung zu vergleichen. Darauf aufbauend gehen wir der Frage nach, ob und wo es aktuell ungenutzte regionale Arbeitspotenziale gibt und inwiefern räumliche Mobilität in dem zuletzt angespannten Arbeitsmarkt eine Entlastung bieten kann.

106 Branchenwechsel von Beschäftigten

Michael Hartmann

Statistik der Bundesagentur für Arbeit, Deutschland
Michael.Hartmann@arbeitsagentur.de

SECTION: Economic, Social and Market Statistics (Session 6): Labour Markets and Social Security II

Ende 2024 hat die Statistik der BA ihre Berichterstattung um die Branchenmobilität von Beschäftigten erweitert. Mit der neuen Mobilitätsstatistik können nun auch die Übergänge zwischen den Branchen dargestellt werden, also z.B. in welchem Umfang neue Arbeitskräfte aus anderen Branchen rekrutiert werden bzw. Arbeitskräfte in andere Branchen wechseln. Das ermöglicht es insbesondere, Fragen im Zusammenhang mit der Transformation der Wirtschaft u. a. infolge der Digitalisierung und Dekarbonisierung tiefer zu beleuchten. Der Vortrag erläutert das Messkonzept und stellt die Analysemöglichkeiten vor, die die neuen Mobilitätsdaten bieten.

107 The Employment Statistics of Severely Disabled People: Description and Research Potential

Karolin Hiesinger, Laura Pohlan, Franka Vetter

Institut für Arbeitsmarkt- und Berufsforschung (IAB), Deutschland
Laura.Pohlan@iab.de

SECTION: Economic, Social and Market Statistics (Session 6): Labour Markets and Social Security II

Falling birth rates combined with an ageing population pose major challenges for industrialized countries in terms of the future working population, economic productivity and the long-term sustainability of social security systems. Germany is particularly affected by demographic change: Between 2001 and 2024, the proportion of employees subject to social insurance contributions and over the age of 55 rose significantly from 9.7 percent to 24.0 percent (Federal Employment Agency, 2024).

Ageing is often accompanied by age-related illnesses that have enduring effects. In addition, the proportion of people with (severe) disabilities increases dramatically with age, as disabilities usually develop in the course of life. In view of an ageing society and the current shortage of skilled workers, which is likely to worsen in the coming years, promoting the integration and maintaining the employability of employees despite disability is of central importance.

The literature on the situation of severely disabled people in the labor market has so far made little use of administrative data. We present the Employment Statistics of Severely Disabled People (BsbM), a unique administrative database of the German Federal Employment Agency, which is collected as part of the process of administering firm compliance with the disabled worker quota. Our aim is to provide an overview of the BsbM data and its research potential. In particular, we examine the development of firms' compliance with the disability quota over time and identify factors associated with a higher likelihood of compliance. We also discuss research opportunities offered by the dataset and its potential for further refinement.

108 Häuser oder Wohnungen: Kann der Neubau seniorengerechter Wohnungen den Neubedarf an EFH reduzieren?

Christian von Malottki

BPD Immobilienentwicklung - Research
c.malottki@bpd-de.de

SECTION: DGD (Session 2): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen II

Deutschlands Haushaltszahl wächst und verursacht deshalb weiteren Wohnungsbedarf. Gleichzeitig sinkt die durchschnittliche Haushaltsgröße, so dass gemessen in Quadratmetern eigentlich genügend Wohnfläche da wäre. Ältere ziehen aber kaum um, belegen als kleine Haushalte die großen Wohnungen und Häuser, ziehen den Durchschnitt der Wohnfläche pro Person nach oben und belegen gleichzeitig flächenintensive Wohnformen, die Familienhaushalte in Befragungen als ihre gewünschte Wohnform angeben. Auch wenn hier gelegentlich politische „Zwangsmaßnahmen“ gefordert werden, so muss davon ausgegangen werden, dass diese nicht durchsetzbar sind. Soll die Deckung des Wohnungsbedarfs nicht zu Lasten der Lebensqualität der jüngeren Generation gehen, so müsste gemäß einem klassischen Lebenszyklusmodell der zusätzliche Wohnungsbedarf weiterhin durch den Zubau von EFH für 35-Jährige gedeckt werden. Unter dem Gesichtspunkt des sparsamen Umgangs mit Grund und Boden erscheint das ebenso wenig sinnvoll wie unter dem Aspekt der Eignung des Wohnungsbestands für seniorengerechtes Wohnen. Eine der zentralen Strategien der Wohnungswirtschaft sollte deshalb in den nächsten Jahren darin bestehen, durch attraktiven seniorengerechten Wohnraum als angebotsseitigen Stimulus die Umzugsquote Älterer zu erhöhen. So erscheint in der nächsten Kohorte auch die Umzugsbereitschaft höher als früher. Die Frage ist nun, in welchem quantitativen Verhältnis die beiden Wohnprodukte „Seniorengerechter Geschosswohnungsbau“ und „Häuser für Familien“ zueinander stehen sollen und wie sich dieses Verhältnis je nach Änderung der Eingabeparameter (Umzugsquote Älterer, Bildung von Familien bei Jüngeren, demographische Parameter) verändert. Entscheidender Hebel ist hier die Menge an „vorzeitig freigezogenen marktfähigen EFH“. Basierend auf Strukturdaten des Mikrozensus und Stromgrößen der BPD-Wohnwunschbefragung wird so ein kleines Modell präsentiert, welche die Auswirkungen von Änderungen der Umzugsquote Älterer auf das Angebot an Bestands-EFH und damit das ideale Verhältnis von Häusern und seniorengerechten Wohnungen im Neubau zeigt. Diskutiert werden die Chancen und Herausforderungen der Regionalisierung sowie die Prioritätensetzung bei der Wahl der relevanten Parameter für ein möglichst schlankes Modell.

109 (Nicht-)Erwerbsbeteiligung im Kontext der demografischen Schrumpfung: ein Blick auf die Datenlage

Timon Hellwagner

Institut für Arbeitsmarkt- und Berufsforschung (IAB), Deutschland
timon.hellwagner@iab.de

SECTION: DGD (Session 2): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen II

Das Erwerbspersonenpotenzial auf dem deutschen Arbeitsmarkt ist seit Beginn der 2010er Jahre beständig gewachsen, steht nun aber vor einem demografisch bedingten Wendepunkt. In der mittleren und langen Frist ist von einem Rückgang des Arbeitskräfteangebots auszugehen, wenngleich das Ausmaß dieses Rückgangs mit entsprechenden Maßnahmen abgeschwächt werden kann. Neben höheren Nettozuzügen können auch steigende Beteiligungsquoten dazu beitragen, die demografisch bedingte Schrumpfung abzufedern. Dabei hängt die mögliche Hebung ungenutzter Potenziale unter der bereits in Deutschland ansässigen Bevölkerung wesentlich vom Ausmaß der (Nicht-)Erwerbsbeteiligung ab. Mit den auf dem Mikrozensus basierenden Erwerbsquoten stehen konsistente Informationen zur Verfügung, die sämtliche Erwerbsformen abdecken – gleichzeitig aber das Ausmaß der Erwerbstätigkeit etwa im Vergleich zur Erwerbstätigenrechnung im Rahmen der Volkswirtschaftlichen Gesamtrechnung deutlich unterschätzen. Außerdem enthalten die Zeitreihen Strukturbrüche, wie zuletzt in der jüngeren Vergangenheit. Die Kombination beider Aspekte erschwert zum einen die Beurteilung verbleibender Potenziale, zum anderen die Prognose künftiger Beteiligungsquoten und damit Einblicke, inwiefern verbleibende Potenziale gehoben werden (könnten). Vor diesem Hintergrund stellt der Vortrag darauf ab, diese Probleme nicht nur aufzuzeigen, sondern die Informationen aus dem Mikrozensus mit weiteren Datenquellen zu kombinieren und darauf basierende Beteiligungsquoten – nach Staatsangehörigkeit, Geschlecht und Altersgruppen disaggregiert – zu präsentieren, kritisch zu diskutieren und verbleibende Potenziale im Kontext der demografischen Schrumpfung einzuordnen.

110 Modellierung des Arbeitsangebotes von Gesundheitsberufen für ein Fachkräftemonitoring des Bundesministeriums für Gesundheit

Jan Philipp Schroer, Alexander Schur

Bundesinstitut für Berufsbildung, Deutschland
jan-philipp.schroer@bibb.de

SECTION: DGD (Session 2): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen II

Der Bedarf an Fachkräften in den Gesundheitsberufen nimmt aufgrund der demografischen Entwicklung sowie geänderter Verhaltensweisen seit Jahren zu und wird voraussichtlich auch in den kommenden Jahren weiter steigen. Wie können betroffene Berufe und wenn möglich, die von den Berufen erbrachten Gesundheitsleistungen in einer möglichst kleinteiligen Differenzierung im Gesundheitsbereich erfasst werden und wie können diese notwendigen Bedarfe befriedigt werden? Hierzu muss das Gesundheitssystem mit Fachkräften in der richtigen Anzahl, mit der richtigen Qualifikation, am richtigen Ort und in der richtigen Zusammensetzung ausgestattet werden.

Um mögliche Fachkräfteengpässe zu lösen, müssen sie frühzeitig erkannt werden. Hier setzt das 2024 vom Bundesministerium für Gesundheit (BMG) beauftragte BMG-Fachkräftemonitoring an. Die Modellierung einer Basisprojektion für das BMG-Fachkräftemonitoring geschieht im Rahmen der Qualifikations- und Berufsprojektionen (QuBe-Projekt), welche unter der gemeinsamen Leitung des Bundesinstituts für Berufsbildung (BIBB) und des Instituts für Arbeitsmarkt und Berufsforschung (IAB) in Zusammenarbeit mit der Gesellschaft für Wirtschaftliche Strukturforchung (GWS) durchgeführt werden.

Dabei ist das Modell in eine Arbeitsangebots- und Arbeitsnachfrageseite unterteilt. Während sich erstere der Anzahl zukünftig verfügbarer Erwerbspersonen widmet, lässt letztere u.a. Rückschlüsse auf die Entwicklung des Versorgungsbedarfs nach Einrichtungsarten zu. Auf die Angebotsseite soll im Folgenden detaillierter eingegangen werden. Für die Modellierung der Angebotsseite stellt die Gesundheitspersonalrechnung (GPR) des Statistischen Bundesamtes die wichtigste Datenquelle dar. Ihr werden die Erwerbstätigendaten der im Gesundheitswesen Beschäftigten entnommen und durch Daten zu erwerbstätigen Fachärzten und -ärztinnen aus der Ärztestatistik der Bundesärztekammer sowie aus dem Bundesarztregister und dem Honorarbericht der Kassenärztlichen Bundesvereinigung ergänzt. In Abstimmung mit dem BMG wurden 55 Berufe des Gesundheitswesens, darunter 16 ärztliche Berufe, identifiziert, die nach Personen und Vollzeitäquivalenten sowie Einrichtungsarten bis 2050 fortgeschrieben werden. Im Vergleich mit der Basisprojektion können Sensitivitätsanalysen und Szenarienrechnungen durchgeführt werden, die eine Abschätzung von Folgewirkungen von Maßnahmen zur Behebung von Passungsproblemen erlauben.

111 QuBe - Projektion des zukünftigen Arbeitsangebots bis 2040

Michael Kalinowski

Bundesinstitut für Berufsbildung, Deutschland
kalinowski@bibb.de

SECTION: DGD (Session 2): Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen II

Zur fundierten Projektion des zukünftigen Arbeitskräfteangebots nach Qualifikationsniveaus und Berufsfeldern ist eine detaillierte Modellierung des Bildungs- und Migrationsgeschehens notwendig. Im Rahmen der BIBB-IAB-Qualifikations- und Berufsprojektionen (QuBe) wurde hierzu ein Angebotsmodell entwickelt, das auf der Bevölkerungsprojektion des IAB basiert.

Das Modell unterscheidet zwischen Personen im Bildungssystem und jenen außerhalb und bildet Bildungswege, Abschlüsse, demografische Entwicklungen sowie Migration differenziert ab. Das sogenannte „berufliche Bildungssystem“ modelliert die Übergänge von Ausbildungseinrichtungen in den Arbeitsmarkt unter Berücksichtigung von Erfolgsquoten und Übergangswahrscheinlichkeiten. Die Migrationskomponente basiert primär auf Daten des Mikrozensus, wobei Qualifikationsstrukturen von Zu- und Fortzügen sowie Einbürgerungen modelliert werden.

Für die Projektion der Bestände im Bildungssystem erfolgt eine Vorausberechnung der Schülerinnen- und Schülerzahlen sowie der Studierenden. Die daraus resultierenden Absolventenzahlen werden unter Verwendung von Erfolgsquoten ermittelt.

Das Kohorten-Komponenten-Modell erlaubt eine differenzierte Betrachtung der demografischen und qualifikatorischen Entwicklung der Erwerbsbevölkerung, unterteilt nach Deutschen und Nicht-deutschen, Geschlecht, Alter, Beruf und Bildungsabschluss.

Insgesamt ermöglicht das Modell eine Darstellung der jährlichen Zu- und Abgänge am Arbeitsmarkt, einschließlich der Beiträge aus dem Bildungssystem und der Migration und liefert eine empirische

Grundlage zur Bewertung zukünftiger Fachkräftepotenziale und unterstützt langfristige Planungen in Arbeitsmarkt-, Bildungs- und Migrationspolitik.

112 Robust Statistical Decisions: With Application to Development Economics

Jörg Stoye

Cornell University, USA
stoye@cornell.edu

SECTION: Empirical Economics and Applied Econometrics (Session 2)

This talk summarizes work with coauthors (<https://arxiv.org/abs/2312.17623>, <https://arxiv.org/abs/2408.09187>, https://joseluismontielolea.com/epsilon_minimax_v2.pdf) that brings statistical decision theory to problems where robustness, e.g. to model uncertainty or limitations of external validity, is a major concern. The language used is that of partial identification; the application discussed is to optimal treatment choice and optimal experimental design in development economics. Part of the contribution is computational: Minimax decision rules can be hard to discover, and we leverage recent progress in computer science toward automated such discovery.

113 Covariate Balancing and the Equivalence of Weighting and Doubly Robust Estimators of Average Treatment Effects

Tymon Słoczyński (2), Derya Uysal (1), Jeffrey Wooldridge (3)

1: LMU Munich, Germany; 2: Brandeis University; 3: Michigan State University
derya.uysal@econ.lmu.de

SECTION: Empirical Economics and Applied Econometrics (Session 2)

We show that when the propensity score is estimated using a suitable covariate balancing procedure, the commonly used inverse probability weighting (IPW) estimator, augmented inverse probability weighting (AIPW) with linear conditional mean, and inverse probability weighted regression adjustment (IPWRA) with linear conditional mean are all numerically the same for estimating the average treatment effect (ATE) or the average treatment effect on the treated (ATT). Further, suitably chosen covariate balancing weights are automatically normalized, which means that normalized and unnormalized versions of IPW and AIPW are identical. For estimating the ATE, the weights that achieve the algebraic equivalence of IPW, AIPW, and IPWRA are based on propensity scores estimated using the inverse probability tilting (IPT) method of Graham, Pinto and Egel (2012). For the ATT, the weights are obtained using the covariate balancing propensity score (CBPS) method developed in Imai and Ratkovic (2014). These equivalences also make covariate balancing methods attractive when the treatment is confounded and one is interested in the local average treatment effect.

114 Heterogeneous net treatment effects

Lennard Maßmann (1,3), Eva-Maria Oeß (2)

1: Universität Duisburg-Essen, Deutschland; 2: Universität zu Köln, Deutschland; 3: Ruhr Graduate School in Economics, Essen, Deutschland
lennard.massmann@uni-due.de

SECTION: Empirical Economics and Applied Econometrics (Session 2)

We introduce a novel methodology for estimating heterogeneous net treatment effects under unit-varying treatment and cost effects. Our approach is designed for optimal assignment of a binary treatment that induces a cost-benefit trade-off: First, it enables identification of the target population, for which the treatment effect is larger than or equal to the cost effect. Second, it allows for direct prioritization of the treatment assignment via the effect size. Using a generalized random forest, we minimize a joint loss function based on the local difference between the two effects. We formally show that our approach achieves a lower mean squared error compared to separate effect estimation and subsequent differencing, if the treatment and cost effects are correlated. In a simulation study, we confirm these findings for finite samples. Additionally, we discuss two empirical applications. In the first example, we use semi-synthetic marketing data to evaluate customer conversion by balancing increased sales against discount offers. In the second example, we use data from a large nonprofit organization to analyze the net effect of a fundraising campaign to increase pledge payments while avoiding donor attrition.

115 Realized Covariance Modeling by Logarithmic Transformations and Dimensionality Reduction

Ilya Archakov (1), Roxana Halbleib (2), Jasper Rennspies (2)

1: York University, Canada; 2: University of Freiburg, Germany
jasper.rennspies@vwl.uni-freiburg.de

SECTION: Statistics in Finance (Session 2)

We propose an approach to model realized covariance matrices by transforming them and selecting only persistent series to model. The proposed models can capture the persistence and the common dynamics in the volatilities and the correlations, while ensuring the positive-definiteness of the covariance matrix and modeling parsimoniously to avoid the curse of dimensionality such that we can deal with a large cross-section of stocks. We use the γ -transformation of [archakov2021new] to ensure the positive-definiteness of the realized correlation matrix and construct models that exploit the empirical persistence and common dynamics in the γ -transformed realized correlations. In a forecasting setting we find that modeling a selection of less than 10

116 Estimation of Non-Gaussian Factors Using Higher-order Multi-cumulants in Weak Factor Models

Wanbo Lu (1), Kris Boudt (2,3,4), Guanglin Huang (1)

1: SWUFE; 2: Universiteit Gent, Belgium; 3: Vrije Universiteit Brussel; 4: Vrije Universiteit Amsterdam
kris.boudt@ugent.be

SECTION: Statistics in Finance (Session 2)

When factors are weak, covariance-based factor analysis methods tend to exhibit poor performance. To address this issue in the presence of non-Gaussian factors, we propose a novel approach that utilizes the eigenvalue decomposition of the product of higher-order multi-cumulant matrices and their transposes. We derive the asymptotic properties of this Higher-order multi-cumulant Factor Analysis (HFA) within the framework of weak factors that are non-Gaussian and where error terms are Gaussian. Simulation results demonstrate that HFA substantially enhances the accuracy of both factor selection and estimation when compared to traditional covariance-based methods. Additionally, we apply HFA to the FRED-MD dataset, detecting and estimating factors to improve forecasts of the monthly S&P 500 equity premium.

117 Combining Portfolio Rules to Improve Prediction of Global Minimum Variance Portfolio Weights

Vasyl Golosnoy (1), Bastian Gribisch (2), Wolfgang Schmid (3), Miriam Isabel Seifert (1)

1: Ruhr-Universität Bochum, Deutschland; 2: Universität zu Köln; 3: bastian.gribisch@statistik.uni-koeln.de
vasyl.golosnoy@rub.de

SECTION: Statistics in Finance (Session 2)

We consider the prediction of the global minimum variance portfolio (GMVP) weights based on realized covariance matrices computed from high-frequency intraday returns of risky assets. As the multivariate high-dimensional time series process of covariance matrices is rather complex and hard to estimate without substantial simplifications of the model structure, there exist various competing approaches for predicting the GMVP weights.

Our major contribution is the development of a novel approach for combining several given GMVP prediction rules in order to determine a low dimensional time-varying vector of these rules' proportions for the GMVP.

We provide statistical results on realized rule proportions and suggest a feasible low-dimensional approach to forecast the proportions based on a set of pre-determined GMVP prediction rules. Our findings are illustrated in an empirical study where we forecast the GMVP weights based on 265 risky assets by combining various popular portfolio rules.

118 Auf der Suche nach dem „Sweet Spot“ zwischen Nutzen, Risiko und Komfort bei der Nutzung amtlicher Daten – Ermöglicht partielle Datensynthetisierung die Anreicherung von Remote Access-Daten?

Yannik Garcia Ritz, Veronika Chakraverty

Statistisches Bundesamt, Deutschland
Yannik.GarciaRitz@destatis.de

SECTION: Methodology of Statistical Surveys (Session 4)

Formal anonymisierte Daten, bei denen direkte Identifikatoren und Hilfsmerkmale entfernt sind, bieten das größte Analysepotenzial für die Nutzung amtlicher Mikrodaten durch die unabhängige Wissenschaft. Der Zugang dazu ist wenig komfortabel, denn er erfolgt an Gastwissenschaftsarbeitsplätzen der Statistischen Ämter des Bundes und der Länder oder über kontrollierte Datenfernverarbeitung ohne Dateneinsicht. Am Institutsarbeitsplatz können Datennutzende ausschließlich mit faktisch anonymisierten Daten, wie Scientific Use Files (SUFs), arbeiten. Deonymisierung ist hier nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft möglich. Im Vergleich zu formal anonymisierten Daten sind jedoch strengere Anonymisierungsmethoden (bspw. Vergrößerung, Entfernung weiterer Variablen) erforderlich, da die Daten die Statistischen Ämter verlassen. Dementsprechend eingeschränkt ist das Analysepotential. Um Daten mit größerem Analysepotenzial „außer Haus“ anzubieten, haben die Forschungsdatenzentren der Statistischen Ämter kürzlich „Remote Access“ als alternativen Zugangsweg geschaffen. Faktische Anonymität wird hier durch erweiterte technisch-organisatorische und vertragliche Maßnahmen sowie Anonymisierungsmaßnahmen sichergestellt. Diese sind zwar weniger streng als bei Offsite-SUFs, führen jedoch ebenfalls zu methodischen Einschränkungen. Beispielsweise ist die Bereitstellung kleinräumiger oder sensibler Mikrodaten weiterhin nicht möglich. Folglich werden Remote-SUFs zwischen formal anonymen Daten und Offsite-SUFs eingeordnet. Die Suche nach dem „Sweet Spot“ zwischen Anonymität, Zugangskomfort und Datennutzen ist also mit dem neuen Zugangsweg keinesfalls beendet.

Die vorliegende Arbeit setzt hier an und untersucht CART-basierte partielle Datensynthetisierung als alternative/ergänzende Anonymisierungsmethode für Remote-SUFs. Ausgewählte sensible Variablen werden hierbei mit Hilfe von Imputationsverfahren durch künstliche Werte ersetzt, die statistische Zusammenhänge (auch zu nicht-synthetisierten Variablen) bewahren. Konkretes Ziel der Arbeit ist, für den Mikrozensus 2019 einen mit kleinräumigen und weiteren sensiblen Variablen angereicherten Remote-SUF zu erzeugen und trotzdem faktische Anonymität zu gewährleisten. Im Fokus stehen ein möglicher Anstieg des Nutzens, aber auch die Aufdeckungsrisiken im Vergleich zum bereits veröffentlichten Remote-SUF, bei dem traditionelle Anonymisierungsmethoden angewendet wurden. Neben bereits etablierten Metriken zur Evaluation des Kosten-Risiko-Verhältnisses werden zusätzlich neu entwickelte Kennzahlen (Raab et al., 2024) angewendet.

119 Modellierung des Wohnraumbestands im Mikrosimulationsmodell MikroSim

Sarah Bohnensteffen (1), Pascal Gläske (1), Ralf Münnich (2)

1: Statistisches Bundesamt, Deutschland; 2: Universität Trier
sarah.bohnensteffen@destatis.de

SECTION: Methodology of Statistical Surveys (Session 4)

Das Mikrosimulationsmodell MikroSim ermöglicht die dynamische Simulation kleinräumiger Entwicklungen von Haushalten und Individuen in Deutschland. Die Entwicklung der demografischen Struktur, der sozioökonomischen Situation der Haushalte sowie Außen- und Binnenwanderungsdynamiken sind dabei wichtige Treiber des lokalen Wohnraumbedarfs. Zur Realisierung eines Wohnraummoduls ist jedoch auch die Modellierung der Entwicklung des Wohnraumbestands unerlässlich, welcher die Möglichkeiten realisierter Wohnsituationen der Haushalte angebotsseitig begrenzt. Besonders in großen Städten ist die Abbildung der Wohnraumverfügbarkeit notwendige Voraussetzung für eine natürliche Begrenzung des Bevölkerungszuwachses im Simulationsmodell. In diesem Beitrag zeigen wir Möglichkeiten der Modellierung des Wohnraumbestands auf Basis von Daten der Gebäude- und Wohnungszählung in Kombination mit den Baustatistiken, aus denen die Fortschreibung des Gebäudebestands auf Gemeindeebene erstellt wird. Zur Modellierung der örtlichen Entwicklung des Wohngebäudebestands werden Verfahren Maschinellen Lernens auf Clustern mit klassischen Zeitreihenanalysen verglichen und die Verfügbarkeit relevanter regionaler Prädiktoren diskutiert.

120 Modellrekalibrierung für räumlich-dynamische Mikrosimulationen

Julian Ernst, Jan Weymeirsch, Ralf Münnich

Uni Trier, Deutschland
ernstj@uni-trier.de

SECTION: Methodology of Statistical Surveys (Session 4)

Mittels Mikrosimulationen können kleinräumige Bevölkerung mit einer Vielzahl an Charakteristiken wie Bildungsstand oder Haushaltskomposition fortgeschrieben werden. Zur Modellierung der demografischen Entwicklungen werden, wie in Makroansätzen, Übergangswahrscheinlichkeiten für Mortalität, Fertilität und Migration benötigt. Diese können zum einen auf Umfragedaten geschätzt werden, welche eine Abbildung von Populationsheterogenität über Alter und Geschlecht hinaus erlauben. Aufgrund von begrenzten Stichprobenumfängen, mangelnder regionaler Differenzierung sowie möglicher Verzerrungen bezüglich der interessierenden Variable kann dies zu unerwünschter und systematischer Abweichung von regional beobachteten Raten führt. Andererseits liegen aus der amtlichen Statistik regionale Aggregatdaten nach Alter und Geschlechtsgruppen vor, welche für die Schätzung verwendet werden könnten, aber keine Informationen über weitere Charakteristiken wie Bildung enthalten, welche die Mikrosimulation in erster Linie interessant machen. Mittels ex post-Anpassung des Übergangmodells durch ein Kalibrierungsmodell werden Modellwahrscheinlichkeiten mit bekannten Randverteilungen kombiniert, um regional plausible Übergänge bei gleichzeitigem Erhalt der gewünschten Modellkomplexität zu erzeugen. Der Beitrag

untersucht verschiedene Datenintegrationsstrategien und deren Anwendung im Kontext räumlicher dynamischer Mikrosimulationen.

