

La Nauze, Andrea; Tan, Tze Yong

Working Paper

A Comment on "Informed Enforcement: Lessons from Pollution Monitoring in China" by Sebastian Axbard and Zichen Deng

I4R Discussion Paper Series, No. 144

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: La Nauze, Andrea; Tan, Tze Yong (2024) : A Comment on "Informed Enforcement: Lessons from Pollution Monitoring in China" by Sebastian Axbard and Zichen Deng, I4R Discussion Paper Series, No. 144, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/302050>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



No. 144

I4R DISCUSSION PAPER SERIES

A Comment on “Informed Enforcement: Lessons from Pollution Monitoring in China” by Sebastian Axbard and Zichen Deng

Andrea La Nauze

Tze Yong Tan

August 2024

I4R DISCUSSION PAPER SERIES

I4R DP No. 144

A Comment on “Informed Enforcement: Lessons from Pollution Monitoring in China” by Sebastian Axbard and Zichen Deng

Andrea La Nauze^{1,2,3}, Tze Yong Tan⁴

¹Deakin University, Geelong/Australia

²University of Queensland, Brisbane/Australia

³CESifo, Munich/Germany

⁴Monash University, Melbourne/Australia

AUGUST 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

A comment on “Informed Enforcement: Lessons from Pollution Monitoring in China” by Sebastian Axbard and Zichen Deng

Andrea La Nauze*

Tze Yong Tan†

June 7, 2024

Abstract

Axbard and Deng (2024) exploit the rollout of new pollution monitors in China in 2015 in 177 medium-size cities to study the effect of air-quality monitors on enforcement actions by local governments and air quality. In their main difference-in-difference analysis, they identify the change in the probability of enforcement for firms that are close to versus further away from the monitor. They find that being within 10km of a monitor increases the probability that a firm receives any enforcement action by 0.0033 (standard error 0.00056) relative to a mean of 0.0046. Computationally, we successfully reproduce the main claims of the paper. We observe minor coding anomalies that do not have a material impact. We find that the main result on all enforcement is robust to all robustness checks: (1) randomization inference (2) alternative fixed effects and (3) multiple hypothesis testing.

KEYWORDS: Replication, Reproducibility, Robustness, Accountability, Regulatory Enforcement, Pollution, China

JEL CODES: K32, L51, O13, P25, P28, Q52, Q53

*Deakin University, University of Queensland, and CESifo. Email: a.lanauze@deakin.edu.au

†Corresponding Author. Department of Economics and SoDa Laboratories, Monash University, Australia. Email: TzeYong.Tan@monash.edu

1 Introduction

[Axbard and Deng \(2024\)](#), henceforth AD, investigate the introduction of air-quality monitors in medium-sized Chinese cities in 2015. Before the introduction of these monitors, no air-quality monitoring was undertaken in these cities. In their main analysis, the authors compare the incidence of enforcement actions by local authorities between firms close to versus further away from these monitors before and after the monitors were installed.

The main datasets for the analysis are records on enforcement collected by the Institute of Public and Environmental Affairs 2010-2017. AD link enforcement records to the Annual Survey of Industrial Firms to firm and firm location. We assume the location of monitors comes from the Ministry of Environmental Protection where air pollution data are sourced.

The authors describe the main results as follows “The results show an increase in the probability of enforcement by 72 percent for firms located within 10 kilometers (km) of a monitor...”.

In the present report prepared for the Institute for Replication ([Brodeur et al. 2024](#)), we investigate whether AD’s analytical results are computationally reproducible and further test robustness to: (1) randomization inference (2) alternative fixed effects and (3) multiple hypothesis testing.

We successfully reproduced AD’s main results (Panel A, Table 1) using their codes however we noted very small discrepancies between the sample size that is reported and the sample size used to generate the estimates.

In terms of robustness we find no meaningful differences in the main result (column 1 Table 1 Panel A) from employing randomization inference, alternative fixed effects, or multiple hypothesis testing. We note that heterogeneity results reported in columns 2-4 of Table 1 Panel A are sensitive to the fixed effects employed but are robust to randomization inference and controlling for multiple hypothesis testing.

