# The Comparative Effectiveness of Public Sponsored Training Revisited: The Merits of Using Rich Administrative Data

Martin Biewen[1] Bernd Fitzenberger[2] Aderonke Osikominu[3] Marie Paul[4]

[1]University of Tübingen, IZA, DIW
[2]Albert–Ludwigs–University Freiburg, ZEW, IZA, IFS
[3]Albert–Ludwigs–University Freiburg, IZA
[4]née Waller, Albert–Ludwigs–University Freiburg

April 2011

# Structure of the Presentation

- Summary of the paper
- Sensitivity Analysis "The Merits of Using Rich Administrative Data" as a focus of this presentation

# Summary of Paper: Introduction

- Training is an important tool of active labor market policy in many countries
- Many recent papers on employment or wage effects (see Card, Kluwe, and Weber 2010)
- In Germany 22 bn Euro on ALMP, 1.5 m entries into training (2002)
- Different schemes: short and long, class-room and practical
- Policy shift from long term training enhancing human capital to short activation programs (and partly backwards)

# Summary of Paper: Contribution

- Evaluation of training schemes: West Germany, 2000-2003
- Effects on employment probability and earnings
- Differential effects of different programs (short-term activation vs. long-term human capital enhancement; theoretical vs. practical)
- Effect heterogeneity w.r.t. different subgroups (gender, age groups, educational groups)
- Develop formal test for effect heterogeneity
- Analyze to what extend features of the data, the particular specification and the dynamic approach influence evaluation outcome: **focus of this presentation.**

# Summary of Paper: Approach

- Data: German administrative data (IEBS)
- Programs: Short-term training (STT), Classroom further training (CFT), Practical further training (PFT)
- Account for dynamic framework (Sianesi, 2004, 2008); distinguish programs starting after elapsed unemployment: 0 to 3 months (stratum 1), 4 to 6 months (stratum 2), and 7 to 12 months (stratum 3)
- Direct comparison of different programs (Imbens, 2000; Lechner, 2001)
- Combine exact matching and propensity score matching, local linear matching.
- Provide tests if individual characteristics and pre-treatment outcomes balanced

# Summary of Paper: Results

- Significant positive employment effects of short-term training and classroom further training for those who started training not too early during unemployment, 5 to 10 percentage points for men, higher for women
- Similar magnitude for short-term training as for long-term training, but shorter lock-in effects
- In cases with positive employment effects also earnings gains, a bit larger for long term training
- Women but not men benefit from practical training
- Average effects may hide significant effects for subgroups

# The Integrated Employment Biographies Sample (IEBS)

- Administrative data
- 2.2% random sample of individuals' data drawn from four administrative processes:
  - **Employment data**: Employment spells based on social security records (1990-2004)
  - **Benefit data**: Transfer payments by Federal Labor Office (1990-2005)
  - **Program data**: Participants in ALMP (2000-2005)
  - **Job search data**: Information on job searchers from labor offices (1990-2004)
- Spell structure: 17 m. spells, 1.4 m. individuals

