

# An Approach Combining Theory, Selection and Empirics Provides Evidence of Regularities in the Bias of Observational Methods

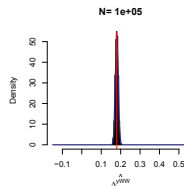
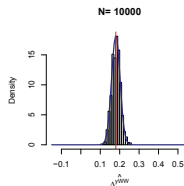
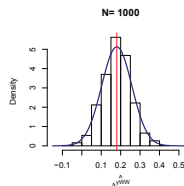
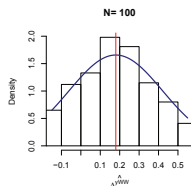
**Sylvain Chabé-Ferret**

Toulouse School of Economics, INRA and IAST

November 29, 2018

# Observational Methods are unreliable

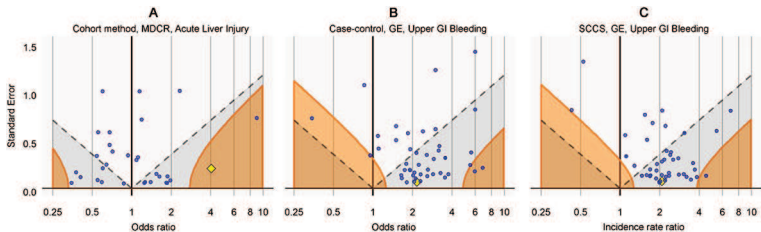
- ▶ Reliable Observational Methods would enable causal inference at a low cost
- ▶ Problem: Observational Methods are unreliable because their bias is unknown
- ▶ The bias of RCTs is much better known



# Making Observational Methods reliable

Characterizing the distribution of the bias of Observational Methods would make them more reliable

- ▶ Correct confidence intervals including uncertainty about bias
- ▶ Choice of method ex ante

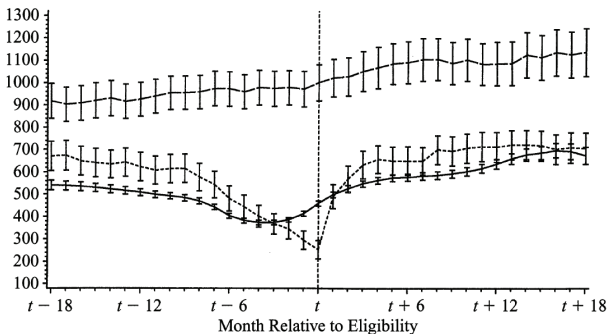


## My proposal: combining 3 steps

1. Derive general properties using stylized theoretical models
2. Derive quantitative properties using simulations of calibrated models
3. Estimate the bias of Observational Methods from real data

# An example: DID Matching and JTPs

- ▶ DID Matching emerges as least biased method when compared with RCTs
- ▶ Intuitive explanation: DID captures fixed effects, Matching captures transitory shocks



# Why does DID Matching work?

## 1. Theoretical results

- ▶ Intuitive story is wrong
- ▶ Fallacy of alignment bias
- ▶ Symmetric DID undoes time varying selection bias

## 2. Simulations

- ▶ Fallacy of alignment bias is sizable
- ▶ Symmetric DID resists to failure of symmetry

## 3. Empirical results confirm predictions

## Literature

- ▶ Overall Approach: Chabé-Ferret (2015), Hill (2008), Eckles and Bakshy (2017)
- ▶ Theory: Heckman (1978), Nickell (1981), Heckman and Robb (1985)
- ▶ Simulations: Heckman, LaLonde, and Smith (1999), Huber, Lechner, and Wunsch (2013), Hatfield and Daw (2018)
- ▶ Empirics:
  - ▶ Within Study Comparisons: LaLonde (1986), Fraker and Maynard (1987), Glazerman et al (2003), Wong et al (2017), Chabé-Ferret et al (2018)
  - ▶ Sensitivity analysis: Ashenfelter and Card (1985), Anderson et al (2013), Caliendo et al (2014)
  - ▶ Meta-Analysis: Benson and Hartz (2000), Hemkens et al (2016)
  - ▶ Placebo-based approach: Schuemie et al (2013,2018)

# Outline of the talk

Theoretical results

Simulations

Experimental evidence

Conclusion



# Outline of the talk

Theoretical results

Simulations

Experimental evidence

Conclusion

# The model

$$Y_{i,t}^0 = \delta_t + \mu_i + U_{it}$$

$$\text{with } U_{i,t} = \rho U_{i,t-1} + v_{i,t}$$

$$D_{i,k}^* = \theta_i + \gamma Y_{i,k-1}^0 + \gamma^f v_{i,k}$$

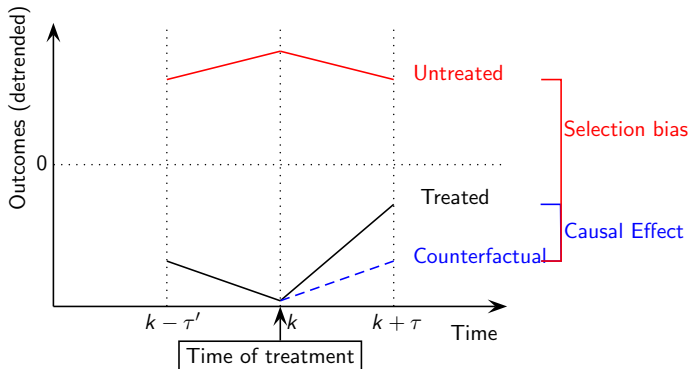
$$D_{i,t} = \mathbb{1}[t \geq k] \mathbb{1}[D_{i,k}^* \geq 0]$$

