

Nutzung personenbezogener Daten in der Forschung: Chancen und Herausforderungen

Prof. Dr. Joachim Winter

Ludwig-Maximilians-Universität München

Datenschutz am Mittag

18. November 2024





Joachim Winter

- Lehrstuhl für empirische Wirtschaftsforschung an der LMU
- Schwerpunkte: Datenerhebung und -auswertung, Entscheidungen privater Haushalte (Konsum, Sparverhalten, Gesundheit)
- Mitglied des Rats für Sozial- und Wirtschaftsdaten (RatSWD) seit 2021
- Mitglied des Wissenschafts- und Innovationsbeirats der Registermodernisierung



Rat für Sozial- und Wirtschaftsdaten (RatSWD)

- Beirat der Bundesregierung (seit 2004)
- Mitglieder aus Datenproduktion und Wissenschaft (gewählt)
- Beratung der Politik beim Aufbau der deutschen Forschungsdateninfrastruktur
- Unterstützung der Datennutzerin der Wissenschaft
- Stellungnahmen, Handreichungen, Konferenzen
- KonsortSWD: entwickelt als Teil der NFDI Dienste der Forschungsdateninfrastruktur



Forschungspotenziale von Mikrodaten

- Deutschland steht vor erheblichen Herausforderungen, wenn Mikrodaten ihr **Potenzial** für eine evidenzbasierte Politik (Datenstrategie der Bundesregierung!), aber auch für die Grundlagenforschung entfallen sollen.
- Nutzungsmöglichkeiten werden aufgrund rechtlicher und technischer Beschränkungen nicht ausgeschöpft.
- Die **Beispiele europäischer Nachbarn** zeigen, wie eine hochwertige datenschutzkonforme Datenlandschaft aussehen kann (auch unter Gültigkeit der GDPR).



Quellen und Nutzung von Mikrodaten

- „Mikrodaten“ beziehen sich auf einzelne Personen oder Unternehmen; sie sind besonders zu schützen
 - Registerdaten
 - Befragungen – amtliche Statistik oder Wissenschaft (z.B. Mikrozensus, Sozio-ökonomisches Panel)
 - Sonstige „administrative“ Daten (z.B. Abrechnungs- und Behandlungsdaten im Bereich der Gesundheit)
- Umfangreiche gesetzliche Regelungen definieren und beschränken Zugang und Möglichkeiten der Verknüpfung



Praxisbeispiel zur Nutzung von Mikrodaten: Einzeldaten der Steuerstatistik („Tax Payer Panel“)

- Welche Steuerpflichtigen reichen keine Erklärung ein, obwohl sie Anspruch auf Erstattung haben?
- Welche (oft trivialen) Gestaltungsmöglichkeiten werden nicht ausgeschöpft?
- Um welche Beträge geht es? Wer zahlt zu viel Steuern? Wie sind die Verteilungswirkungen?
- Datenzugang in Deutschland gut, aber keine Verknüpfung mit anderen Daten möglich



F+ FOLGE DER BÜROKRATIE

Immer mehr Deutsche zahlen zu viele Steuern

Von Luisa Wallossek 10.04.2024, 15:34 Lesezeit: 4 Min.



Praxisbeispiel zur Nutzung von Mikrodaten: Erwerbsverlaufsdaten der BA

- Zugang zu bestimmten BA-Daten möglich, weil deren Nutzung zur Evaluierung arbeitsmarktpolitischer Maßnahmen im SGB vorgesehen ist
- Verknüpfung mit Befragungsdaten (z.B. SOEP) möglich (Zustimmung der SOEP-Befragten)
- Geschlechtsunterschiede in der Entlohnung
- Effekte von Heirat und Kindern auf den Einkommensverlauf von Frauen („child penalty“)
- Beschäftigungswirkungen des Mindestlohns

Praxisbeispiel zur Nutzung von Mikrodaten: Ungleichheitsforschung

FISCAL STUDIES, vol. 42, no. 1, pp. 147–170 (2021) 0143-5671

Geographic Inequality in Income and Mortality in Germany*

PETER REDLER,[†] AMELIE WUPPERMANN,[‡] JOACHIM WINTER,[†]
HANNES SCHWANDT[§] and JANET CURRIE[◇]



Praxisbeispiel zur Nutzung von Mikrodaten: Wechsel Faulheit

American Economic Review 2021, 111(9): 2737–2781
<https://doi.org/10.1257/aer.20170471>

Inattention and Switching Costs as Sources of Inertia in Medicare Part D[†]