# Benchmark Specification

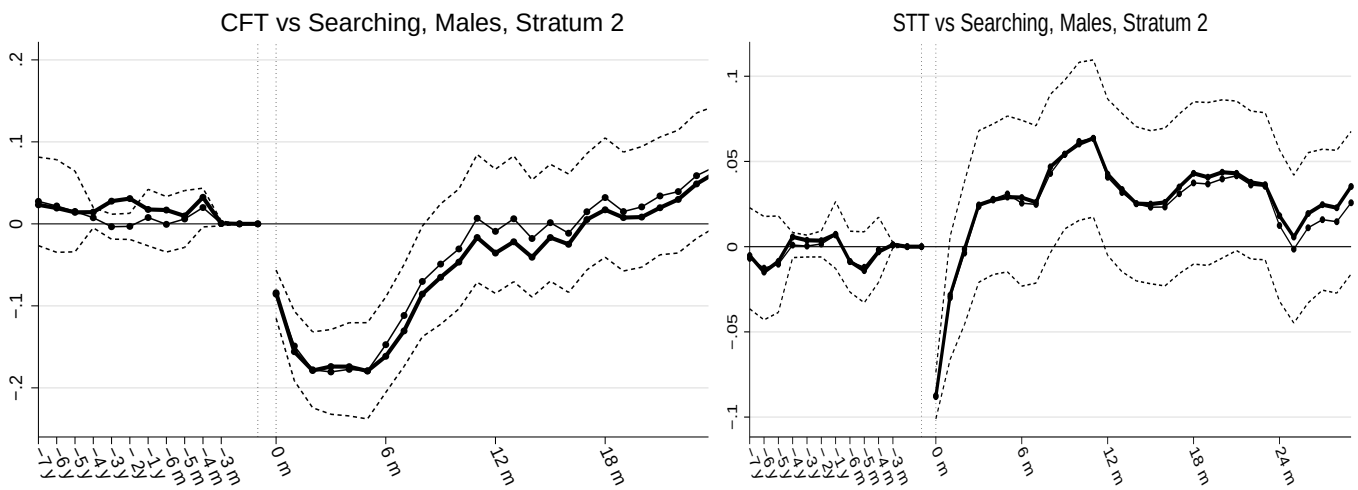
- Matching requirements:
  - Equality in the elapsed time in open unemployment
  - Equality in the previous employment history: exactly matching on the nine sequences (1000), (1001), (1010), (1011), (1100), (1101), (1110), (1111), and (0000).
  - Similarity in the pairwise propensity score (includes employment and earnings history, benefit entitlement, local labor market characteristics, rich personal characteristics etc.).
  - Similarity in the calendar date of the beginning of unemployment
- Specification search for each of the 34 groups (economic considerations, significance, balancing tests). Typically 20 to 35 covariates in PS.

# Sensitivity Analysis

- To what extent do features of our data and specification choices influence the outcome of our evaluation? What is important when choosing the econometric evaluation strategy?
- Start with benchmark specification and then sequentially drop specification features and/or information in the data.
- Aspects studied:
  - Employment History
  - Rich Personal Information and Specification Search
  - Comparison to MTG and Information on Other Programs
  - Future Participation in Other Programs

# Sensitivity Analysis 1: Employment History

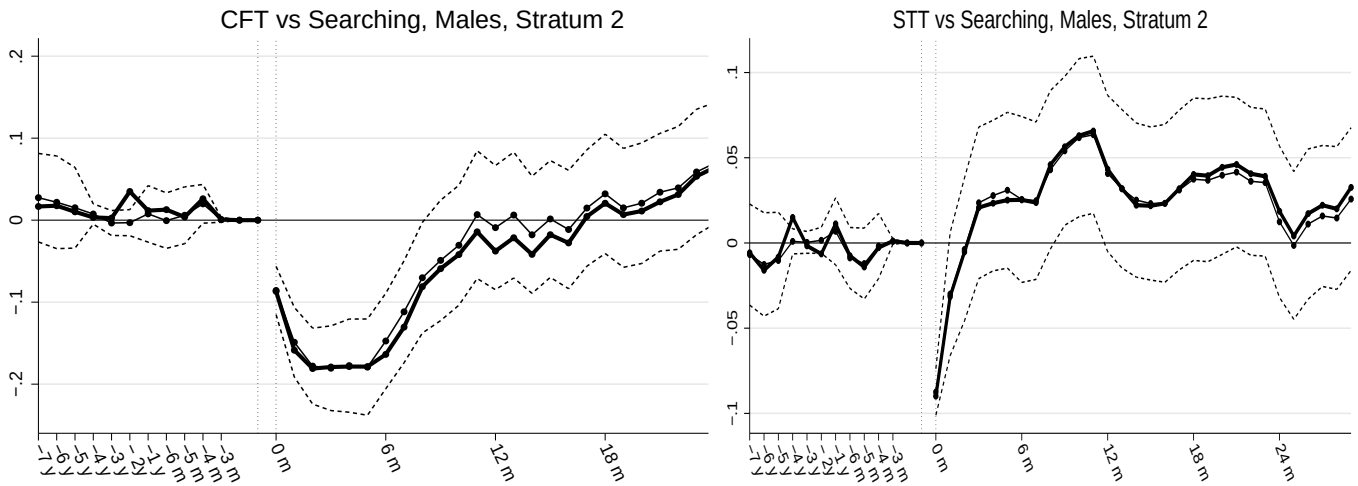
## Step 1: No exact matching on employment histories




Requirement to match *exactly* on employment sequence dropped.

