

Wagner, Joachim

Article

Daten des IAB-Betriebspanels und Firmenpaneldaten aus Erhebungen der Amtlichen Statistik – substitutive oder komplementäre Inputs für die Empirische Wirtschaftsforschung?

Journal for Labour Market Research

Provided in Cooperation with:

Institute for Employment Research (IAB)

Suggested Citation: Wagner, Joachim (2014) : Daten des IAB-Betriebspanels und Firmenpaneldaten aus Erhebungen der Amtlichen Statistik – substitutive oder komplementäre Inputs für die Empirische Wirtschaftsforschung?, Journal for Labour Market Research, ISSN 2510-5027, Springer, Heidelberg, Vol. 47, Iss. 1-2, pp. 63-70, <https://doi.org/10.1007/s12651-013-0129-y>

This Version is available at:

<https://hdl.handle.net/10419/178319>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Daten des IAB-Betriebspanels und Firmenpaneldaten aus Erhebungen der Amtlichen Statistik – substitutive oder komplementäre Inputs für die Empirische Wirtschaftsforschung?

Joachim Wagner

Online publiziert: 7. Februar 2013
© Institut für Arbeitsmarkt- und Berufsforschung 2013

Zusammenfassung Repräsentative Längsschnittdaten für Firmen (Betriebe bzw. Unternehmen), die für wissenschaftliche Auswertungen genutzt werden können, stammen in Deutschland vor allem aus zwei Quellen, dem *IAB-Betriebspanel* und den *AFiD-Panel*, wobei das Akronym AFiD für Amtliche Firmendaten für Deutschland steht, die von den Statistischen Ämtern erhoben werden. In diesem Vortrag werden sechs Aspekte für einen Vergleich dieser beiden Datenquellen herangezogen. Es zeigt sich, dass es sich hierbei um komplementäre und nicht um substitutive Inputs für die Empirische Wirtschaftsforschung handelt. Beide haben jeweils spezifische Stärken und Schwächen, die sie zur Analyse von konkreten Fragestellungen mehr oder weniger gut geeignet oder auch völlig ungeeignet machen. Daraus folgt, dass der Werkzeugkasten für die empirischen Forschung mit Längsschnittdaten für Firmen in Deutschland sowohl das IAB-Betriebspanel als auch die AFiD-Panel enthalten muss, denn nur so ist sicher gestellt, dass ein möglichst breites Spektrum von Fragestellungen mit jeweils adäquaten Daten untersucht werden kann

Are data from the IAB Establishment Panel and firm panel data from Official Statistics substitutes or complements for empirical economic research?

Prof. Dr. J. Wagner (✉)
Institut für Volkswirtschaftslehre/Empirische
Wirtschaftsforschung, Leuphana Universität Lüneburg,
Postfach 2440, 21314 Lüneburg, Deutschland
e-mail: wagner@leuphana.de
url: <http://www.leuphana.de/joachim-wagner.html>

Prof. Dr. J. Wagner
IZA, Bonn, Deutschland

Abstract Representative longitudinal data for German firms (establishments or enterprises) that can be used for scientific analyses are mainly from two sources, the IAB Establishment Panel and the AFiD-Panel based on surveys performed by the Statistical Offices. This paper compares the data from these two sources. It is argued that the two types of firm data are complements and not substitutes for empirical economic analyses. Both have specific strengths and weaknesses, and both are well suited to investigate some but not all topics, and some topics cannot be analyzed with one type of data at all. Therefore, both types of data are needed as inputs for empirical economic research.

1 Motivation

Repräsentative Längsschnittdaten für Firmen (Betriebe bzw. Unternehmen) stammen in Deutschland vor allem aus zwei Quellen:¹

- Seit 1993 wird im Auftrag des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit (IAB) jährlich eine repräsentative geschichtete Stichprobe aller Betriebe mit mindestens einer sozialversicherungspflichtig beschäftigten Person befragt. Die Angaben aus den

¹Neben diesen inhaltlich sehr breiten Paneldaten sind folgende thematisch fokussierte Längsschnittdaten zu nennen: Das Betriebshistorik-Panel, für das prozessproduzierte Daten aus den Individualangaben zu den sozialversicherungspflichtig Beschäftigten auf Betriebsebene aggregiert und über die Jahre verknüpft werden, und das detaillierte Angaben zur Zusammensetzung der Belegschaft nach Alter, Qualifikation, Beruf, Geschlecht, Nationalität und Entlohnung enthält (Spengler 2008); die Bilanzstatistik der Deutschen Bundesbank mit Angaben aus Unternehmensbilanzen (Stöss 2001); die MiDi-Datenbank der Deutschen Bundesbank zur Direktinvestitionsverflechtung deutscher Unternehmen (Lipponer 2003).

Befragungswellen werden über die Jahre verknüpft und bilden das *IAB-Betriebspanel* (Fischer et al. 2009).

- Die Statistischen Landesämter erheben regelmäßig in Betrieben und Unternehmen aus allen Teilen der Wirtschaft Daten zu einer Vielzahl von Themen. Diese Angaben können über die einzelnen Wellen einer Erhebung sowie über verschiedene Erhebungen miteinander verknüpft und für ganz Deutschland zusammengeführt werden. Die so entstehenden Längsschnittdaten für Betriebe und Unternehmen sind die so genannten *AFiD-Panel*, wobei das Akronym AFiD für Amtliche Firmendaten für Deutschland steht (Malchin und Voshage 2009).

Wissenschaftler können das IAB-Betriebspanel und die AFiD-Panel für Analysen nutzen. Der Zugang ist jeweils projektspezifisch vertraglich zu regeln und erfolgt über das Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (Kohlmann 2005) bzw. über die FDZ des Statistischen Bundesamtes und der Statistischen Ämter der Länder (Zühlke et al. 2004). Aus Nutzersicht ist hierbei wichtig, dass bei der Verwendung der Daten des IAB-Betriebspanels keine Gebühren anfallen; die Kosten für die Nutzung eines AFiD-Paneldatensatzes mit zahlreichen Wellen können dagegen insbesondere dann, wenn projektspezifische Erweiterungen erforderlich sind, erheblich sein.

