Nicht-experimentelle Methoden zur Kausalanalyse:
Am Beispiel der Evaluierung öffentlicher Gründungsförderung

Hanna Hottenrott – Technische Universität München

FT Eval Veranstaltung: "Was können ökonometrische Methoden, was können sie nicht?„

Österreichisches Institut für Wirtschaftsforschung

3. September 2021

Start-up subsidies?

NSF SBIR/STTR
SMALL BUSINESS INNOVATION RESEARCH
SMALL BUSINESS TECHNOLOGY TRANSFER

UK
OPEN for BUSINESS
Government Grants & Funding for Small Businesses
The UK Government understands how crucial businesses are for our economy
Funding May Be Available For Your Small Business

Bayerisches Staatsministerium für Wirtschaft, Energie und Technologie

Gründerland. Bayern

Bundesministerium für Wirtschaft und Energie

eXIST

LfA FÖRDERBANK BAYERN
Beratung. Finanzierung. Erfolg.

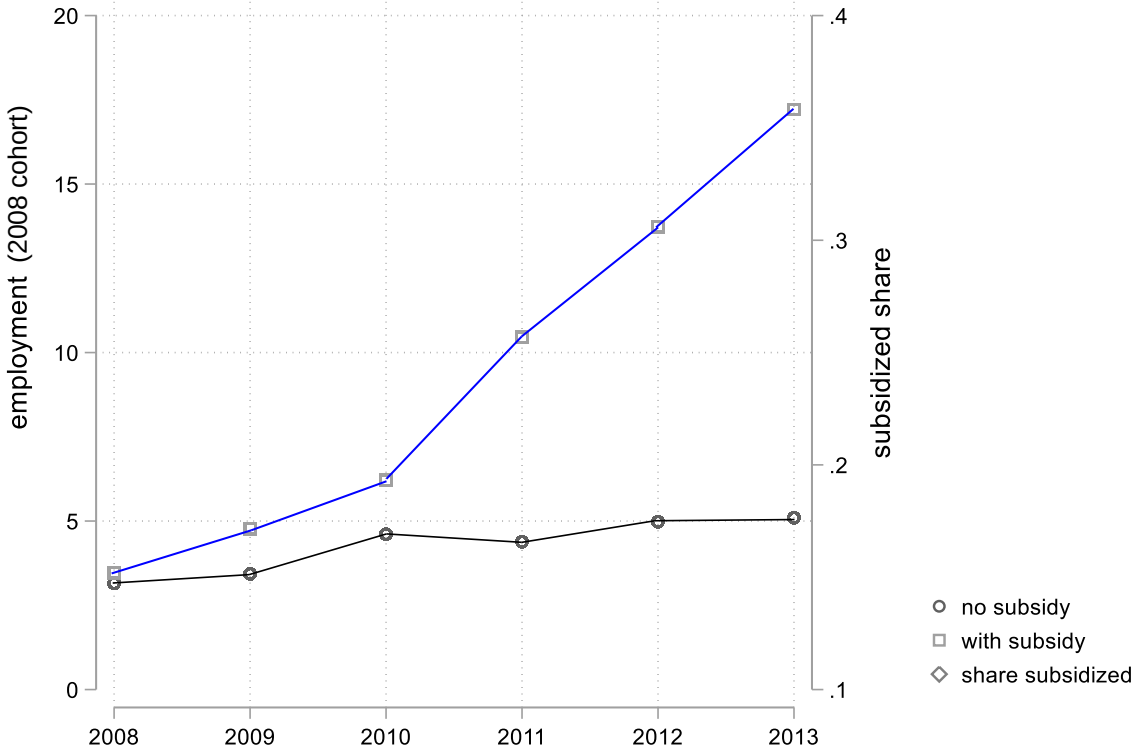
START UP LOANS

KFW
Bank aus Verantwortung

The main research question

Do subsidies help startups to perform better?