# Sensitivity Analysis 1: Employment History

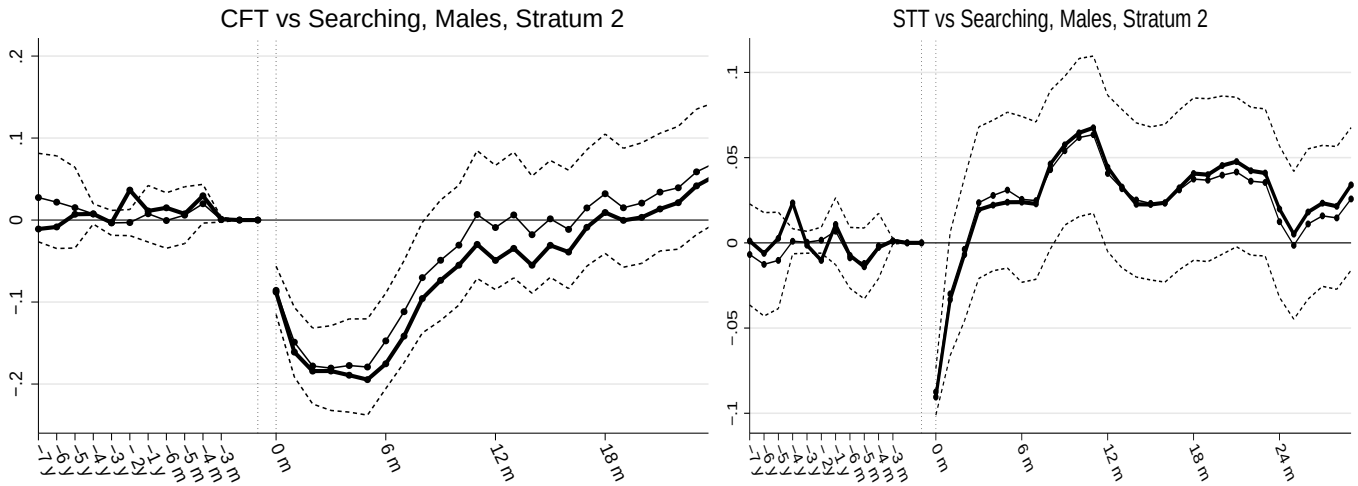
## Step 2: In addition, no history sequence dummies in propensity score




Variables with four year history sequence in addition dropped.

# Sensitivity Analysis 1: Employment History

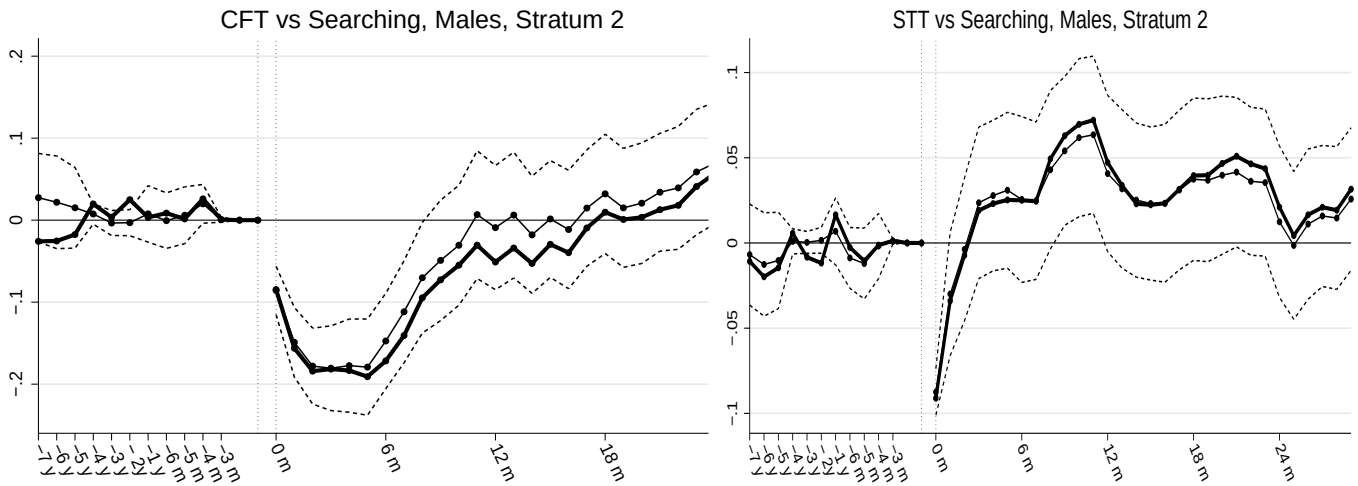
## Step 3: In addition, the seven-year history information is dropped from in propensity score




