



An Ode to Giants and Some Facts About Publishing in the JPE

John A. List
U. Chicago, ANU, and NBER
[@johnlist.bsky.social](https://bsky.app/profile/johnlist.bsky.social)
[@Econ_4_Everyone](https://bsky.app/profile/econ4everyone)



An Ode to Orley

	Before Orley	After Orley
Submissions	765.33	1,471.13
Acceptance Rate	15.96%	7.72%
Impact Factor	1.280	3.139

Average Submissions, Acceptance Rate and Impact Factor in the 15 years immediately before and immediately after Orley Ashenfelter's tenure as AER Chief Editor



Our Other Prominent Panelists

Mary Elizabeth is creating huge value handling the behavioral sciences at Nature. She is truly bringing “*humanity*” to Nature (and maybe even economics!).

Erzo Luttmer, just started in 2023, but his track record as a reporter and research record both signal that he will continue the Ashenfelter tradition



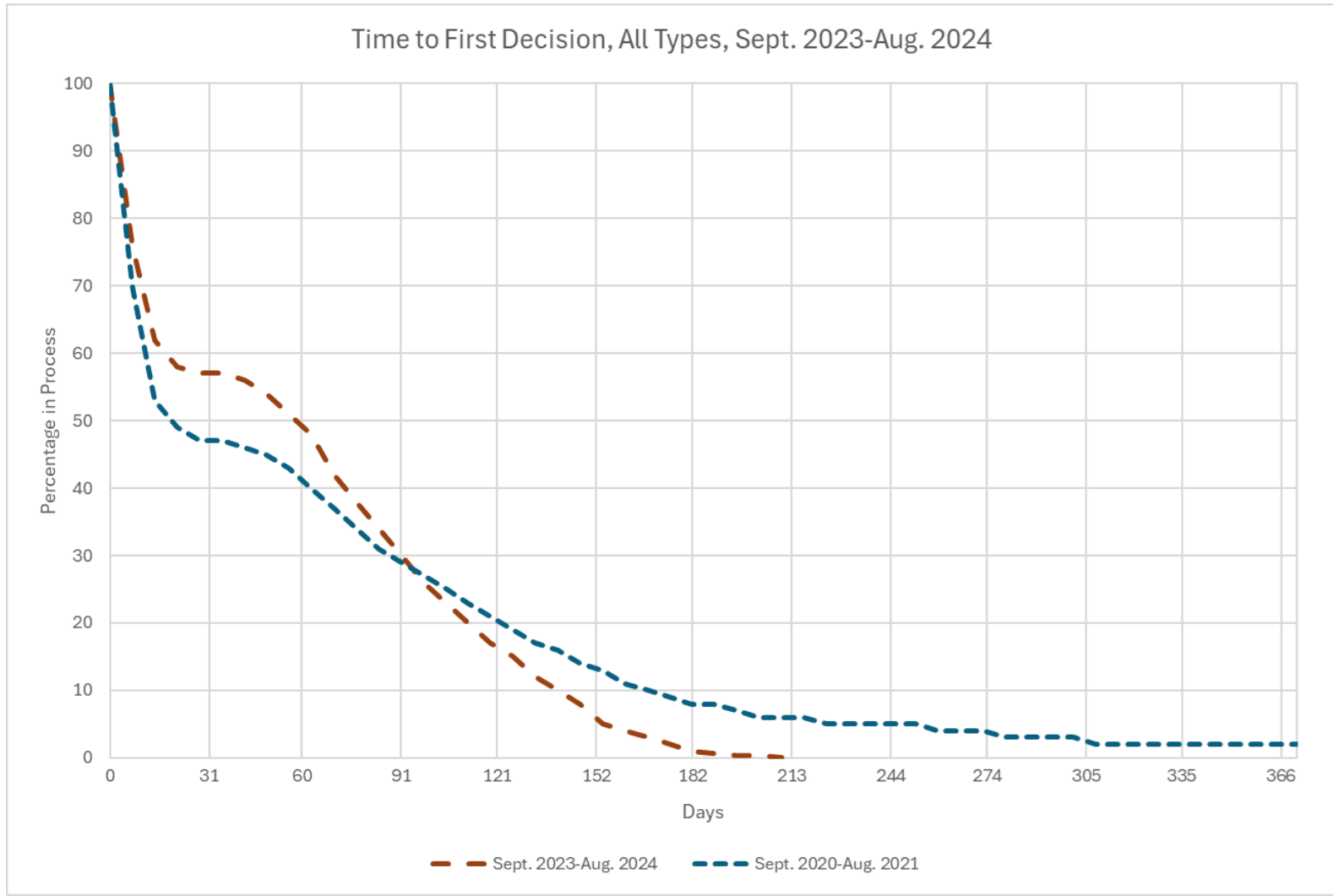
Some JPE Facts

2022: 1270 submissions 70 published articles
2023: 1317 submissions 79 published articles
2024: 1469 submissions 92 published articles