121 Wie die Vergangenheit zählt. Zur Geschichte des Statistischen Bundesamtes (GeStat)

Kerstin Brückweh (1,2), Svetlana Burmistr (1), Paul Treffenfeldt (1,2)

1: Leibniz-Institut für Raumbezogene Sozialforschung, Deutschland; 2: Europa-Universität Viadrina Frankfurt (Oder)
paul.treffenfeldt@leibniz-irs.de

SECTION: Methodology of Statistical Surveys (Session 4)

Das Statistische Bundesamt ist die zentrale Behörde zur Produktion und Verbreitung quantitativen amtlichen Wissens über die deutsche Gesellschaft. Die Inhalte und Formen dieser quantitativen Beschreibung unterliegen hierbei ebenso einem historischen Wandel wie die Institution und die Akteur:innen, die dieses Wissen produzieren. Die Erforschung des historischen Wandels ist der Gegenstand eines Kooperationsprojekts zwischen dem Statistischen Bundesamt und dem Forschungsschwerpunkt „Zeitgeschichte und Archiv“ am Leibniz-Institut für Raumbezogene Sozialforschung in Erkner bei Berlin. Das Projekt widmet sich sowohl Fragen der „klassischen“ historischen Behördenforschung zu personellen und strukturellen Kontinuitäten zwischen Nationalsozialismus und Nachkriegszeit, als auch zeithistorischen Fragestellungen, die gegenwartsbezogen die Entwicklung der amtlichen Statistik und des Statistischen Bundesamtes zu ihrer heutigen Form untersuchen. Besonders bedeutend für diese jüngeren Entwicklungen ist einerseits die einschneidende methodische Neuorientierung des Statistischen Bundesamtes im Kontext der Volkszählungsproteste in den 1980er Jahren, die auch als Reaktion auf den Einsatz neuer Informationstechnologie verstanden werden kann. Andererseits prägen die Auswirkungen der deutschen Vereinigung und die europäische Integration bis heute die Formen der amtlichen statistischen Beschreibung und die personelle und strukturelle Ausstattung des Statistischen Bundesamtes.

Zur Untersuchung dieser historischen Fragestellungen erprobt das Forschungsprojekt einen innovativen kooperativen Zugang zur historischen Behörden- und Institutionenforschung, in welchem zum einen das spezifische Handlungs- und Erfahrungswissen der Mitarbeiter:innen für das Projekt nutzbar gemacht wird, und zum anderen ein reflexiv angelegtes Wissenschaftskommunikationskonzept implementiert wird, das den kritischen Austausch zwischen Mitarbeiter:innen und Forschenden auch schon während der Projektlaufzeit ermöglicht.

Im Paper werden die Anlage des Forschungsprojekts und die möglichen Potenziale einer solchen zeithistorischen Forschung zu Behörden und Institutionen näher beleuchtet. Es werden die Prozesse, die für eine Untersuchung des Statistischen Bundesamtes und der Amtlichen Statistik aus geschichtswissenschaftlicher Perspektive relevant sind, charakterisiert und der methodische und theoretische Zugriff auf die Thematik beschrieben.

122 Quantile Granger Causality in the Presence of Instability

Alexander Mayer (1), **Dominik Wied** (2), Victor Troster (3)

1: Università Ca' Foscari, Venice; 2: Universität zu Köln; 3: Universitat de les Illes Balears
dwied@uni-koeln.de

SECTION: Statistical Theory and Methods (Session 2)

We propose a new framework for assessing Granger causality in quantiles in unstable environments, for a fixed quantile or over a continuum of quantile levels. Our proposed test statistics are consistent against fixed alternatives, they have nontrivial power against local alternatives, and they are pivotal in certain important special cases. In addition, we show the validity of a bootstrap procedure when asymptotic distributions depend on nuisance parameters. Monte Carlo simulations reveal that the proposed test statistics have correct empirical size and high power, even in absence of structural breaks. Moreover, a procedure providing additional insight into the timing of Granger causal regimes based on our new tests is proposed. Finally, an empirical application in energy economics highlights the applicability of our method as the new tests provide stronger evidence of Granger causality.

123 Simultaneous Inference Bands for Autocorrelations

Uwe Hassler (1), Marc-Oliver Pohle (2), **Tanja Zahn** (1)

1: Goethe University Frankfurt, Germany; 2: Heidelberg Institute for Theoretical Studies, Germany
tzahn@wiwi.uni-frankfurt.de

SECTION: Statistical Theory and Methods (Session 2)

Sample autocorrelograms typically come with significance bands (non-rejection regions) for the null hypothesis of temporal independence. These bands have two shortcomings. First, they build on pointwise intervals and suffer from joint undercoverage (overrejection) under the null hypothesis. Second, if this null is clearly violated one would rather prefer to see confidence bands to quantify estimation uncertainty. We propose and discuss both simultaneous significance bands and simultaneous confidence bands for time series and series of regression residuals. They are as easy to construct as their pointwise counterparts and at the same time provide an intuitive and visual quantification of sampling uncertainty as well as valid statistical inference. For regression residuals, we show that for static regressions the asymptotic variances underlying the construction of the bands are as for observed time series and for dynamic regressions (with lagged endogenous regressors) we show how they need to be adjusted. We study theoretical properties of simultaneous significance bands and two types of simultaneous confidence bands (sup-t and Bonferroni) and analyse their finite-sample performance in a simulation study. Finally, we illustrate the use of the bands in an application to monthly US inflation and residuals from Phillips curve regressions.

124 A Perturbation Robust Test Against Spurious Long Memory

Vivien Less, Philipp Sibbertsen

Leibniz Universität Hannover, Deutschland
less@statistik.uni-hannover.de

SECTION: Statistical Theory and Methods (Session 2)

We propose a semiparametric score-type testing procedure to detect spurious long memory under a perturbed fractional framework. The test statistic is based on the weighted sum of the partial derivatives of the local Whittle with noise estimator. We show consistency of the test against the alternatives of smooth trend and random level shift processes. In addition, we derive the limiting distribution of the test. The finite sample properties of the test are examined in a Monte Carlo simulation study. An empirical example on the squared returns and the realised volatilities from the Verizon Communications stock is conducted, and shows the usefulness of the procedure

125 Kleinräumige Daten aus der Lohn- und Einkommensteuerstatistik – Potenziale für Analysen

Silke Joebges

Statistisches Amt, Landeshauptstadt München, Deutschland
silke.joebges@muenchen.de

SECTION: Regionalstatistik/VDSSt: Regionale Verteilung von Einkommen

Für die Landeshauptstadt München stehen erstmals untergemeindliche Daten aus der Lohn- und Einkommensteuerstatistik zur Verfügung. Auf der Ebene der 25 Stadtbezirke und 108 Stadtbezirksteile eröffnen sich damit neue Möglichkeiten zur kleinräumigen Analyse von Einkommensverhältnissen und sozialen Strukturen. In diesem Beitrag werden die Ergebnisse der Einkommensteuerstatistik für München vorgestellt und die bestehenden Herausforderungen und Limitationen bei der Nutzung skizziert.

126 Das Nettoäquivalenz-Resteinkommen als Sozialindikator für Miethaushalte

Andrea Schultz

Stadt Leipzig, Deutschland
andrea.schultz@leipzig.de

SECTION: Regionalstatistik/VDSSt: Regionale Verteilung von Einkommen

Angesichts zunehmend angespannter Wohnungsmärkte stellt sich die Frage, wie soziale Folgen für Miethaushalte gut abgebildet werden können. Dieser Vortrag untersucht die Eignung des Nettoäquivalenz-Resteinkommens als Indikator zur Messung der finanziellen Lage von Mieterinnen und Mietern. Dieser Indikator ermöglicht es, absolute Belastungssituationen präzise abzubilden. Im Gegensatz dazu kann die Mietbelastungsquote zu ungenauen Einschätzungen oder gar Fehleinschätzungen führen. Anhand der Städte München und Leipzig werden die Stärken und Schwächen beider

Indikatoren verdeutlicht. Ziel des Vortrags ist es, zu demonstrieren, dass das Nettoäquivalenz-Resteinkommen ein effektiver Indikator für Städte- und Regionalvergleiche ist, um die finanzielle Lage von Miethaushalten zu erfassen. Der Mikrozensus liefert alle vier Jahre die notwendigen Daten für die Berechnung. Diese Datenquelle ermöglicht es, das Nettoäquivalenz-Resteinkommen deutschlandweit auf Kreisebene oder regionalen Anpassungsschichten zu berechnen und für das Monitoring der Lebenslagen von Mieterinnen und Mietern zu nutzen. Der Vortrag soll für die Anwendung des Nettoäquivalenz-Resteinkommens als Kennwert für ein Monitoring der Lebenslagen von Mieterinnen und Mietern werben.

127 Das neue System of National Accounts 2025 - Neuerungen, Änderungen, Auswirkungen

Albert Braakmann

Statistisches Bundesamt (emeritus), Deutschland
alwies15@gmail.com

SECTION: Economic, Social and Market Statistics (Session 7): National Accounts, Welfare Measurement

Die aktualisierte, weltweit gültige Methodik zur Erstellung der Volkswirtschaftlichen Gesamtrechnungen, das System of National Accounts 2025, wurde von der Statistischen Kommission der Vereinten Nationen im März 2025 angenommen. Der Beitrag beleuchtet die wesentlichen konzeptionellen Neuerungen bzw. Änderungen, insbesondere im Hinblick auf die Auswirkungen für die Messung von Wirtschaftswachstum (BIP) und Staatsdefizit, und geht auf Hintergründe ein. Diese Neuerungen bzw. Änderungen bilden den Startpunkt für die derzeitige Überarbeitung des Europäischen Systems Volkswirtschaftlicher Gesamtrechnungen, das rechtsverbindlich von den Mitgliedstaaten der EU anzuwenden ist.

128 Strukturwandel in der Industrie aus unterschiedlichen Perspektiven: Produktionsindex der Konjunkturstatistik und Produktion in den Volkswirtschaftlichen Gesamtrechnungen

Stefan Linz

Statistisches Bundesamt, Deutschland
stefan.linz@destatis.de

SECTION: Economic, Social and Market Statistics (Session 7): National Accounts, Welfare Measurement

Seit Mitte der 2010er-Jahre verläuft die Entwicklung des Produktionswerts im Verarbeitenden Gewerbe laut Volkswirtschaftlicher Gesamtrechnungen (VGR) deutlich dynamischer als der monatlich ermittelte Produktionsindex der Konjunkturstatistik. Die wachsende Diskrepanz ist Ausdruck des Strukturwandels im Verarbeitenden Gewerbe.

Während der Produktionsindex primär die Entwicklung des Realwertes der Warenproduktion abbildet, berücksichtigt der Produktionswert in der VGR auch die zunehmenden nichtindustriellen Tätigkeiten der Industrieunternehmen. Zum einen werden die Produkte der Industrieunternehmen hybrider, da Waren verstärkt mit produktbegleitenden Dienstleistungen gebündelt werden. Andererseits verändern sich Geschäftsmodelle: Viele Industrieunternehmen lagern Teile ihrer Fertigung

aus und konzentrieren sich vermehrt auf produktbezogene Tätigkeiten wie Entwicklung, Design und Vermarktung. In der ausgeprägtesten Form treten sogenannte fabriklose Warenhersteller (Factoryless Goods Producers, FGPs) auf. Diese geben die Fertigung vollständig an Dritte ab, behalten jedoch die Kontrolle über immaterielle Vermögenswerte und damit über das Ergebnis des Produktionsprozesses.

Der Beitrag zeigt, wie sich die strukturellen Veränderungen – teils nur indirekt – in verschiedenen Indikatoren der amtlichen Statistik widerspiegeln und welche Herausforderungen bei der statistischen Abbildung der Produktionsentwicklung entstehen.

129 Erkenntnisse aus 5 Jahren Arbeit der Large Cases Unit

Klaus Pötzsch, Maximilian Rotsche

Statistisches Bundesamt, Deutschland
klaus.poetzsch@destatis.de

SECTION: Economic, Social and Market Statistics (Session 7): National Accounts, Welfare Measurement

Die Large Cases Unit (LCU) der Statistischen Ämter des Bundes und der Länder hat im Jahr 2020 ihre Arbeit aufgenommen. Sie untersucht seitdem große multinationale Unternehmensgruppen und ihre Untergliederungen auf Kohärenz der in der amtlichen Statistik vorliegenden Daten. Einbezogen in die Prüfungen werden auch externe Datenquellen. Die Aufklärung erkannter Inkohärenzen dient dazu, die Qualität der Volkswirtschaftlichen Gesamtrechnungen und der zugrunde liegenden Wirtschaftsstatistiken zu sichern. Der Vortrag beschreibt das grundlegende methodische Vorgehen der LCU und informiert über die wichtigsten Erkenntnisse der bisherigen LCU-Arbeit.

130 Regionale Sterblichkeitsentwicklung nach Todesursachen in Deutschland vor, während und nach der COVID-19-Pandemie

Michael Mühlichen, Pavel Grigoriev, Sebastian Klüsener

Bundesinstitut für Bevölkerungsforschung (BiB), Deutschland
michael.muehlichen@bib.bund.de

SECTION: DGD (Session 3): Regional, socioeconomic, and sex differences in morbidity and mortality

Während der COVID-19-Pandemie wurde in fast allen Ländern der Erde ein deutlicher Anstieg der Sterblichkeit verzeichnet. In Deutschland ist die Lebenserwartung drei Jahre in Folge gesunken (2020–2022). Regional fiel der Rückgang der Lebenserwartung in Deutschland sehr unterschiedlich aus. Gleiches gilt für das Ausmaß des anschließenden Erholungsprozesses. Auf Basis der amtlichen Todesursachenstatistik untersucht dieser Beitrag, welche Faktoren die regionalen Gradienten der Sterblichkeit bestimmt haben. Neben der durchschnittlichen Lebenserwartung wird hierbei auf standardisierte Sterberaten und Dekompositionsmethoden zurückgegriffen, um auf Ebene der 400 Kreise (harmonisiert für den aktuellen Gebietsstand) von 1992 bis 2023 regionale Unterschiede nach Alter, Geschlecht und Grundleiden auszuwerten. Unsere Ergebnisse zeigen deutliche regionale Unterschiede in der Übersterblichkeit seit Beginn der COVID-19-Pandemie sowohl in der Höhe als auch in ihrem zeitlichen Auftreten. Auch der anschließende Erholungsprozess ist regional bislang sehr unterschiedlich vorangeschritten. Bei den Frauen hat die Lebenserwartung im Jahr 2023 nur in Sachsen und Rheinland-Pfalz wieder

das Vorpandemieniveau erreicht, bei den Männern nur in Baden-Württemberg und Schleswig-Holstein. Die regionale Verteilung der Übersterblichkeit im Jahr 2020 ist überwiegend mit COVID-19 und hohem Alter assoziiert. In den Jahren 2021 und 2022 haben jedoch auch jüngere Altersgruppen in weiten Teilen Deutschlands zum Anstieg der Sterblichkeit beigetragen. Zudem ist das Ausmaß der Grippesterblichkeit ab Ende 2021 gestiegen. Der Wiederanstieg der Lebenserwartung im Jahr 2023 ist vor allem von der Altersgruppe 50-64 getragen. Todesursachenspezifische Auswertungen für 2023 sind aktuell in Arbeit und werden rechtzeitig zur Statistischen Woche abgeschlossen sein.

131 Ergebnisse unserer aktuellen Arbeit an einer Pflegebedarfsprognose. Eine weitere Komponente im System der Raumordnungsprognose des BBSR

Steffen Maretzke

Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) Bonn, Deutschland
steffen.maretzke@bbr.bund.de

SECTION: DGD (Session 3): Regional, socioeconomic, and sex differences in morbidity and mortality

Das BBSR Bonn veröffentlicht regelmäßig regionalisierte Prognoseergebnisse, basierend auf einer Bevölkerungsprognose. Die aktuelle zensusbereinigte Bevölkerungsprognose des BBSR aus dem Jahr 2024 (Maretzke/Hoymann/Schlömer 2024) wurde bereits durch eine Haushalts-, Erwerbspersonen- und Wohnungsmarktprognose aufgewertet. Inzwischen arbeitet das BBSR an einer regionalisierten Pflegebedarfsprognose.

Im Rahmen dieses Vortrages werden regionalisierte Aussagen zum Niveau wie zur Entwicklung des Pflegebedarfs in den Regionen Deutschlands präsentiert. Die diesen Analysen zugrunde liegenden Daten wurden im Rahmen einer Sonderaufbereitung der Pflegestatistik generiert, in Kooperation mit dem Forschungsdatenzentrum Kamenz. Aufbereitet wurden die Daten in folgender Struktur:

- Daten der Jahre 2017, 2019, 2021, (2023 möglich)
- Pflegebedürftige in ambulanter und stationärer Betreuung
- Differenzierung nach fünf Pflegegraden

Im Zuge der Sonderaufbereitung dieser Daten wurde auf der regionalen Ebene der Kreise, Raumordnungsregionen und Länder angestrebt, die Pflegebedarfsdaten hinsichtlich der Alters- und Geschlechterstruktur der zu Pflegenden möglichst differenziert verfügbar zu machen. Dafür wurden unterschiedlichste Altersgruppierungen und Zusammenfassungen von Pflegegraden getestet. Im Ergebnis dieser Analysen liegen für die Bundesländer wesentlich differenziertere alters- und geschlechtsspezifische Pflegebedarfszahlen als für die Kreise vor.

Da die 2023er Pflegedaten vom FDZ Kamenz noch aufbereitet werden, können im Rahmen des Vortrages voraussichtlich nur regionale Muster und erste Aussagen zur Veränderung der alters- und geschlechtsspezifischen Pflegebedarfsquoten (Pflegebedürftige je Einwohner) im Zeitraum 2017-2021 sowie Eckdaten von 2023 diskutiert werden. Auf Basis der alters- und geschlechtsspezifischen Pflegebedarfsquoten von 2017-2023 wird im Weiteren ein differenziertes Annahmegerüst für den Prognosezeitraum 2022-2045 auf Kreisbasis entwickelt, differenziert nach der Art der Pflege und dem Pflegegrad. Diese differenzierten Pflegequoten in Verbindung mit den Ergebnissen der BBSR-Bevölkerungsprognose ermöglichen die Berechnung einer Pflegebedarfsprognose für den Zeitraum 2022-2045 auf Kreisebene. In Abhängigkeit von der Verfügbarkeit der 2023er Pflegebedarfsdaten werden auch deren Analyseergebnisse in die Präsentation einfließen.

132 Untersuchung des Einflusses sozialer Netzwerke und kognitiver Funktionen auf die Gesamtmortalität älterer Erwachsener unter Verwendung der Daten der NAKO Gesundheitsstudie

Elena Rakuša (1), Daniela Georges (2), Gabriele Doblhammer (1)

1: Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE), Deutschland; 2: Universitätsmedizin Rostock, Institut für Arbeits-, Sozial- und Umweltmedizin, Deutschland
elena.rakusa@dzne.de

SECTION: DGD (Session 3): Regional, socioeconomic, and sex differences in morbidity and mortality

Hintergrund: Soziale Integration und kognitive Funktionen sind entscheidend für das Überleben im hohen Alter, ihre gemeinsame Wirkung auf die Mortalität ist jedoch kaum erforscht. Die vorliegende Studie untersucht deshalb, welchen Einfluss soziale Netzwerke auf die Beziehung zwischen semantischem Gedächtnis, exekutiven Funktionen und Mortalität haben.

Methoden: Es wurden Daten der NAKO Gesundheitsstudie aus der Basiserhebung von 2014 bis 2019 verwendet, wobei Personen ab 60 Jahren eingeschlossen wurden. Der Vitalstatus wurde bis 31. Dezember 2023 erhoben. Kognitive Fähigkeiten wurden anhand des semantischen Gedächtnisses und der exekutiven Funktionen gemessen, die soziale Integration über den Social Network Index (SNI). Mittels Cox-Proportional-Hazards-Modellen wurden Hazard Ratios (HR), 95 %-Konfidenzintervalle (CI) und Interaktionseffekte geschätzt.

Ergebnisse: Von 55.178 eingeschlossenen Personen verstarben 2.524 (4,57 %) während des Beobachtungszeitraums. Moderate (HR = 0,72, CI = 0,64–0,81) und hohe exekutive Funktionen (HR = 0,67, CI = 0,55–0,81) waren mit signifikant niedrigeren Mortalitätsrisiken assoziiert; beim semantischen Gedächtnis war der Trend ähnlich, jedoch insignifikant. Personen mit moderater Integration in soziale Netzwerke (HR = 0,63, 95 % CI = 0,54–0,73) wiesen im Vergleich zu isolierten Personen ein geringeres Mortalitätsrisiko auf. Die Interaktionen wiesen mit Ausnahme der hohen exekutiven Gedächtnisfunktion U-förmige Muster auf. Insbesondere in Kombination mit moderater sozialer Integration war die Mortalität bei niedrigen und moderaten kognitiven Fähigkeiten deutlich reduziert (z. B. niedrige exekutive Funktion: HR = 0,54, CI = 0,38–0,75; moderates semantisches Gedächtnis: HR = 0,72, CI = 0,59–0,88).

Schlussfolgerung: Die Befunde unterstreichen die kritische Rolle der sozialen Integration für ältere Menschen mit kognitiven Einschränkungen. Eine Förderung der sozialen Integration könnte das Sterberisiko in dieser Gruppe deutlich senken. Zukünftige Arbeiten sollten die einzelnen Dimensionen der Netzwerke und deren unterschiedliche Effekte auf kognitive Bereiche untersuchen, um zu klären, warum vor allem moderate Integration protektiv wirkt.

Ausblick: Regional- und geschlechtsspezifische Auswertungen laufen und werden auf der Statistischen Woche vorgestellt.

133 Soziodemografische und gesundheitliche Determinanten sportlicher Freizeitaktivität im mittleren und höheren Erwachsenenalter: Eine Analyse mit Daten des Deutschen Alterssurveys (DEAS) 2023

Sonja Nowossadeck, Svenja M. Spuling

Deutsches Zentrum für Altersfragen (DZA) Berlin, Deutschland
sonja.nowossadeck@dza.de

SECTION: DGD (Session 3): Regional, socioeconomic, and sex differences in morbidity and mortality

Hintergrund: Regelmäßige sportliche Aktivität fördert die körperliche und psychische Gesundheit und ist insbesondere im höheren Lebensalter von Bedeutung. Gleichzeitig zeigen Studien, dass soziale Ungleichheiten und gesundheitliche Einschränkungen das Bewegungsverhalten beeinflussen. Ziel der Analyse ist es, den Zusammenhang zwischen soziodemografischen Merkmalen, ausgewählten Erkrankungen und der sportlichen Freizeitaktivität bei Erwachsenen ab 43 Jahren zu untersuchen. Daten und Methoden: Datengrundlage ist die DEAS-Erhebung 2023 mit 4.992 Personen ab 43 Jahren. Erfasst wurde die Häufigkeit sportlicher Aktivität in der Freizeit (z. B. Wandern, Fußball, Gymnastik, Schwimmen). Als sportlich aktiv gelten Personen, die mindestens einmal pro Woche Sport treiben. Zur Analyse der Zusammenhänge wurden stufenweise logistische Regressionsmodelle berechnet.

Ergebnisse: 54 % der Befragten sind mindestens einmal wöchentlich sportlich aktiv. Personen ab 76 Jahren weisen eine deutlich geringere Sportbeteiligung auf (OR: 0,390) als Jüngere. Für das Geschlecht zeigen sich keine signifikanten Unterschiede. Einkommensarme Personen (OR: 0,428) sowie Personen in ländlichen Regionen (OR: 0,727) sind seltener sportlich aktiv. Unter den betrachteten Erkrankungen stehen insbesondere Herz-Kreislauf-Erkrankungen (OR: 0,547) und Diabetes (OR: 0,374) in signifikant negativem Zusammenhang mit der Sportaktivität. Für Krebs, mobilitätseinschränkende und seelische Erkrankungen zeigen sich keine signifikanten Effekte.

Diskussion: Die Ergebnisse verdeutlichen soziale und gesundheitliche Ungleichheiten in der sportlichen Freizeitaktivität. Insbesondere ältere, einkommensarme und chronisch kranke Personen sind seltener sportlich aktiv. Daraus ergibt sich ein Bedarf, gesundheitsfördernde Maßnahmen gezielt auf diese Gruppen auszurichten, um Teilhabechancen zu verbessern und gesundheitlicher Ungleichheit entgegenzuwirken.

134 Panel data models with randomly generated groups

Jean-Pierre Florens (1), Anna Simoni (2)

1: Toulouse School of Economics, France; 2: CREST - CNRS, Italien
anna.simoni@ensae.fr

SECTION: Empirical Economics and Applied Econometrics (Session 3)

We consider a dynamic panel data model that accounts for a latent group structure across individuals which is constant over time. Differently from the previous literature, we adopt a structural modeling that assumes that the individual effects are generated from a finite mixture with an unknown number of components and unknown parameters for each components. We first

establish identification of this model. Then, we specify a prior for the number of components, the parameters of the mixture as well as for the coefficients of the dynamic and exogenous covariates. This extends the mixture of finite mixtures model to panel data settings. We establish asymptotic frequentist properties for the posterior of the parameters of interest as well as for the number of components. A Monte Carlo exercise illustrates finite sample properties.

135 Inference on Directionally Differentiable Functions of Nonparametric Estimators

Jan Scherer

Universität Bonn, Deutschland
s3jnsche@uni-bonn.de

SECTION: Empirical Economics and Applied Econometrics (Session 3)

We study inference on parameters of the form $\phi(\theta_0)$, where ϕ is a known directionally differentiable transformation and θ_0 is an unknown parameter. We focus on settings, where θ_0 is an unknown function estimated using some nonparametric estimator $\hat{\theta}_n$. As many nonparametric estimators do not converge in distribution, existing extensions to the Delta method are not applicable in our setting. We propose to use strong approximations to the distribution of $\hat{\theta}_n$ as an alternative concept to convergence in distribution. Further, we present a notion of directional differentiability which is sufficiently flexible to handle the irregularity of nonparametric estimators. These concepts enable us to derive a new Delta method which approximates the distribution of the plug-in estimator $\phi(\hat{\theta}_n)$. Since these distributional approximations are rarely pivotal, we suggest a simulation-based estimator and provide conditions for its consistency. Confidence intervals based on this estimator are shown to provide local size control under conditions on the directional derivative of ϕ . We illustrate the applicability of our results in three examples and study its finite sample performance in a simulation study.

