Roth (2022, AER)

Student Presentation in Master's Thesis Workshop 1

Yasuyuki Matsumura (Kyoto University)

Last Updated: May 10, 2025

https://yasu0704xx.github.io

AER: Insights 2022, 4(3): 305–322
https://doi.org/10.1257/aeri.20210236

Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends†

By JONATHAN ROTH*

This paper discusses two important limitations of the common practice of testing for preexisting differences in trends ("pre-trends") when using difference-in-differences and related methods. First, conventional pre-trends tests may have low power. Second, conditioning the analysis on the result of a pretest can distort estimation and inference, potentially exacerbating the bias of point estimates and under-coverage of confidence intervals. I analyze these issues both in theory and in simulations calibrated to a survey of recent papers in leading economics journals, which suggest that these limitations are important in practice. I conclude with practical recommendations for mitigating these issues. (*JEL* A14, C23, C51)

This slide is available on
https://github.com/yasu0704xx/ArticleReview.

Testing for Pre-Trends

- When using DiD and related methods, reserchers often test for pretreatment differences in trends ("pre-trends") as a way of assessing the plausibility of the parallel trends assumption.
- These tests are remarkably common: 70 publications in *AER*, *AEJ Applied Economics*, *AEJ Economic Policy* between 2014 and June 2018 use an "event-study" plot to visually test for pre-trends.

Limitations with Pretesting for Pre-Trends

- Low Power: Conventional pretests may have low power, meaning that preexisting trends that produce meaningful bias in the treatment effects estimates may not be detected with substantial probability.
- Distortion in Estimation: Conditioning the analysis on the result of a pre-trends test induces distortions to estimation and inference from pretesting. In other words, the draws of the data that survive a pretest are a selected sample from the true DGP. Because of this selection, the bias caused by a violation of parallel trends can actually be worse conditional on passing the pretest.
- These results imply that pre-trends tests may be ineffective in avoiding biases from violations of parallel trends and can even exacerbate these biases.

Related Literature: Pretesting, Model-Selection

- Issues arising with pretesting or model-selection setup:
  - Keynes (1939) [36], Friedman (1940) [26]
  - Giles and Giles (1993) [29], Leeb and Potscher (2005) [42], Lee et al. (2016) [41], Andrews (2018) [2]
- Publication bias matters:
  - Rothstein, Sutton and Borenstein (2005) [50], Christensen and Miguel (2016) [14], Snyder and Zhuo (2018) [52], Andrews and Kasy (2019) [3]
- Low power of the pre-trends tests to detect meaningful violations of parallel trends:
  - Freyaldenhoven, Hansen and Shapiro (2019) [24], Kahn-Lang and Lang (2020) [35], Bilinski and Hatfield (2018) [8]
  - This paper also provides theoretical and empirical evidence on additional statistical distortions from pretesting.
  - The alternative approach can be found in Bilinski and Hatfield (2018) [8] and Freyaldenhoven, Hansen and Shapiro (2019).

Related Literature: DiD

- Failure of standard TWFE models:
  - Borusyak and Jaravel (2016) [9], Sun and Abraham (2021) [53], de Chaisemartin and D'Haultfoeuille (2020) [16], Goodman-Bacon (2021) [31], Callaway and Sant'Anna (2021) [12], Athey and Imbens (2022) [5]
  - This paper points out that conventional pretests may do a poor job detecting violations of the relevant parallel trends assumption.

Contents

Survey of Recent Papers

Sample

- Candidates: 70 empirical papers with the phrase "event study" in *AER, AEJ Applied Economics, AEJ Economic Policy* between 2014 and June 2018.
- The final sample of 12 papers:
  - Bailey and Goodman-Bacon (2015) [6], Bosch and Campos-Vazquez (2014) [10], Deryugina (2017) [17], Deschenes et al. (2017) [19], Fitzpatrick and Lovenheim (2014) [22], Gallagher (2014) [27], He and Wang (2017) [33], Kuziemko et al. (2018) [37], Lafortune et al. (2018) [39], Markevich and Zhuravskaya (2018) [45], Tewari (2014) [54], Ujhelyi (2014) [56]
- The above 12 studies show an event plot with pointwise CIs that alows ..... (続きここから)