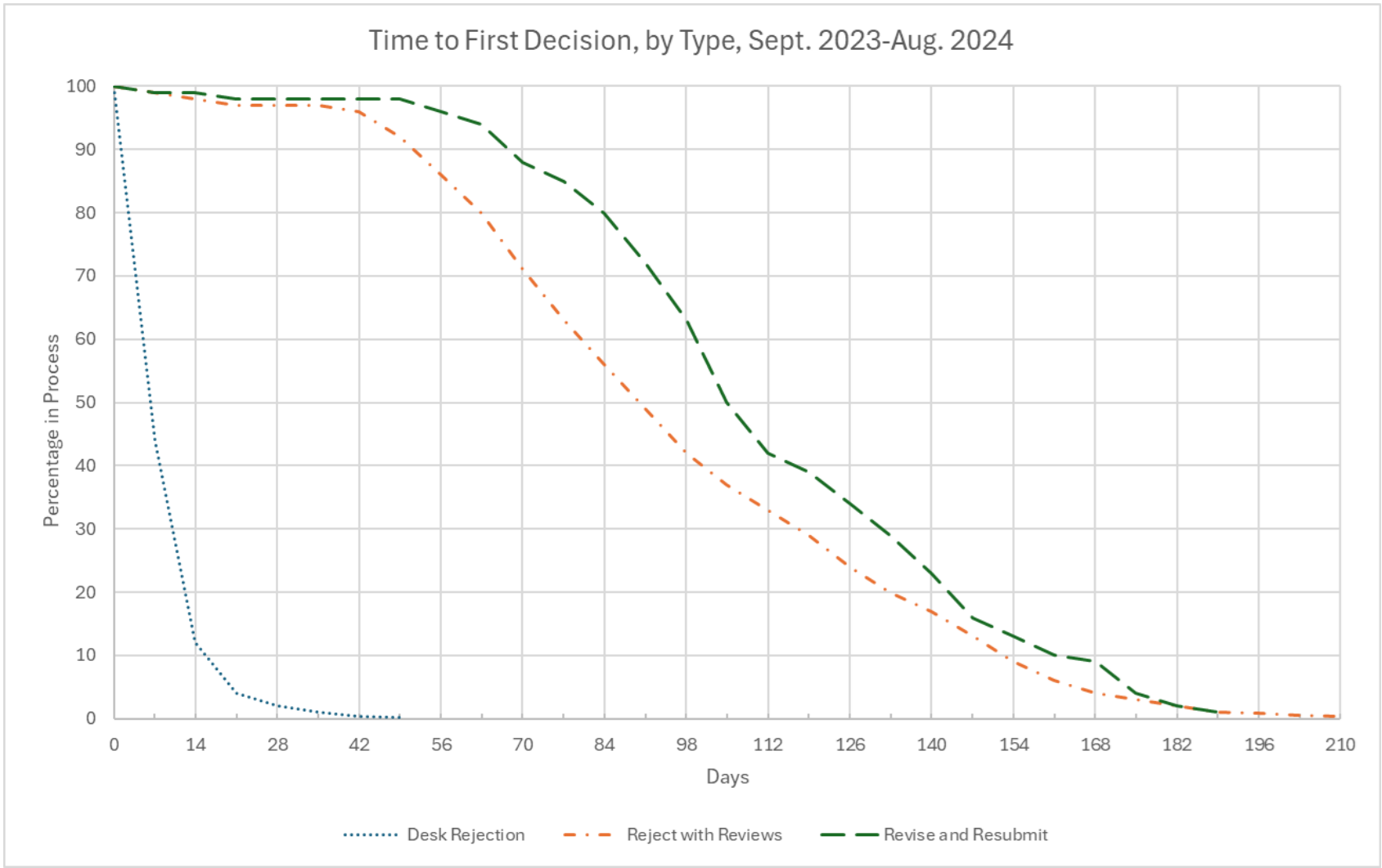
	Outcome of 1st-Round Decisions	Mean Days to Decision	Median Days to Decision
Desk Rejection	42%	9	7
Reject with Reviews	52%	96	90
Revise	6%	111	106