$$\mathbb{E}[Y_{i,t}^0 | D_{i,k}^*, Y_{i,k-1}] \text{ linear}$$

$$\text{Corr}(\mu_i, \theta_i) = \rho_{\theta, \mu}$$

- ▶ Both permanent (when  $\rho_{\theta, \mu} \neq 0$ ) and transitory (when  $\gamma \neq 0$  and  $\rho \neq 0$ ) confounders
- ▶ Both limited (when  $\gamma^f = 0$ ) and full (when  $\gamma^f = \frac{\gamma}{\rho}$ ) information
- ▶ Both self-selection in a JTP and eligibility criteria

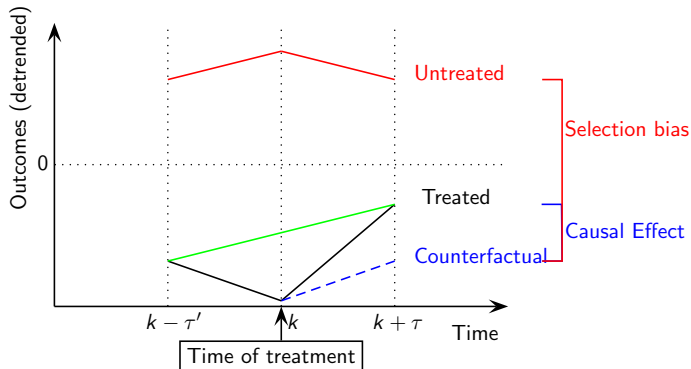
# The estimators and their asymptotic biases



Estimators

Bias

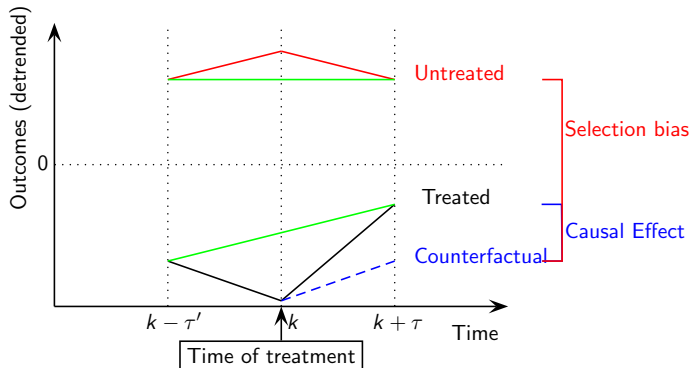
# The estimators and their asymptotic biases



Estimators

Bias

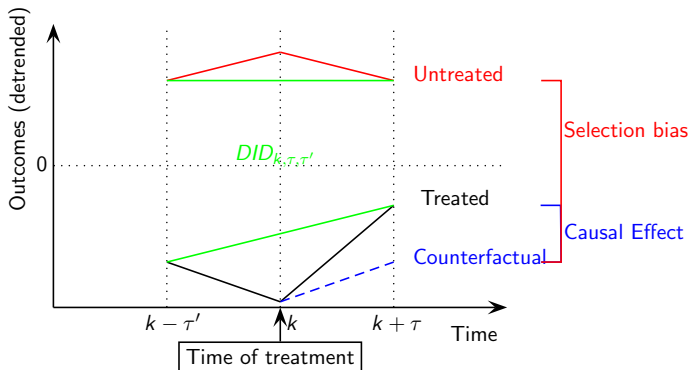
# The estimators and their asymptotic biases



Estimators

Bias

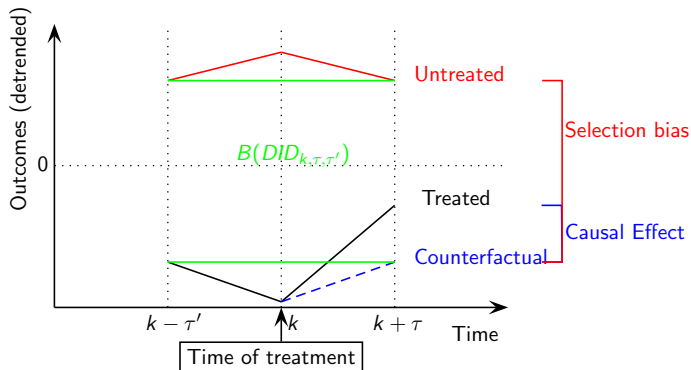
# The estimators and their asymptotic biases



Estimators

Bias

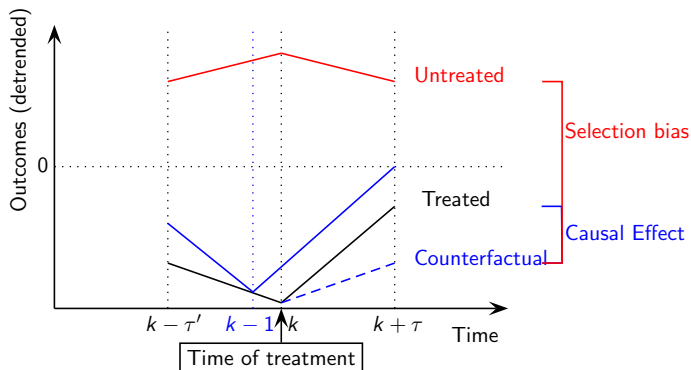
# The estimators and their asymptotic biases



