
Evaluation von Maßnahmen aktiver Arbeitsmarktpolitik in der Bundesagentur für Arbeit

Susanne Rässler

*Kompetenzzentrum Empirische Methoden, IAB
Produkt- und Programmanalyse, Bundesagentur für Arbeit
Professur für Computational Statistics, Frankfurt School of Finance & Management*

„Reformbilanz der aktiven Arbeitsmarktpolitik“, Bonn, 21.03.2007

Gliederung

- **Einleitung**
- **Evaluationsprobleme und ihre klassische Lösung**
- **Wirkungsanalyse in der Bundesagentur für Arbeit**
- **Operationalisierungen und Daten**
- **Beispielhafte empirische Auswertungen**
- **Fazit**

Gliederung

- **Einleitung**
- Evaluationsprobleme und ihre klassische Lösung
- Wirkungsanalyse in der Bundesagentur für Arbeit
- Operationalisierungen und Daten
- Beispielhafte empirische Auswertungen
- Fazit

Einleitung

„Angesichtes immenser Ausgaben für die aktive Arbeitsmarktpolitik ist ihre wissenschaftlich fundierte und aussagekräftige Evaluation seit längerer Zeit mehr als überfällig.“ (Wolfgang Franz, 2005, AStA 88, 155)

- **Aufwendungen Bundesagentur und Bund im Jahr 2006 für**
 - Leistungen der aktiven Arbeitsförderung für SGBIII: 11,1 Mrd. Euro
 - Arbeitslosengeld I: 23,0 Mrd. Euro
 - Arbeitslosengeld II und Sozialgeld: 26,4 Mrd. Euro

- **Wie das Gesetz (§282 SGB III, §55 SGB II) es befiehlt: Ist aktive Arbeitsmarktpolitik effektiv?**
 - Verbessert sie die Vermittlungsaussichten?
 - Erhöht sie die Beschäftigungsfähigkeit?
 - Wie hoch sind die volkswirtschaftlichen Nettoeffekte?

Gliederung

- Einleitung
- **Evaluationsprobleme und ihre klassische Lösung**
- Wirkungsanalyse in der Bundesagentur für Arbeit
- Operationalisierungen und Daten
- Beispielhafte empirische Auswertungen
- Fazit

Evaluation: Wenn das so einfach wäre...

- **Enorme Heterogenitäten:** Teilnehmer, Maßnahmen, Träger, Regionen
- **Kausalität:** Vielzahl parallel durchgeführter gesetzlicher Änderungen
- **Selektionsprozesse:** Selbst- und Fremdselektion
- **Endogenität:** Arbeitsmarktlage und Arbeitsmarktpolitik
- **Datenverfügbarkeit und Datenqualität:** Aufbau geeigneter Datenbasen nötig

Klassisches Vorgehen

- Sei $Y(0)$ das Ergebnis bei Nichtteilnahme, $Y(1)$ bei Teilnahme
- **Rubins Kausalmodell der potentiellen Ergebnisse:** Individueller kausaler Teilnahme-Effekt ergibt sich aus Differenz $\tau_i = Y_i(1) - Y_i(0)$
- **Matching:** „Das Arbeitspferd der Evaluationsforschung“
- **Teilnehmer- und Kontrollgruppe:** Suche zu jedem Teilnehmer einer Maßnahme einen in Charakteristika X identischen oder möglichst ähnlichen Nichtteilnehmer
- **Üblich:** Schätzung des durchschnittlichen Maßnahmeneffekts für die Maßnahmenteilnehmer (ATET = Average Treatment Effect on the Treated)
- Exaktes \sim , Propensity Score \sim , Mahalanobis-Distanz Matching

Gliederung

- Einleitung
- Evaluationsprobleme und ihre klassische Lösung
- **Wirkungsanalyse in der Bundesagentur für Arbeit**
- Operationalisierungen und Daten
- Beispielhafte empirische Auswertungen
- Fazit

Wirkungsanalyse in der Bundesagentur für Arbeit

Das Projekt „**TrEffeR**“ (**T**reatment **E**ffects and **P**Rediction): Wirkungsanalyse zur Unterstützung der operativen Steuerung in der Bundesagentur für Arbeit (BA)

