

“Is propensity score matching a substitute for experiments?” Is Not the Right Question

Daniel S. Nagin

Carnegie Mellon University

12th Annual Jerry Lee Conference

April 24, 2012

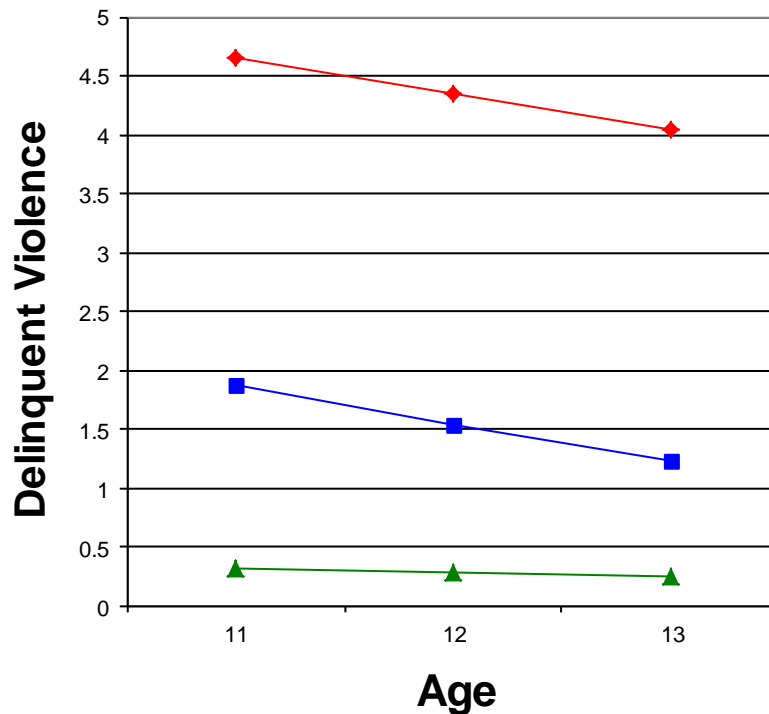
The Right Question is: What Does Propensity Score Matching Do Well?

- Transparency –Weaknesses Open to View
- Keeping Time in Order
- Helps to Insure that “Apples are being compared to apples not tomatoes.”
- Explicitly Tests for Balance on Observables
- Keeps You from Peeking

Illustrative Example

- Research Question: What is the effect of first-time gang membership on violent delinquency? Does this effect depend upon the history of violent delinquency?
- Data: Montreal Longitudinal Study (N=1027)
- Haviland, Nagin, and Rosenbaum (2007)
Psychological Methods

Baseline Measurements: Trajectories of Violent Delinquency from Age 11 to 13 for Sub-sample with NO Gang Involvement over this Period