Ein wesentlicher Unterschied zwischen den beiden Paneldatensätzen besteht in der Art ihrer Entstehung und den damit verbundenen Kosten. Während die AFiD-Panel auf vorhandene Datenbestände zurückgreifen können und diese Daten lediglich neu aufbereitet (d.h. unter Verwendung entsprechender Identifikatoren über die Wellen einer Erhebung und dann über mehrere Erhebungen verknüpft) werden müssen, ist für das IAB-Betriebspanel jedes Jahr eine eigenständige Erhebung durch Interviewer eines Befragungsinstituts sowie eine anschließende Aufbereitung der erhobenen Daten erforderlich. Auf der Grundlage meiner umfangreichen Erfahrungen aus zwanzig Jahren Arbeit mit Firmenpaneldaten versuche ich in diesem Beitrag zu zeigen, warum das IAB-Betriebspanel und die AFiD-Panel komplementäre Inputs für die Empirische Wirtschaftsforschung bilden. Beide Datensätze haben spezifische Stärken und Schwächen, die sie zur Analyse von konkreten Fragestellungen jeweils mehr oder weniger gut geeignet machen. Dies soll hier – oft unter Verweis auf eigene empirische Studien von mir – gezeigt werden.

2 Sechs Aspekte für einen Vergleich von IAB-Betriebspanel und AFiD-Panel

2.1 Fallzahlen

Für das IAB-Betriebspanel wird jährlich eine nach Größenklassen und Wirtschaftszweigen geschichtete Stichprobe al-

ler Betriebe mit mindestens einer sozialversicherungspflichtig beschäftigten Person befragt; zurzeit umfasst diese Stichprobe rund 16.000 Betriebe. Bei den AFiD-Panel handelt es sich vielfach um Totalerhebungen der Grundgesamtheit (teils mit einer Abschneidegrenze z. B. bezüglich der Anzahl der tätigen Personen wie beim Monatsbericht für Betriebe des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden (Konold 2007)) oder um sehr große Stichproben [wie etwa bei der Kostenstrukturerhebung im Verarbeitenden Gewerbe, im Bergbau sowie in der Gewinnung von Steinen und Erden (Fritsch et al. 2004) oder bei der Strukturerhebung im Dienstleistungssektor Vogel 2009].

Das IAB-Betriebspanel ist aufgrund der Stichprobengröße für kleinräumige Analysen und für Studien zu eng abgegrenzten Bereichen der Wirtschaft nicht geeignet. Hierfür eignen sich die AFiD-Panel sehr viel besser; sie erlauben z. B. Analysen für Betriebe aus spezifischen Industrien in einzelnen Bundesländern [wie eine Untersuchung der Komponenten der Arbeitsplatzdynamik im Industriezweig „Herstellung von Hütten- und Walzwerkeinrichtungen“ in Niedersachsen von 1978 bis 1990 in Gerlach und Wagner (1995)] oder Analysen auf der Ebene von Kreisen bzw. kreisfreien Städten (wie der Arbeitsplatzentwicklung in der Landeshauptstadt Hannover in den Bereichen Spitzentechnik, höherwertige Technik und in der sonstigen Industrie von 1978 bis 1994 in Wagner (1995)).

- Bei dem Aspekt „Fallzahlen“ liegt der Vorteil bei den AFiD-Panel.

2.2 Fragenprogramm

Das Fragenprogramm der Erhebungen der Amtlichen Statistik, die die Basis der AFiD-Panel bilden, ist jeweils gesetzlich fixiert und umfasst vielfach nur wenige Merkmale. Eine Folge davon ist, dass es nicht möglich ist, die Erhebungen kurzfristig als Reaktion auf aktuelle Entwicklungen und Datenanforderungen zu ergänzen. Ferner werden ausschließlich Fakten erhoben; subjektive Einschätzungen und Erwartungen sind nicht Gegenstand der amtlichen Erhebungen.

Im Gegensatz dazu liegt die Entscheidung über das Fragenprogramm des IAB-Betriebspanels in der Hand der für die Umfrage verantwortlichen Wissenschaftler, wobei diese Entscheidung im Einvernehmen mit den am IAB-Betriebspanel beteiligten Institutionen (IAB, IWH und Länderministerien) getroffen wird. Dieses Fragenprogramm ist sehr umfangreich, eine Reaktion auf aktuelle Geschehnisse ist möglich und subjektive Einschätzungen und Erwartungen der Entscheidungsträger im Betrieb nehmen einen breiten Raum ein.

Ein Beispiel für eine aus ökonomischer Sicht wichtige Angabe, die sich im IAB-Betriebspanel – aber nicht

in den AFiD-Panel – findet, ist die Information über die Existenz eines Betriebsrates. Die entsprechende Angabe im IAB-Betriebspanel hat die empirische Mitbestimmungsforschung in Deutschland in erheblichem Umfang gefördert [vgl. z. B. Addison et al. (2003) oder Schank et al. (2004)] sowie die umfassenden Überblicke in Jirjahn (2011) und Stegmeier (2012)]; sie bildet damit eine wichtigen Grundlage für eine evidenzbasierte Diskussion der wirtschafts- und gesellschaftspolitisch relevanten Frage nach den ökonomischen Folgen der betrieblichen Mitbestimmung der Arbeitnehmer.

