Simen Owen Academic Proceedings Series

4th International Conference on Media, Economy, Communication and Intelligence Management (MECI 2026)

Article

Green Finance Policy and Corporate Investment Efficiency: Evidence from a Quasi-Natural Experiment Using Difference-in-Differences and Panel Data

Boshen Jin[1,*]

[1] University of California, Irvine, Irvine, USA
* Correspondence: Boshen Jin, University of California, Irvine, Irvine, USA

Abstract: Green finance policies can change firms' financing conditions and investment choices, but their effects are difficult to separate from regional shocks and measurement choices. This study reanalyses a publicly released panel of Chinese listed companies around the 2017 green finance reform and innovation pilot program. The analysis audits exposure coding, fixes firms' provincial membership, excludes provinces entering later pilot rounds, and estimates difference-in-differences models with firm and year fixed effects. The resulting sample contains 18,753 firm-year observations from 2,100 companies in 29 provinces, including five exposed provinces. The main coefficient on the published absolute investment residual is −0.00419. Its province-clustered 95% confidence interval is [−0.01022, 0.00185], and the null-imposed wild-cluster-bootstrap p-value is 0.403. Clustering the same specification by firm instead produces a p-value of 0.031, illustrating the importance of aligning inference with policy variation. Event-study estimates, alternative samples, and leave-one-exposed-province-out checks identify additional uncertainty, including substantial dependence on Guangdong. The evidence is consistent with a reduction in the investment-inefficiency proxy, but does not establish a precisely estimated average provincial effect. The study provides a reproducible framework for evaluating policy exposure, generated outcomes, and small-treated-cluster inference in green finance research.

Keywords: Green finance; Investment efficiency; Difference-in-differences; Panel data; Clustered inference

Received: 30 August 2026
Revised: 20 October 2026
Accepted: 02 November 2026
Published: 05 November 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Green finance connects environmental objectives with the allocation of corporate capital. Lending standards, bond markets, disclosure requirements, and local financial infrastructure can alter which projects receive funding. These changes may improve investment decisions when firms face financing constraints or weak monitoring. They can also impose adjustment costs, redirect investment toward administratively favored projects, or change the timing of expenditure. A policy assessment therefore needs to examine the allocation of investment alongside its volume. Evidence from green bond issuance documents links between sustainable financing and corporate behavior, while leaving open how geographically targeted policies affect the wider population of listed firms [1].

Corporate investment efficiency is difficult to observe directly. The value of a project depends on uncertain future cash flows, environmental effects, and financing conditions. Accounting studies commonly measure deviations from an estimated investment benchmark. Richardson's framework separates expected investment from unexplained expenditure and provides a practical way to study capital allocation [2]. A smaller

residual can indicate improved alignment with measured opportunities, but it can also reflect changes in omitted opportunities or accounting practices. Research on financial reporting quality likewise links information environments with departures from expected investment [3]. These measures are useful empirical proxies whose interpretation depends on the underlying model.

China's 2017 green finance reform and innovation program provides a policy setting with several pre-policy and post-policy years. Selected areas in Guangdong, Zhejiang, Jiangxi, Guizhou, and Xinjiang received a package of financial and institutional initiatives. Published research examines investment outcomes using listed-company panels, including an openly released replication archive [4]. That archive enables a new assessment of the same underlying observations. It also exposes choices that are difficult to evaluate from regression tables alone, including whether treatment corresponds to a pilot locality or its entire province, whether companies change location, and whether later pilot entrants remain in the comparison group.

The statistical challenge extends beyond assigning a policy date. A dataset can contain thousands of companies while its policy variation comes from only a few regional units. Firms in the same province may share credit conditions, investment cycles, and regulatory implementation. Treating their errors as independent across firms can produce confidence intervals that understate this common uncertainty. Bertrand, Duflo, and Mullainathan show how serial dependence complicates difference-in-differences inference [5]. Regional dependence adds a related concern when policy exposure is coded geographically. The effective number of policy comparisons is not the number of accounting observations.

This paper asks how the evidence on investment efficiency changes when the released panel is subjected to an explicit exposure and inference audit. The analysis preserves the published investment-residual outcome, identifies companies with changing provincial membership, and constructs a stable-location sample. It excludes two provinces associated with later pilot rounds and estimates firm- and year-fixed-effect models. Province-level clustering, a restricted wild bootstrap, an event study, alternative outcome and sample definitions, and province-deletion diagnostics are used to describe uncertainty. The treatment variable remains a province-level exposure proxy; exact participation in a local pilot area is not observed for every firm.

The empirical contribution is a reproducible reanalysis with a clearly bounded estimand. It distinguishes a negative coefficient on an inefficiency proxy from a demonstrated causal improvement, separates reproducibility of an earlier specification from robustness of its interpretation, and examines the dependence of inference on policy-level clustering. The main estimate is negative, but its regional confidence interval includes zero and the wild-bootstrap p-value is substantially larger than the corresponding firm-clustered value. This contrast motivates closer attention to treatment geography and shared regional shocks. The resulting workflow can support subsequent studies with more precise project, location, and credit-exposure data. The resulting findings therefore emphasize transparent measurement and uncertainty in geographically targeted finance.

2. Related Work

2.1. Green Finance and Corporate Capital Allocation

Green finance research studies several instruments that differ in eligibility, incentives, and monitoring. Geographic reform programs combine credit guidance, information sharing, financial product development, and public support. Zhang examines financing responses to China's green finance innovation and reform pilot zones, establishing the relevance of corporate financing as an intermediate policy outcome [6]. Liu and Wang investigate green innovation around the pilot program [7]. These outcomes can help explain investment responses, but financing and patent activity are not interchangeable with efficient capital allocation. A firm can obtain additional funds without selecting

better projects, and innovation expenditure can increase before any financial return becomes observable.

Green credit studies further indicate that financing responses depend on borrower characteristics. Yu and colleagues examine the demand for green finance in relation to constraints on green innovation [8]. Xu and Li consider asymmetric changes in debt financing costs and maturity across enterprise types [9]. Liu, Wang, and Cai connect green credit policy, property rights, and debt financing [10]. Together these studies motivate heterogeneous adjustment channels. Credit restrictions may reduce discretionary expansion by some firms while tightening funding for productive investment by others. The present analysis accordingly treats the sign of the average residual response as an empirical question and avoids assuming that less investment always means greater efficiency.