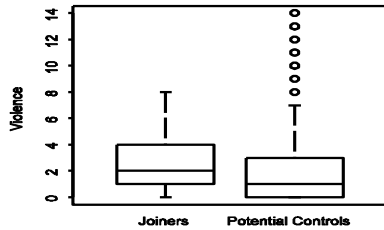
31% of Chronics
Join Gangs
at Age 14

15% of Decliners Join
Gangs
at Age 14

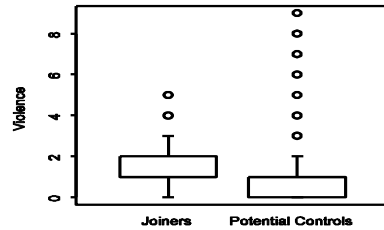
7% of Lows
Join Gangs
at Age 14

Twelve Covariates Comparing Gang Joiners at 14 with Potential Controls

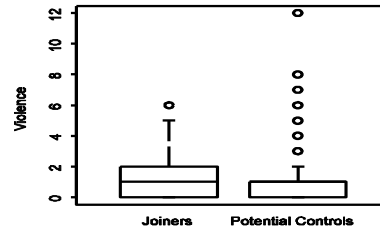
Violence Age 10, P=.0087



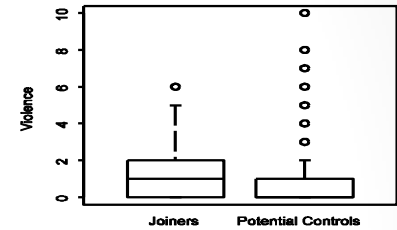
Violence Age 11, P=.0028



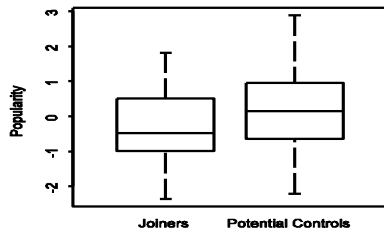
Violence Age 12, P=.044



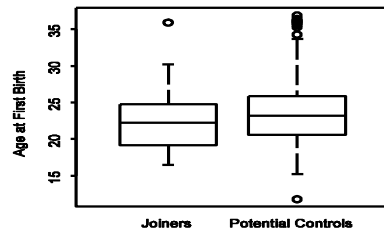
Violence Age 13, P=.0032



