Fixed-effects regression for within-subjects designs: Causal effects of a transition to online course evaluations

LMU|LA
Loyola Marymount
University

Bryce Mason, Ph.D.
Director of IR

Ryan Johnson, Ph.D.
Research Associate

- School of Ed. enjoying benefits of online collection

  - Tens of thousands in labor savings

  - Real-time access to reports for faculty and administration

  - Instructional time savings

  - No more missing packets / 0% response rates

- Expanded adoption up to other deans

  - Concerned about impact on scores, response rates

  - Initial results suggested no impact on rates

  - IR initiated study with strong design to discern impacts

LMU|LA Loyola Marymount University

Simple Means of Response Rates Suggest No Effect
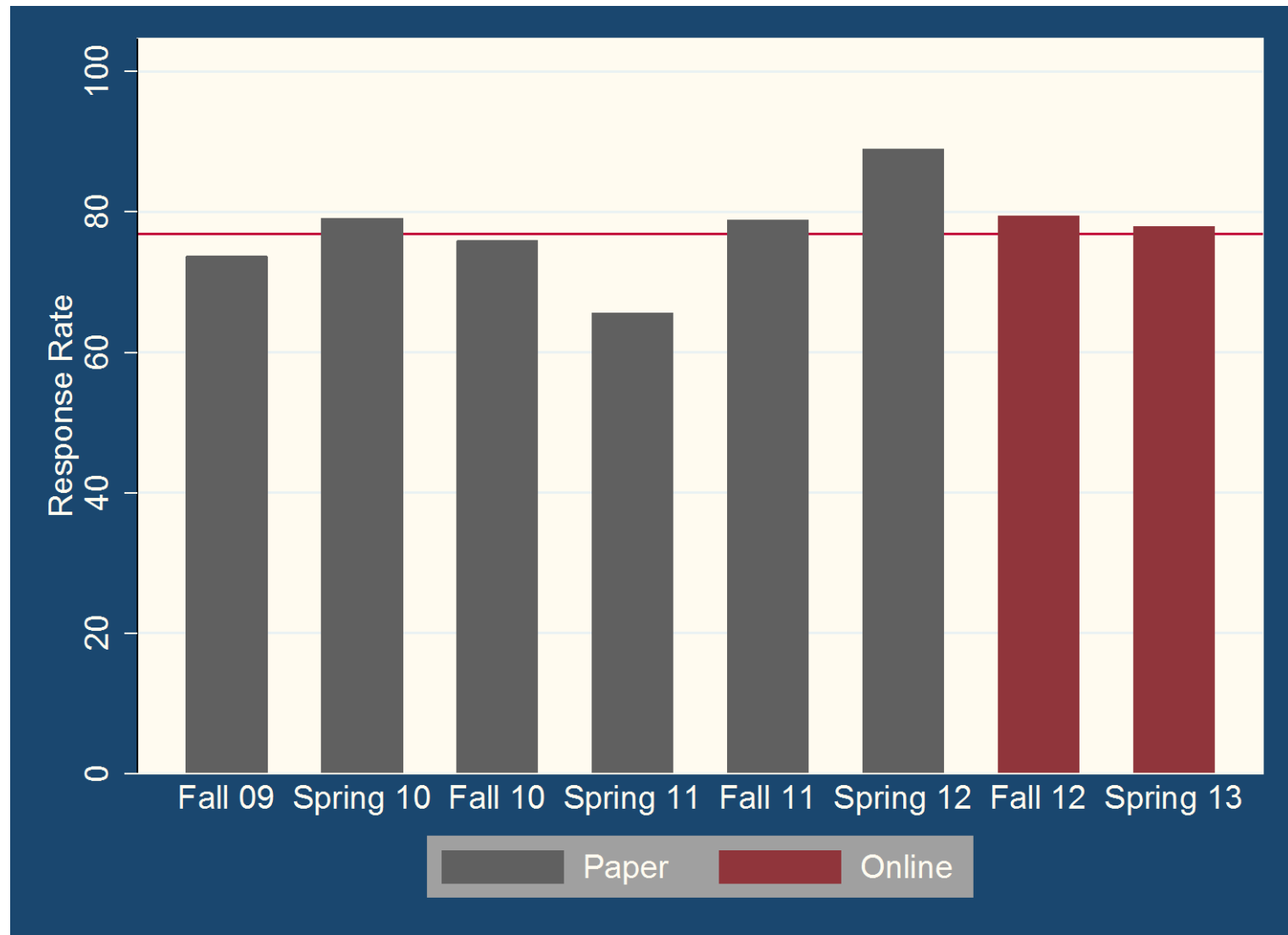
- Evaluations from Fall 2009 to Spring 2013

