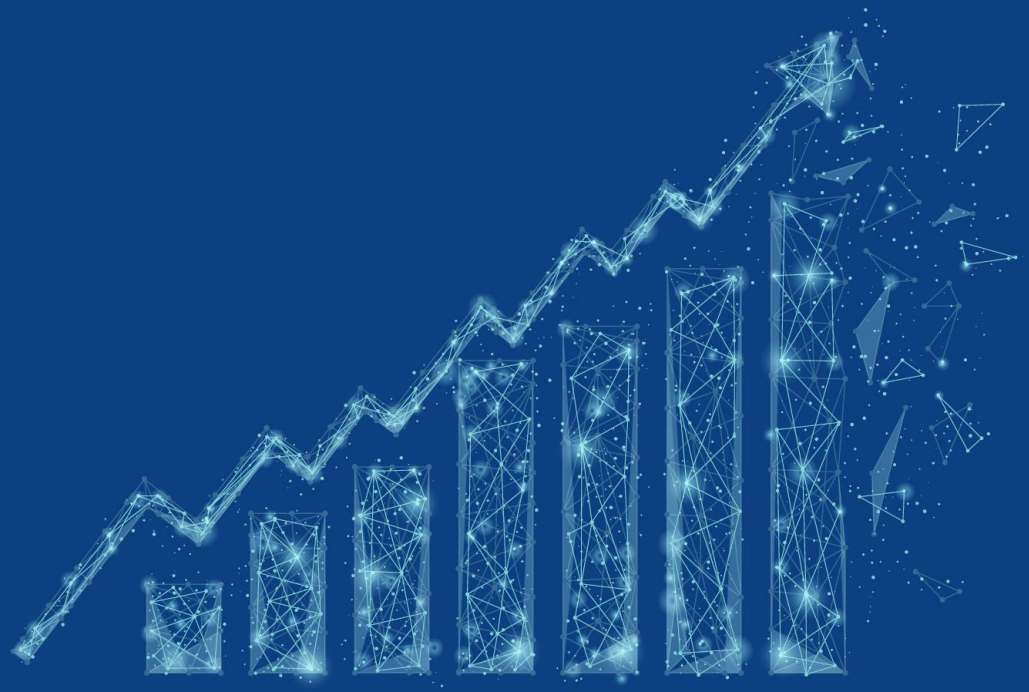
GERMAN DATA SCIENCE DAYS 2024

Economic Data Science



Dr. Sebastian Wichert,
Head of Research Data Center (EBDC)
ifo Institute, Munich

BIG DATA ifo Economics

CESifo cluster
Economic Research with Firm Data

ifo INSTITUTE

LMU-ifo EBDC Economics & Business Data Center

KonsortSWD
New Options for Access to Firm Data

ifo is a leading economic research institute in Germany

Our mission: Shaping the economic debate

Key themes

Economic growth and sustainability

Digital transformation of the economy

Public sector function and taxation

European integration, globalization and inter-system competition

Equality of opportunity and integration

Key facts

Leading German economic research institute on research performance, media coverage and policy impact

Non-profit organization + connection to LMU Munich

Headed by Prof. Dr. Clemens Fuest

9 research areas + research data center

approx. 220 employees

ifo is a leading economic research institute in Germany

Our mission:

tagesschau

Sendung verpasst?

