

Are inspections going to “waste”?

A national field test of machine learning vs. expert judgement to improve EPA regulatory compliance

Jesse Buchsbaum, Michael Greenstone, and Katherine Meckel^{*}

Abstract

With the U.S. Environmental Protection Agency, we developed a machine learning model to predict sites where inspections would uncover severe violations of hazardous waste regulations. We estimate that using our model to target inspections will increase the “hit rate” by 46%. As is often the case, the model’s data are highly selected (representing about ~2% of sites), suggesting that classic selection bias concerns make our estimate’s relevance to the full population unknown. We therefore conducted a national field test of the model’s versus the EPA’s inspection targets; the model’s relative performance was even better, increasing the hit rate by 79%.

FULL DRAFT COMING SOON

^{*} Contact: Buchsbaum: Energy and Environment Lab, University of Chicago. Greenstone: Energy Policy Institute and Department of Economics, University of Chicago and NBER. Meckel: Department of Economics, UCSD, NBER and CESifo. We are indebted to our partners at the EPA and each of the regional offices.

Thank you to Leslie Hodel for invaluable guidance, support, and management of this project. We are grateful to Christina Seybolt, Bridget Pals, Michael Galparin, Sarah Armstrong, and Aditya Jain for excellent research assistance. We thank seminar participants at TWEEDs, UC Berkeley, and UCLA for conversations and suggestions that have greatly improved the quality of this project. In addition, we are grateful to the Smith Richardson Foundation, Arnold Ventures, and the Abdul Latif Jameel Poverty Action for their generous grant support of this project, as well as philanthropic funding from the Dunn family. All views and errors are solely ours.