Information (on employment and earnings) going back to at most three years remains.

# Sensitivity Analysis 1: Employment History

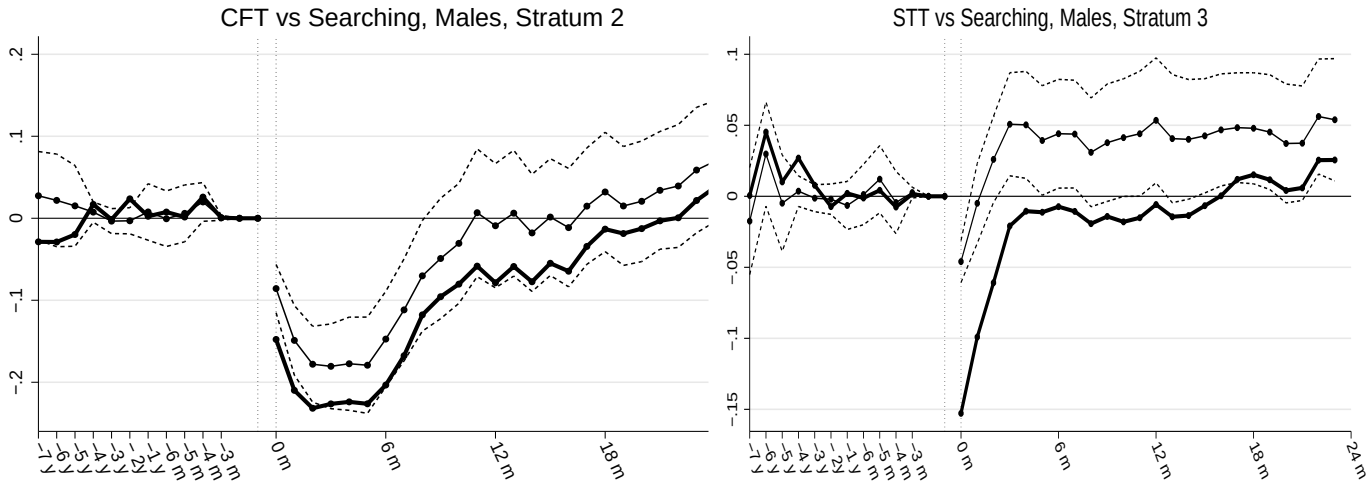
## Step 4: In addition, everything related to benefit information dropped




Non information on benefit claim, earlier benefit receipt etc. remains.