Estimators

Bias

# The estimators and their asymptotic biases

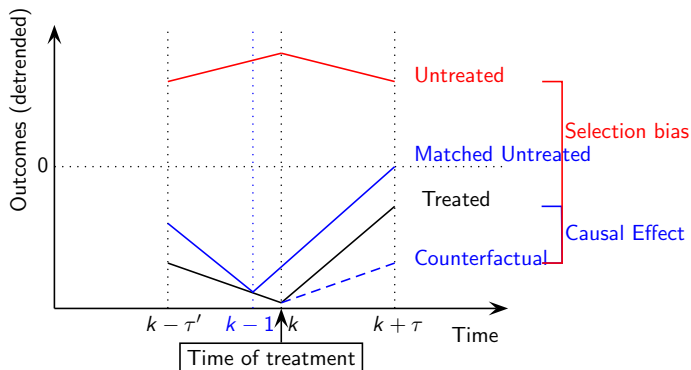


Estimators

Bias



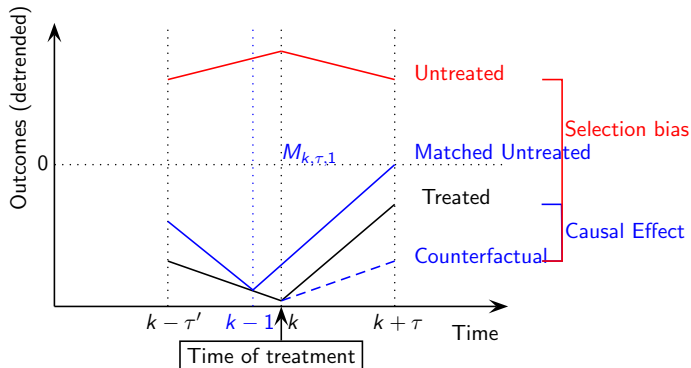
# The estimators and their asymptotic biases



Estimators

Bias

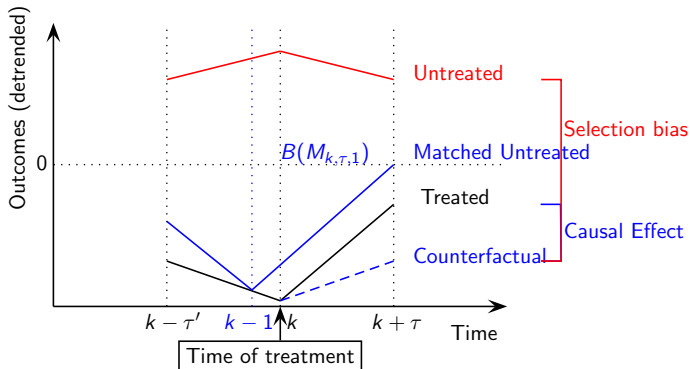
# The estimators and their asymptotic biases



Estimators

Bias

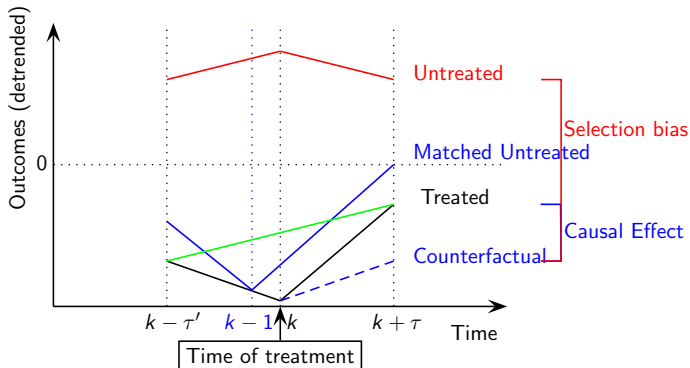
# The estimators and their asymptotic biases



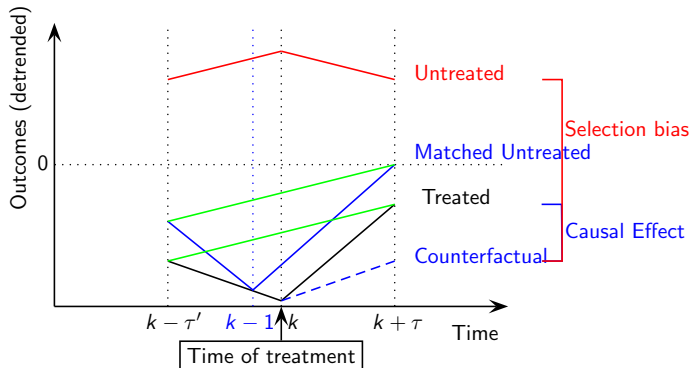
Estimators

Bias

# The estimators and their asymptotic biases



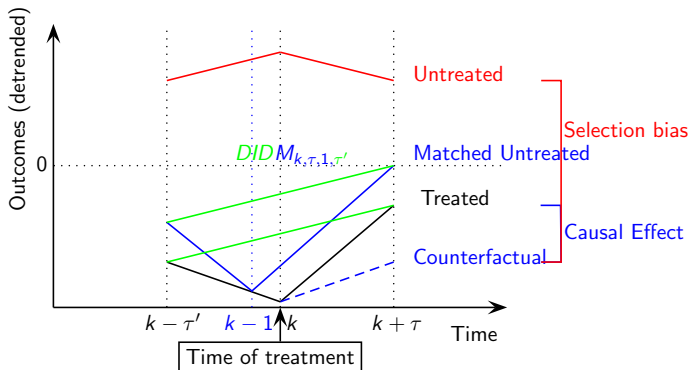
# The estimators and their asymptotic biases



