

# The Use of Administrative Data to Evaluate ALMP: Lessons from Germany

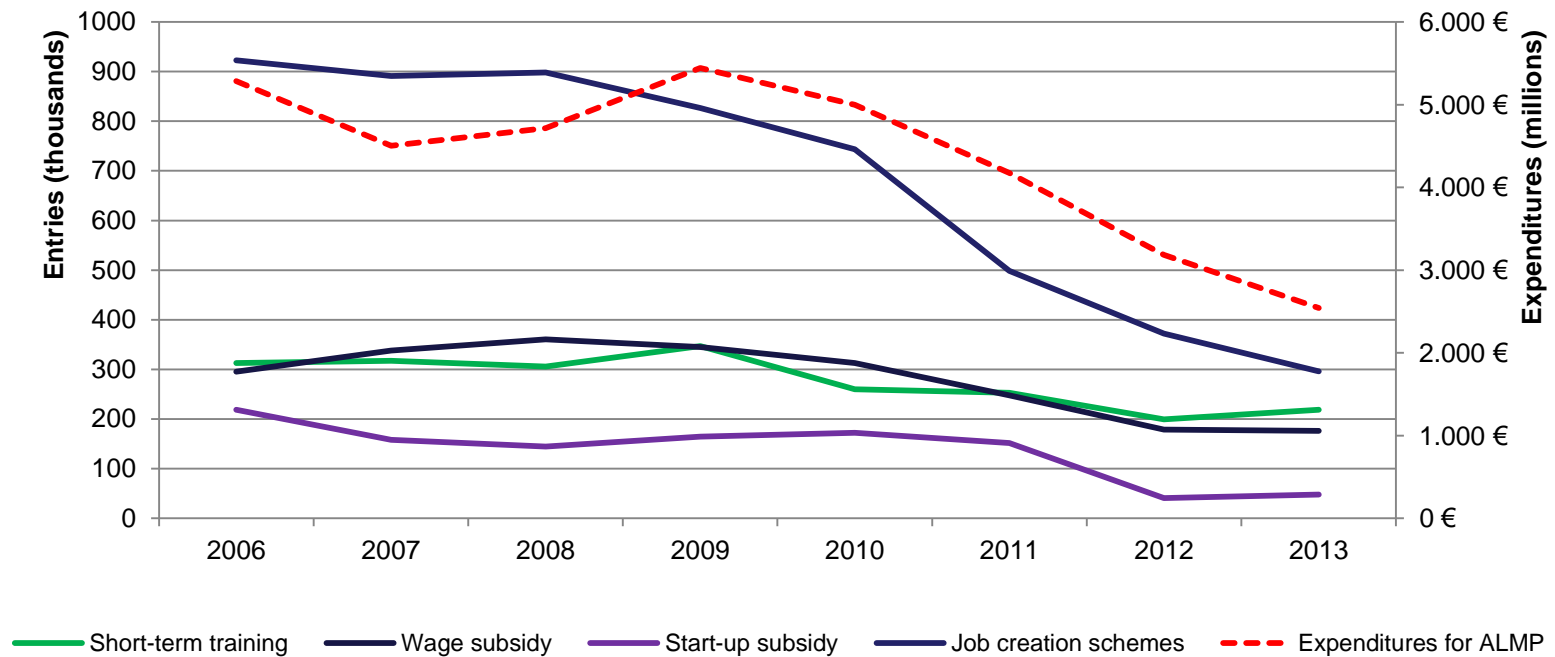
Steffen Künn

December 9, 2014

Barcelona

# Introduction

- Active Labor Market Policy (ALMP) in Germany



Source: Statistic of the Federal Employment Agency

- Key question: Are the programs successful?

# **Access to Administrative Data in Germany: A Brief Review**

## Access to Admin Data in Germany: A Brief Review

- Initiated by the scientific community, the German Federal Ministry of Education and Research (BMBF) appointed a commission to examine the data infrastructure in Germany and to make proposals how it could be improved with respect to scientific research
  - At this time, admin data were (in general) not accessible for researcher
  - Format and documentation of the data prevented scientific research



