Online Appendix

A Systematic Review Process

My initial sample consists of all articles registered in Web of Science as published from 2015 onwards in a Top 5 economics journal (specifically *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*). I obtained bibliographic information on this set of 3732 articles from Web of Science on 28 July 2023. This bibliographic information is then loaded into ASReview, an interface that employs machine learning and text classification to assist with managing systematic literature reviews by sorting abstracts from most to least relevant (van de Schoot et al. 2021). I then manually reviewed the abstracts, classifying them as relevant if the abstract makes some claim that a phenomenon or relationship is either negligible or nonexistent. After reviewing 2987 abstracts, 50 consecutive abstracts were assessed to be irrelevant, and thus the remaining 745 articles are discarded as irrelevant based on ASReview's relevance probability ranking.[1] The abstract reviews yield 603 potentially relevant records, at which point all articles published prior to 2020 are discarded, ensuring that the sample reflects only the most recent practice in the economics literature and has the highest probability of reproducibility while still keeping the number of (attempted) reproductions down to a feasible level.[2] 287 potentially relevant articles published from 2020-2023 arise from this first phase of the systematic search.

I then examine the abstracts of each of these 287 potentially relevant articles, isolating every null claim made in each abstract and discarding an article if, upon further inspection, its abstract does not in fact make an identifiable null claim. This

[1] This is an intended feature of ASReview – the probability ranking permits early cessation of the review process with a strong reassurance that the most relevant articles still remain in the sample (van de Schoot et al. 2021).

[2] The additional articles from 2015-2019 help ensure the quality of the relevance probability ranking, and thus the irrelevance of discarded articles.

step produces 556 null claims across 285 articles. For each of these null claims, I attempt to locate the estimate(s) used to support that claim within the article. I discard a claim if it is not defended by at least one statistically insignificant estimate, otherwise storing the main estimate(s) being used to defend that claim. I discard articles if no null claims remain after this discarding process. This step yields my intermediate sample of 2346 estimates across 279 claims in 158 articles. Thereafter, I attempt to reproduce every estimate in the intermediate sample. Estimates are discarded when data is not available for reproduction or the reproduction is not conformable to my final analysis. After such discarding, my final sample consists of 876 estimates across 135 null claims in 81 articles.

B Final Sample

All publications included in the final sample are cited in these references. All publications in the final sample also are part of the intermediate sample. These references also cite repositories wherein the data for the final sample's articles are stored, when applicable. Data for articles without a separate repository is linked to the publisher's online version of the article itself. Bagues & Campa (2020), which is in the final sample, makes use of data from Casas-Arce & Saiz (2015), which is not in the final sample. Historical datasets in Bureau of Labor Statistics (2022) are cited at the direction of Gertler, Huckfeldt, & Trigari (2020).

References

Abebe, Girum et al. (2021). "Anonymity or distance? Job search and labour market exclusion in a growing African city". *The Review of Economic Studies* 88.3, pp. 1279–1310. DOI: 10.1093/restud/rdaa057.

Acemoglu, Daron, Giuseppe De Feo, Giacomo De Luca, et al. (2021). *Replication data for: War, socialism, and the rise of fascism: An empirical exploration*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/CLJTSC.

— (2022). "War, socialism, and the rise of fascism: An empirical exploration". *The Quarterly Journal of Economics* 137.2, pp. 1233–1296. DOI: 10.1093/qje/qjac001.

Acemoglu, Daron, Giuseppe De Feo, and Giacomo Davide De Luca (2020). "Weak states: Causes and consequences of the Sicilian mafia". *The Review of Economic Studies*, pp. 537–581. DOI: 10.1093/restud/rdz009.

Ager, Philipp, Leah Boustan, and Katherine Eriksson (2021a). *Data and code for: The intergenerational effects of a large wealth shock: White Southerners after the Civil War*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E138741V1.

Ager, Philipp, Leah Boustan, and Katherine Eriksson (2021b). "The intergenerational effects of a large wealth shock: White Southerners after the Civil War". *American Economic Review* 111.11, pp. 3767–3794. DOI: 10.1257/aer.20191422.

Akhtari, Mitra, Diana Moreira, and Laura Trucco (2022). "Political turnover, bureaucratic turnover, and the quality of public services". *American Economic Review* 112.2, pp. 442–493. DOI: 10.1257/aer.20171867.

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva (2022). *Replication Package for Immigration and Redistribution*. Dataset V1. Geneva, Switzerland: Zenodo. DOI: 10.5281/zenodo.5997521.

— (2023). "Immigration and redistribution". *The Review of Economic Studies* 90.1, pp. 1–39. DOI: 10.1093/restud/rdac011.

Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden (2020). "Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians?" *Journal of Political Economy* 128.5, pp. 1753–1788. DOI: 10.1086/705551.

Andrabi, Tahir et al. (2020a). *Data and code for: Upping the ante: The equilibrium effects of unconditional grants to private schools*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E118805V1.

— (2020b). "Upping the ante: The equilibrium effects of unconditional grants to private schools". *American Economic Review* 110.10, pp. 3315–3349. DOI: 10.1257/aer.20180924.

Arbatli, Cemal Eren et al. (2020). "Diversity and conflict". *Econometrica* 88.2, pp. 727–797. DOI: 10.3982/ECTA13734.

Asher, Sam and Paul Novosad (2020a). *Data and code for: Rural roads and local economic development*. Dataset V2. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E109703V2.

Asher, Sam and Paul Novosad (2020b). "Rural roads and local economic development". *American Economic Review* 110.3, pp. 797–823. DOI: 10.1257/aer.20180268.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, et al. (2020). "Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services". *American Economic Review* 110.5, pp. 1355–1394. DOI: 10.1257/aer.20180326.

Ashraf, Nava, Oriana Bandiera, Scott S. Lee, et al. (2020). *Data and code for: Losing prosociality in the quest for talent*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E111683V1.

Ashraf, Nava, Natalie Bau, Corinne Low, et al. (2019). *Replication data for: 'Negotiating a better future: How interpersonal skills facilitate intergenerational investment'*. Dataset V3. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/IJE4RJ.

— (2020). "Negotiating a better future: How interpersonal skills facilitate intergenerational investment". *The Quarterly Journal of Economics* 135.2, pp. 1095–1151. DOI: 10.1093/qje/qjz039.

Ashraf, Nava, Natalie Bau, Nathan Nunn, et al. (2020). "Bride price and female education". *Journal of Political Economy* 128.2, pp. 591–641. DOI: 10.1086/704572.

Attanasio, Orazio and Elena Pastorino (2020). "Nonlinear pricing in village economies". *Econometrica* 88.1, pp. 207–263. DOI: 10.3982/ECTA13918.

Bagues, Manuel and Pamela Campa (2020). "Women and power: Unpopular, unwilling, or held back? A comment". *Journal of Political Economy* 128.5, pp. 2010–2016. DOI: 10.1086/705669.

Balán, Pablo et al. (2022). "Local elites as state capacity: How city chiefs use local information to increase tax compliance in the Democratic Republic of the Congo". *American Economic Review* 112.3, pp. 762–797. DOI: 10.1257/aer.20201159.

Bandiera, Oriana et al. (2021a). *Replication data for: 'The allocation of authority in organizations: A field experiment with bureaucrats'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/OJQWO2.

— (2021b). "The allocation of authority in organizations: A field experiment with bureaucrats". *The Quarterly Journal of Economics* 136.4, pp. 2195–2242. DOI: 10.1093/qje/qjab029.

Bazzi, Samuel, Gabriel Koehler-Derrick, and Benjamin Marx (2019). *Replication data for: 'The institutional foundations of religious politics: Evidence from Indonesia'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/OY4SM9.

— (2020). "The institutional foundations of religious politics: Evidence from Indonesia". *The Quarterly Journal of Economics* 135.2, pp. 845–911. DOI: 10.1093/qje/qjz038.

Becker, Sascha O. et al. (2020a). *Data and code for: Forced migration and human capital: Evidence from post-WWII population transfers*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E115202V1.

— (2020b). "Forced migration and human capital: Evidence from post-WWII population transfers". *American Economic Review* 110.5, pp. 1430–1463. DOI: 10.1257/aer.20181518.

Beraja, Martin et al. (2023a). "AI-tocracy". *The Quarterly Journal of Economics* 138.3, pp. 1349–1402. DOI: 10.1093/qje/qjad012.

— (2023b). *Replication data for: 'AI-tocracy'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/GCOVGX.