136 Group-Specific Heterogeneity in Short Binary Outcome Panels

Björn Höppner

Universität Bonn, Deutschland
b.hoepner@uni-bonn.de

SECTION: Empirical Economics and Applied Econometrics (Session 3)

This paper presents identification and estimation results in the presence of latent group-specific heterogeneity in short binary outcome panels when covariates are available. Specifically, we assume that each unit can be assigned to a time-constant group and model the joint distribution of the binary outcomes conditional on the covariates across all time periods as a mixture model. In our baseline model, we allow the component weights to be fully flexible functions of all covariates over time, while we assume that the binary outcomes are independent over time given the covariates, and the component distributions depend on the contemporaneous covariates only. We present nonparametric identification results for the component weights and distributions under weak conditions. For instance, when the component weights are positive on the entire covariate support and the component distributions exhibit sufficient variation at a single point in the covariate space, the component distributions are nonparametrically identified on the entire support of the

covariates. We leverage an inherent exclusion restriction of the model to solve the intra-component label switching problem without further assumptions. In our baseline model, the number of time periods needed for identification depends on the number of groups in a less stringent way than in a setting without covariates. Specifically, identification is possible with as few as two time periods. This dependence changes in a dynamic panel or under additional exclusion restrictions. Our identification results carry over to settings with continuous outcomes or discrete outcomes with support larger than 2. Additionally, we develop a semiparametric estimator and study its asymptotic properties.

137 Decomposing Price-Energy in Germany and China's Divergent Electricity Markets using STL and Multivariate Functional Data Analysis

Miao Yu, Philipp Sibbertsen

Leibniz University Hannover, Germany
miao.yu@statistik.uni-hannover.de

SECTION: High-dimensional Time Series and Network Data

This study develops a data-driven framework that integrates Functional Data Analysis (FDA) with Seasonal-Trend decomposition using Loess (STL) to investigate how energy production dynamics interact with electricity prices in Germany and China: two contrasting markets governed by marginal cost pricing and policy-anchored regimes, respectively. By decomposing time series into trend-cycle-residual components, we isolate long-term structural forces from short-term variability and policy-induced distortions. Functional Principal Component Analysis (FPCA) reveals nation-specific modes: fossil fuels in Germany and nuclear energy in China function as flexible stabilizers, mitigating renewable intermittency within their respective pricing architectures. Multivariate Functional Response Analysis (MFRA) further quantifies time-varying elasticity coefficients, revealing that Germany's liberalized market structure amplifies the coupling between renewable generation and electricity prices, whereas China's coal-indexed benchmark pricing dampens this linkage. In Germany, where excess renewable capacity during critical periods, particularly summer photovoltaic surpluses, paradoxically increases overall system costs, highlighting the need for investments in grid flexibility rather than continued unilateral capacity expansion. Conversely, China's hybrid system exhibits that although wind and solar technologies achieve rapid cost reductions, rigid pricing mechanisms for hydropower and nuclear energy distort market signals, inadvertently sustaining fossil gas's marginal pricing influence. Finally, prediction results validate the robustness of the estimated coefficient functions.

138 Go with the Flow? Forecasting Regime Switches in Wind Direction

Chris Fotso (1), Andrew Harvey (2), **Yeliz Özer** (1), Dario Palumbo (3), Philipp Sibbertsen (1)

1: Leibniz Universität Hannover, Institut für Statistik, Deutschland; 2: Faculty of Economics, University of Cambridge, United Kingdom; 3: Homerton College, University of Cambridge, United Kingdom
oezer@statistik.uni-hannover.de

SECTION: High-dimensional Time Series and Network Data

Wind energy plays a crucial role in the transition to green energy. Although wind turbine technology continues to improve and repowering takes place, maintenance costs for wind turbines remain a major issue. Horizontal-axis wind turbines with azimuth engines adjust the nacelle and blades to the wind direction to maximize energy capture. However, frequent directional changes despite motion reversals can cause significant cable damage over time.

In this paper, we extend regime-switching models for circular variables by incorporating seasonality and allowing for more than two regimes in the joint distribution of wind direction and wind speed. The model is built on a dynamic mixture framework with score-driven updates, where regime probabilities are determined via a softmax transformation of latent variables. By explicitly modelling the joint density across multiple regimes, the approach captures switching behaviour and seasonal effects.

Using high-frequency data coming from wind farms located in two different wind zones in Germany, we test our approach and demonstrate its ability to forecast regime switches in wind direction. Our model offers wind farm operators a practical tool to optimize nacelle adjustment strategies, for example during summer months with frequent directional shifts and lower wind speeds. This has the potential not only to reduce maintenance costs but also to improve competitiveness in spot market trading. While we focus on wind energy applications, the approach is general and can be applied to other contexts, such as modelling pollutant dispersion in the atmosphere.

139 Sparsity and Fusion Penalization in Large Vector Autoregressions

Daniel Dzikowski, Carsten Jentsch

TU Dortmund University, Deutschland
daniel.dzikowski@tu-dortmund.de

SECTION: High-dimensional Time Series and Network Data

We consider a general high-dimensional VAR setup that covers periodic or panel VARs, among others. The high-dimensionality of these models causes considerable problems when estimating them with standard methods such as OLS. Therefore we propose the use of appropriate regularization methods such as the fused lasso to control the dimension of the models. The fused lasso combines the standard lasso penalty term with a fusion penalty term. While the lasso penalty shrinks parameter estimates towards zero, the fusion penalty is designed to shrink two parameter estimates towards each other. Accordingly, the application of the fused lasso is particularly suitable in periodic VAR applications with sparse periodic structures or in panel VAR applications with similar cross-sectional dependence structures. We consider estimation and inference in general high-dimensional VAR setups using the fused lasso. Further, we provide a method for suitable choices of the tuning

parameters that also accounts for the dependencies in the data. A key result of the paper is to provide valid inference on many target parameter in high-dimensional VARs. For this purpose, we use the asymptotic properties of desparsified variants of the regularized estimators with fused lasso. Finally, we demonstrate the practical usefulness of our approach by a real data application and illustrate its performance via Monte Carlo simulations.

140 Probabilistic forecasting and forecast reconciliation for wind power production

Sven Pappert, Antonia Arsova

TU Dortmund University, Germany
arsova@statistik.tu-dortmund.de

SECTION: High-dimensional Time Series and Network Data

Forecast reconciliation is applied to ensure that forecasts for multiple time series on different levels of a hierarchy conform to the linear restrictions prescribed by the hierarchy. When reconciling probabilistic forecasts, this linear restriction is to be enforced on the distributional level. Building upon the approach of Panagiotelis et al. (European Journal of Operational Research, 306(2):693–706, 2023), who construct their reconciled forecasts as a linear function of the base-level forecasts, we construct the reconciled forecasts using a feedforward multilayer perceptron (MLP) neural network. One aim of our work is to study under which circumstances a linear reconciliation strategy is sufficient (and optimal) and when non-linear generalizations, such as the MLP, are needed. As an empirical application, the different reconciliation strategies are compared when applied to probabilistic one-step ahead forecasts for German wind-power production at three different spatial hierarchical levels.

141 Geheimhaltung georeferenzierter Ergebnisse der Lohn- und Einkommensteuerstatistik durch lokale Aggregation mit Hilfe des Quadtree-Verfahrens

Michael Buchner (1), Robert Garthoff (2)

1: Statistisches Bundesamt, Deutschland; 2: Statistisches Landesamt des Freistaates Sachsen, Deutschland
michael.buchner@destatis.de

SECTION: Methodology of Statistical Surveys (Session 5)

Seit dem Berichtsjahr 2019 liegen die Ergebnisse der Lohn- und Einkommensteuerstatistik in georeferenzierter Form auf Basis von Gitterzellen mit 100m-Rasterweite vor. Während sich damit zahlreiche neue Auswertungsmöglichkeiten ergeben, erhöht sich gleichzeitig der Geheimhaltungsbedarf erheblich. Die lokale Aggregation – und anschließende Disaggregation der zusammengefassten Werte in das Ausgangsraster – mit Hilfe der Erzeugung hierarchischer Quadtrees erweist sich für die Lohn- und Einkommensteuerstatistik als besonders vielversprechender Ansatz. Er ist zum einen kompatibel zum Verfahren der Zellspernung, welches für rein tabellarische Auswertungen bislang genutzt wird. Zum anderen ermöglicht das Quadtree-Verfahren die Bereitstellung der den kartografischen Darstellungen zugrunde liegenden Daten in einer für den Nutzer grundsätzlich leicht nachvollziehbaren Form. In der praktischen Anwendung des Quadtree-Verfahrens auf die Daten der Lohn- und Einkommensteuerstatistik ergeben sich jedoch zwei methodische Herausforderungen.

Zum einen sind die Daten lokal sehr unterschiedlich konzentriert, wodurch zum Teil sehr starke Aggregationen, mit entsprechenden Informationsverlusten, nötig werden. Zum anderen werden die Ergebnisse der Lohn- und Einkommensteuerstatistik auch auf Ebene von Verwaltungseinheiten sehr intensiv genutzt. Dies betraf bislang Auswertungen bis auf Gemeindeebene. Mit der Georeferenzierung der Statistik sind nun aber auch kleinräumige Auswertungen unterhalb der Gemeindeebene, auf Basis unterschiedlicher Städtesystematiken, möglich. Damit stellt sich in zunehmendem Maße das Problem der Aufdeckung von Geheimhaltungsfällen durch geografische „Verschneidung“ von Auswertungen auf Gitterzellenbasis einerseits sowie auf Basis von Verwaltungseinheiten andererseits. Gerade dieser letzte Punkt wird jedoch von der derzeit verfügbaren Standardsoftware, die bislang für die Umsetzung des Quadtree-Verfahrens genutzt wird, nicht berücksichtigt. Für die Geheimhaltung georeferenzierter Ergebnisse der Lohn- und Einkommensteuerstatistik wird deshalb derzeit ein eigener Ansatz entwickelt, der aus einer Kombination des Quadtree-Verfahrens mit einer begrenzten Anzahl an Aggregationen und einem datenverändernden Verfahren für die restlichen bzw. zusätzlichen Geheimhaltungsfälle besteht. In einem Werkstattbericht sollen die Methodik des Ansatzes, die verwendeten Instrumente sowie erste Ergebnisse vorgestellt werden.

142 Veröffentlichungspotenziale von georeferenzierten Daten der Lohn- und Einkommenssteuerstatistik am Beispiel der kleinräumigen Einkommensverteilung in Nordrhein-Westfalen

Christoph Alfken, Sören Görner

IT.NRW – Statistisches Landesamt Nordrhein-Westfalen, Deutschland
christoph.alfken@it.nrw.de

SECTION: Methodology of Statistical Surveys (Session 5)

In den letzten Jahren hat die Debatte über die wachsende Einkommensungleichheit an Intensität zugenommen. Dabei wird häufig auf die Differenzen zwischen sozialen Gruppen verwiesen. Einkommensdisparitäten manifestieren sich jedoch nicht nur auf sozialer, sondern auch auf räumlicher Ebene. So konzentrieren sich Personen mit ähnlichen Einkommen in bestimmten Gebieten. Diese räumliche Segregation lässt sich auf verschiedenen Maßstabsebenen beobachten, etwa zwischen ost- und westdeutschen Bundesländern oder urbanen und ländlichen Gemeinden. Auch innerhalb von Städten und Gemeinden zeigt sich dieses Phänomen, da bestimmte Einkommensgruppen in unterschiedlichen Stadtteilen konzentriert sind. Bisher fehlen jedoch flächendeckende und einheitliche Daten, um diese räumlichen Muster detailliert zu analysieren.

Das vorliegende Projekt zielt darauf ab, amtliche Einkommensdaten auf möglichst kleinräumiger Ebene bereitzustellen. Hierfür werden georeferenzierte Daten aus der Lohn- und Einkommenssteuerstatistik verwendet und auf 1-km-Gitterzellen aggregiert. Aufgrund der sensiblen Natur dieser Daten stellte die Wahrung der Geheimhaltung bislang ein Hindernis für eine kleinräumige Veröffentlichung dar. Durch die Anwendung der Methode der Kerndichteschätzung können jedoch die gesetzlichen Anforderungen an die statistische Geheimhaltung erfüllt werden, während gleichzeitig ein hoher Informationsgehalt der Daten gewahrt bleibt, da die räumliche Struktur der Originaldaten erhalten bleibt.

Das Projekt nutzt die Daten der Lohn- und Einkommenssteuerstatistik aus dem Jahr 2020 für Nordrhein-Westfalen und berechnet die mittleren Einkommen je Gitterzelle aus den zu versteuernden Gesamteinkünften. Damit liegen erstmals amtliche Daten zur kleinräumigen Einkommensverteilung flächendeckend für ein Bundesland vor.

Eine erste Analyse für Nordrhein-Westfalen offenbart prägnante geografische Muster in der Einkommensverteilung. Neben den bekannten räumlichen Disparitäten innerhalb des Ruhrgebiets,

lassen sich auch Korrelationen zwischen Einkommen und Raumnutzung feststellen. So finden sich häufig Konzentrationen niedriger Einkommen in der Nähe großer Industrie- und Gewerbegebiete oder Großinfrastrukturen während sich höhere Einkommen in weniger verdichteten Gebieten am Rand der Ballungsgebiete konzentrieren. Die Daten bieten somit das Potenzial, räumliche Einkommensdisparitäten auf kleinräumiger Ebene zu untersuchen und gezielte politische Maßnahmen zu unterstützen bzw. zu evaluieren

143 'Fitness for use' von demographischen Rasterdaten mit künstlichem Messfehler

Martin Möhler

Destatis, Deutschland
martin.moehler@destatis.de

SECTION: Methodology of Statistical Surveys (Session 5)

In der europäischen Zensusrunde 2021/22 produzierten die Mitgliedsstaaten demographische Aggregate in kleinen geographischen Rasterzellen. In solchen Datenprodukten ist es aufgrund der hohen räumlichen Granularität schwer, die Vertraulichkeit statistischer Angaben sicherzustellen. Erschwert wird diese Aufgabe dadurch, dass Rasterzellen-Aggregate sich auch mit Aggregaten für administrative Gebiete verschneiden lassen, um Ergebnisse für noch kleinere Areale zu errechnen (sog. geographisches Differenzbildungsrisiko).

Deutschland und andere EU-Mitgliedsstaaten schützen Rasterdaten durch stochastische Überlagerung mit der sog. Cell-Key-Methode (CKM), ein Verfahren zur Sicherstellung der statistischen Geheimhaltung, das kleine zufällige Störterme verwendet, um Differenzbildungsrisiken zu begrenzen. Wir betrachten den Schutzeffekt von CKM hinsichtlich des Risikos geographischer Differenzbildung sowie die Auswirkungen der Methode auf analytische Validität („Fitness for use“) eines demographischen Rasterprodukts der Auflösung 1ha. Analytische Validität wird anhand einer exemplarischen Auswertung zur Erreichbarkeit von Apotheken untersucht.

144 Regionale Auswertung von anonymisierten georeferenzierten Daten

Lorena Gril, Ulrich Rendtel

Freie Universität Berlin, Deutschland
lorena.gril@fu-berlin.de

SECTION: Methodology of Statistical Surveys (Session 5)

Georeferenzierte Daten werden aus Gründen des Datenschutzes häufig anonymisiert. Dies geschieht entweder durch Aggregation zu größeren räumlichen Einheiten (wie Verwaltungseinheiten höherer Ordnung oder Gitterzellen mit größeren Kantenlängen) oder durch den Einsatz stochastischer Verfahren, die die ursprünglichen Koordinaten gezielt überlagern. Ziel einer Analyse ist es, die räumliche Verteilung eines interessierenden Merkmals zu visualisieren, zum Beispiel in Form von Karten. Herkömmliche Analysemethoden berücksichtigen oft nicht den Anonymisierungsprozess und behandeln anonymisierte Koordinaten als tatsächliche Koordinaten. Ein statistisches Messfehlermodell ermöglicht jedoch wesentlich effizientere Analysen, indem es den Einfluss der Anonymisierung explizit berücksichtigt. Ziel des Vortrags ist es, Ergebnisse zur effizienten Nutzung von anonymi-

sierten georeferenzierten Daten zu präsentieren. Neben der Vorstellung der entwickelten Methoden werden auch empirische Ergebnisse aus einer Bevölkerungs- und Gesundheitsbefragung sowie der regionalen Verteilung der Einkommenssteuerpflichtigen in Berlin präsentiert.

145 Stochastically interpretable distributional regression with neural networks

Malte Jahn

HSU Hamburg, Deutschland
jahnma@hsu-hh.de

SECTION: Statistical Theory and Methods (Session 3)

A generalization of conventional mean regression models is obtained by introducing additional regression equations for higher moments of the conditional distribution of the response variable. In this paper, it is shown how artificial neural networks can be used to construct corresponding efficiently parameterized distributional regression models based on the fact that all conditional moments depend on the same set of explanatory variables. In particular, the focus is on explicitly maximizing a suitable likelihood function subject to the neural network parameters. The size of the network is kept in a reasonable balance to the number of available observations, allowing for inference in terms of the calculation of (approximate) confidence and prediction intervals, most importantly for the marginal effects of the considered explanatory variables on the moment parameters. This contributes significantly to the (stochastic) interpretability of the model. The practical usefulness is demonstrated by analyzing several real-world data sets and comparing the proposed framework to existing distributional regression frameworks.

146 Assessing Loss Preferences on the Basis of Noisy Forecast Errors

Matei Demetrescu (1), **Christoph Hanck** (2), Hajo Holzmann (3)

1: TU Dortmund; 2: Universität Duisburg-Essen; 3: Universität Marburg
christoph.hanck@vwl.uni-due.de

SECTION: Statistical Theory and Methods (Session 3)

The framework of (RES 2005) allows quantifying the asymmetry of unknown loss functions based on observed forecasts as well as for rationality testing under asymmetric loss. To this end, they formalized a class of loss functions characterized by the asymmetry and the tail weight parameter, which nest asymmetric linear and asymmetric quadratic loss functions. Given the shape of these loss functions, they derive tractable moment conditions for optimality of forecasts. In particular, the "generalized" forecast error should have zero conditional expectation given all predictors available. Based on these moments, GMM estimation of the asymmetry parameter is possible. Under conditions, GMM is consistent and asymptotically normal. Formulas for standard errors allow to conduct inference, typically about the asymmetry parameter.

However, forecast errors are likely to be noisy (e.g. from estimating the forecast models), such that Elliott et al.'s assumptions are only approximately met. We follow West (1996) in modelling the estimation noise and assess, theoretically and in simulations, the impact of such noise on the estimators of loss function parameters. We show that consistency of GMM of loss function

parameters is typically not affected, but Elliott et al.'s standard errors underestimate the true variability of the estimators. Like in West's analysis of the Diebold-Mariano test, the correct standard errors depend on the forecast model and on the estimation scheme.

While one often observes forecasts, information on how the forecasts were generated is seldom available, such that adjusting standard errors is not often feasible. Notwithstanding, our second contribution is to propose a conditional moment test for the parameters of the loss function that is robust to estimation noise, and find it to have good size. Size control implies somewhat reduced power. Yet, the test has the further advantage of being able to identify subsamples where the null is violated.

147 Inference in Panel SVARs with Cross-Sectional Dependence of Unknown Form

Lennart Empting (1), Saskia Öztürk (2), Simone Maxand (3), Konstantin Wagner (4)

1: Universität Duisburg-Essen; 2: Universität Göttingen; 3: Europa-Universität Viadrina Frankfurt (Oder); 4: Humboldt Universität Berlin
lennart.empting@vwl.uni-due.de

SECTION: Statistical Theory and Methods (Session 3)

Moving-block bootstrap procedures have become a preferred method to determine the sampling uncertainty of vector autoregressive (VAR) model estimation, most prominently visualized as confidence bands around impulse response functions (IRF). In this study, we extend these inferential methods for multivariate time series by the cross-section dimension and compile recursive-design moving-block bootstrap procedures for proxy-identified panel VAR models and their structural IRF. The procedures resample blocks of estimated error terms either in (i) the temporal, (ii) the cross-sectional, or (iii) both dimensions jointly. Their asymptotic assessment and a finite-sample Monte-Carlo study both suggest the preferred use of panel-block resampling in both dimensions when confronted with data properties typically found in empirical panel VAR applications.

148 The pvars R-Package: VAR modeling for heterogeneous panels

Lennart Empting

Uni Duisburg-Essen, Deutschland
lennart.empting@vwl.uni-due.de

SECTION: Statistical Theory and Methods (Session 3)

pvars offers a seamless implementation of vector autoregressive (VAR) methods for heterogeneous panel data. The R-package comprises panel cointegration rank tests which can account for cross-sectional dependence and for structural breaks in the deterministic term. The implemented panel SVAR models can be estimated under these specifications with pooled cointegrating vectors and identified by various panel identification procedures. In this article, we review these methods and present their

modular implementation in R. Two empirical illustrations reproduce examples from the literature step-by-step and guide the pvars user into conducting own analyses.

149 Einkommensverteilung und Armut in Deutschland: Aktuelle Verteilungsdaten auf dem Prüfstand

Judith Niehues

Institut der deutschen Wirtschaft, Deutschland
niehues@iwkoeln.de

SECTION: Economic, Social and Market Statistics (Session 8)

Für die Bewertung der Verteilungssituation in Deutschland und zur Ableitung von Handlungsoptionen zur Bekämpfung von Armut sind belastbare Daten zur (Einkommens-)Verteilung unerlässlich. Mit Blick auf die Einkommensverteilung gibt es in Deutschland drei verschiedene Mikrodatsätze, die typischerweise zur Herleitung von Verteilungs- und Armutskenziffern herangezogen werden: Das Sozio-oekonomische Panel (SOEP), die Erhebung „Leben in Europa“ (EU-SILC) sowie der Mikrozensus des Statistischen Bundesamtes. Neben datensatzspezifischen Vor- und Nachteilen teilen die drei Datensätze die Gemeinsamkeit, dass sie allesamt in den letzten Jahren von Erhebungsumstellungen betroffen waren. Hervorzuheben ist in diesem Zusammenhang die umfassende Neuaufstellung des Mikrozensus im Jahr 2020, in dessen Folge die Zeitreihen des Mikrozensus ab dem Jahr 2020 separat ausgewiesen werden. Im Zuge der Neukonzeption wurde zudem die bislang separat durchgeführte Erhebung des EU-SILC ab dem Erhebungsjahr 2020 als Unterstichprobe in den Mikrozensus integriert. Da sich durch die Umstellung die Stichprobe des EU-SILC umfassend verändert hat, ist ein Vergleich der Daten der Erhebungsjahre ab dem Jahr 2020 (Einkommensreferenzjahr 2019) mit den Vorjahren nicht möglich. Auch im SOEP haben Stichprobenveränderungen Einflüsse auf Verteilungsbetrachtungen, beispielsweise wenn spezielle Migrations-, Hocheinkommens- oder Hochvermögendenstichproben neu in die Erhebung integriert werden, um auch Bevölkerungsgruppen zu berücksichtigen, die in Befragungen typischerweise unterrepräsentiert sind. Im Zuge eines Wechsels des Befragungsinstituts und vor dem Hintergrund coronabedingter Befragungseinschränkungen kam es im Jahr 2021 zudem zu einem deutlichen Rückgang der Stichprobe, der einen Einkommensvergleich zwischen den betreffenden Jahren erschwert.

Im Beitrag wird beleuchtet, inwiefern die Erhebungsumstellungen eine konsistente Interpretation der Zeitreihen zur Einkommensverteilung erschweren und weshalb Schlussfolgerungen zum Einfluss der Coronapandemie auf die Verteilungssituation daher praktisch nicht möglich sind. Zum anderen werden unterschiedliche Verteilungstrends und stark divergierende Befunde zu Armutsrisikogruppen herausgestellt und die resultierenden Herausforderungen für evidenzbasierte Wirtschaftspolitik diskutiert.

150 Schätzen und Testen der Übersterblichkeiten während der Corona-Pandemie im Zeitraum 2020 bis 2022

Michael Weba

Goethe-Universität Frankfurt, Deutschland
weba@wiwi.uni-frankfurt.de

SECTION: Economic, Social and Market Statistics (Session 8)

Die Übersterblichkeit beschreibt die Differenz zwischen der Gesamtzahl aller während einer Pandemie beobachteten Sterbefälle (unabhängig von den Ursachen) und der Anzahl der Sterbefälle, die man unter „normalen“ Bedingungen erwartet hätte.

In der Literatur sind verschiedene Ansätze zur Schätzung der Übersterblichkeit und der Festlegung von Vergleichswerten behandelt worden, z.B. Verfahren, die auf versicherungsmathematischen Grundsätzen basieren. Die Aufsätze von

De Nicola / Kauermann (J.Roy.Stat.Soc.A, 2025).und Kuhbandner / Reitzner (Cureus 2023) diskutieren ausführlich die Schätzung von Übersterblichkeiten und bieten einen umfangreichen Überblick über die vorhandene Literatur.

In dem Vortrag wird ein wahrscheinlichkeitstheoretischer Ansatz vorgestellt, um Übersterblichkeiten zu schätzen; dieser Ansatz ermöglicht es auch, einen verteilungsfreien Test anzugeben, um Hypothesen über Unter- oder Übersterblichkeiten zu überprüfen. Betrachtet werden sowohl jährliche als auch monatliche Sterbedaten .

151 Internationalität der Unternehmensgruppen am Wirtschaftsstandort Hamburg

Christian Salwiczek

Statistisches Amt für Hamburg und Schleswig-Holstein, Deutschland
christian.salwiczek@statistik-nord.de

SECTION: Economic, Social and Market Statistics (Session 8)

Vor einigen Jahren wurde die „Unternehmensgruppe“ als eigene Identität im statistischen Unternehmensregister etabliert. Generell stehen Unternehmensgruppen häufig im Fokus wirtschaftspolitischer Diskussionen, da sie allgemein einen großen Einfluss auf das Wirtschaftsgeschehen haben.

In den letzten Jahren gewann neben allgemeinen wirtschaftsstatistischen Analysen vermehrt der Aspekt „Internationale Verflechtungen“ massiv an Bedeutung. Dies zeigte sich u.a. während der Brexit-Phase, in Zusammenhang mit dem laufenden Ukraine-Krieg oder der aktuellen Zoll-Politik der Trump-Administration.

Im Rahmen eines Inhouse-Analyseprojektes fand nun eine entsprechende Analyse bei Statistik Nord statt. Hierbei wurde nicht nur eine umfassende geobasierte Auswertung auf traditionellen Wege vorgenommen, sondern auch unterschiedliche KI-Tools (z.B. BankGPT, LLMoin) eingesetzt und evaluiert.

Wie kann bspw. auf Basis unterschiedlicher (Verwaltungs-) Datenquellen die aktuelle Struktur von Unternehmensgruppen erfasst werden? Welche internationalen Verflechtungen gibt es im Bereich der Hamburger Wirtschaft? Lassen sich ggf. branchenspezifische Effekte (z.B. hinsichtlich der Standort-Wahl) feststellen? Diese und zahlreiche weitere Fragestellungen wurden im Rahmen des Projekts bearbeitet und sollen vorgestellt werden.