TABLE 1—SUMMARY OF PRE-PERIOD EVENT-STUDY COEFFICIENTS

Paper	# pre-periods	# significant	Max |t|	Joint *p*-value	|t| for slope
Bailey and Goodman-Bacon (2015)	5	0	1.674	0.540	0.381
Bosch and Campos-Vazquez (2014)	11	2	2.357	0.137	0.446
Deryugina (2017)	4	0	1.090	0.451	1.559
Deschenes et al. (2017)	5	1	2.238	0.014	0.239
Fitzpatrick and Lovenheim (2014)	3	0	0.785	0.705	0.977
Gallagher (2014)	10	0	1.542	0.166	0.855
He and Wang (2017)	3	0	0.884	0.808	0.720
Kuziemko et al. (2018)	2	0	0.474	0.825	0.474
Lafortune et al. (2017)	5	0	1.382	0.522	1.390
Markevich and Zhuravskaya (2018)	3	0	0.850	0.591	0.676
Tewari (2014)	10	0	1.061	0.948	0.198
Ujhelyi (2014)	4	1	2.371	0.003	1.954

Notes: This table provides information about the pre-period event-study coefficients in the papers reviewed. The table shows the number of pre-period coefficients in the event study, the number of the pre-period coefficients that are significant at the 95 percent level, the maximum t-stat among those coefficients, the *p*-value for a chi-squared test of joint significance, and the t-stat for the slope of the linear trend through the pre-period coefficients. See Section I for more detail on the sample of papers reviewed.

Theoretical Analysis

Practical Recommendations

Appendix

Canonical Two-Period DiD

- Suppose that we have panel data $\{\{(Y_{it}, D_{it})\}_{t=0}^{T}\}_{i=1}^{n}$, where i and t index units and time periods, respectively. $Y_{it} \in \mathbb{R}$ and $D_{it} \in \{0, 1\}$ denote the outcome and treatment, respectively, for unit i in period t.
- Consider the simplest case where $T = 1$, no units are treated at $t = 0$ (i.e., $D_{i0} = 0$ for any i), and some but not all units become treated at $t = 1$.
- Then, only D_{i1} is relevant, and we simply write $D_i = D_{i1}$.
- Let $Y_{it}(d)$ denote the potential outcome given $D_i = d$.

Parallel Trends in Canonical DiD

- Parallel trends (PT) is a key assumption for point identification of ATT in DiD. [1]
- In a canonical two-period DiD model, PT means

$$\mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|D_i = 1] = \mathbb{E}[Y_{i1}(0) - Y_{i0}(0)|D_i = 0]$$

  In words, the untreated potential outcome is required to have the same average path over time between the treated and untreated groups.
- Under assumptions of no anticipation and PT, ATT is identified as

$$\begin{aligned}\tau_{\text{ATT}} &= \mathbb{E}[Y_{i1}(1) - Y_{i1}(0)|D_i = 1] \\ &= \mathbb{E}[Y_{i1} - Y_{i0}|D_i = 1] - \mathbb{E}[Y_{i1} - Y_{i0}|D_i = 0]\end{aligned}$$

[1] Note that identification of ATT also requires another assumption, namely, "no anticipation," or "no anticipatory effects of treatment," which we do not focus on here.

Pre-Trends Tests

- Since $Y_{i1}(0)|D_i = 1$ is counterfactual (i.e., canoot be observed), PT cannot be directly examined.
- Instead, reseachers implement pre-trends tests, testing for preexisting differences in trends, so that they can assess the plausibility of the PT assumption.
- Note that this "palusibility" is neither a necessary nor sufficient condition for PT.

References

References i

- Abadie, Alberto. 2005. “ Semiparametric Difference-in-Differences Estimators. ” *Review of Economic Studies* **72**(1): 1-19.
- Andrews, Isaiah. 2018. “ Valid Two-Step Identification-Robust Confidence Sets for GMM. ” *Review of Economics and Statistics* **100**(2): 337-48.
- Andrews, Isaiah, & Maximilian Kasy. 2019. “ Identification of and Correction for Publication Bias. ” *American Economic Review* **109**(8): 2766-94.

References ii

- Ashenfelter, Orley. 1978. “ Estimating the Effect of Training Programs on Earnings. ” *Review of Economics and Statistics* **60**(1): 47-57.
- Athey, Susan, & Guido W. Imbens. 2022. “ Design-Based Analysis in Difference-in-Differences Settings with Staggered Adoption. ” *Journal of Econometrics* **226**(1): 62-79.
- Bailey, Martha J., & Andrew Goodman-Bacon. 2015. “ The War on Poverty ’ s Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans. ” *American Economic Review* **105**(3): 1067-104.

References iii


Bailey, Martha J., & Andrew Goodman-Bacon. 2019. “Replication data for: The War on Poverty’s Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans.” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E112871V1.