Key themes

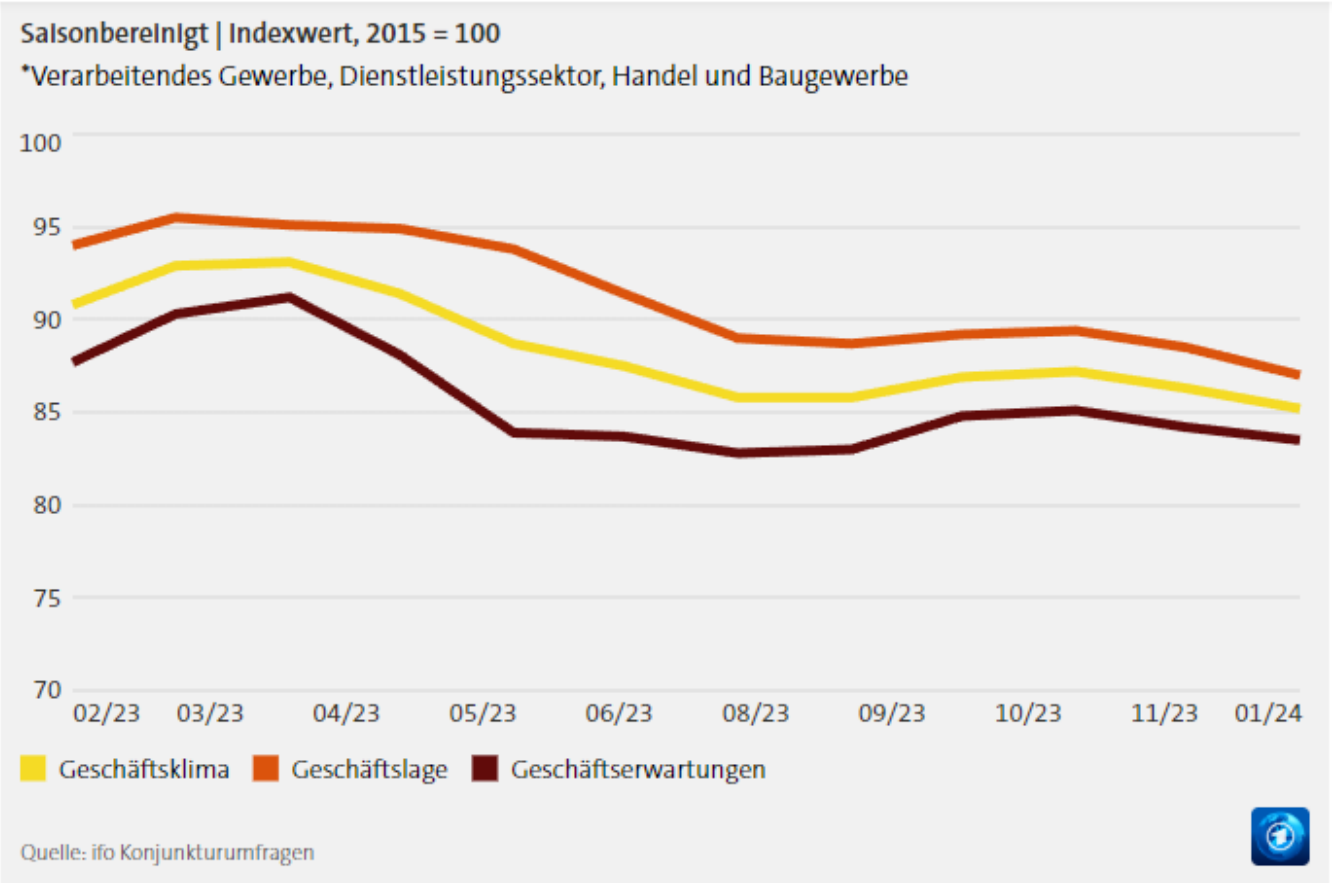
Ec

Di

Pu

Eu
int

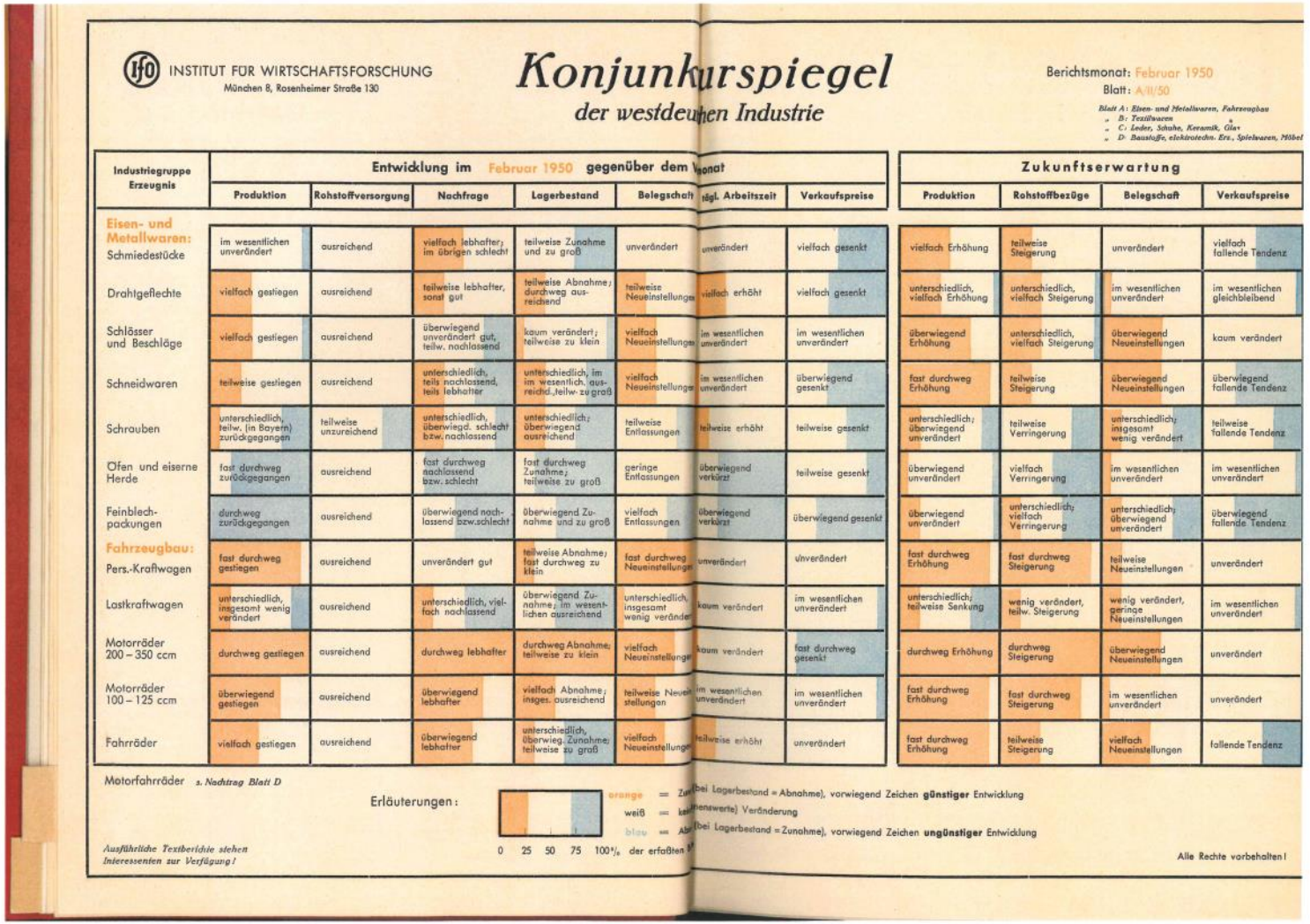
Eq



itute on
and policy

LMU Munich

Economics and (Big) Data Science – a long history until today



INSTITUT FÜR WIRTSCHAFTSFORSCHUNG
München 8, Rosenheimer Straße 130

Konjunkturspiegel
der westdeutschen Industrie

Berichtsmonat: Februar 1950
Blatt: A/II/50

Blatt A: Eisen- und Metallwaren, Fahrzeugbau
„ B: Textilwaren
„ C: Leder, Schuhe, Keramik, Glas
„ D: Baustoffe, elektrotechn. Erz., Spielwaren, Möbel