Weitere Beispiele für regelmäßig im IAB Betriebspanel erhobene Angaben, für die entsprechende Informationen in Erhebungen der amtlichen Statistik fehlen, sind die Verbreitung von betrieblichen Bündnissen für Beschäftigung und Standortsicherung, Tarifbindung, Gewinn- und Kapitalbeteiligung, Prozess- und Produktinformationen und Frauen in Führungspositionen. Diese Liste ließe sich lange fortsetzen; die Anzahl sehr informativer und aus wissenschaftlicher sowie politischer Perspektive relevanter Studien auf der Basis dieser Daten ist entsprechend hoch.

Eine weitere Stärke des IAB-Betriebspanels verglichen mit den AFiD-Daten besteht darin, dass das Befragungsprogramm sehr flexibel an laufende Entwicklungen angepasst werden kann. Damit können kurzfristig Informationen für eine repräsentative Stichprobe von Betrieben bereitgestellt werden, die dann in die Analyse aktueller wirtschaftspolitisch relevanter Fragen und eine fundierte Politikberatung einfließen können. Hierbei können im Gegensatz zu den Standardbefragungen der Amtlichen Statistik auch Erwartungen und subjektive Einschätzungen erhoben und ausgewertet werden. Beispiele hierfür sind die erwartete Entwicklung des Umsatzes und die Einschätzung des technischen Standes der Produktionsanlagen.

Die Amtliche Statistik ist, was das Befragungsprogramm betrifft, allerdings nicht vollständig unflexibel. Neben den regelmäßigen Erhebungen mit einem festen Fragenprogramm gibt es auch die in § 7 des Bundesstatistikgesetzes (BstatG) geregelten „Erhebungen für besondere Zwecke“. Hierbei handelt es sich um Bundesstatistiken ohne Auskunftspflicht bei höchstens 20.000 Befragten, die durchgeführt werden können zur „Erfüllung eines kurzfristig auftretenden Datenbedarfs für Zwecke der Vorbereitung und Begründung anstehender Entscheidungen oberster Bundesbehörden . . . , wenn eine oberste Bundesbehörde eine solche Bundesstatistik fordert“ oder zur „Klärung wissenschaftliche-methodischer Fragestellungen auf dem Gebiet der Statistik“.

Beispiele für solche Sondererhebungen zu aktuellen Fragestellungen sind die in Kooperation mit dem Statistischen Amt der Europäischen Gemeinschaften (Eurostat) durchgeführte Piloterhebung zum Thema „International Sourcing“

(Zwania 2008) sowie die Befragung zum „Zugang kleiner und mittlerer Unternehmen zu Finanzmitteln“ (Söllner 2011). Hierbei handelt es sich allerdings jeweils um einmalige Erhebungen (auch wenn der § 7 BstatG in Absatz 5 Wiederholungsbefragungen ausdrücklich zulässt) bei einem jeweils spezifisch abgegrenzten Berichtskreis (bei der Umfrage zum Zugang zu Finanzmitteln z. B. um „unabhängige Unternehmen mit 10 bis 49 Beschäftigten der nicht-finanziellen Wirtschaft, die vor 2006 gegründet wurden“; Söllner 2011, S. 621). Verlaufsanalysen und Analysen von kausalen Ursache-Wirkungs-Zusammenhängen sind mit diesen Daten nicht möglich.² Da nur bis zum Zeitpunkt der einmaligen Erhebung überlebende Firmen befragt werden konnten ist zudem mit einem „survivor bias“ zu rechnen. Dies alles führt zu starken Einschränkungen bzgl. der Repräsentativität und Belastbarkeit der Befunde auf der Basis der Daten der Amtlichen Statistik aus diesem Typ von Erhebungen.

- Bei dem Aspekt „Fragenprogramm“ liegt der Vorteil bei dem IAB-Betriebspanel.

2.3 Teilnahme- und Auskunftspflicht vs. Panelmortalität und Antwortverweigerungen

Die AFiD-Panel entstehen aus Datensätzen, die in Befragungen von Betrieben oder Unternehmen mit Auskunftspflicht erhoben werden; die teilnehmenden Firmen sind also gesetzlich verpflichtet, umfassend und wahrheitsgemäß zu antworten. Antwortverweigerungen in einer Befragungswelle bei der gesamten Befragung (Panelmortalität) oder bei einzelnen Merkmalen (item non-response) stellen demnach hier kein Problem dar. Hiermit verbundene Selektionsprobleme bestehen daher nicht.

Die Situation beim IAB-Betriebspanel ist grundlegend verschieden. Die Teilnahme der Betriebe ist freiwillig, sie kann beim Erstkontakt oder bei Folgebefragungen verweigert werden. Bei einzelnen Fragen, die von den befragten Personen in dem Betrieb als sensibel oder als Geschäftsgeheimnis betrachtet werden (hierzu zählt in vielen Fällen z. B. bereits der Umsatz), kommt es zu Antwortverweigerungen. Dies führt dazu, dass in manchen Fällen gar nicht versucht wird, „harte“ Indikatoren zu erheben und stattdessen auf eher „weiche“ Informationen zurückgegriffen wird. Anders als z. B. in den Kostenstrukturerhebungen der Amtlichen Statistik finden sich daher im IAB-Betriebspanel auch

²Das Analysepotenzial von Daten aus solchen „Erhebungen für besondere Zwecke“ kann erheblich gesteigert werden, wenn sie mit den Daten der AFiD-Panel verknüpft werden. Ein Beispiel hierfür ist die Studie von Wagner (2011a), für die die Daten der Verlagerungsumfrage mit dem AFiD-Panel Industrieunternehmen verknüpft wurde. Dieser neue Typus von amtlichen Firmendaten, bei dem Daten aus regelmäßigen und einmaligen Erhebungen zu spezifischen Fragestellungen miteinander verknüpft werden, wird als *AFiDextended* bezeichnet (Wagner 2010).

keine Angaben, die die Berechnung einer Renditekennziffer ermöglichen. Stattdessen findet sich eine Frage nach einer subjektiven Einschätzung der Ertragslage als „sehr gut/gut/befriedigend/ausreichend/mangelhaft“. Bei Fragen, für deren exakte Beantwortung umfangreiche Recherchen erforderlich wären, wird vielfach um ungefähre Angaben („ca. × Prozent“) oder um Schätzungen gebeten.