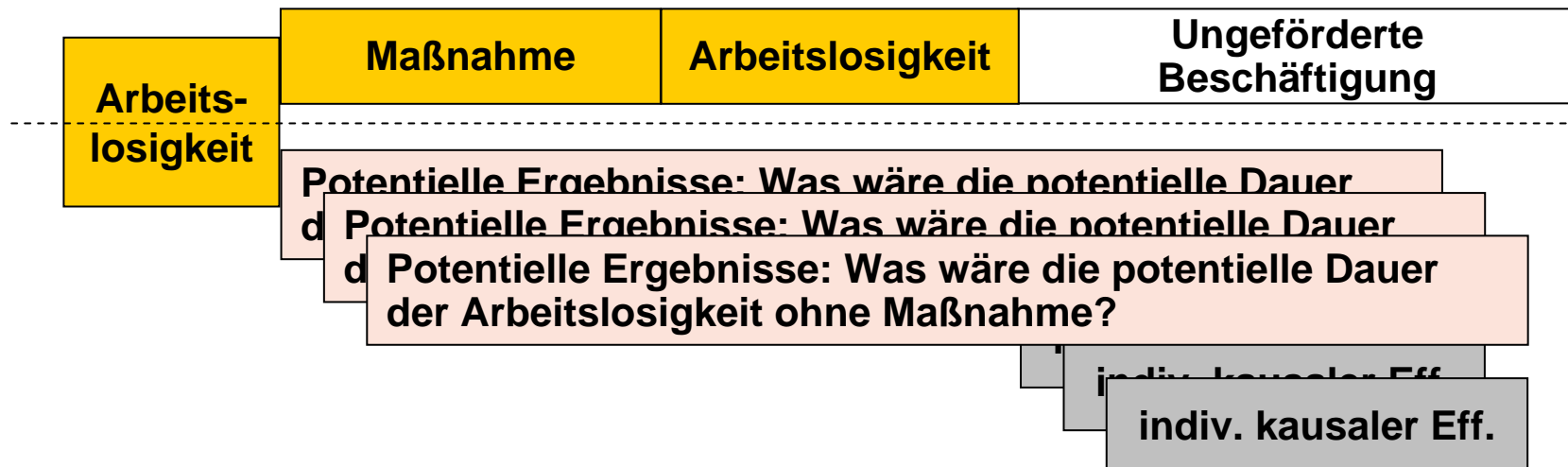
- Kooperation von BA, IAB und der amerikanischen Universität Harvard
- Leitung: Rässler / Rubin
- **Ziel 1:** Möglichst zeitnahe und umfassende retrospektive Wirkungsanalyse / Evaluation aller Instrumente der aktiven Arbeitsmarktpolitik
- **Ziel 2:** Entwicklung eines Targeting-Tools zur Unterstützung der Vermittler

Problematik durchschnittlicher Fördereffekte

- **Matching-Diskrepanzen:** Teilnehmer und Kontrollpersonen sind in ihren X Variablen im Durchschnitt weitestgehend „ausgeglichen“, aber nicht notwendigerweise die individuellen Paare
- **Problematik der Matching-Diskrepanzen:** Aussagen lediglich für die jeweiligen Gruppen statistisch valide
- Ergebnisse müssen auf **disaggregierter** Ebene verfügbar sein für Aggregationen beliebiger Art, z.B. Analysen für Jugendliche, Personen mit Migrationshintergrund, Geringqualifizierte, Ältere,...
- ⇒ **Schätzung individueller Maßnahmeeffekte** unbedingt nötig!

Idee: Potentielle Ergebnisse sind „Missing Data“

- Schätzung individueller kausaler Effekte



- Schätzung von $Y_i(0) | T_i=1, X=x$, der Dauer der Arbeitslosigkeit einer Person i ohne Maßnahme bei gegebenen X Variablen
- Berücksichtigung von Kosten geplant

Die Lösung des Statistikers

- Schätzung **individueller kausaler Effekte** in zwei Schritten:
- **Schritt 1:** Matching zur Selektionskontrolle
- **Schritt 2:** Parametrischer Bayes-Ansatz zur Schätzung der potentiellen Ergebnisse
- Mittels statistischer Simulation können auch Bandbreiten/Varianzen geschätzt werden
- King et al. (2000). Making the Most of Statistical Analyses: Improving Interpretation and Presentation, *American Journal of Political Sciences*, 44, 341-355.
- Rubin, D.B. (2007). The Design versus the Analysis of Observational Studies for Causal Effects: Parallels with the Design of Randomized Trials, *Statistics in Medicine*, forthcoming.

Schritt 1: Matching

- Matching in hoch differenzierten Untergruppen
- **Schätzung der Teilnahmewahrscheinlichkeiten:** der Propensity Scores
 $p(X) = P(T=1 | X=x)$
- Innerhalb der Gruppen Matching nach dem Verfahren des Nächsten Nachbarn mit caliper $0,25 \cdot s(p_T)$ und mit Zurücklegen
- Mahalanobis-Distanz Matching auf
 - die linearen propensity scores
 - die zu den linearen propensity scores orthogonalisierten erklärenden Variablen X
- Verwendung des Ergebnisses einer Gruppe nur, wenn
 - mindestens 50% der Teilnehmer und
 - mindestens 20 Teilnehmer absolut gematcht werden konnten.

Schritt 2: Parametrisches Modell

- Erneutes Schätzen der Propensity Scores von TN und Kontrollpersonen (KP)
- **Klassifikation:** Einteilen der KP anhand der neuen Propensity Score Verteilung in Subgruppen
- **Modellbasierte Ergänzung** der potentiellen Ergebnisse über geeignetes Regressionsmodell innerhalb jeder Klasse der KP:

$$Y_C(0) = X_C\beta + \varepsilon, \quad \varepsilon \sim N(0; \sigma^2)$$

- Bestimmung von $\hat{\beta}$ und $\hat{\sigma}$ mittels OLS
- Innerhalb der zugehörigen Klassen der TN Simulation (250 mal) von $Y_T(0)$ durch

$$1) \quad \beta \sim N(\hat{\beta}; \hat{\sigma}^2 (X_C' X_C)^{-1})$$

$$2) \quad \hat{Y}_T(0) = X_T \beta$$

Ergebnis: Nulldatensatz

- Bildung eines Nulldatensatzes ausgehend von den zuvor ermittelten Matching-Ergebnissen
- **Nulldatensatz** enthält für jeden Maßnahmeteilnehmer i
 - sein beobachtetes Ergebnis bei Maßnahmeteilnahme $Y_i(1)$
 - sein geschätztes potentielltes Ergebnis bei Nichtteilnahme $\hat{Y}_i(0)$
- **Verteilung der individuellen kausalen Effekte** der Maßnahmeteilnehmer:

$$\tau_{ik} = Y_i(1) - \hat{Y}_{ik}(0) \text{ für } k = 1, \dots, 250$$

Vorteile dieses Ansatzes

- Bereitstellung von individuellen Fördereffekten für die Maßnahmeteilnehmer
- Ermöglicht Auswertungen auf beliebiger Ebene
- Ausschließlich Verwendung von Informationen von gematchten Nichtteilnehmern, die dem jeweiligen Teilnehmer bzgl. der Propensity Scores ähnlich sind.
- **Wenig Annahmen** bzgl. der funktionalen Form nötig, da das Imputationsmodell für jede Klasse gesondert gerechnet wird.
- **Berücksichtigung der Unsicherheit** durch 250malige Wiederholung und Simulation einer Verteilung von $Y_T(0)$