# Sensitivity Analysis 1: Employment History

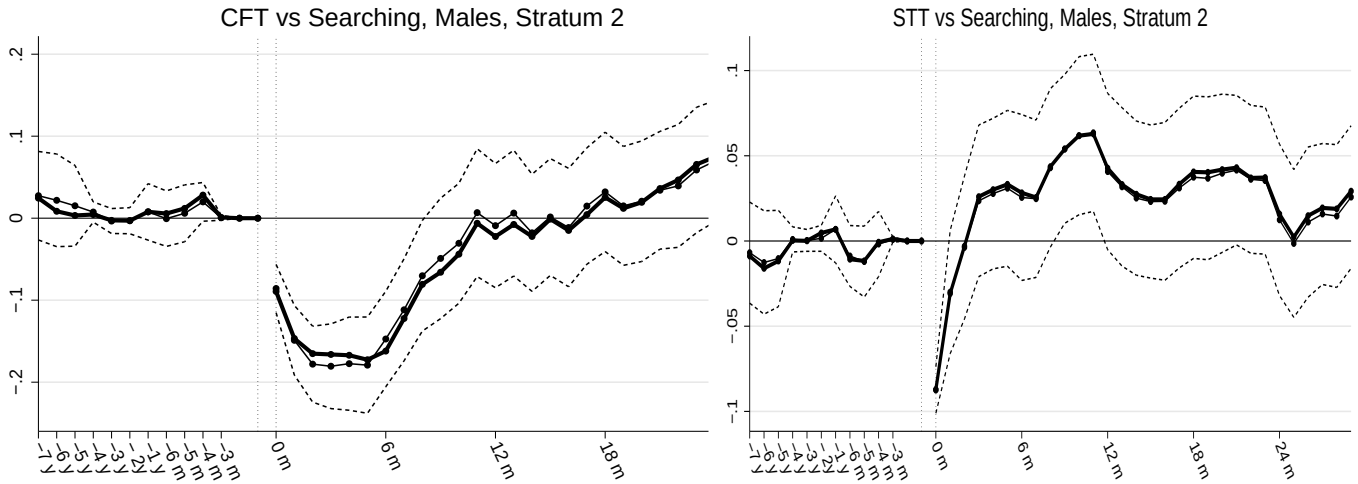
## Step 5: In addition, no exact matching on unemployment duration




Within a stratum, those with longer and shorter elapsed unemployment durations may be matched.

# Sensitivity Analysis 2: Rich Personal Information and Specification Search

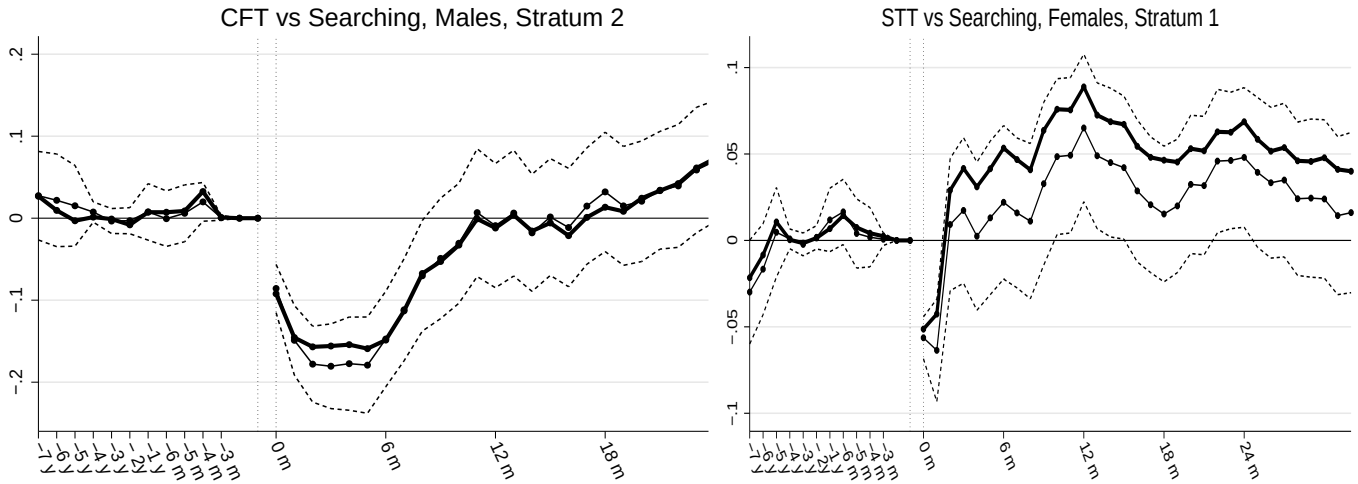
## Step 1: No personality variables




The following variables are dropped from PS: dropout or penalties in past, signs of lack of motivation in past, program with psychosocial component in past, wish to change occupation, number of job proposals.