- Bei dem Aspekt „Teilnahme- und Auskunftspflicht“ liegt der Vorteil bei den AFiD-Panel.

2.4 Kombinierbarkeit von Firmen- und Personendaten

Zahlreiche ökonomische Fragestellungen können überzeugend nur empirisch untersucht werden, wenn neben detaillierten Informationen für die Firmen auch Informationen für die in ihnen tätigen Personen vorliegen. Datensätze, die auf der Mikroebene Informationen von Firmen und Beschäftigten kombinieren, werden als „linked employer-employee (LEE)“ Daten bezeichnet. Aus rechtlichen Gründen ist es nicht möglich, die Betriebs- und Unternehmensdaten aus Erhebungen der Amtlichen Statistik mit Daten zu den darin arbeitenden Personen zu verknüpfen.³ Die Situation ist beim IAB-Betriebspanel ganz anders – hier liegt ein LEE-Datensatz vor, der die Daten aus Erhebungen des Betriebspanels und Angaben zu allen in den Betrieben sozialversicherungspflichtig Beschäftigten enthält [vgl. zum LIAB – das Akronym steht für „linked employer-employee dataset at the IAB“ – Alda et al. (2005)].

Ein Beispiel für die Notwendigkeit eines LEE-Paneldatensatzes für eine überzeugende empirische Untersuchung ist der Zusammenhang zwischen der Exporttätigkeit einer Firma und der Entlohnung. Studien mit Firmendaten aus zahlreichen Ländern finden einen positiven Zusammenhang zwischen Durchschnittslohn und Exporten – der Lohn pro Kopf ist in Firmen mit Exporten höher als in Firmen, die ihre Produkte ausschließlich im jeweiligen Inland verkaufen. Diese „Exporteur-Lohnprämie“ bleibt bestehen, wenn in Regressionsschätzungen für beobachtbare Firmeneigenschaften (wie Größe oder Industriezweig) und in Modellen mit Paneldaten und fixen Firmeneffekten zusätzlich für unbeobachtete zeitinvariante Firmeneigenschaften kontrolliert wird. Erst LEE-Paneldaten wie die LIAB-Daten ermöglichen es zu überprüfen, ob sich diese Lohnprämien auch feststellen lassen, wenn für beobachtete und unbeobachtete Eigenschaften der Beschäftigten kontrolliert wird – ob also

„gleiche“ Individuen in exportierenden Firmen höhere Löhne erhalten als in „sonst gleichen“ nicht-exportierenden Firmen. Mit den LIAB-Daten finden Schank et al. (2007), dass dies in Deutschland tatsächlich der Fall ist. In einer weiteren Studie mit den LIAB-Daten finden die Autoren, dass diese Lohnprämie in exportierenden Betrieben bereits in den Jahren vor dem Beginn der Exporttätigkeit gezahlt wird und dass sie in den Jahren nach dem Exportstart nicht ansteigt. Die höheren Löhne in exportierenden Betrieben sind also eine Folge der Selbstselektion von produktiveren, besser zahlenden Firmen auf die Exportmärkte und nicht eine Folge der Exporttätigkeit (vgl. Schank et al. 2010).

- Bei dem Aspekt „Kombinierbarkeit von Firmen- und Personendaten“ liegt der Vorteil beim IAB-Betriebspanel.

2.5 Kombinierbarkeit mit Firmendaten aus externen Quellen

Auch bei einem breit angelegten Fragenprogramm wie dem des IAB-Betriebspanels – und erst recht bei einem eher schmalen Spektrum der erhobenen Informationen, wie es bei den Umfragen der Amtlichen Statistik die Regel ist – gibt es immer wieder spezifische Informationen, die in den verfügbaren Firmenpanel-Datensätzen nicht enthalten sind. Aus Kosten- und Zeitgründen ist es nicht möglich, für die Analyse einer spezifischen Fragestellung ein eigenes Panel zu erheben. Eine themenspezifische Ergänzung des Fragenprogramms des IAB-Betriebspanels über mehrere Wellen hinweg oder eine „Erhebungen für besondere Zwecke“ durch das Statistische Bundesamt mit mehreren Wellen ist ebenfalls kein gangbarer Weg. Was tun?

Ein in vielen Fällen erfolgversprechender Weg ist ein „Andocken“ von Informationen aus externen Quellen an vorhandene Firmenpanel-Daten. *Technisch* erfordert dies, dass in den Firmenpaneldaten und in den Daten aus der externen Informationsquelle ein für jede Firma identischer Identifikator vorhanden ist, der diese Zusammenführung ermöglicht. Hierbei kann es sich z. B. um die Umsatzsteuer-Nummer eines Unternehmens oder um die Nummer des Handelsregister-Eintrags zusammen mit der Angabe über das zuständige Registergericht handeln. Möglich – wenn auch in vielen Fällen sehr viel komplizierter und mit einem hohen zeitlichen Aufwand verbunden – ist auch eine Zusammenführung anhand des Firmennamens und der Firmenanschrift.⁴ *Rechtlich* muss diese Datenzusammenführung zulässig sein.

³Eine Ausnahme ist die schon als kombinierte Erhebung von Firmen- und Personeninformationen konzipierte Verdienststrukturerhebung (bzw. früher: Gehalts- und Lohnstrukturerhebung; vgl. Hafner und Lenz (2008)), die allerdings nur Angaben zu einer Stichprobe der tätigen Personen enthält. Die verschiedenen Wellen dieser Erhebung sind ferner kein Panel.