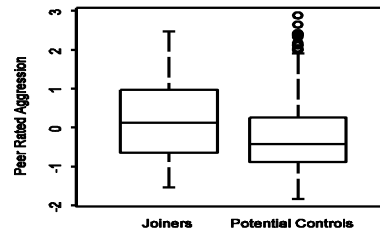
Popularity, Age 11, P=.0015



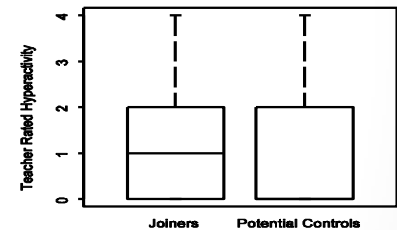
Mother's Age, P=.020



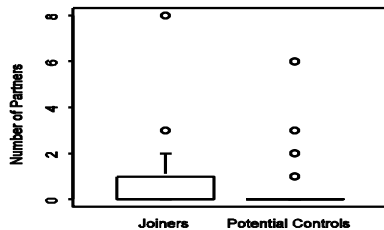
Aggression, Age 11, P=.0081



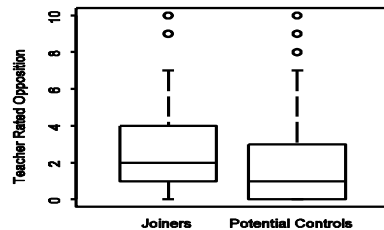
Hyperactivity, Age 11, P=0.017



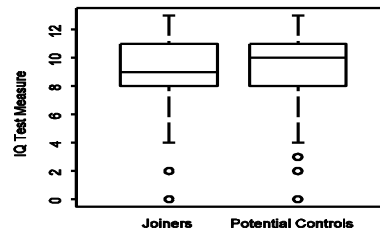
Sex Partners, Age 13, P=.0024



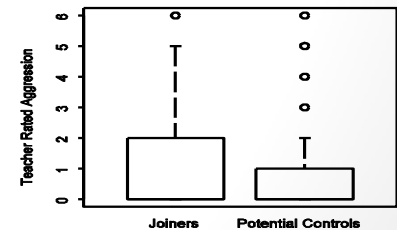
Opposition, Age 11, P=.034



IQ, P=.89



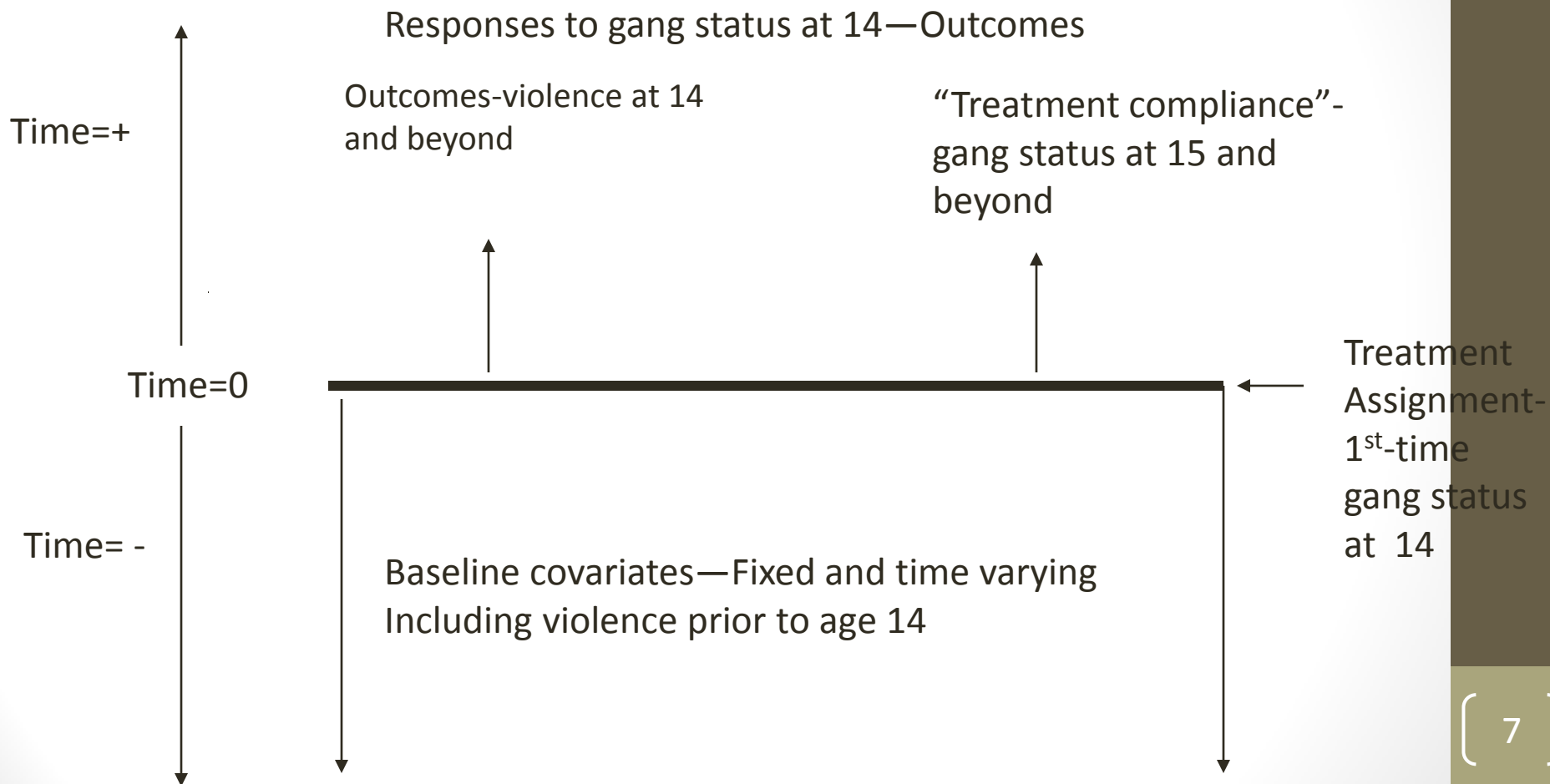
Aggression, Age 11, P=.27