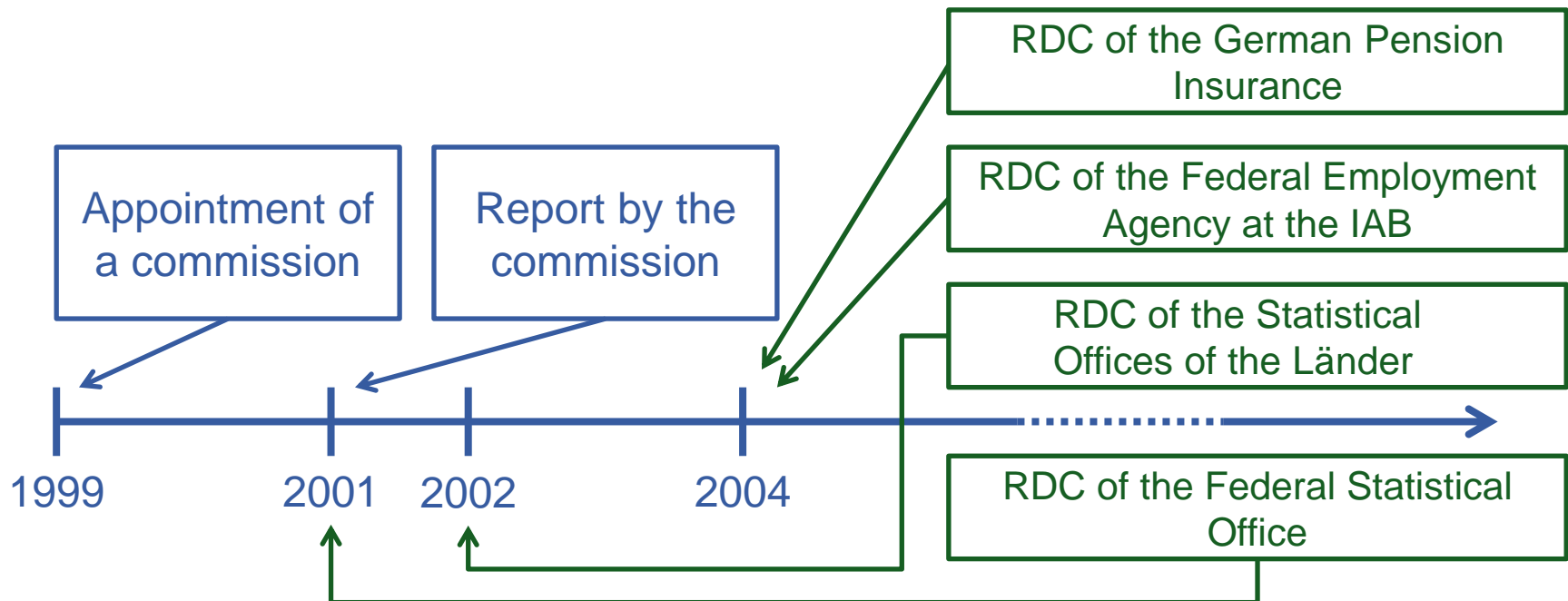
## Access to Admin Data in Germany: A Brief Review

- Report released: In total, 35 recommendations to improve the data infrastructure in Germany
- In particular, two recommendations had a significant impact!



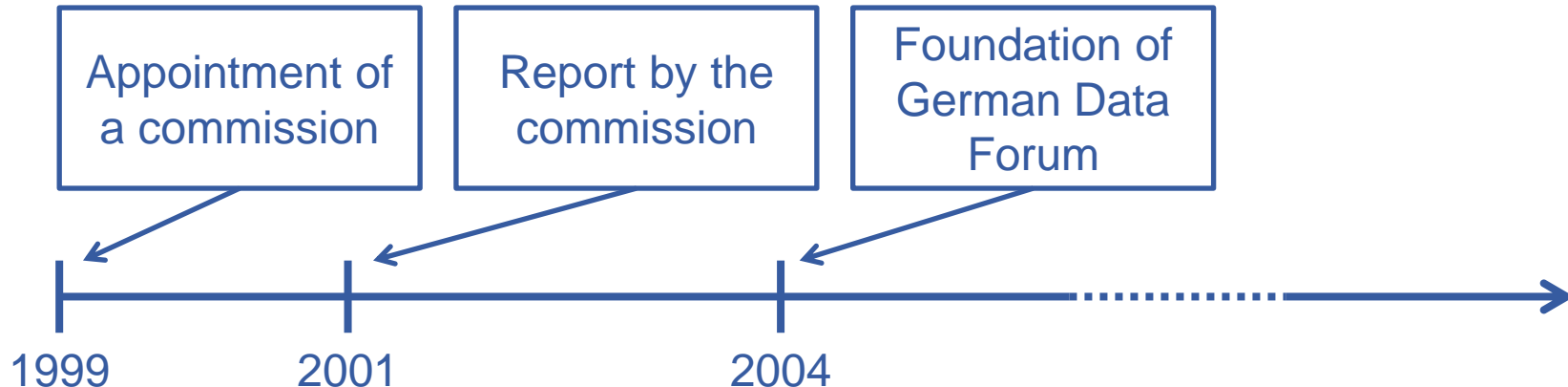
## Access to Admin Data in Germany: A Brief Review