Bergquist, Lauren Falcao and Michael Dinerstein (2020a). “Competition and entry in agricultural markets: Experimental evidence from Kenya”. *American Economic Review* 110.12, pp. 3705–3747. DOI: 10.1257/aer.20171397.

— (2020b). *Data and code for: Competition and entry in agricultural markets: Experimental evidence from Kenya*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E119743V1.

Berkouwer, Susanna B. and Joshua T. Dean (2022a). “Credit, attention, and externalities in the adoption of energy efficient technologies by low-income households”. *American Economic Review* 112.10, pp. 3291–3330. DOI: 10.1257/aer.20210766.

— (2022b). *Data and code for: Credit, attention, and externalities in the adoption of energy efficient technologies by low-income household*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E166661V1.

Bessone, Pedro et al. (2021a). *Replication data for: 'The economic consequences of increasing sleep among the urban poor'*. Dataset V2. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/GJ9QPC.

— (2021b). “The economic consequences of increasing sleep among the urban poor”. *The Quarterly Journal of Economics* 136.3, pp. 1887–1941. DOI: 10.1093/qje/qjab013.

Blakeslee, David, Ram Fishman, and Veena Srinivasan (2020a). *Replication package for: Way down in the hole: Adaptation to long-term water loss in rural India*. Dataset V1. Nashville, TN, U.S.A.: American Economic Association. URL: https://www.aeaweb.org/journals/dataset?id=10.1257/aer.20180976.

— (2020b). “Way down in the hole: Adaptation to long-term water loss in rural India”. *American Economic Review* 110.1, pp. 200–224. DOI: 10.1257/aer.20180976.

Bold, Tessa et al. (2022a). *Data and code for: Market access and quality upgrading: Evidence from four field experiments*. Dataset V1. Ann Arbor, MI, U.S.A.:

Inter-university Consortium for Political and Social Research. DOI: 10.3886/E158401V1.

Bold, Tessa et al. (Aug. 2022b). "Market access and quality upgrading: Evidence from four field experiments". *American Economic Review* 112.8, pp. 2518–2552. DOI: 10.1257/aer.20210122.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani (2021a). *Data and code for: Labor rationing*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E141441V1.

— (2021b). "Labor rationing". *American Economic Review* 111.10, pp. 3184–3224. DOI: 10.1257/aer.20201385.

Brocas, Isabelle and Juan D. Carrillo (2021). "Steps of reasoning in children and adolescents". *Journal of Political Economy* 129.7, pp. 2067–2111. DOI: 10.1086/714118.

— (2024). *"Steps of reasoning in children and adolescents"*. Dataset 695096b. San Francisco, CA, U.S.A.: Github. URL: https://github.com/labelinstitute/dev_DM/tree/main/Levels.

Brodeur, Abel, Nikolai Cook, and Anthony Heyes (2020). "Methods matter: p-hacking and publication bias in causal analysis in economics". *American Economic Review* 110.11, pp. 3634–3660. DOI: 10.1257/aer.20190687.

— (2022). *Data and code for: Methods matter: P-hacking and publication bias in causal analysis in economics*. Dataset V2. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E120246V1.

Brownback, Andy and Sally Sadoff (2020). "Improving college instruction through incentives". *Journal of Political Economy* 128.8, pp. 2925–2972. DOI: 10.1086/707025.

Bryan, Gharad, James J. Choi, and Dean Karlan (2020). *Replication data for: 'Randomizing religion: The impact of Protestant evangelism on economic outcomes'*.

Dataset V3. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/RNGHDV.

Bryan, Gharad, James J. Choi, and Dean Karlan (2021). "Randomizing religion: The impact of Protestant evangelism on economic outcomes". *The Quarterly Journal of Economics* 136.1, pp. 293–380. DOI: 10.1093/qje/qjaa023.

Bureau of Labor Statistics, United States Census Bureau (2022). *Survey of Income and Program Participation Datasets*. Datasets 1990-2008. Suitland, MD, U.S.A.: United States Census Bureau. DOI: 10.7910/DVN/OQNZYE.

Byrne, David P, Leslie A Martin, and Jia Sheen Nah (2022a). "Price discrimination by negotiation: A field experiment in retail electricity". *The Quarterly Journal of Economics* 137.4, pp. 2499–2537. DOI: 10.1093/qje/qjac021.

— (2022b). *Replication data for: 'Price discrimination by negotiation: A field experiment in retail electricity'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/KRHAWJ.

Campbell, Douglas L and Karsten Mau (2020). *Replication files for "On 'Trade induced technical change: The impact of Chinese imports on innovation, IT, and productivity'"*. Dataset V1. Geneva, Switzerland: Zenodo. DOI: 10.5281/zenodo.3972652.

— (2021). "On "Trade induced technical change: The impact of Chinese imports on innovation, IT, and productivity"". *The Review of Economic Studies* 88.5, pp. 2555–2559. DOI: 10.1093/restud/rdab037.

Caprettini, Bruno and Hans-Joachim Voth (2022). *Replication data for: 'New Deal, new patriots: How 1930s government spending boosted patriotism during WWII'*. Dataset V2. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/3A8CBI.

— (2023). "New Deal, new patriots: How 1930s government spending boosted patriotism during World War II". *The Quarterly Journal of Economics* 138.1, pp. 465–513. DOI: 10.1093/qje/qjac028.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti (2022). "Goals and gaps: Educational careers of immigrant children". *Econometrica* 90.1, pp. 1–29. DOI: 10.3982/ECTA17458.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, Afra Sial, et al. (2021). *Replication package for: "Who chooses commitment? Evidence and welfare implications"*. Dataset V1. Geneva, Switzerland: Zenodo. DOI: 10.5281/zenodo.5173081.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky (2021). "Who chooses commitment? Evidence and welfare implications". *The Review of Economic Studies* 89.3, pp. 1205–1244. DOI: 10.1093/restud/rdab056.

Casas-Arce, Pablo and Albert Saiz (2015). "Women and power: Unpopular, unwilling, or held back?" *Journal of Political Economy* 123.3, pp. 641–669. DOI: 10.1086/680686.

Chew, Soo Hong, Wei Huang, and Xiaojian Zhao (2020). "Motivated false memory". *Journal of Political Economy* 128.10, pp. 3913–3939. DOI: 10.1086/709971.

Chodorow-Reich, Gabriel, Plamen T. Nenov, and Alp Simsek (2021a). *Data and code for "Stock market wealth and the real economy: A local labor market approach"*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E123521V1.

— (2021b). "Stock market wealth and the real economy: A local labor market approach". *American Economic Review* 111.5, pp. 1613–1657. DOI: 10.1257/aer.20200208.

Corno, Lucia, Eliana La Ferrara, and Justine Burns (2022a). *Data and code for: 'Interaction, stereotypes, and performance. Evidence from South Africa*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E174501V1.

Corno, Lucia, Eliana La Ferrara, and Justine Burns (2022b). "Interaction, stereotypes, and performance: Evidence from South Africa". *American Economic Review* 112.12, pp. 3848–3875. DOI: 10.1257/aer.20181805.

DellaVigna, Stefano et al. (2021). *Data and code for: "Estimating social preferences and gift exchange at work"*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E148481V1.

— (2022). "Estimating social preferences and gift exchange at work". *American Economic Review* 112.3, pp. 1038–1074. DOI: 10.1257/aer.20190920.

Derenoncourt, Ellora and Claire Montialoux (2020). *Replication data for: 'Minimum wages and racial inequality'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/MHNS1S.

— (2021). "Minimum wages and racial inequality". *The Quarterly Journal of Economics* 136.1, pp. 169–228. DOI: 10.1093/qje/qjaa031.

Dhar, Diva, Tarun Jain, and Seema Jayachandran (2022a). *Data and code for: Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E149882V1.

— (2022b). "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India". *American Economic Review* 112.3, pp. 899–927. DOI: 10.1257/aer.20201112.

Djourelova, Milena (2023a). *Data and code for: Persuasion through slanted language: Evidence from the media coverage of immigration*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E182482V1.

— (2023b). "Persuasion through slanted language: Evidence from the media coverage of immigration". *American Economic Review* 113.3, pp. 800–835. DOI: 10.1257/aer.20211537.

Egger, Dennis et al. (2022). "General equilibrium effects of cash transfers: Experimental evidence from Kenya". *Econometrica* 90.6, pp. 2603–2643. DOI: 10.3982/ECTA17945.