Cochran's Advice on how to proceed: "How should the study be conducted if it were possible to do it by controlled experimentation?"

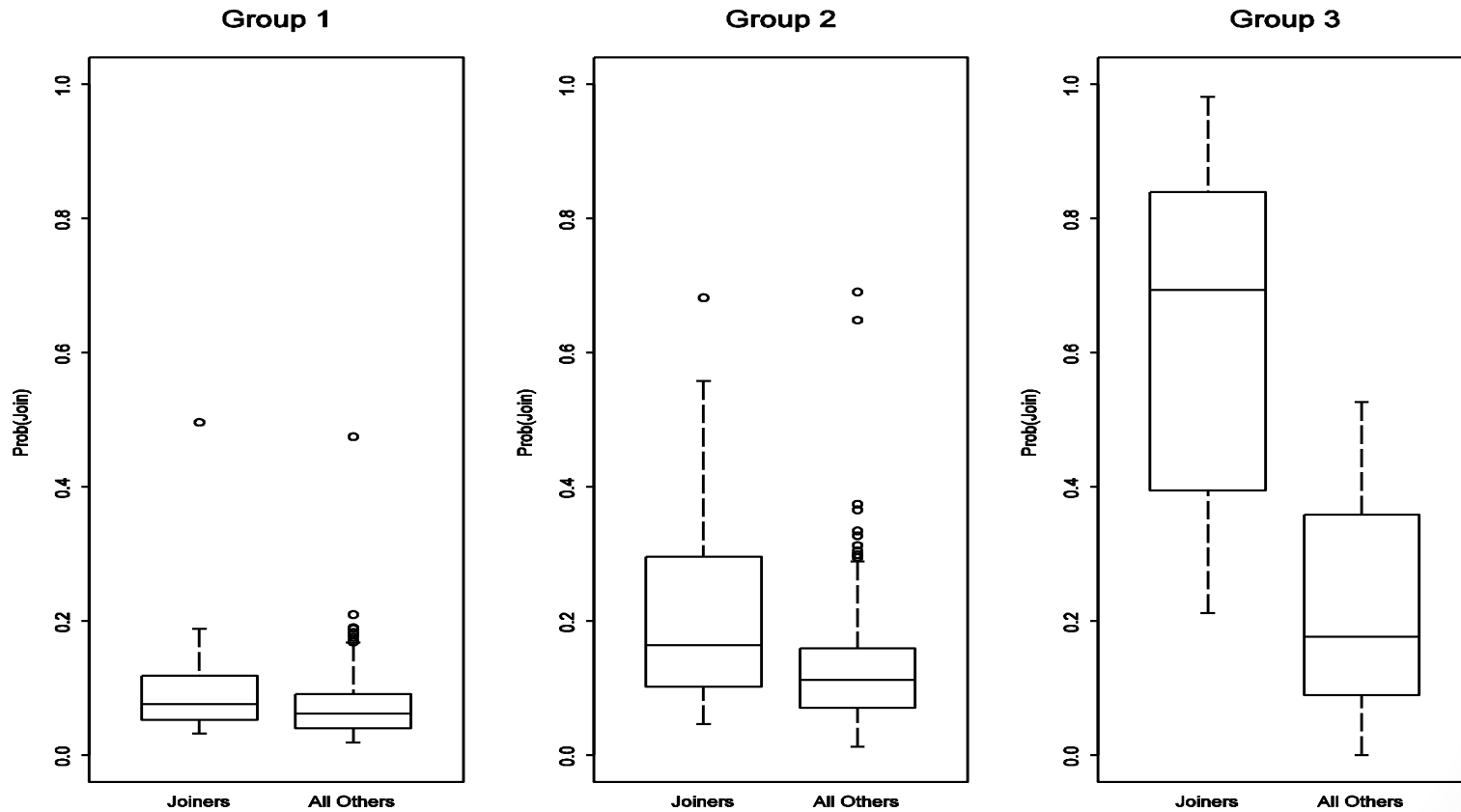
- Well defined treatment—what is the effect of first-time gang membership at age 14 on violence at age 14 and beyond?
- Good baseline measurements on the treated (gang members at 14) and controls (non-gang members at 14)—provided by trajectory groups
- Randomize treatment to create comparability (i.e. balance) on all covariates between treated and controls—provided by propensity score matching