Environmental investment provides another distinct outcome. Yang and colleagues examine the scale and efficiency of firms' environmental investment around the reform program [11]. Their focus highlights the importance of defining what the investment measure includes. General corporate capital expenditure, pollution-control expenditure, and low-carbon innovation can move differently after a financing reform. An accounting residual based on general investment does not measure avoided emissions or the social return to environmental expenditure. The current paper evaluates the released general investment proxy and maintains this distinction when discussing policy implications.

2.2. Information, Financing Frictions, and Investment Measures

Financial reporting can affect how investors assess projects and how managers use internal funds. Chen and colleagues study reporting quality and investment efficiency among private firms in emerging markets [12]. Their analysis supports attention to information frictions beyond public equity markets. Agency theory identifies a separate source of inefficient decisions when managers' incentives differ from those of capital providers [13]. Myers and Majluf explain how asymmetric information can change financing and investment decisions [14], while Stiglitz and Weiss show why imperfect information can generate credit rationing [15]. These mechanisms make the direction of a financing reform conditional on existing constraints and governance.

Environmental regulation adds another layer. Porter and van der Linde discuss conditions under which environmental requirements can stimulate innovation and offset some compliance costs [16]. Such responses need not appear immediately in annual investment data. Gillingham and Stock emphasize the variation in the costs of reducing greenhouse gas emissions across approaches and contexts [17]. A reduction in an accounting residual therefore cannot establish that a policy chooses the lowest-cost environmental investments. Project-level outcomes, financing terms, and environmental performance would be needed to evaluate that broader claim.

Market-based green finance also involves investor demand. Tang and Zhang study shareholder outcomes around green bond issuance [18]. Zerbib examines how pro-environmental preferences are reflected in bond prices [19]. Pástor, Stambaugh, and Taylor model sustainable investing in equilibrium [20]. This literature explains why financing conditions may respond to environmental preferences even without a geographic policy change. It also identifies a competing time-varying influence: national sustainable-investment trends may affect firms differently according to industry and investor composition. Year effects absorb common shocks, but they cannot absorb every region-specific response to those shocks.

2.3. Identification and Inference in Policy Panels

Recent difference-in-differences research separates treatment timing, comparison groups, and effect aggregation. Callaway and Sant'Anna develop methods for multiple periods and treatment cohorts [21]. Sun and Abraham address event-study estimation with heterogeneous treatment effects [22]. Goodman-Bacon analyses the comparisons embedded in two-way fixed-effect regressions when adoption dates vary [23]. The retained five-province design has a common exposure date, reducing the specific problem

of using earlier-treated groups as controls for later-treated groups. Excluding later pilot provinces helps preserve that structure. It does not remove nonrandom pilot selection, spillovers, or unobserved regional trends.

Event-study coefficients provide evidence about observed pre-policy differences, but their interpretation requires care. Roth shows that pretrend tests can have low power and that selecting a specification after such a test can distort inference [24]. Rambachan and Roth develop sensitivity analysis for departures from parallel trends [25]. The present study reports the lead coefficients and their joint test without using nonrejection as a guarantee of identification. It does not implement a formal parallel-trends sensitivity region, and therefore does not claim robustness to a quantified class of trend violations.

Clustered inference is central when policy exposure varies across regions. Cameron, Gelbach, and Miller develop bootstrap approaches for clustered errors [26]; Cameron and Miller explain practical issues in choosing the clustering level and dealing with few clusters [27]. Abadie's semiparametric estimator and Sant'Anna and Zhao's doubly robust approach provide alternatives for conditional comparison designs [28,29]. Such estimators require their own overlap and nuisance-model assumptions. They do not create additional independent policy assignments. Finally, de Chaisemartin and D'Haultfœuille show how heterogeneous effects complicate two-way fixed-effect interpretation [30]. These contributions motivate an explicit audit of exposure, aggregation, and uncertainty before interpreting a negative policy coefficient.

3. Methodology

3.1. Public Panel and Exposure Audit

The empirical material is the S1 replication archive accompanying Wang and Shao's published study. It was downloaded from the publisher on October 12, 2026. The archive contains a Stata data file and analysis commands; the original accounting information was assembled by the source authors. The released file contains 19,649 observations, 2,204 distinct companies, and annual records from 2012 through 2022. This study performs a secondary analysis of that release. It introduces no simulated observations and does not claim independent collection of the underlying company accounts. The archive is identified by the supplementary-material identifier 10.1371/journal.pone.0313861.s001.

Table 1 reports the sample construction. Company identifiers and years are checked for duplicates, and every released policy interaction is verified against the product of the exposure and post-policy indicators. Fifty-three companies change recorded province at least once. Their 434 observations are excluded to ensure that location, exposure membership, and clustering remain stable within a company. Gansu and Chongqing are then excluded because later pilot initiatives would complicate their use as untreated comparisons over the full observation window. This removes 51 companies and 462 observations. The final panel has 18,753 observations from 2,100 companies: 642 exposed and 1,458 comparison companies.

Table 1. Auditable Sample Construction. Firms Moving between Provinces Are Removed Before Later-Adopter Provinces Are Excluded. All Retained Firms Have a Fixed Province and Exposure Label.

Step	Firm-years	Unique firms
Published replication panel	19,649	2,204
Remove 53 province movers	19,215	2,151
Exclude Gansu and Chongqing	18,753	2,100
Main exposed group	5,728	642

Main comparison group	13,025	1,458

The final sample covers 29 provincial units, including municipalities and autonomous regions. Five units are coded as exposed. Their 5,728 observations are compared with 13,025 observations from the other 24 units. Let i denote a company, t a year, and g(i) its fixed province. The exposed set P contains Guangdong, Zhejiang, Jiangxi, Guizhou, and Xinjiang. Exposure and the common post-policy indicator are defined in Eq. (1). The product Z is the variable called TESTGFP in the release. Its identity is checked numerically rather than inferred from a variable label.

$$D_i = 1\{g(i) \in P\}, \quad P_t = 1\{t \geq 2017\} \qquad \textbf{(1)}$$
$$Z_{it} = D_i P_t$$

The initial pilots were located in selected areas within these provinces. The released TEST variable nevertheless covers companies across each exposed province. The estimand therefore concerns broad provincial exposure, which can include direct participation, local financial spillovers, and unrelated provincial changes. Firms outside the pilot locality are not verified recipients of policy support. Exact loan access or project eligibility cannot be recovered from this indicator. Figure 1 presents the data audit, estimation sequence, and common policy timeline. The first post-policy year is 2017; omitting that transition year is a sensitivity check because the program began during the year.