Eichenbaum, M S, B K Johannsen, and S T Rebelo (2020). "Monetary policy and the predictability of nominal exchange rates". *The Review of Economic Studies* 88.1, pp. 192–228. DOI: 10.1093/restud/rdaa024.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova (2020). "Social media and protest participation: Evidence from Russia". *Econometrica* 88.4, pp. 1479–1514. DOI: 10.3982/ECTA14281.

Exley, Christine L and Judd B Kessler (2022a). *Replication data for: 'The gender gap in self-promotion'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/YSWKHY.

— (2022b). "The gender gap in self-promotion". *The Quarterly Journal of Economics* 137.3, pp. 1345–1381. DOI: 10.1093/qje/qjac003.

Exley, Christine L., Muriel Niederle, and Lise Vesterlund (2020). "Knowing when to ask: The cost of leaning in". *Journal of Political Economy* 128.3, pp. 816–854. DOI: 10.1086/704616.

Fajgelbaum, Pablo D et al. (2020a). *Replication data for: 'The return to protectionism'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/KSOVSE.

— (2020b). "The return to protectionism". *The Quarterly Journal of Economics* 135.1, pp. 1–55. DOI: 10.1093/qje/qjz036.

Fé, Eduardo, David Gill, and Victoria Prowse (2022). "Cognitive skills, strategic sophistication, and life outcomes". *Journal of Political Economy* 130.10, pp. 2643–2704. DOI: 10.1086/720460.

Fehr, Dietmar, Günther Fink, and B. Kelsey Jack (2022). "Poor and rational: Decision-making under scarcity". *Journal of Political Economy* 130.11, pp. 2862–2897. DOI: 10.1086/720466.

Flückiger, Matthias et al. (2022a). *Replication package for: Roman transport network connectivity and economic integration*. Dataset V1. Geneva, Switzerland: Zenodo. DOI: 10.5281/zenodo.4788227.

— (2022b). "Roman transport network connectivity and economic integration". *The Review of Economic Studies* 89.2, pp. 774–810. DOI: 10.1093/restud/rdab036.

Fuster, Andreas, Greg Kaplan, and Basit Zafar (2021). "What would you do with $500? Spending responses to gains, losses, news, and loans". *The Review of Economic Studies* 88.4, pp. 1760–1795. DOI: 10.1093/restud/rdaa076.

— (2022). *Replication package for: "What would you do with $500? Spending responses to gains, losses, news, and loans"*. Dataset V1. Geneva, Switzerland: Zenodo. DOI: 10.5281/zenodo.4115399.

Gertler, Mark, Christopher Huckfeldt, and Antonella Trigari (2020). "Unemployment fluctuations, match quality, and the wage cyclicality of new hires". *The Review of Economic Studies* 87.4, pp. 1876–1914. DOI: 10.1093/restud/rdaa004.

Giorcelli, Michela and Petra Moser (2020). "Copyrights and creativity: Evidence from Italian opera in the Napoleonic age". *Journal of Political Economy* 128.11, pp. 4163–4210. DOI: 10.1086/710534.

Grosjean, Pauline, Federico Masera, and Hasin Yousaf (2020). *Replication data for: 'Inflammatory political campaigns and racial bias in policing'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/A3B9HE.

— (2023). "Inflammatory political campaigns and racial bias in policing". *The Quarterly Journal of Economics* 138.1, pp. 413–463. DOI: 10.1093/qje/qjac037.

Guarnieri, Eleonora and Ana Tur-Prats (2023a). "Cultural distance and conflict-related sexual violence". *The Quarterly Journal of Economics* 138.3, pp. 1817–1861. DOI: 10.1093/qje/qjad015.

— (2023b). *Replication data for: 'Cultural distance and conflict-related sexual violence'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/B3LJGQ.

Hartley, Robert Paul, Carlos Lamarche, and James P. Ziliak (2022). "Welfare reform and the intergenerational transmission of dependence". *Journal of Political Economy* 130.3, pp. 523–565. DOI: 10.1086/717893.

Hau, Harald, Yi Huang, and Gewei Wang (2020). "Firm response to competitive shocks: Evidence from China's minimum wage policy". *The Review of Economic Studies* 87.6, pp. 2639–2671. DOI: 10.1093/restud/rdz058.

Hazell, Jonathon et al. (2022a). *Replication data for: 'The slope of the Phillips curve: Evidence from U.S. states'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/OQNZYE.

— (2022b). "The slope of the Phillips curve: Evidence from U.S. states". *The Quarterly Journal of Economics* 137.3, pp. 1299–1344. DOI: 10.1093/qje/qjac010.

He, Guojun, Shaoda Wang, and Bing Zhang (2020a). *Replication data for: 'Watering down environmental regulation in China'*. Dataset V3. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/LVS8VX.

— (2020b). "Watering down environmental regulation in China". *The Quarterly Journal of Economics* 135.4, pp. 2135–2185. DOI: 10.1093/qje/qjaa024.

Huber, Kilian (2021). "Are bigger banks better? Firm-level evidence from Germany". *Journal of Political Economy* 129.7, pp. 2023–2066. DOI: 10.1086/714120.

Jack, William et al. (2023a). "Credit access, selection, and incentives in a market for asset-collateralized loans: Evidence from Kenya". *Review of Economic Studies* 90.6, pp. 3153–3185. DOI: 10.1093/restud/rdad026.

— (2023b). *Replication package for: Credit access, selection, and incentives in a market for asset-collateralized loans: Evidence from Kenya.* Dataset V2. Geneva, Switzerland: Zenodo. DOI: 10.5281/zenodo.7594227.

Jordà, Òscar et al. (2021). "Bank capital redux: Solvency, liquidity, and crisis". *The Review of Economic Studies* 88.1, pp. 260–286. DOI: 10.1093/restud/rdaa040.

Kelly, Morgan, Joel Mokyr, and Cormac Ó Gráda (2023). "The mechanics of the Industrial Revolution". *Journal of Political Economy* 131.1, pp. 59–94. DOI: 10.1086/720890.

Kline, Patrick, Evan K Rose, and Christopher R Walters (2022a). *Replication data for: 'Systemic discrimination among large U.S. employers'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/HLO4XC.

— (2022b). "Systemic discrimination among large U.S. employers". *The Quarterly Journal of Economics* 137.4, pp. 1963–2036. DOI: 10.1093/qje/qjac024.

Kosse, Fabian et al. (2020). "The formation of prosociality: Causal evidence on the role of social environment". *Journal of Political Economy* 128.2, pp. 434–467. DOI: 10.1086/704386.

Kranz, Sebastian and Peter Pütz (2022a). *Data and code for: Methods matter: p-hacking and publication bias in causal analysis in economics: Comment*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E159221V1.

— (2022b). "Methods matter: p-hacking and publication bias in causal analysis in economics: Comment". *American Economic Review* 112.9, pp. 3124–3136. DOI: 10.1257/aer.20210121.

Le Pennec, Caroline and Vincent Pons (2022). *Replication data for: 'How do campaigns shape vote choice? Multicountry evidence from 62 elections and 56 TV debates'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/XMDFQO.

— (2023). "How do campaigns shape vote choice? Multicountry evidence from 62 elections and 56 TV debates". *The Quarterly Journal of Economics* 138.2, pp. 703–767. DOI: 10.1093/qje/qjad002.

Lee, Kenneth, Edward Miguel, and Catherine Wolfram (2020). "Experimental evidence on the economics of rural electrification". *Journal of Political Economy* 128.4, pp. 1523–1565. DOI: 10.1086/705417.

Li, Xiaomin and Colin F Camerer (2022a). "Predictable effects of visual salience in experimental decisions and games". *The Quarterly Journal of Economics* 137.3, pp. 1849–1900. DOI: 10.1093/qje/qjac025.

— (2022b). *Replication data for: 'Predictable effects of visual salience in experimental decisions and games'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/9LCYKG.

Mayshar, Joram, Omer Moav, and Luigi Pascali (2022). "The origin of the state: Land productivity or appropriability?" *Journal of Political Economy* 130.4, pp. 1091–1144. DOI: 10.1086/718372.

Moreira, Diana, Mitra Akhtari, and Laura Trucco (2021). *Data and code for: Political turnover, bureaucratic turnover, and the quality of public services*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E150323V1.

Moscona, Jacob and Karthik A Sastry (2022). *Replication data for: 'Does directed innovation mitigate climate damage? Evidence from U.S. agriculture'*. Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/5ELEPA.

— (2023). "Does directed innovation mitigate climate damage? Evidence from U.S. agriculture". *The Quarterly Journal of Economics* 138.2, pp. 637–701. DOI: 10.1093/qje/qjac039.