- Foundation of Research Data Centers
  - Founded by data producer
  - Main Tasks:
    - Prepare, document and provide survey and admin data for scientific research
    - Creation of scientific use files, access to sensible data in a secure environment, remote access



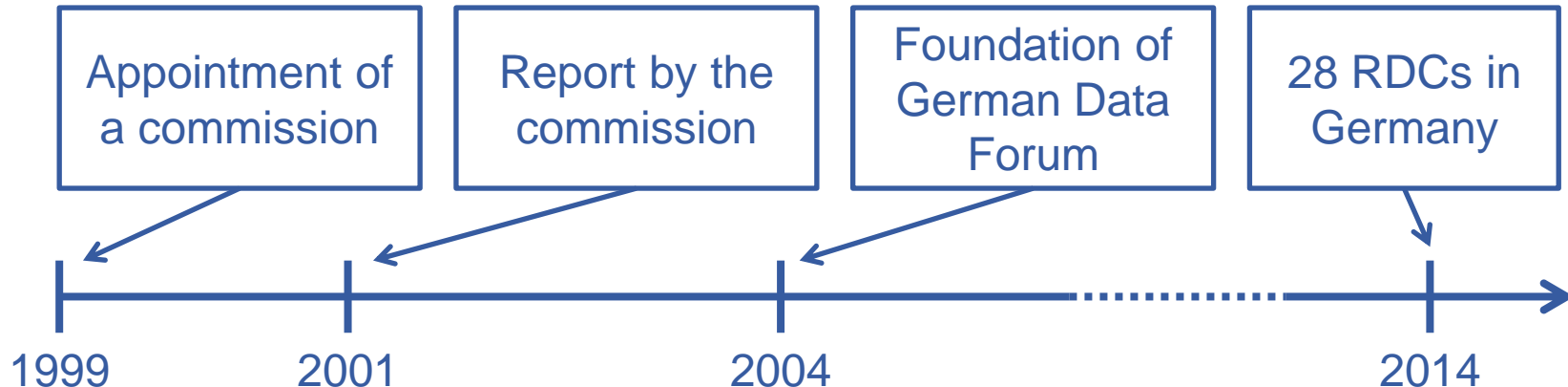
## Access to Admin Data in Germany: A Brief Review

- Foundation of the „German Data Forum“
  - Platform for exchange between data users and producers
  - Main Tasks
    - Secure and improve the data infrastructure in Germany
    - Evaluation and consulting of RDCs



## Access to Admin Data in Germany: A Brief Review

- Today: Well developed data infrastructure in Germany
  - In total, 28 RDCs exist (providing 323 datasets)
  - Many administrative records are available for scientific research, e.g.,
    - Social security data
    - Data of the Federal Employment Agency on unemployment, benefits and participation in ALMP
    - Firm level data
    - Pension data
    - ....