Industriegruppe Erzeugnis	Entwicklung im Februar 1950 gegenüber dem Vormonat: Produktion	Rohstoffversorgung	Nachfrage	Lagerbestand	Belegschaft	tägl. Arbeitszeit	Verkaufspreise	Zukunftserwartung: Produktion	Rohstoffbezüge	Belegschaft	Verkaufspreise
Eisen- und Metallwaren:											
Schmiedestücke	im wesentlichen unverändert	ausreichend	vielfach lebhafter; im übrigen schlecht	teilweise Zunahme und zu groß	unverändert	unverändert	vielfach gesenkt	vielfach Erhöhung	teilweise Steigerung	unverändert	vielfach fallende Tendenz
Drahtgeflechte	vielfach gestiegen	ausreichend	teilweise lebhafter, sonst gut	teilweise Abnahme; durchweg ausreichend	teilweise Neueinstellungen	vielfach erhöht	vielfach gesenkt	unterschiedlich, vielfach Erhöhung	unterschiedlich, vielfach Steigerung	im wesentlichen unverändert	im wesentlichen gleichbleibend
Schlösser und Beschläge	vielfach gestiegen	ausreichend	überwiegend unverändert gut, teilw. nachlassend	kaum verändert; teilweise zu klein	vielfach Neueinstellungen	im wesentlichen unverändert	im wesentlichen unverändert	überwiegend Erhöhung	unterschiedlich, vielfach Steigerung	überwiegend Neueinstellungen	kaum verändert
Schneidwaren	teilweise gestiegen	ausreichend	unterschiedlich, teils nachlassend, teils lebhafter	unterschiedlich, im wesentlich. ausreichd., teilw. zu groß	vielfach Neueinstellungen	im wesentlichen unverändert	überwiegend gesenkt	fast durchweg Erhöhung	teilweise Steigerung	überwiegend Neueinstellungen	überwiegend fallende Tendenz
Schrauben	unterschiedlich, teilw. (in Bayern) zurückgegangen	teilweise unzureichend	unterschiedlich, überwiegd. schlecht bzw. nachlassend	unterschiedlich; überwiegend ausreichend	teilweise Entlassungen	teilweise erhöht	teilweise gesenkt	unterschiedlich; überwiegend unverändert	teilweise Verringerung	unterschiedlich; insgesamt wenig verändert	teilweise fallende Tendenz
Ofen und eiserne Herde	fast durchweg zurückgegangen	ausreichend	fast durchweg nachlassend bzw. schlecht	fast durchweg Zunahme; teilweise zu groß	geringe Entlassungen	überwiegend verkürzt	teilweise gesenkt	überwiegend unverändert	vielfach Verringerung	im wesentlichen unverändert	im wesentlichen unverändert
Feinblech-packungen	durchweg zurückgegangen	ausreichend	überwiegend nachlassend bzw. schlecht	überwiegend Zunahme und zu groß	vielfach Entlassungen	überwiegend verkürzt	überwiegend gesenkt	überwiegend unverändert	unterschiedlich; vielfach Verringerung	unterschiedlich; überwiegend unverändert	überwiegend fallende Tendenz
Fahrzeugbau:											
Pers.-Kraftwagen	fast durchweg gestiegen	ausreichend	unverändert gut	teilweise Abnahme; fast durchweg zu klein	fast durchweg Neueinstellungen	unverändert	unverändert	fast durchweg Erhöhung	fast durchweg Steigerung	teilweise Neueinstellungen	unverändert
Lastkraftwagen	unterschiedlich, insgesamt wenig verändert	ausreichend	unterschiedlich, vielfach nachlassend	überwiegend Zunahme; im wesentlichen ausreichend	unterschiedlich, insgesamt wenig verändert	kaum verändert	im wesentlichen unverändert	unterschiedlich; teilweise Senkung	wenig verändert, teilw. Steigerung	wenig verändert, geringe Neueinstellungen	im wesentlichen unverändert
Motorräder 200 – 350 ccm	durchweg gestiegen	ausreichend	durchweg lebhafter	durchweg Abnahme; teilweise zu klein	vielfach Neueinstellungen	kaum verändert	fast durchweg gesenkt	durchweg Erhöhung	durchweg Steigerung	überwiegend Neueinstellungen	unverändert
Motorräder 100 – 125 ccm	überwiegend gestiegen	ausreichend	überwiegend lebhafter	vielfach Abnahme; insges. ausreichend	teilweise Neueinstellungen	im wesentlichen unverändert	im wesentlichen unverändert	fast durchweg Erhöhung	fast durchweg Steigerung	im wesentlichen unverändert	unverändert
Fahrräder	vielfach gestiegen	ausreichend	überwiegend lebhafter	unterschiedlich, überwieg. Zunahme; teilweise zu groß	vielfach Neueinstellungen	teilweise erhöht	unverändert	fast durchweg Erhöhung	teilweise Steigerung	vielfach Neueinstellungen	fallende Tendenz

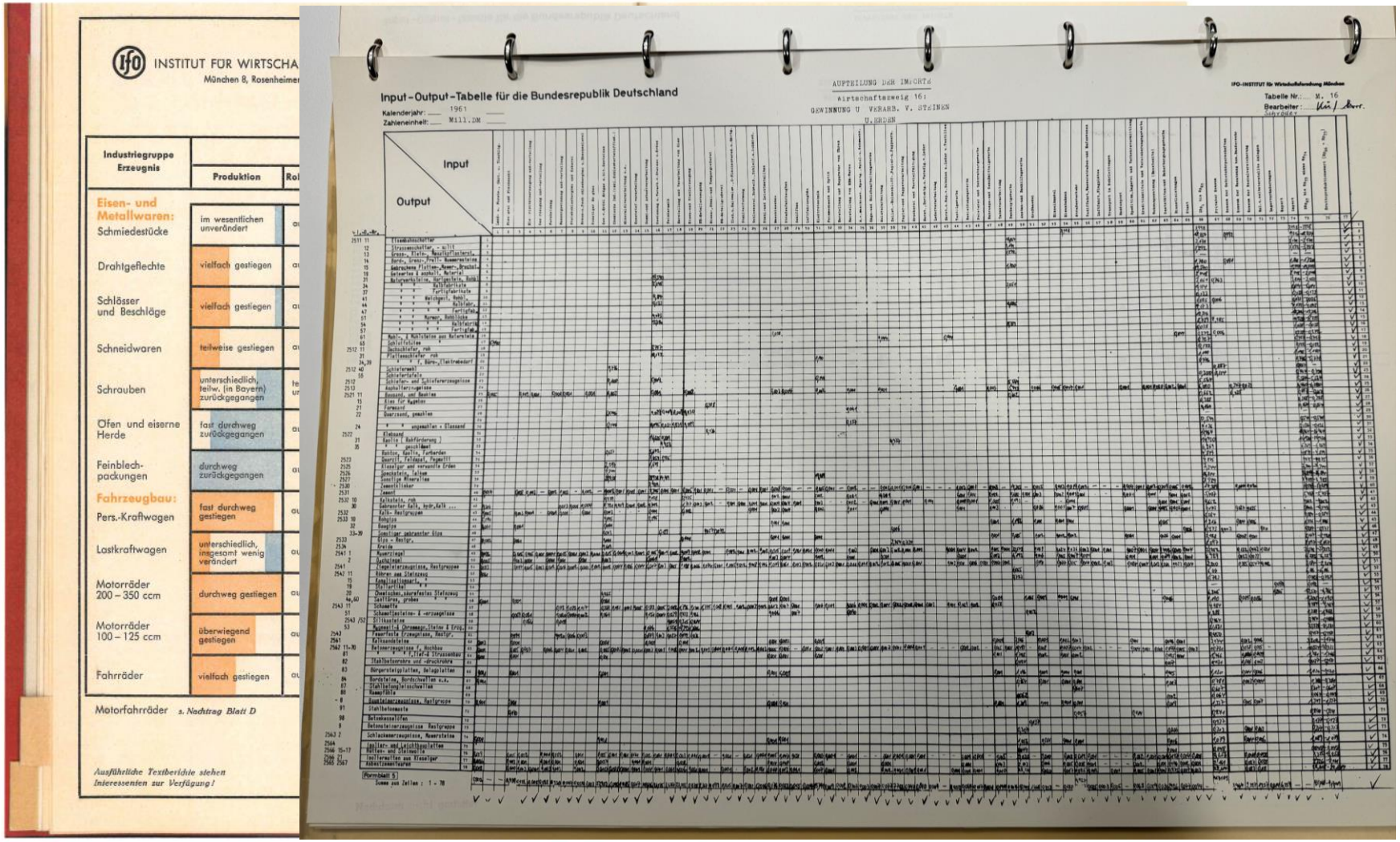
Motorfahrräder *s. Nachtrag Blatt D*

Erläuterungen: orange = Zu… (bei Lagerbestand = Abnahme), vorwiegend Zeichen **günstiger** Entwicklung
weiß = kei… (meßwerte) Veränderung
blau = Ab… (bei Lagerbestand = Zunahme), vorwiegend Zeichen **ungünstiger** Entwicklung
0 25 50 75 100% der erfaßten M…

Ausführliche Textberichte stehen Interessenten zur Verfügung!