Mueller, Andreas I., Johannes Spinnewijn, and Giorgio Topa (2020). *Data and codes for: "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias"*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E120501V1.

— (2021). "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias". *American Economic Review* 111.1, pp. 324–363. DOI: 10.1257/aer.20190808.

Okeke, Edward N. (2023a). *Data and code for: "When a doctor falls from the sky: The impact of easing doctor supply constraints on mortality"*. Dataset V1. Ann

Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E181581V1.

Okeke, Edward N. (2023b). "When a doctor falls from the sky: The impact of easing doctor supply constraints on mortality". *American Economic Review* 113.3, pp. 585–627. DOI: 10.1257/aer.20210701.

Romero, Mauricio, Justin Sandefur, and Wayne Sandholtz (2018). *Partnership schools for Liberia*. Dataset V4. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/5OPIYU.

Romero, Mauricio, Justin Sandefur, and Wayne Aaron Sandholtz (Feb. 2020). "Outsourcing education: Experimental evidence from Liberia". *American Economic Review* 110.2, pp. 364–400. DOI: 10.1257/aer.20181478.

Sadoff, Sally, Anya Samek, and Charles Sprenger (2020). "Dynamic inconsistency in food choice: Experimental evidence from two food deserts". *The Review of Economic Studies* 87.4, pp. 1954–1988. DOI: 10.1093/restud/rdz030.

Sánchez de la Sierra, Raúl (2021). "Whither formal contracts?" *Econometrica* 89.5, pp. 2341–2373. DOI: 10.3982/ECTA16083.

Sarsons, Heather et al. (2021). "Gender differences in recognition for group work". *Journal of Political Economy* 129.1, pp. 101–147. DOI: 10.1086/711401.

Stantcheva, Stefanie (2021a). *Replication data for: 'Understanding tax policy: How do people reason?'* Dataset V1. Cambridge, MA, U.S.A.: Harvard Dataverse. DOI: 10.7910/DVN/OAHUIP.

— (2021b). "Understanding tax policy: How do people reason?" *The Quarterly Journal of Economics* 136.4, pp. 2309–2369. DOI: 10.1093/qje/qjab033.

Tabellini, Marco (2020). "Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration". *The Review of Economic Studies* 87.1, pp. 454–486. DOI: 10.1093/restud/rdz027.

Weidmann, Ben and David J. Deming (2021). “Team players: How social skills improve team performance”. *Econometrica* 89.6, pp. 2637–2657. DOI: 10.3982/ECTA18461.

Weigel, Jonathan et al. (2022). *Replication data for: Local elites as state capacity: How city chiefs use local information to increase tax compliance in the D.R. Congo.* Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E147561V1.

C Intermediate Sample

The following publications are included in the intermediate sample, but are not included in the final sample.

References

Abaluck, Jason et al. (2021). "Mortality effects and choice across private health insurance plans". *The Quarterly Journal of Economics* 136.3, pp. 1557–1610. DOI: 10.1093/qje/qjab017.

Abdulkadiroğlu, Atila et al. (2020). "Do parents value school effectiveness?" *American Economic Review* 110.5, pp. 1502–1539. DOI: 10.1257/aer.20172040.

Aggeborn, Linuz and Mattias Öhman (2021). "The effects of fluoride in drinking water". *Journal of Political Economy* 129.2, pp. 465–491. DOI: 10.1086/711915.

Akcigit, Ufuk, Salomé Baslandze, and Francesca Lotti (2023). "Connecting to power: Political connections, innovation, and firm dynamics". *Econometrica* 91.2, pp. 529–564. DOI: 10.3982/ecta18338.

Alexander, Diane (2020). "How do doctors respond to incentives? Unintended consequences of paying doctors to reduce costs". *Journal of Political Economy* 128.11, pp. 4046–4096. DOI: 10.1086/710334.

Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P Vasquez (2022). "The effects of joining multinational supply chains: New evidence from firm-to-firm linkages". *The Quarterly Journal of Economics* 137.3, pp. 1495–1552. DOI: 10.1093/qje/qjac006.

Allcott, Hunt et al. (2022). "Are high-interest loans predatory? Theory and evidence from payday lending". *The Review of Economic Studies* 89.3, pp. 1041–1084. DOI: 10.1093/restud/rdab066.

Alvarez, Fernando and David Argente (2022). "On the effects of the availability of means of payments: The case of Uber". *The Quarterly Journal of Economics* 137.3, pp. 1737–1789. DOI: 10.1093/qje/qjac008.

Ang, Desmond (2020). "The effects of police violence on inner-city students". *The Quarterly Journal of Economics* 136.1, pp. 115–168. DOI: 10.1093/qje/qjaa027.

Angelucci, Manuela and Daniel Bennett (2021). "Adverse selection in the marriage market: HIV testing and marriage in rural Malawi". *The Review of Economic Studies* 88.5, pp. 2119–2148. DOI: 10.1093/restud/rdaa088.

Aucejo, Esteban and Jonathan James (2021). "The path to college education: The role of math and verbal skills". *Journal of Political Economy* 129.10, pp. 2905–2946. DOI: 10.1086/715417.

Backus, Matthew (2020). "Why is productivity correlated with competition?" *Econometrica* 88.6, pp. 2415–2444. DOI: 10.3982/ecta12926.

Bahaj, Saleem, Angus Foulis, and Gabor Pinter (2020). "Home values and firm behavior". *American Economic Review* 110.7, pp. 2225–2270. DOI: 10.1257/aer.20180649.

Bald, Anthony et al. (2022). "The causal impact of removing children from abusive and neglectful homes". *Journal of Political Economy* 130.7, pp. 1919–1962. DOI: 10.1086/719856.

Beerli, Andreas et al. (2021). "The abolition of immigration restrictions and the performance of firms and workers: Evidence from Switzerland". *American Economic Review* 111.3, pp. 976–1012. DOI: 10.1257/aer.20181779.

Berger, David, Ian Dew-Becker, and Stefano Giglio (2020). "Uncertainty shocks as second-moment news shocks". *The Review of Economic Studies* 87.1, pp. 40–76. DOI: 10.1093/restud/rdz010.

Bianchi, Nicola et al. (2023). "Career spillovers in internal labour markets". *The Review of Economic Studies* 90.4, pp. 1800–1831. DOI: 10.1093/restud/rdac067.

Biasi, Barbara and Heather Sarsons (2022). “Flexible wages, bargaining, and the gender gap”. *The Quarterly Journal of Economics* 137.1, pp. 215–266. DOI: 10.1093/qje/qjab026.

Bleemer, Zachary (2022). “Affirmative action, mismatch, and economic mobility after California’s proposition 209”. *The Quarterly Journal of Economics* 137.1, pp. 115–160. DOI: 10.1093/qje/qjab027.

Britto, Diogo G., Paolo Pinotti, and Breno Sampaio (2022). “The effect of job loss and unemployment insurance on crime in Brazil”. *Econometrica* 90.4, pp. 1393–1423. DOI: 10.3982/ecta18984.

Bulman, George et al. (2021). “Parental resources and college attendance: Evidence from lottery wins”. *American Economic Review* 111.4, pp. 1201–1240. DOI: 10.1257/aer.20171272.

Butters, R. Andrew, Daniel W. Sacks, and Boyoung Seo (2022). “How do national firms respond to local cost shocks?” *American Economic Review* 112.5, pp. 1737–1772. DOI: 10.1257/aer.20201524.

Cantoni, Enrico and Vincent Pons (2021). “Strict ID laws don’t stop voters: Evidence from a U.S. nationwide panel, 2008–2018”. *The Quarterly Journal of Economics* 136.4, pp. 2615–2660. DOI: 10.1093/qje/qjab019.

Card, David et al. (2020). “Are referees and editors in economics gender neutral?” *The Quarterly Journal of Economics* 135.1, pp. 269–327. DOI: 10.1093/qje/qjz035.

Chodorow-Reich, Gabriel and Johannes Wieland (2020). “Secular labor reallocation and business cycles”. *Journal of Political Economy* 128.6, pp. 2245–2287. DOI: 10.1086/705717.

Cloyne, James, Clodomiro Ferreira, and Paolo Surico (2020). “Monetary policy when households have debt: New evidence on the transmission mechanism”. *The Review of Economic Studies* 87.1, pp. 102–129. DOI: 10.1093/restud/rdy074.

