

# Jacob (Jake) Carlson

jacob.carlson@g.harvard.edu

<https://jscarlson.github.io>

## EDUCATION

**Harvard University**, Cambridge, Massachusetts  
Ph.D. Economics, August 2022 – May 2028 (expected)

**Yale University**, New Haven, Connecticut  
B.A. Statistics and Data Science, August 2016 – May 2020  
B.A. Ethics, Politics, and Economics, August 2016 – May 2020  
*Summa cum laude* with distinction in both subjects

## HONORS AND AWARDS





- Presidential Scholar (Harvard University)
- Phi Beta Kappa (Yale University)
- Forethought Foundation for Global Priorities Research Undergraduate Thesis Prize (2020)

## WORKING PAPERS

- Carlson, Jacob. “Making Interpretable Discoveries from Unstructured Data: A High-Dimensional Multiple Hypothesis Testing Approach.” Working paper (2025).
- Carlson, Jacob and Melissa Dell. “A Unifying Framework for Robust and Efficient Inference with Unstructured Data.” arXiv:2505.00282 (2025).
- Andrews, Isaiah, Harvey Barnhard, and Jacob Carlson. “True and Pseudo-True Parameters.” Working paper (2024).

## PUBLICATIONS

- Carlson, Jacob, Tom Bryan, and Melissa Dell. “Efficient OCR for Building a Diverse Digital History.” *ACL* (2024). Oral presentation.
- Dell, Melissa, Jacob Carlson, Tom Bryan, Emily Silcock, Abhishek Arora, Zejiang Shen, Luca D’Amico-Wong, Quan Le, Pablo Querubin, Leander Heldring. “American Stories: A Large-Scale Structured Text Dataset of Historical U.S. Newspapers.” *NeurIPS Datasets and Benchmarks Track* (2023).
- Bryan, Tom, Jacob Carlson, Abhishek Arora, Melissa Dell. “EfficientOCR: An Extensible, Open-Source Package for Efficiently Digitizing World Knowledge.” *EMNLP System Demonstrations* (2023).
- Carlson, Jacob, Trevor Incerti, and P. M. Aronow. “Dyadic Clustering in International Relations.” *Political Analysis* (2024).
- Shen, Zejiang, Ruochen Zhang, Melissa Dell, Benjamin Lee, Jacob Carlson, and Weining Li. “LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis.” *ICDAR* (2021). Oral presentation.

ACADEMIC POSITIONS	<b>Harvard University</b> , Cambridge, Massachusetts Research Assistant May 2023 – August 2024 <ul style="list-style-type: none"> <li>Assisted Isaiah Andrews and Davide Viviano with theoretical and applied econometrics projects</li> </ul>
	<b>Harvard University</b> , Cambridge, Massachusetts Data Science Fellow in Deep Learning September 2020 – May 2022 <ul style="list-style-type: none"> <li>Developed deep learning methods and tools for social science data curation with Melissa Dell</li> </ul>
	<b>Yale University</b> , New Haven, Connecticut Research Assistant May 2018 – December 2018 <ul style="list-style-type: none"> <li>Evaluated effectiveness of evidence-based policies across the U.S. with Yusuke Narita</li> </ul>
	<b>Yale Law School</b> , New Haven, Connecticut Research Assistant May 2017 – May 2019 <ul style="list-style-type: none"> <li>Researched mandatory minimum sentencing trends across DOJ administrations with Kate Stith</li> </ul>
TEACHING	<b>Harvard University</b> , Cambridge, Massachusetts <ul style="list-style-type: none"> <li><i>Causal Inference in Economics</i> (ECON 2148), Graduate, Spring 2025</li> <li><i>Time Series</i> (STAT 242), Graduate, Fall 2024, Fall 2025</li> <li><i>Unleashing Novel Data at Scale</i> (ECON 2355), Graduate, Spring 2021</li> </ul>
PRESENTATIONS	<b>ASSA 2022 Virtual Annual Meeting</b> <ul style="list-style-type: none"> <li>Presented “LayoutParser” at “Nonstandard Data Sources” paper session (January 8, 2022)</li> </ul>
SOFTWARE	<ul style="list-style-type: none"> <li>NetworkTargeting  (with Davide Viviano): <a href="https://github.com/dviviano/NetworkTargeting">https://github.com/dviviano/NetworkTargeting</a></li> <li>EffOCR  (with Tom Bryan and Melissa Dell): <a href="https://pypi.org/project/efficient-ocr">https://pypi.org/project/efficient-ocr</a></li> <li>dcr   (with Trevor Incerti and P.M. Aronow): <a href="https://github.com/jscarlson/dcr">https://github.com/jscarlson/dcr</a>, <a href="https://github.com/jscarlson/stata-dcr">https://github.com/jscarlson/stata-dcr</a></li> </ul>
ACADEMIC SERVICE	<i>Referee</i> : Journal of Political Economy (JPE)
LAST UPDATED	December 17, 2025