Example: Employment growth



The evaluation question

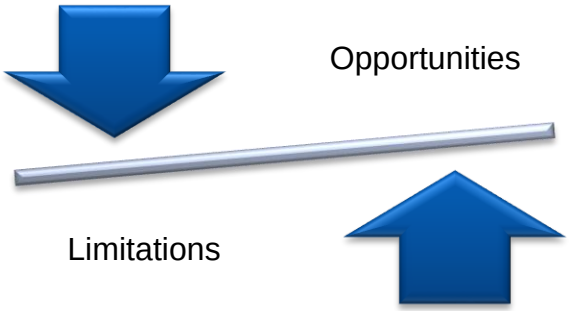
How would a start-up have performed in absence of the subsidy?

Answer:

We will never know.

Solution?

- **Econometric methods for causal interference in non-experimental settings**
- Example implementations based on two recent articles



Research design

We cannot observe how a start-up would performed in absence of the subsidy

But:

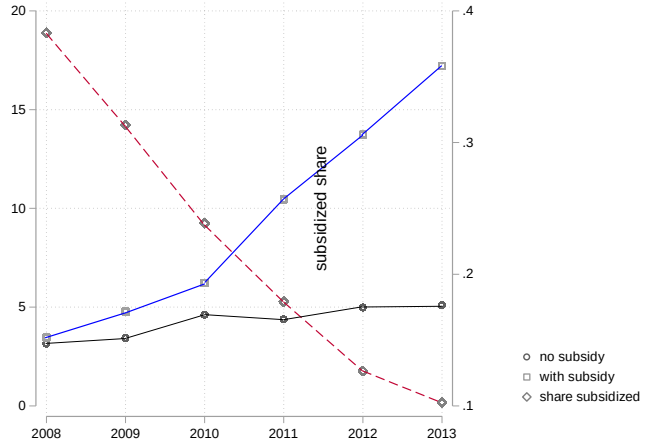
We can try to *estimate* the counter factual situation

to get an **approximation of the treatment effect** based on difference between the treated firms' outcomes and the counterfactuals

Which method to pick?

- Depends on research setting & available data
  - Do we know successful and unsuccessful applicants?
  - Do we know details about the selection process, e.g. evaluation scores of proposals?
  - Do we observe firms before and after the receipt?

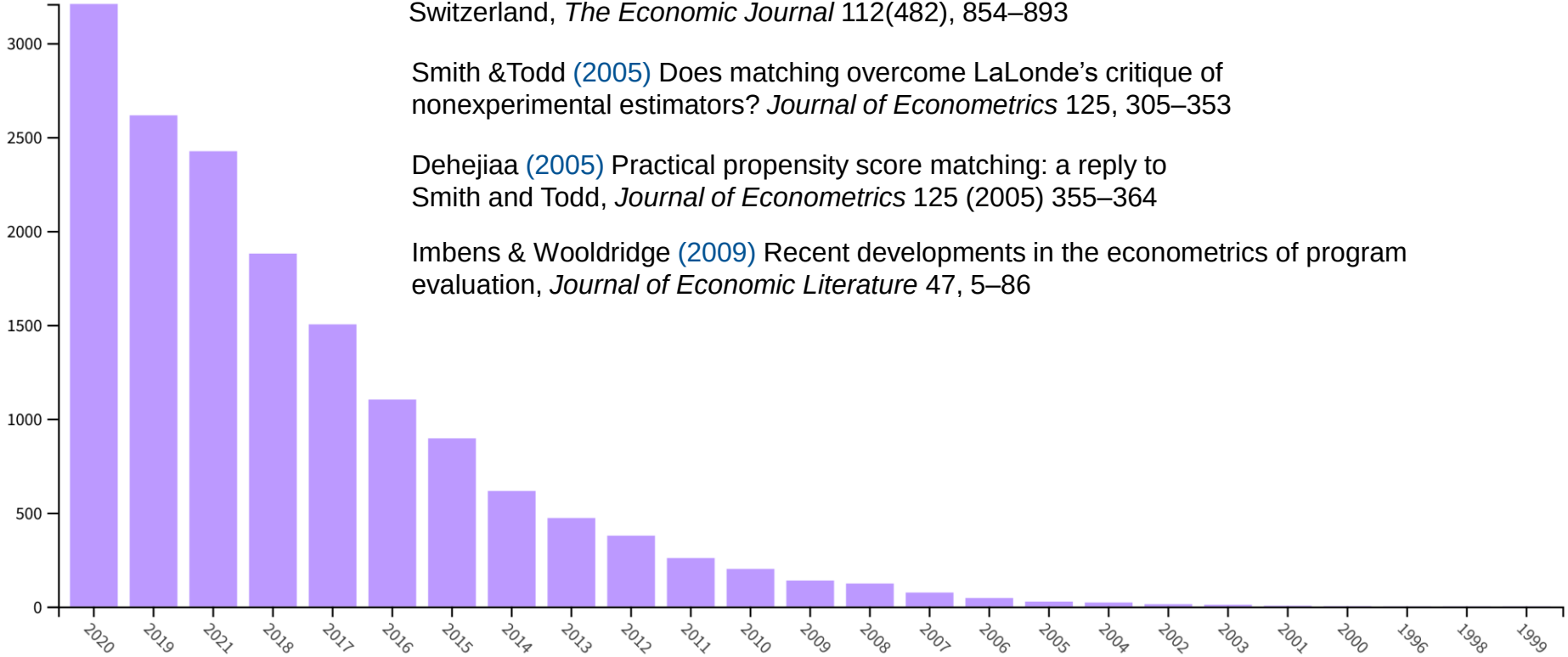
- In our case
  - No information about application process or selection
  - No information about rejected applicants
  - In many cases, we do not observe the firms *before* the subsidy



