



Corinna Kleinert



# WELCHE DATEN HABEN WIR UND WELCHE BENÖTIGEN WIR IN DER ARBEITSMARKTFORSCHUNG?



16. Februar 2018

SAMF-Jahrestagung | Berlin

# Was Sie erwartet

2

## Inhalte

- Veränderungen und Konstanten auf Arbeitsmärkten
- Vorhandene Daten: Überblick und Analysepotenziale
- Erkenntnislücken und Innovationen

## Blickwinkel

- Empirische Arbeitsmarktforschung
- Mikroebene
- Soziologie
- Vergleich mit Bildungsforschung: NEPS

# Wozu sollten Daten dienen?

3

## Berichterstattung

- Kontinuität und Wandel über die Zeit: Trenddaten
- Abbildung neuer Entwicklungen

## Forschung

- Kausale Prozesse und Mechanismen
- Erwerbskarrieren und Betriebshistorien: Längsschnittdaten
- Zusammenspiel von Angebot und Nachfrage: Multi-Akteur-Daten

## Beratung

- Wirkungsforschung und Implementation
- Maßnahmeevaluation, Experimente, qualitative Begleitstudien

# Arbeitsmärkte im zyklischen Wandel

4

## **Arbeitsnachfrage**

- Konjunkturzyklen
- Saisonale Schwankungen
- Regionale Märkte

## **Arbeitsangebot**

- Mobilität und Migration
- Regionale Differenzen

# Arbeitsmärkte im strukturellen Wandel

5

## Arbeitsnachfrage

- Technologischer Wandel (Digitalisierung, Arbeit 4.0)
- Freisetzungsprozesse, Veränderung von Arbeitsanforderungen und Jobprofilen, Entstehen neuer Tätigkeiten und Berufe
- Veränderungen in der Organisation von Arbeit: internet-, projektbasiert
- Internationale Verflechtung

## Arbeitsangebot

- Bildungsexpansion und Tertiarisierung
- Demographischer Wandel und Alterung
- Zunahme und veränderte Steuerung internationaler Migration

# Welche Daten gibt es?

## Administrative Daten

Integrierte Erwerbsbiographien  
Betriebshistorikpanel

## Betriebsdaten

IAB-Stellenerhebung  
Verdienst(struktur)-  
erhebungen  
IAB-Betriebspanel

## Trenddaten Erwerbstätige

Mikrozensus  
BIBB-IAB- und BIBB-BAUA  
Verdienst(struktur)erhebungen

## Multi-Akteur-Daten

Linked-Employer-Employee-  
Daten des IAB (LIAB)  
LPP - Linked Personnel Panel  
Panel WeLL  
SOEP-LEE

## Internationale Vergleiche

Europ. Arbeitskräfteerhebung  
Europ. Work. Condition Survey  
PIAAC\*

## Längsschnittdaten Erwerbsverläufe

SOEP (Haushalte)  
ALWA / NEPS (Individuen)\*  
PASS (Bedarfsgemeinschaften)\*  
PIAAC-L (Individuen und Partner)\*

## Einzelstudien

...

\* Verlinkung mit Integrierten Erwerbsbiographien verfügbar oder geplant

# Wandel der Datenlandschaft

7

## Was wurde bisher erreicht?

- Steigende Speicher- und Verarbeitungskapazität
- Administrative Daten: valide und reliable Informationen zu Erwerbskarrieren und Löhnen (Berufe, Branchen, Regionen)
- Verbesserung von Datenverfügbarkeit und Zugängen (SUFs, JoSuA)
- Record Linkage: Anreicherungen von Befragungen mit administrativen Längsschnittdaten
- Räumliche Informationen: Georeferenzierung

**Die Datenlandschaft im Bereich Arbeitsmarkt ist weitaus „dichter“ als im Bereich der Bildung**

**Teilweise sind die Datenstrukturen und -angebote aber noch nicht ausreichend bekannt und genutzt**

# Brauchen wir neue Datenstrukturen?

8

## Lessons learnt from NEPS: Don't try too much at the same time!

- Längsschnitt- und Mehrebenendesigns harmonisieren nur bedingt
- Linked-Employer-Employee-Daten geben keine Auskunft über Freelancer, Selbständige, Arbeitslose und Nicht-Erwerbstätige
- Panelstudien sind nicht gut geeignet, um Aussagen über Trends zu machen
- Administrative Daten sind relativ erkenntnisarm und spiegeln in Teilen eher Nachfrage als Angebot wieder (Berufe, Bildung)
- Individuen können Maßnahmen, Leistungsbezug nicht gut wiedergeben
- Befragte machen nicht alles mit: Belastungsgrenzen

**Am Anfang sollten klar formulierte Funktionen, Ziele, Akteure und Populationen stehen, die Design vorgeben**



# Welche Daten fehlen uns?

9

## **Löhne und Arbeitszeiten in einem Datensatz**

- Wichtig vor allem für Berichterstattung und Dauerbeobachtung
- Herausforderung für bestehende Datenstrukturen

## **Neue Erwerbsformen im Längsschnitt**

- Sehr schnelle Statusänderungen zu erwarten
- Herausforderung für Erhebungsfrequenz und Befragtenbelastung
- Integration in bestehende Datenstrukturen kaum möglich