Estimators

Bias

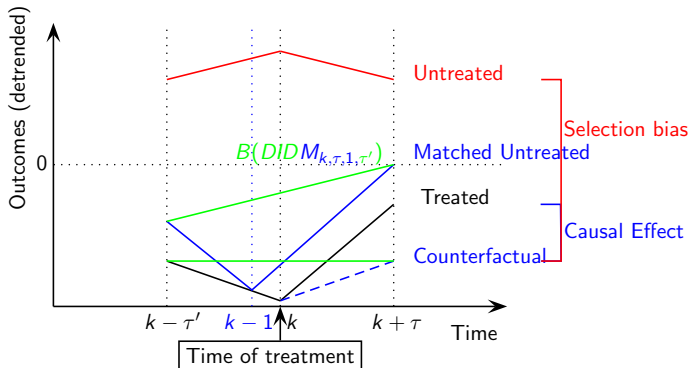
# The estimators and their asymptotic biases



Estimators

Bias

# The estimators and their asymptotic biases



Estimators

Bias

# Conditions for consistency

## Definition (Consistency)

An estimator  $E_{k,\tau,\tau'}$  is consistent  $\Leftrightarrow B(E_{k,\tau,\tau'}) = 0$ ,

- ▶  $\forall k \geq 3$
- ▶  $\forall \tau \geq 1$
- ▶ For  $\tau' \in [2 \dots \max\{2, k - 1\}]$  with either:
  - ▶  $\forall \tau'$
  - ▶ For  $\tau' = f(\tau)$



# Main theoretical result

## Theorem (Consistency of M, DID and DIDM)

$\forall k \geq 3, \forall \tau \geq 1$ , for  $\tau' \in [2 \dots \max\{2, k - 1\}]$ ,

(i)  $B(M_{k,\tau,1}) = 0 \Leftrightarrow \rho_{\theta,\mu} = \gamma^f = 0$  or  $\rho_{\theta,\mu} = \rho = 0$ .

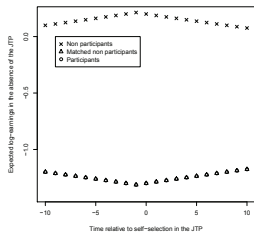
(ii)  $B(DID_{k,\tau,\tau'}) = 0 \Leftrightarrow$

$$\begin{cases} \rho = 0 \text{ or } \gamma = \gamma^f = 0 \\ \text{or} \\ \sigma_{U_0}^2 = \frac{\sigma^2}{1-\rho^2} \text{ and } \tau' = \tau + 1 + \frac{\ln \bar{\rho}}{\ln \rho}, \text{ with } \bar{\rho} = \rho + \frac{\gamma^f}{\gamma}(1 - \rho^2). \end{cases}$$

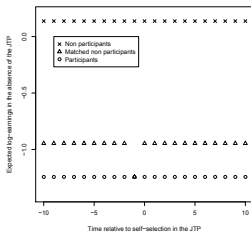
(iii)  $B(DIDM_{k,\tau,1,\tau'}) = 0 \Leftrightarrow$

$$\begin{cases} \rho_{\theta,\mu} = \gamma^f = 0 \text{ or } \rho = 0 \\ \text{or} \\ \sigma_{U_0}^2 = \frac{\sigma^2}{1-\rho^2} \text{ and } \tau' = \tau + 1 + \frac{\ln \rho^*}{\ln \rho}, \text{ with } \rho^* = \rho - \gamma^f \sigma^2 \frac{\frac{\sigma^2}{1-\rho^2} + \sigma_\mu^2}{\rho \sigma_\mu a \frac{\sigma^2}{1-\rho^2}}. \end{cases}$$

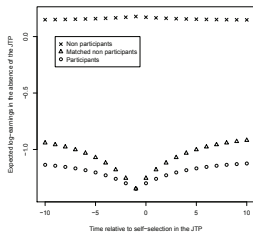
# Illustration



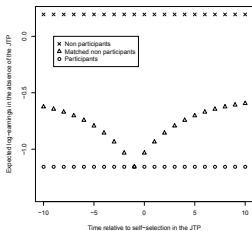
(a)  $\rho_{\theta, \mu} = \gamma^f = 0$



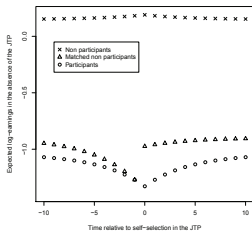
(b)  $\gamma = \gamma^f = \rho = 0$



(c)  $\rho_{\theta, \mu}, \rho, \gamma \neq 0, \gamma^f = 0$



(d)  $\rho \neq 0, \gamma = \gamma^f = 0$



(e)  $\rho_{\theta, \mu}, \rho, \gamma \neq 0, \gamma^f = \frac{\gamma}{\rho}$

# Outline of the talk

Theoretical results

**Simulations**

Experimental evidence

Conclusion

## The model

$$Y_{i,t}^0 = a + b \frac{18+t}{10} + c \left( \frac{18+t}{10} \right)^2 + (\delta + r_t d) E_i + \mu_i + \beta_i t + U_{it}$$

$$\text{with } U_{i,t} = \rho U_{i,t-1} + m_1 v_{i,t-1} + m_2 v_{i,t-2} + v_{i,t}$$

$$D_{i,k}^{*\iota} = \frac{\alpha_j}{r} - c_{i,k} - \mathbb{E} [Y_{i,k}^0 | \mathcal{I}_{i,k}^\iota]$$