Coibion, Olivier, Yuriy Gorodnichenko, and Tiziano Ropele (2020). "Inflation expectations and firm decisions: New causal evidence". *The Quarterly Journal of Economics* 135.1, pp. 165–219. DOI: 10.1093/qje/qjz029.

Cook, Cody et al. (2021). "The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers". *The Review of Economic Studies* 88.5, pp. 2210–2238. DOI: 10.1093/restud/rdaa081.

D'Acunto, Francesco, Daniel Hoang, et al. (2023). "IQ, expectations, and choice". *The Review of Economic Studies* 90.5, pp. 2292–2325. DOI: 10.1093/restud/rdac075.

D'Acunto, Francesco, Ulrike Malmendier, et al. (2021). "Exposure to grocery prices and inflation expectations". *Journal of Political Economy* 129.5, pp. 1615–1639. DOI: 10.1086/713192.

Dahl, Gordon B, Andreas Kotsadam, and Dan-Olof Rooth (2020). "Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams". *The Quarterly Journal of Economics* 136.2, pp. 987–1030. DOI: 10.1093/qje/qjaa047.

Daruich, Diego, Sabrina Di Addario, and Raffaele Saggio (2023). "The effects of partial employment protection reforms: Evidence from Italy". *Review of Economic Studies* 90.6, pp. 2880–2942. DOI: 10.1093/restud/rdad012.

De Neve, Jan-Emmanuel et al. (2021). "How to improve tax compliance? Evidence from population-wide experiments in Belgium". *Journal of Political Economy* 129.5, pp. 1425–1463. DOI: 10.1086/713096.

DellaVigna, Stefano et al. (2022). "Evidence on job search models from a survey of unemployed workers in Germany". *The Quarterly Journal of Economics* 137.2, pp. 1181–1232. DOI: 10.1093/qje/qjab039.

Deryugina, Tatyana and David Molitor (2020). "Does when you die depend on where you live? Evidence from Hurricane Katrina". *American Economic Review* 110.11, pp. 3602–3633. DOI: 10.1257/aer.20181026.

Deshpande, Manasi and Michael Mueller-Smith (2022). "Does welfare prevent crime? The criminal justice outcomes of youth removed from SSI". *The Quarterly Journal of Economics* 137.4, pp. 2263–2307. DOI: 10.1093/qje/qjac017.

Dobbie, Will, Andres Liberman, et al. (2021). "Measuring bias in consumer lending". *The Review of Economic Studies* 88.6, pp. 2799–2832. DOI: 10.1093/restud/rdaa078.

Dobbie, Will and Jae Song (2020). "Targeted debt relief and the origins of financial distress: Experimental evidence from distressed credit card borrowers". *American Economic Review* 110.4, pp. 984–1018. DOI: 10.1257/aer.20171541.

Doran, Kirk, Alexander Gelber, and Adam Isen (2022). "The effects of high-skilled immigration policy on firms: Evidence from Visa lotteries". *Journal of Political Economy* 130.10, pp. 2501–2533. DOI: 10.1086/720467.

Eliason, Paul J et al. (2020). "How acquisitions affect firm behavior and performance: Evidence from the dialysis industry". *The Quarterly Journal of Economics* 135.1, pp. 221–267. DOI: 10.1093/qje/qjz034.

Feigenberg, Benjamin and Conrad Miller (2022). "Would eliminating racial disparities in motor vehicle searches have efficiency costs?" *The Quarterly Journal of Economics* 137.1, pp. 49–113. DOI: 10.1093/qje/qjab018.

Figlio, David et al. (2023). "Diversity in schools: Immigrants and the educational performance of U.S.-born students". *Review of Economic Studies*. DOI: 10.1093/restud/rdad047.

Foote, Christopher L, Lara Loewenstein, and Paul S Willen (2020). "Cross-sectional patterns of mortgage debt during the Housing Boom: Evidence and implications". *The Review of Economic Studies* 88.1, pp. 229–259. DOI: 10.1093/restud/rdaa034.

Friedrich, Benjamin U and Martin B Hackmann (2021). "The returns to nursing: Evidence from a parental-leave program". *The Review of Economic Studies* 88.5, pp. 2308–2343. DOI: 10.1093/restud/rdaa082.

Ganong, Peter and Pascal Noel (2020). "Liquidity versus wealth in household debt obligations: Evidence from housing policy in the Great Recession". *American Economic Review* 110.10, pp. 3100–3138. DOI: 10.1257/aer.20181243.

Giglio, Stefano et al. (2021). "Five facts about beliefs and portfolios". *American Economic Review* 111.5, pp. 1481–1522. DOI: 10.1257/aer.20200243.

Gopinath, Gita et al. (2020). "Dominant currency paradigm". *American Economic Review* 110.3, pp. 677–719. DOI: 10.1257/aer.20171201.

Gray-Lobe, Guthrie, Parag A Pathak, and Christopher R Walters (2023). "The long-term effects of universal preschool in Boston". *The Quarterly Journal of Economics* 138.1, pp. 363–411. DOI: 10.1093/qje/qjac036.

Greenberg, Kyle et al. (2022). "Army service in the all-volunteer era". *The Quarterly Journal of Economics* 137.4, pp. 2363–2418. DOI: 10.1093/qje/qjac026.

Grennan, Matthew and Ashley Swanson (2020). "Transparency and negotiated prices: The value of information in hospital-supplier bargaining". *Journal of Political Economy* 128.4, pp. 1234–1268. DOI: 10.1086/705329.

Grennan, Matthew and Robert J. Town (2020). "Regulating innovation with uncertain quality: Information, risk, and access in medical devices". *American Economic Review* 110.1, pp. 120–161. DOI: 10.1257/aer.20180946.

Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya (2021). "3G internet and confidence in government". *The Quarterly Journal of Economics* 136.4, pp. 2533–2613. DOI: 10.1093/qje/qjaa040.

Han, Jin Soo et al. (2021). "When does regulation distort costs? Lessons from fuel procurement in US electricity generation: Comment". *American Economic Review* 111.4, pp. 1356–1372. DOI: 10.1257/aer.20200679.

Helm, Ines (2020). "National industry trade shocks, local labour markets, and agglomeration spillovers". *The Review of Economic Studies* 87.3, pp. 1399–1431. DOI: 10.1093/restud/rdz056.

Herrera, Helios, Guillermo Ordoñez, and Christoph Trebesch (Feb. 2020). "Political booms, financial crises". *Journal of Political Economy* 128.2, pp. 507–543. DOI: 10.1086/704544.

Hoffman, Mitchell and Steven Tadelis (2021). "People management skills, employee attrition, and manager rewards: An empirical analysis". *Journal of Political Economy* 129.1, pp. 243–285. DOI: 10.1086/711409.

Hvidberg, Kristoffer B, Claus T Kreiner, and Stefanie Stantcheva (2023). "Social positions and fairness views on inequality". *Review of Economic Studies* 90.6, pp. 3083–3118. DOI: 10.1093/restud/rdad019.

Jäger, Simon, Benjamin Schoefer, and Jörg Heining (2021). "Labor in the boardroom". *The Quarterly Journal of Economics* 136.2, pp. 669–725. DOI: 10.1093/qje/qjaa038.

Jäger, Simon, Benjamin Schoefer, Samuel Young, et al. (2020). "Wages and the value of nonemployment". *The Quarterly Journal of Economics* 135.4, pp. 1905–1963. DOI: 10.1093/qje/qjaa016.

Jäger, Simon, Benjamin Schoefer, and Josef Zweimüller (2023). "Marginal jobs and job surplus: A test of the efficiency of separations". *The Review of Economic Studies* 90.3, pp. 1265–1303. DOI: 10.1093/restud/rdac045.

Kehrig, Matthias and Nicolas Vincent (2021). "The micro-level anatomy of the labor share decline". *The Quarterly Journal of Economics* 136.2, pp. 1031–1087. DOI: 10.1093/qje/qjab002.

Kreisman, Daniel and Jonathan Smith (2023). "Distinctively black names and educational outcomes". *Journal of Political Economy* 131.4, pp. 877–897. DOI: 10.1086/722093.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet (2021). "Gender differences in job search: Trading off commute against wage". *The Quarterly Journal of Economics* 136.1, pp. 381–426. DOI: 10.1093/qje/qjaa033.

Levy, Ro'ee (2021). "Social media, news consumption, and polarization: Evidence from a field experiment". *American Economic Review* 111.3, pp. 831–870. DOI: 10.1257/aer.20191777.

Lindqvist, Erik, Robert Östling, and David Cesarini (2020). "Long-run effects of lottery wealth on psychological well-being". *The Review of Economic Studies* 87.6, pp. 2703–2726. DOI: 10.1093/restud/rdaa006.