By FLORIAN HEISS, DANIEL MCFADDEN, JOACHIM WINTER,
AMELIE WUPPERMANN, AND BO ZHOU*

Consumers' health plan choices are highly persistent even though optimal plans change over time. This paper separates two sources of inertia, inattention to plan choice and switching costs. We develop a panel data model with separate attention and choice stages, linked by heterogeneity in acuity, i.e., the ability and willingness to make diligent choices. Using data from Medicare Part D, we find that inattention is an important source of inertia but switching costs also play a role, particularly for low-acuity individuals. Separating the two stages and allowing for heterogeneity is crucial for counterfactual simulations of interventions that reduce inertia. (JEL D12, G22, H51, I13, I18, L65)



Was derzeit eingeschränkt oder nicht möglich ist

- Größtes Defizit: Verknüpfung von Mikrodaten aus verschiedenen Quellen
- **Ungleichheit:**
Mikrozensus & Befragungsdaten (Familienhintergrund)
Mikrozensus & Abrechnungsdaten (Gesundheit)
- **Steuerkomplexität:**
Steuerdaten & Befragungsdaten (Einfluss der Bildung)
- **Wechselhaftigkeit bei Verträgen:**
Im Gesundheitsbereich nicht möglich



Status Quo: Forschungsdatenzentren (FDZ)

- Forschungsdatenzentren ermöglichen die Nutzung solcher Daten durch die Wissenschaft
- Angesiedelt bei datenproduzierenden Einrichtungen (Statistische Ämter, BA, RV, Forschungsinstitute)
- Zugangswege:
 - Gesicherte Arbeitsplätze im FDZ
 - Remote Execution/Access (mit Outputkontrolle)
 - Lieferung der Daten an die Nutzenden
- Gründung erstes FDZ 2001; aktuell 41 FDZ, die vom RatSWD akkreditiert sind





Wie geht es weiter? Das Forschungsdatengesetz

- Ziel: Verbesserung und Vereinfachung des Zugangs zu Daten für öffentliche und private Forschung

Meilensteine und Zeitplan

- Stakeholder-Konsultation (u.a. RatSWD)
- Eckpunktepapier des BMBF ([download](#))
- Haus- und Ressortabstimmung
- Referentenentwurf (11.10.2024)
- Kabinettsbefassung (ab November)
- Parlamentarisches Verfahren (Ziel: Anfang 2025)



Eckpunkte

- Sektorenübergreifender Ansatz
- Aber: separates Gesundheitsdatennutzungsgesetz
- Ein „Micro Data Center“ soll den Zugang zu Daten der öffentlichen Hand öffnen, vereinfachen und erweitern und Verknüpfungsmöglichkeiten schaffen
- Der Datenschutz soll forschungsfreundlicher gestaltet werden und die besonderen Bedürfnisse der Forschung besser berücksichtigen
- Die Auffindbarkeit von Forschungsdaten soll verbessert werden



Referentenentwurf

- Artikelgesetz mit Stammgesetz und Änderungen in weiteren Bundesgesetzen (Bundesstatistikgesetz, SGB II und X, Verwaltungsdatenverwendungsgesetz, Statistikregistergesetz, Hochschulstatistikgesetz, ...)
- Zentrale Regelungen:
 - Aufbau und Einrichtung eines Deutschen Zentrums für Mikrodaten („DZM“) am Statistischen Bundesamt
 - Verbesserungen im Datenschutzrecht für Datenzugang und -nutzung zugunsten der Forschung (u.a. „federführende Datenschutzaufsichtsbehörden“)
 - Verbesserung der Auffindbarkeit von Forschungsdaten (Verpflichtung zur Bereitstellung von „Metadaten“)



Weitgehender Datenzugang für die Forschung

- §4 (2): Forscherinnen und Forscher können Zugang zu folgenden Daten beantragen:
 1. Daten, die von registerführenden Behörden des Bundes und der Länder zum Zwecke der Registerführung erhoben oder bezogen werden
(...)
- §4 (1): Voraussetzungen:
 1. Angehörigkeit zu einer akkreditierten Forschungseinrichtung
 2. Erforderlichkeit für das konkrete Forschungsvorhaben



Einschätzung

- Sehr forschungsfreundlicher Ansatz (von Details abgesehen)
- Erste Reaktionen legen nahe, dass noch Klärungsbedarf in verschiedener Hinsicht besteht, z.B.
 - Verhältnis des DZM zu bestehenden Forschungsdatenzentren (FDZ)
 - Implikationen der möglichen kommerziellen Nutzung von über das DZM bereitgestellten Daten
 - Umsetzung der Akkreditierung von Forschungseinrichtungen
- Was macht die neue Regierung?

Vielen Dank für Ihr Interesse!

Anmerkungen gerne an winter@lmu.de