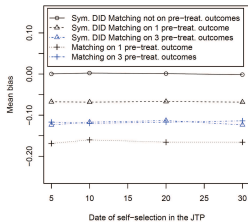
$$\text{with } c_{i,k} = c_i + \beta_x E_i - \left( a + b \frac{18+k}{10} + c \left( \frac{18+k}{10} \right)^2 + (\delta + r_k d) E_i \right)$$

- ▶  $\iota = f$ : full information
- ▶  $\iota = l$ : limited information
- ▶  $\iota = b$ : Bayesian updating

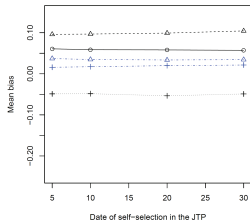
## Parameterization

	RIP (MaCurdy, 1982)	HIP (Guvenen, 2007, 2009)
$\rho$	0.99	0.821
$m_1$	-0.4	0
$m_2$	-0.1	0
$\sigma^2$	0.055	0.055
$\sigma_\mu^2$	0	0.022
$\sigma_\beta^2$	0	0.00038
$\sigma_{\mu,\beta}$	0	-0.002

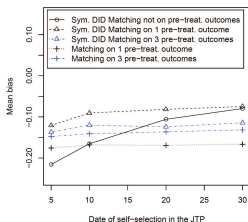
# Results: RIP



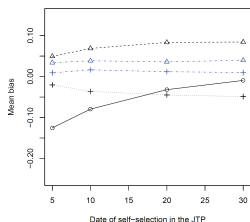
(a) Full Info, Long Run



(b) Limited Info, Long Run

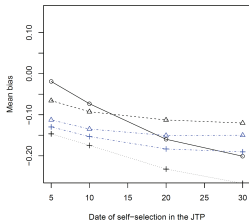


(c) Full Info, Short Run

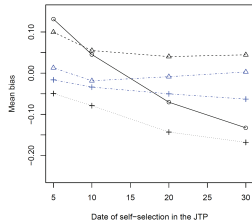


(d) Limited Info, Short Run

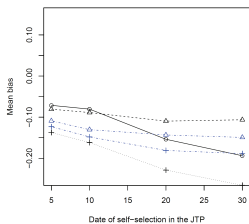
# Results: HIP



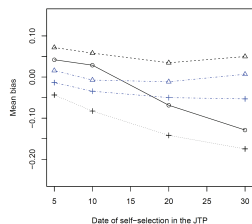
(a) Full Info, Long Run



(b) Bayes, Long Run



(c) Full Info, Short Run



(d) Bayes, Short Run

# Outline of the talk

Theoretical results

Simulations

**Experimental evidence**

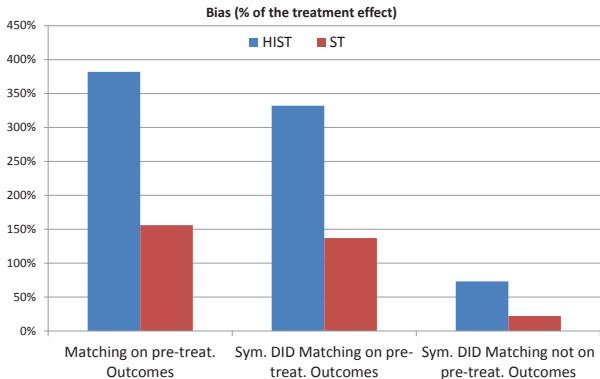
Conclusion



# Evidence on the bias of observational methods

- ▶ Within Study Comparisons
- ▶ Between Study Comparisons
- ▶ Within Study Sensitivity Analysis

# Results of Within Study Comparisons for JTPs



HIST

AC

# Meta-analysis of Within Study Comparisons for JTPs

Glazerman, Levy, and Myers (2003)

RESULTS SHOWING THE EFFECT OF NONEXPERIMENTAL APPROACH ON BIAS IN EARNINGS IMPACTS

Explanatory Variable	Model Specification						
	1	2	3	4	5	6	
Intercept	4,467*** (657)		4,687*** (1,231)		5,775*** (1,120)		
Statistical method							
Regression	-1,583** (729)	-1,516** (705)	-1,476** (675)	-1,416** (706)	-3,225*** (1,195)	-3,572*** (1,284)	
Matching	-478* (715)	-1,268** (794)	-807* (692)	-1,427*** (799)	-2,463* (1,375)	-3,178** (1,508)	
Regression × Matching			951 (1,422)	1,320 (1,484)			
Difference-in-Differences	-1,874 (763)	-1,596 (816)	-1,859 (718)	-1,568 (813)	-3,532*** (1,253)	-3,231** (1,336)	
Regression × Differences-in-Differences				2,325 (1,455)	2,676* (1,600)		
Matching × Differences-in-Differences					1,889 (1,477)	1,774 (1,547)	
Selection correction	2,508* (1,248)	2,376 (1,305)	4,619 (1,048)	2,441 (1,299)	3,291*** (1,163)	3,072** (1,284)	
Comparison group strategy							
Geographic match			-357 (973)	-646 (1,182)	-673 (957)	-581 (1,160)	
National data set			1,145 (1,062)	1,695 (1,536)	915 (1,043)	1,668 (1,479)	
Control group from another site			-1,762 (1,011)		NA	-2,124** (995)	-1,346 (2,863)
Study dummies included	No	Yes	No	Yes	No	Yes	

NOTE: Dependent variable is the absolute value of the bias in annual earnings, expressed in 1996 dollars. Standard errors are in parentheses. All explanatory variables are dummy variables. Sample size is 69 bias estimate types.