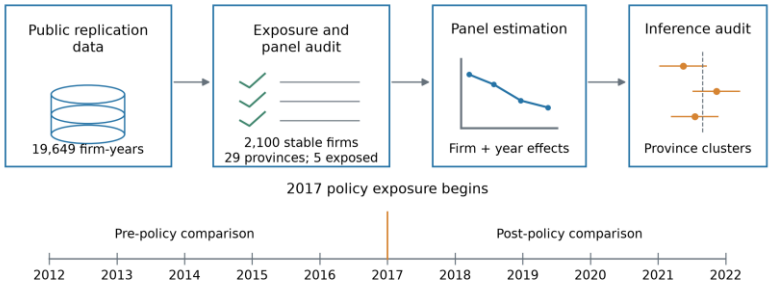


Figure 1. Reanalysis Workflow and Policy Timeline. Data-store and Interval Icons Represent the Public Panel and Clustered Inference. Exposure Is a Province-Level Proxy for Local Pilot Areas; the First Post-Policy Year Is 2017.

3.2. Outcome and Covariates

The dependent variable Y is the released Absinv measure, interpreted as an absolute investment-residual proxy. A lower value represents a smaller deviation from the source study's expected-investment benchmark. The benchmark's original inputs are not all present in the public file, so its first-stage regression is not re-estimated. The release also contains a signed variable inv and an alternative residual Absinvre. In 38 observations, Absinv differs from the absolute value of inv by more than 0.0000001. This is consistent with nonidentical preprocessing or precision, but the precise explanation is not documented. The analysis retains Absinv as released and does not silently replace it with a reconstructed value.

Eight released covariates are used in a sensitivity specification. They represent the largest shareholder's ownership share, independent-director share, leverage, book-to-market ratio, operating-income growth, asset turnover, logged enterprise age, and earnings volatility. Their archive names are Top, Indep, Lever, BM, Grow, ATO, AGE, and EQ. Earnings volatility refers to variation in earnings before interest, taxes, depreciation, and amortization relative to assets. These definitions provide a variable map rather than an independent reconstruction of the source accounting series. Table 2 reports their empirical distributions together with the outcome. No additional winsorization, imputation, or deletion based on estimated effect size is performed.

Table 2. Descriptive Statistics in the Main Sample (18,753 Firm-Years). SD Denotes Standard Deviation. Outcomes and Controls Retain the Publication Archive's Preprocessing. No Additional Winsorization Is Applied.

Variable	Mean	SD	Median	Minimum	Maximum
Absinv	0.03870	0.05066	0.02327	0.00025	0.54997
Top	0.34070	0.14788	0.31970	0.07110	0.76680
Indep	0.37567	0.05448	0.36360	0.30000	0.60000
Lever	0.43260	0.19650	0.42872	0.04673	0.86561
BM	0.34286	0.15822	0.31842	0.05328	0.90302
Grow	0.37497	0.96908	0.13623	-0.68539	10.28792
ATO	0.62698	0.43340	0.53014	0.06882	3.11720
AGE	2.48585	0.59427	2.56495	1.09861	3.43399
EQ	0.02401	0.03214	0.01424	0.00072	0.31394

The main specification excludes current covariates. Financing and governance variables can themselves change after the policy, so conditioning on them could alter the estimand or absorb part of a policy response. The covariate-adjusted model is reported as a specification comparison, not automatically treated as more causally credible. All eight variables are observed in the main sample, allowing the comparison to use exactly the same rows. The reported estimates remain in the normalized residual's original units. Neither their signs nor their magnitudes establish changes in firm value or environmental welfare.

3.3. Target Comparison and Fixed-Effect Estimation

The desired causal quantity is the average treatment effect on the treated (ATT), measured as the average change in Y attributable to provincial exposure among exposed companies in post-policy periods, as written in Eq. (2). Y(1) and Y(0) denote potential outcomes with and without that exposure. This target is distinguished from the coefficient actually identified by the panel regression. The target requires a meaningful exposure definition, no anticipation, a suitable untreated comparison, and assumptions about spillovers. Because exact local participation is not measured, a direct effect on participating firms is outside the identified target.

$$\tau_{ATT} = E[Y_{it}(1) - Y_{it}(0) \mid D_i = 1, t \geq 2017] \quad \textbf{(2)}$$

Parallel trends requires that, absent exposure, changes in the mean outcome would have been comparable across exposed and comparison companies. Equation (3) states this requirement for an adjacent pair of years. It concerns unobserved untreated potential outcomes after the program begins. Similar observed pre-policy changes are informative but cannot verify the post-policy requirement. Firm effects account for time-invariant company differences; year effects account for shocks common to all companies. Provincial credit cycles and other geographically concentrated interventions can still violate the comparison assumption.

$$E[Y_{it}(0) - Y_{i,t-1}(0) \mid D_i = 1] = E[Y_{it}(0) - Y_{i,t-1}(0) \mid D_i = 0] \quad \textbf{(3)}$$

Equation (4) gives the estimating model. The main analysis sets the covariate vector X to be empty. The adjusted analysis uses the eight current covariates described above. Company effects α and year effects δ are included in both models. The parameter τ measures the conditional difference-in-differences association under the selected weighting implicit in ordinary least squares. In an unbalanced panel, that weighting need not equal an equal-company or equal-province average. A balanced-panel sensitivity analysis makes the role of observation availability more visible without assuming that balanced firms are representative.

$$Y_{it} = \alpha_i + \delta_t + \tau Z_{it} + X_{it}^{\top}\beta + u_{it} \quad \textbf{(4)}$$

Fixed effects are removed by alternating within-group projections until the largest update is below 10^{-11}. Let M_F denote the residual-maker for the fixed effects, A the matrix containing Z and any included covariates, and a tilde denote multiplication by M_F. Equation (5) defines the resulting least-squares estimator. This implementation avoids allocating a dense matrix with thousands of company indicators. All calculations use double-precision arithmetic. Unique identifiers, exposure identities, convergence, finite estimates, and outcome decomposition are checked by executable assertions.