Martínez, Isabel Z., Emmanuel Saez, and Michael Siegenthaler (2021). "Intertemporal labor supply substitution? Evidence from the Swiss income tax holidays". *American Economic Review* 111.2, pp. 506–546. DOI: 10.1257/aer.20180746.

Mertens, Thomas M. and John C. Williams (2021). "What to expect from the lower bound on interest rates: Evidence from derivatives prices". *American Economic Review* 111.8, pp. 2473–2505. DOI: 10.1257/aer.20181461.

Miller, Sarah, Norman Johnson, and Laura R Wherry (2021). "Medicaid and mortality: New evidence from linked survey and administrative data". *The Quarterly Journal of Economics* 136.3, pp. 1783–1829. DOI: 10.1093/qje/qjab004.

Mueller-Smith, Michael and Kevin T. Schnepel (2020). "Diversion in the criminal justice system". *The Review of Economic Studies* 88.2, pp. 883–936. DOI: 10.1093/restud/rdaa030.

Mullainathan, Sendhil and Ziad Obermeyer (2022). "Diagnosing physician error: A machine learning approach to low-value health care". *The Quarterly Journal of Economics* 137.2, pp. 679–727. DOI: 10.1093/qje/qjab046.

Murphy, Richard and Felix Weinhardt (2020). "Top of the class: The importance of ordinal rank". *The Review of Economic Studies* 87.6, pp. 2777–2826. DOI: 10.1093/restud/rdaa020.

Norris, Samuel, Matthew Pecenco, and Jeffrey Weaver (2021). "The effects of parental and sibling incarceration: Evidence from Ohio". *American Economic Review* 111.9, pp. 2926–2963. DOI: 10.1257/aer.20190415.

Prager, Elena and Matt Schmitt (2021). "Employer consolidation and wages: Evidence from hospitals". *American Economic Review* 111.2, pp. 397–427. DOI: 10.1257/aer.20190690.

Sandvik, Jason J. et al. (2020). "Workplace knowledge flows". *The Quarterly Journal of Economics* 135.3, pp. 1635–1680. DOI: 10.1093/qje/qjaa013.

Shapiro, Bradley T., Günter J. Hitsch, and Anna E. Tuchman (2021). "TV advertising effectiveness and profitability: Generalizable results from 288 brands". *Econometrica* 89.4, pp. 1855–1879. DOI: 10.3982/ecta17674.

Wasserman, Melanie (2023). "Hours constraints, occupational choice, and gender: Evidence from medical residents". *The Review of Economic Studies* 90.3, pp. 1535–1568. DOI: 10.1093/restud/rdac042.

Weaver, Jeffrey (2021). "Jobs for sale: Corruption and misallocation in hiring". *American Economic Review* 111.10, pp. 3093–3122. DOI: 10.1257/aer.20201062.

D Effect Size Benchmarking

Table A1 shows the values of σ and r for a selected sample of ten highly-cited and recent results from the economics literature that represent plausibly large effects. I term this the 'benchmarking sample'. All articles in this sample have publicly available replication repositories and are published between 2015-2020. I isolate one main claim of each article and the primary estimate used to defend this claim. The benchmarking sample thus consists of ten articles, each with one claim and one estimate defending that claim. Online Appendix E provides citations for all articles in the benchmarking sample, along with associated replication repositories (when applicable).

Two features of Table A1 are worth noting. First, though σ and r are quite positively correlated and always share the same sign, they do not necessarily monotonically correspond. Second, though the estimates in this benchmarking sample are all statistically significant under the standard NHST framework, their effect sizes are also quite small in general.

Article	Setting	Outcome Variable	Exposure Variable	Initial p-Value	σ	r	Location
Acemoglu & Restrepo (2020)	Difference-in-differences analysis of U.S. commuting zones, 1990-2007	Employment rates (continuous)	Industrial robot exposure (continuous)	0.000	-0.206	-0.16	Table 7, Panel A, US exposure to robots, Model 3
Acemoglu et al. (2019)	Difference-in-differences analysis of countries, 1960-2010	Short-run log GDP levels (continuous)	Democratization (binary)	0.001	0.005	0.255	Table 2, Democracy, Model 3
Berman et al. (2017)	African 0.5 $\times$ 0.5 longitude-latitude cells with mineral mines, 1997-2010	Conflict incidence (binary)	Log price of main mineral (continuous)	0.012	0.521	0.007	Table 2, ln price x mines > 0, Model 1
Deschênes, Greenstone, & Shapiro (2017)	Difference-in-differences analysis of U.S. counties, 2001-2007	Nitrogen dioxide emissions (continuous)	Nitrogen dioxide cap-and-trade participation (binary)	0.000	-0.134	-0.468	Table 2, Panel A, NOx, Model 3
Haushofer & Shapiro (2016)	Experiment with low-income Kenyan households, 2011-2013	Non-durable consumption (continuous)	Unconditional cash transfer (binary)	0.000	0.376	0.195	Table V, Non-durable expenditure, Model 1
Benhassine et al. (2015)	Experiment with families of Moroccan primary school-aged students, 2008-2010	School attendance (binary)	Educational cash transfer to fathers (binary)	0.000	0.18	0.252	Table 5, Panel A, Attending school by end of year 2, among those 6-15 at baseline, Impact of LCT to fathers
Bloom et al. (2015)	Field experiment with Chinese workers, 2010-2011	Attrition (binary)	Voluntarily working from home (binary)	0.002	-0.397	-0.196	Table VIII, Treatment, Model 1
Duflo, Dupas, & Kremer (2015)	Experiment with Kenyan primary school-aged girls, 2003-2010	Reaching eighth grade (binary)	Education subsidy (binary)	0.023	0.1	0.125	Table 3, Panel A, Stand-alone education subsidy, Model 1
Hanushek et al. (2015)	OECD adult workers, 2011-2012	Log hourly wages (continuous)	Numeracy skills (continuous)	0.000	0.091	0.316	Table 5, Numeracy, Model 1
Oswald, Proto, & Sgroi (2015)	UK students, piece-rate laboratory task	Productivity (continuous)	Happiness (continuous)	0.018	0.753	0.244	Table 2, Change in happiness, Model 4

Note: Effect sizes and initial standard NHST p-values of each estimate are reported. Each original estimate can be found in its respective article at the specified location. Some articles are reproduced using data from repositories (Hanushek 2016; Benhassine et al. 2019; Berman et al. 2019; Deschênes, Greenstone, & Shapiro 2019; Duflo, Dupas, & Kremer 2019), whereas others are reproduced using files linked to the publisher's online webpage for the article.

Table A1: Effect Size Benchmarking

E Benchmarking Sample

All articles and associated replication repositories (when applicable) of the benchmarking sample are provided here.

References

Acemoglu, Daron, Suresh Naidu, et al. (2019). "Democracy does cause growth". *Journal of Political Economy* 127.1, pp. 47–100. DOI: 10.1086/700936.

Acemoglu, Daron and Pascual Restrepo (2020). "Robots and jobs: Evidence from US labor markets". *Journal of Political Economy* 128.6, pp. 2188–2244. DOI: 10.1086/705716.

Benhassine, Najy et al. (2015). "Turning a shove into a nudge? A "labeled cash transfer" for education". *American Economic Journal: Economic Policy* 7.3, pp. 86–125. DOI: 10.1257/pol.20130225.

— (2019). *Replication data for: Turning a shove into a nudge? A "labeled cash transfer" for education*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E114579V1.

Berman, Nicolas et al. (2017). "This mine is mine! How minerals fuel conflicts in Africa". *American Economic Review* 107.6, pp. 1564–1610. DOI: 10.1257/aer.20150774.

— (2019). *Replication data for: This mine is mine! How minerals fuel conflicts in Africa*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E113068V1.

Bloom, Nicholas et al. (2015). "Does working from home work? Evidence from a Chinese experiment". *The Quarterly Journal of Economics* 130.1, pp. 165–218. DOI: 10.1093/qje/qju032.

Deschênes, Olivier, Michael Greenstone, and Joseph S. Shapiro (2017). "Defensive investments and the demand for air quality: Evidence from the NOx Budget Program". *American Economic Review* 107.10, pp. 2958–2989. DOI: 10.1257/aer.20131002.

Deschênes, Olivier, Michael Greenstone, and Joseph S. Shapiro (2019). *Replication data for: Defensive investments and the demand for air quality: Evidence from the NOx Budget Program*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E112938V1.