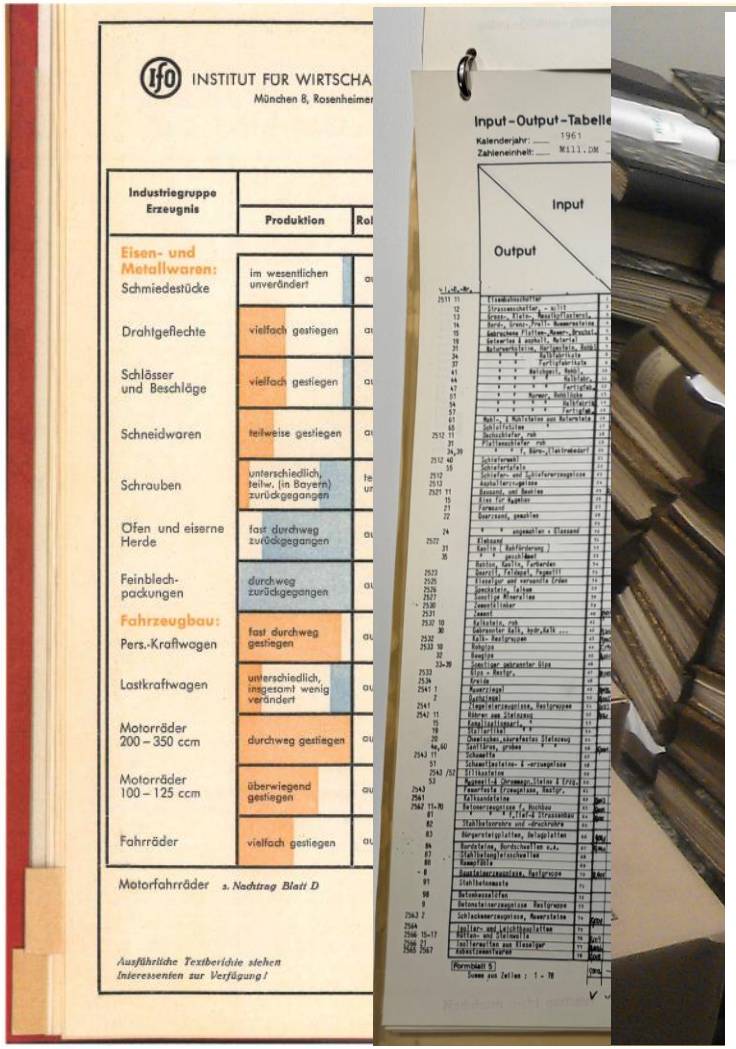
Alle Rechte vorbehalten!

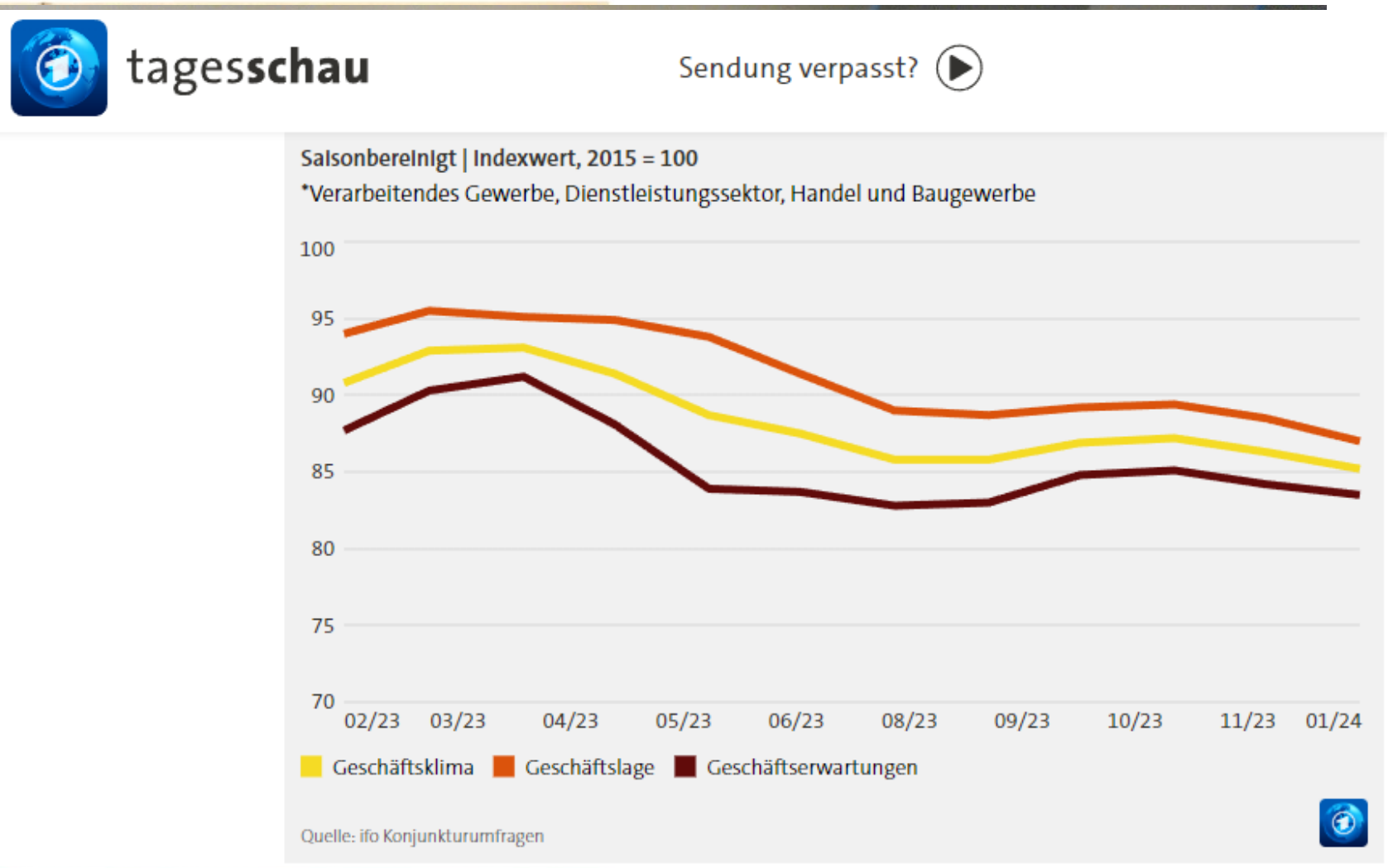
Economics and (Big) Data Science – a long history until today