  - New form implemented Fall 2009

  - 7 instructor-related questions; 1 global item; 5-pt Likert

  - Online collection began Fall 2012 (same form)

- Course registration database for class size

- Course catalogs 2009-10 to 2012-13

  - Built map of course name / number changes

  - Worked with associate dean

  - Provided largest "same course" sample over time

LMU|LA Loyola Marymount University

- Remap old course numbers to current naming
- Find instructor-course combos before/after online
- Collapse individual data to section level
- 3 outcomes related to global item
  - Section mean
  - Section standard deviation
  - Response rate
  - Other items correlated with global item 0.78<p<0.90!
- 461 sections
  - 113 instructor/course combos; 2-17 sections each
  - 81 unique courses; 84 unique instructors

- Within-subjects design desirable

  - Examine changes over time within each combination

  - Outcome changes unlikely due to curriculum / instruction

  - Immune to mix shifts over time / new courses or programs

- Fixed effects regression

  - Control for all course/instructor factors that do not change

  - Allows other control variables that vary over time

LMU|LA Loyola Marymount University

- Intended to provide intuition only!
- Best-fit line for each instructor-course pairing (in reality it's its own level shift)
- Slope averaged across all lines (not estimated to be its own RV)
- Indicator used to model a level shift due to online collection at and beyond Fall 2012
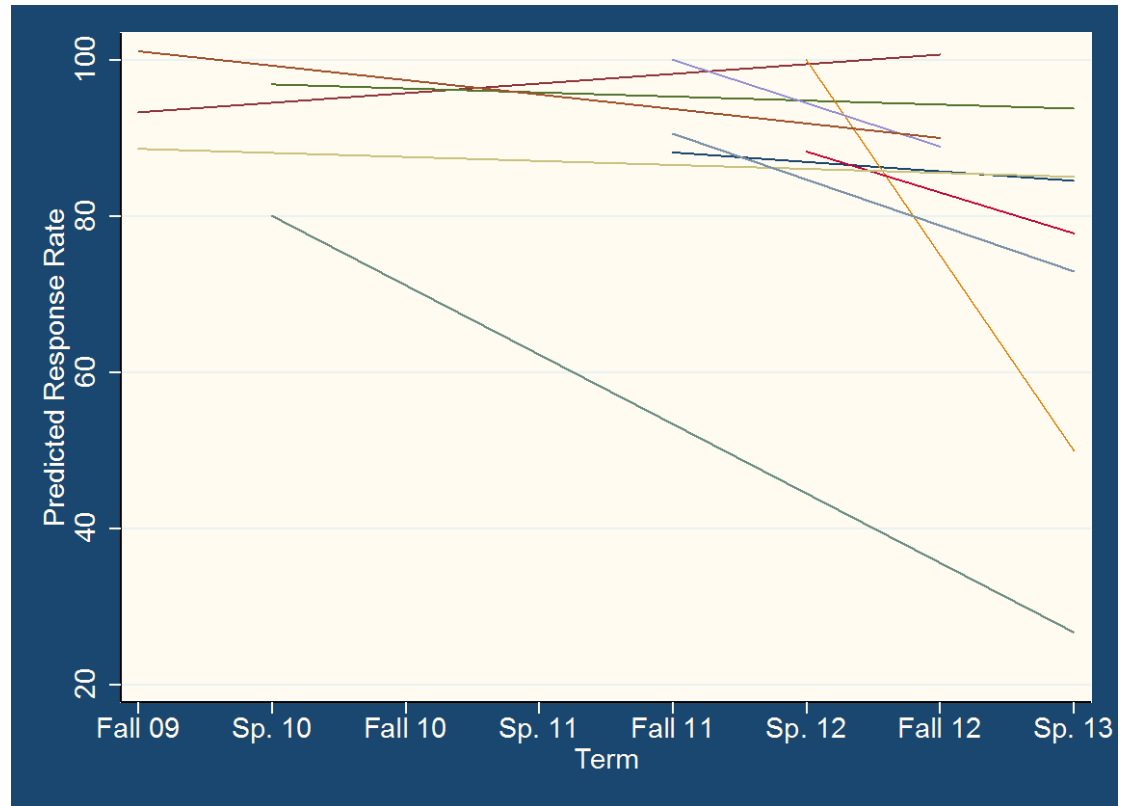
- Superior panel estimators not advisable/possible

  - Unbalanced (inst.-course combos missing) (complicating)

  - 1 to 4 sections at $T_{i,j}$ (-16% sample to solve) (unwilling)

  - Sections not weighted equally (class size) (unable)