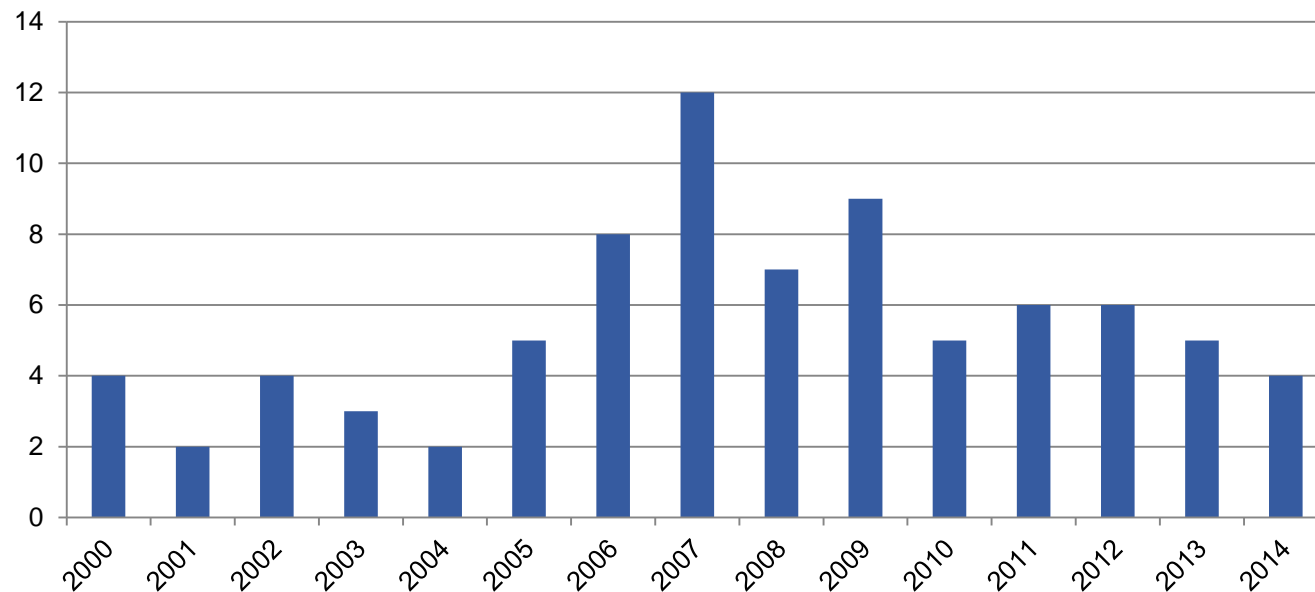




## Importance of Admin Data for Scientific Research

- The improvement of the data infrastructure in Germany and hence better access to admin data triggered a massive increase in research studies based admin data

**IZA Discussion Papers using admin data for ALMP evaluation**

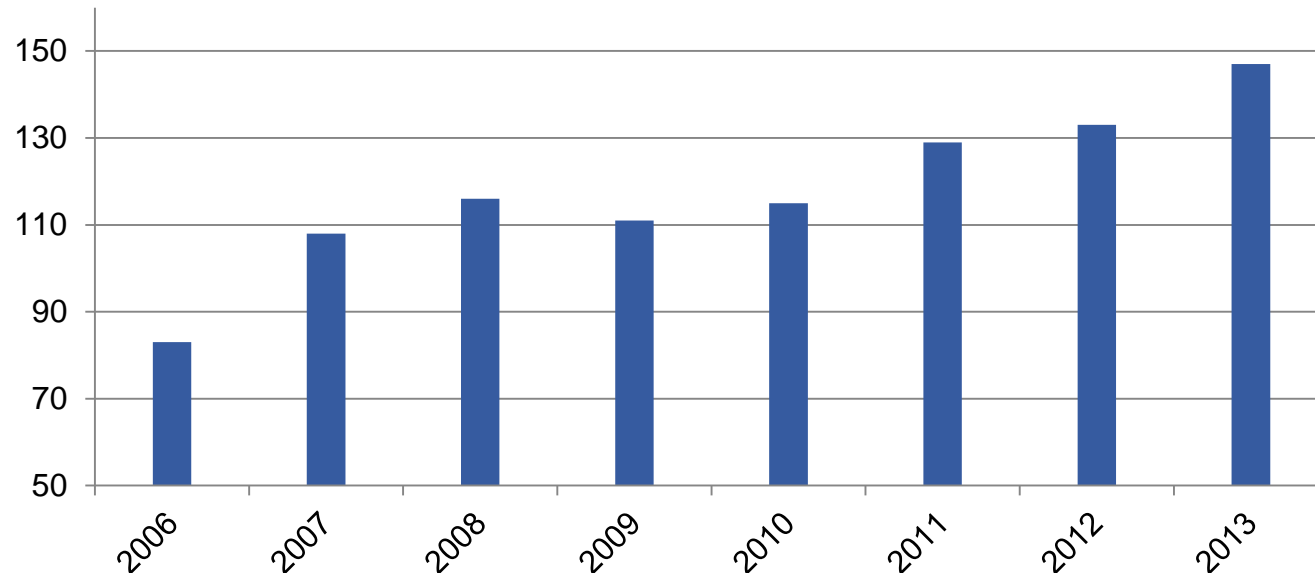


Source: Own calculations.

## Importance of Admin Data for Scientific Research

- The improvement of the data infrastructure in Germany and hence better access to admin data triggered a massive increase in research studies based admin data