→ Regression discontinuity designs (RDD) or Difference-in-Differences (DiD) not implementable

→ **Matching techniques** to find suitable counter factuals

Popularity of Propensity Score Matching (all fields)



Gerfin & Lechner (2002) A microeconometric evaluation of theactive labour market policy in Switzerland, *The Economic Journal* 112(482), 854–893

Smith &Todd (2005) Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* 125, 305–353

Dehejiaa (2005) Practical propensity score matching: a reply to Smith and Todd, *Journal of Econometrics* 125 (2005) 355–364

Imbens & Wooldridge (2009) Recent developments in the econometrics of program evaluation, *Journal of Economic Literature* 47, 5–86

Research method: Propensity Score Matching

- ✓ Minimum data requirements:
  - Comprehensive data with ideally **some time-series properties**
    - **Recipients** and **non-recipients**
    - **Founder** (team) and **firm characteristics**

- ✓ Here:
  - Founding cohorts 2005-2012 (IAB/ZEW Start-up Panel)
  - Information on receipt of subsidy
  - Information on **relevant performance outcomes**: employees, turnover, R&D spending, investment, innovation success
  - Information on **relevant founder and firm characteristics** (Lechner, M., Wunsch, C., 2013. Sensitivity of matching-based program evaluations to the availability of control variables. Labour Econ. 21, 111–121)

Idea

Pick for each treated the most similar un-treated firm

Treatment group S = 1

Founders: Dr. Andreas Sichert, Dr.-Ing. Andreas Schuster, Richard Aumann
Year: 2008
Activity: waste heat recovery

Founders: Dr. Andre Lodwig, Dr. Hans Oswald
Year: 2007
Activity: audiological diagnostics

Control group S = 0

Founders: Daniel Quinger, Dr. Michael Geppert, Tobias Mayer
Year: 2008
Activity: energy storage applications

Founders: Dr. Joachim Wiest, Prof. Dr. Bernhard Wolf, Dr. Helmut Herz, Herbert Zuleger
Year: 2007
Activity: system solutions for microphysiometry

Propensity Score Matching

- **Collect attributes that predict treatment**

  Founder(s)‘ academic background, age, industry experience, entrepreneurial experience, professional experience, team (composition, e.g. gender, academic background), enture age, revenue, profit, ex-ante financing structure, patents, market penetration (e.g. export), R&D activity, capacity utilization, location characteristics

- **Estimate propensity to receive treatment**

$$\Pr(S = 1 \mid X) = G(\beta \boldsymbol{X})$$

- Calculate the Mahalanobis distance between a treated and a control observation (Gerfin & Lechner 2002)

- **For each treated firm pick closest neighbor(s)** (Rosenbaum and Rubin, 1983)