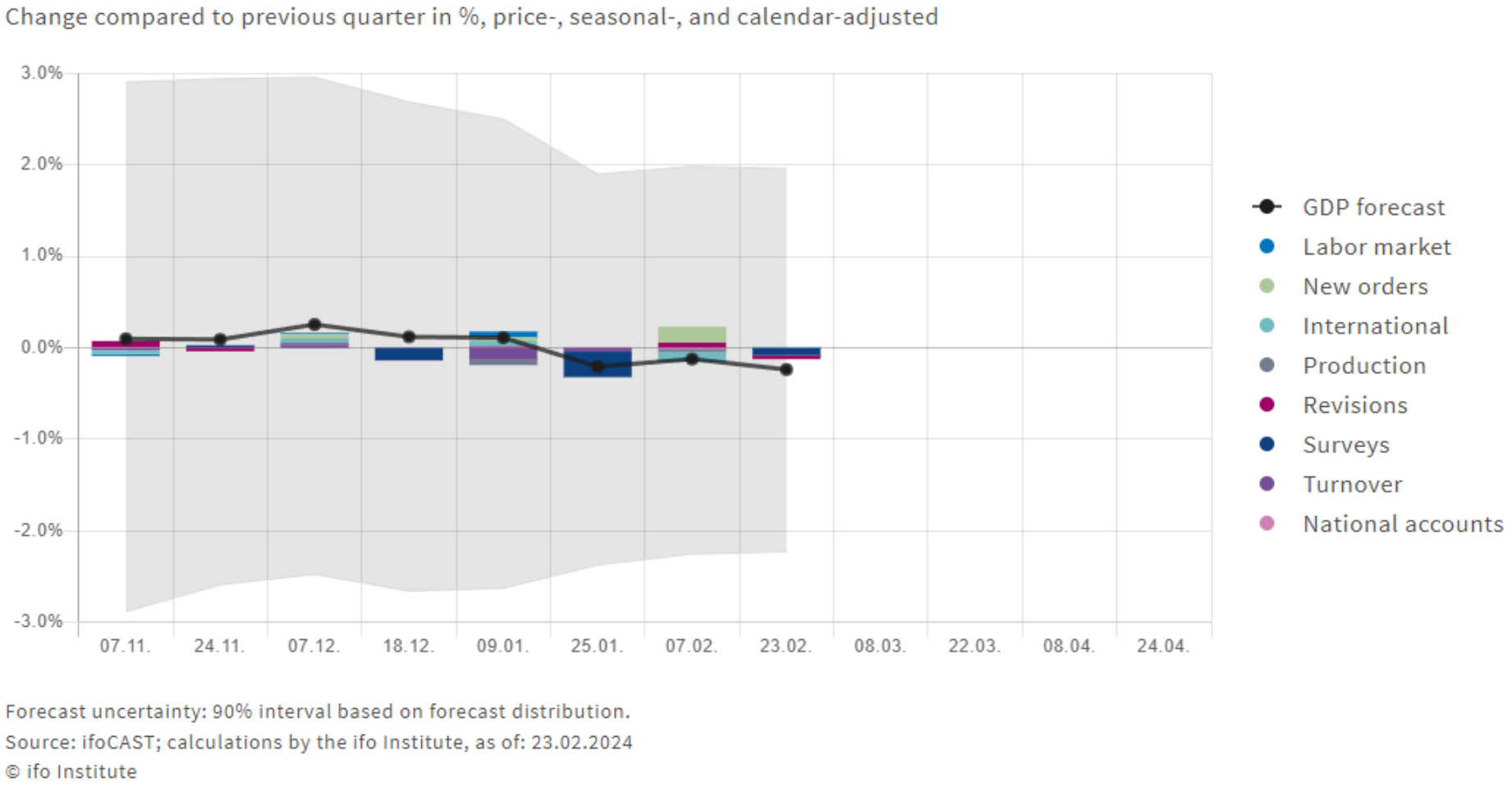
Economics and (Big) Data Science – a long history until today

Economics and (Big) Data Science – a long history until today





ifoCast –ML-based Back-, Now- and Forecasting



A Novel Trade Agreement Database

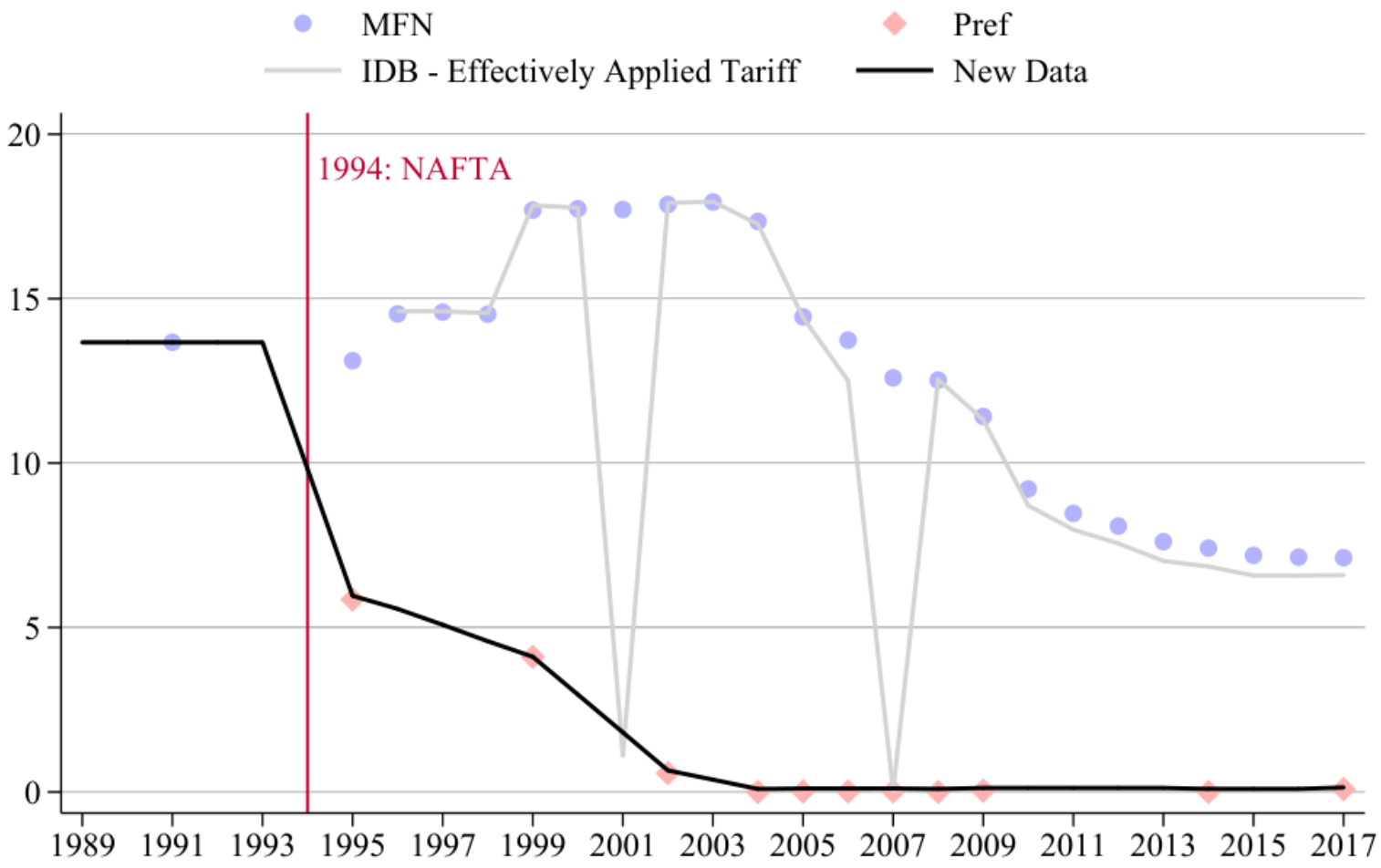
Starting point: Existing official trade data is massively flawed