Bilinski, Alyssa, & Laura A. Hatfield. 2018. “Seeking Evidence of Absence: Reconsidering Tests of Model Assumptions.” arXiv:1805.03273v1.


Borusyak, Kirill, & Xavier Jaravel. 2016. “Revisiting Event Study Designs.” Unpublished manuscript.

References iv


Bosch, Mariano, & Raymundo M. Campos-Vazquez. 2014. “ The Trade-Offs of Welfare Policies in Labor Markets with Informal Jobs: The Case of the ‘ Seguro Popular ’ Program in Mexico. ” *American Economic Journal: Economic Policy* **6**(4): 71-99.


Bosch, Mariano, & Raymundo M. Campos-Vazquez. 2019. “ Replication data for: The Trade-Offs of Welfare Policies in Labor Markets with Informal Jobs: The Case of the ‘ Seguro Popular ’ Program in Mexico. ” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E114886V1.

References v

- Callaway, Brantly, & Pedro H.C. Sant'Anna. 2021. " Difference-in-Differences with Multiple Time Periods. " *Journal of Econometrics* **225**(2): 200-30.
- Cartinhour, Jack. 1990. " One-Dimensional Marginal Density Functions of a Truncated Multivariate Normal Density Function. " *Communications in Statistics—Theory and Methods* **19**(1): 197-203.
- Christensen, Garret S., & Edward Miguel. 2016. " Transparency, Reproducibility, and the Credibility of Economics Research. " NBER Working Paper No. 22989.
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. New York: Routledge.

References vi

- de Chaisemartin, Clement, & Xavier D'Haultfoeuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* **110**(9): 2964-96.
- Deryugina, Tatyana. 2017. "The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance." *American Economic Journal: Economic Policy* **9**(3): 168-98.
- Deryugina, Tatyana. 2019. "Replication data for: The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E114620V1.

References vii

- Deschenes, Olivier, Michael Greenstone, & Joseph S. Shapiro. 2017. “ Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program. ” *American Economic Review* **107**(10): 2958-89.
- Desch nes, Olivier, Michael Greenstone, & Joseph S. Shapiro. 2019. “ Replication data for: Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program. ” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E112938V1.

References viii

- Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, & Matthew J. Notowidigdo. 2018. “ The Economic Consequences of Hospital Admissions. ” *American Economic Review* **108**(2): 308-52.
- Fitzpatrick, Maria D., & Michael F. Lovenheim. 2014. “ Early Retirement Incentives and Student Achievement. ” *American Economic Journal: Economic Policy* **6**(3): 120-54.
- Fitzpatrick, Maria D., & Michael F. Lovenheim. 2019. “ Replication data for: Early Retirement Incentives and Student Achievement. ” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E114865V1.

References ix

- Freyaldenhoven, Simon, Christian Hansen, & Jesse M. Shapiro. 2019. " Pre-event Trends in the Panel Event-Study Design. " *American Economic Review* **109**(9): 3307-38.
- Freyaldenhoven, Simon, Christian Hansen, Jorge Perez Perez, & Jesse M. Shapiro. 2021. " Visualization, Identification, and Estimation in the Linear Panel Event-Study Design. " NBER Working Paper No. 29170.
- Friedman, Milton. 1940. " Review: Business Cycles in the United States of America, 1919-1932. " *American Economic Review* **30**(3): 657-60.

References x

- Gallagher, Justin. 2014. "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States." *American Economic Journal: Applied Economics* **6**(3): 206-33.
- Gallagher, Justin. 2019. "Replication data for: Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E113898V1.
- Giles, Judith A., & David E.A. Giles. 1993. "Pre-test Estimation and Testing in Econometrics: Recent Developments." *Journal of Economic Surveys* **7**(2): 145-97.

References xi

- Goodman-Bacon, Andrew. 2018. “ Public Insurance and Mortality: Evidence from Medicaid Implementation. ” *Journal of Political Economy* **126**(1): 216-62.
- Goodman-Bacon, Andrew. 2021. “ Difference-in-Differences with Variation in Treatment Timing. ” *Journal of Econometrics* **225**(2): 254-77.
- Heckman, James J., Hidehiko Ichimura, & Petra E. Todd. 1997. “ Matching as An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. ” *Review of Economic Studies* **64**(4): 605-54.

References xii


He, Guojun, & Shaoda Wang. 2017. “Do College Graduates Serving as Village Officials Help Rural China?” *American Economic Journal: Applied Economics* **9**(4): 186-215.


He, Guojun, & Shaoda Wang. 2019. “Replication data for: Do College Graduates Serving as Village Officials Help Rural China?” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E113682V1.