Gliederung

- Einleitung
- Evaluationsprobleme und ihre klassische Lösung
- Wirkungsanalyse in der Bundesagentur für Arbeit
- **Operationalisierungen und Daten**
- Beispielhafte empirische Auswertungen
- Fazit

Operationalisierungen

- **Geförderte:** Maßnahmeteilnahmen 2001 bis 2005
- **Maßnahmen:** Hochdifferenziert bis auf Maßnahmennummer möglich
- **X Variablen:**
 - **Aktuelle sozio-ökonomische Merkmale**
wie Alter, Nationalität, Geschlecht, Schulbildung, Berufsausbildung, Wunsch BKZ (einstellig), ...
 - **18-Monate-Historie**
mit Informationen zu vergangenen Arbeitslosigkeitsperioden, Maßnahmeteilnahmen, Leistungsbezügen, Sperrzeiten, ...
 - **regionale Arbeitsmarktsituation**
mit Informationen zu Arbeitslosen, offenen Stellen und Einstellungen für die 3stellige BKZ

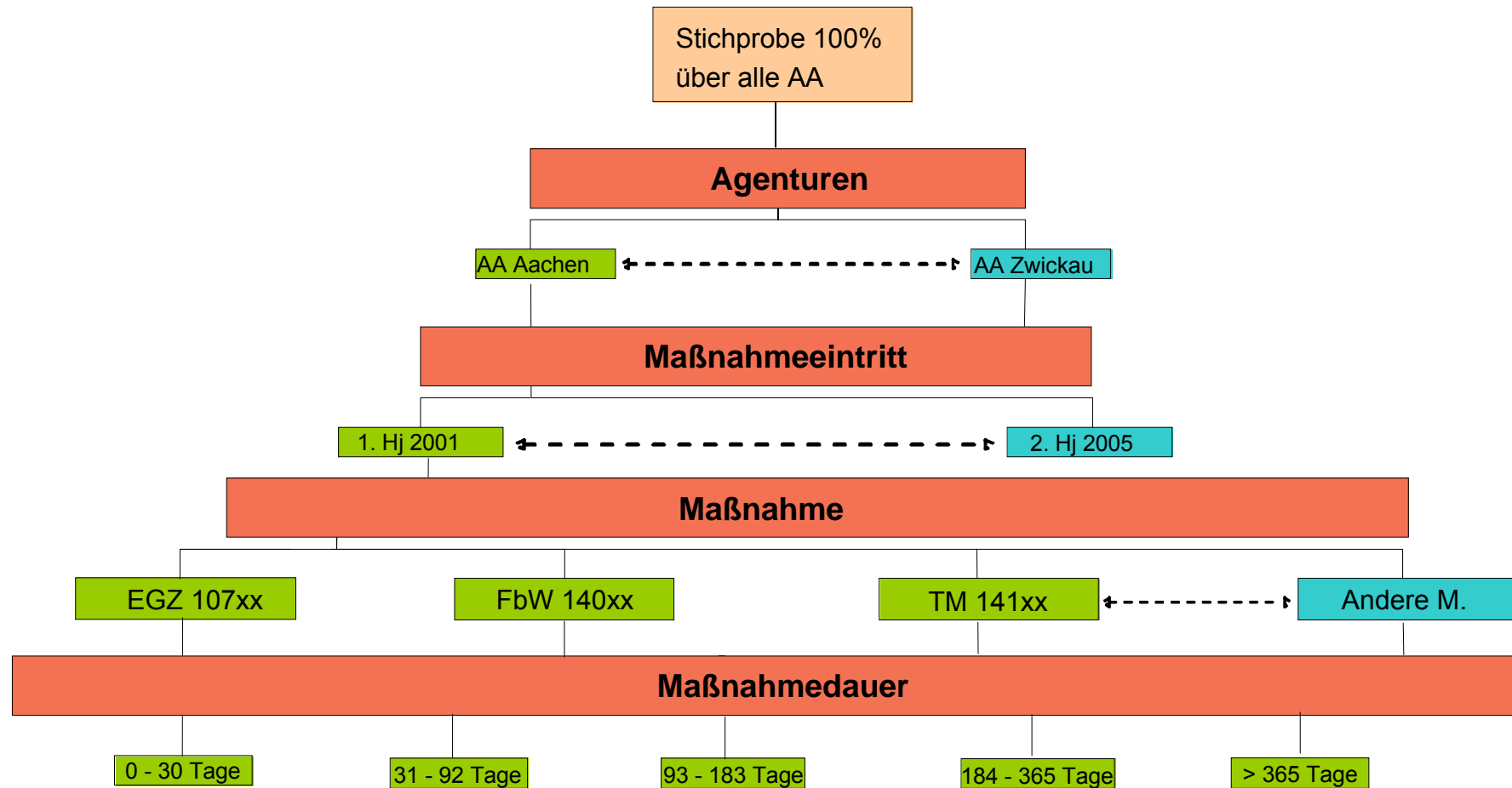
Ergebnisvariablen

- Vier Ergebnisvariablen zur Abbildung der Integrationsleistung bzw. der Vermeidung von Arbeitslosigkeit
- Verbleibsrisiko in Arbeitslosigkeit (inkl. Maßnahmeteilnahme):
Verbleib in faktischer Arbeitslosigkeit (VifA)
- Verbleibschancen in (geförderter und ungeförderter) sozialversicherungspflichtiger Beschäftigung:
Verbleib in sozialversicherungspflichtiger Beschäftigung (VisvB)
- Kumulierte Dauer der Arbeitslosigkeit (inkl. Maßnahmeteilnahme):