Note: Decisions made September 2023-August 2024

Some JPE Facts



Some JPE Facts



4 JPE Tips

- As you might guess, rejection is a bit like disequilibrium analysis: there are several reasons to have a paper rejected
- What I have learned from my editing chores is that four tips can help (JEEM, ERE, AER, JEP, JEL, MS, JPE, JPE-Micro)
- Caveat: I handle mostly empirical work!

4 Tips

1. First-order topic (now or a path to the future)
 - When we go down the research rabbit hole, we tend to think it is much broader than it really is
 - Chicago litmus test: if you visit 3 random offices will all 3 people say that this is a contribution worthy of their attention (think Stokey, Reny, and Mogstad)
 - It is the author's job, not the editor or reviewers, to make the case of your topic's import



4 Tips

2. Establish causality/credibility beyond a reasonable doubt and control the false positive rate



2. Controlling False Positives:

An Example: MHT

Multiple hypothesis testing refers to any instance in which a family of hypotheses are carried out at once and one has to decide which hypotheses to reject.

A common, yet potentially critical, flaw that leads to an excess of false positives

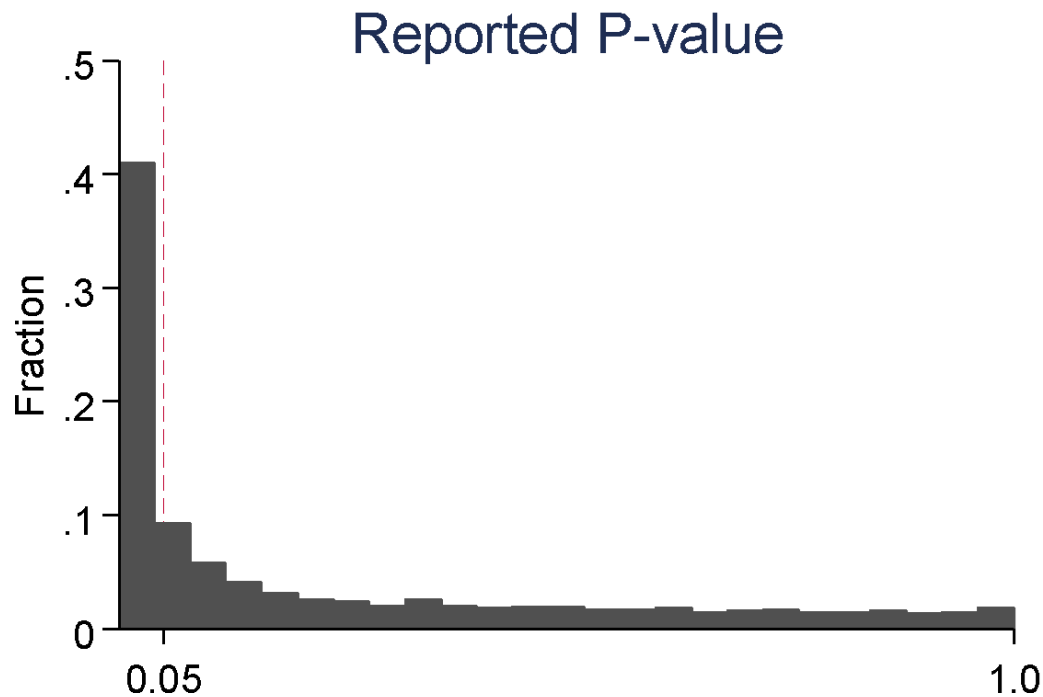
Just because you have your plan pre-registered does not mean you can ignore