Kahn-Lang, Ariella, & Kevin Lang. 2020. “The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications.” *Journal of Business & Economic Statistics* **38**(3): 613-20.

- Keynes, J.M. 1939. “ Professor Tinbergen ’ s Method. ” *Economic Journal* **49**(195): 558-77.
- Kuziemko, Ilyana, Katherine Meckel, & Maya Rossin-Slater. 2018. “ Does Managed Care Widen Infant Health Disparities? Evidence from Texas Medicaid. ” *American Economic Journal: Economic Policy* **10**(3): 255-83.
- Kuziemko, Ilyana, Katherine Meckel, & Maya Rossin-Slater. 2019. “ Replication data for: Does Managed Care Widen Infant Health Disparities? Evidence from Texas Medicaid. ” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E114651V1.

Lafortune, Julien, Jesse Rothstein, & Diane Whitmore Schanzenbach. 2018. "School Finance Reform and the Distribution of Student Achievement." *American Economic Journal: Applied Economics* **10**(2): 1-26.

Lafortune, Julien, Jesse Rothstein, & Diane Whitmore Schanzenbach. 2019. "Replication data for: School Finance Reform and the Distribution of Student Achievement." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E113709V1.

References xv

- Lee, Jason D., Dennis L. Sun, Yuekai Sun, & Jonathan E. Taylor. 2016. " Exact Post-selection Inference, with Application to the Lasso. " *Annals of Statistics* **44**(3): 907-27.
- Leeb, Hannes, & Benedikt M. Potscher. 2005. " Model Selection and Inference: Facts and Fiction. " *Econometric Theory* **21**(1): 21-59.
- Lovenheim, Michael F., & Alexander Will n. 2019. " The Long-Run Effects of Teacher Collective Bargaining. " *American Economic Journal: Economic Policy* **11**(3): 292-324.
- Manjunath, B.G., & Stefan Wilhelm. 2012. " Moments Calculation for the Doubly Truncated Multivariate Normal Density. " arXiv:1206.5387.

- Markevich, Andrei, & Ekaterina Zhuravskaya. 2018. "The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire." *American Economic Review* **108**(4-5): 1074-117.
- Markevich, Andrei, & Ekaterina Zhuravskaya. 2019. "Replication data for: The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E113121V1.

- Mathematics Stack Exchange. 2013. Stack Exchange Inc. https://math.stackexchange.com/q/455809 (accessed August 1, 2013).
- Rambachan, Ashesh, & Jonathan Roth. 2022. "A More Credible Approach to Parallel Trends." Unpublished manuscript.
- Roth, Jonathan. 2022. "Replication data for: Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E151982V1.

References xviii

- Rothstein, Hannah R., Alexander J. Sutton, & Michael Borenstein. 2005. "Publication Bias in Meta-Analysis." In *Publication Bias in Meta-Analysis*, ed. Hannah R. Rothstein, Alexander J. Sutton, & Michael Borenstein, 1-7. West Sussex, England: John Wiley & Sons.
- Sant'Anna, Pedro H.C., & Jun Zhao. 2020. "Doubly Robust Difference-in-Differences Estimators." *Journal of Econometrics* **219**(1): 101-22.
- Snyder, Christopher M., & Ran Zhuo. 2018. "Sniff Tests in Economics: Aggregate Distribution of Their Probability Values and Implications for Publication Bias." Unpublished manuscript.

- Sun, Liyang, & Sarah Abraham. 2021. " Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. " *Journal of Econometrics* **225**(2): 175-99.
- Tewari, Ishani. 2014. " The Distributive Impacts of Financial Development: Evidence from Mortgage Markets during US Bank Branch Deregulation. " *American Economic Journal: Applied Economics* **6**(4): 175-96.
- Tewari, Ishani. 2019. " Replication data for: The Distributive Impacts of Financial Development: Evidence from Mortgage Markets during US Bank Branch Deregulation. " American Economic Association [publisher], Inter-university Consortium

for Political and Social Research [distributor]. https://doi.org/10.3886/E116384V1.


Ujhelyi, Gergely. 2014. "Civil Service Rules and Policy Choices: Evidence from US State Governments." *American Economic Journal: Economic Policy* **6**(2): 338-80.


Ujhelyi, Gergely. 2019. "Replication data for: Civil Service Rules and Policy Choices: Evidence from US State Governments." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E114862V1.

 Wolfers, Justin. 2006. “ Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results. ” *American Economic Review* **96**(5): 1802-20.