kumulierte Dauer faktischer Arbeitslosigkeit (kDfA)
- Kumulierte Dauer der (geförderter und ungeförderter) sozialversicherungspflichtiger Beschäftigung:
kumulierte Dauer sozialversicherungspflichtiger Beschäftigung (kDsvB)

Anforderungen an Kontrollpersonen

- Die Arbeitslosigkeitsdauer der Kontrollperson ist mindestens so lang wie die des Maßnahmeteilnehmers bei Eintritt in die Maßnahme.
- Teilnehmer und Kontrollperson haben dasselbe Geschlecht.
- Teilnehmer und Kontrollperson gehören zur selben Altersklasse.
- Die Anfangszeitpunkte der Arbeitslosigkeit von Teilnehmer und Kontrollperson liegen nicht mehr als 60 Tage auseinander.
- Keine Maßnahmeteilnahme der Kontrollperson in der betrachteten Episode.

Aufbau der „Match-Pakete“ & Automatisierung



Zwischen Traum und Alptraum

- **Datenbasis**
 - 17 Mio. Arbeitslosen-Episoden ohne Maßnahmen
 - 7 Mio. Maßnahmeteilnahmen
 - 115 differenzierte Maßnahmen
 - 178 Agenturbezirke

- **Über 43.000 Match-Pakete nach**
 - Maßnahmeart, Agentur, Dauer der Maßnahme
 - 10 Halbjahresschnitte: 1. Halbjahr 2001, ..., 2. Halbjahr 2005

- **aktueller Beobachtungszeitraum** bis Ende Oktober 2006

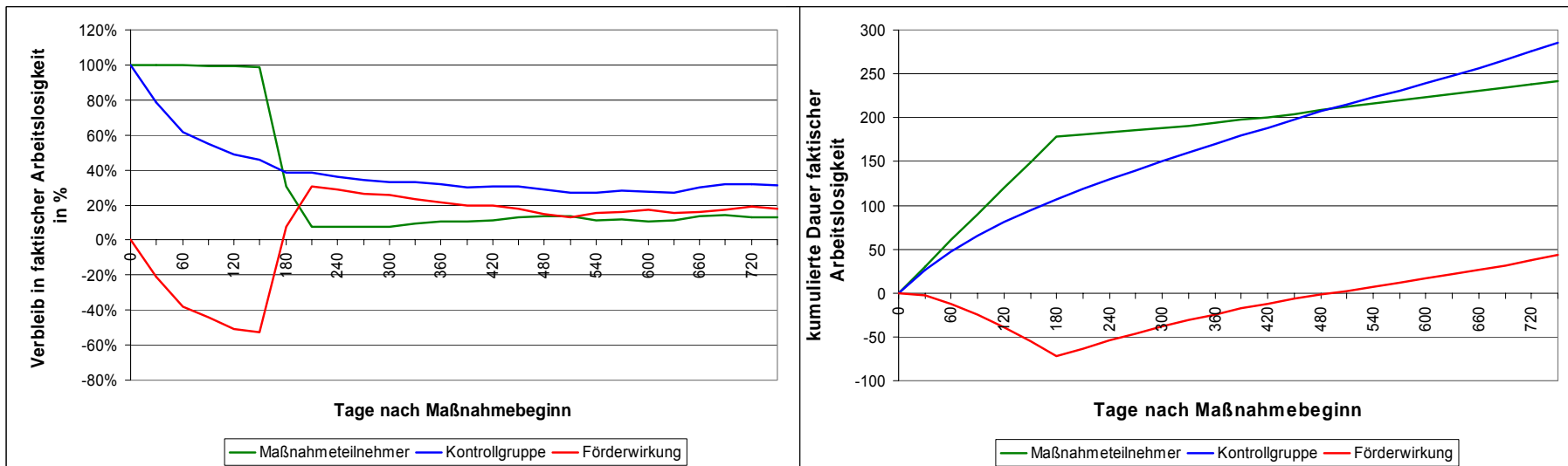
- **Laufzeit:** etwa 3 Wochen auf 20 PCs mit Server