Tariff data is granular and complex:

- **160 x 159 x 5,000 x 30 = 5.6 bn observations**
- Tariff, trade, sanction documents = Legal docs with mix of tables, text in different languages

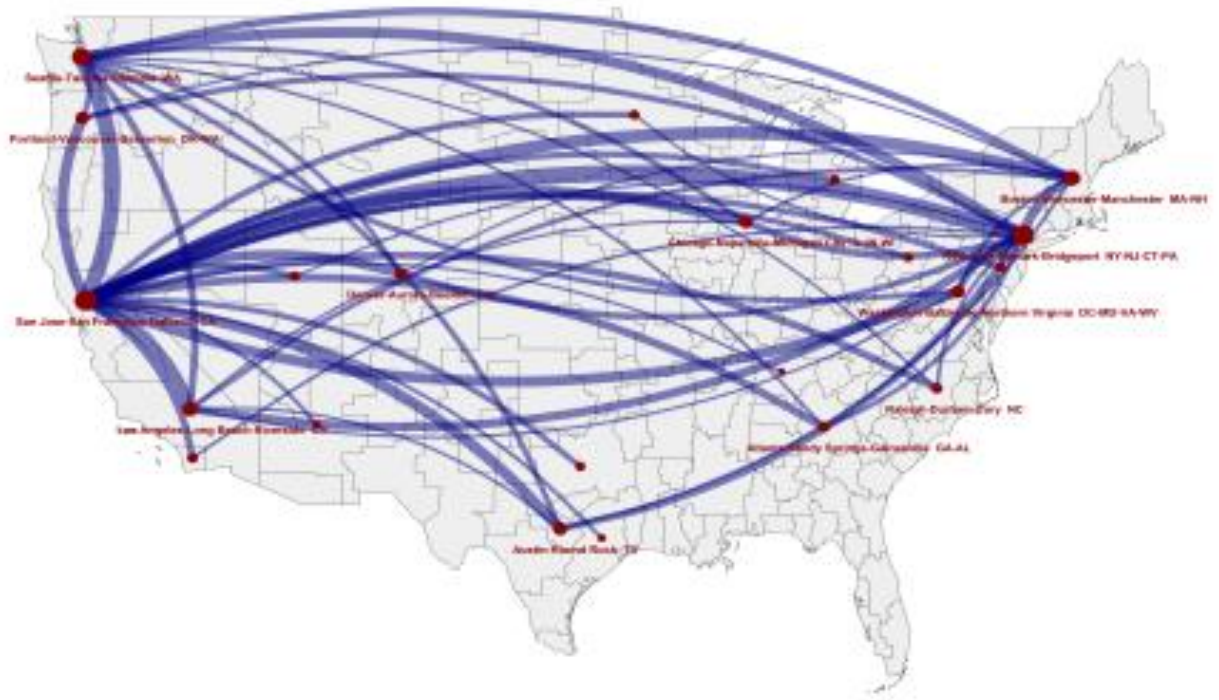
Methods: Web-scraping, NLP methods (structural topic models, text similarity), K-clustering, LASSO, causal inference

Result: Trade elasticities are highly sensitive to messy data



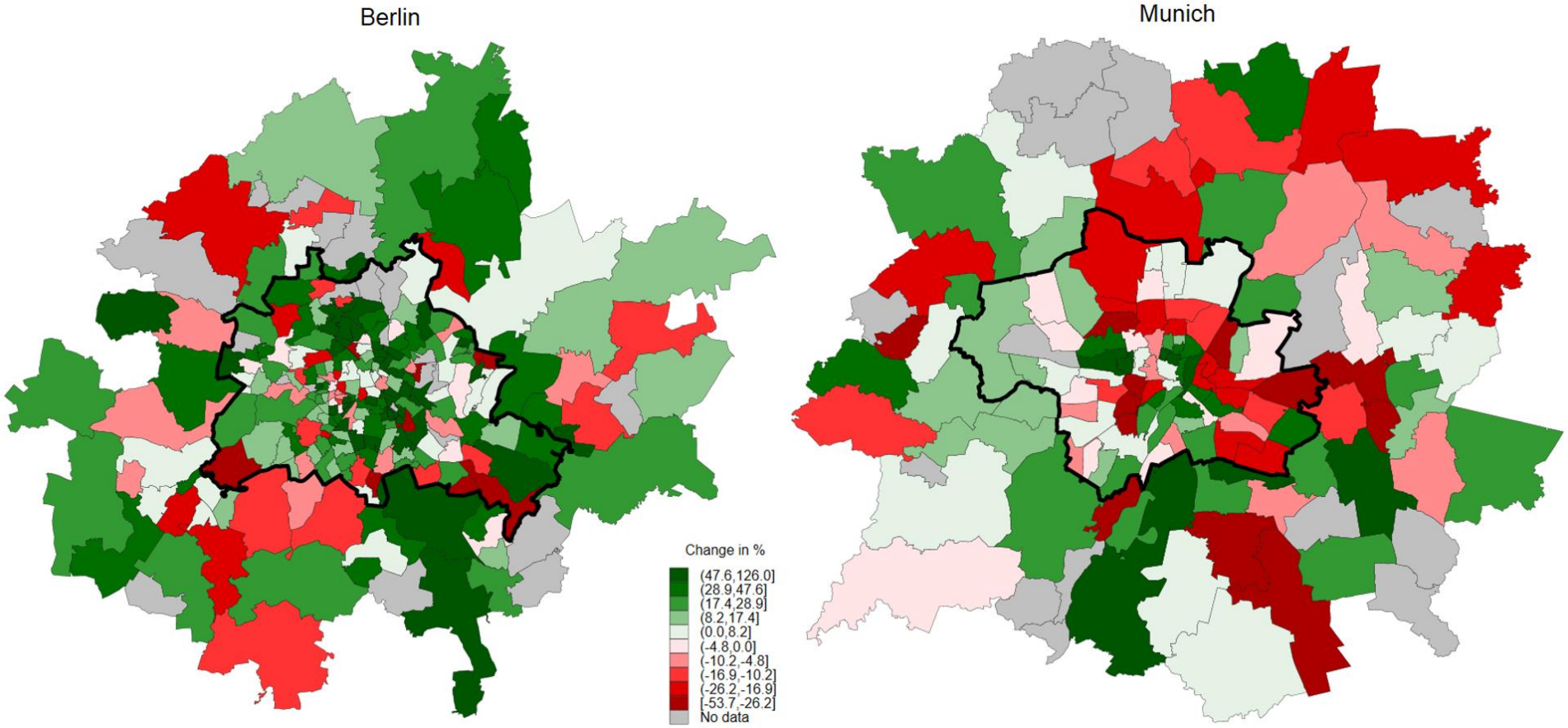
Measuring collaboration and the (still existing) role of distance using GitHub