- Pooled OLS

  - Indicator variable for every instructor-course combination

  - Model online evaluation with time indicator variable

  - Analytic weights allowed (outcomes are averages)

  - Robust & clustered standard errors to relax assumptions

  - Use all available data

LMU|LA Loyola Marymount University

$$\begin{Bmatrix} \mu_{i,t} \\ \sigma_{i,t} \\ R_{i,t} \end{Bmatrix} = \beta_0 + \beta_1 \text{Online}_{i,t} + \beta_2 \overline{\text{Interest}}_{i,t} + \sum_{j=2}^{113} \beta_{3,j} \text{Line}_{i,j,t} + \epsilon_{i,t}$$

Variable	Mean Rating		S.D. of Ratings		Response Rate	
	β	p	β	p	β	p
Online	0.06	0.14	-0.02	0.50	-12.4	0.00
Interest	0.63	0.00	-0.41	0.00	2.5	0.41
R^2	72.9%		58.9%		49.7%	

- No effect on ratings or distribution of ratings
- Clear negative impact on response rate

LMU|LA Loyola Marymount University

Panel Estimator Method Yields Similar Result

```
. xtreg rate q17 online, fe vce(robust)

Fixed-effects (within) regression              Number of obs     =        388
Group variable: id                             Number of groups  =        113

R-sq:  within  = 0.3026                         Obs per group: min =          2
       between = 0.0423                                        avg =        3.4
       overall = 0.2229                                        max =          8

                                                F(2,112)          =      38.29
corr(u_i, Xb)  = 0.0118                          Prob > F          =     0.0000

                                (Std. Err. adjusted for 113 clusters in id)
------------------------------------------------------------------------------