Duflo, Esther, Pascaline Dupas, and Michael Kremer (2015). "Education, HIV, and early fertility: Experimental evidence from Kenya". *American Economic Review* 105.9, pp. 2757–2797. DOI: 10.1257/aer.20121607.

— (2019). *Replication data for: Education, HIV, and early fertility: Experimental evidence from Kenya*. Dataset V1. Ann Arbor, MI, U.S.A.: Inter-university Consortium for Political and Social Research. DOI: 10.3886/E112899V1.

Hanushek, Eric A. (2016). *Data for: Returns to skills around the world: Evidence from PIAAC*. Dataset V1. Amsterdam, The Netherlands: Mendeley Data. DOI: 10.17632/nmsxzyjfkk.1.

Hanushek, Eric A. et al. (2015). "Returns to skills around the world: Evidence from PIAAC". *European Economic Review* 73, pp. 103–130. DOI: 10.1016/j.euroecorev.2014.10.006.

Haushofer, Johannes and Jeremy Shapiro (2016). "The short-term impact of unconditional cash transfers to the Poor: Experimental evidence from Kenya". *The Quarterly Journal of Economics* 131.4, pp. 1973–2042. DOI: 10.1093/qje/qjw025.

Oswald, Andrew J., Eugenio Proto, and Daniel Sgroi (2015). "Happiness and productivity". *Journal of Labor Economics* 33.4, pp. 789–822. DOI: 10.1086/681096.

F SSPP Data

The SSPP survey was posted publicly to the SSPP website, and any interested respondent was free to take the survey. The survey was also publicly disseminated on Twitter/X by the SSPP. 58 of the 62 survey respondents (93.5%) are members of the SSPP's Superforecaster Panel, which is a sample of researchers that are pre-selected by SSPP and are paid a semi-annual flat rate for completing a sufficient proportion of the surveys that are posted to the SSPP website each month. The remaining four respondents are not part of the Superforecaster Panel, and are not incentivized to take the survey.

My SSPP sample is relatively young, with the median respondent being 32.5 years of age (mean = 34.6, SD = 10.8). Though much of the sample has ample experience with making predictions for social science research questions by virtue of being part of the Superforecaster Panel, my sample rates their five-point Likert confidence in their predictions at a median of 2.5 (mean = 2.4, SD = 1). This is sensible, as only nine respondents (14.5%) report conducting prior research on the topics discussed in my survey. The sample is male-dominated, with 53 respondents (85.5%) reporting a masculine gender identity. The SSPP sample also predominantly originates from WEIRD countries (Henrich, Heine, & Norenzayan 2010) – 42 respondents (67.7%) spent the majority of their time prior to starting university education in OECD member states, and 48 respondents (77.4%) have spent the majority of their time since starting university education in OECD member states.

G Equivalence Testing Failure Rate Computation

Let j be an individual partition, and let i index an individual estimate. j represents an individual claim when calculating claim-level ETFRs, whereas j represents an entire article when calculating article-level ETFRs. Each estimate i belongs to exactly one partition j. Because all ETFRs in this paper are calculated for symmetric ROPEs, it is sufficient to define ETFR $R(\epsilon, \tau, L)$ as a function of ROPE length $\epsilon > 0$, effect size measure $\tau \in \{\sigma, r\}$, and aggregation level L. Further, because the ECI approach described in Definition 4.3 yields identical results to the TOST procedure described in Definition 4.2, I approach ETFR calculation by defining the exact 95% ECI's outer bound $\text{ECIOB}_{i,j}(\tau)$ for each effect size measure τ of every estimate i. Let M_j represent the number of estimates i belonging to partition j, and let M be the total number of partitions j. One can then calculate the ETFR as

$$R(\epsilon, \tau, L) = \sum_{j=1}^{M} \sum_{i=1}^{M_j} \frac{\mathbb{1}\left[|\text{ECIOB}_{i,j}(\tau)| > \epsilon\right]}{M_j M}. \tag{A1}$$

I also calculate claim-level ETFRs that apply an inverse weighting approach, ensuring that each article receives the same weight in the sample. Let $W_{j,k}$ be equal to 1 divided by the number of claims that belong to claim j's article, and let k be an individual article. Then the inverse-weighted claim-level ETFR can be written as

$$R_{\text{Wgt.}}(\epsilon, \tau) = \frac{1}{\sum_{j=1}^{M} W_{j,k}} \sum_{j=1}^{M} W_{j,k} \sum_{i=1}^{M_{j,k}} \frac{\mathbb{1}\left[|\text{ECIOB}_{i,j,k}(\tau)| > \epsilon\right]}{M_{j,k}}, \tag{A2}$$

where $M_{j,k}$ is now the number of estimates belonging to claim j in article k, and M is now the total number of articles.

I measure precision using standard errors of the mean for the unweighted ETFRs in Equation A1 and standard errors of the weighted mean for the weighted ETFRs in Equation A2. The standard error of the mean for an ETFR is

$$\mathrm{SE}\left[R(\epsilon, \tau, L)\right] = \frac{\mathrm{SD}\left[R(\epsilon, \tau, L)\right]}{\sqrt{M}}, \tag{A3}$$

where $\mathrm{SD}\left[R(\epsilon, \tau, L)\right]$ is just the within-sample standard deviation of $R(\epsilon, \tau, L)$. Let the ETFR for claim j in article k be defined as

$$R_{j,k}(\epsilon, \tau, L) = \sum_{i=1}^{M_{j,k}} \frac{\mathbb{1}\left[|\mathrm{ECIOB}_{i,j,k}(\tau)| > \epsilon\right]}{M_{j,k}}.$$

Though Gatz & Smith (1995) note that there is no universally-agreed definition for the standard error of the weighted mean, they find that one formulation produces closer estimates to the bootstrap than other competing formulas. In this setting, the square of that optimal formula can be written as

$$\begin{aligned}\left(\mathrm{SE}\left[R_{\mathrm{Wgt.}}(\cdot)\right]\right)^2 = \frac{M}{(1-M)\,M^2}\Bigg[&\sum_{j=1}^{M}\left\{\left[W_{j,k}R_{j,k}(\cdot) - \overline{W}_{j,k}R_{\mathrm{Wgt.}}(\cdot)\right]^2\right\} - \\ &2R_{\mathrm{Wgt.}}(\cdot)\sum_{j=1}^{M}\left\{\left(W_{j,k} - \overline{W}_{j,k}\right)\left[W_{j,k}R_{j,k}(\cdot) - \overline{W}_{j,k}R_{\mathrm{Wgt.}}(\cdot)\right]\right\} + \\ &\left[R_{\mathrm{Wgt.}}(\cdot)\right]^2\sum_{j=1}^{M}\left\{\left[W_{j,k} - \overline{W}_{j,k}\right]^2\right\}\Bigg],\end{aligned}$$

where $\overline{W}_{j,k}$ is the mean inverse weight $W_{j,k}$ across all claims and M is the total number of articles. The results in Section 6.2 show that this standard error derivation corresponds quite closely with simple standard errors for unweighted ETFRs as derived in Equation A3.

H Online Appendix Tables and Figures

This appendix provides table versions of two main figures in Section 6.

	(1)	(2)	(3)	(4)	(5)	(6)
γ_r	-0.046	$\cdot$	-0.02	0.002	0.214	0.228
	(0.016)	$(\cdot)$	(0.017)	(0.02)	(0.023)	(0.028)
Type	Judgment	Judgment	Judgment	Judgment	Prediction	Prediction
Rate	Type I Error	Type II Error	TOST/ECI Failure	TOST/ECI Failure	TOST/ECI Failure	TOST/ECI Failure
Effect Size Measure			σ	r	σ	r

Note: This table provides the numerical estimates displayed in Figure 4.

Table A2: Within-Researcher Estimates of Differences in Predictions/Judgments

	(1)	(2)	(3)	(4)	(5)	(6)
Equivalence Testing Failure Rate	0.361	0.385	0.379	0.633	0.609	0.617
	(0.035)	(0.041)	(0.044)	(0.038)	(0.044)	(0.048)
Effect Size Measure	σ	σ	σ	r	r	r
SSPP Tolerance	0.1065	0.1065	0.1065	0.1295	0.1295	0.1295
Aggregation Level	Claim	Claim	Article	Claim	Claim	Article
Inverse Weighting		x			x	

Note: This table provides the numerical estimates displayed in Figure 6. ROPEs are $[-0.2\sigma, 0.2\sigma]$ and $[-0.1r, 0.1r]$. SSPP tolerance indicates the median SSPP respondent's tolerance for ETFRs for the given ROPE (see Section 5.3).