Figure 3: Inter-regional collaboration of users



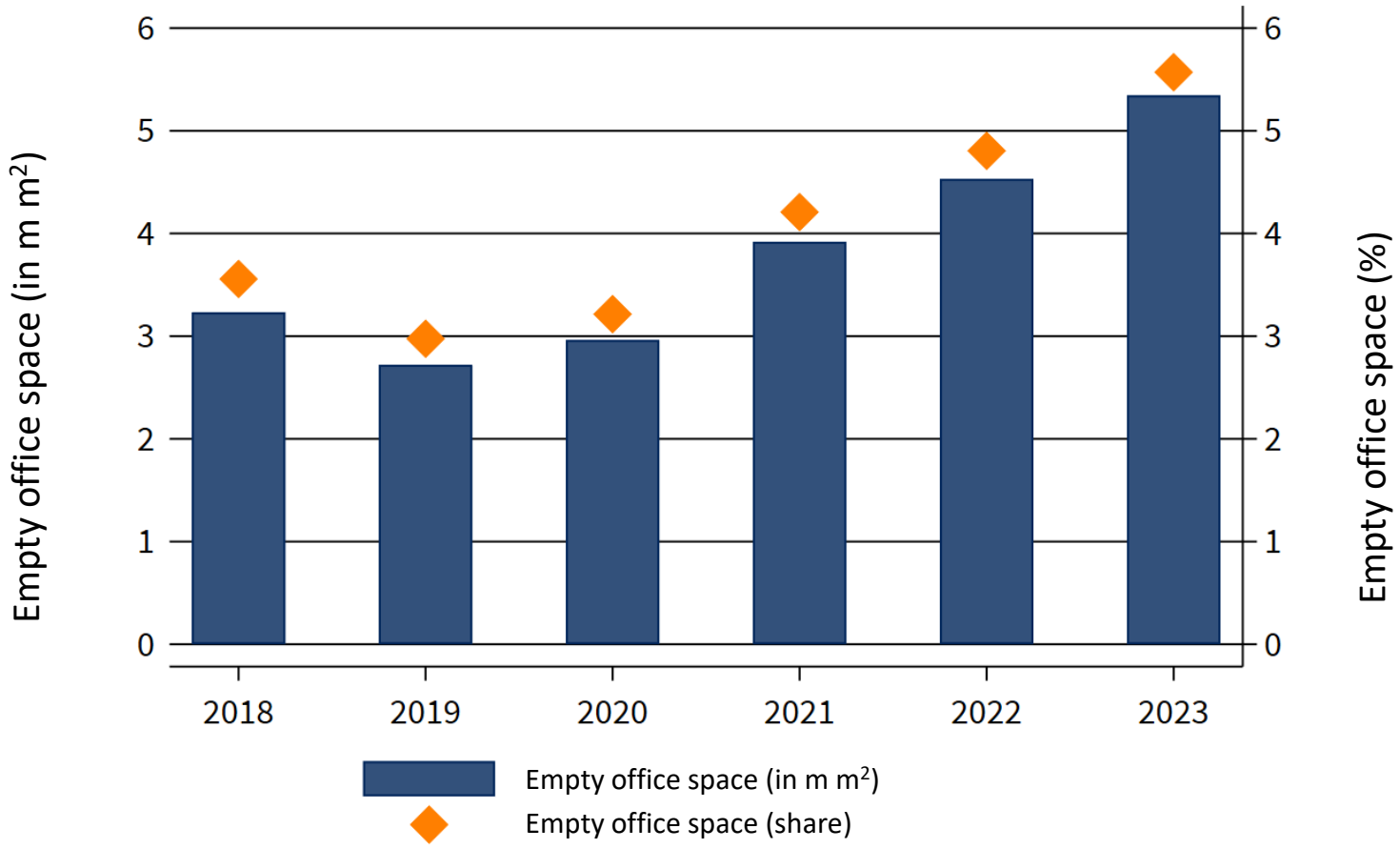
New economic geography and consumption expenditures in cities

Change in credit card expenditures 2019 vs. 2022



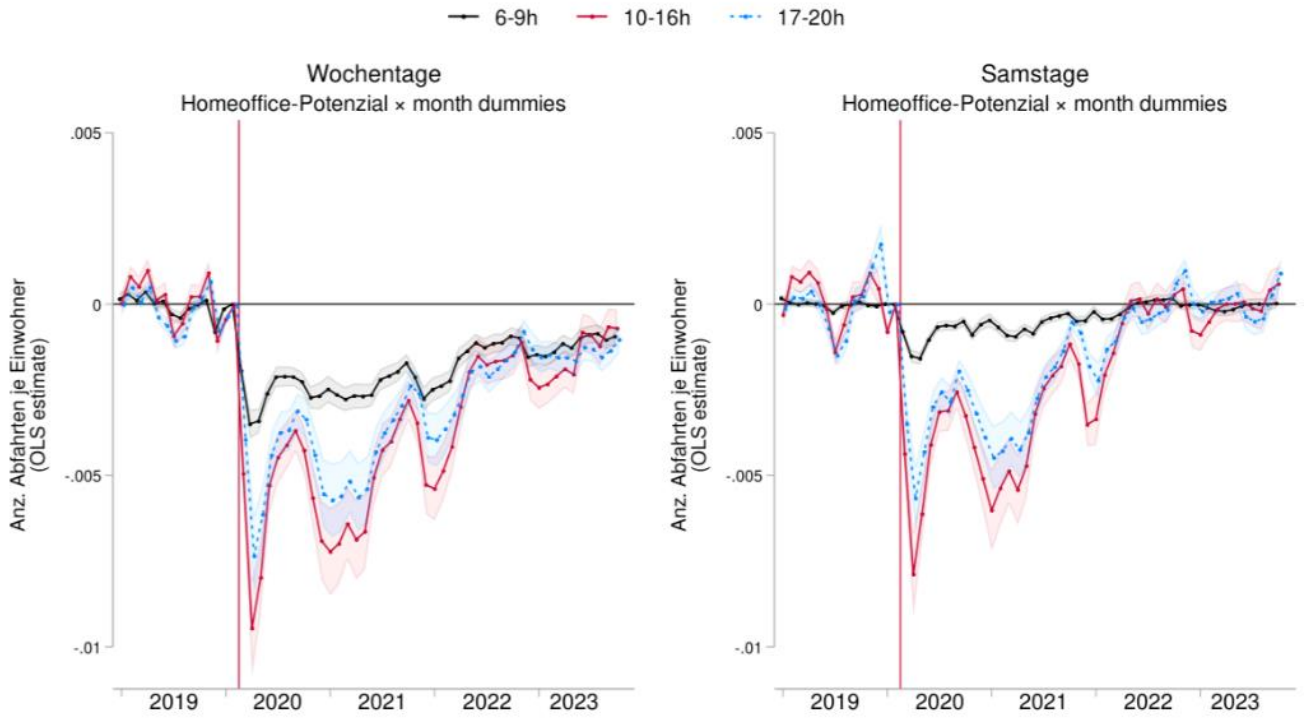
Source: Mastercard GeoInsights; Calculations by ifo Institute.