\*Significantly different from zero at the .10 level. \*\*Significantly different from zero at the .05 level. \*\*\*Significantly different from zero at the .01 level, two-tailed test.

## Between Study Analysis of JTPs (Example)

Card, Kluge, and Weber (2015)

---

	Number Est's. (1)	Median Sample Size (2)	Percent RCT's (3)	Mean Program Effect on Prob. Emp. ( $\times 100$ )		
				Short Term (4)	Medium Term (5)	Longer Term (6)
<u>By Evaluation Design:</u>						
Experimental	166	1,471	100.0	4.4 (28)	2.5 (25)	0.5 (15)
Non-experimental	691	16,000	0.0	0.9 (113)	6.0 (118)	11.0 (53)

# Outline of the talk

Theoretical results

Simulations

Experimental evidence

**Conclusion**

## Conclusions so far

- ▶ The proposed approach seems to work
- ▶ Combining DID and Matching on pre-treatment outcomes is not a great idea

## To do for this paper

- ▶ Simulations with levels
- ▶ Re-analyze JTPA data
- ▶ Nickell bias?

# Further research

## 1. Bias of OM for JTPs

- ▶ One prediction to be tested: controlling for more pre-treatment outcomes could make things better
- ▶ Two empirical results to explain
  - ▶ Conditioning for labor market transitions improves matching
  - ▶ The bias of observational methods increases with time after treatment

## 2. Bias of OM in other applications

- ▶ Collect more estimates of bias
  - ▶ Within Study Comparisons using RCTs with imperfect compliance (Chabé-Ferret et al, 2018)
  - ▶ Test validity of pseudo-experiments
- ▶ Develop simulations and theories
- ▶ Put information into accessible database (SKY: Social Science Knowledge Accumulation Initiative)



ANDERSSON, F., H. J. HOLZER, J. I. LANE, D. ROSENBLUM, AND J. SMITH (2013): "Does Federally-Funded Job Training Work? Nonexperimental Estimates of WIA Training Impacts Using Longitudinal Data on Workers and Firms," Working Paper 19446, National Bureau of Economic Research.

ASHENFELTER, O., AND D. CARD (1985): "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *The Review of Economic Statistics*, 67(4), 648–660.

BENSON, K., AND A. J. HARTZ (2000): "A Comparison of Observational Studies and Randomized, Controlled Trials," *New England Journal of Medicine*, 342(25), 1878–1886, PMID: 10861324.

CALIENDO, M., R. MAHLSTEDT, AND O. A. MITNIK (2014): "Unobservable, but Unimportant? The Influence of Personality Traits (and Other Usually Unobserved Variables) for the Evaluation of Labor Market Policies," IZA Discussion Papers 8337, Institute for the Study of Labor (IZA).

CARD, D., J. KLUVE, AND A. WEBER (2015): "What Works?"

A Meta Analysis of Recent Active Labor Market Program Evaluations,” Working Paper 21431, National Bureau of Economic Research.

CHABÉ-FERRET, S. (2015): “Analysis of the Bias of Matching and Difference-In-Difference under Alternative Earnings and Selection Processes,” *Journal of Econometrics*, 185(1), 110–123.

ECKLES, D., AND E. BAKSHY (2017): “Bias and high-dimensional adjustment in observational studies of peer effects,” *ArXiv e-prints*.

FRAKER, T., AND R. MAYNARD (1987): “The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programs,” *The Journal of Human Resources*, 22(2), 194–227.

GLAZERMAN, S., D. M. LEVY, AND D. MYERS (2003): “Nonexperimental Versus Experimental Estimates of Earnings Impacts,” *The ANNALS of the American Academy of Political and Social Science*, 589(1), 63–93.

GUVENEN, F. (2007): “Learning Your Earning: Are Labor Income

Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687 – 712.

——— (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

HATFIELD, L., AND J. DAW (2018): "Matching and Regression to the Mean in Difference-in-Differences Analysis," *HSR*.

HECKMAN, J. J. (1978): "Longitudinal Studies in Labor Economics: A Methodological Review," Mimeo, University of Chicago.

HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter, and D. Card, vol. 3, chap. 31, pp. 1865–2097. Elsevier, North Holland.

HECKMAN, J. J., AND R. ROBB (1985): "Alternative Methods for Evaluating the Impact of Interventions," in *Longitudinal Analysis of Labor Market Data*, ed. by J. J. Heckman, and B. Singer, pp. 156–245. Cambridge University Press, New-York.