# Sensitivity Analysis 2: Rich Personal Information and Specification Search

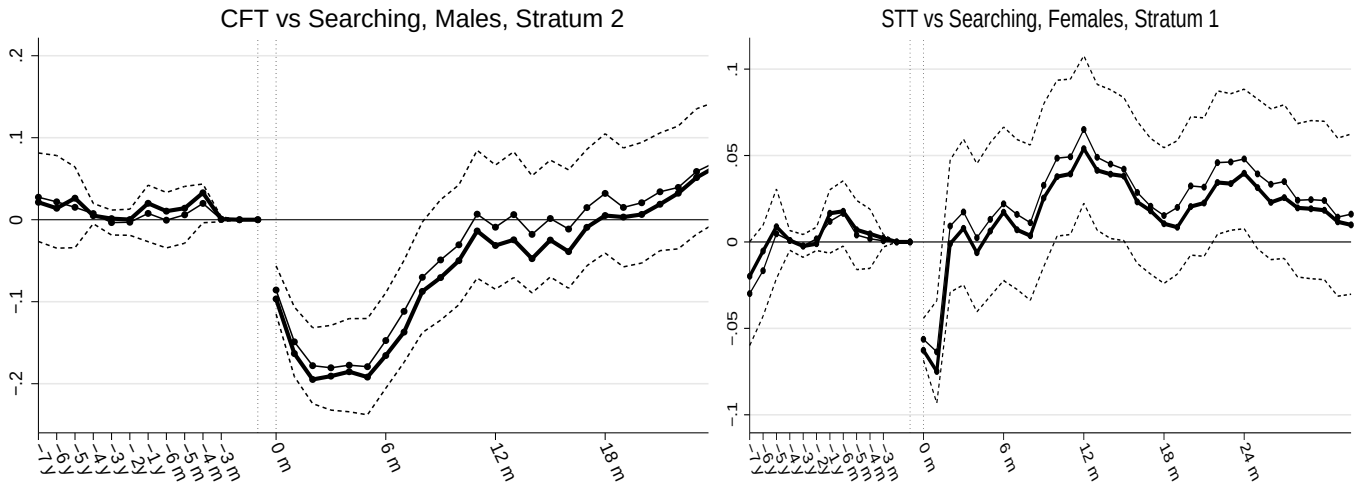
## Step 2: In addition, no rich personal information




The following variables are dropped from PS: information on health, disability, number and age of dependent children, marital status, information on household type, previous part-time employment and reasons why the last job was ended.

# Sensitivity Analysis 2: Rich Personal Information and Specification Search

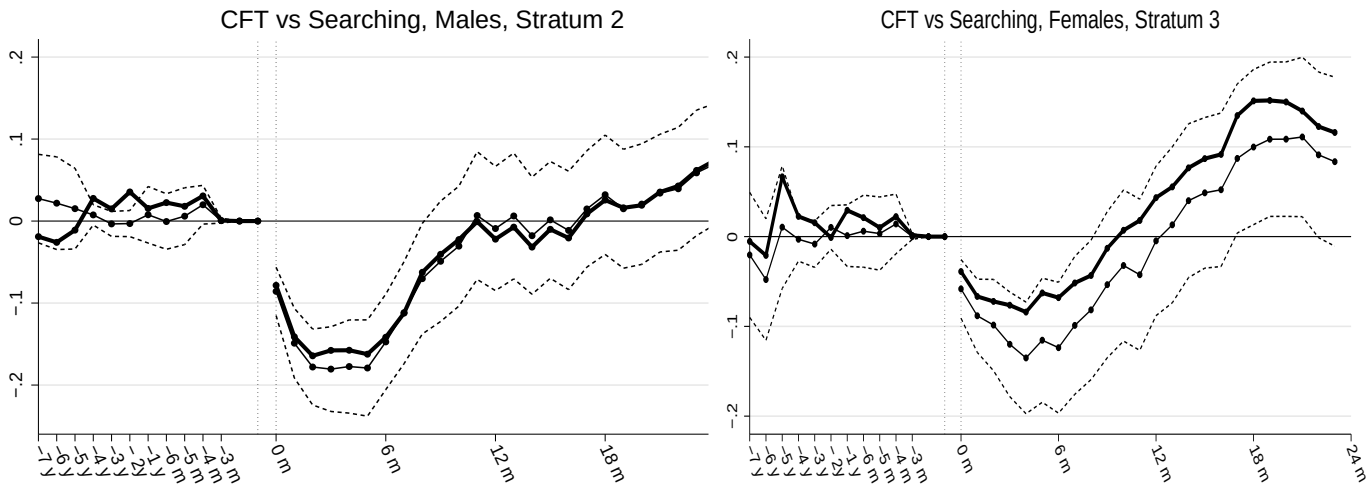
## No detailed specification search




Same basic specification (same variables, no interactions etc.) for each of the 34 groups.