Propensity Score Matching

- **Employ a caliper to avoid "bad matches"** by imposing a threshold of the maximum distance allowed
  - Why? Closest neighbor could still be relatively far away, i.e. not similar enough

- **Combine with exact matching** (founding cohort, sector, region)
  - Why? Some comparisons may simply not be reasonable
  - **PSM** reduces complexity by matching on a single score, but this has drawbacks compared to (Coarsened) **Exact Matching**

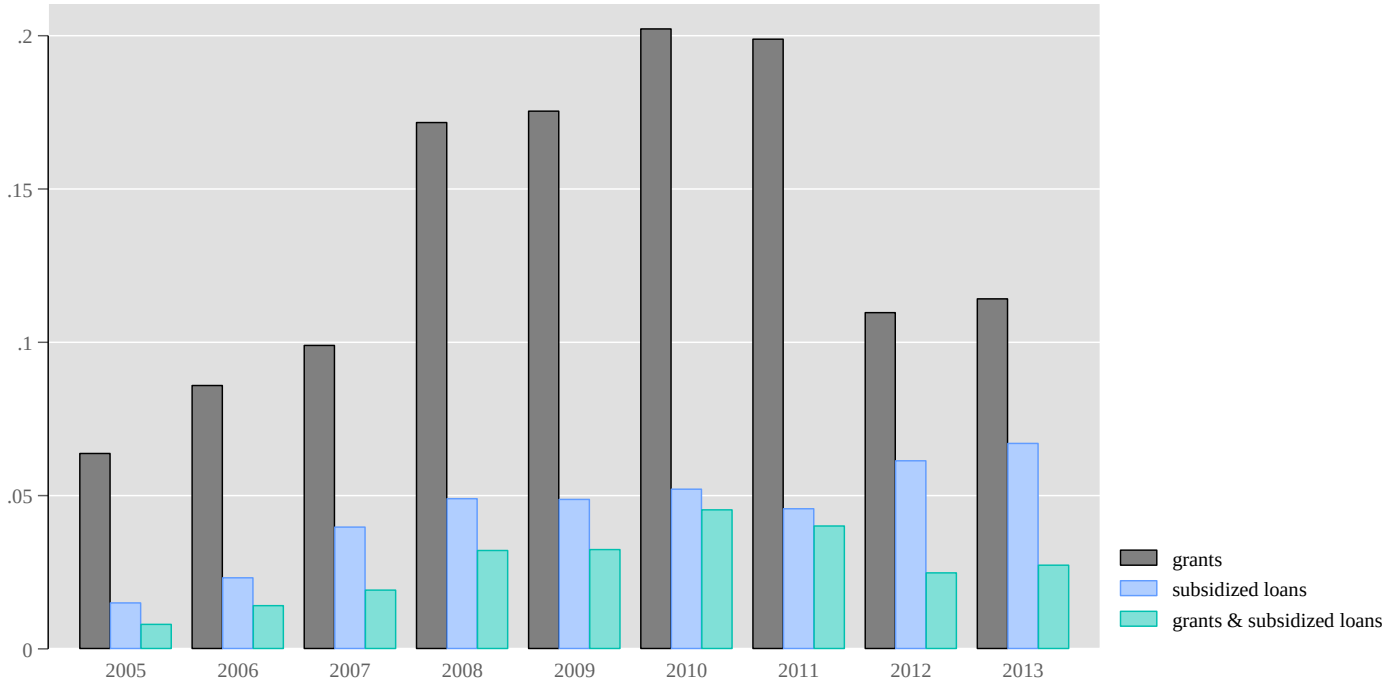
- Conditional independence assumption (Rubin 1977):

$$E(Y|S=1,X) = E(Y|S=0,X)$$

- Estimate the **average treatment effect** as:

$$\alpha^{TT} = E(Y^T|S=1,X=x) - E(Y^C|S=0,X=x)$$

The data


.2
.15
.1
.05
0
2005
2006
2007
2008
2009
2010
2011
2012
2013
grants
subsidized loans
grants & subsidized loans