Working from home and the market for office space in Germany



Source: Colliers; Calculations by ifo Institute.

Less mobility in areas with more working from home potential



Source: Mobile phone-based mobility data from Deutsche Telekom and working from home potential by infas360. Calcuations by ifo Institute. #depatures per month on postcode level.

Several more studies

The deterrent effects of arrests in domestic violence cases (using UK police and emergency call data)

Measuring the effects of climate protests on elections (using i.a. mobility/location data)

The effects of different forms of custody (using Swedish court records)

HR studies using employee data (of course in a GDPR compliant way):

- Fostering psychological safety in teams using field experiments – benefits and measures
- Explaining talent hoarding in firms (and how to avoid it) using team evaluations and survey evidence
- Addressing female underrepresentation in leadership positions using internal job board data

….

Economic Research together with Firms – a Great Opportunity

Initial situation

- **Official statistics not sufficient:** Many (time-critical) questions cannot be answered with official statistics or surveys → *Data is not available at all, too late, or not detailed roughly*
- **Big Data:** Digitization creates "big data" (= huge, unconventional, often unstructured data) in business processes at companies
- **Long-standing expertise:** Economic researchers (@ifo and beyond) are proficient in data analysis and the application of state-of-the-art econometric methods

+ connections to other sciences

Research cooperation with firms

- **Creating a win-win situation:** Cooperation projects that bring together company data and economic, methodological expertise lead to mutual benefits for companies and science
- **Cooperation in new areas:** Pilot projects and long-standing in new fields apart from engineering, direct product development, and marketing is possible - even in (privacy-sensitive) Germany.

Cooperation between Science and Industry – Why and How?

Possible goals & requirements for science:

- Evidence-based scientific research that feeds into policy advice and public debate
- Connection to the world's leading research
- Utilizing synergy effects of the business and science location

Possible goals of business partners:

- Generate new insights to optimize business processes
- Knowledge transfer
- Recruitment data scientists
- CSR/ESG („Data for good“)/Avoiding stronger regulations
- Good PR through joint studies and press

Process:

- Joint development of research questions and, if necessary, regular coordination rounds
- Open-ended and independent research
- Cooperation always on a contractual basis, confidentiality agreement and data protection check + if necessary, anonymization and data processing only in a secure working environment

Initiatives to foster collaboration between science and industry

CESifo cluster
Economic Research with Firm Data

KonsortSWD
New Options for Access to Firm Data

PPA
PREDICTIVE PEOPLE ANALYTICS

nfdi Nationale Forschungsdaten Infrastruktur
Section Industry Engagement

BERD @NFDI

DATAGROUP BUSINESS 2 SCIENCE

Take your chance and collaborate with science!

JOIN OUR TEAM!

Applied Data Scientist/ PostDoc (f/m/d)
Big Data Economics

Shaping the Economic Debate

BIG DATA ifo Economics

THANK YOU!

wichert@ifo.de

www.linkedin.com/in/sebastianwichert

THANK YOU FOR YOUR ATTENTION!

Dr. Sebastian Wichert

E-Mail: wichert@ifo.de

LinkedIn: www.linkedin.com/in/sebastianwichert

Personal website: https://sites.google.com/view/sebastianwichert/home

ifo website: https://www.ifo.de/en

References I

ifo/LMU researchers in bold

- Slide 3 - ifo Business Climate Index: https://www.ifo.de/en/survey/ifo-business-climate-index-germany
- Slide 5 - ifoCast: https://www.ifo.de/en/ifoCAST
- Slide 6 - Tariff data **Feodora Teti**, 30 Years of Trade Policy: Evidence from 5.7 Billion Tariffs, mimeo
- Slide 7 - GitHub data: **Moritz Goldbeck**, Bit by Bit - Colocation and the Death of Distance in Software Developer Networks
- Slide 8 – Mastercard GeoInsights + Infas360 data: **Victor Alipour, Oliver Falck, Simon Krause, Carla Krolage, Sebastian Wichert**, The Future of Work and Consumption in Cities after the Pandemic: Evidence from Germany
- Slide 9 - Colliers office-real estate data: **Simon Krause**, Andreas Trumpp, Tobias Dichtl, Susanne Kiese, Alexander Rutsch, Neue Arbeitswelt, neue Arbeitsorte: Auswirkungen von Homeoffice auf den Büroimmobilienmarkt
- Slide 10 - Telekom mobility data: Unpublished analysis for project „The Future of Work and Consumption in Cities after the Pandemic: Evidence from Germany “ (see slide 8)

References II

ifo/LMU researchers in bold

- Slide 11:
  - Sofia Amaral, Gordon B. Dahl, **Victoria Endl-Geyer**, Timo Hener, **Helmut Rainer**, Deterrence or Backlash? Arrests and the Dynamics of Domestic Violence
  - Marc Fabel, Matthias Flückiger, Markus Ludwig, **Helmut Rainer, Maria Waldinger, Sebastian Wichert,** The Power of Youth: Did the "Fridays for Future" Climate Movement Trickle-Up to Influence Voters, Politicians, and the Media?
  - **Stella Canessa, Anna Hasselqvist, Helmut Rainer**, Joint versus Sole Custody after Divorce: Causal Effects on Child Outcomes and Family Environment
  - **Silvia Castro, Florian Englmaier**, Maria Guadalupe, Fostering Psychological Safety in Teams: Evidence from an RCT, mimeo
  - **Ingrid Hägele**, Talent Hoarding in Organizations
  - **Ingrid Hägele,** The Broken Rung: Gender and the Leadership Gap
- Slide 14:
  - CESifo Cluster “Economic Research with Firm Data”
  - Predictive People Analytics (PPA)
  - KonsortSWD project “New Options for Access to Firm Data”
  - BERD@NFDI Data Marketplace
  - NFDI Section “Industry Engagement”
  - Datagroup Business 2 Science