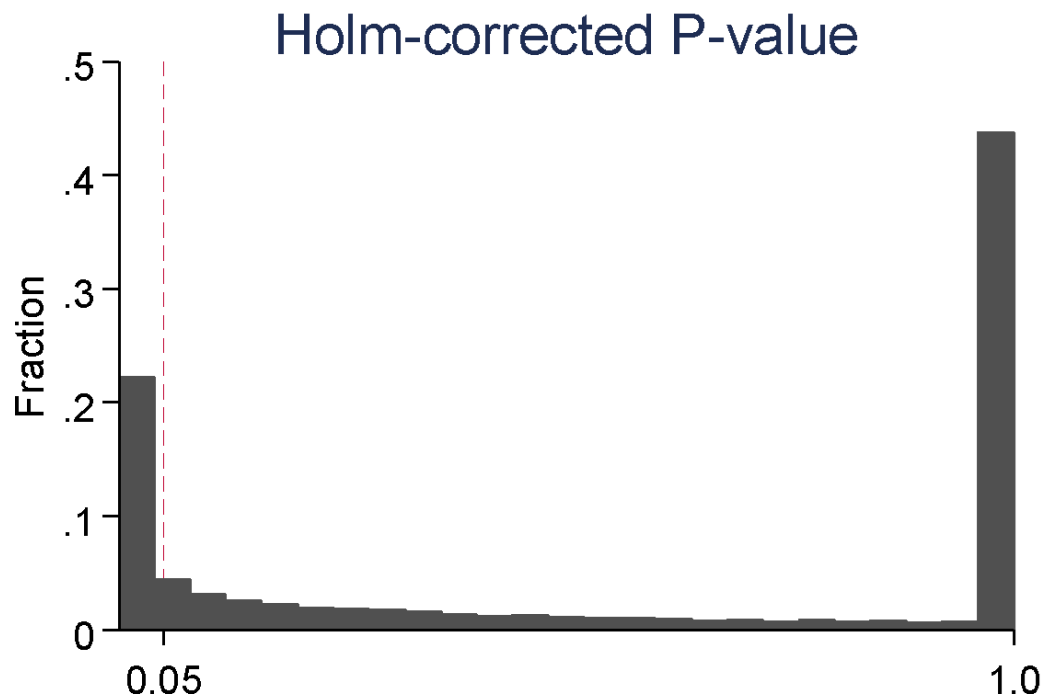
Three Major Cases of MHT

- A. More than one outcome
- B. Heterogeneity (sub-group analysis)
- C. More than one treatment

How Big of a Problem is it?



Corrected P-values



An Approach to MHT

Multiple Hypothesis Testing in
Experimental Economics; List, Shaikh,
and Xu (2019, EE)

- <https://github.com/seidelj/mht>

2019

2021

Multiple Testing with Covariate
Adjustment in Experimental Economics;
List, Shaikh, and Vayalinkal (2023,
Econometric Reviews)