Characteristics before matching

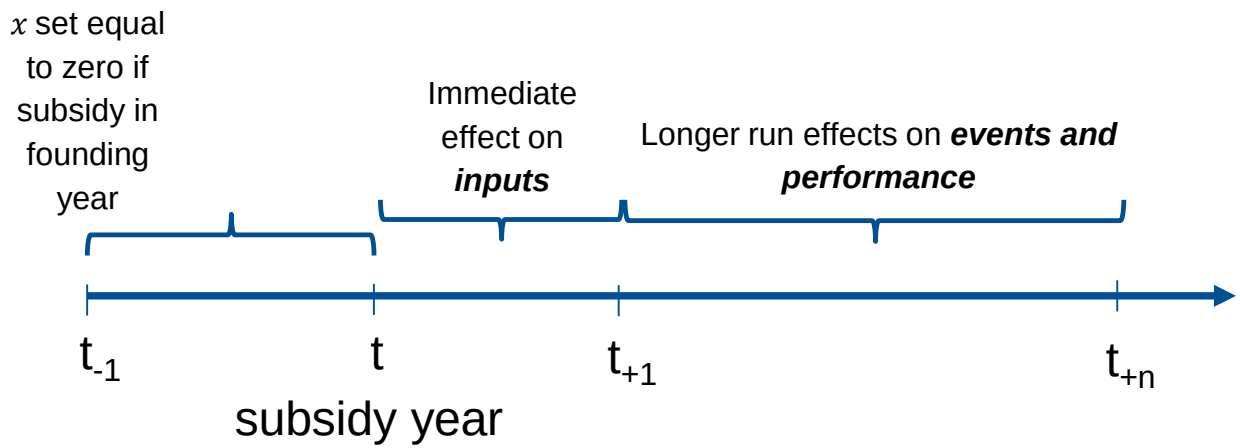
Variables	Potential Control Group N= 4,057 Mean	SD	Treatment Group N= 822 Mean	SD	t-test*
Founder characteristics					
University	0.332	0.471	0.341	0.474	0.623
Vocational training	0.190	0.392	0.178	0.382	0.415
Master craftsman	0.221	0.415	0.210	0.408	0.502
Founder age	44.746	10.079	44.797	9.211	0.893
Industry experience	17.239	9.660	16.591	9.021	**0.076**
Entrepreneurial experience	0.458	0.498	0.416	0.493	**0.029**
Bankruptcy experience	0.072	0.258	0.071	0.256	0.886
Opportunity driven	0.779	0.415	0.755	0.430	0.138
Female	0.128	0.334	0.108	0.311	0.120

Variables	Potential Control Group N= 4,057 Mean	SD	Treatment Group N= 822 Mean	SD	t-test*
Firm characteristics					
Team	0.374	0.484	0.384	0.487	0.589
Start-up age$_{t-1}$	2.796	1.655	2.331	1.524	**0.000**
Limited liability	0.520	0.500	0.557	0.497	**0.051**
ln(Tangible assets)	5.814	4.468	5.733	4.551	0.634
Patent stock	0.183	3.672	0.130	1.586	0.688
Export activity$_{t-1}$	0.151	0.358	0.212	0.409	**0.000**
Capacity utilization$_{t-1}$	84.988	29.852	88.104	28.668	**0.006**
East Germany location	0.131	0.338	0.210	0.408	**0.000**
ln(R&D-Expenditure)$_{t-1}$	1.832	3.946	2.714	4.676	**0.000**
ln(Employees)$_{t-1}$	0.865	0.753	0.833	0.934	0.284
ln(Revenue)$_{t-1}$	7.695	5.620	6.451	5.973	**0.000**
ln(Tangible Investment)$_{t-1}$	5.471	4.817	5.017	5.052	**0.015**
Profit$_{t-1}$	17.724	105.303	6.123	94.865	**0.003**