             |               Robust
        rate |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         q17 |   1.595988   2.855016     0.56   0.577    -4.060861    7.252836
      online |  -13.05868   1.499591    -8.71   0.000    -16.02993   -10.08744
       _cons |    85.3433   12.44181     6.86   0.000     60.69144    109.9952
-------------+----------------------------------------------------------------
     sigma_u |  8.4742692
     sigma_e |  10.994828
         rho |  .37266961   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```
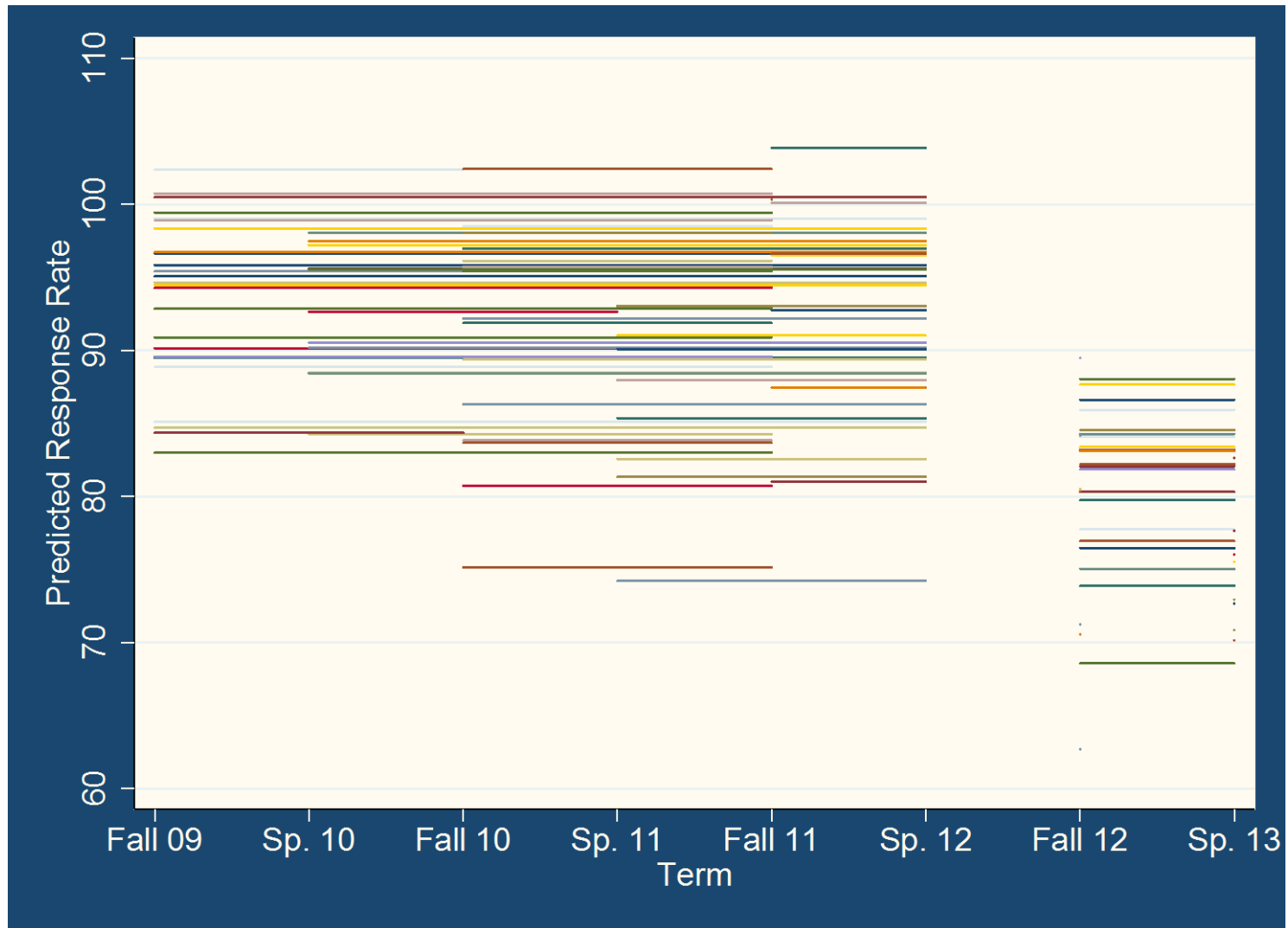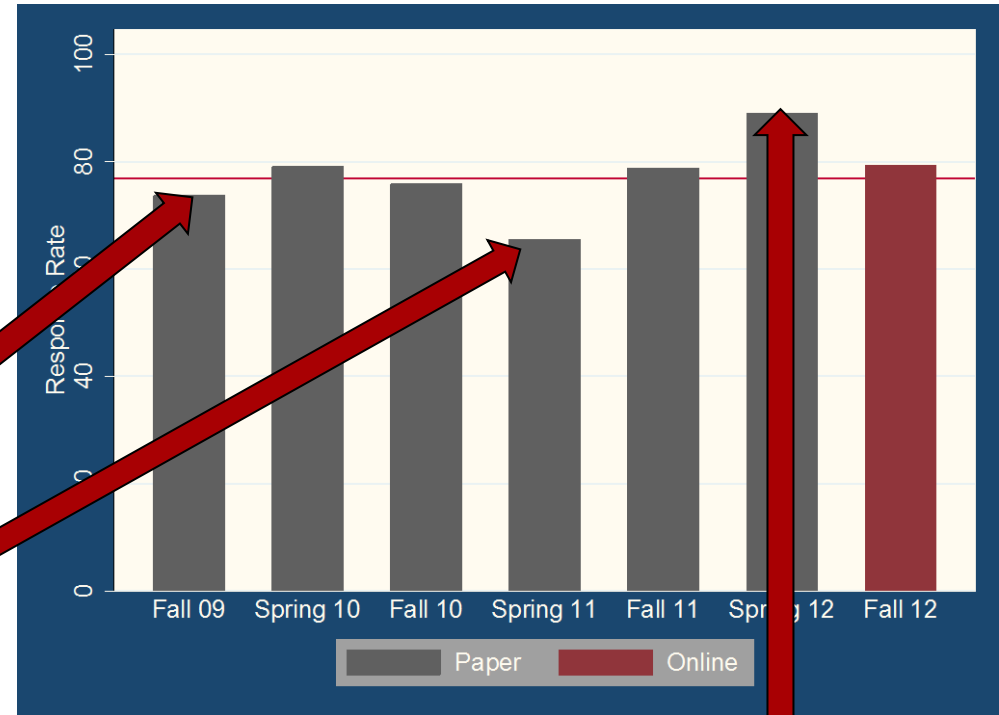
- Result did not square with administrators' expectations

- Investigation yielded fascinating masking of effects

- Individualized studies

- Exam courses

- Clinical coursework added erroneously

- Faculty saw these paper packets and never bothered (0% rate)

Stopped sending paper evaluations to unmeasured sections.

Had staff had the correct course list from Fall 2009, bars would have been consistently higher and demonstrated negative effect more clearly

- Generalization a downside of internally valid designs

- Very specific population

  - Graduate School of Education only

  - Our selection criteria cut 60% of Ed. Sections

- Considerable emphasis on response rate

  - Email alerts to students

  - Space in computer labs dedicated to terminals

  - Encouragement of instructors to convey importance

  - Sub study to look at instructor behavior on response rate

Questions?

Leave business card for a digital copy of paper and slides.

LMU|LA