### New Approved Research Projects at the RDC of the Federal Employment Agency

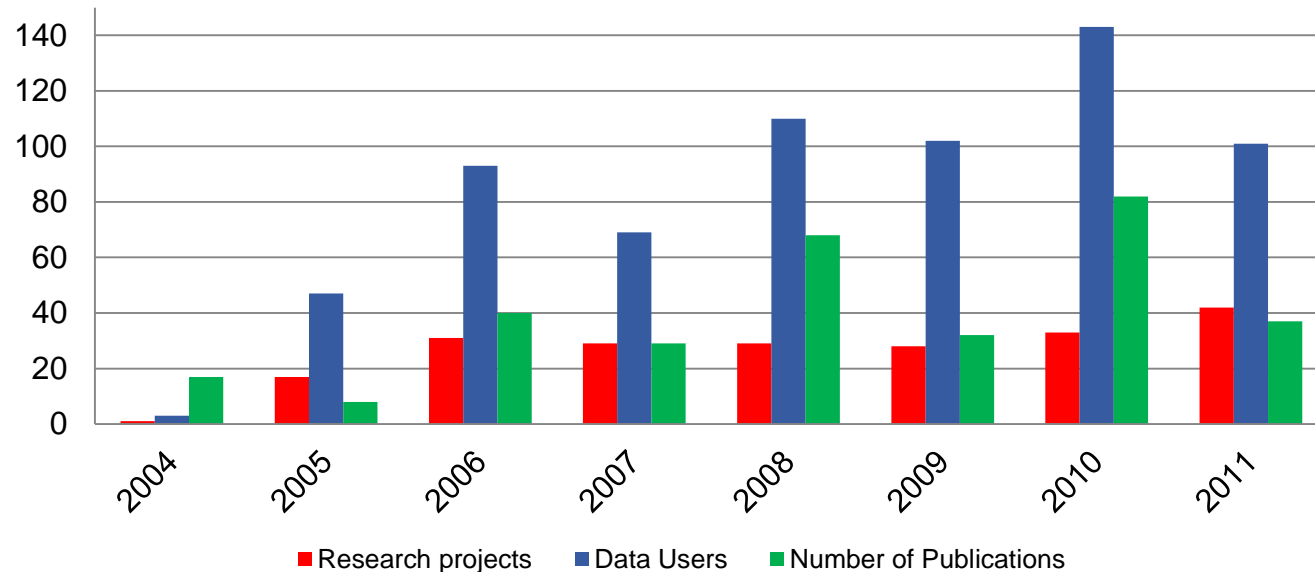


Source: Bender et al. (2014).

## Importance of Admin Data for Scientific Research

- The improvement of the data infrastructure in Germany and hence better access to admin data triggered a massive increase in research studies based admin data

### Data usage at the RDC of the German Pension Insurance



Source: Hansen et al. (2012)

## Are Admin Data Appropriate for ALMP Evaluation?

Key Question:      Do Admin Data Contain Sufficient  
Information to Achieve Reliable Results?

## Introduction to the Evaluation of ALMP

- Different econometric strategies exist to evaluate ALMP programs
- Main idea: Comparing the labor market outcomes of participants and non-participants
- Problem: Participants and non-participants differ! → Comparison without solving the selection problem would lead to biased results
- In the absence of experimental data, the key assumption to identify causal program effects is the conditional independence assumption
  - Comparison btw participants and „comparable“ non-participants in terms of observable characteristics
  - Justification heavily depends on the available information in the data
- Are admin data sufficiently informative to properly address the selection bias and hence estimate reliable effects?

## Advantages of Admin Data for the Evaluation of ALMP

- Compared to survey data, admin data ...
  1. are highly reliable (minimizes measurement error)
  2. are regularly updated and long time periods are observable for the same individual → No attrition!
  3. cover a large number of observations; sometimes 100% of the pop.
  4. are a relatively cheap data source (collected for admin reasons anyway)
  5. allow a clear assignment of treatment and control group as the exact ALMP program is observable in which the job seekers has participated
- A large set of individual information is likely to make the identifying assumption plausible achieving reliable estimates!
  - Labor market history, information on employers
  - Socio-demographics, regional characteristics

## Limitations of Admin Data for the Evaluation of ALMP

- However, admin data do not contain information on individuals' personality, labor market networks and cognitive skills
- Literature shows importance of these items for economic outcomes (e.g. Heckman et al., 2006)
- However, these items are usually not recorded in admin data as not required for administrative tasks
- Are the included information in admin data sufficiently informative or do we need additional variables to reliably estimate program effects?

## The IZA Evaluation Dataset

- Answering this question requires combined admin-survey data
- The *IZA Evaluation Dataset* was created to overcome such data limitations in the field of the evaluation of ALMP (Arni et al., 2014)
- Sample: Inflows into UE (N=18,000) btw June 2007 and May 2008
- The Survey follows individuals for 36 months and collects detailed information on labor market activities, job search behavior, personality, culture, social networks etc.
  - Available soon as SUF: visit <http://idsc.iza.org/iza-ed-survey>
- Main feature: Survey information can be linked to the admin data as provided by the Federal Employment Agency (restricted access)