2 Computational Reproducibility

We used the replication package [here](#). The replication package did not include the raw data or the scraping and cleaning codes.¹ Intermediate data files and analysis scripts were provided. See Table 1 for details. We reproduced the main results in Table 1 Panel A from the data and analysis scripts provided (a successful computational reproduction) see Table 2 for the reproduced results.

We noted one minor coding issue that affected the reporting of observations in the main table. We noticed that the observations reported in the Table are from the full dataset, which differs to the sample used to estimate the coefficients due to singletons being dropped in the estimation procedure. The discrepancy is small (608 observations). We also note that the reported summary statistics in Appendix Table C1 are calculated from the incorrect sample (or at least not the main sample used for estimation), but the statistics do not differ at this level of rounding.

3 Robustness Reproducibility

In each of the following sections, we explore the robustness of the results in Panel A of Table 1 of AD.

3.1 Randomization Inference

We first report robustness to implementing randomization inference by using the command developed by [Heß \(2017\)](#) in Stata. The purpose of this exercise is to ensure that the results remain robust to the arbitrary randomization procedures of the treatment variable. This involves assigning firms randomly to be either close to or further away from the monitor while keeping the outcome variables fixed. The test generates p-values by calculating the proportion of replications where the placebo treatment effect exceeds those of the baseline treatment effect. This exercise is conducted with 1,000 random permutations and at seed 5.

¹The raw data here refers to the files obtained directly from the source without converting them to the Stata dataset format.

The randomization inference-adjusted p-values are reported in Table 3. We did not find any loss of significance in the first four specifications. In the last specification, we find that the likelihood of achieving a treatment effect at least as high as the estimated one by chance alone is 0.461. This is consistent with the baseline estimation, further confirming that being close to a monitor does not impact the probability a firm receives a warning enforcement.

3.2 Fixed Effects

We explore robustness to several alternative specifications of fixed effects. The fixed effects employed by AD are firm, industry by time (quarter), and province by time (quarter). We test robustness to specifications with time (quarter) and firm only (Table 4), firm and year-by-industry (Table 5), and firm and year-by-province fixed effects (Table 6). The treatment effect of being close to a monitor on any enforcement action is robust across all specifications. The treatment effects in columns 2-4 (outcome variables “Suspension”, “Upgrade”, and “Fine”) are less robust, in particular, they are all insignificant in Table 4, a specification that does not allow for differential time trends by industry or province.

3.3 Multiple Hypothesis Testing

Our final set of robustness checks is to account for multiple hypothesis testing. We take the hypotheses in Panel A of Table 1 of AD as one family of 5 hypothesis tests, and compute sharpened false discovery rate q-values according to (Anderson 2008). The original p-values and the sharpened q-values are reported in Table 7. Accounting for multiple hypotheses does not change the conclusions.

4 Conclusion

The main result of AD is robust - being closer to a newly installed air-quality monitor led to a large and statistically significant increase in the likelihood of enforcement action relative to firms located further from a monitor. We find that the

heterogeneity results are less robust to alternative fixed effects specifications. We also found minor coding issues that do not materially affect the results.

References

- Anderson, M. L.: 2008, Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects, *Journal of the American Statistical Association* **103**(484), 1481–1495.
- Axbard, S. and Deng, Z.: 2024, Informed enforcement: Lessons from pollution monitoring in china, *American Economic Journal: Applied Economics* **16**(1), 213–252.
- Brodeur *et al.*: 2024, Mass Reproducibility and Replicability: A New Hope, *I4R Discussion Paper Series* .
- Heß, S.: 2017, Randomization inference with stata: A guide and software, *The Stata Journal* **17**(3), 630–651.

5 Tables

Table 1: Replication Package Contents and Reproducibility

Replication Package Item	Fully	Partial	No
Raw data provided			✓
Analysis data provided	✓		
Cleaning code provided			✓
Analysis code provided	✓		
Reproducible from raw data			✓
Reproducible from analysis data	✓		

Notes: This table summarizes the replication package contents contained in [Axbard and Deng \(2024\)](#)

Table 2: Pollution Monitoring and Enforcement Activities

	Any	Suspension	Upgrade	Fine	Warning
$Mon_{10km} \times Post$	0.0033 (0.00056) [0.000]	0.0014 (0.00045) [0.002]	0.0014 (0.00041) [0.001]	0.0014 (0.00043) [0.002]	-0.000058 (0.00016) [0.711]
Mean Outcome	0.0046	0.0024	0.0025	0.0022	0.00070
Observations	1154688	1154688	1154688	1154688	1154688
Conley SE	0.00040	0.00031	0.00031	0.00030	0.00017

Notes: This table is a reproduction of AD Table 1 Panel A using code in their replication package. Standard errors are clustered on the city (in parentheses). P-values are reported in brackets.