152 Die Entwicklung der E-Mobilität in Deutschland - Eine Analyse regionaler Unterschiede

Benjamin Fuchs, Martin Kastner

Kraftfahrt-Bundesamt, Deutschland
benjamin.fuchs@kba.de

SECTION: Economic, Social and Market Statistics (Session 8)

Im Bundes-Klimaschutzgesetz ist mit dem Erreichen der Klimaneutralität bis 2045 ein ambitioniertes Ziel formuliert. Tatsächlich sanken die Treibhausgasemissionen im Verkehrssektor im Vergleich zu 1990 in absoluter Größe, andere Sektoren konnten im gleichen Zeitraum jedoch deutlich mehr einsparen. Der Anteil des Verkehrssektors an den Emissionen im Vergleich zu 1990 stieg hierdurch um 9 Prozentpunkte auf 22% in 2023.

Ein wichtiger Baustein zur Emissionsreduzierung im Verkehrssektor ist die Erhöhung der Elektrifizierung. Hierzu wurde 2016 mit dem zwischenzeitlich eingestellten Umweltbonus ein finanzieller Anreiz zum Kauf elektrisch betriebener Fahrzeuge eingeführt.

In diesem Aufsatz wird anhand von Daten des Zentralen Fahrzeugregisters des Kraftfahrt-Bundesamtes der Effekt der Einführung des Umweltbonus, dessen Ausweitung in 2020 sowie Reduzierung und Einstellung in 2023 analysiert. Ein Schwerpunkt wird hierbei auf die Untersuchung der regionalen Variation auf kleinräumiger Ebene sowie auf die Zusammensetzung der fahrzeughaltenden Personen gelegt.

Die Einführung des Umweltbonus in 2016 ließ den Anteil von elektrisch betriebenen Fahrzeugen bei den Pkw-Neuzulassungen nur geringfügig ansteigen. Erst eine Erhöhung der Förderung in 2020 um mehr als das Doppelte hatte einen sichtbaren positiven Effekt. Wie erwartet sind zudem Mitnahmeeffekte kurz vor der Reduzierung der finanziellen Förderung in 2023 zu erkennen. Insgesamt tendieren Privatpersonen eher zu reinen Batterieelektrofahrzeugen (BEV), während Neuzulassungen von Plug-in-Hybriden eher bei gewerblichen Personen ansteigen.

Trotz eines zeitweilig sehr starken Anstiegs in den Neuzulassungszahlen, bleibt der Anteil an elektrisch betriebenen Fahrzeugen im Bestand weiterhin überschaubar. Zum 01.01.2025 beträgt der Anteil an BEV am Pkw-Bestand etwa 3%, während Benzin- und Diesel-Pkw diesen mit einem Anteil von ca. 89% weiterhin dominieren.

Mit dem Umweltbonus ist ausschließlich in westdeutschen Bundesländern zeitgleich ein Anstieg an BEV am Pkw-Bestand zu beobachten. Der Anteil an BEV in Ostdeutschland bleibt mit Ausnahme von Berlin und den umliegenden Kreisen weiterhin deutlich unter dem Bundesdurchschnitt.

Der Aufsatz analysiert zudem die gemeinsame regionale Entwicklung der öffentlichen Ladeinfrastruktur und der E-Mobilität.

153 Bereitstellung von Längsschnittgewichten für das Beschäftigtenpanel

Clara Seesing-Coelho (1,2), Christina Jaeger (1), Marcel Preising (1)

1: Statistisches Bundesamt, Deutschland; 2: Universität Hamburg, Deutschland
clara.seesing-coelho@destatis.de

SECTION: Methodology of Statistical Surveys (Session 6)

Die monatliche Verdiensterhebung (VE) ist eine zentrale Datenquelle der amtlichen Statistik zur Erfassung von Löhnen und Arbeitszeiten in Deutschland, bei der rund 55.000 Betriebe verpflichtend Angaben auf Ebene der etwa 9 Millionen Beschäftigten übermitteln. Die Daten stammen direkt aus den betrieblichen Entgeltabrechnungen und zeichnen sich dadurch durch eine vergleichsweise hohe Validität und Vollständigkeit aus. Künftig wird dieser umfangreiche Datensatz der wissenschaftlichen Forschung als Panel über das Forschungsdatenzentrum bereitgestellt – mit großem Potenzial für Analysen zu Lohnentwicklungen und Arbeitsmarktdynamiken.

Die VE ist als geschichtete Stichprobe konzipiert, wodurch sich unterschiedliche Ziehungswahrscheinlichkeiten für strukturell verschiedene Betriebe ergeben. Eine korrekte Gewichtung ist somit für verlässliche Schätzungen unerlässlich. Die Erstellung einer Längsschnittgewichtung, die für Auswertungen über mehrere Berichtsmonate hinweg benötigt wird, stellt dabei eine besondere methodische Herausforderung dar. Zunächst ist eine geeignete Längsschnittpopulation zu definieren, die durch das Panel repräsentiert wird. Die Designgewichte der Betriebe in der Stichprobe werden dann mittels einer generalisierten Kalibrierung so angepasst, dass bekannte Eckwerte dieser Längsschnittpopulation im Initialmonat reproduziert werden. Als Eckwerte werden die Zahlen der Bundesagentur für Arbeit für die nach Bundesland, Betriebsgrößenklassen und Wirtschaftszweig gegliederte Anzahl an Betrieben, die Anzahl an sozialversicherungspflichtigen sowie die Anzahl an geringfügig Beschäftigten verwendet. Zusätzlich wird monatlich ein gewichtungsbasierter Ausgleich für Antwortausfälle vorgenommen. Dazu werden im Zeitverlauf sowohl die Bleibe- bzw. Antwortwahrscheinlichkeiten der Stichprobenbetriebe als auch die Wahrscheinlichkeit für sogenannte Wiedereintritte – also Betriebe, die nach einem vorübergehenden Ausfall erneut melden durch logit-Modelle modelliert und bei der Gewichtung berücksichtigt.

In diesem Beitrag stellen wir unsere methodischen Ansätze zur Bewältigung dieser Herausforderungen vor.

154 EU-SILC: Änderungen im Auswahl- und Gewichtungsverfahren für die deutsche SILC-Stichprobe ab 2026

Charlotte Articus

Statistisches Bundesamt, Deutschland
Charlotte.Articus@destatis.de

SECTION: Methodology of Statistical Surveys (Session 6)

Die europäische Erhebung über Einkommen und Lebensbedingungen (EU-SILC) ist die amtliche Hauptdatenquelle für die Messung von Armut und Lebensbedingungen in Deutschland und der Europäischen Union. Europäische Verordnungen legen definitorische und methodische Standards fest und gewährleisten so die EU-weite Vergleichbarkeit der Ergebnisse.

Seit dem Erhebungsjahr 2020 ist SILC als Substichprobe in den Mikrozensus integriert. Im Zuge der Integration wurden zentrale Bestandteile des Auswahl- und Gewichtungsverfahrens des Mikrozensus auch für SILC übernommen. Um europäische Anforderungen einzuhalten, muss das SILC-Verfahren zukünftig aber konsequent auf ein Paneldesign umgestellt werden.

Der Mikrozensus kombiniert Wiederholungsbefragungen mit einer Flächenstichprobe und einer Querschnittsperspektive: Rotationsviertelweise werden Auswahlbezirke ausgewählt, die für vier Jahre in ihrer jeweils aktuellen Zusammensetzung berücksichtigt werden. In die Fläche einziehende Haushalte werden also ab dem Zeitpunkt ihres Einzugs aufgenommen, ausziehende Haushalte werden nicht einbezogen. Die Stichprobe bildet damit zu jedem Zeitpunkt die Population in ihrer aktuellen Zusammensetzung ab. Sie kann außerdem in jedem Jahr für alle Rotationsviertel als neu gezogene Stichprobe behandelt werden, in der Panelmortalität keine Rolle spielt.

SILC ist als Panel konzipiert, in dem die lückenlose Befragung und die konsequente Weiterverfolgung der mit der ersten Befragung ausgewählten Personen über vier Jahre eine zentrale Rolle spielt. Ein Einbezug von (zuziehenden) Haushalten nach der Erstbefragung ist nicht vorgesehen. Weiterhin sind fortziehende Haushalte und Personen weiterzuverfolgen. Damit sind bei der Gewichtung der Zeitpunkt der Initialziehung eines Rotationsviertels und Panelmortalität zu berücksichtigen.

Während die Weiterverfolgung bei Fortzug im deutschen System bereits seit 2020 umgesetzt ist, wurde für zuziehende Haushalte bisher der Ansatz der Mikrozensus auf SILC übertragen. Ab 2026 erfolgt nun schrittweise die vollständige Umstellung auf das geforderte Paneldesign. Dadurch wird auch eine Anpassung des Gewichtungsverfahrens notwendig, das zukünftig Rotationsviertel getrennt betrachten und Panelmortalität berücksichtigen muss. Im Vortrag werden die laufenden methodischen Arbeiten an dieser Umstellung dargestellt, zentrale Änderungen zum bisherigen System aufgezeigt und Auswirkungen diskutiert.

155 Das Einheitenproblem in Unternehmensstatistiken - Herausforderungen und Auswirkungen auf Schätzungen

Lara Krell, Ralf Münnich

Universität Trier, Deutschland
lara.krell@uni-trier.de

SECTION: Methodology of Statistical Surveys (Session 6)

Das Einheitenproblem beschreibt Unstimmigkeiten zwischen den Erhebungseinheiten, die in Datenquellen wie Unternehmensregistern erfasst werden, und den statistischen Einheiten, die dargestellt werden sollen. Diese Abweichungen können zu Fehlern in den Schätzungen führen, die durch eine falsche Identifizierung, Beschreibung oder Klassifizierung der Einheiten in einer Erhebung oder in einem Datensatz verursacht werden.

In der deutschen amtlichen Unternehmensstatistik werden wirtschaftliche Einheiten häufig auf Basis rechtlicher Einheiten erfasst. Die Kennzahlen der Unternehmensstatistiken sollen laut EU-Definition allerdings auf Unternehmensebene berichtet werden. Werden die rechtlichen Einheiten zu einem Gesamtunternehmen zusammengefasst, so werden deren unterschiedliche Zugehörigkeiten zu Branchen oder Standorten vernachlässigt. Dies kann zu Verzerrungen in den Kennzahlen führen. Betroffen sind insbesondere große und diversifizierte Unternehmen, welche gleichzeitig eine hohe ökonomische Relevanz besitzen.

Es werden Ansätze untersucht, um den Auswirkungen des Einheitenproblems entgegenzuwirken und damit die Verzerrungen zu reduzieren. Ziel ist es, den Einfluss unterschiedlicher Abgrenzungen auf die Qualität der Statistiken zu quantifizieren und methodische Anpassungen zur Verbesserung

der Erfassung wirtschaftlicher Realität zu diskutieren. Hierbei werden Hilfsinformationen sowie Gewichtsadjustierungen genutzt, um die Schätzungen auf Unternehmensebene zu verbessern.

156 Standard procedure for methodological aspects of thematic surveys at the Federal Statistical Office, Switzerland

Jacques Saliba

Swiss Federal Statistical Office, Switzerland
jacques.saliba@bfs.admin.ch

SECTION: Methodology of Statistical Surveys (Session 6)

Thematic surveys are one component of the Swiss population census system introduced in 2010 by the Federal Statistical Office of Switzerland (FSO). The traditional census has since been replaced by registers supplemented by the annual structural survey, thematic surveys and Omnibus surveys. With samples of 10 000 to 40 000 people, thematic surveys make it possible each year to examine in greater depth one of the topics not fully covered by register data or the structural survey, like gathering detailed information on the state of health of the population.

The presentation describes different statistical methods, from the creation of the survey design to the elaboration of variance estimation, applied by the FSO for all thematic surveys conducted between 2018 and 2024 among people and households: Families and Generations Surveys (EFG 2018 and 2023), the Language, Religion and Culture Survey (ELRC 2019), the Basic Education and Continuous Education Microcensus (MZB 2021), the Mobility and Transport Microcensus (MRMT 2021), the Swiss Health Survey (ESS 2022), the Language and Religion Survey (ELR 2024) and the Culture Survey (EC 2024).

During this period, thematic surveys underwent significant changes: some moved from telephone interviews only to principal self-administered electronic questionnaires, which led to the ELRC survey being split into two surveys, ELR and EC, in 2024. In addition, qualitative improvements were achieved at the weighting stage by using newly available administrative data, such as insured income recorded by the old-age and survivors' insurance (AVS). Our presentation thus provides an overview of the procedures for applying statistical methods aimed at standardizing these methods as far as possible for the FSO's thematic surveys, which could also benefit other FSO surveys.

157 Monitoring time-varying tail risk with the extended Generalized Pareto distribution

Carlotta Pacifici (1), Luca Trapin (2)

1: Bocconi University, Italy; 2: University of Bologna, Italy
luca.trapin@unibo.it

SECTION: (Systemic) Risk Modeling

Tail risk in financial markets reflects the probability of extreme financial losses and is typically measured by the tail index of the financial return distribution. A commonly employed method to assess the dynamics of the tail index is the dynamic Peaks over Threshold (PoT) approach, which assumes that exceedances over a high threshold follow a Generalized Pareto (GP) distribution with time-varying parameters. However, this method is sensitive to the choice of the threshold, and deviations from the GP distribution can introduce bias in tail index estimates. To address

this issue, we extend the dynamic PoT model to incorporate an extended GP (EGP) distribution, which accommodates departures from the standard GP distribution. Through simulation studies and an empirical analysis using the S&P500 returns, we find that the EGP model provides more reliable assessments of tail risk dynamics, leading to improvements in the stability and accuracy of tail risk estimates in the dynamic setting.

158 Expected Shortfall LASSO

Sander Barendse

University of Amsterdam, Netherlands, The
s.c.barendse@uva.nl

SECTION: (Systemic) Risk Modeling

We propose an l_1 -penalized estimator for high-dimensional models of Expected Shortfall (ES). The estimator is obtained as the solution to a least-squares problem for an auxiliary dependent variable, which is defined as a transformation of the dependent variable and a pre-estimated tail quantile. Leveraging a sparsity condition, we derive a nonasymptotic bound on the prediction and estimator errors of the ES estimator, accounting for the estimation error in the dependent variable, and provide conditions under which the estimator is consistent. Our estimator is applicable to heavy-tailed time-series data and we find that the amount of parameters in the model may grow with the sample size at a rate that depends on the dependence and heavy-tailedness in the data. In an empirical application, we consider the systemic risk measure CoES and consider a set of regressors that consists of nonlinear transformations of a set of state variables. We find that the nonlinear model outperforms an unpenalized and untransformed benchmark considerably.

159 Nuclear Hedging Portfolios: Shrinking Complexity for Robust Diversification

Alberto Quaini

Erasmus University, Niederlande
quaini@ese.eur.nl

SECTION: (Systemic) Risk Modeling

Minimum-variance portfolios often incur high rebalancing costs and estimation risk, particularly in large asset markets. We propose a novel approach that targets the precision matrix by jointly estimating hedging portfolios under two complementary regularizations. The first controls the complexity of hedging portfolios, while the second shrinks hedging positions without imposing sparsity. Unlike sparse methods, our framework accommodates pervasive factors and multicollinearities, and it produces stable, diversified portfolios without extreme weights with low turnover. In empirical tests across both small and large cross-sections of assets, these portfolios achieve low out-of-sample volatility, and higher net Sharpe ratios than competing methods.

160 Some new null-proportion estimators with plug-in FDR control.

Sebastian Döhler (1), Iqraa Meah (2)

1: Hochschule Darmstadt, Deutschland; 2: Institut national de la santé et de la recherche médicale (Inserm), Paris, France
sebastian.doehler@h-da.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 1)

The Benjamini-Hochberg (BH) procedure is a staple of modern high-dimensional data analysis. This method can be made more powerful by incorporating estimators of the number (or proportion) of null hypotheses, yielding an adaptive BH procedure which still controls the false discovery rate (FDR).

In this talk we present a unified class of estimators, which encompasses existing and new estimators and which can also be extended to discrete tests. While our focus is on presenting the generality and flexibility of the new class of estimators, we also include some analyses on simulated and real data.

161 EEG-based Eye-Tracking: A Benchmark Dataset for Functional Data Analysis with Open Challenges and Baseline Results

Tiago Vasconcelos Afonso (1), Florian Heinrichs (2)

1: Hochschule Darmstadt, Deutschland; 2: FH Aachen, Deutschland
f.heinrichs@fh-aachen.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 1)

Many methods for functional data require the data to be registered, that is, aligned in time. While this assumption simplifies the analysis, it cannot be justified in many applications, especially when working with continuously recorded sensor data, where no "start", "end" or any other landmark exists.

We introduce a novel dataset designed for benchmarking FDA methods, which includes eye-tracking data from over 100 participants, measured simultaneously via camera and EEG headset. The primary task is to reconstruct eye movements from EEG signals. The dataset contains different levels of difficulty (registered and unregistered data, continuous and abrupt movements, four and more directions of movement).

Additionally, we present a comparative analysis of functional neural networks, that are specifically designed for unregistered data, with established methods.

162 Forecasting the wind: From necessity to added value

Jethro Browell

University of Glasgow, Vereinigtes Königreich
jethro.browell@glasgow.ac.uk

SECTION: Statistics in Environmental and Engineering Sciences (Session 1)

Wind power forecasting became a necessity following deployment of the first large-scale wind farms. Today, these forecasts are as important as ever and there are abundant opportunities for forecasts to unlock added value in settings ranging from maintenance scheduling to algorithmic trading. Despite advances in probabilistic forecasting, many operational systems struggle to incorporate uncertainty in a meaningful way, though this is changing. This talk will reflect on the evolution of wind power forecasting and the innovations that have led to improvements in forecast skill and forecast value. Drawing on a decade of work with wind farm operators, traders, TSOs and forecast vendors, I will present examples of how, with a little help, we can add value with wind power forecasts.

163 On the Estimation of Climate Normals and Anomalies

Tommaso Proietti (1), Alessandro Giovannelli (2)

1: University of Rome Tor Vergata, Italien; 2: Università dell'Aquila, Italien
tommaso.proietti@uniroma2.it

SECTION: Statistical Theory and Methods (Session 4)

The quantification of the interannual component of variability in climatological time series is essential for the assessment and prediction of the El Niño - Southern Oscillation phenomenon. This is achieved by estimating the deviation of a climate variable (e.g., temperature, pressure, precipitation, or wind strength) from its normal conditions, defined by its baseline level and seasonal patterns. Climate normals are currently estimated by simple arithmetic averages calculated over the most recent 30-year period ending in a year divisible by 10. The suitability of the standard methodology has been questioned in the context of a changing climate, characterized by nonstationary conditions. The literature has focused on the choice of the bandwidth and the ability to account for trends induced by climate change. The paper contributes to the literature by proposing a regularized real time filter based on local trigonometric regression, optimizing the estimation bias-variance trade-off in the presence of climate change, and by introducing a class of seasonal kernels enhancing the localization of the estimates of climate normals. Application to sea surface temperature series in the Niño 3.4 region and zonal and trade winds strength in the equatorial and tropical Pacific region, illustrates the relevance of our proposal.

164 glmSTARMA - An R package for fitting spatio-temporal models based on generalized linear models

Steffen Maletz (1), Konstantinos Fokianos (2), Roland Fried (1)

1: TU Dortmund University, Germany; 2: University of Cyprus, Cyprus
steffen.maletz@tu-dortmund.de

SECTION: Joint DStatG and TIES Session

The analysis of spatio-temporal data plays a crucial role in many research areas. Some examples include the observation of disease cases in different regions in epidemiology or the measurement of various climate variables such as precipitation and temperature at measuring stations.

We present a flexible framework for modeling such data that efficiently captures spatial and temporal dependencies. Our approach is based on generalized linear models and allows for the integration of autoregressive dependence structures, the inclusion of covariates as well as different distributions to account for the specific data properties.

In addition to methods for parameter estimation, the framework also includes functions for statistical inference. A particular focus is given to a user-friendly implementation.

In this talk, we will present the basic functionalities of the R package and illustrate their application using concrete examples.

165 Bayesian estimation for virtual experiments in metrology featuring semi-parametric modelling using Gaussian processes

Finn Hughes, Manuel Stavridis, Manuel Marschall

Physikalisch-Technische Bundesanstalt, Germany
finn.hughes@ptb.de

SECTION: Joint DStatG and TIES Session

Virtual experiments are a tool commonly used in measurement science to simulate virtual data that mimic real observations. Virtual experiments are particularly useful when real data are sparse and it is impractical to obtain a statistically sufficient sample of real observations needed to conduct a valid data analysis. Constructed by experts, virtual experiments constitute a highly complex representation of a real measurement process where a parametric model is often considered, mapping real physical quantities to virtual observations. However, due to the required accuracy and high resolution of many metrological applications, purely parametric models often do not provide a sufficient representation of the physically meaningful inputs and corresponding outputs of a virtual experiment. For instance, in optical metrology, there can be high-frequency terms representing small imperfections of a machined specimen that cannot be represented by the purely parametric model. Hence, a non-parametric formulation can be employed to improve the representation of a real measurement process. This work proposes a semi-parametric virtual experiment model that operates in conjunction with measurement data and a Bayesian estimation procedure. We introduce a Gaussian process that accompanies our parametric virtual experiment. Heteroscedasticity and correlations between local points, often-ignored factors in data analyses that can have a significant impact on the accuracy of the analysis, are accounted for in the proposed statistical method. We explore estimation techniques to obtain a Bayesian estimator with a quantifiable uncertainty. Once defined, we apply the developed estimation technique to an example from optical metrology

where the current approach in literature, a deterministic multi-step optimisation strategy, can be translated to a semi-parametric model with an unknown link function.

166 Statistical Modeling of Clustering and Seasonality in Return Times of Midlatitude Cyclones

Merle Mendel, Roland Fried

TU Dortmund, Deutschland
merle.mendel@tu-dortmund.de

SECTION: Joint DStatG and TIES Session

Many extreme weather events, such as heavy precipitation or storms, tend to appear in temporal clusters and exhibit seasonal fluctuations, two behaviors that the stationary Poisson process cannot capture. In previous research, the fractional Poisson process (FPP) has been used to model occurrences of midlatitude cyclones, enabling the description of temporal clustering. However, since the standard FPP cannot account for seasonal behavior, most existing studies primarily focus on modeling winter cyclones, with less attention given to the other seasons. To address this limitation, we develop a modified approach that incorporates seasonality into the FPP.

We also evaluate different estimation methods for the parameters of the Mittag-Leffler distribution, which is the distribution of the return times of an FPP. We propose a new quantile-based estimator for the parameters and compare it with the existing estimation methods. The quantile-based estimator outperforms the widely used log-moments estimator in terms of mean squared error, while also offering shorter computing time and better robustness than the maximum likelihood estimator. We illustrate our approach by modeling the return times of midlatitude cyclones using climate reanalysis data.

167 Earth Observation Data and AI for Construction Statistics

Stefan Irrgang (1), Maren Köhlmann (1), Frederik Stellmach (2), Carola Stolle (2)

1: Statistisches Bundesamt, Deutschland; 2: Bundesamt für Kartographie und Geodäsie
Stefan.Irrgang@destatis.de

SECTION: Joint DStatG and TIES Session

Im Rahmen eines von Eurostat geförderten Projekts (Earth Observation Data and AI for Construction Statistics, EO4ConStat) werden Möglichkeiten geprüft, eine automatisierte Erkennung von Baustellen auf Luftbildern zur Qualitätssicherung in der Bautätigkeitsstatistik zu nutzen. Als Datenquellen für die Detektion von Baustellen mit Verfahren des maschinellen Sehens werden digitale Orthofotos genutzt. Neben der Baustelle selbst, sollen auch vordefinierte Bauphasen automatisch erkannt werden. Die anschließende Auswertung von Satellitendaten soll die zeitliche Einordnung von Baubeginn und -fertigstellung ermöglichen. Hierfür werden Veränderungen in den spektralen Eigenschaften der erkannten Baustellen genutzt. Diese lassen sich durch die Verwendung sogenannter spektraler Indizes ebenfalls maschinell auswerten. Die vorliegenden ersten Ergebnisse sind vielversprechend, zeigen aber auch Herausforderungen bei der Abgrenzung zu optisch ähnlichen aber inhaltlich verschiedenen Flächen auf. Sollten sich die getesteten Verfahren als zuverlässig bestätigen, kann die fernkundliche Erkennung von Baustellen zu Zwecken der Qualitätssicherung in der Statistik der Baubeginne und Baufertigstellungen eingesetzt werden. Dazu müssen die

Ergebnisse der maschinellen Erkennung sorgfältig validiert werden. Auch hierzu werden im Projekt geeignete Methoden entwickelt und getestet. Um einen möglichst repräsentativen Querschnitt verschiedener Baustellentypen, ihrer Umgebungen oder auch Bodenarten abzubilden, wurde als Testgebiet die gesamte Landesfläche von Nordrhein-Westfalen ausgewählt. Der Umfang der hierfür auszuwertenden Daten erhöht einerseits den Aufwand für die Prozessierung, ermöglicht aber andererseits eine Abschätzung zu den notwendigen Ressourcen für eine potentielle Anwendung im größeren Rahmen. Das Projekt ist eine Zusammenarbeit mit dem Bundesamt für Kartographie und Geodäsie (BKG) und dem Deutschen Zentrum für Luft- und Raumfahrt (DLR).

168 Multiple Imputation zur Umschätzung der Berufsklassifikationen der Deutschen Arbeitsmarktstatistik

Johannes Ludsteck

Institut für Arbeitsmarkt- und Berufsforschung (IAB) Nürnberg, Deutschland
Johannes.Ludsteck@iab.de

SECTION: Economic, Social and Market Statistics (Session 9): Labour Markets and Social Security III

Berufsklassifikationen sind eine wesentliche Grundlage für die Arbeitsmarktstatistik und für wissenschaftliche Analysen des Arbeitsmarktgeschehen. Im Jahr 2011 wurde die für die deutsche Arbeitsmarktstatistik relevante Klassifikation der Berufe KIDB88 auf eine deutlich aktualisierte Fassung (KIDB2010) umgestellt. Wegen erheblicher Unterschiede zwischen den Klassifikationen war die Erstellung einer eindeutigen Korrespondenz (Umsteigeschlüssel) nicht möglich. Davon sind insbesondere Berufe betroffen, die in den vergangenen 20 Jahren stark an Bedeutung gewonnen haben und sich deswegen stark ausdifferenziert haben, z.B. Software-, Internet- und Datenbankspezialisten.