$$\hat{b} = (\tilde{A}^\top \tilde{A})^{-1} \tilde{A}^\top \tilde{y}, \quad \tilde{A} = M_F A, \quad \tilde{y} = M_F y \tag{5}$$

3.4. Regional Inference and Restricted Bootstrap

Errors are clustered by fixed province in the principal analysis. The cluster score s_g sums transformed regressors multiplied by fitted residuals within province g. Equation (6) defines the covariance estimator, with H equal to the inverse transformed cross-product matrix. The finite-sample multiplier c uses the number of clusters G, observations n, estimated nonabsorbed regressors K, and absorbed fixed-effect rank r_F. This explicitly counts the absorbed effects in the residual degrees of freedom. It is a transparent finite-sample convention, not a claim that such a correction resolves the limited number of exposed regions.

$$\hat{V} = cH\left[\sum_{g=1}^{G} s_g s_g^\top\right]H, \quad H = (\tilde{A}^\top \tilde{A})^{-1} \tag{6}$$
$$s_g = \sum_{(i,t)\in g} \tilde{a}_{it} \hat{u}_{it}, \quad c = \frac{G}{G-1} \frac{n-1}{n - r_F - K}$$

The principal confidence interval follows Eq. (7), using a t critical value with G−1 degrees of freedom and the standard error (SE) from Eq. (6). Thus the main model uses 28 degrees of freedom, while the balanced sample uses 26. Firm-clustered inference is calculated for the identical main regression as a diagnostic comparison. It permits dependence within each firm but does not capture the same cross-company provincial shocks. It is therefore not used to select the preferred significance statement. Increasing the number of firms cannot by itself increase the number of exposed provinces.

$$CI_{0.95} = \hat{\tau} \pm t_{0.975,G-1}\, SE(\hat{\tau}) \tag{7}$$

A null-imposed wild cluster bootstrap supplements the regional t test. For each model, the null restricts τ to zero while retaining its nuisance regressors and fixed effects. The restricted residuals are multiplied by independent province-level Rademacher weights, each taking −1 or +1 with equal probability. The transformed pseudo-outcome in Eq. (8) is regressed on the same transformed full design. Here X_0 contains the restricted covariates, $\hat{\beta}_0$ their fitted coefficients, and $\hat{u}_0$ the restricted residuals; w assigns a common draw to observations in a province. Each bootstrap coefficient is studentized using the same cluster covariance convention. There are 999 draws with seed 112233, and the two-sided p-value uses the plus-one correction in Eq. (9). The bootstrap addresses a specific finite-sample inference problem; five exposed regions still limit its reliability.

$$\tilde{y}^{*(b)} = M_F[X_0 \hat{\beta}_0 + \hat{u}_0 \odot w^{(b)}] \tag{8}$$

$$p_{wild} = \frac{1 + \sum_{b=1}^{B} 1\{|t^{*(b)}| \geq |t_{obs}|\}}{B+1}, \quad B = 999 \tag{9}$$

3.5. Dynamic and Sensitivity Analyses

The event study replaces the pooled policy interaction with exposure-by-year indicators. Equation (10) uses event time k=t−2017 and omits k=−1, corresponding to 2016. Every other observed year is included, producing four pre-policy and six post-policy coefficients. Pointwise regional confidence intervals are reported, and the four pre-policy coefficients are jointly tested using their cluster covariance matrix. The test is an observed-data diagnostic. It is not a rule for choosing the main specification and does not establish equality of unobserved counterfactual trends.

$$Y_{it} = \alpha_i + \delta_t + \sum_{k=-5,\, k\neq -1}^{5} \theta_k D_i 1\{t - 2017 = k\} + u_{it} \tag{10}$$

Sensitivity analyses restore the later-adopter provinces, end the sample in 2019, omit 2017, retain only firms observed in all 11 years, and substitute Absinvre. A direction-based decomposition retains all observations. Equation (11) multiplies the released absolute outcome by indicators for the sign of inv. This preserves the released magnitude despite the documented discrepancies with the signed variable. The two components sum exactly to Y, so their regression coefficients add to the main coefficient when the design and sample are identical. These are unconditional components, not regressions restricted to observations selected by a post-policy residual sign.

$$Y_{it}^{+} = Y_{it}1\{inv_{it} \geq 0\}, \quad Y_{it}^{-} = Y_{it}1\{inv_{it} < 0\} \qquad (11)$$
$$Y_{it} = Y_{it}^{+} + Y_{it}^{-}$$

Finally, each exposed province is removed in turn. This diagnostic measures how the coefficient changes when one regional contribution is absent. The five estimates share most observations and are not independent replications. They can reveal concentration of identifying variation but cannot isolate mechanisms within a province. The implementation uses the archived data with installed numerical libraries and records the input file's cryptographic hash. Together, the data audit, fixed-effect estimator, bootstrap, and sensitivity checks define a reproducible secondary analysis whose claims are limited to the released observations and the stated exposure proxy.

4. Experiments and Results

4.1. Sample Properties and Unadjusted Patterns

The executable audit confirms 19,649 distinct company-year records in the public release and no duplicate identifier-year pairs. Removing provincial movers leaves 19,215 records; excluding the two later-adopter provinces gives the 18,753-record principal sample in Table 1. The exposed group contains 642 stable-location companies, and the comparison group contains 1,458. There are 29 regional clusters and five exposed clusters. These counts describe independent company identifiers and fixed provincial membership. They avoid counting a company that changes exposure status in both treatment categories. No company is removed because its outcome weakens or reverses the estimated policy association.

The main outcome's mean is 0.03870, its median is 0.02327, and its standard deviation is 0.05066. Its minimum and maximum are 0.000254 and 0.549972. The mean exceeds the median, indicating a right-skewed residual magnitude. Table 2 also shows variation in financial characteristics. Mean leverage is 0.43260, mean largest-shareholder ownership is 0.34070, and mean asset turnover is 0.62698. Operating-income growth has a mean of 0.37497 and a standard deviation of 0.96908. This dispersion motivates retaining the archive's preprocessing and reporting the distribution rather than introducing undocumented additional trimming.