Gliederung

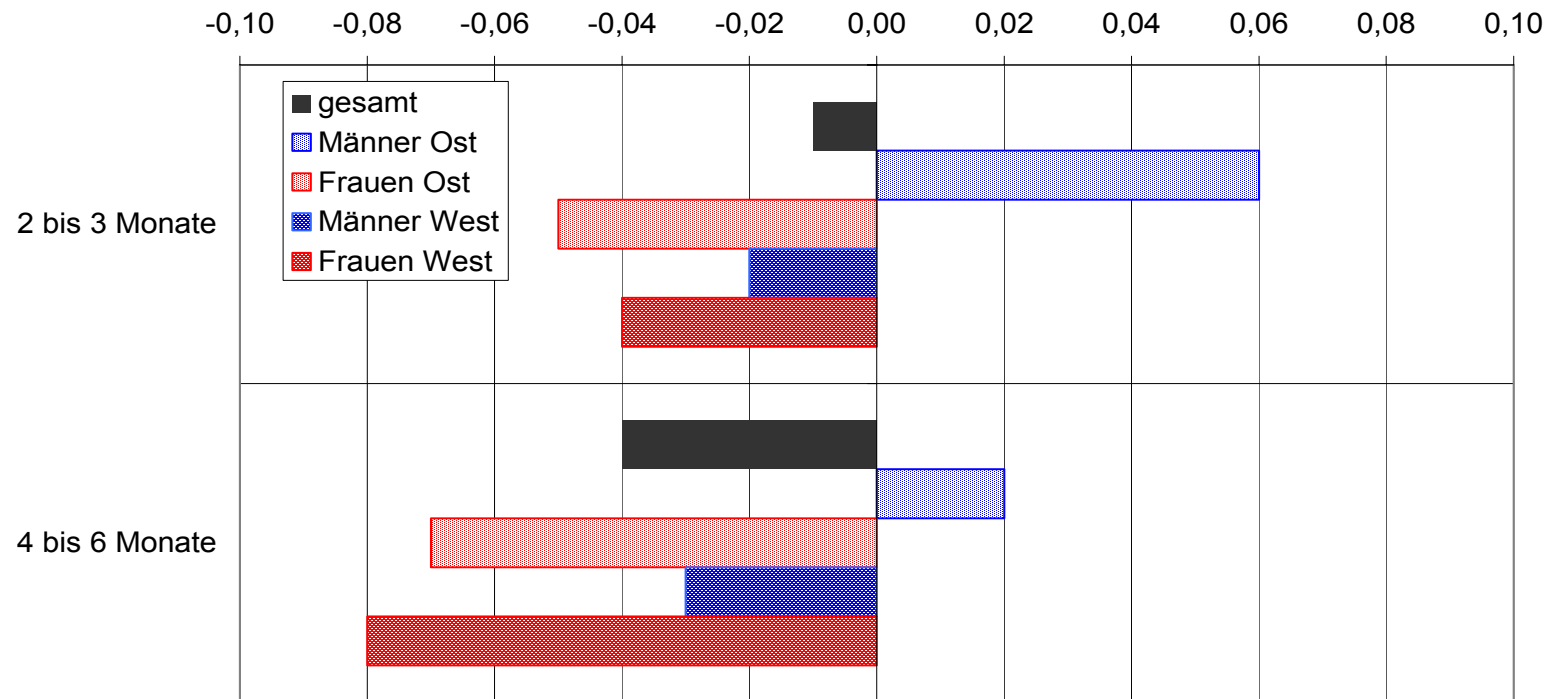
- Einleitung
- Evaluationsprobleme und ihre klassische Lösung
- Wirkungsanalyse in der Bundesagentur für Arbeit
- Operationalisierungen und Daten
- **Beispielhafte empirische Auswertungen**
- Fazit

Fördereffekte für ein ausgewähltes Beispiel

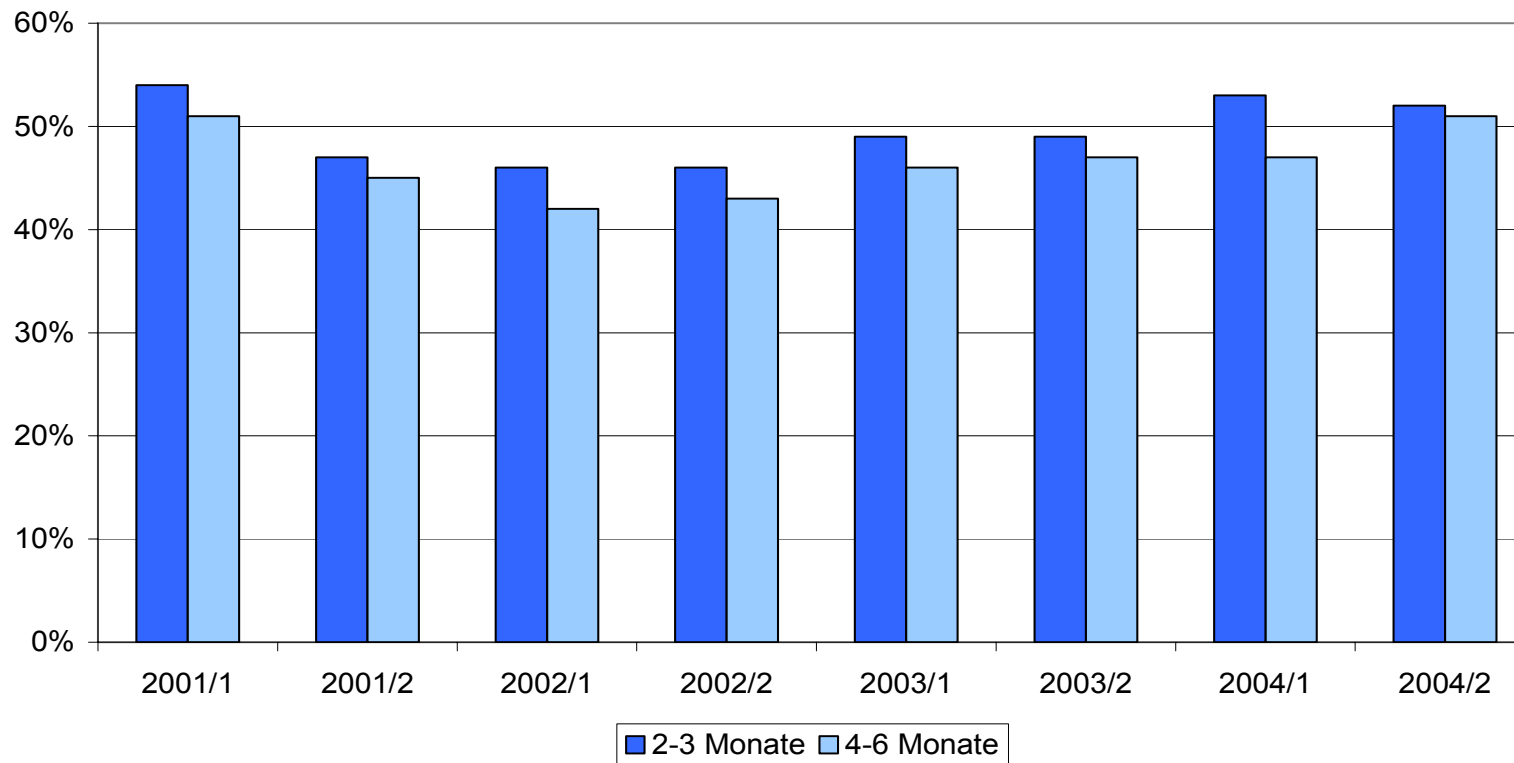
- **Überbrückungsgeld** in einer ausgewählten Arbeitsagentur
- **Fördereffekte:**
 - Am Stichtag in Arbeitslosigkeit oder Maßnahme
 - Kumulierte Dauer in Arbeitslosigkeit oder Maßnahme