# Sensitivity Analysis 3: Comparison with MTG and Information on Other Programs

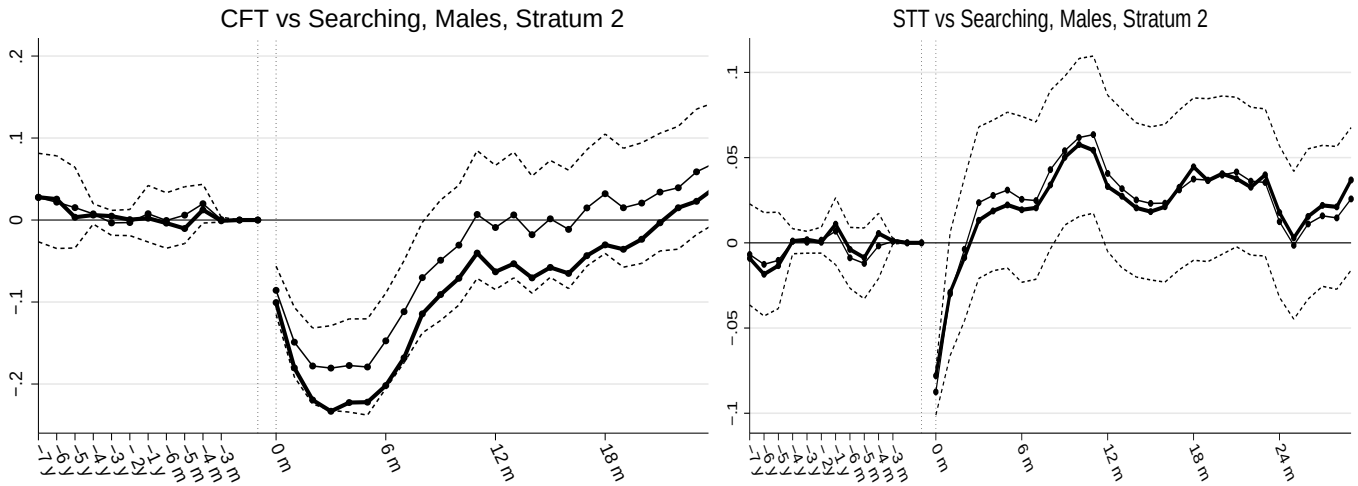
## Comparison with specification similar to MTG (Mueser et al., 2007)




MTG: study representative for what is possible using US administrative data sources (here: Missouri); basic demographic and educational information and region, four quarters of earnings history and employment dummies based on whether earnings were positive.

# Sensitivity Analysis 3: Comparison with MTG and Information on Other Programs

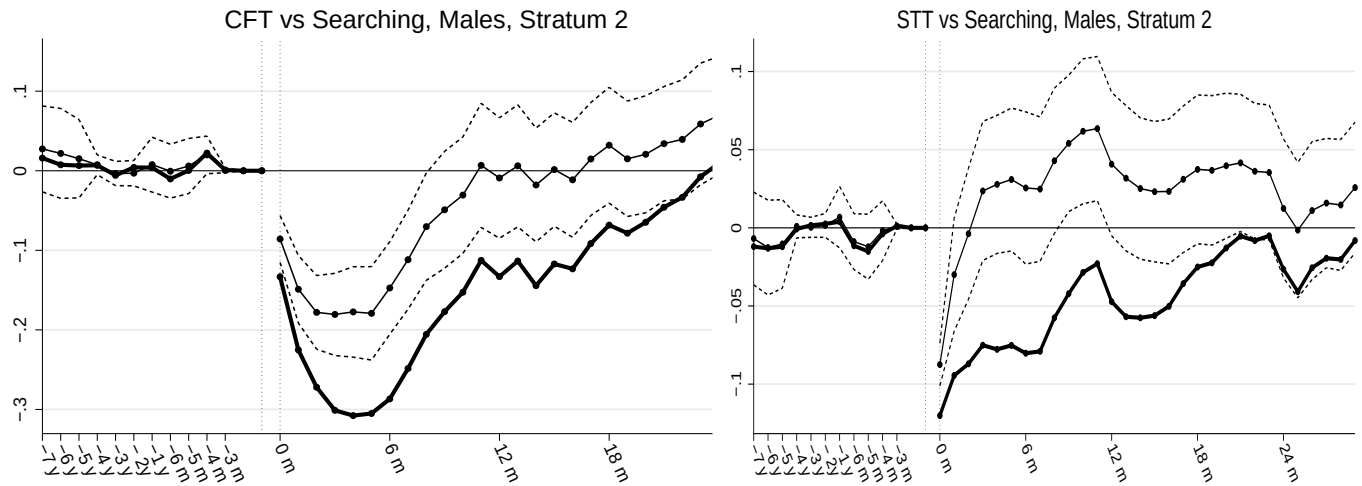
## Ignoring information on other programs