- <https://github.com/vayalinkal/mhtexp2>

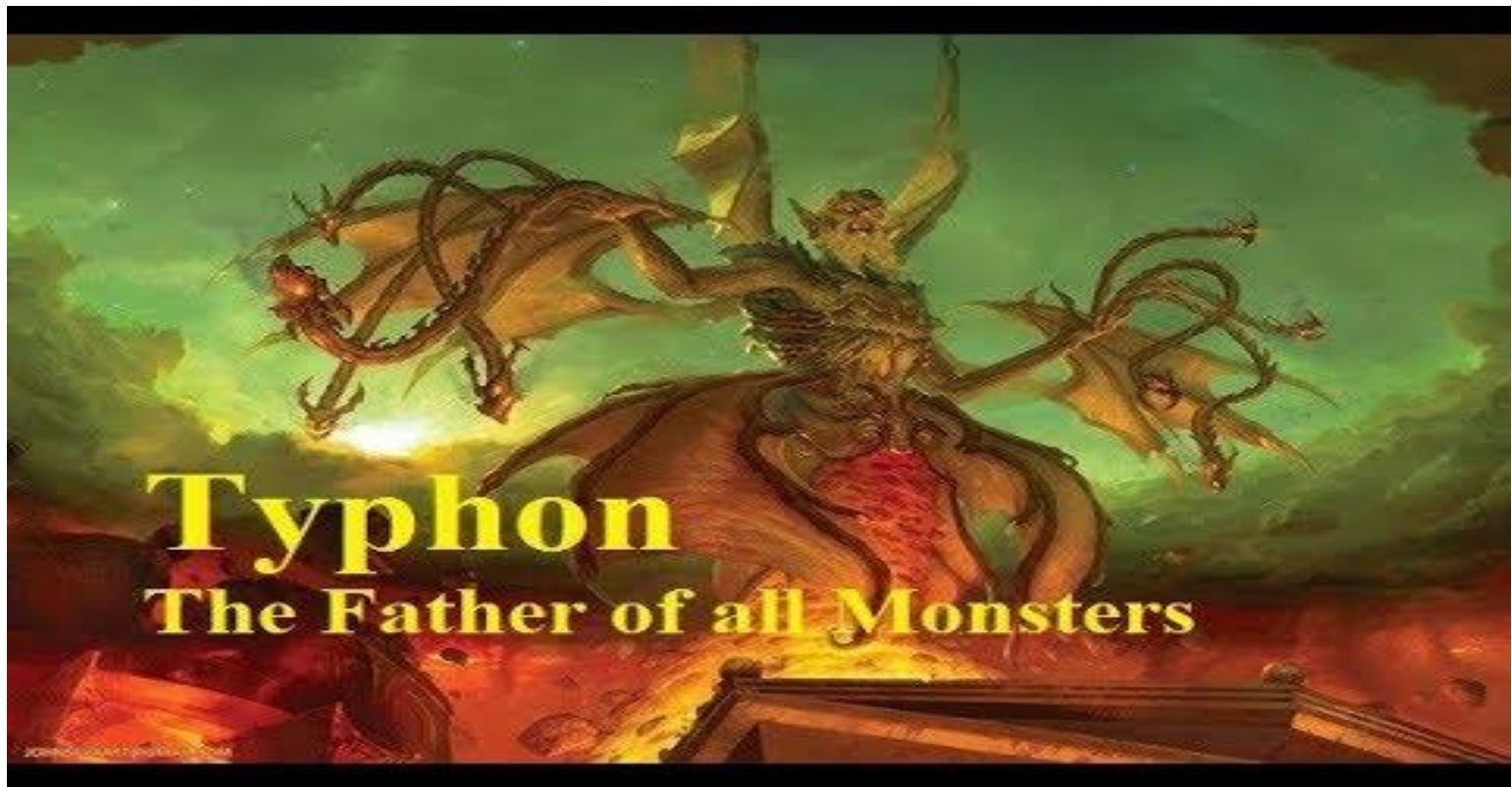


4 Tips

3. Make Proper Inference by Avoiding 2 Monsters

What is the Typhon to Experimentalists Trying to Publish These Days?

EXTERNAL VALIDITY!





External Validity

- More than 9 of 10 experimental studies that I handle at the JPE receive this comment in some form
- If we really want to stop all empiricism in top journals we should maintain our absent-minded complaining about this issue.
 - Experimenters are especially susceptible because identification is usually “cleaner” (with naturally-occurring data, identification tends to be the showdown)
 - The truth is that if you really want, EVERY empirical study can be rejected based on this complaint (time, situation (environment), selection, and/or space will do you in every time)



What to do?

Tackling it Head-On

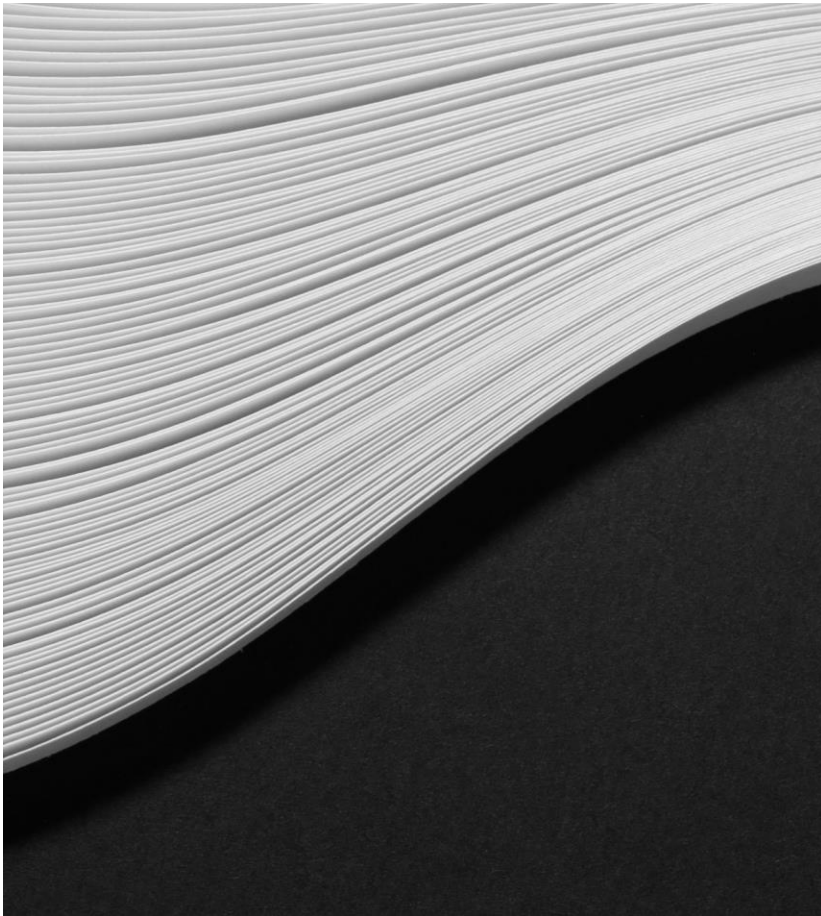
- A satirical take:

List, John A. (2020) “Non est Disputandum de Generalizability? A Glimpse into The External Validity Trial”

Author Onus Probandi

- Much like we have features that give confidence of internal validity in our work, we need the same with EV
- Economic approach to EV provides 4 reporting areas (SANS) that give insight into when preferences, beliefs, or individual constraints might vary importantly across populations of people, settings, situations, and time
- Our papers should report on these 4 SANS areas just as we do with IV

A Litmus Test



- EV Litmus Test: can you confidently say that it is difficult/impossible to find a better setting than the one studied in your paper to test convincingly the relevance of your conjectures?
- Many critics view unique settings as a negative distraction.
- This is the exact opposite way to think of the issue: if the unique setting itself allows you to do the relevant test at hand and no other setting can, then you have found the *perfect* domain for your study.



**There is an even more
important monster around
the corner...Echidna**

SCALING

Scaling



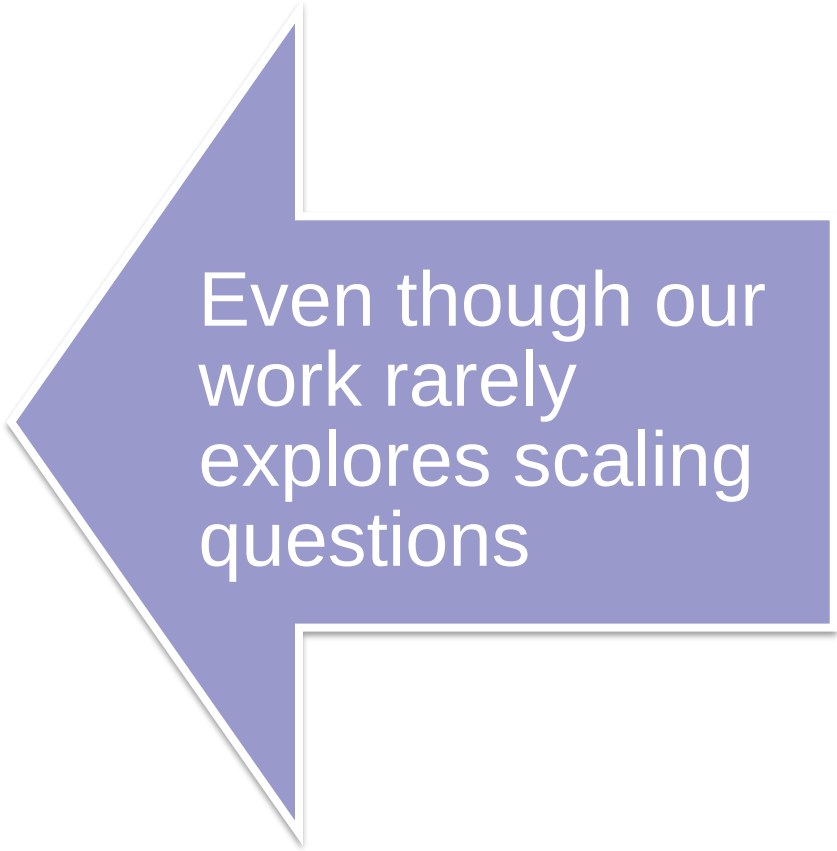
Disturbingly vague: bundles many different concepts that need to be understood before we can give advice to policymakers

External validity is about generalizing treatment effects

Scaling is about generalizing benefit-cost profiles



Speaking to Policymakers



Even though our
work rarely
explores scaling
questions



We still give
policy advice!