Um diese Lücke zu füllen, entwickeln und implementieren wir für die Beschäftigtenstatistik des IAB einen Imputationsalgorithmus zur Umschlüsselung in beide Richtungen (neu nach alt und alt nach neu). Wir schätzen - für jeden 3-Steller der alten Klassifikation - ein Random-Forest-Modell zur Prognose der jeweiligen Zielberufe der neuen Klassifikation. Für die Gegenrichtung erfolgt die Schätzung separat für jeden 5-Steller. Die Hyperparameter der Random-Forests (Anzahl der Bäume und minimum-impurity-decrease) werden mittels 5-facher Kreuzvalidierung bestimmt. Bei der Implementation haben wir besonderes Gewicht auf die Generierung zahlreicher Prädiktoren mit hoher Prognosekraft gelegt. Neben bereits in den Rohdaten vorliegenden Merkmalen wie Alter, Arbeitsort, Geschlecht, Jobstatus (Teilzeit, Mini-Job), Bildung, Spelldauer und (Tages-)Lohnhöhe, leisten generierte Prädiktoren hohe Beiträge zur Prognosegüte. Das sind z.B. (1) Beschäftigungsanteile der fünf wichtigsten 3-Steller-Branchen im Betrieb, (2) Beschäftigungsanteile der fünf wichtigsten Ausländergruppen im Betrieb, (3) Beschäftigungsanteile von Qualifikationsgruppen, TZ-Beschäftigten, Frauen, (4) rollierende Dauern (relativ zur Beschäftigungsdauer) von TZ-Beschäftigung, geringfügiger Beschäftigung und Nicht-Beschäftigung, und (5) Betriebszugehörigkeitsdauer. Die stochastische Imputation wird kombiniert mit deterministischer Extrapolation. D.h. wenn die Codes von alter und neuer Klassifikation für ein Beschäftigungsverhältnis simultan vorliegen, wird der Code der alten Klassifikation in die Zukunft und der Code neuen in die Vergangenheit kopiert bis das Beschäftigungsverhältnis endet. Um zuverlässige Abschätzung des durch die Imputation induzierten Fehlers in Analysemodellen zu ermöglichen, erfolgt die Imputation multipel (mit fünf Replikationen).

169 Veränderte Arbeitszeitwünsche: Abbildung der Realität oder Effekte der Befragungsmethode? – Ergebnisse des Mikrozensus

Martina Rengers

Statistisches Bundesamt, Deutschland
martina.rengers@destatis.de

SECTION: Economic, Social and Market Statistics (Session 9): Labour Markets and Social Security III

Längere oder kürzere Arbeitszeiten? Bessere Work-Life-Balance, 4-Tage-Woche und mehr Zeit für Freunde, Familie, Hobbys, Reisen, ohne finanzielle Einbußen – das wünschen sich viele Erwerbstätige. Doch wie sieht es aus, wenn eine Reduzierung der Arbeitszeit mit einem entsprechend verringerten Einkommen verbunden ist, während umgekehrt eine Erhöhung der Arbeitszeit zu einem Verdienstanstieg führen würde? Gibt es hier einen gesellschaftlichen Wandel hin zu einem stärkeren Wunsch nach Arbeitszeitreduzierung trotz Verdienstaufschlag?

Genau diese Frage ist Gegenstand vieler Studien und Umfragen, ebenso wie auch der Mikrozensus Arbeitszeitwünsche von Erwerbstätigen unter Berücksichtigung entsprechender Verdienstanpassung erfasst. Allerdings hatten bereits frühere Untersuchungen zum Unterschied verschiedener Umfrageergebnisse gezeigt, wie sensibel befragte Personen auf Frageformulierungen, Fragereihenfolge und Abfragekontext bei der Ermittlung von Arbeitszeitwünschen reagieren. Mit der Neugestaltung des Mikrozensus ab 2020 wurde ein Multi-Mode-Erhebungsdesign geschaffen. Seitdem ist es möglich, den Mikrozensus online zu beantworten. Neben CAWI stehen jedoch weiterhin das persönliche Interview CAPI, das telefonische Interview CATI sowie der Papierfragebogen PAPI als Befragungsmethoden zur Verfügung.

Mit der Einführung des neuen Mikrozensus und in den nachfolgenden Berichtsjahren haben sich die Ergebnisse zu den Arbeitszeitwünschen der Erwerbstätigen erheblich verändert und im Vergleich zum Berichtsjahr 2019 sogar umgekehrt. Von den 10% Erwerbstätigen mit Arbeitszeitdiskrepanzen im Jahr 2019, wollten fast 66% ihre Arbeitszeit erhöhen, entsprechend umgekehrt knapp 34% von ihnen die Arbeitszeit verringern. Im Jahr 2023 lag die Diskrepanzenquote bei fast 14% der Erwerbstätigen und nur noch 24% wollten eine Arbeitszeiterhöhung (Unterbeschäftigte), rund 76% dagegen eine Arbeitszeitverringerung (Überbeschäftigte). Das Verhältnis von Unter- und Überbeschäftigten hat sich also umgekehrt.

Nach Corona-Pandemie, Work-Life-Balance-Diskussionen und zunehmender Arbeitsmarktbeteiligung der Generation Z wäre eine stärkere Fokussierung auf mehr Freizeit denkbar. Auffällig ist andererseits, dass es in allen Altersklassen und nach Geschlecht deutliche Unterschiede zwischen den Ergebnissen von CAWI und NICHT CAWI gibt. Sind die veränderten Arbeitszeitwünsche also Abbildung der Realität oder ein Methodeneffekt? Dieser Beitrag versucht, eine Antwort darauf zu finden.

170 Saisonalität am Arbeitsmarkt

Gerald Seidel

Statistik der Bundesagentur für Arbeit, Deutschland

Gerald.Seidel@arbeitsagentur.de

SECTION: Economic, Social and Market Statistics (Session 9): Labour Markets and Social Security III

Die Veränderung des Saisonmusters kann bei verschiedenen soziodemographischen Größen beobachtet werden. In diesem Beitrag wird die Veränderung der Saison am Arbeitsmarkt dargestellt, die damit verbundenen Herausforderungen bei der Saisonbereinigung erörtert und der Versuch unternommen, Erklärungsansätze für die sich verändernde Saison am Arbeitsmarkt beizutragen.

171 Arbeitsmarktkonzentration in Deutschland

Martin Popp (1), Michael Oberfichtner (1,2)

1: Institut für Arbeitsmarkt- und Berufsforschung, Deutschland; 2: Friedrich-Alexander-Universität

Erlangen-Nürnberg

martin.popp@iab.de

SECTION: Poster Session

Arbeitsmarktkonzentration misst, wie stark Neueinstellungen in einem lokalen Arbeitsmarkt auf wenige Betriebe konzentriert sind. Anhand von Sozialversicherungsdaten dokumentieren wir in unserer Studie erstmalig das Ausmaß der Arbeitsmarktkonzentration in Deutschland.

Wir zeigen, dass der durchschnittliche deutsche Arbeitsmarkt für Neueinstellungen – definiert für Kombinationen aus Berufsgruppen, Anforderungsniveaus und Pendlerzonen – stark konzentriert ist ($\text{HHI}=0,257$). Nach den kartellrechtlichen Schwellenwerten der EU weisen 56 Prozent dieser Arbeitsmärkte eine moderate oder hohe Konzentration auf. In diesen Märkten, die häufig nur vergleichsweise klein ausfallen, sind 9 Prozent der Erwerbstätigen beschäftigt. Wir dokumentieren außerdem, dass die durchschnittliche Arbeitsmarktkonzentration in unserem Analysezeitraum der Jahre 2012 bis 2023 relativ stabil geblieben ist.

Beruflich und regional bestehen große Unterschiede in der Arbeitsmarktkonzentration. In den akademischen Berufen, künstlerischen Berufen und Berufen des Verarbeitenden Gewerbes fällt der Anteil der Beschäftigten in konzentrierten Arbeitsmärkten überdurchschnittlich hoch aus. Besonders groß ist der Anteil außerdem in weiten Teilen Nord- und Ostdeutschlands.

Personen mit Spezialisten- oder Expertentätigkeiten sind am häufigsten in Arbeitsmärkten mit moderater oder hoher Konzentration tätig. Gleichzeitig sind Personen in Helfer-Jobs häufiger von moderater oder hoher Konzentration betroffen als Fachkräfte.

Vor dem Hintergrund, dass sich eine hohe Arbeitsmarktkonzentration häufig in niedrigeren Löhnen und einer geringeren Beschäftigungsqualität niederschlägt, verdeutlicht unsere Studie, dass bestehende Richtlinien für den Wettbewerb auf dem Gütermarkt explizit auch auf den Arbeitsmarkt angewandt werden sollten.

172 Bevölkerungsprojektionen mit dem R-Paket propop

Norah Efosa, Adrian Gadiant

Statistik Aargau, Departement Finanzen und Ressourcen, Kanton Aargau, Schweiz
norah.efosa@ag.ch

SECTION: Poster Session

Das Ziel von propop ist die Erstellung von Bevölkerungsprojektionen mit der Kohortenkomponentenmethode. Die Funktion propop::propop() generiert Projektionen für unterschiedliche demografische Gruppen und kann für verschiedene Szenarien sowie auf unterschiedlichen räumlichen Ebenen (z.B. Land, Bezirke oder Gemeinden) angewendet werden.

173 Methodische Grundlagen und Herausforderungen einer bundesweiten Kulturstatistik

Marlon Suárez

Statistisches Bundesamt, Deutschland
marlon.suarez@destatis.de

SECTION: Poster Session

Das Projekt „Bundesweite Kulturstatistik“ widmet sich seit 2014 im Auftrag der Kultusministerkonferenz (KMK) und der Beauftragten der Bundesregierung für Kultur und Medien (BKM) dem Aufbau eines konsistenten kulturstatistischen Datenangebotes. Ziel ist es, die Grundlage für eine evidenzbasierte Kulturpolitik zu stärken und langfristig ein belastbares Monitoring der kulturellen Entwicklung in Deutschland zu etablieren. Das bis Ende 2027 laufende Projekt verfolgt außerdem das Ziel, eine dauerhafte Kulturberichterstattung mit gesetzlicher Grundlage bereitzustellen.

Ein zentrales methodisches Anliegen besteht in der Harmonisierung heterogener Datenquellen und der Operationalisierung des vielschichtigen Kulturbegriffs für die amtliche Statistik. Hierzu wurde im Rahmen des Projekts ein indikatorenbasiertes Konzept entwickelt, das sich insbesondere auf öffentlich verfügbare Daten stützt. Die jährliche Veröffentlichung „Kulturindikatoren“ verdeutlicht, wie anhand statistischer Kennzahlen, etwa zu Ausgaben, Angebotsstrukturen oder Erwerbstätigkeit, kulturrelevante Informationen systematisiert und aufbereitet werden.

Ein ergänzendes methodisches Element bildet die im Jahr 2023/2024 durchgeführte Nutzendenumfrage, deren Ergebnis unter anderem zeigt, welche Anforderungen Kulturakteure an die Datenlage stellen. So besteht seitens der Nutzenden ein hohes Interesse an regional aufgeschlüsselten, aktuellen und vergleichbaren Daten sowie an der Abbildung kultureller Teilhabe.

Auf europäischer Ebene ist das Statistische Bundesamt in die Aktivitäten der Eurostat Expert Group on Culture Statistics eingebunden und bringt dort seine Fachexpertise zur konzeptionellen Abgrenzung und statistischen Erfassung kultureller Aktivitäten ein. Im Mittelpunkt steht derzeit die Überarbeitung des europäischen Kulturstatistik-Frameworks, das unter anderem eine Ausweitung auf die Kreativwirtschaft, die Weiterentwicklung funktionaler Abgrenzungen sowie die Anpassung an die überarbeitete europäische Wirtschaftszweigklassifikation NACE Rev. 2.1 vorsieht. Die dabei diskutierten Konzepte und Definitionen könnten auch Impulse für eine nationale Darstellung wirtschaftlicher Aktivitäten in der Kultur- und Kreativwirtschaft liefern.

174 Methoden für die Integration von Mobilfunkdaten mit weiteren Datenquellen

Gloria Deetjen, Natalie Rosenski

Statistisches Bundesamt (Destatis), Deutschland
Gloria.Deetjen@destatis.de

SECTION: Poster Session

Mobilfunkdaten sind aufgrund ihrer zeitlichen Aktualität sowie ihrer räumlichen Genauigkeit eine wertvolle Datenquelle und kommen bisher vor allem in Machbarkeitsstudien und experimenteller Statistik zum Einsatz. Besonders wertvoll ist die Verknüpfung von Mobilfunkdaten mit weiteren – sowohl neuen digitalen als auch traditionellen – Datenquellen, denn neben Qualitätsverbesserungen besteht auch Potenzial für neue Statistiken. Für eine mögliche Nutzung von Mobilfunkdaten in der amtlichen Statistik sind sowohl ebendiese Qualitätsverbesserungen als auch weiterführende Methodik notwendig. Im ESSnet-Projekt MNO-MINDS (Mobile Network Operator (data) - Methods for Integrating New Data Sources) werden mithilfe einer Feldanalyse die relevantesten Datenquellen identifiziert, die sich für eine Verknüpfung mit Mobilfunkdaten eignen. Zusätzlich wird ein methodischer Referenzrahmen und ein Fragebogen entwickelt, welche eine Verringerung von Messfehlern und Missklassifikation ermöglichen. In dem Beitrag werden einige Wochen vor Abschluss des Projekts viele wertvolle Ergebnisse vorgestellt.

175 Evaluating methods to prevent and detect inattentive responding in web surveys

Lukas Olbrich (1), Joseph W. Sakshaug (1,2,3), Eric Lewandowski (4)

1: IAB Nürnberg; 2: LMU München; 3: Universität Mannheim; 4: NYU
lukas.olbrich@iab.de

SECTION: Poster Session

Inattentive respondents pose a substantial threat to data quality in web surveys. In this study, we evaluate methods for preventing and detecting inattentive responding and investigate its impacts on substantive research. We use data from two large-scale non-probability surveys fielded in the US. Our analysis consists of four parts: First, we experimentally test the effect of asking respondents to commit to providing high-quality responses at the beginning of the survey on various data quality measures (attention checks, item nonresponse, break-offs, straightlining, speeding). Second, we conducted an additional experiment to compare the proportion of flagged respondents for two versions of an attention check item (instructing them to select a specific response vs. leaving the item blank). Third, we propose a timestamp-based cluster analysis approach that identifies clusters of respondents who exhibit different speeding behaviors and in particular likely inattentive respondents. Fourth, we investigate the impact of inattentive respondents on univariate, regression, and experimental analyses. First, our findings show that the commitment pledge had no effect on the data quality measures. As indicated by the timestamp data, many respondents likely did not even read the commitment pledge text. Second, instructing respondents to leave the item blank instead of providing a specific response significantly increased the rate of flagged respondents. Third, the timestamp-based clustering approach efficiently identified clusters of likely inattentive respondents and outperformed a related method, while providing additional insights on speeding

behavior throughout the questionnaire. Fourth, we show that inattentive respondents can have substantial impacts on substantive analyses.

176 pRDM: A note on a popular collusion marker

Christian Mueller

Constructor University, Deutschland
cmueller01@constructor.university

SECTION: Poster Session

Statistical tools for detecting fraud in procurement processes have gained well-deserved attention. Detecting fraud not only saves considerable expenditures but also helps to avoid legal fees. Statistical “markers” that indicate misconduct hence are efficient and established means of protecting procurers. However, despite their popularity sometimes only little is known about their statistical properties and hence their potential to deliver correct signals. In this note, we will discuss the popular so-called relative distance measure (RDM). We show that it is prone to false-positive bias and suggest an amendment, pRDM, that improves its performance by up to one third.

177 Harnessing the power of Input-Output analysis for sustainability. A simulation study based on US data

Ulf von Kalckreuth

Deutsche Bundesbank, Deutschland
ulf.von-kalckreuth@bundesbank.de

SECTION: Poster Session

Measuring carbon contents reliably, for products, firms and industries, is key for identifying climate change related transition risks. Phase 3 of the G20 Data Gaps Initiative requests the collection of emission data and multiregional Input-Output (IO) tables to enable the calculation of aggregate carbon contents. What sectoral distinctions do we need – and at what level of granularity? Do we need information on technology? How can statistical data be used in carbon accounting? Based on IO tables and company-level data from the United States (US), I construct a micro simulation environment that can act as a laboratory for answering these questions. The database consists of almost 5000 units located (with few exceptions) in the US and Canada. The analysis focuses on indirect emissions and carbon contents. For levels of aggregation typical of IO tables, the within-sector heterogeneity of carbon contents is very high in some industries. Still, averages can be very useful for company-level carbon accounting. Statistical data can provide consistent starting values for inputs in cases where direct information from providers is missing. Specifically, they may be used to approximate indirect emissions of suppliers, when company-level information on their direct emissions is available. This will be the standard case in the European Union (EU), once upcoming reporting requirements are in place..

178 Das R-Paket „deseats“ zur simultanen Schätzung von Trend und Saisonalität in Zeitreihen anhand datengesteuerter lokal gewichteter Regression

Dominik Schulz

Universität Paderborn, Deutschland
dominik.schulz@uni-paderborn.de

SECTION: Poster Session

Das bereits auf dem Comprehensive R Archive Network (CRAN) veröffentlichte R-Paket „deseats“ (deseasonalize time series) dient der vollkommen datengesteuerten Komponentenerlegung von Zeitreihen in Trend-, Saison- und Restkomponente anhand lokal gewichteter Regression, deren benötigte Bandbreite zur simultanen Schätzung von Trend und Saisonalität anhand eines iterativen Plug-In-Algorithmus, getauft „DeSeaTS“, zuvor aus den Daten geschätzt wird. Bei dieser Bandbreitenauswahl werden kurzfristige Abhängigkeitsstrukturen im Fehlerterm erlaubt wie z.B. in Form von ARMA-Modellen, ohne dass jedoch eine solch explizite Fehlermodellform unterstellt wird, da notwendige Größen in der Bandbreite vollumfänglich nichtparametrisch ermittelt werden. Während etablierte Ansätze wie das Berliner Verfahren 4.1 (BV4.1), X-12-ARIMA und TRAMO-SEATS ausschließlich im Sinne volkswirtschaftlich wichtiger vierteljährlicher oder monatlicher Zeitreihen Anwendung finden, ist DeSeaTS auf Zeitreihen mit allen ganzzahligen saisonalen Frequenzen anwendbar.

Eine Erweiterung der Idee auf ein saisonales semiparametrisches ARMA-Modell bietet die Möglichkeit, die Residuen der Komponentenerlegung durch ein ARMA-Modell zu analysieren, um anschließend Punkt- und Intervallprognosen zu bestimmen.

Empirische Anwendungen von DeSeaTS auf real beobachtete Zeitreihen aus verschiedenen Bereichen, wie unter anderem der Makroökonomie und der Umwelt, zeigen adäquate Komponentenerlegungen, Saisonbereinigungen und plausible Punkt- und Intervallprognosen. Auch unter sich langsam verändernden Saisonkomponenten bleiben die Ergebnisse, obwohl in der Bandbreitenformel in DeSeaTS eine exakt periodische Saisonkomponente unterstellt wird, qualitativ gut. Eine abschließende Simulationsstudie bestätigt einerseits die Güte der Bandbreitenauswahl durch DeSeaTS und vergleicht die Performance zur Schätzung von Trend-, Saison- und kombinierter Trend-Saison-Komponente zwischen DeSeaTS, dem BV4.1-Basismodell, X-12-ARIMA, TRAMO-SEATS, STL und der R-internen Funktion „decompose“, wobei in allen Fällen nach Möglichkeit entweder in den Programmen auswählbare Automatisierungen oder in der Literatur empfohlene Einstellungen genutzt werden. Es zeigt sich anhand der Simulationsstudie, dass DeSeaTS mit lokal kubischem Trend für Trend, Saison und Trend-Saison konsistente Schätzer liefert, auch unter sich langsam verändernder Saisonalität, und anhand des MSE-Kriteriums entweder ähnlich gut oder teilweise besser funktioniert zur Schätzung dieser Komponenten als etablierte Zerlegungsmethoden.

179 NeDaMo - Neue Daten für Mobilitätsanalysen

Markus Zwick, Natalie Rosenski, Johannes Volk

Statistisches Bundesamt (Destatis), Deutschland
Gloria.Deetjen@destatis.de

SECTION: Poster Session

Mobilfunkdaten erlauben wesentlich erweiterte Erkenntnisse im Bereich von Mobilitätsanalysen, weil sie hochaktuell, kleinräumig und dynamisch vorliegen. Bisherige Forschungsarbeiten dazu haben jedoch ergeben, dass die Kundenstruktur zwischen den in Deutschland aktiven Mobilfunknetzbetreibern systematische Differenzen aufweist, etwa nach Alter, Geschlecht und Wohnort. Weil die Daten der einzelnen Mobilfunkunternehmen zusätzlich Non-Probability-Stichproben sind, ist eine verlässliche Hochrechnung auf die Mobilität der Gesamtbevölkerung nicht ohne Zusatzinformationen möglich. Das vom Bundesministerium für Digitalisierung und Staatsmodernisierung (BMDS) im Rahmen der Innovationsinitiative „mFUND“ geförderte Projekt „Neue Daten für Mobilitätsanalysen“ (NeDaMo) zielt darauf ab, diese Verzerrungen in Mobilfunkdaten mit einer maßgeschneiderten Erhebung nach §7 Bundesstatistikgesetz auszugleichen. Dazu werden auf freiwilliger Basis bis zu 20.000 Einheiten zu ihrer Mobilfunknutzung und ihrem Mobilitätsverhalten befragt. Anschließend werden in Kombination mit weiteren Statistiken Verfahren zur statistisch validen Nutzung von Mobilfunkdaten entwickelt, um diese Datenquelle zur qualitativ hochwertigen Nutzung im Rahmen des Bundesverkehrswege- und Mobilitätsplan zu erschließen. Zuletzt werden qualitativ optimierte Verflechtungsmatrizen im Rahmen von Mikrosimulationen pilothaft für ausgewählte Regionen zur Prognose von Mobilitätsszenarien genutzt.

Projektsteckbrief bei „mFUND“:

<https://www.bmv.de/SharedDocs/DE/Artikel/DG/mfund-projekte/nedamo.html>

180 Hotspotanalyse der Automobilindustrie in Deutschland

Dirk Brunken (2), Corinna Mundzeck (2), Christian Salwiczek (2), Julian Schulz (1)

1: Landesamt für Statistik Niedersachsen, Deutschland; 2: Statistisches Amt für Hamburg und Schleswig-Holstein, Deutschland

julian.schulz@statistik.niedersachsen.de

SECTION: Regional Statistics (Session 1)

Der Vortrag präsentiert den Werkstattbericht zu einer räumliche Hotspotanalyse der Automobil- und Automobilzuliefererindustrie in Deutschland auf Basis von Daten aus dem Unternehmensregister der amtlichen Statistik. Ziel der Analyse ist die Identifikation von Clustern mit überdurchschnittlicher Unternehmensdichte, losgelöst von administrativen Grenzen. Mittels raumstatistischer Methoden werden geografische Konzentrationen dieser Industriebranchen sichtbar gemacht und in Form einer Hotspotkarte dargestellt. Aufbauend auf diesen identifizierten Hotspotflächen erfolgt eine weiterführende Analyse der dort ansässigen Unternehmen. Die Ergebnisse liefern Einblicke in die aktuelle Struktur und regionale Bedeutung der Automobilwirtschaft in Deutschland. Eine Veröffentlichung der Auswertungen als kleinräumige Informationen im Statistikportal von Bund und Ländern ist für Ende des Jahres geplant.

181 Methodische Ansätze zur Modellierung der Lage in Mietspiegeln

Georg Wiegleb

Landeshauptstadt Magdeburg, Deutschland
georg.wiegleb@stat.magdeburg.de

SECTION: Regional Statistics (Session 1)

Die Lage einer Mietwohnung zählt zu den zentralen Einflussfaktoren auf die Miethöhe – zugleich besteht wenig Einigkeit darüber, wie sie sachgerecht und empirisch fundiert in Mietpreismodellen abgebildet werden kann. §19 der Mietspiegelverordnung verweist auf „vor Ort feststellbare Faktoren“ wie Bebauungsdichte, Infrastruktur oder Begrünung als Grundlage der Lageermittlung. Die Verwendung adressgenauer Indikatoren – etwa Laufwege, Distanzen, Kerndichten oder Flächenanteile im Umkreis – kann eine geographisch hochauflösende und plausible Annäherung an diese Vorgabe darstellen. Gleichzeitig ergeben sich je nach Datenmenge, Berücksichtigung des räumlichen Bezugs und Modellintegration methodische Herausforderungen. Der Vortrag stellt zwei mögliche Herangehensweisen zum Umgang mit adressgenauen Lageindikatoren vor: zum einen die vorangestellte Bündelung lagebezogener Informationen zu stabilen, geographisch kohärenten Einheiten. Alternativ dazu kann eine direkte Einbettung der Indikatoren bzw. zusammenfassenden Größen als kontinuierliche erklärende Größen in ein Regressionsmodell erfolgen. Beide Wege werden hinsichtlich ihrer Anforderungen, Modellintegration, Anwendbarkeit und Verständlichkeit im Kontext der Mietspiegelerstellung gegenübergestellt. Ziel ist es, den methodisch herausfordernden Aspekt der Lagemodellierung durch eine datenbasierte und strukturierte Vorgehensweise nachvollziehbar und transparent zu gestalten.

182 HILDA - Ein Chatbot für die Landesdatenbank NRW

Tobias Wolfanger

IT.NRW, Deutschland
tobias.wolfanger@it.nrw.de

SECTION: Regional Statistics (Session 1)

HILDA (Hilfreiche Interaktive Landesdatenbank-Assistentin) ist ein bei IT.NRW entwickelter Chatbot der es Datennutzenden ermöglicht statistische Fragestellungen direkt an die Landesdatenbank NRW (LDB), das Informationssystem des Statistischen Landesamts NRW, zu richten. Durch die Verwendung eines großen Sprachmodells werden die auf die Anfrage passenden Tabellen in der Datenbank identifiziert und auf die für die Beantwortung der Frage relevanten Inhalte reduziert. Der neue KI-gestützte Ansatz basiert auf der Idee, statistische Informationen möglichst einfach zugänglich zu machen und Barrieren, wie die Navigation in einem komplexen Datenbanksystem, unübersichtliche Tabellenstrukturen oder das Erfordernis zur Kenntnis statistischer Fachbegriffe aus dem Weg zu räumen.