- HEMKENS, L. G., D. G. CONTOPOULOS-IOANNIDIS, AND J. P. A. IOANNIDIS (2016): "Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey," *BMJ*, 352.
- HUBER, M., M. LECHNER, AND C. WUNSCH (2013): "The Performance of Estimators Based on the Propensity Score," *Journal of Econometrics*, 175(1), 1 – 21.
- LALONDE, R. J. (1986): "Evaluating the Econometric Evaluation of Training Programs with Experimental Data," *American Economic Review*, 76, 604–620.
- MACURDY, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18(1), 83–114.
- NICKELL, S. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49(6), 1417–1426.
- SCHUEMIE, M. J., G. HRIPCSAK, P. B. RYAN, D. MADIGAN, AND M. A. SUCHARD (2018): "Empirical confidence interval calibration for population-level effect estimation studies in

observational healthcare data," *Proceedings of the National Academy of Sciences*, 115(11), 2571–2577.

SCHUEMIE, M. J., P. B. RYAN, W. DUMOUCHEL, M. A. SUCHARD, AND D. MADIGAN (2013): "Interpreting observational studies: why empirical calibration is needed to correct p-values," *Statistics in Medicine*, 33(2), 209–218.

WONG, V. C., J. C. VALENTINE, AND K. MILLER-BAINS (2017): "Empirical Performance of Covariates in Education Observational Studies," *Journal of Research on Educational Effectiveness*, 10(1), 207–236.

## Estimators: definition

$$M_{k,\tau,1} = \mathbb{E} [\mathbb{E} [Y_{i,k+\tau} | D_{i,k} = 1, Y_{i,k-1}] \\ - \mathbb{E} [Y_{i,k+\tau} | D_{i,k} = 0, Y_{i,k-1}] | D_{i,k} = 1]$$

$$DID_{k,\tau,\tau'} = \mathbb{E} [Y_{i,k+\tau} - Y_{i,k-\tau'} | D_{i,k} = 1] \\ - \mathbb{E} [Y_{i,k+\tau} - Y_{i,k-\tau'} | D_{i,k} = 0]$$

$$DIDM_{k,\tau,1,\tau'} = \mathbb{E} [\mathbb{E} [Y_{i,k+\tau} - Y_{i,k-\tau'} | D_{i,k} = 1, Y_{i,k-1}] \\ - \mathbb{E} [Y_{i,k+\tau} - Y_{i,k-\tau'} | D_{i,k} = 0, Y_{i,k-1}] | D_{i,k} = 1].$$

Back

## Bias: definition

$$B(M_{k,\tau,1}) = \mathbb{E} \left[ \mathbb{E} \left[ Y_{i,k+\tau}^0 \mid D_{i,k} = 1, Y_{i,k-1}^0 \right] \right. \\ \left. - \mathbb{E} \left[ Y_{i,k+\tau}^0 \mid D_{i,k} = 0, Y_{i,k-1}^0 \right] \mid D_{i,k} = 1 \right]$$

$$B(DID_{k,\tau,\tau'}) = \mathbb{E} \left[ Y_{i,k+\tau}^0 - Y_{i,k-\tau'}^0 \mid D_{i,k} = 1 \right] \\ - \mathbb{E} \left[ Y_{i,k+\tau}^0 - Y_{i,k-\tau'}^0 \mid D_{i,k} = 0 \right]$$

$$B(DIDM_{k,\tau,1,\tau'}) = \mathbb{E} \left[ \mathbb{E} \left[ Y_{i,k+\tau}^0 - Y_{i,k-\tau'}^0 \mid D_{i,k} = 1, Y_{i,k-1}^0 \right] \right. \\ \left. - \mathbb{E} \left[ Y_{i,k+\tau}^0 - Y_{i,k-\tau'}^0 \mid D_{i,k} = 0, Y_{i,k-1}^0 \right] \mid D_{i,k} = 1 \right]$$

# Consistency of Matching: sketch of proof

By linearity of conditional expectations:

$$\mathbb{E} [Y_{i,t}^0 | D_{i,k}^*, Y_{k-1}^0] = \mathbb{E} [Y_{i,t}^0] + \theta_{Y_{k+\tau}^0, D_k^*} (D_{i,k}^* - \mathbb{E} [D_{i,k}^*]) \\ + \theta_{Y_{k+\tau}^0, Y_{k-1}^0} (Y_{i,k-1}^0 - \mathbb{E} [Y_{i,k-1}^0])$$

$$\theta_{Y_{k+\tau}^0, D_k^*} = \frac{\overbrace{\sigma_{Y_{k+\tau}, D_k^*}^2 \sigma_{Y_{k-1}}^2 - \sigma_{Y_{k-1}, D_k^*} \sigma_{Y_{k-1}, Y_{k+\tau}}}^{\text{num}_{k,\tau}}}{\sigma_{D_k^*}^2 \sigma_{Y_{k-1}}^2 - \sigma_{Y_{k-1}, D_k^*}^2}$$

$$B(M_{k,\tau,1}) = \theta_{Y_{k+\tau}^0, D_k^*} \mathbb{E} [\mathbb{E} [D_{i,k}^* | D_{i,k} = 1, Y_{k-1}^0] - \mathbb{E} [D_{i,k}^* | D_{i,k} = 0, Y_{k-1}^0]]$$

$$\text{num}_{k,\tau} = \underbrace{\rho^\tau \left( \gamma^f \sigma^2 \left[ \sigma_\mu^2 + \sigma_{U_{k-1}}^2 \right] - \rho_{\theta,\mu} \sigma_\theta \sigma_\mu \rho \sigma_{U_{k-1}}^2 \right)}_{F(k)} + \underbrace{\rho_{\theta,\mu} \sigma_\theta \sigma_\mu \sigma_{U_{k-1}}^2}_{G(k)}$$