⁴Hierbei kann einerseits nach sorgfältiger Aufbereitung der Angaben (etwa einer Vereinheitlichung der Schreibweise von Umlauten, Abkürzungen etc.) Software für ein (fehlertolerantes) Matching von Strings eingesetzt werden, andererseits ist ein Umweg über die Ermittlung und Zufügung von Handelsregister-Informationen (siehe www.handelsregister.de) möglich.

Für die AFiD-Panel sind beide Bedingungen in viele Fällen erfüllt. *Rechtlich* regelt dies das BstatG in § 13a. Hier wird u. a. ausdrücklich gestattet, Firmendaten aus Erhebungen der amtlichen Statistik mit Daten aus „allgemein zugänglichen Quellen“ zusammenzuführen. „Allgemein zugängliche Quellen“ sind hierbei nicht nur kostenlos im Internet verfügbare oder in anderer Form publizierte Daten sondern auch Informationen aus Datenbanken kommerzieller Anbieter, deren Nutzung kostenpflichtig ist.⁵ *Technisch* ist dies dann ohne großen Aufwand möglich, wenn die externen Daten einen der Identifikatoren enthalten, der auch im Unternehmensregister-System (URS) enthalten ist; hierzu zählen z.B. die Umsatzsteuernummer und die Handelsregisternummer zusammen mit der Angabe des Handelsregistergerichts. Kombinierte Daten aus Erhebungen der Amtlichen Statistik und aus externen Datenquellen werden als *AFIDplus*-Daten bezeichnet (Wagner 2010).

Ein Beispiel für die Nutzung solcher kombinierter Daten sind Studien zur Verbreitung und zu den Folgen der Drittelmitbestimmung im Aufsichtsrat in Gesellschaften mit beschränkter Haftung (GmbH) mit 500 bis 2000 Beschäftigten. Diese GmbHs fallen (vernachlässigt man Details und Ausnahmen) unter das Drittelmitbestimmungsgesetz, wonach ein Drittel der Sitze im Aufsichtsrat von Arbeitnehmern eingenommen wird. Informationen über die tatsächliche Umsetzung dieses Gesetzes (Gibt es in den GmbHs einen Aufsichtsrat? Sind ein Drittel der Sitze mit Arbeitnehmern besetzt?) sind aus keiner Erhebung der Amtlichen Statistik verfügbar. In einer aufwändigen telefonischen Erhebung und unter Verwendung von Angaben aus der kommerziellen *Hoppenstedt Datenbank Großunternehmen* wurden diese Angaben für GmbHs der genannten Größenklasse für 2007/2008 aufbereitet und die neu erhobenen Daten wurden dann in Arbeitspapieren publiziert. Über die Angabe der Handelsregisternummer und des Handelsregistergerichts wurden diese allgemein zugänglichen Informationen dann unter Verwendung des Unternehmensregister-Systems mit Unternehmensdaten der Amtlichen Statistik aus den Kostenstrukturerhebungen im Verarbeitenden Gewerbe bzw. der Strukturerhebung im Dienstleistungssektor verknüpft. Mit diesen kombinierten Daten konnten erstmals durch einen direkten Vergleich von mitbestimmten und nicht mitbestimmten Unternehmen Untersuchungen zur Verbreitung und zu den ökonomischen Folgen der Drittelmitbestimmung im Aufsichtsrat durchgeführt werden (vgl. Boneberg

2011 und Wagner 2011b). Dies zeigt exemplarisch den hohen Zusatznutzen einer Kombination von AFiD-Panel und Firmendaten aus externen Quellen.

Ein weiteres Beispiel ist ein im Sommer 2012 begonnenes eigenes Projekt, in dem Angaben zur Kreditwürdigkeit von Unternehmen aus den Datenbeständen des Verbands der Vereine Creditreform (VVC) mit dem AFID-Panel Industrieunternehmen verknüpft wurden. Für diese Verknüpfung wurde die sowohl in der VVC-Datenbank als auch im Unternehmensregister verfügbare Kennnummer aus der Hoppenstedt-Datenbank genutzt. Diese so verknüpften Daten ermöglichen z. B. empirische Untersuchungen zu Zusammenhängen von Exporten und Finanzierungsbeschränkungen für deutsche Industrieunternehmen (vgl. Wagner 2012).

Eine Verknüpfung mit externen Informationen ist auch für die Daten aus dem IAB-Betriebspanel möglich. Ein Beispiel hierfür ist die Studie von Bellmann und Kirchhof (2006), in der Angaben zu Unternehmenszusammenschlüssen oder Unternehmensübernahmen aus der Datenbank *Thomsen One Banker* mit dem IAB-Betriebspanel verknüpft wurden. Die Identifikation der Firmen in den beiden Datenbeständen war hierbei allerdings nur über einen Abgleich der Firmennamen und der Adressen möglich. Erfahrungsgemäß ist ein solches Vorgehen sehr zeitaufwändig und damit sehr teuer. Die Studie von Bellmann und Kirchhof (2006) belegt aber die Machbarkeit.

- Bei dem Aspekt „Kombinierbarkeit mit Firmendaten aus externen Quellen“ ist kein eindeutiger Vorteil für eines der beiden hier betrachteten Panels erkennbar.

2.6 Internationale Vergleichbarkeit der Daten

Die empirische Wirtschaftsforschung nutzt Firmenpaneldaten für unterschiedliche Zwecke. Hierzu zählen die Aufdeckung bisher nicht bekannter Fakten, Regelmäßigkeiten und Verhaltensmuster, der Test von Hypothesen aus theoretischen Überlegungen und formalen Modellen sowie die Evaluation wirtschaftspolitischer Maßnahmen. Die Ergebnisse dieser empirischen Forschung liefern dann Beiträge zu einem besseren Verständnis realer wirtschaftlicher Abläufe, sie können Hinweise für eine realitätsorientierte theoretische Modellierung ökonomischer Zusammenhänge liefern und sie können die Grundlage für eine evidenzbasierte Politikberatung bilden.