Keeping Time in Order: Treatment, Baseline Covariates, & Outcomes



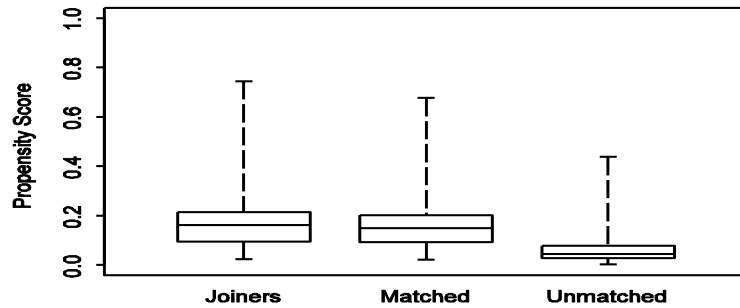
Comparing “Apples to Apples”

(Propensity for gang joining by trajectory group before matching)

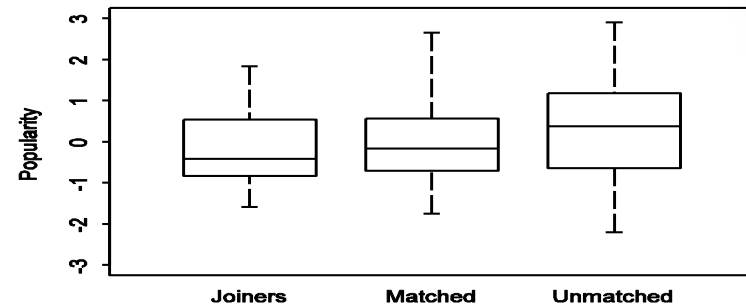


Checking for Balance

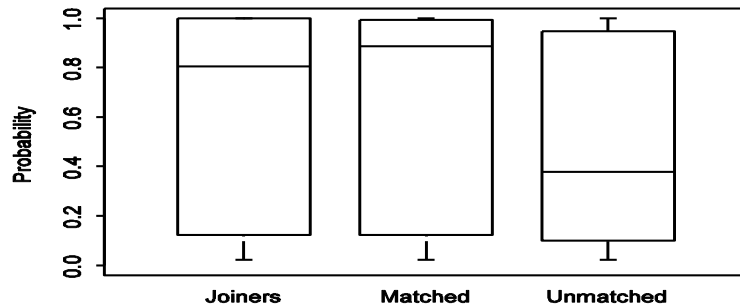
Propensity Score



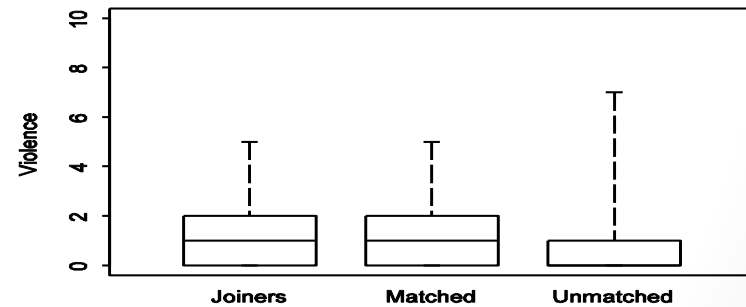
Peer Rated Popularity



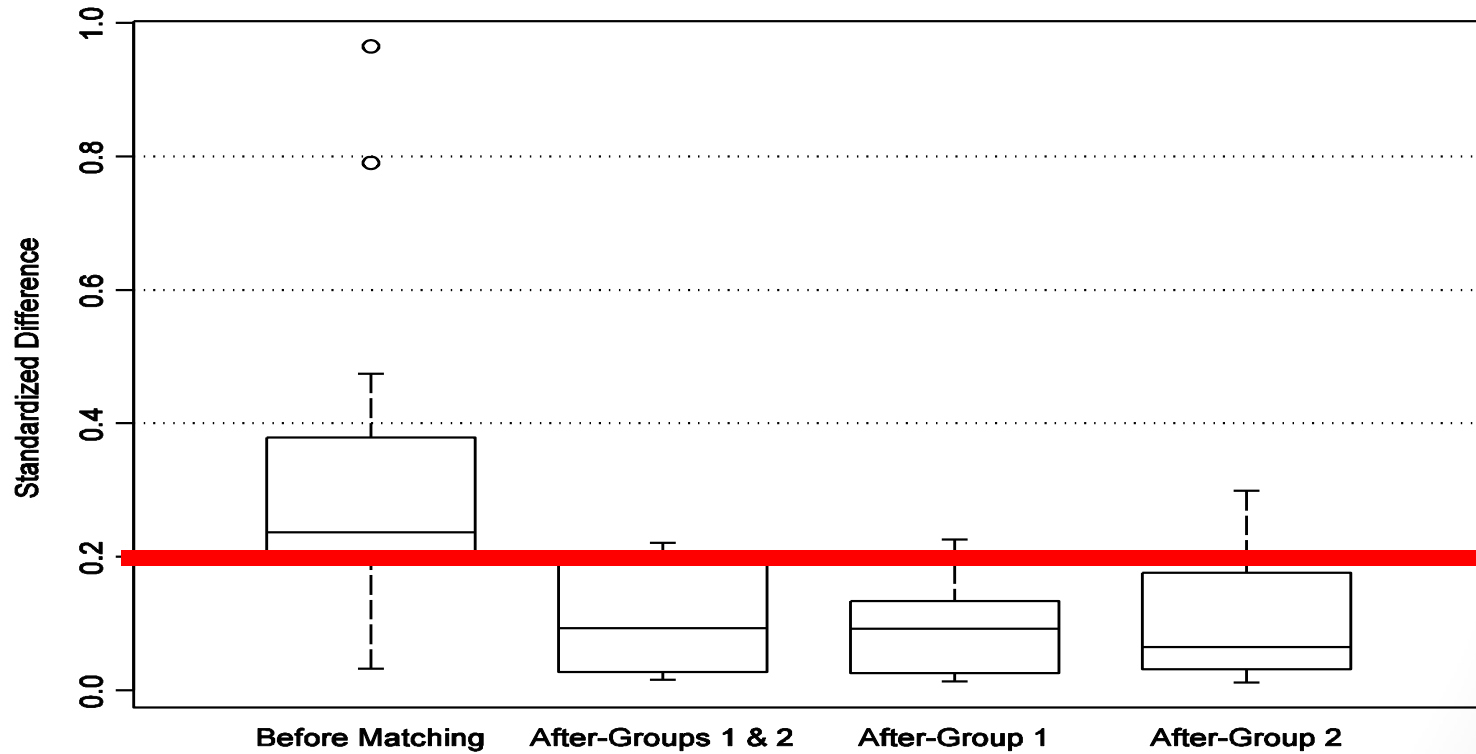
Prob(Group 2 | 1 or 2)



Violence Age 11



Standardized differences across the 15 variables used in matching



Not Peeking Before Testing Treatment Effect: “Intent to Treat” Effects of First-time Gang Membership at 14 on Violence at age 14 to 17

Age	Group	Significance Level
14	Low Declining	.008 .033
15	Low Declining	.034 .086
16	Low Declining	.044 .753
17	Low Declining	.070 .530

Concluding Thoughts

- PM is not a substitute for experiments
- “Hidden Bias” is an unavoidable hazard of analyzing non-experimental data
- PM is an alternative to regression for analyzing non-experimental data
- Chief advantage is that it is a good reminder that the world does not operate according to a regression model