Outcomes before matching

	Potential Control Group		Treatment Group		
Variables	Mean	SD	Mean	SD	t-test*
$(R\&D\text{-}Exp/\ Employees)_{t+1}$	402	1,029	494	1,173	**0.023**
$(R\&D\text{-}Emp/\ Employees)_{t+1}$	11.200	26.318	15.003	28.326	**0.000**
$ln(R\&D\text{-}Expenditure)_{t+1}$	3.208	4.835	4.178	5.277	**0.000**
$ln(R\&D\text{-}Employees)_{t+1}$	0.220	0.482	0.347	0.596	**0.000**
$ln(Employees)_{t+1}$	1.340	0.637	1.504	0.746	**0.000**
$ln(Revenue)_{t+1}$	11.035	3.559	11.260	3.534	**0.098**
$ln(Tangible\ Investment)_{t+1}$	6.254	4.606	7.001	4.441	**0.000**

Plausible timing?



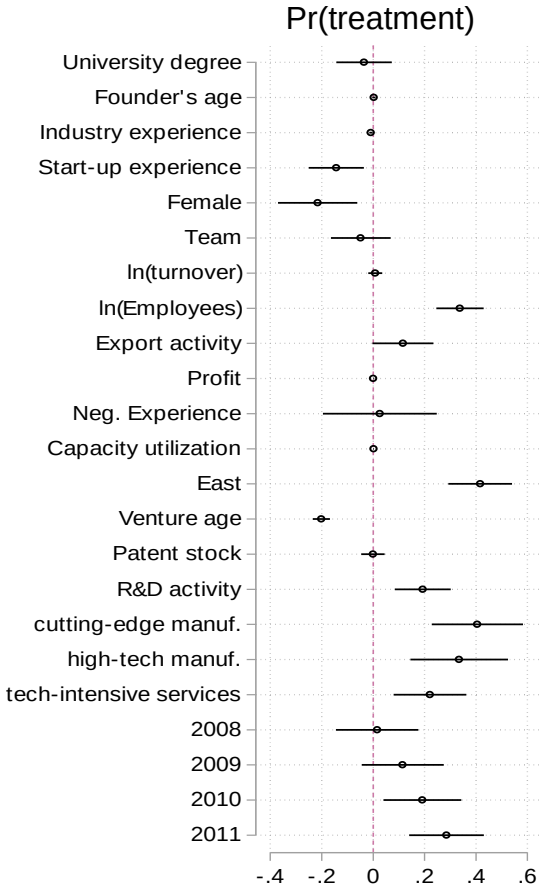
- time-variant founder & venture characteristics

$$\Pr(S = 1 \mid X = x)$$
with $S =$ **grant**

- R&D activities
- Investment
- Employment

- Innovation
- Revenue
- Employment
- Merger
- Acquisition
- Venture Capital
- Bankruptcy