In vielen Fällen ist es dabei sehr hilfreich, wenn die empirischen Ergebnisse nicht ausschließlich auf Daten aus einem Land beruhen. Zeigen sich Zusammenhänge in identischer oder sehr ähnlicher Form in vielen Ländern, dann können diese Regelmäßigkeiten als stilisierte Fakten gelten und eine breite Gültigkeit beanspruchen. Zeigen sich in Studien mit identisch spezifizierten empirischen Modellen und

⁵Das Statistische Bundesamt nutzt diese Möglichkeit selbst durch den Ankauf von Informationen darüber, ob es sich bei einem Unternehmen um ein unabhängiges Unternehmen, ein Gruppenoberhaupt, eine gruppenabhängige Einheit mit deutschem Gruppenoberhaupt oder um eine Einheit mit ausländischem Gruppenoberhaupt handelt. Diese Informationen sind u. a. die Basis für die Statistik über auslandskontrollierte Unternehmen (Inward-FATS); siehe hierzu Weche Gelübcke (2011).

bei Verwendung identisch definierter Variablen Unterschiede zwischen Ländern, dann führt dies zu der Frage, warum es eben diese Unterschiede gibt. Entsprechende Untersuchungen können dann sehr viel zum besseren Verständnis ökonomischer Zusammenhänge beitragen.

Die Daten der AFiD-Panel liegen in sehr vielen Fällen in identischer Form (Definition der erhobenen Merkmale, Abgrenzung der Grundgesamtheit der Erhebung) für zahlreiche Länder vor. Dies ermöglicht international vergleichende Studien auf der Basis identisch spezifizierter empirischer Modelle mit identisch definierten Variablen. Der Zugang zu den Firmendaten in den einzelnen Ländern ist dabei in der Regel nur Forschern aus den jeweiligen Ländern möglich, was eine ex-ante Koordination entsprechender empirischer Studien erfordert.

Ein Beispiel für eine solche international vergleichende Studie auf der Basis von Daten des Typs der AFiD-Panel aus 14 Ländern ist die Untersuchung der Produktivitätsunterschiede zwischen exportierenden und nicht-exportierenden Industriefirmen, die von der International Study Group on Exports and Productivity (ISGEP) (2008) durchgeführt wurde. Auf der Basis vergleichbarer Firmendaten und identisch spezifizierter empirischer Modelle finden die Autoren, dass exportierende Firmen auch bei Kontrolle beobachteter und unbeobachteter Firmeneigenschaften produktiver sind als nicht exportierende Firmen, dass es eine Selbstselektion produktiverer Firmen auf die Exportmärkte gibt und dass sich kaum Hinweise auf positive Produktivitätseffekte der Exporte nachweisen lassen. Dies ist das aus zahlreichen unabhängig voneinander durchgeführten empirischen Untersuchungen aus vielen Ländern vertraute Bild (vgl. Wagner 2007).

Darüber hinaus können die Autoren aber zeigen, dass sich die Produktivitätsunterschiede zwischen Exporteuren und Nicht-Exporteuren zwischen den Ländern in der Höhe deutlich voneinander unterscheiden – und zwar ohne dass diese Unterschiede durch unterschiedlich definierte Variablen, unterschiedlich abgegrenzte Stichproben oder unterschiedlich spezifizierte empirische Modelle verursacht wurden. In einer Meta-Analyse finden die Verfasser, dass die Produktivitätsunterschiede – wie theoretisch erwartet – höher sind in Ländern mit einer geringeren Partizipationsrate der Firmen am Export, mit einer restriktiveren Handelspolitik, mit geringerem Pro-Kopf Einkommen, mit weniger effizienten Regierungen und schlechterer Regulierungsqualität und in Ländern mit einem höheren Anteil der Exporte zu weiter entfernten Märkten. Damit liefert diese international vergleichende Studie einen wichtigen Beitrag zum Verständnis eines zentralen Themas aus der Literatur zur mikroökonomischen Analyse internationaler Firmentätigkeit.

Die internationale Vergleichbarkeit von Daten aus Erhebungen der Amtlichen Statistik ermöglicht international vergleichende Studien auf vergleichbarer Basis. Damit ist es

möglich, Ergebnisse von Studien mit Daten aus verschiedenen Ländern zu replizieren und aus Unterschieden und Gemeinsamkeiten der Ergebnisse zu lernen. Solche Replikationsstudien stellen einen sehr wertvollen Beitrag zum wissenschaftlichen Fortschritt dar – auch wenn dies nicht immer von allen am Wissenschaftsprozess beteiligten Akteuren (Autoren, Gutachtern und Herausgebern von Zeitschriften) gesehen wird.⁶

Beim IAB-Betriebspanel ist diese internationale Vergleichbarkeit im Hinblick auf die Definition von erhobenen Merkmalen der Betriebe und von Einschätzungen und Erwartungen der Entscheidungsträger nicht in diesem Umfang gegeben. Damit lassen sich mit diesen Daten Replikationsstudien oder international vergleichende Untersuchungen nicht im gleichen Umfang durchführen. In vielen Fällen sind aber Firmenpanelerhebungen in verschiedenen Ländern in Teilbereichen der erhobenen Informationen hinreichend ähnlich, um eine tragfähige Basis für international vergleichende Studien zu bieten. Im Bereich der Mitbestimmungsforschung z. B. existieren in Frankreich und Großbritannien mit RESPONSE und WERS zwei für internationale Vergleiche durchaus geeignete Datensätze. Ferner besteht grundsätzlich die Möglichkeit, einzelne Module des IAB-Betriebspanels mit dem Ziel international vergleichender Studien abzustimmen.

- Bei dem Aspekt „Internationale Vergleichbarkeit“ liegt der Vorteil bedingt durch die Standardisierung des Fragenprogramms bei den AFiD-Panel.