Der Vortrag stellt die auf dem Retrieval-Augmented Generation (RAG) basierende Funktionsweise von HILDA live vor und ordnet die bei der Entwicklung getroffenen Designentscheidungen vor dem Hintergrund des Spannungsfelds zwischen generativer KI und dem Selbstverständnis der amtlichen Statistik als verlässliche Produzentin qualitativ hochwertiger statistische Informationen ein.

183 Die Entwicklung der Energiebilanzierung in den Bundesländern

Tom Hoger, Tobias Diergardt, Dennis Quandt

Geschäftsstelle Länderarbeitskreis Energiebilanzen, Statistisches Landesamt Baden-Württemberg, Deutschland
tom.hoger@stala.bwl.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 2)

Energiebilanzen geben Auskunft über die energiewirtschaftlichen Aufkommen, Umwandlungen und Verbräuche von Energie aufgeteilt nach Energieträgern und Wirtschaftssektoren. Auf Grundlage von Energiebilanzen werden CO₂-Bilanzen erstellt. Ein Großteil der Daten für die Energiebilanz wird von der amtlichen Statistik erhoben und ist durch das Energiestatistikgesetz legitimiert. Somit bilden Energiebilanzen eine objektive Grundlage für politische Entscheidungen zur Zusammensetzung und Sicherstellung der Energieversorgung sowie Bewertung von Klimazielen. Seit über 60 Jahren werden in den Bundesländern Energiebilanzen erstellt. In Folge der Ölpreiskrise 1973 wird seit Ende der 1970er Jahre die Erstellung der Bilanzen gemeinsam im Länderarbeitskreis Energiebilanzen koordiniert und eine einheitliche Methodik verfolgt. Die letzten 60 Jahre sind in Deutschland durch erhebliche energiepolitische, wirtschaftliche und technologische Änderungen gekennzeichnet. Der Vergleich der Daten der Energiebilanzen mit geschichtlichen Ereignissen der letzten 60 Jahre am Beispiel Baden-Württemberg zeigt, dass sich globale Ereignisse und politische Entscheidungen (z.B. Kohlekrise 1957, Ölpreiskrise 1973, Nuklearkatastrophe von Tschernobyl 1986, Erneuerbare-Energien-Gesetz 2000, Atomausstieg 2000, Nuklearkatastrophe von Fukushima 2011, Covid-19-Pandemie 2020) in den Daten der Energiebilanzen wiederfinden lassen und methodische Anpassungen erforderlich machten. Derzeit sind Fragen zur sicheren, nachhaltigen und ressourcenschonenden Energieversorgung, Energiewende, Energiepreiskrise und Dunkelflaute gesellschaftlich und politisch so präsent wie nie. Aktuell und zukünftig steht die Energiebilanzierung vor neuen Herausforderungen und Chancen (z.B. Smart-Meter-Daten, Wasserstoffherhebung, Big Data, künstliche Intelligenz, Datenaktualität, Desinformation).

184 Measuring the carbon footprints of portfolios. Emission intensities and financed emissions indicators from Multi-Regional IO and Air Emission Accounts data

Ulf von Kalckreuth

Deutsche Bundesbank, Deutschland
ulf.von-kalckreuth@bundesbank.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 2)

Total carbon content of output, defined as the sum of direct and upstream indirect emissions, is the most comprehensive measure for transition risk, as it captures all emissions due to the production of this output and yields a direct metric for the sensitivity of emission of carbon taxes. This paper has a focus on financial portfolios and is meant to be a hands-on guide. It describes how industry level data on production interlinkages and on direct emissions can be used to generate estimates for indirect emission intensities for aggregate portfolios. Essentially, this is done by attributing industry level intensities to entities and computing weighted averages of these intensities.

The first part of this essay explains the nature of indirect emissions and how they are constituted by production interlinkages in the context of a linear production model. The second part describes how to approximate these expressions based on aggregate data. A third part looks at results for industries, groups of counties, groups of products, with a closer look on Germany, and the components of carbon content for the output or Euro area countries are computed as a time series. A final section presents an important application: the weighted average carbon intensity of banks in Germany. Emission intensities are matched with outstanding bank loans according to industry compounds, as provided by the German balance sheet statistics. It turns out that bank lending in Germany is characterised strong and persistent bias towards high emission industries. This bias is entirely due to a disproportionate reliance of the energy, water supply and disposal industries on bank financing.

185 Sensibilität für versunkene Kosten am Beispiel des Mausverhaltens in der Restaurant-Row-Aufgabe.

Jekaterina Zukovska

Humboldt-Universität zu Berlin, Deutschland
zukovskj@hu-berlin.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 2)

Versunkene Kosten sind irreversible Investitionen – etwa bereits aufgewendete Zeit oder Ressourcen – die durch zukünftige Entscheidungen nicht mehr rückgängig gemacht werden können. Aus rational-ökonomischer Sicht sollten solche Kosten keinen Einfluss auf aktuelles Entscheidungsverhalten haben. Dennoch zeigen zahlreiche Studien, dass sowohl Menschen als auch Tiere in ihrem Verhalten sensitiv auf versunkene Kosten reagieren. Besonders ausgeprägt ist dies im Umgang mit Zeitkosten, wenn Individuen Aktivitäten fortsetzen, obwohl diese objektiv nicht mehr lohnend erscheinen – allein aufgrund der bereits investierten Zeit.

In dieser Studie untersuchen wir die zeitbezogene Sensitivität gegenüber versunkenen Kosten bei Mäusen mithilfe eines automatisierten, verhaltensbasierten Entscheidungsparadigmas – der sogenannten Restaurant-Row-Aufgabe. In diesem Paradigma bewegen sich Mäuse durch ein ringförmiges Labyrinth, in dem sie vier „Restaurants“ in einer festgelegten Reihenfolge nacheinander besuchen. Jedes Restaurant bietet Belohnungen in unterschiedlicher Menge. Die Tiere können nicht frei wählen, welches Restaurant sie aufsuchen, wohl aber entscheiden, ob sie das aktuelle Angebot annehmen oder zum nächsten Restaurant weitergehen.

Jedes Restaurant besteht aus einer Angebots- und einer Wartezone. In der Angebotszone wird akustisch signalisiert, wie lange die Wartezeit bis zur Belohnung beträgt (1–30 Sekunden). In der Wartezone können die Mäuse die Wartezeit durchhalten oder abbrechen.

Unsere Ergebnisse zeigen, dass die Entscheidung der Tiere nicht nur vom verbleibenden Countdown, sondern auch von der bereits verbrachten Wartezeit beeinflusst wird – ein klarer Hinweis auf Sensitivität gegenüber versunkenen Kosten.

Die automatisierte Versuchsanordnung, entwickelt im Winterlab, erlaubt eine effiziente Datenerhebung, flexible Anpassung der Versuchsparameter und eine einfache Übertragbarkeit auf andere Spezies.

186 Goodness-of-fit Tests for Exponential and Weibull Distributions in Progressively Type-II Censored Data

Anja Bettina Schmiedt (1), Narayanaswamy Balakrishnan (2), Erhard Cramer (3)

1: OTH Regensburg University of Applied Sciences, Regensburg, Germany; 2: McMaster University, Hamilton, Ontario, Canada; 3: RWTH Aachen University, Aachen, Germany
anja.schmiedt@oth-regensburg.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 2)

In progressive Type-II censoring, n items are subjected to a life test. Upon the first failure, a predefined number of items are randomly selected and removed from the experiment. This process continues with each subsequent failure until the m -th failure occurs, at which point the life test concludes, and the remaining items are censored. While exponential distributions are widely used as underlying lifetime distributions, Weibull distributions are also prevalent, necessitating the evaluation of their fit based on available data. Although goodness-of-fit tests for exponentiality are well-established in progressively Type-II censored data, those for Weibull distributions have received limited attention.

One popular approach for testing exponentiality relies on the spacings of the observed failure times. In this talk, we follow this approach and introduce novel goodness-of-fit tests utilizing scale-invariant statistics derived from the Mahalanobis norm of normalized failure times. This leads us to the Euclidean norm (among others) of spacings in progressively Type-II censored data based on the standard uniform distribution. We present results from a power study that evaluates our novel tests against existing benchmarks across various alternative distributions and censoring plans, demonstrating superior performance in scenarios with small to moderate sample sizes. Additionally, we extend our methodology to approximate goodness-of-fit tests for Weibull distributions, ensuring robustness concerning the approximated significance level under unknown shape parameters. Finally, we illustrate the practical applicability of our tests through a data example and discuss further research questions that arise from our approach.

187 Visualisierung schiefer Verteilungen am Beispiel von Einkommen und Miete

Michael Neutze, Tobias Lentz

Statistisches Bundesamt, Deutschland
tobias.lentz@destatis.de

SECTION: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 1)

Die regionale Einkommensverteilung ist ein zentraler Indikator, um die Lebensverhältnisse in den verschiedenen Regionen Deutschlands zu messen. Untersuchungen zeigen jedoch, dass in Regionen mit höheren verfügbaren Einkommen die Kosten für Wohnraum höher sind. Im Mieterland Deutschland sind für einen großen Teil der Bevölkerung die Mietpreise relevante Fixkosten, die die für den Vermögensaufbau und den privaten Konsum verfügbaren Finanzmittel maßgeblich beeinflussen.

Wir betrachten daher sowohl das 2022 verfügbare Einkommen der privaten Haushalte auf Kreisebene, das auf den Volkswirtschaftlichen Gesamtrechnungen der Länder basiert, als auch Bestandsmieten in

den deutschen Gemeinden auf Basis des Zensus 2022. Beide Indikatoren weisen schiefe Verteilungen auf. Bei der Visualisierung dieser Daten werden zumeist Choroplethenkarten verwendet. Dies ist zur Darstellung der geografischen Verteilung sehr nützlich, aus methodischer Sicht ergeben sich allerdings Probleme. Für Nutzende sind z.B. einzelne Ausprägungen verschiedener Regionen nur schwer vergleichbar und die Bevölkerungsgröße – die einen maßgeblichen Einfluss auf die regionalen Mietpreise haben kann – ist nicht direkt ersichtlich.

Mit dem Beeswarm-Diagramm existiert eine Darstellungsform, die die Datenvisualisierung schiefer regionalstatistischer Verteilungen bereichern kann. Der Beeswarm ist eine univariate Visualisierung, die sich methodisch am Boxplot orientiert. Gegenüber dem klassischen Boxplot zeichnet er sich durch eine ästhetischere Gestaltung und eine leichtere Zugänglichkeit auch für statistisch nicht geschulte Nutzende aus. Wir zeigen mit interaktiven Beeswarm-Diagrammen für das verfügbare Einkommen und die Bestandsmieten, dass diese Darstellungsform eine sehr sinnvolle Ergänzung im visuellen Vokabular der Datenvisualisierung ist.

Im Vortrag soll die technische Implementierung und Optimierung der auf der JavaScript-Bibliothek d3 basierenden Webanwendung gezeigt und diskutiert werden, wie diese interaktiven Darstellungsformen die Informationsvermittlung komplexer regionalstatistischer Ergebnisse bereichern können. Anwendungsbeispiel sind regionalisierte amtliche Daten zu verfügbarem Einkommen und Mieten, zwei zentralen Variablen für das Vermögen und den Konsum privater Haushalte.

188 Einkommensverteilung in SILC

Gerit Offermann

Destatis, Deutschland
gerit.offermann@destatis.de

SECTION: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 1)

EU-SILC ist die Erhebung für vergleichbare Daten auf europäischer Ebene über Einkommen, Armut sowie Lebensbedingungen. Die Erhebung ist seit 2020 in den Mikrozensus integriert und repräsentativ für die gesamte Bevölkerung in Deutschland. Für die Vergleichbarkeit auf europäischer Ebene wurden gemeinsame Zielvariablen und Erhebungsstandards festgelegt. Die Harmonisierung der europäischen Zielvariablen spiegelt nicht immer landestypische Einkommensarten ausreichend wider. Dies betrifft insbesondere die Sozialleistungen, so dass ein ungenaues Bild entsteht. In Deutschland gibt es Leistungen mit versicherungsähnlichem Charakter, so wie Arbeitslosengeld, Leistungen mit einkommensunabhängigem Charakter, so wie Kindergeld, oder Sozialleistungen, die den Grundbedarf der Bevölkerung abdecken sollen, so wie Bürgergeld oder Kinderzuschläge. Die Möglichkeit die Sozialleistungen entsprechend ihrem Charakter darzustellen ist Inhalt dieser Präsentation.

Im Statistischen Bundesamt werden zusätzlich zu den veröffentlichten europäischen Zielvariablen differenziertere Einkommenskomponenten auf Personen- und Haushaltsebene vor. Um das Bild der Einkommen aus Sozialleistungen in Deutschland zu vervollständigen, werden die Einkommenskomponenten neu zusammengefasst. Dies ermöglicht die nationalen Besonderheiten verschiedener Einkommensbestandteile besser abbilden können. Für diese andere Aufgliederung des Haushaltseinkommens werden vergleichende Analysen durchgeführt. Hierbei werden verschiedene Gruppen von Einkommensempfängern gebildet und die Zusammensetzung der jeweiligen Einkommensquellen vorgestellt. Es werden regionale Unterschiede und Gemeinsamkeiten dargestellt. Neben einer Vorstellung der Einkommensanteile werden die Verteilung der Vorkommen thematisiert. Darüber hinaus werden Gleichheitsmaße für verschiedene Gruppen gegenübergestellt.

Ziel ist es, die in Deutschland spezifischen Gruppen von Sozialleistungen besser abzubilden. Im Zuge dessen kann auch die Bedeutung verschiedener Sozialleistungen für unterschiedliche Haushaltstypen klarer herausgearbeitet werden. Die Präsentation vervollständigt damit das Bild über die bereits veröffentlichten Einkommensverteilungen aus SILC.

189 Wer trägt die Einkommensteuerlast in Deutschland?

Robert Buhles, **Roland Jeske**

Hochschule Kempten, Deutschland
robertbuhles@gmx.de

SECTION: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 1)

Die Einkommensteuer stellt eine der wichtigsten Einnahmequellen des Staates in Deutschland dar. Der Vortrag untersucht die Daten der amtlichen Statistik seit der deutschen Wiedervereinigung. Mithilfe aus den klassierten Daten ermittelter Quantile lassen sich sowohl Querschnitts- als auch Längsschnittanalysen tätigen, die aufzeigen, welche Gruppen von Steuerzahlern zu welcher Zeit die Einkommensteuerlast tragen. Zudem werden Analysen der Entwicklung der Durchschnittssteuersätze seit 1961 der einzelnen Quantile getätigt. Für diese Analyse wurden die Daten seit 1961 der amtlichen Statistik herangezogen.

Zur Erklärung der Entwicklung der Einkommensteuerlast und der Durchschnittssteuersatzbelastung werden wirtschaftliche Faktoren wie BIP, Arbeitslosenquote und Veränderungen der Einkommensteuertarife durch den Gesetzgeber herangezogen.

190 20 Jahre Code of Practice: Kritische Würdigung und Reformbedarf der europäischen Qualitätskriterien

Susanne Hagenkort-Rieger (1), Walter J. Radermacher (2), Kerstin Schneider (3)

1: Wirtschaftsministerium NRW, Deutschland; 2: LMU München, Deutschland; 3: Universität Wuppertal, Deutschland
wjrr@outlook.de

SECTION: Panel Discussion: 20 Jahre Code of Practice: Kritische Würdigung und Reformbedarf der europäischen Qualitätskriterien

Der im Jahr 2005 veröffentlichte Verhaltenskodex für europäische Statistiken (Code of Practice – CoP) definiert einheitliche Qualitätsstandards für alle statistischen Stellen im Europäischen Statistischen System. Ziel ist es, dass die amtlichen Statistiken in den europäischen Mitgliedstaaten nach gemeinsamen Qualitätsstandards und frei von äußerer Einflussnahme erstellt werden. Zudem garantiert der CoP die Vergleichbarkeit der Daten auf EU-Ebene und bildet somit eine wichtige Grundlage für politisches, wirtschaftliches und gesellschaftliches Handeln innerhalb der Europäischen Union. Der Verhaltenskodex hat somit eine hohe Bedeutung für die Sicherung des Vertrauens in die amtliche Statistik.

In der Paneldiskussion soll der Beitrag des CoP als wichtige Maßnahme der Qualitätssicherung kritisch gewürdigt werden. Dazu zählt auch die Frage, inwiefern der CoP als wichtiges Alleinstellungsmerkmal zur Sicherung der Relevanz der europäischen Statistiken beiträgt. Schließlich soll auch der Reformbedarf des CoP skizziert werden. Dabei sind sowohl die inhaltliche Ausgestaltung als auch die Akzeptanz zu berücksichtigen.

191 Melderegister vs. Zensus zur Ermittlung der Einwohnerzahl

Andreas Gleich

Stadt Augsburg, Deutschland
Andreas.Gleich@augzburg.de

SECTION: Regional Statistics (Session 2)

Die zentralen Erkenntnisse aus dem Zensus 2022 gleichen denen vom Zensus 2011. Als Hauptursache für die größten Probleme beim registergestützten Zensus kann die Haushaltebefragung auf Basis einer Stichprobe identifiziert werden.

Für das Zensusziel 1 = Ermittlung der Einwohnerzahl wird ein Vergleich der kommunalen Einwohnerzahlen aus dem Melderegister und der staatlichen Einwohnerzahl vor und nach dem Zensus gezogen und auf die unterschiedlichen Abweichungen der einzelnen Einwohnerzahlen eingegangen. Aus diesen Erkenntnissen, den Empfehlungen des Statistischen Beirats, des RatSWD und den Ergebnissen aus dem Registerzensuserprobungsgesetz wird zur Ermittlung einer gemeinsamen Einwohnerzahl von Kommunen, Ländern und Bund ein Melderegisterzensus vorgeschlagen, woraus sich viele Vorteile ergeben: - keine Erhebungsstellen in den Kommunen, keine kommunalen Erhebungsbeauftragten (zu aufwendig, zu teuer, nicht leistbar), - Bestand und Fortschreibung stammen aus demselben System (Melderegister), damit geringere Abweichungen zwischen Einwohnerzahlen aus dem Bestand und deren zwischenzeitlichen Fortschreibung mit Melderegisterbewegungen (v.a. Geburten, Sterbefälle, Um-, Zu-, Wegzüge, Abmeldung von Amtswegen), - kann jährlich erfolgen (was auch den ursprünglichen EU-Vorgaben entspricht), - Planungen auch über administrative (Kommunal-/Bezirks-/Landes-) Grenzen hinweg, basierend auf denselben aktuellen vergleichbaren Zahlen.

192 Der Zensus 2022 - eine besondere Datenquelle zum Thema Wohnen und Haushalte

Andreas van de Loo

IT.NRW, Deutschland
andreas.vandeLoo@it.nrw.de

SECTION: Regional Statistics (Session 2)

Der Zensus 2022 bietet auch eine besondere Datenbasis zur Wohn- und Haushaltssituation in Deutschland. Diese ermöglicht unterschiedliche Auswertungen und Analysen - von der Haushaltszusammensetzung über Eigentums- und Mietquoten bis hin zu Mietpreisen, Wohnflächen, Leerstandsquoten oder der Energieversorgung. So können beispielweise regionale Unterschiede und Entwicklungen im Bereich Wohnen und Haushalte betrachtet werden.

Der Vortrag gibt einen Überblick über die Bandbreite der verfügbaren Analysen sowie bisherigen Veröffentlichungen zu den Themen Wohnen und Haushalte auf Basis der Zensusergebnisse für Nordrhein-Westfalen und Impulse für weiterführende Fragestellungen im Themenfeld Wohnen und Haushalte.

193 Der Zensus 2022 – ein Datenschatz

Leonie Grauwinkel

IT.NRW, Deutschland
leonie.grauwinkel@it.nrw.de

SECTION: Regional Statistics (Session 2)

Der Zensus 2022 stellt die größte Erhebung in der amtlichen Statistik dar. Für die Themenbereiche Bevölkerung, Gebäude, Wohnungen, Haushalte und Familien liefert er eine breite Datengrundlage. Zwei wesentliche Ziele dieser Statistik sind die Ermittlung aktueller Bevölkerungszahlen sowie umfangreicher Informationen zum Wohnraum in Deutschland. Die besonderen Stärken der Zensusdaten zeigen sich in der Auswertbarkeit auf regional tiefer Ebene und der Verknüpfbarkeit unterschiedlicher Themenbereiche, was die Gewinnung von Erkenntnissen für komplexe, übergreifende Fragestellungen ermöglicht.

Im Mittelpunkt des Beitrags stehen dieses Analysepotential der Zensusdaten sowie die Informationsangebote für Datennutzende. Exemplarisch werden Veröffentlichungsprodukte aus Nordrhein-Westfalen und bundesweit verfügbare Datenbanken vorgestellt.

194 Asymptotic studies of adapted threshold detectors based on density processes

Duohong Sun, Ansgar Steland

Rheinisch-Westfälische Technische Hochschule Aachen, Deutschland
duohong.sun@rwth-aachen.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 3)

Control statistics, originally developed by Walter A. Shewhart, are widely used to monitor the quality of processes in various fields, including industry, healthcare, and machine learning. These statistics give an alarm when observed data exceed a threshold, traditionally set as a constant value to maintain a desired false alarm rate. However, one area of research in control statistics remains under-explored: When monitoring a sequence of observations, there may be additional information that potentially affects the law of the observations and should be considered by the design of the monitoring process. Precisely, we would like to change the design by using adapted thresholds, which are functions of the additional information.

So far, we have introduced several classes of adaptive threshold functions for continuous observations, including constant, proportional, and dominated classes. Our focus is on the proportional class, which adjusts sensitivity based on the external information, making it particularly effective in detecting rare but critical cases. We derive an estimator for this threshold function using kernel density estimation and establish its consistency. Further, we prove the asymptotic normality of the estimator using the functional delta method, which allows for deriving confidence intervals and error variances. Finally, we conduct simulations to evaluate the performance of the estimator. This work provides a foundation for future research on adaptive monitoring methods in various applications.

195 EWMA control charts for exponential data — numerical methods and transformations

Sven Knoth

Helmut-Schmidt-Universität Hamburg, Deutschland
knoth@hsu-hh.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 3)

The inherent simplicity of the exponential distribution allows explicit solutions of the average run length (ARL) integral equation for various control charts (here EWMA charts), cf. to some handy methods published in the last 30 years. On the other hand, the omnipresent Markov chain approximation method works feebly (an example will be given, for illustration). Nonetheless, it was used, for example, in prominent papers and elsewhere. Deploying the link between the exponential and the chi-square distribution with two degrees of freedom allows the usage of results (collocation applied to the ARL integral equation) for EWMA S^2 charts, where an R package is available. Finally, Monte Carlo studies can be utilized. Thus, one goal of the talk is comparing all these numerical algorithms and promoting efficient ones.

Having all these methods, we can judge whether transformations of the exponential proposed in dozen papers, in particular $X \hat{1}/3.6$ for achieving a bell-shape density function, which mimics the normal case, or other ones like $e^{\lambda X}$ and $\Phi^{-1}(e^{\lambda X})$ to get (exact) beta (uniform in the in-control case) and normal distribution, respectively, are appropriate convenience wrappers or simply diminish the detection performance. Eventually, some recommendations regarding the usage of these transformations will be given.

196 Detecting change-points of multivariate time series using the Wasserstein distance between the empirical marginal distributions

Anton Imm, Ansgar Steland

RWTH Aachen University, Deutschland
imm@stochastik.rwth-aachen.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 3)

In this talk we are interested in detecting change-points in multivariate nonstationary time series in a nonparametric setting. Firstly, we construct and discuss statistics which depend on the pairwise Wasserstein distance between the empirical marginal distributions of the vectors and from these, we construct CUSUM-statistics to detect change-points. The latter statistics depend on the partial sum process of the Wasserstein statistics and we show that this weakly converges to a Gaussian limit. A Bahadur representation result allows us to consider the asymptotic behavior of the empirical distribution function instead of the quantile function, which characterizes the one dimensional Wasserstein distance. A simulation study shows how well the significance level is retained under the null hypothesis of no change. Lastly, an outlook towards the power of the tests will be given.

197 EWMA control charts for the correlation coefficient

Maik Ulmer

RWTH Aachen/HSU Hamburg, Deutschland
ulmer@stochastik.rwth-aachen.de

SECTION: Statistics in Environmental and Engineering Sciences (Session 3)

In this talk we are monitoring the correlation between two normally distributed random variables through Exponentially Weighted Moving Average (EWMA) control charts. For this control chart, we calculate symmetric control limits around the initial value for a given Average Run Length (ARL), as well as ARL unbiased control limits. We determine these control limits not only through Monte Carlo simulation but also by numerically solving an integral equation for the ARL. In this context, we consider four different numerical methods: Markov chain approximation, Gauss-Legendre Nyström method, and plain and piecewise collocation methods. Afterwards, we examine the out-of-control behavior of the control chart with symmetric control limits compared to those with unbiased control limits for various underlying correlations. The results show that the closer the underlying correlation is to zero or the larger the sample size becomes, the more symmetric the resulting ARL curves are and the less we need the elaborated unbiased ARL control limits design. Lastly, an outlook will be given towards an application of the control chart to data from a bridge project.

198 Makroökonomische Einkommensverteilung: Hintergrund und Anforderungen aus dem SNA 2025 und der G20 Data Gap Initiative

Ricarda Senger, Stefan Hauf

Statistisches Bundesamt, Deutschland
ricarda.senger@destatis.de

SECTION: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 2)

Während das Bruttoinlandsprodukt (BIP) eine zentrale Kennzahl zur Messung des wirtschaftlichen Wachstums ist, zeigt es nicht, wer von diesem Wachstum profitiert und wie sich die wirtschaftliche Ungleichheit entwickelt. Zur Beurteilung des materiellen Wohlergehens und der wirtschaftlichen Nachhaltigkeit sind daher über das BIP hinausgehende makroökonomische Verteilungsanalysen erforderlich.

Als Reaktion auf die globale Finanzkrise ab 2007 und auf die Empfehlungen der Stiglitz-Sen-Fitoussi-Kommission wurde 2009 die G20 Data Gap Initiative ins Leben gerufen, um identifizierte politikrelevante Datenlücken zu schließen. Dazu gehören insbesondere die Verbesserung der Daten zur Einkommens-, Konsum- und Sparverteilung sowie zur Vermögensverteilung. Auch das neue System of National Accounts (SNA) 2025 unterstreicht die Bedeutung detaillierter Verteilungsrechnungen, an denen sich das künftige Europäische System Volkswirtschaftlicher Gesamtrechnungen (ESVG) orientieren wird.

