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