3 Schlussfolgerungen

Die Diskussion von sechs Aspekten für einen Vergleich von IAB-Betriebspanel und AFiD-Panel hat deutlich gemacht, dass es sich bei diesen beiden Datensätzen um komplementäre und nicht um substitutive Inputs für die Empirische Wirtschaftsforschung handelt. Beide Datensätze haben spezifische Stärken und Schwächen, die sie zur Analyse von konkreten Fragestellungen jeweils mehr oder weniger gut geeignet oder auch völlig ungeeignet machen. Daraus folgt, dass der Werkzeugkasten für die empirischen Forschung mit Längsschnittdaten für Firmen in Deutschland sowohl das IAB-Betriebspanel als auch die AFiD-Panel enthalten muss, denn nur so ist sicher gestellt, dass ein möglichst breites Spektrum von Fragestellungen mit einem jeweils adäquaten Datensatz untersucht werden kann.⁷

⁶Vgl. Wagner (2011c, p. 406ff.) für eine ausführliche Diskussion dieses Punktes.

⁷Diese Schlussfolgerung ist übrigens identisch mit der aus einem Vergleich von zwei „Vorgänger-Projekten“ des IAB-Betriebspanels und der AFiD-Panel. Gerlach und Wagner (1995) vergleichen Paneldaten für niedersächsische Industriebetriebe aus der amtlichen Statistik (be-

Der Befund, dass beide Firmenpanel Stärken und Schwächen aufweisen, legt die Frage nahe, ob sich die jeweiligen Stärken noch verstärken und die jeweiligen Schwächen weiter abschwächen lassen, wenn es gelingen könnte, beide Panel zu kombinieren. Eine solche „Fusion“ steht vor so hohen Hürden, dass es müßig erscheint, hierüber weiter zu spekulieren:

Erstens ist die Beobachtungseinheit in beiden Firmenpanels nicht (immer) identisch. Bei dem IAB-Betriebspanel sind es lokale Produktionseinheiten (Betriebe), bei den AFiD-Panel in sehr vielen Fällen kleinste rechtlich selbständige Einheiten (Unternehmen). Beide Firmentypen sind nur im Fall von Einbetriebsunternehmen identisch – und eine Aggregation der Betriebsangaben auf die Unternehmensebene im Fall von Mehrbetriebsunternehmen ist im IAB-Betriebspanel technisch nicht möglich, da nicht alle Betriebe der Mehrbetriebsunternehmen befragt werden.⁸

Zweitens werden beide Datensätze von unterschiedlichen Datenproduzenten bereit gestellt, und die Zusammenführung von Einzeldaten für Firmen über die Grenzen von Datenproduzenten hinweg erfordert in Deutschland (zumindest gilt dies heute, d.h. im Jahr 2012) eine schriftliche Zustimmung der Firmen. Diese Zustimmung zu erhalten ist nicht unmöglich,⁹ aber dies ist mit einem so hohen Aufwand verbunden, dass im Fall einer Kombination von IAB-Betriebspanel und Betriebsdaten aus Erhebungen der Amtlichen Statistik nach meiner Einschätzung zu erwarten ist, dass die Kosten den erwarteten Zusatznutzen übersteigen.

Wir werden also als empirische Wirtschaftsforscher weiter mit den Stärken und Schwächen von IAB-Betriebspanel und AFiD-Panel leben – nicht *müssen*, sondern *dürfen*, und dies hoffentlich noch für lange Zeit!

Danksagung Ich danke Lutz Bellmann für sehr hilfreiche Anmerkungen zu früheren Fassungen.

Literatur

Addison, J.T., Bellmann, L., Schnabel, C., Wagner, J.: German works councils old and new: incidence, coverage and determinants. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **123**(3), 339–358 (2003)

geschrieben in Wagner 2000) und Daten aus dem Hannoveraner Firmenpanel, einer Befragung einer Stichprobe von niedersächsischen Industriebetrieben, die in vieler Hinsicht dem IAB-Betriebspanel vergleichbar ist (vgl. Gerlach et al. 2003).

⁸Die Situation ist hier im Betriebs-Historik-Panel (Spengler 2008) anders; hier sind alle Betriebsstätten eines Mehrbetriebsunternehmens enthalten (soweit darin mindestens eine Person sozialversicherungs-pflichtig beschäftigt ist).

⁹Vgl. das Projekt KombiFiD – Kombinierte Firmendaten für Deutschland, in dem Unternehmensdaten aus den AFiD-Panel der Amtlichen Statistik mit zu Unternehmensdaten aggregierten Angaben aus dem Betriebs-Historik-Panel des IAB und Daten aus den MiDi- und Ustan-Datenbanken der Deutschen Bundesbank verknüpft wurden (vgl. www.kombifid.de).