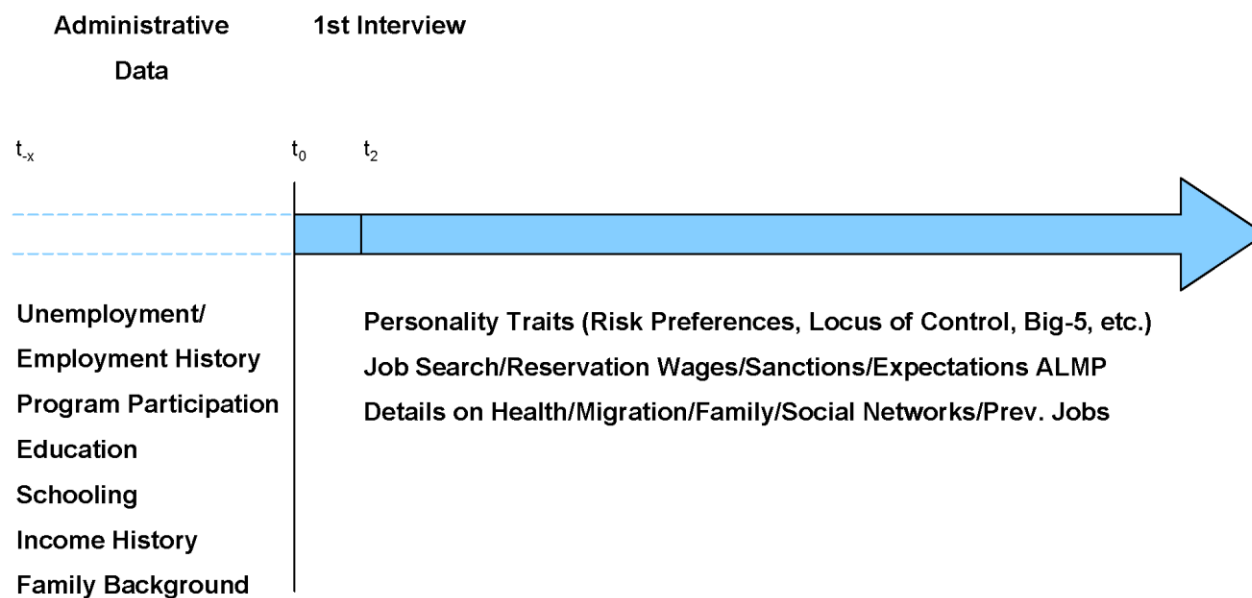
# The IZA Evaluation Dataset



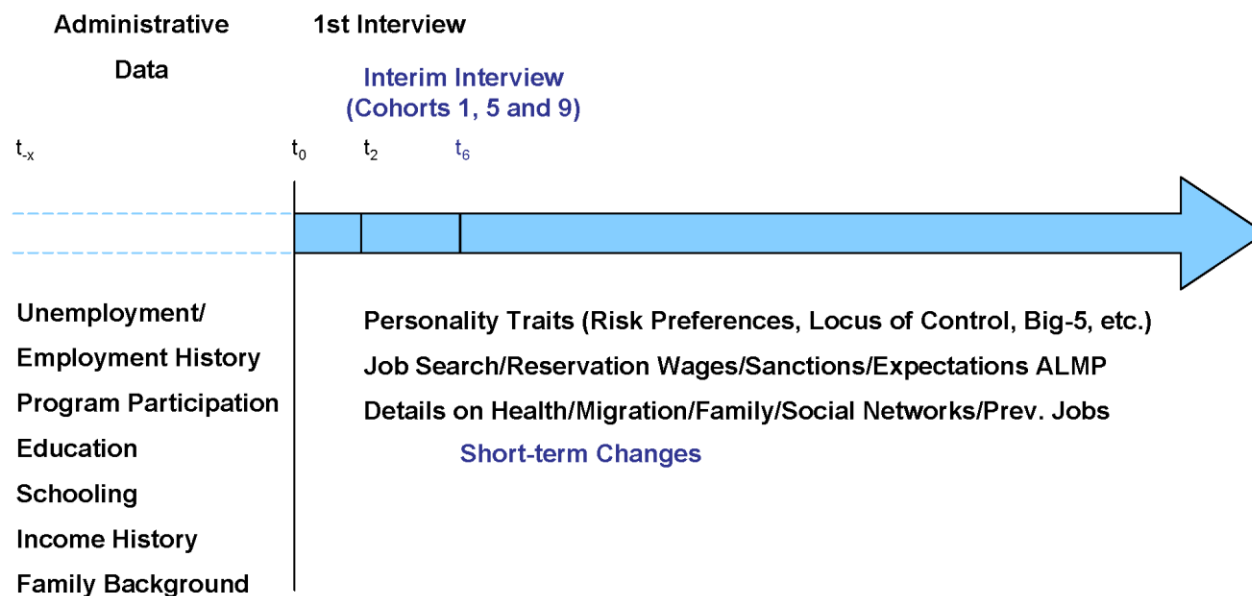
# The IZA Evaluation Dataset



# The IZA Evaluation Dataset



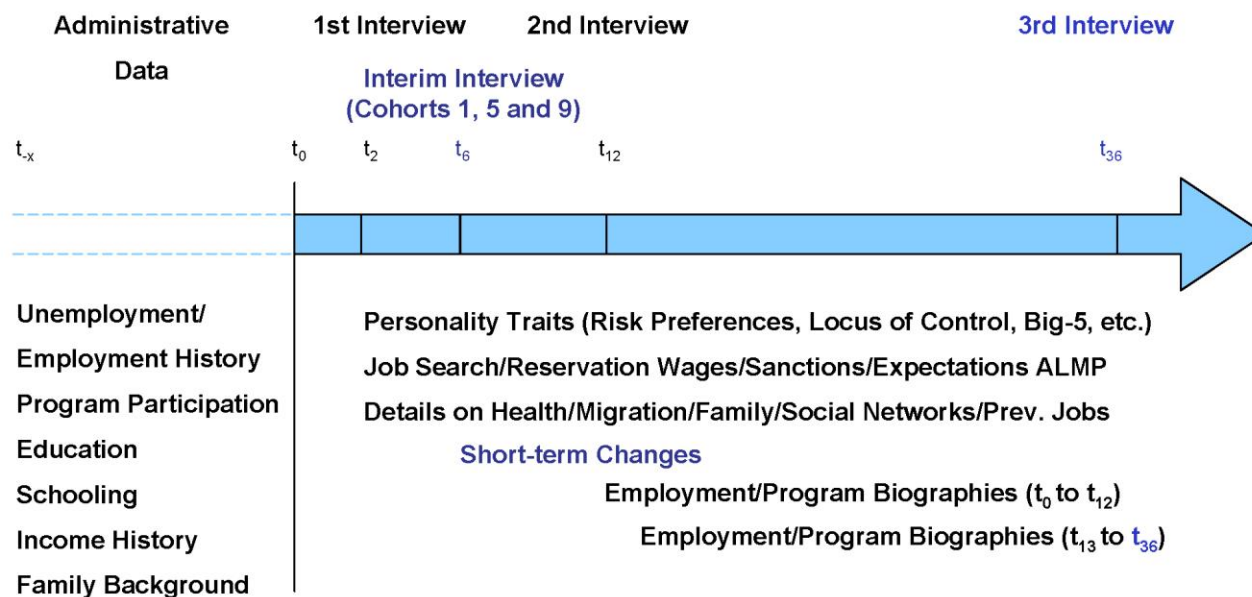
# The IZA Evaluation Dataset



# The IZA Evaluation Dataset



# The IZA Evaluation Dataset



## Do Admin Data Contain Sufficient Information?

- Based on the IZA ED, Caliendo/Mitnik/Mahlstedt (2014) investigate whether usually unobserved characteristics such as personality traits, attitudes, expectations, and job search behavior affect the estimation of program effects (ST training, wage subsidy)
- The availability of these items do not make a significant difference in the estimation of program effects, compared to an analysis relying on admin data only (including detailed labor market history)
- This suggests that **admin data seem to be rich enough** to estimate reliable program effects (STT/WS) allowing policy conclusions
- Their results are confirmed by similar studies:
  - Caliendo/Künn/Weißenberger (2014): Start-up Subsidies
  - Arni/Caliendo/Künn/Mahlstedt (2014): Prediction of LTU

## Conclusion

- The German experience
  - A well developed data infrastructure is key for scientific research (affects number of research studies significantly)
  - Admin data provide a relatively cheap data source for the evaluation of ALMP (as it is collected for admin purposes anyway) → Efficient use!
  - Recent studies have confirmed that admin data contain sufficient information (in particular LM history) to reliably evaluate ALMP programs
- However, the conclusion might differ for other countries, e.g., due to
  - Lower quality of admin data
  - Less available information in admin data
  - Institutional/cultural differences: Personality might not be sufficiently reflected by human capital or labor market experience



Dr. Steffen Künn  
Senior Research Associate

IZA, P.O. Box 7240,  
53072 Bonn, Germany  
Phone: +49 (0) 228 - 38 94 533  
Fax: +49 (0) 228 - 38 94 180  
E-mail: [kuenn@iza.org](mailto:kuenn@iza.org)