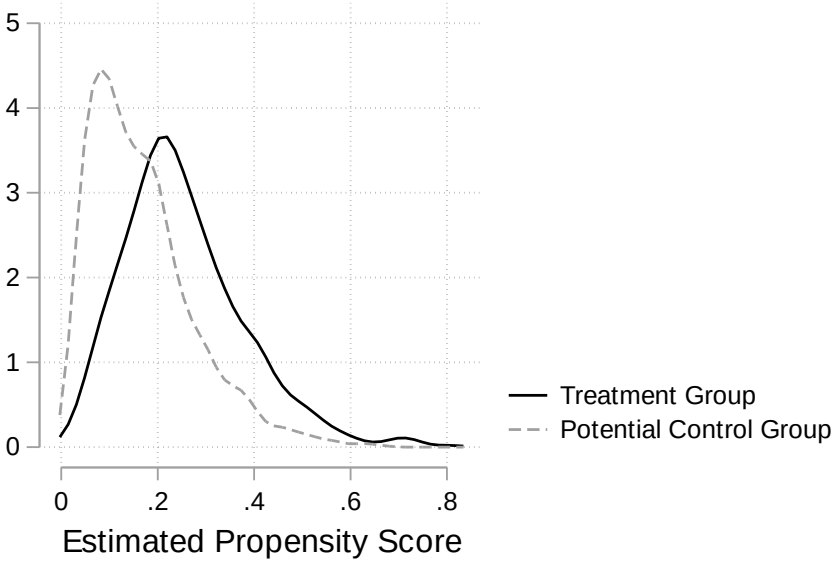
Propensity Score Estimation

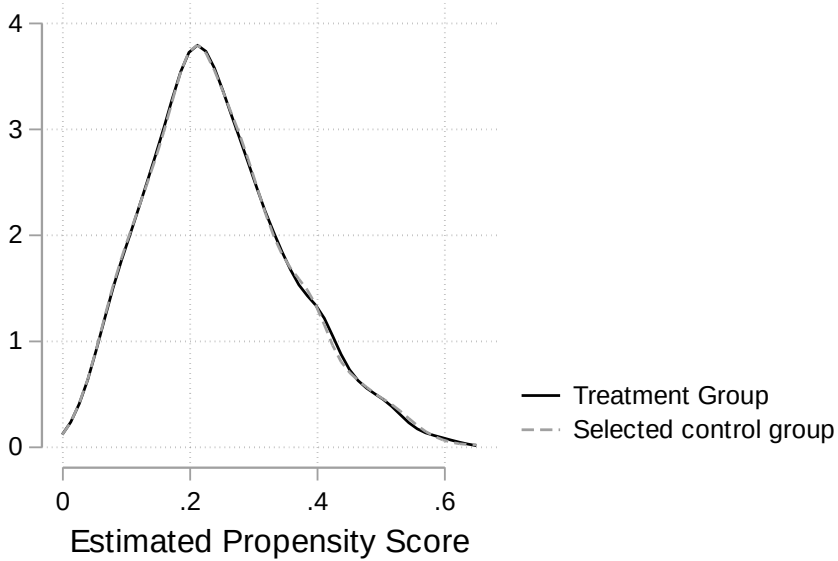


Several of the predictors explain subsidy receipt

→ Significant differences between treated and untreated firms that could also explain differences in performance

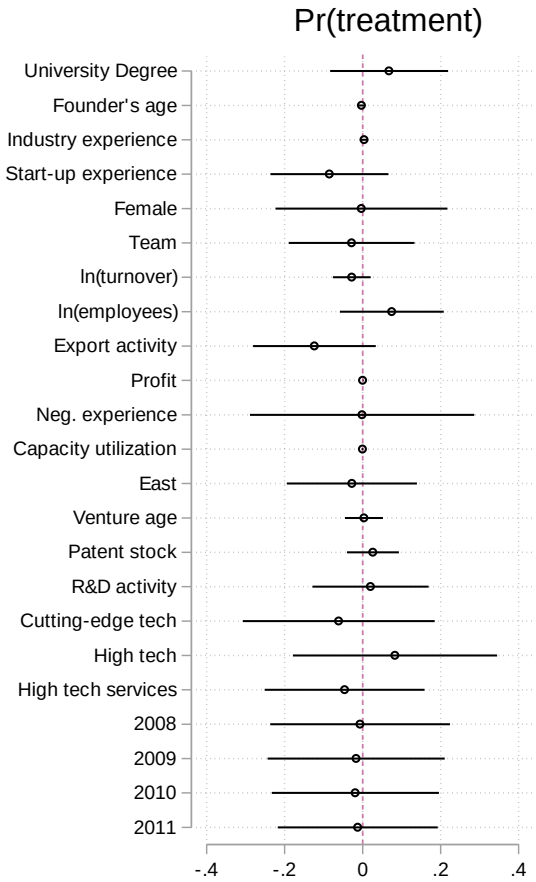
Matching outcome





Propensity score **& all covariates** should be balanced after matching!

Propensity Score Estimation (after matching)