Vor diesem Hintergrund entwickelt das Statistische Bundesamt derzeit eine makroökonomische Einkommensverteilung. Diese wird gemäß den methodischen Empfehlungen des SNA nach für den Sektor private Haushalte (S.14) berechnet. Die Volkswirtschaftlichen Gesamtrechnungen (VGR) dienen als zentrale Datengrundlage und werden durch Mikrodaten ergänzt, die mit den

makroökonomischen Aggregaten harmonisiert werden. Ziel dieser Anpassung ist eine konsistente und international vergleichbare Erfassung aller Einkommensbestandteile über die Zeit.

Die Verteilungsergebnisse sollen nach Dezilen sowie, sofern möglich, nach weiteren relevanten soziodemografischen Merkmalen ausgewertet werden. Um Einkommensunterschiede zwischen Haushalten mit unterschiedlicher Größe und Zusammensetzung vergleichbar zu machen, wird eine Äquivalenzskala verwendet. Das SNA sieht schließlich vor, eine gemeinsame Verteilung von Einkommen, Konsum und Vermögen zu erstellen, um daraus multivariate Kennzahlen wie das Vermögen-Einkommen-Verhältnis oder die Konsumquote abzuleiten.

199 Makroökonomische Einkommensverteilung: Überblick des methodischen Vorgehens

Ricarda Senger, Felix Lederer, Frederik Wiyneck, Aaron Hennig

Statistisches Bundesamt, Deutschland
ricarda.senger@destatis.de

SECTION: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 2)

Die makroökonomische Einkommensverteilung soll gemäß methodischen Empfehlungen von Eurostat und der OECD im Rahmen der Data Gap Initiative (DGI) 3 sowie des neuen System of National Accounts (SNA) 2025 berechnet werden. Grundlage dafür sind die Einkommensaggregate der Volkswirtschaftlichen Gesamtrechnungen (VGR), insbesondere das Einkommen aus selbstständiger und nichtselbstständiger Arbeit, Vermögenseinkommen, Sozialleistungen, Steuer- und Sozialabgaben sowie sonstige Transferzahlungen, aus denen sich das verfügbare Einkommen ableitet. Mit dem Aufbau der makroökonomischen Einkommensverteilungsrechnung als neuer Bestandteil der VGR beschäftigt sich seit 2024 ein abteilungsübergreifendes Projektteam des Statistischen Bundesamtes. Zur Ermittlung der Verteilungsinformationen über die einzelnen Einkommenskomponenten werden Mikrodaten, insbesondere aus EU-SILC und der Lohn- und Einkommensteuerstatistik, mit den makroökonomischen Aggregaten harmonisiert. Eine methodische Herausforderung besteht darin, dass die Berechnungen für private Haushalte erfolgen sollen, da nur EU-SILC Daten auf Haushaltsebene liefert, während die Lohn- und Einkommensteuerstatistik Steuerpflichtige abbildet. Zudem bestehen quantitative Abweichungen zwischen den Makro- und Mikrodaten, die auf Unterschiede in fachlichen Konzepten sowie auf Stichproben- und Nichtstichprobenfehler (bei EU-SILC) oder eine kleinere Grundgesamtheit als in der VGR (bei der Lohn- und Einkommensteuerstatistik) zurückzuführen sind. Darüber hinaus fehlen in den Mikrodaten bestimmte makroökonomische Komponenten vollständig, beispielsweise schattenwirtschaftliche Aktivitäten oder sozialer Sachtransfer. Zusätzlich erschweren Zeitreihenbrüche in den Mikrodaten die Analyse. Um diese Datenlücken zu schließen, sind statistische Anpassungsmethoden erforderlich.

Die Verteilungsergebnisse sollen nicht nur nach Einkommensgruppen (z. B. Dezile), sondern auch nach weiteren soziodemografischen Merkmalen differenziert werden, die für wirtschaftliche Analysen und politische Maßnahmen von Interesse sind. Allerdings ergeben sich hierbei Unschärfen, da nicht alle relevanten Merkmale in den Mikrodatenquellen gleichermaßen verfügbar sind und fehlende Merkmale imputiert werden müssen. Vorgesehen ist, dass erste experimentelle Ergebnisse aus dem Projekt Ende 2025 vorliegen.

200 Microfoundations in measurement and theory

Junyi Zhu

Deutsche Bundesbank, Germany
iamthr@gmail.com

SECTION: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 2)

This paper clarifies that, to address the aggregation problem, macroeconomic models must return to the empirical aggregates of national accounts (NA), thereby reconciling theoretical constructs with observed macro outcomes. However, some fundamental micro–macro gaps emerge between the bottom-up, microfounded macro and macro measurement. The primary gap stems from the divergence between closed aggregation—rooted in micro reality and theory—and the open aggregation approach employed by NA, which implicitly imposes macro restrictions that lack full microfoundations (Schlicht, 1985). In practice, two barriers—information friction and policy-oriented objectives—prevent macro measurement from adopting a closed aggregation framework. Using the aggregate production function and the contrast between Real GDP and Physical Value Added as illustrative examples, we demonstrate these barriers and rationalize NA's reliance on open aggregation. A microfounded explanation is achievable only when the corresponding micro restrictions are satisfied, ensuring consistency between bottom-up aggregation and macro measurement; otherwise, NA aggregates diverge from those derived from micro data. Recognizing these gaps, we argue that many current efforts to redistribute the aggregate data exhibit a paradoxical regression to representative agent thinking. Tentative solutions include the development of satellite accounts and/or the imposition of top-down micro–macro restrictions—drawing on index number theory and practice—on the bottom-up aggregation process.

201 The Joint Distribution of Income, Wealth and Consumption in Germany

Julia Le Blanc (1), Tobias Schmidt (2), Anna Stelzer (3)

1: European Commission - Joint Research Centre, Italien; 2: Deutsche Bundesbank, Deutschland; 3: Oesterreichische Nationalbank
Tobias.Schmidt@bundesbank.de

SECTION: Micro- and Macroeconomic Perspectives on the Distribution of Income, Wealth, and Consumption (Session 2)

How wealth, income and consumption are jointly distributed is important to assess the levels and trends in inequality in households' resources.

The goal of this paper is to empirically estimate the joint distribution of income, consumption and wealth. We use micro-data from three waves of the Panel on Household Finances (PHF). While wealth and income are directly collected in this wealth survey, total consumption is calculated using information on active savings flows and net income, and non-durable consumption is imputed. Our main findings can be summarized as follows: Inequality is a pervasive phenomenon in Germany although it affects consumption, wealth and income differently. As in other countries, the income and consumption distributions are more equal, as measured by the Gini coefficient, than the wealth distribution. A multivariate extension of the Gini coefficient suggests that bivariate inequality in

net wealth and income or consumption, respectively, is lower than the more pronounced inequality in net wealth in isolation, which suggests that measuring inequality in net wealth and the more equally distributed income or consumption together results in a more equal joint distribution. In contrast, the multivariate Gini in income and consumption is higher than both univariate Gini coefficients. Analysing inequality in terms of a trivariate Gini in turn suggests again that inequality measured in all three variables lies somewhere in between the more equal distributions of income and consumption and the more unequal distribution of net wealth. This pattern seems quite stable over all three waves of our data.

These results improve our understanding of inequality in the largest Euro area economy and furthermore add new dimensions to measuring inequality which has been impeded by data limitations in the past.

Social Programme:

For activities of the social programme, prior registration via ConfTool is necessary. Please note the respective registration deadline and the maximum number of participants. For late registration please send an e-mail to statistische-woche@dstatg.de.

- **Monday, Sep 01, 2025:**

City tour – 6 pm

Duration: 1 hour

Meeting point: Kurhausplatz 1, 65189 Wiesbaden

Participation fee per person: €9.50

Language: German

Relaxed Get-Together – Karim’s Brasserie – 6:30 pm

Meeting point: Webergasse 7, 65183 Wiesbaden

Participation fee per person: 5€ - in return you will receive a voucher for 5€, which you can (exclusively) redeem at the venue on the same evening. Please note that these costs cannot be refunded in case of not showing up or not using the voucher.

- **Tuesday, Sep 02, 2025:**

Reception by the City of Wiesbaden at the Town Hall accompanied by the Destatis Choir – 7 pm

Duration: 1.5 hours Meeting point: Schlossplatz 6, 65183 Wiesbaden

Language: German

Relaxed Get-Together – Rathauskeller Andechser – 8:30 pm

Meeting point: Schlossplatz 6, 65183 Wiesbaden

- **Wednesday, Sep 03, 2025:**

Conference Dinner – Burg Crass – 6 pm

Meeting point: Freygäßchen 1, 65343 Eltville am Rhein

Participation fee: free of charge for participants of the Statistical Week 2025.

- **Thursday, Sep 04, 2025:**

Relaxed Get-Together – Das Wohnzimmer – 6:30 pm

Meeting point: Schwalbacher Straße 51, 65183 Wiesbaden

Participation fee per person: 5€ - in return you will receive a voucher for 5€, which you can (exclusively) redeem at the venue on the same evening. Please note that these costs cannot be refunded in case of not showing up or not using the voucher.

Wiesbaden Casino – 9 pm

Meeting point: Kurhausplatz 1, 65189 Wiesbaden

- **Friday, Sep 05, 2025:**

City Tour – 1:45 pm

Duration: 2 hours

Meeting point: Marktplatz 1, 65183 Wiesbaden

Participation fee per person: €14

Language: German

Rahmenprogramm:

Für Aktivitäten des Rahmenprogramms ist eine vorherige Anmeldung über das ConfTool notwendig. Bitte beachten Sie den jeweiligen Anmeldeschluss bzw. die maximale Teilnehmerzahl. Zur nachträglichen Anmeldung senden Sie bitte eine E-Mail an statistische-woche@dstatg.de.

● Montag, 01.09.2025:**Stadtführung – 18:00 Uhr**

Dauer: 1 Stunde

Treffpunkt: Kurhausplatz 1, 65189 Wiesbaden

Teilnahmegebühr pro Person: 9.50€

Sprache: Deutsch

Gemütliches Beisammensein – Karim's Brasserie – 18:30 Uhr

Treffpunkt: Webergasse 7, 65183 Wiesbaden

Teilnahmegebühr pro Person: 5€ – im Gegenzug erhalten Sie einen Verzehrutschein über 5€, den Sie (ausschließlich) am gleichen Abend in der Gaststätte einlösen können. Bitte beachten Sie, dass diese Kosten bei Nichterscheinen oder Nichtnutzung nicht erstattet werden können.

● Dienstag, 02.09.2025:**Empfang der Stadt Wiesbaden im Rathaus begleitet vom Destatis-Chor – 19:00 Uhr**

Dauer: 1.5 Stunden

Treffpunkt: Schlossplatz 6, 65183 Wiesbaden

Sprache: Deutsch

Gemütliches Beisammensein – Rathauskeller Andechser – 20:30 Uhr

Treffpunkt: Schlossplatz 6, 65183 Wiesbaden

● Mittwoch, 03.09.2025:**Conference Dinner – Burg Crass – 18:00 Uhr**

Treffpunkt: Freygäßchen 1, 65343 Eltville am Rhein

Teilnahmegebühr: für Teilnehmende der Statistischen Woche 2025 kostenfrei.

● Donnerstag, 04.09.2025:**Gemütliches Beisammensein – Das Wohnzimmer – 18:30 Uhr**

Treffpunkt: Schwalbacher Straße 51, 65183 Wiesbaden

Teilnahmegebühr pro Person: 5€ – im Gegenzug erhalten Sie einen Verzehrutschein über 5€, den Sie (ausschließlich) am gleichen Abend in der Gaststätte einlösen können. Bitte beachten Sie, dass diese Kosten bei Nichterscheinen oder Nichtnutzung nicht erstattet werden können.

Spielbank Wiesbaden – 21:00 Uhr

Treffpunkt: Kurhausplatz 1, 65189 Wiesbaden

● Freitag, 05.09.2025:**Stadtführung – 13:45 Uhr**

Dauer: 2 Stunden

Treffpunkt: Marktplatz 1, 65183 Wiesbaden

Teilnahmegebühr pro Person: 14€

Sprache: Deutsch

Participants / Teilnehmende

Abesadze, Tamar; Constructor University; tabesadze23@gmail.com
Ade, Lorenz; Statistisches Bundesamt; lorenz.ade@destatis.de
Alau, Lena; Statistisches Bundesamt; lena.alau@destatis.de
Alfken, Christoph; IT.NRW – Statistisches Landesamt Nordrhein-Westfalen; christoph.alfken@it.nrw.de
Altmos, Klaus; Destatis; klaus.altmos@destatis.de
Andrä, Diana; Stadt Dortmund, Dortmunder Statistik; dandrae@stadtdo.de
Appelhanz, Shanna; Statistisches Bundesamt; Shanna.Appelhanz@destatis.de
Arsova, Antonia; TU Dortmund University; arsova@statistik.tu-dortmund.de
Articus, Charlotte; Statistisches Bundesamt; Charlotte.Articus@destatis.de
Bach, Maximilian; Statistisches Bundesamt; maximilian.bach@destatis.de
Bach, Paul; Institut für Qualitätssicherung und Transparenz im Gesundheitswesen; paul.bach@iqtig.org
Balczynski, Oliver; IT.NRW; oliver.balczynski@it.nrw.de
Barendse, Sander; University of Amsterdam; s.c.barendse@uva.nl
Bauer, Lukas; University of Freiburg; lukas.bauer@vwl.uni-freiburg.de
Baumstark, Johannes; Bayerisches Landesamt für Statistik; johannes.baumstark@statistik.bayern.de
Bayer, Florian; Bitkom e.V.; f.bayer@bitkom.org
Beck, Martin; ehemals Statistisches Bundesamt; martin@mgm-beck.de
Becker, Angelika; Verband der Chemischen Industrie e.V.; becker@vci.de
Becker, Sebastian; Bundesministerium für Arbeit und Soziales; Sebastian.Becker@bmas.bund.de
Behmann, Mathias; Stadtmagistrat Innsbruck; mathias.behmann@innsbruck.gv.at
Benner, Niklas; RWI – Leibniz-Institut für Wirtschaftsforschung, Essen; nbenner@rwi-essen.de
Berg, Gabriele; Statistisches Bundesamt; gabriele.berg@destatis.de
Berger, Ursula; Ludwig-Maximilians-Universität München; ursula.berger@lmu.de
Bickel, Michael; Statistisches Bundesamt; Michael.Bickel@destatis.de
Birkigt, Holger; Statistisches Bundesamt; holger.birkigt@destatis.de
Bleninger, Sara; Bayerisches Landesamt für Statistik; sara.bleninger@statistik.bayern.de
Böhm, Ulrike; Statistisches Bundesamt; ulrike.boehm@destatis.de
Bohnensteffen, Sarah; Statistisches Bundesamt; sarah.bohnensteffen@destatis.de
Böhringer-Becker, Tanja; Statistisches Bundesamt; tanja.boehringer-becker@destatis.de
Borg, Sandra; Statistisches Bundesamt; sandra.borg@destatis.de
Borgs, Christian; IT.NRW; christian.borgs@it.nrw.de
Boudt, Kris; Universiteit Gent; kris.boudt@ugent.be
Braakmann, Albert; Statistisches Bundesamt (emeritus); alwies15@gmail.com
Brand, Ruth; Statistisches Bundesamt; ruth.brand@destatis.de
Brennecke, Cederic; Statistisches Bundesamt; cederic.brennecke@destatis.de
Brenzel, Hanna; Statistisches Bundesamt; Hanna.Brenzel@destatis.de
Browell, Jethro; University of Glasgow; jethro.browell@glasgow.ac.uk
Brownlees, Christian; Universität Pompeu Fabra; christian.brownlees@upf.edu
Brückweh, Kerstin; Leibniz-Institut für Raumbezogene Sozialforschung; paul.treffenfeldt@leibniz-irs.de
Brunßen, Frank; Statistisches Bundesamt; frank.brunssen@destatis.de
Buchholz, Christine; Hochschule Bonn-Rhein-Sieg; christine.buchholz@h-brs.de
Buchner, Michael; Statistisches Bundesamt; michael.buchner@destatis.de
Buhl, Yannik; Statistisches Bundesamt; yannik.buhl@destatis.de
Burgard, Jan Pablo; Universität Trier; burgardj@uni-trier.de
Burmistr, Svetlana; Leibniz Institut für Raumbezogene Sozialforschung; svetlana.burmistr@leibniz-irs.de
Buscher, Lioba; LH Dresden; lbuscher@dresden.de

Carius-Munz, Lisa; Statistisches Bundesamt; lisa.carius-munz@destatis.de
Chakraverty, Veronika; Statistisches Bundesamt; veronika.chakraverty@destatis.de
Chen, Li; Universität Paderborn; lichen@campus.uni-paderborn.de
Chlumsky, Jürgen; Statistisches Bundesamt (ehem.); jchlumsky@t-online.de
Chrysopoulou-Tseva, Kleio; Humboldt-Universität zu Berlin; kleio.chrysopoulou.tseva@hu-berlin.de
Chukhrova, Nataliya; Universität Heidelberg; nataliya.chukhrova@medma.uni-heidelberg.de
Coutu, Melanie; Statistisches Bundesamt; melanie.coutu@destatis.de
Cramer, Katharina; IT.NRW; katharina.cramer@it.nrw.de
Dagdoug, Mehdi; Mc Gill University; mehdi.dagdoug@mcgill.ca
Dechent, Jens; Statistisches Bundesamt; jens.dechent@destatis.de
Deetjen, Gloria; Statistisches Bundesamt (Destatis); Gloria.Deetjen@destatis.de
della Peruta, Simona; Statistisches Bundesamt; simona.dellaPeruta@destatis.de
Demetrescu, Matei; TU Dortmund; mdeme@statistik.tu-dortmund.de
Deschermeier, Philipp; Institut der deutschen Wirtschaft; deschermeier@iwkoeln.de
Dittenberger, Gerhard; Statistisches Bundesamt; Gerhard.Dttenberger@destatis.de
Dittrich, Stefan; Statistisches Bundesamt; stefan.dittrich@destatis.de
Döhler, Sebastian; Hochschule Darmstadt; sebastian.doehler@h-da.de
Dorka, Katrin; Statistisches Bundesamt; katrin.dorka@destatis.de
Dr. Roder, Alexandra; Statistisches Bundesamt; Alexandra.Roder@destatis.de
Dressel, Felix; Statistisches Landesamt Baden-Württemberg; Felix.Dressel@stala.bwl.de
Dröge, Lennart; Universität Augsburg; lennart.droege@uni-a.de
Droßard, Ralf; Statistisches Bundesamt; ralf.drossard@destatis.de
Düker, Martha; Deutsche Bundesbank; martha.dueker@bundesbank.de
Dumpert, Florian; Statistisches Bundesamt; florian.dumpert@destatis.de
Dyckerhoff, Rainer; Universität zu Köln; rainer.dyckerhoff@statistik.uni-koeln.de
Dzikowski, Daniel; TU Dortmund University; daniel.dzikowski@tu-dortmund.de
Eckendorff, Roman; Statistikservice der Bundesagentur für Arbeit; eckendorff@gmx.de
Efosa, Norah; Statistik Aargau, Kanton Aargau; norah.efosa@ag.ch
Empting, Lennart; Uni Duisburg-Essen; lennart.empting@vwl.uni-due.de
Emrich, Katharina; Krebsregister Rheinland-Pfalz am IDG; emrich@idg-rlp.de
Erdemsiz, Baran; Statistisches Bundesamt; baran.erdemsiz@destatis.de
Esslinger, André; Helmut-Schmidt-Universität; esslinger@hsu-hh.de
Forster, Michael; ehemals IT.NRW; family_forster@t-online.de
Forster, Pascal; Bayerisches Landesamt für Statistik; pascal.forster@statistik.bayern.de
Frahm, Gabriel; Helmut-Schmidt-Universität; frahm@hsu-hh.de
Franzen, Maren; IT.NRW, Statistisches Landesamt Nordrhein-Westfalen; maren.franzen@it.nrw.de
Fried, Roland; TU Dortmund; roland.fried@tu-dortmund.de
Fries, Jan; Amt für Statistik und Stadtforschung Wiesbaden; dr.jan.fries@wiesbaden.de
Frommeyer, Dominik; Statistisches Bundesamt; dominik.frommeyer@destatis.de
Fuchs, Benjamin; Kraftfahrt-Bundesamt; benjamin.fuchs@kba.de
Fürnrohr, Michael; Uni Bamberg; michael.fuernrohr@statistik.bayern.de
Garcia Ritz, Yannik; Statistisches Bundesamt; Yannik.GarciaRitz@destatis.de
Giar, Katharina; Statistisches Bundesamt; Katharina.Giar@destatis.de
Gleich, Andreas; Stadt Augsburg; Andreas.Gleich@augzburg.de
Gliszczynska, Karolina; Universität Duisburg-Essen; karolina.gliszczynska@uni-due.de
Goemans, Pascal; FernUniversität in Hagen; pascal.goemans@fernuni-hagen.de
Golosnoy, Vasyli; Ruhr-Universität Bochum; vasyli.golosnoy@rub.de
Gondrom, Sabine; go2data; contact@go2data.de
Gores, Thorsten; Statistisches Bundesamt; thorsten.gores@destatis.de
Görner, Sören; Statistisches Landesamt NRW; soeren.goerner@it.nrw.de

Gramüller-Heinz, Frank; SAS Institute GmbH; frank.gramueller@sas.com
Grauwinkel, Leonie; IT.NRW; leonie.grauwinkel@it.nrw.de
Grehling, Tom; Statistisches Bundesamt; tom.grehling@destatis.de
Gril, Lorena; Freie Universität Berlin; lorena.gril@fu-berlin.de
Große, Franziska; Landesamt für Statistik Niedersachsen; franziska.grosse@statistik.niedersachsen.de
Gruber, Brigitte; Statistisches Bundesamt; brigitte.gruber@destatis.de
Guglielmelli, Mattia; SAS Institute GmbH; mattia.guglielmelli@sas.com
Guskov, Dmitry; Amt für Statistik und Stadtforschung, Landeshauptstadt Wiesbaden; dmitry.guskov@wiesbaden.de
Habla, Heike; Statistisches Bundesamt; heike.habla@destatis.de
Haid, Philipp; Universität Augsburg; philipp1.haid@uni-a.de
Halbleib, Roxana; University of Freiburg; roxana.halbleib@vwl.uni-freiburg.de
Hamzaoui, Patrice; Statistisches Bundesamt; patrice.hamzaoui@destatis.de
Hanck, Christoph; University of Duisburg-Essen; christoph.hanck@vwl.uni-due.de
Hannemann, Tobias; Amt für Statistik Berlin-Brandenburg; tobias.hannemann@statistik-bbb.de
Hartmann, Michael; Statistik der Bundesagentur für Arbeit; Michael.Hartmann@arbeitsagentur.de
Hassler, Uwe; Goethe-Universität Frankfurt; hassler@wiwi.uni-frankfurt.de
Hauf, Stefan; Statistisches Bundesamt; stefan.hauf@destatis.de
Haußmann, Michael; Stadt Freiburg i.Br.; michael.haussmann@freiburg.de
Heiden, Sebastian; DStatG; heiden@dstatg.de
Heiden, Anja; Universitätsstadt Siegen - GB 3 / SeFo; a.heiden@siegen.de
Heinrichs, Florian; FH Aachen; f.heinrichs@fh-aachen.de
Hellwagner, Timon; Institute for Employment Research (IAB); timon.hellwagner@iab.de
Hembach-Stunden, Katharina; GWS mbH; hembach-stunden@gws-os.com
Hennig, Aaron; Statistisches Bundesamt; Aaron.Hennig@destatis.de
Herget, Anna; Bayerisches Landesamt für Statistik; Anna.Herget@statistik.bayern.de
Herkner, Thomas; Privat; thomasherker@outlook.com
Herter, Michael; dai data analytics institute AG; michael.herter@dai.institute
Heuser, Annelie; Amt für Statistik und Stadtforschung Wiesbaden; annelie.heuser@wiesbaden.de
Hoette, Kerstin; KEDGE Business School; kerstin.hoette@posteo.de
Hoffmann, Stefanie; Statistisches Bundesamt; stefanie.hoffmann@destatis.de
Hoffmann, Christina; Statistisches Bundesamt; Christina.Hoffmann@destatis.de
Hofmann, Bernd; Statistik der Bundesagentur für Arbeit; bernd.hofmann5@arbeitsagentur.de
Hoga, Yannick; Universität Duisburg-Essen; yannick.hoga@vwl.uni-due.de
Hoger, Tom; Geschäftsstelle Länderarbeitskreis Energiebilanzen, Statistisches Landesamt Baden-Württemberg; tom.hoger@stala.bwl.de
Hohlstein, Franziska; Statistisches Bundesamt; franziska.hohlstein@destatis.de
Holzendorf, Volker; Stadtverwaltung Jena; volker.holzendorf@jena.de
Hönig, Tillman; IW Consult; hoenig@iwkoeln.de
Höppner, Björn; Universität Bonn; b.hoepfner@uni-bonn.de
Huang, Chenhui; Duisburg-Essen Uni; charlotte0203666@gmail.com
Hughes, Finn Ernie; Physikalisch-Technische Bundesanstalt; finn.hughes@ptb.de
Imm, Anton; RWTH Aachen University; imm@stochastik.rwth-aachen.de
Irrgang, Stefan; Statistisches Bundesamt; Stefan.Irrgang@destatis.de
Jackisch, Annett; Statistikamt Nord; Annett.Jackisch@statistik-nord.de
Jaeger, Christina; Statistisches Bundesamt; christina.jaeger@destatis.de
Jäger, Solveigh; Bundesverband der Deutschen Industrie (BDI); s.jaeger@bdi.eu
Jahn, Malte; HSU Hamburg; jahnma@hsu-hh.de
Jentsch, Carsten; TU Dortmund; jentsch@statistik.tu-dortmund.de
Jeschke, Anke; Stadt Frankfurt am Main; anke.jeschke@stadt-frankfurt.de
Jeske, Roland; Hochschule Kempten; roland.jeske@hs-kempten.de