Table 3: Randomization Inference Robustness

	Any	Suspension	Upgrade	Fine	Warning
$Mon_{10km} \times Post$	0.0033 (0.00056)	0.0014 (0.00045)	0.0014 (0.00041)	0.0014 (0.00043)	-0.000058 (0.00016)
Unadjusted p-value	0.000	0.002	0.001	0.002	0.711
RI-adjusted p-value	0.000	0.000	0.000	0.000	0.461
Mean Outcome	0.0046	0.0024	0.0025	0.0022	0.0007
Observations	1154688	1154688	1154688	1154688	1154688

Notes: Table reports unadjusted p-values and randomization inference (RI) adjusted p-values for the coefficients $Mon_{10km} \times Post$ (robustness check of AD Table 1 Panel A).

Table 4: Fixed Effects Robustness I

	Any	Suspension	Upgrade	Fine	Warning
$Mon_{10km} \times Post$	0.0025 (0.00068) [0.000]	0.00074 (0.00058) [0.204]	0.00074 (0.00053) [0.168]	0.00059 (0.00055) [0.287]	-0.00068 (0.00016) [0.679]
Mean Outcome	0.0046	0.0024	0.0025	0.0022	0.00070
Observations	1154688	1154688	1154688	1154688	1154688
Firm FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	No	No	No	No	No
Province-Year FE	No	No	No	No	No

Notes: Table reports coefficients and standard errors for $Mon_{10km} \times Post$ with firm and time fixed effects (robustness check of AD Table 1 Panel A). Standard errors are clustered on the city (in parentheses). P-values are reported in brackets.

Table 5: Fixed Effects Robustness II

	Any	Suspension	Upgrade	Fine	Warning
$Mon_{10km} \times Post$	0.0031 (0.00060) [0.000]	0.0013 (0.00052) [0.015]	0.0013 (0.00046) [0.005]	0.0012 (0.00049) [0.017]	-0.000043 (0.00016) [0.785]
Mean Outcome	0.0046	0.0024	0.0025	0.0022	0.00070
Observations	1154688	1154688	1154688	1154688	1154688
Firm FE	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	No	No
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
Province-Year FE	No	No	No	No	No

Notes: Table reports coefficients and standard errors for $Mon_{10km} \times Post$ with firm and industry-by-year fixed effects (robustness check of AD Table 1 Panel A). Standard errors are clustered on the city (in parentheses). P-values are reported in brackets.

Table 6: Fixed Effects Robustness III

	Any	Suspension	Upgrade	Fine	Warning
$Mon_{10km} \times Post$	0.0027 (0.00063) [0.000]	0.00088 (0.00050) [0.080]	0.0014 (0.00040) [0.001]	0.00074 (0.00047) [0.116]	-0.000066 (0.00016) [0.689]
Mean Outcome	0.0046	0.0024	0.0025	0.0022	0.00070
Observations	1154688	1154688	1154688	1154688	1154688
Firm FE	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	No	No
Industry-Year FE	No	No	No	No	No
Province-Year FE	Yes	Yes	Yes	Yes	Yes

Notes: Table reports coefficients and standard errors for $Mon_{10km} \times Post$ with firm and province-by-year fixed effects (robustness check of AD Table 1 Panel A). Standard errors are clustered on the city (in parentheses). P-values are reported in brackets.

Table 7: Multiple Hypothesis Testing Robustness

	Any	Suspension	Upgrade	Fine	Warning
p-value	0.000	0.002	0.001	0.002	0.711
sharpened q-value	0.001	0.002	0.002	0.002	0.166

Notes: Table reports original p-values and Sharpened q-values that account for multiple hypotheses from the null hypothesis that the coefficient on $Mon_{10km} \times Post$ is zero in AD Table 1 Panel A.