Förderwirkung VifA360 für FbW - berufliche Weiterbildung Qualifikationserweiterung im Jahr 2004



Anteil positiver Effekte gemäß VifA360 für FbW - berufliche Weiterbildung Qualifikationserweiterung



Beispiel des Analysepotentials

- **Maßnahme:** Förderung der beruflichen Weiterbildung (Qualifikationserweiterung) mit einer Dauer von 4 bis 6 Monaten
- **Gesucht:** Beste Fördereffekte auf die Vermeidung von Arbeitslosigkeit 1 Jahr nach Maßnahmebeginn im Jahr 2004
- **Ergebnis:** Verheiratete Deutsche in Fertigungsberufen mit einem Schulabschluss bis Fachoberschule, die in den 18 Monaten vor Maßnahmebeginn zwischen 1 und 117 Tage arbeitslos waren:

		Ausbildung ^{*)}	
		keine Ausbildung	Berufsausbildung
regionale Arbeitslosen- quote	unter 10%	0.13	0.02
	10% bis unter 30%	0.06	0.24
	30% und mehr	0.07	0.13

^{*)} wegen zu geringer Fallzahlen ohne die Ausprägung „Studium“

Gliederung

- Einleitung
- Evaluationsprobleme und ihre klassische Lösung
- Wirkungsanalyse in der Bundesagentur für Arbeit
- Operationalisierungen und Daten
- Beispielhafte empirische Auswertungen
- **Fazit**