Some Recent Quotes



“Policymakers who want to lock down on colluders need to follow this science”

“Our work measuring trust has deep policy implications for social media regulation”

“This study will transform the lives of all the homeless if policymakers listen”

To give policy advice, we need an entirely different research program.



Option C Thinking = Policy-Based Evidence

Option C Thinking

Envisioning how the scaled
situational features depart from the
experimental setting and designing
appropriately

➔ Policy-Based Evidence

List, JA. “Optimally Generate Policy-
Based Evidence Before Scaling,”
Nature, 626, pp. 491-99, 2024.



4 Tips

4. Understand your contribution and state it clearly

Type	Definition	Examples	How to Address
Theoretical Gap	discrepancy between existing theories or models and observed phenomena	lack of studies examining the relationship between X and Y	conduct further research to test existing theories or develop new theoretical frameworks
Methodological Gap	insufficiency or inadequacy in the methods used to investigate a research question	absence of studies utilizing qualitative methods in the field	review and refine research methodologies, consider alternative approaches, or combine methods to address limitations
Empirical Gap	missing data or evidence needed to fully understand or explain a phenomenon	limited research on the long-term effects of treatment X	gather additional data through experiments, surveys, or longitudinal studies to fill gaps in knowledge
Conceptual Gap	lack of clarity or consensus regarding key concepts or definitions in the field	variation in definitions of 'success' across studies	clarify definitions through consensus-building efforts, standardization of terms, or development of clear conceptual frameworks
Temporal Gap	lack of research over a certain period, leaving a discontinuity in understanding	sparse studies examining the impact of recent technological advancements	conduct studies to address current gaps and ensure research keeps pace with the latest developments in the field
Spatial Gap	absence of research in specific geographical areas, limiting generalizability	few studies exploring the cultural differences in consumer behavior across regions	expand research efforts to include underrepresented geographical areas, ensuring a more comprehensive understanding of the phenomenon
Literature Gap	failure to address existing knowledge gaps or build upon prior research adequately	neglecting to consider recent advancements in the literature when designing a study	conduct a thorough literature review to identify existing gaps and build upon prior research findings to contribute to the advancement of knowledge

A Pressing Issue

Tyranny of the
Top 5



Some Discussion Issues

Tyranny of the Top 5

Referees acting as coauthors

One of the major changes ~ 25 years

Referees are much more strident on insisting that someone else's paper is written how they want it written. This is not about science. This is about style.

Other Issues

Tyranny of the Top 5

Referees acting as coauthors

GEN AI
Reviewers/Authors

One of the major changes ~ 25 years

Referees are much more strident on insisting that someone else's paper is written how they want it written. This is not about science. This is about style.

References

- List, John A. "Field Experiments: Here Today Gone Tomorrow?" *The American Economist*, forthcoming.
- List, John A. "Optimally Generate Policy-Based Evidence Before Scaling," *Nature*, 626: 491–499 (2024).
- List, John A, "The Voltage Effect," *Business Economics*, 58(1): 3-8 (2023).
- List, John A, "Experimental Economics: Theory and Practice," U. Chicago Press, 2025
- List, John A., Azeem Shaikh, and Atom Vayalinkal, ""Multiple Testing with Covariate Adjustment in Experimental Economics," *Journal of Applied Econometrics*, 38(6): 920-39 (2023).
- List, John A. "How Field Experiments in Economics can Complement Psychological Research in Judgement Biases," *Current Directions in Psychological Science* (2021), 30 (5) , pp. 431-436.
- List, John A. (2020) "Non est Disputandum de Generalizability? A. Glimpse into The External Validity Trial" NBER working paper.
- Gosnell, Greer, John A. List, & Robert Metcalfe, "[The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains](#)," *Journal of Political Economy*, (2020), 128(4), pp. 1195-1233.
- List, John A., Azeem Shaikh, and Yang Xu, "Multiple Hypothesis Testing in Experimental Economics," *Experimental Economics*, (2019), 22(4), pp. 773-793.
- Czibor, Eszter, David Jimenez-Gomez, & John A. List, "[The Dozen Things Experimental Economists Should Do \(More Of\)](#)," *Southern Economic Journal*, (2019), 86(2), pp. 371-432.