Figure 2 shows the unadjusted annual mean Absinv within each group. Both groups have a pronounced increase in 2015 and 2016 followed by a decline in 2017. The comparison-group mean rises from 0.03782 in 2012 to 0.06366 in 2016 and then falls to 0.04383 in 2017. The exposed-group mean is 0.04389 in 2012, 0.07573 in 2016, and 0.04528 in 2017. The series therefore do not describe an improvement unique to exposed companies. A common investment cycle is visible, while the size of the exposed-comparison gap also changes.

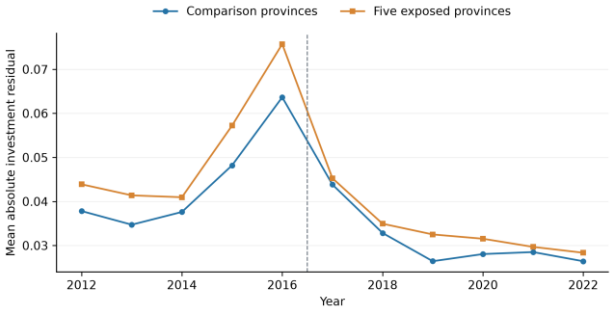


Figure 2. Unadjusted Mean Absolute Investment Residual by Group and Year in the Main Sample. Lower Values Indicate Smaller Published Residuals. The Dashed Line Separates 2016 and 2017; Changes in Group Composition Remain Possible in This Unbalanced Panel.

Pooling observations within the pre-policy and post-policy periods gives comparison means of 0.04476 and 0.03118, and exposed means of 0.05243 and 0.03398. The corresponding unadjusted change-in-changes is approximately −0.00487. These four means use different numbers of observations across periods because the panel is unbalanced. They are descriptive and are not substitutes for the company-fixed-effect coefficient. Their negative sign provides context for the regressions while leaving open whether regional shocks, sample composition, or broad policy exposure account for the difference.

4.2. Baseline Estimation and the Clustering Contrast

Table 3 begins with a reproducibility check of the published regression structure. It uses the complete released sample, eight current covariates, province, broad-industry and year effects, and firm-clustered inference. The policy coefficient is −0.00586 with a standard error of 0.00186. This calculation verifies that the archive supports the reported direction and approximate baseline magnitude. It does not validate every measurement or causal assumption in that structure. In particular, province and industry effects do not remove time-invariant differences between individual companies, and firm clustering does not capture shared provincial disturbances.

Table 3. Baseline and Inference Audit. The Published-Specification Check Uses Province, Broad-Industry and Year Effects with Firm Clustering; Main Models Use Firm and Year Effects. Province Clustering Is Used Except Where Explicitly Labeled. Wild P-Values Use 999 Null-Imposed Rademacher Draws.

Specification	Estimate	SE	t-test p	Wild p	Rows
Published-specification check	-0.00586	0.00186	0.002	-	19,649
Main; province cluster	-0.00419	0.00295	0.166	0.403	18,753
Eight current controls	-0.00412	0.00302	0.183	0.417	18,753
Main; firm cluster	-0.00419	0.00194	0.031	-	18,753

The principal model uses the audited sample, company and year effects, and no current covariates. Its coefficient is −0.00419. A negative coefficient denotes a smaller absolute residual relative to the comparison trajectory under the model assumptions. The province-clustered standard error is 0.00295, giving a 95% interval from −0.01022 to 0.00185. The conventional regional t-test p-value is 0.166, and the restricted wild-bootstrap p-value is 0.403. The data therefore do not provide a precisely estimated average reduction at conventional thresholds under the principal inference approach. The interval also permits economically different positive and negative responses.

The estimate should not be described as a 0.419% or a 4.19% improvement in investment efficiency. Its scale is an accounting residual ratio rather than a percentage efficiency score. Multiplying the coefficient by 100 gives −0.419 percentage points in that ratio, but does not establish a percentage change in firm value, investment returns, or environmental outcomes. Relative to the overall outcome mean, its absolute magnitude is

approximately one-tenth; this is a descriptive scale comparison and not an independently identified causal effect. The uncertainty interval must accompany any discussion of magnitude.

Adding the eight current covariates produces a coefficient of −0.00412 and a province-clustered standard error of 0.00302. The conventional p-value is 0.183 and the wild-bootstrap p-value is 0.417. Similar point estimates across these two models suggest that this particular adjustment set does not account for the entire negative association. They do not show that all relevant confounders are controlled. The variables can contain policy responses, and the archive does not provide a complete model of regional credit supply, project opportunities, or other concurrent interventions.

The same principal regression, clustered by company, has an unchanged coefficient but a smaller standard error of 0.00194 and a p-value of 0.031. Its interval excludes zero, whereas the province-clustered interval includes zero. This is an inference change, not an effect-size change: the outcome, observations, and regression design are identical. The difference indicates that the assumed dependence structure materially affects the significance conclusion. With exposure assigned to five provinces, the firm-level result cannot be used to claim that thousands of independent policy experiments support the negative coefficient.

4.3. Event-Study Evidence

Figure 3 reports the dynamic estimates relative to 2016. The four pre-policy coefficients range from −0.00725 in 2014 to −0.00323 in 2015. Their pointwise 95% regional intervals all include zero. The joint test yields F=1.538 with four numerator and 28 denominator degrees of freedom, corresponding to p=0.218. The observed data do not reject zero leads under this test. The intervals remain wide, however, and the consistent negative signs indicate that the choice of 2016 as a reference matters for visual interpretation. Nonrejection is compatible with undetected differences in counterfactual trends.

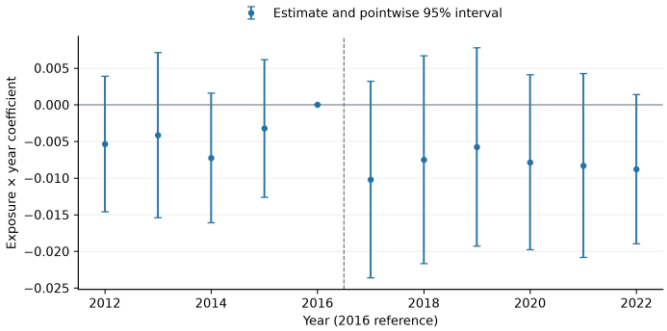


Figure 3. Firm- and Year-Fixed-Effect Event Study with 2016 Omitted. Error Bars Are Pointwise 95% Province-Clustered T Intervals, Using 28 Degrees of Freedom. The Reference-Year Zero Is Imposed, Not Estimated. Intervals Are Not Simultaneous Confidence Bands.