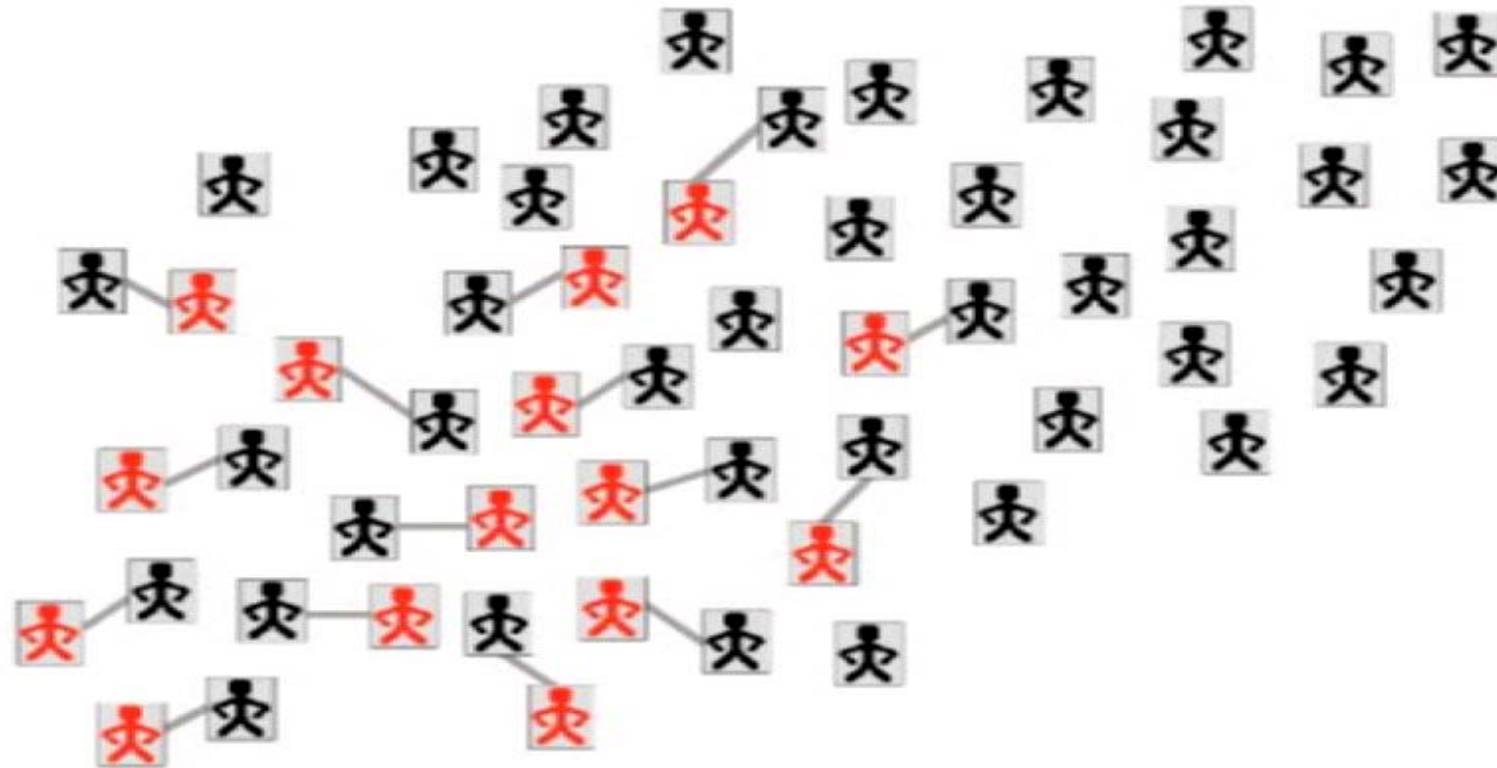
Fazit

- **Ist aktive Arbeitsmarktpolitik generell effektiv?**
 - Hängt vom Erfolgskriterium Y ab
 - Hängt vom Kontrollgruppenkonzept ab
 - ⇒ Variiert stark individuell und maßnahmenspezifisch

- **Enormes Analyse- und Lernpotential:** Ergebnisse stehen den Arbeitsagenturen dezentral und bis auf einzelne Maßnahmen auswertbar zur Verfügung
 - ⇒ Für welche Individuen wirken welche Maßnahmen am besten?
 - ⇒ Was ist „Good bzw. Best Practice“?

- **Prognose:** Die zu erwartenden Maßnahmeeffekte werden auf Grundlage dieser Erkenntnisse für neue Kunden geschätzt!

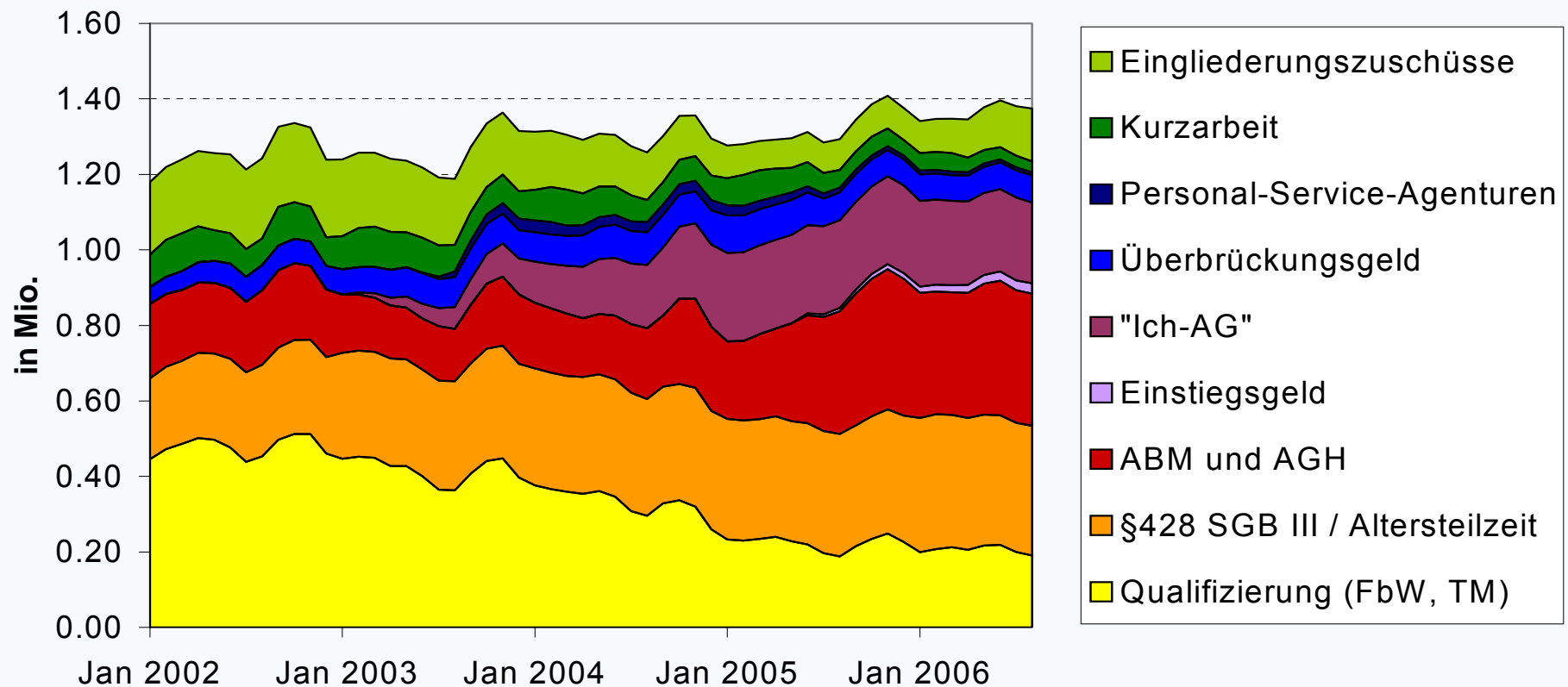
- **Antwort an Wolfgang Franz:** Rässler, S. (2006). Der Einsatz von Missing Data Techniken in der Arbeitsmarktforschung des IAB, AStA



Vielen Dank!