- Alda, H., Bender, S., Gartner, H.: The linked employer-employee dataset created from the IAB-establishment panel and the process-produced data of the IAB (LIAB). *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **125**(2), 327–336 (2005)
- Bellmann, L., Kirchhof, K.: Akquisitionen und Unternehmenszusammenschlüsse im IAB-Betriebspanel. In: Bellmann, L., Wagner, J. (Hrsg.) *Betriebsdemographie. Beiträge zur Arbeitsmarkt- und Berufsforschung*, Bd. 305, S. 191–203. Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit, Nürnberg (2006)
- Boneberg, F.: One-third co-determination at company supervisory boards and its economic consequences. New evidence for employment. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **131**(1), 107–131 (2011)
- Fischer, G., Janik, F., Müller, D., Schmucker, A.: The IAB establishment panel—things users should know. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **129**(1), 133–148 (2009)
- Fritsch, M., Görzig, B., Hennchen, O., Stephan, A.: Cost structure surveys for Germany. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **124**(4), 557–566 (2004)
- Gerlach, K., Wagner, J.: Die Heterogenität der Arbeitsplatzdynamik innerhalb der Industrie. Zum Verhältnis von Belegschafts- und betriebsfluktuation im Verarbeitenden Gewerbe Niedersachsens (1978–1990). In: Semlinger, K., Frick B. (Hrsg.) *Betriebliche Modernisierung in personeller Erneuerung. Personalentwicklung, Personalaustausch und Betriebliche Fluktuation*, S. 39–57. Sigma, Berlin (1995)
- Gerlach, K., Hübler, O., Meyer, W.: The Hannover firm panel (HFP). *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **123**(3), 463–470 (2003)
- Hafner, H.-P., Lenz, R.: The German structure of earnings survey: methodology, data access and research potential. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **128**(3), 489–500 (2008)
- International Study Group on Exports and Productivity (ISGEP): Understanding cross-country differences in exporter premia: comparable evidence for 14 countries. *Rev. World Econ.* **144**(4), 596–635 (2008)
- Jirjahn, U.: Ökonomische Wirkungen der Mitbestimmung in Deutschland. Ein Update. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **131**(1), 3–57 (2011)
- Kohlmann, A.: The research data centre of the federal employment service in the institute for employment research. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **125**(3), 437–447 (2005)
- Konold, M.: New possibilities for economic research through integration of establishment-level panel data of German official statistics. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **127**(2), 321–334 (2007)
- Lipponer, A.: The Deutsche Bundesbank's FDI micro database. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **123**(4), 593–600 (2003)
- Malchin, A., Voshage, R.: Official firm data for Germany. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **129**(3), 501–513 (2009)
- Schank, T., Schnabel, C., Wagner, J.: Works councils—sand or grease in the operation of German firms? *Appl. Econ. Lett.* **11**(3), 159–161 (2004)
- Schank, T., Schnabel, C., Wagner, J.: Do exporters really pay higher wages? First evidence from German linked employer-employee data. *J. Int. Econ.* **72**(1), 52–74 (2007)
- Schank, T., Schnabel, C., Wagner, J.: Higher wages in exporting firms: self-selection, export effect, or both? First evidence from linked employer-employee data. *Rev. World Econ.* **146**(2), 303–322 (2010)
- Söllner, R.: Der Zugang kleiner und mittlerer Unternehmen zu Finanzmitteln. *Wirtschaft. Stat.* **7**, 619–628 (2011)
- Spengler, A.: The establishment history panel. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **128**(3), 501–509 (2008)
- Stegmeier, J.: Effects of works councils on firm-provided further training in Germany. *Br. J. Ind. Relat.* **50**(4), 667–689 (2012)

- Stöss, E.: Deutsche Bundesbank's corporate balance sheet statistics and areas of application. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **121**(1), 131–137 (2001)
- Vogel, A.: The German business services statistics panel 2003 to 2007. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **129**(3), 515–522 (2009)
- Wagner, J.: Arbeitsplatzdynamik, Technologieintensität und Betriebsgrößenstruktur in der Hannoverschen Industrie 1978 bis 1994. Beiträge zur Regionalen Entwicklung, Bd. 50. Kommunalverband Großraum Hannover, Hannover (1995)
- Wagner, J.: Firm panel data from German official statistics. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **120**(1), 143–150 (2000)
- Wagner, J.: Exports and productivity: a survey of the evidence from firm-level data. *World Econ.* **30**(1), 60–82 (2007)
- Wagner, J.: The research potential of new types of enterprise data based on surveys from official statistics. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **130**(1), 133–142 (2010)
- Wagner, J.: Offshoring and firm performance: self-selection, effects on performance, or both? *Rev. World Econ.* **147**(2), 217–247 (2011a)
- Wagner, J.: One-third codetermination at the company supervisory boards and firm performance in German manufacturing industries: first direct evidence from a new type of enterprise data. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **131**(1), 91–106 (2011b)
- Wagner, J.: From estimation results to stylized facts. Twelve recommendations for empirical research in international activities of heterogeneous firms. *Economist* **159**(4), 389–412 (2011c)
- Wagner, J.: Credit constraints and exports: evidence for German manufacturing enterprises. University of Lüneburg Working Paper Series in Economics, Bd. 251, October (2012)
- Weche Gelübcke, J.P.: Ownership patterns and enterprise groups in German structural business statistics. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **131**(4), 635–647 (2011)
- Zühlke, S., Zwick, M., Scharnhorst, S., Wende, T.: The research data centres of the Federal Statistical Office and the statistical offices of the Länder. *Schmollers Jahrb. Wirtschaft.- Soz.wiss.* **124**(4), 567–578 (2004)
- Zwania, J.: Verlagerung wirtschaftlicher Aktivitäten. Theoretischer Hintergrund zur Erhebung. *Wirtschaft. Stat.* **6**, 477–482 (2008)

Prof. Dr. Joachim Wagner ist Professor für Volkswirtschaftslehre/Empirische Wirtschaftsforschung an der Leuphana Universität Lüneburg. Er ist Research Fellow des IZA und Mitherausgeber von *Schmollers Jahrbuch/Journal of Applied Social Science Studies*, *Jahrbücher für Nationalökonomie und Statistik/Journal of Economics and Statistics* und *Journal for Labour Market Research/Zeitschrift für Arbeitsmarktforschung*. Er ist Mitglied mit Rat für Sozial- und Wirtschaftsdaten (RatSWD). Joachim Wagner hat eine große Anzahl Aufsätze in nationalen und internationalen referierten Journals publiziert; Schwerpunkte seiner Forschungen sind Bestimmungsgründe und Folgen internationaler Firmentätigkeit sowie Untersuchungen zur Arbeitsmarktökonomik und zum System der industriellen Beziehungen. Die meisten seiner Studien nutzen Firmenpaneldaten für mikroökonomische Untersuchungen.