The 2017 coefficient is −0.01020, with a 95% interval of [−0.02363, 0.00322]. Subsequent coefficients remain negative, including −0.00749 in 2018 and −0.00879 in 2022. Their pointwise 95% intervals also include zero. The 2022 conventional p-value is 0.088, which does not establish a 5% finding and is one of several year-specific comparisons. No simultaneous confidence band or multiple-testing adjustment is implemented. The dynamic graph consequently supports a descriptive pattern of smaller post-policy gaps relative to 2016 while providing limited precision for individual-year effects.

The dynamic estimates should not be averaged visually and equated with the pooled coefficient. They compare each year with one reference year, while the pooled regression uses the pre-policy and post-policy observations with weights determined by the unbalanced design and fixed-effect projection. The 2016 outcome peak contributes to the dynamic contrast. This explains why post-policy event coefficients can have larger

negative magnitudes than the pooled estimate without creating a contradiction. Neither representation resolves the uncertainty about whether the counterfactual regional trajectory would have evolved similarly after 2017.

4.4. Sample and Outcome Sensitivity

Table 4 and Figure 4 report design-based sensitivity calculations. Restoring Gansu and Chongqing yields −0.00419, a conventional p-value of 0.162, and a wild-bootstrap p-value of 0.396. The close estimate indicates that these observations have little influence on the pooled coefficient in this release. Their exclusion nevertheless clarifies the comparison group because later exposure creates a substantive design problem. Ending the main sample in 2019 gives −0.00386 and a wild p-value of 0.589. This shorter window reduces exposure to pandemic-period shocks but also reduces post-policy information; it does not sharpen the evidence for an average effect.

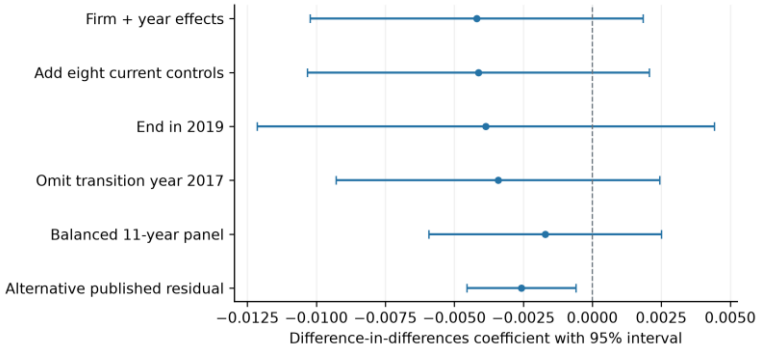


Figure 4. Main and Sensitivity Estimates. Error Bars Are 95% Province-Clustered T Intervals; the Balanced Sample Has 27 Clusters and 26 Degrees of Freedom, While Other Displayed Models Have 29 Clusters and 28 Degrees of Freedom. The Dashed Line Marks Zero.

Table 4. Design Sensitivity, Using Firm and Year Effects with Province Clustering. The Alternative Outcome Is Absinvre as Released. Over-direction and Under-Direction Outcomes Partition the Published Absinv by the Sign of Inv and Retain Zeros.

Specification	Estimate	SE	t-test p	Wild p	Rows
Restore later-adopter provinces	-0.00419	0.00292	0.162	0.396	19,215
End in 2019	-0.00386	0.00405	0.348	0.589	13,764
Omit 2017	-0.00342	0.00286	0.243	0.557	16,904
Balanced 11-year panel	-0.00170	0.00205	0.414	0.637	6,622
Alternative published residual	-0.00256	0.00096	0.013	0.058	18,753
Over-direction component	-0.00093	0.00247	0.710	0.812	18,753
Under-direction component	-0.00326	0.00085	0.001	0.043	18,753

Omitting the transition year 2017 gives −0.00342 with a conventional p-value of 0.243 and a wild p-value of 0.557. Retaining only companies observed in every year gives 6,622 records from 602 firms in 27 provinces. The balanced-panel estimate is −0.00170, with a conventional p-value of 0.414 and a wild p-value of 0.637. Its smaller magnitude indicates sensitivity to the observation structure and retained company population. Balancing also changes the estimand, so the difference cannot be attributed solely to eliminating attrition bias. The original release does not identify why each company-year is absent.

Substituting the released alternative residual gives −0.00256. Its conventional regional p-value is 0.013, but the restricted wild p-value is 0.058. This comparison demonstrates that outcome construction and inference can affect threshold-based conclusions even when signs agree. The alternative variable is taken directly from the archive, and its complete first-stage construction is not independently reproduced. It is therefore a sensitivity outcome with additional measurement uncertainty. Agreement in sign across the two variables is informative, but does not establish that either is a complete measure of efficient investment.

The unconditional direction decomposition yields −0.00093 for the over-direction component and −0.00326 for the under-direction component. They sum to the principal coefficient within numerical precision, as required by their construction. The respective wild p-values are 0.812 and 0.043. The latter is an exploratory component result across several related analyses; it should be interpreted with multiplicity and generated-outcome uncertainty in mind. It is not evidence that every initially underinvesting company benefits, nor a separate proof of a financing mechanism. The main aggregate estimate remains imprecise under regional inference.

4.5. Dependence on Exposed Provinces

Table 5 and Figure 5 show leave-one-exposed-province-out results. Removing Guangdong changes the coefficient to −0.00005, with an interval spanning −0.00327 to 0.00318. Removing Zhejiang instead yields −0.00616, with an interval of [−0.01117, −0.00116]. Other omissions produce estimates between −0.00429 and −0.00460. Guangdong contributes 2,920 observations and 333 companies, while Zhejiang contributes 2,111 observations and 235 companies. These two provinces account for most exposed observations, making their different contributions important for the pooled comparison.