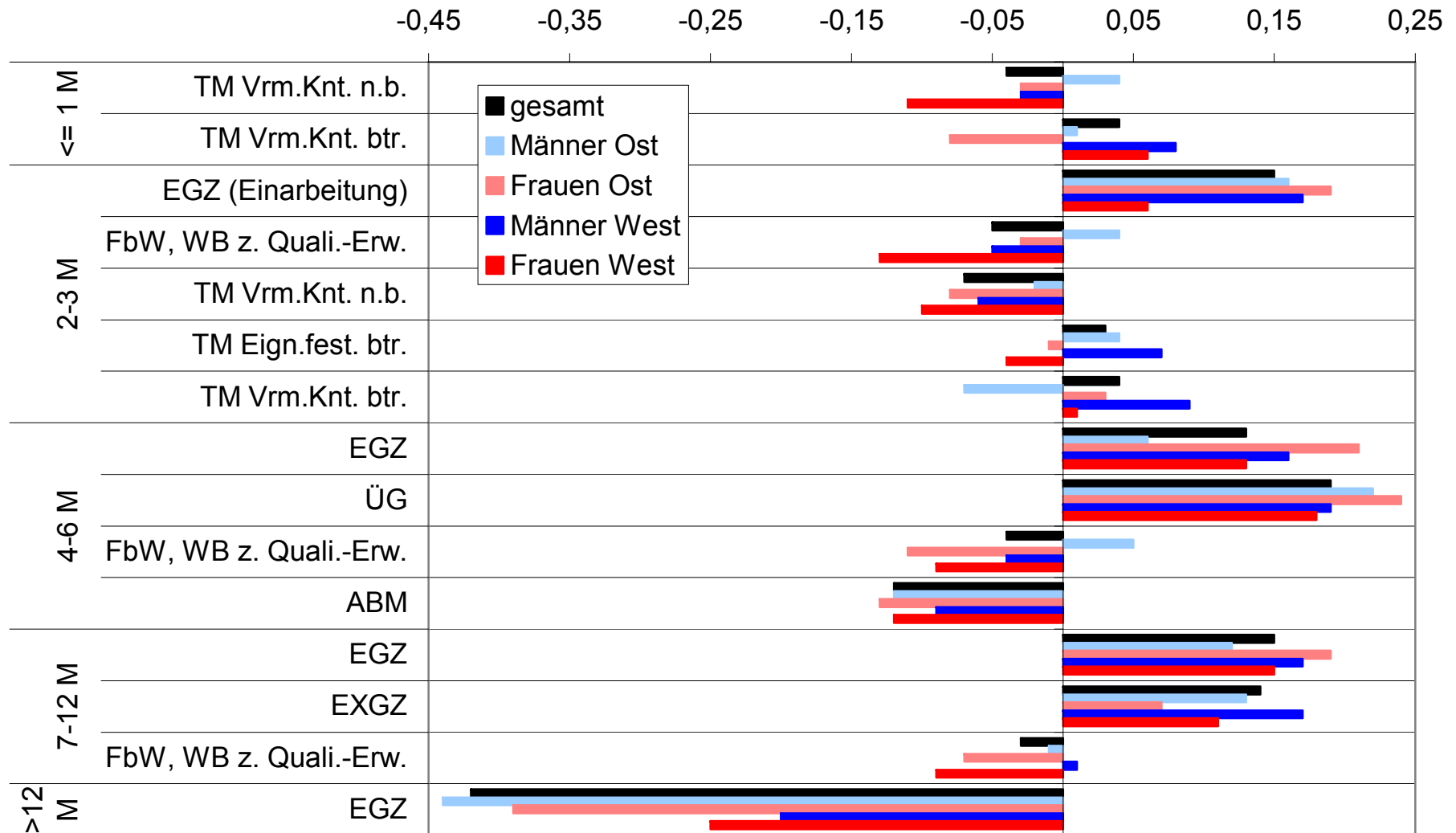
„Ist also die aktive Arbeitsmarktpolitik nicht „gut“ und verschleiert sie die „Wahre“? Einmal abgesehen davon, dass vor einem gültigen Urteil die angesprochenen Evaluationsergebnisse abgewartet werden sollten, empfiehlt es sich, folgenden Aspekt nicht völlig aus den Augen zu verlieren. Soll die Gesellschaft beispielsweise arbeitslose Jugendliche letztlich ihrem Schicksal überlassen und ihrem Weg in den Drogenkonsum, die Kriminalität und den politischen Radikalismus ebnen?“ (Wolfgang Franz, 2005, AStA 88, 156)

Förderung durch arbeitsmarktpolitische Instrumente - Bestand der Geförderten in Mio. -



Quelle: BA-Statistik

Förderwirkung VifA720 für ausgewählte Maßnahmen - Personen über 50



Anteil positiver Effekte nach kDfA360 für ausgewählte Maßnahmen