## **Multi-Akteur-Studien mit qualitativen Informationen**

- Zwei mögliche Designs: Individuen > Betriebe oder Betriebe > Beschäftigte
- Problem bei beiden: Längsschnitt- versus Mehrebenenlogik
- Sehr anspruchsvoll, teuer und belastend, v.a. für (Groß-)Betriebe

## **Systematischer Methodenmix**

- Anreicherung quantitativer Längsschnitte mit qualitativen Längsschnittstudien
- Nutzbarmachung von qualitativen Daten für Sekundäranalysen

# Die Zukunft(?)

10

## Stärkere Verzahnung bestehender Datenstrukturen

- Koordinierte Verknüpfung amtsstatistischer Datenquellen zu Betrieben und Beschäftigten/Bevölkerung (Beispiel KombiFid)
- Entlastung von Betrieben und Individuen
- Wichtigste Voraussetzung: datenschutzrechtliche Änderungen, Schaffung einer dauerhaften gesetzlichen Grundlage
- Wünschenswert, aber in weiter Ferne: internationale Verknüpfung

## Nutzung neuer technischer Möglichkeiten

- Befragungen per App auf digitalen Geräten (Beispiel IAB Smart-App)
- Erlaubt Verknüpfung mit automatisch erfassten Daten: Bewegungen, Zeitnutzung, soziale Netzwerke, Nutzung sozialer Medien...
- Unter der Voraussetzung gezielter Forschungsfragen und Einhaltung forschungsethischer Leitlinien

11

Danke für Ihre Aufmerksamkeit!

# Literatur

12

- Biewen, Elena, Anja Gruhl, Christopher Gürke, Tanja Hethey-Maier und Emanuel Weiß (2012): Kombinierte Firmendaten für Deutschland. Möglichkeiten und Konsequenzen der Zusammenführung von Unternehmensdaten unterschiedlicher Datenproduzenten. FDZ Methodenreport 05/2012. Nürnberg.
- Bonin, Holger (2017): Berichterstattung über die Arbeitswelt in Deutschland: Datengrundlagen für Branchenbetrachtungen. Kurzexpertise. Bundesministerium für Arbeit und Soziales FB488. Bonn.
- Dütsch, Matthias, Ralf K. Himmelreicher und Clemens Ohlert (2017): Zur Berechnung von Bruttostundenlöhnen - Verdienst(struktur)erhebung und Sozioökonomisches Panel im Vergleich. SOEPpapers on Multidisciplinary Panel Data Research, No. 911. Berlin.
- Schmucker, Alexandra, Heiko Stüber und Silke Hamann (2017): Sekundärdaten für die Organisationsforschung – Datenangebot und Nutzungsmöglichkeiten. S. 17-36 in S. Liebig et al. (Hrsg.), Handbuch Empirische Organisationsforschung, Springer Reference Wirtschaft.
- Schneider, Hilmar (2009): Mehr und bessere Daten für die Arbeitsmarktforschung. Vorschläge zur effizienten Erschließung von bislang ungenutztem Potenzial der amtlichen Statistik. RatSWD Working Paper No. 106. Berlin.

# Informationen in wichtigen Datensätzen

Tabelle 2.2 Merkmals-Matrix zur Teilauswahl der zehn wichtigsten untersuchten Datensätze

Datensatzname	MZ	SOEP	NEPS	PIAAC	WIF	BIBB-BAuA	BAuA	EWCS	VSE	IAB BP	
Datentyp*	Bevölkerung/Haushalte				Erwerbstätige				Betriebe		
Fallzahlen	820.000	30.000	9.700	5.000	19.400	20.000	20.030	2.000	71.000	16.000	
Stellen der Klassifikation	3-Steller	2-Steller	2-Steller	4-Steller	4-Steller	5-Steller	2-Steller	2-Steller	5-Steller	5-Steller	<b>Summe</b>
a) Entlohnung/Leistungspolitik [...]	2	13	10	6	14	2	1	6	20	4	<b>78</b>
b) Arbeitsplatzsicherheit [...]	15	35	25	17	28	29	34	33	13	11	<b>240</b>
c) Tätigkeiten/Gesundheit/[...]	17	21	36	22	6	24	19	26	2	5	<b>178</b>
d) Vielfalt im Betrieb	2	1	0	2	8	1	1	4	2	13	<b>34</b>
e) Arbeitsbeziehungen/[...]	0	3	0	0	4	0	1	2	2	3	<b>15</b>
<b>Summe</b>	<b>36</b>	<b>73</b>	<b>71</b>	<b>47</b>	<b>60</b>	<b>56</b>	<b>56</b>	<b>71</b>	<b>39</b>	<b>36</b>	<b>233</b>

Quelle: Bonin 2017

# Neue Datenstrukturen: Voraussetzungen

14

## Der Aufwand für qualitativ hochwertige Surveys wird gerne unterschätzt

- Zielpopulation und Stichprobe: Wer soll befragt werden? Aus welchen Informationen kann eine repräsentative Stichprobe gezogen werden? Wie groß muss diese sein, um das Ziel zu erreichen?
- Ausschöpfung und Selektivität: Wie kann das Programm so gestaltet werden, dass alle Teilnehmergruppen gleichermaßen dabei bleiben?
- Operationalisierung und Itementwicklung: Wie können die theoretisch relevanten Konstrukte valide und reliabel gemessen werden?