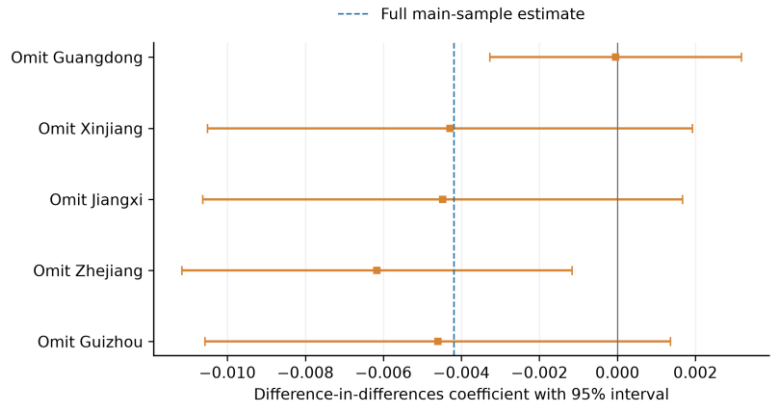


Figure 5. Leave-one-exposed-province-out Estimates with Pointwise 95% Province-Clustered T Intervals. The Blue Dashed Line Is the Full Main-Sample Estimate. Each Omission Leaves Four Exposed Provinces and 28 Total Clusters.

Table 5. Leave-one-exposed-province-out Audit. Each Model Keeps the Remaining Comparison Provinces and Uses Firm and Year Effects with Province Clustering. These Overlapping Estimates Are Diagnostics, Not Independent Tests.

Omitted province	Estimate	95% lower	95% upper	Rows
Guangdong	-0.00005	-0.00327	0.00318	15,833
Xinjiang	-0.00429	-0.01051	0.00193	18,506
Jiangxi	-0.00448	-0.01063	0.00167	18,473
Zhejiang	-0.00616	-0.01117	-0.00116	16,642
Guizhou	-0.00460	-0.01057	0.00137	18,583

The deletion exercise reveals concentration rather than a verified provincial mechanism. Removing a province changes the composition of exposed companies, regional shocks, and regression weights simultaneously. The overlapping estimates cannot be counted as five independent confirmations or selected according to which confidence interval excludes zero. The near-zero estimate without Guangdong and the stronger negative estimate without Zhejiang indicate that a single average conceals material geographic dependence. More granular pilot-boundary and financing data would be needed to determine whether that dependence reflects implementation, industry structure, regional investment cycles, or exposure misclassification.

5. Discussion

The reanalysis separates the direction of an estimated association from the precision of a policy claim. The principal coefficient is negative, but its province-clustered confidence interval includes zero and the restricted wild-bootstrap p-value is 0.403. This result is compatible with a reduction in the released investment-inefficiency proxy, as well as with smaller or opposite average responses. It does not demonstrate that the program had no effect. The sample contains only five exposed provinces, and substantial variation remains concentrated in two large provincial groups.

The contrast between company and province clustering has a practical implication for financial policy research. A large accounting panel can produce apparently precise estimates when common regional disturbances are not represented in the covariance model. The same outcome and regression here yield different significance conclusions under those two dependence assumptions. Reporting both estimates clarifies the issue, while aligning principal inference with the geographic exposure variable maintains a consistent statistical argument. Neither clustered t tests nor a wild bootstrap eliminate uncertainty from nonrandom pilot selection.

Measurement also limits interpretation. The public archive enables verification of sample identities, estimation, and reported directions, but does not contain every input needed to reproduce the original expected-investment model. The provincial exposure proxy extends beyond verified local pilot participation. A smaller absolute residual may capture better project allocation, changing opportunities, or altered accounting relationships. It does not directly establish lower emissions, higher social returns, or improved capital productivity. The under-direction decomposition is suggestive, but its exploratory status and dependence on generated outcomes require a restrained interpretation.

A useful next study would link companies to stable historical addresses and precise pilot boundaries, then measure credit access and project outcomes independently of the policy indicator. Such data could separate direct participation from spillovers and permit more informative mechanism tests. Additional regional cohorts would require estimators that preserve suitable untreated comparisons across adoption dates. The present evidence supports that more detailed evaluation agenda while showing why negative coefficients, pretrend nonrejection, and large observation counts are insufficient foundations for a confident average causal claim.

6. Conclusion

This study evaluates the relationship between green finance policy exposure and corporate investment efficiency through a reproducible reanalysis of a public company panel. The analysis retains the source study's investment-residual proxy and examines treatment coding, provincial mobility, later policy entrants, fixed-effect specification, and clustered uncertainty. The principal sample contains 18,753 firm-year observations from 2,100 companies in 29 provinces, of which five are classified as exposed to the 2017 program.

The main company- and year-fixed-effect coefficient is −0.00419. Its province-clustered 95% confidence interval spans −0.01022 to 0.00185, and the null-imposed wild-bootstrap p-value is 0.403. The negative estimate is consistent with a smaller absolute investment residual, but the principal inference does not establish a precisely estimated average provincial reduction. Clustering the identical regression by company yields a p-value of 0.031, demonstrating that the dependence assumption changes the significance conclusion without changing the estimated coefficient.

Event-study estimates and sensitivity analyses provide context for this uncertainty. The joint pre-policy test does not reject zero leads, yet cannot verify the counterfactual trend assumption. Alternative samples retain negative coefficients with limited precision. Geographic deletion diagnostics show substantial dependence on Guangdong and a

different contribution from Zhejiang. The direction-based components preserve an exact additive accounting relationship, while their interpretation remains exploratory.

The substantive conclusion is bounded by the available data. Provincial exposure is broader than confirmed pilot participation, and the expected-investment model cannot be fully reconstructed from the release. The analysis therefore supports cautious interpretation of a negative association and demonstrates an auditable evaluation workflow. Future evidence should connect precise policy boundaries, credit conditions, project choices, and environmental outcomes. Expanding those measurements can improve the assessment of green finance without treating a large firm-level sample as a substitute for independent policy variation or overlooking uncertainty in the investment-efficiency proxy and its construction under the observed provincial design.