<http://www.iza.org>

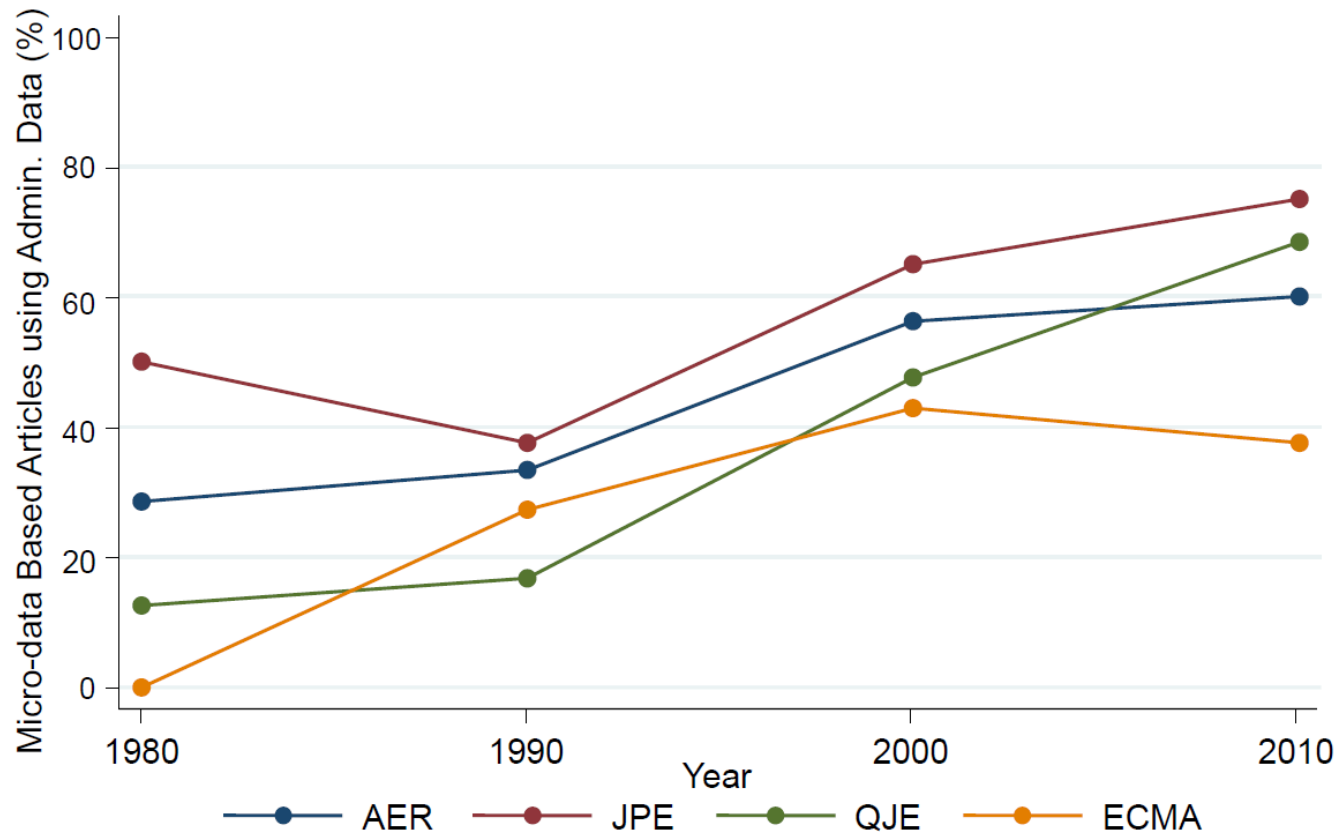
## References

- Arni, P., M. Caliendo, S. Künn and K. F. Zimmermann (2014): „The IZA Evaluation Dataset Survey: A Scientific Use File”, *IZA Journal of European Labor Studies*, 3-6.
- Arni, P., M. Caliendo, S. Künn and R. Mahlstedt (2014): „Predicting the Risk of Long-Term Unemployment: What can we learn from Personality Traits, Beliefs and other Behavioral Variables?“, Working Paper, Bonn.
- Bender, S. , A. Schmucker, I. Dieterich, I. Gunsellmann, D. Müller, S. Seth and V. Zakrocki (2014): “FDZ-Jahresbericht 2011-2013”. FDZ Methodenreport 08/2014, Nuremberg.
- Caliendo, M., S. Künn and M. Weißenberger (2014): „The Long-Term Effects of Start-Up Subsidies and the Role of Personality Traits“, Working Paper, Bonn.
- Caliendo, M., O. A. Mitnik and R. Mahlstedt (2014): “Unobservable, but Unimportant? The Influence of Personality Traits (and Other Usually Unobserved Variables) for the Evaluation of Labor Market Policies”, IZA Discussion Paper 8337, Bonn.
- Chetty, R. (2012): „Time Trends in the Use of Administrative Data for Empirical Research”, NBER Summer Institute, July 2012.
- Hansen, I., R. K. Himmelreicher, D. Mai and F. Röder (2012): „Entwicklung des Datenangebots und deren Nachfrage in neun Jahren Forschungszentrum der Rentenversicherung (2004 bis 2012)“, Working Paper No. 212, RatSWD Working Paper Series.
- Heckman, J.J., J. Stixrud and S. Urzua (2006): “The Effects Of Cognitive and Noncognitive Abilities On Labor Market Outcomes and Social Behavior“, *Journal of Labor Economics*, 24(3), 411-482.

# Appendix

## International Evidence

- Source: Chetty (2012)
- Increased usage of admin data over the last decades.



## International Evidence

- Source: Chetty (2012)
- While at the same time, usage of survey data decreased