Joebges, Silke; Landeshauptstadt München; silke.joebges@muenchen.de
Johannssen, Arne; Universität Hamburg; arne.johannssen@uni-hamburg.de
Jordan, Christopher; Statistisches Bundesamt; Christopher.Jordan@destatis.de
Jung, Johanna; Statistisches Bundesamt (Destatis); Johanna.Jung@destatis.de
Kadic, Alma; Stadt Neumünster; alma.kadic@neumuenster.de
Kalinowski, Michael; Bundesinstitut für Berufsbildung; kalinowski@bibb.de
Kasten, Berit; Manage Now GmbH; berit.kasten@manage-now.de
Kastner, Martin; Kraftfahrt-Bundesamt; martin.kastner@kba.de
Kaszubowska, Joanna; Stadt Freiburg - Amt für Bürgerservice und Informationsmanagement; joanna.kaszubowska@freiburg.de
Kauermann, Göran; LMU München; goeran.kauermann@lmu.de
Kaufmann, Violetta; Amt für Statistik Berlin-Brandenburg; Violetta.Kaufmann@statistik-bbb.de
Keller, Manuel; Stadt Dortmund; mkeller@stadtdo.de
Kirchner, Robert; Deutsche Bundesbank; robert.kirchner@bundesbank.de
Kladroba, Andreas; Stifterverband für die Deutsche Wissenschaft; andreas.kladroba@stifterverband.de
Klaus, Anton; Statistik der Bundesagentur für Arbeit; anton.klaus@arbeitsagentur.de
Klenke, Jens; Universität Duisburg-Essen; jens.klenke@vwl.uni-due.de
Klinke, Sigbert; Humboldt-Universität zu Berlin; sigbert@wiwi.hu-berlin.de
Knapp, Romy; Statistisches Bundesamt; romy.knapp@destatis.de
Knoth, Sven; Helmut-Schmidt-Universität Hamburg; knoth@hsu-hh.de
Koch, Cornelia; Statistisches Bundesamt; cornelia.koch@destatis.de
Köhler, Andreas; IT.NRW, Statistisches Landesamt; andreas.koehler@it.nrw.de
Konietzko, Thorsten; Statistisches Bundesamt; thorsten.konietzko@destatis.de
Kösler, Julia; Statistisches Landesamt Baden-Württemberg; julia.koesler@stala.bwl.de
Koslik, Jan-Ole; Universität Bielefeld; jan-ole.koslik@uni-bielefeld.de
Kovac, Daniela; Statistisches Bundesamt; daniela.kovac@destatis.de
Kramer, Walter; TU Dortmund; walterk@statistik.tu-dortmund.de
Kraus, Petra; Hauptverband der Deutschen Bauindustrie e.V.; petra.kraus@bauindustrie.de
Krell, Lara; Universität Trier; lara.krell@uni-trier.de
Krentz, Ariane; Statistisches Landesamt Baden-Württemberg; Ariane.Krentz@stala.bwl.de
Krinitz, Jonas; Gesellschaft Wirtschaftliche Strukturforchung; krinitz@gws-os.com
Kuhn, Sarah; IAB (Bundesagentur für Arbeit); sarah.kuhn@iab.de
Kühn, Lisa; Wissenschaftszentrum Berlin (WZB); lisa.kuehn@ratswd.de
Kühn, Rachel; Statistisches Bundesamt; rachel.kuehn@destatis.de
Lange, Tatjana; HS Merseburg; tanja.28.lange@gmail.com
Lederer, Felix; Statistisches Bundesamt; felix.lederer@destatis.de
Leerhoff, Holger; Amt für Statistik Berlin-Brandenburg; holger.leerhoff@statistik-bbb.de
Lentz, Tobias; Statistisches Bundesamt; tobias.lentz@destatis.de
Lenz, Tanja; Bundesverband Baustoffe - Steine und Erden e.V.; t.lenz@bvbaustoffe.de
Lenz, Jürgen; Kreis Mettmann; juergen.lenz@kreis-mettmann.de
Limberg, Heiko; Statistisches Bundesamt; heiko.limberg@destatis.de
Linz, Stefan; Statistisches Bundesamt; stefan.linz@destatis.de
Lopez Menendez, Philip; Statistisches Bundesamt; philip.lopezmenendez@destatis.de
Ludsteck, Johannes; Institut für Arbeitsmarkt- und Berufsforschung (IAB) Nürnberg; Johannes.Ludsteck@iab.de
Ludwig, Matthias; European Commission - Eurostat; matthias.ludwig@ec.europa.eu
Lugert, Patricia; Bundesamt für Verbraucherschutz und Lebensmittelsicherheit; patricia.lugert@bvl.bund.de
Lüüs, Hans-Peter; Statistisches Bundesamt; hans-peter.lueues@destatis.de
Maisack, Mike; destatis; mike.maisack@web.de
Maletz, Steffen; Technische Universität Dortmund; steffen.maletz@tu-dortmund.de

Mannigel, Alice; Statistikamt Nord; alice.mannigel@statistik-nord.de
Maretzke, Steffen; RS 1 Raumentwicklung; steffen.maretzke@bbr.bund.de
Maßmann, Lennard; Universität Duisburg-Essen; lennard.massmann@uni-due.de
Massmann, Michael; WHU Otto Beisheim School of Management; michael.massmann@whu.edu
Meer, Uwe; Stadt Wolfsburg; uwe.meer@stadt.wolfsburg.de
Mendel, Merle; TU Dortmund; merle.mendel@tu-dortmund.de
Miller, Bernhard; GESIS - Leibniz-Institut für Sozialwissenschaften; bernhard.miller@gesis.org
Mödinger, Patrizia; Statistisches Bundesamt; patrizia.moedinger@destatis.de
Möhler, Martin; Destatis; martin.moehler@destatis.de
Moritz, Steffen; Statistisches Bundesamt; steffen.moritz@destatis.de
Mosler, Karl; Universität zu Köln; kmosler@uni-koeln.de
Moys, Gunnar; Techniker Krankenkasse; dr.gunnar.moys@tk.de
Mrani Alaoui, Meryam; Statistisches Bundesamt; meryam_0207@icloud.com
Mueller, Christian; Constructor University; cmueller01@constructor.university
Mühlenhaupt, Jörg; IT.NRW, Statistisches Landesamt NRW; joerg.muehlenhaupt@it.nrw.de
Mühlhan, Jannek; Destatis; jannek.muehlhan@destatis.de
Mühlichen, Michael; Bundesinstitut für Bevölkerungsforschung (BiB); michael.muehllichen@bib.bund.de
Müller, Cornelia; Bundesinstitut für Bau-, Stadt- und Raumforschung; cornelia.mueller@bbr.bund.de
Müller-Funk, Ulrich; Universität Münster; funk@wiwi.uni-muenster.de
Mundzeck, Corinna; Statistisches Amt für Hamburg und Schleswig-Holstein; corinna.mundzeck@statistik-nord.de
Münnich, Ralf; Universität Trier; muennich@uni-trier.de
Muth, Mareike; Statistisches Bundesamt; mareike.muth@destatis.de
Neumann, Alexander; Hochschule Wirtschaft und Recht, Berlin; neumannalexander993@gmail.com
Neurauter, Stefanie; Landeshauptstadt Wiesbaden; stefanie.neurauter@wiesbaden.de
Niehues, Judith; Institut der deutschen Wirtschaft; niehues@iwkoeln.de
Nordhausen, Klaus; University of Helsinki; klaus.nordhausen@helsinki.fi
Nowossadeck, Enno; Robert Koch-Institut; nowossadecke@rki.de
Nowossadeck, Sonja; DZA Berlin; sonja.nowossadeck@dza.de
Nüsing, Linus; Universität Konstanz; linus.nuesing@uni-konstanz.de
Ochsmann, Andrea; Stadt Recklinghausen; andrea.ochsmann@recklinghausen.de
Offermann, Gerit; Destatis; gerit.offermann@destatis.de
Okhrin, Yarema; University of Augsburg; yarema.okhrin@wiwi.uni-augsburg.de
Okhrin, Ostap; TU Dresden; ostap.okhrin@tu-dresden.de
Olbrich, Lukas; IAB Nürnberg; lukas.olbrich@iab.de
Oltmanns, Erich; Statistisches Bundesamt; erich.oltmanns@destatis.de
Oruc, Bayram; Statistisches Bundesamt; Bayram.Oruc@destatis.de
Österling, Meike; Magistrat Stadt Wetzlar; meike.oesterling@wetzlar.de
Otto, Philipp; University of Glasgow; philipp.otto@glasgow.ac.uk
Otto, Sven; University of Cologne; sven.otto@uni-koeln.de
Overmann, Gina Lena; IT.NRW; ginalena.overmann@it.nrw.de
Özer, Yeliz; Leibniz Universität Hannover; oezer@statistik.uni-hannover.de
Öztürk, Saskia; Georg-August Universität Göttingen; saskia.oeztuerk@uni-goettingen.de
Pappert, Sven; TU Dortmund; pappert@statistik.tu-dortmund.de
Plöger, Amelie; IT.NRW; amelie.ploeger@it.nrw.de
Pohlan, Laura; Institut für Arbeitsmarkt- und Berufsforschung (IAB); Laura.Pohlan@iab.de
Pohle, Marc-Oliver; Karlsruhe Institute of Technology; pohle@kit.de
Popp, Martin; Institut für Arbeitsmarkt- und Berufsforschung; martin.popp@iab.de
Pötzsch, Klaus; Statistisches Bundesamt; klaus.poetzsch@destatis.de
Preisling, Marcel; Statistisches Bundesamt; marcel.preisling@destatis.de

Proietti, Tommaso; University of Rome Tor Vergata; tommaso.proietti@uniroma2.it
Purrucker, Rosanna; Bayerisches Landesamt für Statistik; Rosanna.Purrucker@statistik.bayern.de
Qiu, Peihua; University of Florida; pqiu@ufl.edu
Quaini, Alberto; Erasmus University; quaini@ese.eur.nl
Quandt, Dennis; Geschäftsstelle Länderarbeitskreis Energiebilanzen, Statistisches Landesamt Baden-Württemberg; dennis.quandt@stala.bwl.de
Radermacher, Walter J.; LMU München; wjr@outlook.de
Rakuša, Elena; Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE); elena.rakusa@dzne.de
Ramic, Arijana Amina; Statistisches Bundesamt; arijanaamina.ramic@destatis.de
Rees, Peter; Statistisches Bundesamt; Peter.Rees@destatis.de
Reichold, Karsten; TU Wien; karsten.reichold@tuwien.ac.at
Reimann-Peters, Malte; Statistik der Bundesagentur für Arbeit; malte.reimann-peters@arbeitsagentur.de
Reincke, Ulrich; SAS Institute GmbH; Ulrich.Reincke@sas.com
Ren, Rui; University of Augsburg; rui.ren.research@gmail.com
Rendtel, Ulrich; Freie Universität Berlin; urlich.rendtel@fu-berlin.de
Rengers, Martina; Statistisches Bundesamt; martina.rengers@destatis.de
Rennspies, Jasper; Universität Freiburg; jasper.rennsbies@vwl.uni-freiburg.de
Resin, Johannes; Goethe-Universität Frankfurt; resin@econ.uni-frankfurt.de
Richter, Ann-Kathrin; Statistisches Landesamt NRW (IT.NRW); ann-kathrin.richter@it.nrw.de
Rimpler, René; Zentralverband des Deutschen Handwerks; rimpler@zdh.de
Rink, Anke; Statistisches Bundesamt; anke.rink@destatis.de
Rohde, Johannes; IT.NRW - Statistisches Landesamt; johannes.rohde@it.nrw.de
Röhrig, Marc; Statistisches Bundesamt; marc.roehrig@destatis.de
Römer, Ellen; Statistisches Bundesamt; ellen.roemer@destatis.de
Rommelspacher, Simon; Statistisches Bundesamt; simon.rommelspacher@destatis.de
Rosenbaum, Claudia; Springer Gabler; claudia.rosenbaum@springer.com
Rosteck, Veronika; Springer-Verlag GmbH; veronika.rosteck@springer.com
Rotsche, Maximilian; Destatis; maximilian.rotsche@destatis.de
Rudloff, Karin; Bundesagentur für Arbeit; karin.rudloff@arbeitsagentur.de
Ruhmann, Iris; Springer-Verlag GmbH; iris.ruhmann@springer.com
Ruhnau, Magdalene; SAS Institute Deutschland GmbH; magdalene.ruhnau@sas.com
Saliba, Jacques; Swiss Federal Statistical Office; jacques.saliba@bfs.admin.ch
Salwiczek, Christian; Statistisches Amt für Hamburg und Schleswig-Holstein; christian.salwiczek@statistik-nord.de
Santoro, Chiara; IT.NRW; chiara.santoro@it.nrw.de
Schäfer, Josef; Technical University Dortmund; jugschaefer@aol.com
Schaller, Jannik; Deutsche Bundesbank; jannik.schaller@bundesbank.de
Scherer, Jan; Universität Bonn; s3jnsche@uni-bonn.de
Schiele, Stefan; Bayerisches Landesamt für Statistik; Stefan.Schiele@statistik.bayern.de
Schiesberg, Sara; Statistisches Bundesamt; sara.schiesberg@destatis.de
Schimpke, Michael; Capgemini Deutschland GmbH; michael.schimpke@capgemini.com
Schirbaum, Anna Maria; Statistisches Landesamt - it.nrw; anna.schirbaum@it.nrw.de
Schlieker, Pinar; Stadt Dortmund; pschlieker@stadtdo.de
Schmid, Wolfgang; Europa-Universität; milesle@icloud.com
Schmid, Timo; Otto-Friedrich-Universität Bamberg; timo.schmid@uni-bamberg.de
Schmidt, Tobias; Deutsche Bundesbank; Tobias.Schmidt@bundesbank.de
Schmidt, Fabian; TU Dortmund; f.schmidt@statistik.tu-dortmund.de
Schmidt, Peter; Statistisches Bundesamt; peter.schmidt@destatis.de
Schmidtke, Kerstin; Statistisches Landesamt NRW (IT.NRW); kerstin.schmidtke@it.nrw.de

Schmiedt, Anja Bettina; Ostbayerische Technische Hochschule Regensburg; anja.schmiedt@oth-regensburg.de
Schmitz, Martin; European Central Bank; martin.schmitz@ecb.europa.eu
Schneider, Gaby; Goethe University Frankfurt; schneider@math.uni-frankfurt.de
Schneider, Daniel Christoph; Max Planck Institute for Demographic Research; schneider@demogr.mpg.de
Schneider, Nicholas-Sebastian; Statistisches Bundesamt; nicholas-sebastian.schneider@destatis.de
Schnurbus, Joachim; Universität Passau; joachim.schnurbus@uni-passau.de
Scholtus, Sander; Statistics Netherlands; s.scholtus@cbs.nl
Schrader, Klaus; GdW Bundesverband deutscher Wohnungs- und Immobilienunternehmen; schra-der@gdw.de
Schroer, Jan Philipp; Bundesinstitut für Berufsbildung; jan-philipp.schroer@bibb.de
Schüller, Daniela; Stadtverwaltung Koblenz; Daniela.Schueller@stadt.koblenz.de
Schüller, Katharina; STAT-UP Statistical Consulting & Data Science GmbH; katharina.schueller@stat-up.com
Schultz, Andrea; Stadt Leipzig; andrea.schultz@leipzig.de
Schulz, Dominik; Universität Paderborn; dominik.schulz@uni-paderborn.de
Schulz, Julian; Landesamt für Statistik Niedersachsen; julian.schulz@statistik.niedersachsen.de
Schulz, Christian; Universität Duisburg-Essen; christian.schulz@vwl.uni-due.de
Schumann, Carsten; Statistisches Bundesamt; carsten.schumann@destatis.de
Schur, Alexander Christian; Bundesinstitut für Berufsbildung; alexander.schur@bibb.de
Schürmann, Daniel; Stadt Dortmund; dschuermann@stadtdo.de
Schüssler, Rainer Alexander; University of Münster; raineralexanderschuessler@gmail.com
Schütze, Florian; Helmut-Schmidt-Universität; schuetzf@hsu-hh.de
Schwabbacher, Wolfram; Deutsche Gesetzliche Unfallversicherung e.V.; wolfram.schwabbacher@dguv.de
Schweitzer, Michael; TU Dortmund; michael.schweitzer@tu-dortmund.de
Schwengler, Barbara; Institut für Arbeitsmarkt- und Berufsforschung; Barbara.Schwengler@iab.de
Schwenzl, Nina; Bayerisches Landesamt für Statistik; nina.schwenzl@statistik.bayern.de
Schwickerath, Marco; Statistisches Bundesamt; marco.schwickerath@destatis.de
Seesing-Coelho, Clara; Statistisches Bundesamt; clara.seesing-coelho@destatis.de
Seidel, Gerald; Statistik der Bundesagentur für Arbeit; Gerald.Seidel@arbeitsagentur.de
Seifert, Ingo; Stadt Leipzig; ingo.seifert@leipzig.de
Senger, Ricarda; Statistisches Bundesamt; ricarda.senger@destatis.de
Sibbertsen, Philipp; Leibniz Universität Hannover; sibbertsen@statistik.uni-hannover.de
Siems, Inken; Universität Trier; inkensiems@gmx.de
Simoni, Anna; CREST - CNRS; anna.simoni@ensae.fr
Söllner, René; Statistisches Bundesamt; rene.soellner@destatis.de
Speich, Wolf-Dietmar; Statistisches Landesamt des Freistaates Sachsen; wolf-dietmar.speich@statistik.sachsen.de
Steger, Almut; Deutsche Bundesbank; almut.steger@t-online.de
Stephan, Andreas; Linnaeus University, Department of Forestry and Wood Technology; andre-as.stephan@lnu.se
Stoye, Jörg; Cornell University; stoye@cornell.edu
Suárez, Marlon; Statistisches Bundesamt; marlon.suarez@destatis.de
suffert, martin; Manage Now GmbH; martin.suffert@manage-now.de
Sun, Duohong; Rheinisch-Westfälische Technische Hochschule Aachen; duohong.sun@rwth-aachen.de
Tesching, Karin; Bayerisches Landesamt für Statistik; karin.tesching@statistik.bayern.de
Thees, Nicole; Stadtverwaltung Trier; nicole.thees@trier.de
Thien-Seitz, Uta; LH München; uta.thien@muenchen.de
Thobe, Ines; GWS mbH; thobe@gws-os.com
Thurow, Maria; TU Dortmund University; maria.thurow@tu-dortmund.de

Trapin, Luca; University of Bologna; luca.trapin@unibo.it
Treffenfeldt, Paul; Leibniz-Institut für Raumbezogene Sozialforschung; p.treff@outlook.de
Tuekam, Victor; LMU Munich; tuekam@lmu.de
Uhrich, Stefanie; Statistisches Bundesamt; stefanie.uhrich@destatis.de
Ulmer, Maik; RWTH Aachen/ HSU Hamburg; ulmer@stochastik.rwth-aachen.de
Uysal, Derya; LMU Munich; derya.uysal@econ.lmu.de
Vallé, Jennifer; Statistisches Bundesamt; jennifer.valle@destatis.de
Vallizadeh, Ehsan; Statistik der Bundesagentur für Arbeit; ehsan.vallizadeh@arbeitsagentur.de
van de Loo, Andreas; IT.NRW; andreas.vandeLoo@it.nrw.de
Vanella, Patrizio; aQua-Institut; patrizio.vanella@aqua-institut.de
Verhoeven, Sascha; Stadt Heidelberg; sascha.verhoeven@heidelberg.de
Volk, Johannes; Statistisches Bundesamt; johannes.volk@destatis.de
Völker, Julia; Statistisches Bundesamt, Destatis; julivoelker@web.de
von Kalckreuth, Ulf; Deutsche Bundesbank; ulf.von-kalckreuth@bundesbank.de
von Malottki, Christian; BPD Immobilienentwicklung; c.malottki@bpd-de.de
von Reumont, Frederik; Universität zu Köln; f.von-reumont@uni-koeln.de
Vorgriemler, Daniel; Statistisches Bundesamt; Daniel.vorgriemler@destatis.de
Voshage, Ramona; Hochschule für Wirtschaft und Recht Berlin; ramona.voshage@hwr-berlin.de
Wagner, Stephan; Statistikamt Nord; stephan.wagner@statistik-nord.de
Waschipky, Martin; Stadt Leipzig, Amt für Statistik und Wahlen; martin.waschipky@leipzig.de
Waßenberg, Sebastian; Information und Technik Nordrhein Westfalen - Statistisches Landesamt; sebastian.wassenberg@it.nrw.de
Weba, Michael; Goethe-Universität Frankfurt; weba@wiwi.uni-frankfurt.de
Weigert, Alexander; Statistisches Bundesamt; alexander.weigert@destatis.de
Weiß, Christian; Helmut Schmidt University; weissc@hsu-hh.de
Weiß, Sebastian; Bayerisches Landesamt für Statistik; Sebastian.Weiss@statistik.bayern.de
Weißbrodt, Thomas; Bundesagentur für Arbeit; thomas.weissbrodt@arbeitsagentur.de
Wermuth, Jan-Lukas; Goethe Universität Frankfurt; wermuth@econ.uni-frankfurt.de
Wickenhöfer, Lena; Agentur für Arbeit; Lena.Wickenhoefer@arbeitsagentur.de
Wied, Dominik; Universität zu Köln; dwied@uni-koeln.de
Wiegleb, Georg; Landeshauptstadt Magdeburg; georg.wiegleb@stat.magdeburg.de
Windmann, Michael; LMU München; michael.windmann@lmu.de
Winkler, Dorothee; Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR); dorothee.winkler@bbr.bund.de
Wynck, Frederik; Statistisches Bundesamt; frederik.wynck@destatis.de
Wójcik, Ewa-Juditha; Statistisches Bundesamt; ewa-juditha.wojcik@destatis.de
Wolfanger, Tobias; IT.NRW; tobias.wolfanger@it.nrw.de
Wondrak, Stefan; Statistisches Bundesamt; stefan.wondrak@destatis.de
Würbach, Ariane; Leibniz-Institut für Bildungsverläufe (LIfBi); ariane.wuerbach@lifbi.de
Yadegar, Edesa; Statistisches Bundesamt; edesa.yadegarshirabad@destatis.de
Yanzuk, Christopher; Statistisches Bundesamt; Christopher.yanzuk@destatis.de
Yu, Miao; Leibniz University Hannover; miao.yu@statistik.uni-hannover.de
Zahn, Tanja; Goethe University Frankfurt; tzahn@wiwi.uni-frankfurt.de
Zeigermann, Oliver; Techniker Krankenkasse; oliver.zeigermann@tk.de
Zemann, Christian; Statistik der Bundesagentur für Arbeit; Christian.Zemann@arbeitsagentur.de
Zenk, Johanna; Institut für Arbeitsmarkt- und Berufsforschung (IAB); johanna.zenk@iab.de
Zhu, Junyi; Deutsche Bundesbank; iamthr@gmail.com
Ziebarth, Gerhard; Aktionskreis Stabiles Geld; gerhard.ziebarth@gmail.com
Zika, Gerd; Institut für Arbeitsmarkt- und Berufsforschung (IAB); gerd.zika@iab.de
Zimmermann, Markus; Statistisches Bundesamt; markus.zimmermann2@destatis.de
Zins, Bernhard Stefan; Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur

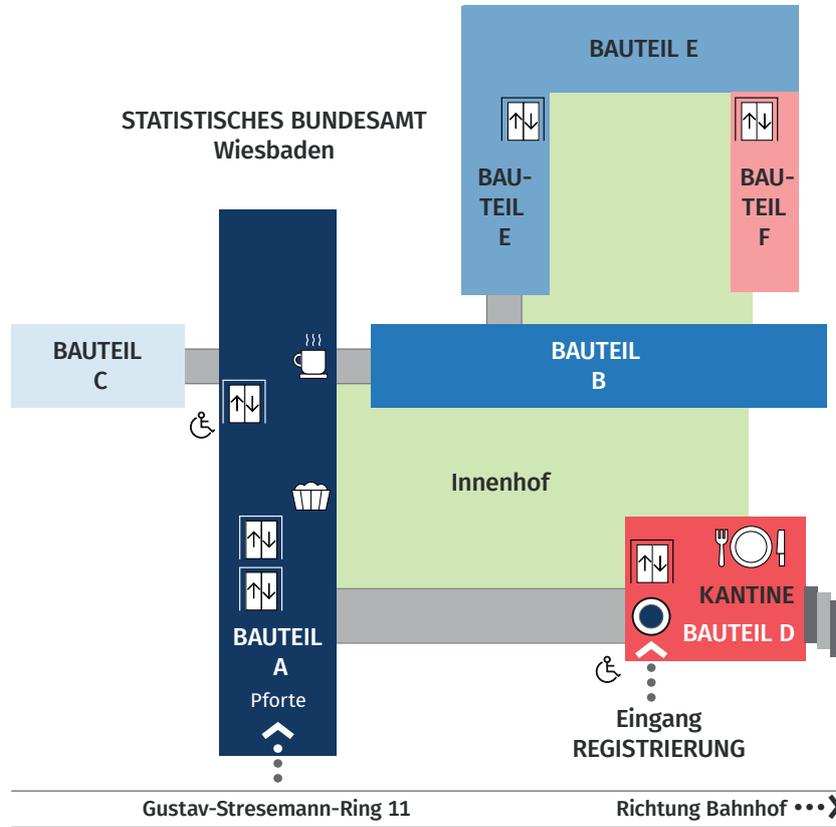
für Arbeit (BA); stefan.zins@iab.de

Zou, Jing; Technische Universität Dresden; jing.zou@tu-dresden.de

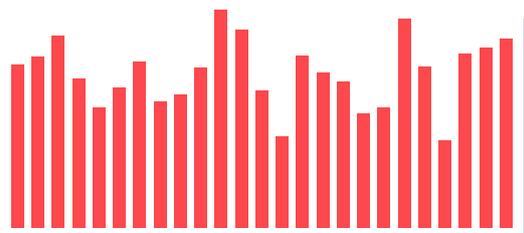
Zukovska, Jekaterina; Humboldt-Universität zu Berlin; zukovskj@hu-berlin.de

Zwick, Markus; Statistisches Bundesamt; markus.zwick@destatis.de

Building Plans / Gebäudepläne



REGISTRIERUNG	— D – Foyer EG	Konferenzräume	— A. 13.207
Tagungsbüro	— D – Foyer EG — B. 00.102		— A. 03.205
Mittagessen	— D – Kantine 1. OG		— A. 03.212
Kaffeepause	— A – Foyer EG		— A. 00. Bibliothek
Poster Session	— D – Foyer EG		— E. 03.112
Informationsstände	— A – Foyer EG		— F. 04.208
		Tutorium	— B. 03.104
		Vorstandssitzung	— A. 12.209

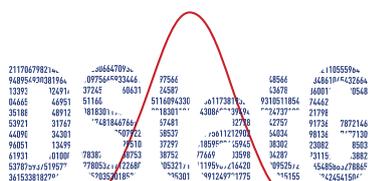


Statistische Woche
02. – 05. September

Wiesbaden
2025



Alle Informationen zur Statistischen Woche 2025
für Android, IOS und als Web App unter
statistische-woche.de/app



Deutsche Statistische Gesellschaft

D  **STATIS**
Statistisches Bundesamt