## Consistency of DID Matching: sketch of proof

$$B(DIDM_{k,\tau,1,\tau'}) = 0 \Leftrightarrow \text{num}_{k,\tau} - \text{num}_{k,-\tau'} = 0$$

$$\text{num}_{k,\tau} - \text{num}_{k,-\tau'} = B(\tau, \tau') + \rho^{2(k-\tau')} C(\tau, \tau')$$

$$B(\tau, \tau') = H(\tau') + \rho^\tau I$$

$$H(\tau') = \sigma_\mu \rho_{\theta,\mu} \sigma_\theta \rho^{\tau'-1} \frac{\sigma^2}{1-\rho^2}$$

$$I = \frac{\sigma^2}{1-\rho^2} (\gamma^f \sigma^2 - \rho \sigma_\mu \rho_{\theta,\mu} \sigma_\theta) + \gamma^f \sigma^2 \sigma_\mu^2$$

$$C(\tau, \tau') = J(\tau') + \rho^\tau K(\tau')$$

$$J(\tau') = \sigma_\mu \rho_{\theta,\mu} \sigma_\theta \rho^{\tau'-1} \left( \sigma_{U_0}^2 - \frac{\sigma^2}{1-\rho^2} \right)$$

$$K(\tau') = \gamma^f \sigma^2 - \left( \sigma_{U_0}^2 - \frac{\sigma^2}{1-\rho^2} \right) \sigma_\mu \rho_{\theta,\mu} \sigma_\theta \rho$$

## Consistency of DID Matching: sketch of proof (cont'd)

$$B(\tau, f(\tau)) = \rho^\tau L(\tau)$$

$$L(\tau) = \rho \sigma_\mu \rho_{\theta, \mu} \sigma_\theta \frac{\sigma^2}{1 - \rho^2} (\rho^{f(\tau) - \tau - 2} - 1) + \gamma^f \sigma^2 \left( \frac{\sigma^2}{1 - \rho^2} + \sigma_\mu^2 \right)$$

$$C(\tau, f(\tau)) = \rho^{\tau + 2(f(\tau) - 1)} M(\tau)$$

$$M(\tau) = \left( \sigma_{U_0}^2 - \frac{\sigma^2}{1 - \rho^2} \right) N(\tau)$$

$$N(\tau) = \sigma_\mu \rho_{\theta, \mu} \sigma_\theta \rho (\rho^{-(f(\tau) + \tau)} - 1) + \gamma^f \sigma^2$$

Back

## Consistency of DID: sketch of proof

$$B(DID_{k,\tau,\tau'}) = 0 \Leftrightarrow \sigma_{Y_{k+\tau}^0, D_k^*} - \sigma_{Y_{k-\tau'}^0, D_k^*} = 0$$

$$\sigma_{Y_{k+\tau}^0, D_k^*} - \sigma_{Y_{k-\tau'}^0, D_k^*} = P(\tau, \tau') + \rho^{2(k-\tau')} Q(\tau, \tau')$$

$$P(\tau, \tau') = \gamma(\rho^{\tau+1} - \rho^{\tau'-1}) \frac{\sigma^2}{1 - \rho^2} + \gamma^f \rho^\tau \sigma^2$$

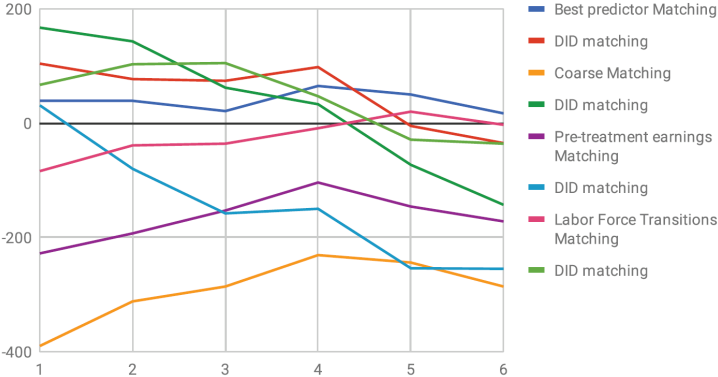
$$Q(\tau, \tau') = \gamma \left( \sigma_{U_0}^2 - \frac{\sigma^2}{1 - \rho^2} \right) (\rho^{\tau+1} \rho^{2(\tau'-1)} - \rho^{\tau'-1})$$

$$P(\tau, f(\tau)) = \rho^{\tau+1} \left( \gamma \rho \frac{\sigma^2}{1 - \rho^2} + \gamma^f \sigma^2 - \gamma \rho^{f(\tau) - \tau - 1} \frac{\sigma^2}{1 - \rho^2} \right)$$

$$Q(\tau, f(\tau)) = \rho^{\tau+1+2(f(\tau)-1)} \gamma \left( \sigma_{U_0}^2 - \frac{\sigma^2}{1 - \rho^2} \right) \left( 1 - \rho^{-f(\tau) - \tau + 1} \right).$$

# Results of HIST

## Bias of Matching and DID matching in HIST



# Results of Ashenfelter and Card

Difference Trainees-Controls in Earnings