Table A3: Main Equivalence Testing Failure Rate Estimates

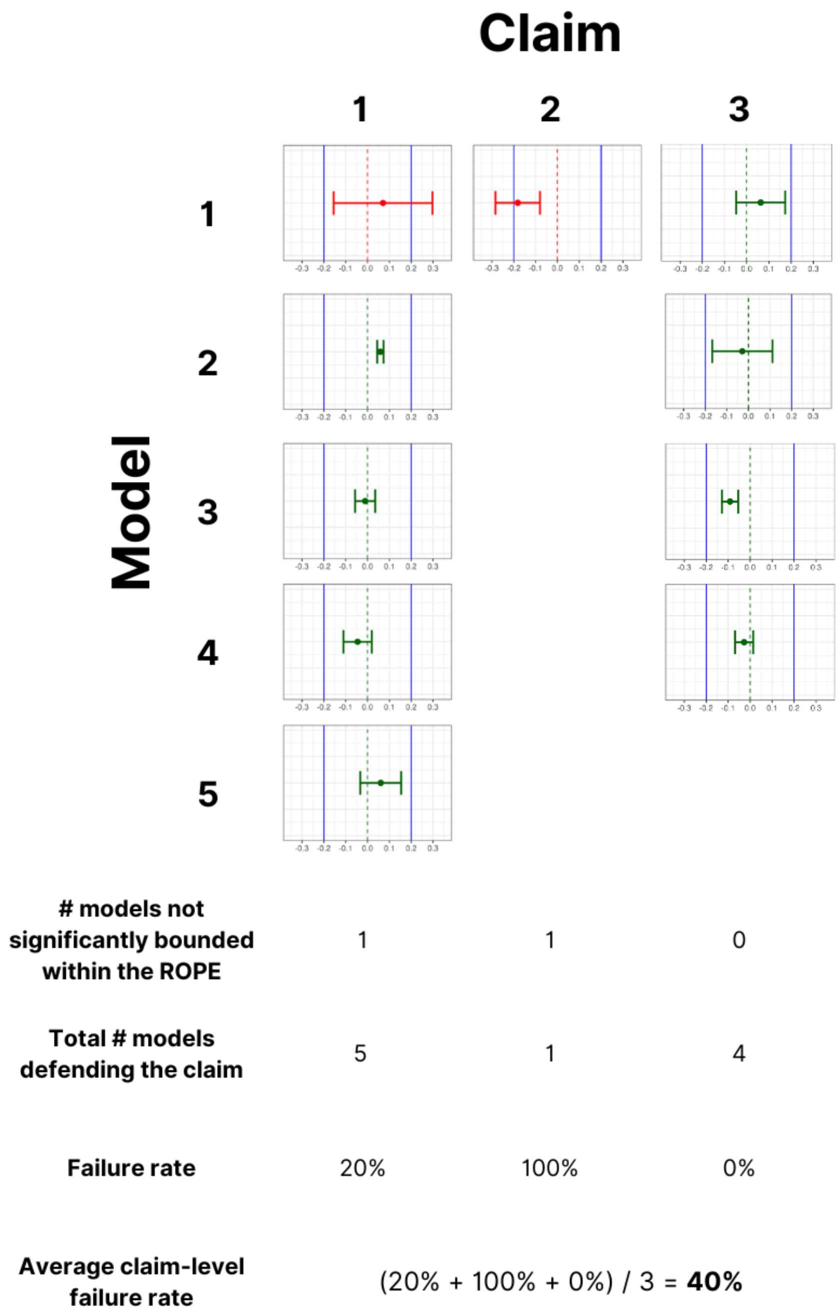

Figure A1: Visualization of Equivalence Testing Failure Rate Computation

I Robustness Checks

This appendix reports extended robustness checks on the main results in Section 6.2.

	Estimates	Claims	Articles	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Initially Insignificant Estimates	790	134	80	0.34 (0.035)	0.36 (0.041)	0.353 (0.044)	0.618 (0.039)	0.587 (0.045)	0.594 (0.05)
Panel B: Initially Significant Estimates	86	32	26	0.576 (0.087)	0.625 (0.088)	0.622 (0.091)	0.719 (0.081)	0.756 (0.078)	0.756 (0.084)
Effect Size Measure				σ	σ	σ	r	r	r
SSPP Tolerance				0.1065	0.1065	0.1065	0.1295	0.1295	0.1295
Aggregation Level				Claim	Claim	Article	Claim	Claim	Article
Inverse Weighting					x			x	

Note: Estimates with initial standard NHST p-values ≥ 0.05 (before conformability changes, if applicable) are removed from the sample in Panel A, and constitute the entire sample in Panel B. ROPEs are $[-0.2\sigma, 0.2\sigma]$ and $[-0.1r, 0.1r]$. SSPP tolerance indicates the median SSPP respondent's tolerance for ETFRs for the given ROPE (see Section 5.3).

Table A4: ETFR Robustness – Initial Estimate Significance

	Estimates	Claims	Articles	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: CYCD Removed	675	105	63	0.342 (0.04)	0.362 (0.046)	0.356 (0.049)	0.62 (0.044)	0.617 (0.049)	0.628 (0.054)
Panel B: CYBD Removed	563	91	59	0.36 (0.045)	0.37 (0.049)	0.369 (0.054)	0.621 (0.047)	0.558 (0.053)	0.562 (0.058)
Panel C: BYCD Removed	563	124	74	0.398 (0.038)	0.417 (0.043)	0.409 (0.047)	0.651 (0.04)	0.631 (0.046)	0.64 (0.051)
Panel D: BYBD Removed	653	119	73	0.365 (0.038)	0.39 (0.043)	0.386 (0.046)	0.634 (0.04)	0.625 (0.046)	0.629 (0.052)
Effect Size Measure				σ	σ	σ	r	r	r
SSPP Tolerance				0.1065	0.1065	0.1065	0.1295	0.1295	0.1295
Aggregation Level				Claim	Claim	Article	Claim	Claim	Article
Inverse Weighting					x			x	

Note: Panels denote whether estimates corresponding to continuous/binary outcome/exposure variables (respectively) are removed from the sample. For example, 'CYBD removed' implies that estimates corresponding to a continuous outcome variable and a binary exposure variable are removed from the sample. ROPEs are $[-0.2\sigma, 0.2\sigma]$ and $[-0.1r, 0.1r]$. SSPP tolerance indicates the median SSPP respondent's tolerance for ETFRs for the given ROPE (see Section 5.3).

Table A5: ETFR Robustness – Regressor Type Combination

	Estimates	Claims	Articles	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Replicable Estimates Removed	803	123	74	0.388	0.406	0.399	0.618	0.607	0.615
				(0.038)	(0.043)	(0.047)	(0.04)	(0.046)	(0.051)
Panel B: Non-Conformable Estimates Removed	807	130	77	0.358	0.37	0.365	0.65	0.626	0.636
				(0.036)	(0.041)	(0.044)	(0.038)	(0.044)	(0.049)
Effect Size Measure				σ	σ	σ	r	r	r
SSPP Tolerance				0.1065	0.1065	0.1065	0.1295	0.1295	0.1295
Aggregation Level				Claim	Claim	Article	Claim	Claim	Article
Inverse Weighting					x			x	

Note: Estimates are non-replicable if my best attempts to replicate the exact published estimates using the article's replication repository do not succeed. Estimates are 'non-conformable' if the models that produce them require conformability modifications before inclusion in the final sample. ROPEs are $[-0.2\sigma, 0.2\sigma]$ and $[-0.1r, 0.1r]$. SSPP tolerance indicates the median SSPP respondent's tolerance for ETFRs for the given ROPE (see Section 5.3).

Table A6: ETFR Robustness – Replicability/Conformability

References

Gatz, Donald F. and Luther Smith (1995). "The standard error of a weighted mean concentration—I. Bootstrapping vs other methods". *Atmospheric Environment* 29.11, pp. 1185–1193. DOI: 10.1016/1352-2310(94)00210-c.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan (2010). "The weirdest people in the world?" *Behavioral and Brain Sciences* 33.2–3, pp. 61–83. DOI: 10.1017/s0140525x0999152x.

van de Schoot, Rens et al. (2021). "An open source machine learning framework for efficient and transparent systematic reviews". *Nature Machine Intelligence* 3.2, pp. 125–133. DOI: 10.1038/s42256-020-00287-7.