References

1. C. Flammer, "Corporate green bonds," *Journal of Financial Economics*, vol. 142, no. 2, pp. 499–516, 2021, doi: 10.1016/j.jfineco.2021.01.010.
2. S. Richardson, "Over-investment of free cash flow," *Review of Accounting Studies*, vol. 11, no. 2–3, pp. 159–189, 2006, doi: 10.1007/s11142-006-9012-1.
3. G. C. Biddle, G. Hilary, and R. S. Verdi, "How does financial reporting quality relate to investment efficiency?," *Journal of Accounting and Economics*, vol. 48, no. 2–3, pp. 112–131, 2009, doi: 10.1016/j.jacceco.2009.09.001.
4. H. Wang and R. Shao, "Does the green finance policy affect the efficiency of corporate investment? Evidence from China's green finance reform and innovation pilot zones," *PLOS ONE*, vol. 19, no. 11, Art. no. e0313861, 2024, doi: 10.1371/journal.pone.0313861.
5. M. Bertrand, E. Duflo, and S. Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?," *The Quarterly Journal of Economics*, vol. 119, no. 1, pp. 249–275, 2004, doi: 10.1162/003355304772839588.
6. T. Zhang, "Can green finance policies affect corporate financing? Evidence from China's green finance innovation and reform pilot zones," *Journal of Cleaner Production*, vol. 419, Art. no. 138289, 2023, doi: 10.1016/j.jclepro.2023.138289.
7. S. Liu and Y. Wang, "Green innovation effect of pilot zones for green finance reform: Evidence of quasi natural experiment," *Technological Forecasting and Social Change*, vol. 186, Art. no. 122079, 2023, doi: 10.1016/j.techfore.2022.122079.
8. C. Yu, X. Wu, D. Zhang, S. Chen, and J. Zhao, "Demand for green finance: Resolving financing constraints on green innovation in China," *Energy Policy*, vol. 153, Art. no. 112255, 2021, doi: 10.1016/j.enpol.2021.112255.
9. X. Xu and J. Li, "Asymmetric impacts of the policy and development of green credit on the debt financing cost and maturity of different types of enterprises in China," *Journal of Cleaner Production*, vol. 264, Art. no. 121574, 2020, doi: 10.1016/j.jclepro.2020.121574.
10. X. Liu, E. Wang, and D. Cai, "Green credit policy, property rights and debt financing: Quasi-natural experimental evidence from China," *Finance Research Letters*, vol. 29, pp. 129–135, 2019, doi: 10.1016/j.frl.2019.03.014.
11. Q. Yang, S. Ming, R. Zhang, and H. Yan, "Green finance and corporate environmental investment: 'Scale Up' or 'Efficiency Up'?," *PLOS ONE*, vol. 19, no. 2, Art. no. e0297456, 2024, doi: 10.1371/journal.pone.0297456.
12. F. Chen, O.-K. Hope, Q. Li, and X. Wang, "Financial Reporting Quality and Investment Efficiency of Private Firms in Emerging Markets," *The Accounting Review*, vol. 86, no. 4, pp. 1255–1288, 2011, doi: 10.2308/accr-10040.
13. M. C. Jensen and W. H. Meckling, "Theory of the firm: Managerial behavior, agency costs and ownership structure," *Journal of Financial Economics*, vol. 3, no. 4, pp. 305–360, 1976, doi: 10.1016/0304-405X(76)90026-X.
14. S. C. Myers and N. S. Majluf, "Corporate financing and investment decisions when firms have information that investors do not have," *Journal of Financial Economics*, vol. 13, no. 2, pp. 187–221, 1984, doi: 10.1016/0304-405X(84)90023-0.
15. J. E. Stiglitz and A. Weiss, "Credit rationing in markets with imperfect information," *American Economic Review*, vol. 71, no. 3, pp. 393–410, 1981.
16. M. E. Porter and C. van der Linde, "Toward a New Conception of the Environment-Competitiveness Relationship," *Journal of Economic Perspectives*, vol. 9, no. 4, pp. 97–118, 1995, doi: 10.1257/jep.9.4.97.
17. K. Gillingham and J. H. Stock, "The Cost of Reducing Greenhouse Gas Emissions," *Journal of Economic Perspectives*, vol. 32, no. 4, pp. 53–72, 2018, doi: 10.1257/jep.32.4.53.
18. D. Y. Tang and Y. Zhang, "Do shareholders benefit from green bonds?," *Journal of Corporate Finance*, vol. 61, Art. no. 101427, 2020, doi: 10.1016/j.jcorpfin.2018.12.001.
19. O. D. Zerbib, "The effect of pro-environmental preferences on bond prices: Evidence from green bonds," *Journal of Banking & Finance*, vol. 98, pp. 39–60, 2019, doi: 10.1016/j.jbankfin.2018.10.012.
20. Ľ. Pástor, R. F. Stambaugh, and L. A. Taylor, "Sustainable investing in equilibrium," *Journal of Financial Economics*, vol. 142, no. 2, pp. 550–571, 2021, doi: 10.1016/j.jfineco.2020.12.011.
21. B. Callaway and P. H. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, vol. 225, no. 2, pp. 200–230, 2021, doi: 10.1016/j.jeconom.2020.12.001.

22. L. Sun and S. Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, vol. 225, no. 2, pp. 175–199, 2021, doi: 10.1016/j.jeconom.2020.09.006.
23. A. Goodman-Bacon, "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, vol. 225, no. 2, pp. 254–277, 2021, doi: 10.1016/j.jeconom.2021.03.014.
24. J. Roth, "Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends," *American Economic Review: Insights*, vol. 4, no. 3, pp. 305–322, 2022, doi: 10.1257/aeri.20210236.
25. A. Rambachan and J. Roth, "A More Credible Approach to Parallel Trends," *Review of Economic Studies*, vol. 90, no. 5, pp. 2555–2591, 2023, doi: 10.1093/restud/rdad018.
26. A. C. Cameron, J. B. Gelbach, and D. L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, vol. 90, no. 3, pp. 414–427, 2008, doi: 10.1162/rest.90.3.414.
27. A. C. Cameron and D. L. Miller, "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, vol. 50, no. 2, pp. 317–372, 2015, doi: 10.3368/jhr.50.2.317.
28. A. Abadie, "Semiparametric Difference-in-Differences Estimators," *The Review of Economic Studies*, vol. 72, no. 1, pp. 1–19, 2005, doi: 10.1111/0034-6527.00321.
29. P. H. Sant'Anna and J. Zhao, "Doubly robust difference-in-differences estimators," *Journal of Econometrics*, vol. 219, no. 1, pp. 101–122, 2020, doi: 10.1016/j.jeconom.2020.06.003.
30. C. de Chaisemartin and X. D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, vol. 110, no. 9, pp. 2964–2996, 2020, doi: 10.1257/aer.20181169.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.