Ignores all information on programs other than the one in question; this mimics the situation in MTG who do not know if controls have participated in other programs.

# Sensitivity Analysis 4: Future Participation

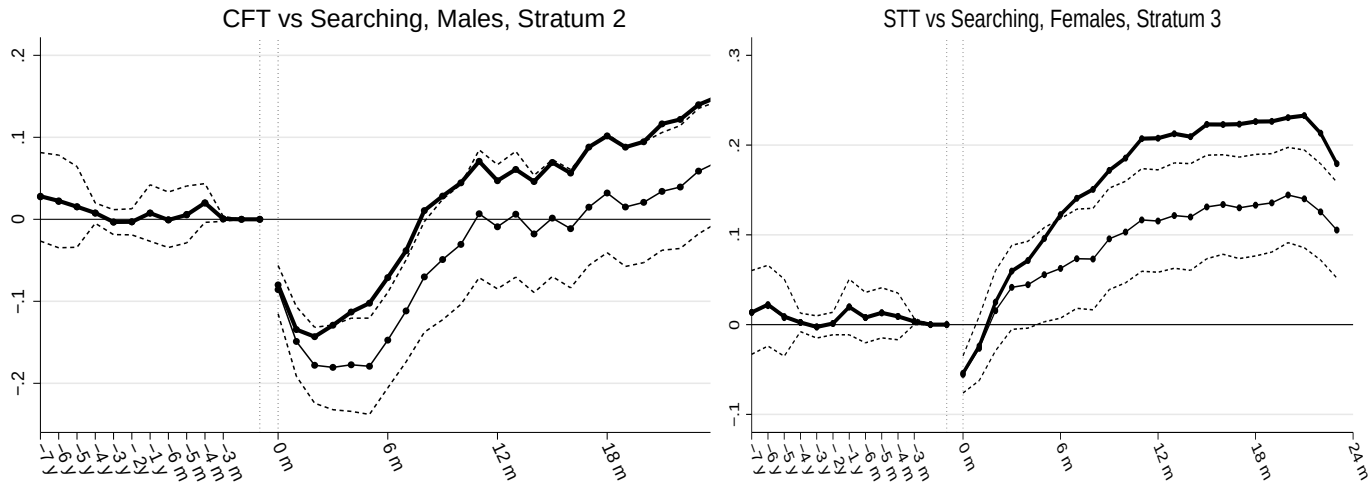
## Excluding future participants from the beginning of stratum




Exclude all individuals from the beginning on who will eventually (within the same unemployment spell and 35 months) participate in training or another intensive active labor market program.
Corresponds to control group design: never treated.

# Sensitivity Analysis 4: Future Participation

## Excluding future participants from the month they enter a program




Exclude future participants from the control group from the month they enter a program onwards. (To check whether our effects are systematically driven up because control group members are locked in future programs.)

# Conclusion

- Most important feature of the employment history to be controlled for is the unemployment duration prior to treatment, other aspects of employment history surprisingly little additional impact
- Availability of rich personal information and information on training programs other than the ones in focus important.
- Variations in the dynamic definition of the control group may have strong consequences for evaluation results.