Results

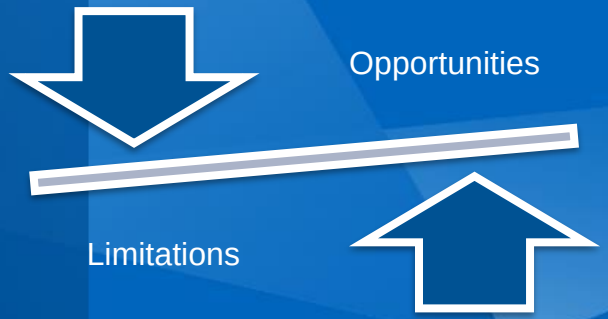
Outcomes after matching

	Selected Control Group N= 732		Treatment Group N= 732		
Outcome variables	Mean	SD	Mean	SD	t-test of mean difference
$(R\&D\text{-}Exp/\ Employees)_{t+1}$	465.111	1,070.221	464.538	1,118.555	**0.992**
$(R\&D\text{-}Emp/\ Employees)_{t+1}$	11.738	26.208	14.028	27.890	**0.106**
$ln(R\&D\text{-}Expenditure)_{t+1}$	3.624	4.979	3.800	5.124	**0.506**
$ln(R\&D\text{-}Employees)_{t+1}$	0.236	0.511	0.299	0.543	0.024
$ln(Employees)_{t+1}$	1.349	0.640	**1.428**	0.700	**0.026**
$ln(Revenue)_{t+1}$	10.791	4.019	11.132	3.546	0.085
$ln(Tangible\ Investment)_{t+1}$	5.995	4.672	6.916	4.403	0.000

Before matching:

$ln(Employees)_{t+1}$	**1.340**	**0.637**	**1.504**	**0.746**	**0.000**

Conclusions



Strengths and pitfalls I

- **Differences between groups in post treatment outcome variables can be causally attributed to the subsidy**

  if we assume that we observe all relevant differences between treated and untreated firms
- Propensity score matching → simple method to reduce selection bias

Key advantages:

- Modest data requirements
- Quasi-randomization
- Simplifies matching on many characteristics → single score

Points to remember when using matching methods

- Critically reflect the results
  - Not only propensity score should be balanced after matching, but also all predictors
  - If not the case: combine with caliper and/or exact matching
- **What *unobservable* factors may be important (but are not proxied for)?**

Addressing “matching on observables problem”

- **Combine PSM with fixed effects models** (Arkhangeslky & Imbens 2018, The role of the propensity score in fixed effect models, NBER Working Paper #24814)

Panel A: POLS (unmatched)

	VC	GVC	BA	IVC	CVC
Subsidy(t)	0.0034***	0.0028***	0.0016***	0.0011***	0.0007**
	(0.0006)	(0.0005)	(0.0004)	(0.0004)	(0.0003)
Obs.	55051	55330	55659	55589	55837

Panel B: POLS (matched)

	VC	GVC	BA	IVC	CVC
Subsidy(t)	0.0026***	0.0021***	0.0015***	0.0003	0.0007*
	(0.0008)	(0.0007)	(0.0005)	(0.0006)	(0.0004)
Obs.	24978	25104	25285	25212	25323

Panel C: Within (matched)

	VC	GVC	BA	IVC	CVC
Subsidy(t)	0.0058**	0.0044**	0.0031*	0.0013	0.0011
	(0.0025)	(0.0021)	(0.0018)	(0.0017)	(0.0010)
Groups	3953	3955	3963	3961	3961
Obs.	24978	25104	25285	25212	25323

Question here:
Are start-up subsidies causally linked to follow-on financing?

Captures influence of time-constant unobservables

Strengths and pitfalls II

- What other ***external* factors** could affect the result? Are there other “treatments“ happening?
- How plausible is the **timing**?

- How sensitive is the *size of treatment effect* to the given sample?
  - How plausible is the effect size? How meaningful is the average treatment effect?
  - What about treated firms that were not “matched”?
  - Discrete versus continuous treatment

- How sensitive are the results to using different matching approaches:
  - E.g. **Coarsened Exact Matching or Kernel matching** *(see e.g. Todd P.E. (2010) Matching Estimators. In: Durlauf S.N., Blume L.E. (eds) Microeconometrics. The New Palgrave Economics Collection)*

- Methods for **continuous treatments** → generalized propensity scores *(see e.g. Imbens 2000, The role of the propensity score in estimating dose-response functions, Biometrika, 87(3), 706–710)*

Thank you!

Contact details

Professor Dr. Hanna Hottenrott
Professorship Economics of Innovation

Technical University of Munich
TUM School of Management

Arcisstrasse 21
80333 Munich

hanna.hottenrott@tum.de
www.professors.wi.tum.de/eoi