Curbing Bureaucratic Information Manipulation: Evidence from a Statistics Reform in China

Yongwei Nian

Bocconi University

EEA-ESEM 2023

Motivation

- Accurate economic data: key to policy-making and academic research, but countries may have the incentive to manipulate (Wallace 2016; Martinez 2022)
  - Why? To paint a positive portrait of the country...
- Within countries, lower-level governments also manipulate if evaluated by the data (Xiong 2018)
  - E.g., China relies on GDP growth to promote local leaders

"Data published by the government are not reliable, especially for GDP; they are just for reference"

— Li Keqiang, former premier of China

RQ: Can top-down monitoring decrease data manipulation?

- **A statistics reform in China as a shock to monitoring**
  - Initiated by the central govt. in 2009 ▸ Detailed background
  - Increased punishment, but mainly relied on survey teams in 40% of counties pre-deployed in 2005 to detect manipulation ⇒ DiD
  - Identification further strengthened by using random assignment of earlier survey teams as IV

- **Disaggregated data at the county and politician level**
  - Focus on manipulation of GDP growth rate: the overarching economic indicator for politicians (Li and Zhou 2005; Xu 2011)
  - Measured as discrepancy b/w reported GDP growth and light growth (Henderson, Storeygard, and Weil 2012)

Preview of key findings

- **GDP growth manipulation: $\downarrow$ by 0.6 pp.**
  - $\approx$ 5% of mean reported GDP growth
  - Key mechanism: disciplining effect of the reform
- **Politicians exert more efforts in stimulating the economy**
  - Government policies became more conducive to growth
  - Bank loans and deposits expanded
  - Firm entry increased, especially for productive firms
  - Citizen attitudes towards government improved
- **Consistent with economic tournament model** ▸ Theory
  - Effort shift: GDP growth manipulation $\rightarrow$ stimulating the economy

Relation to literature

- **GDP manipulation in autocracies** (Wallace 2016; Xiong 2018; Martinez 2022)
  - Aggregate analysis; correlational; only existence of manipulation
  - This paper: county-level quasi-experimental design; monitoring ↘ manipulation
- **The role of monitoring in disciplining politicians** (Olken 2007; Ferraz and Finan 2008; Avis, Ferraz, and Finan 2018; Vannutelli 2021)
  - Settings with electoral and judicial accountability
  - This paper: politicians accountable to superiors; weak legal system ⇒ different incentives
- **Real consequences of corruption (misbehaviors)** (Giannetti et al. 2021; Colonnelli and Prem 2022; Ajzenman 2021)
  - Corruption can affect govt. policies and resource allocation
  - This paper: similar findings but focus on an overlooked and seemingly harmless aspect—GDP manipulation

Data

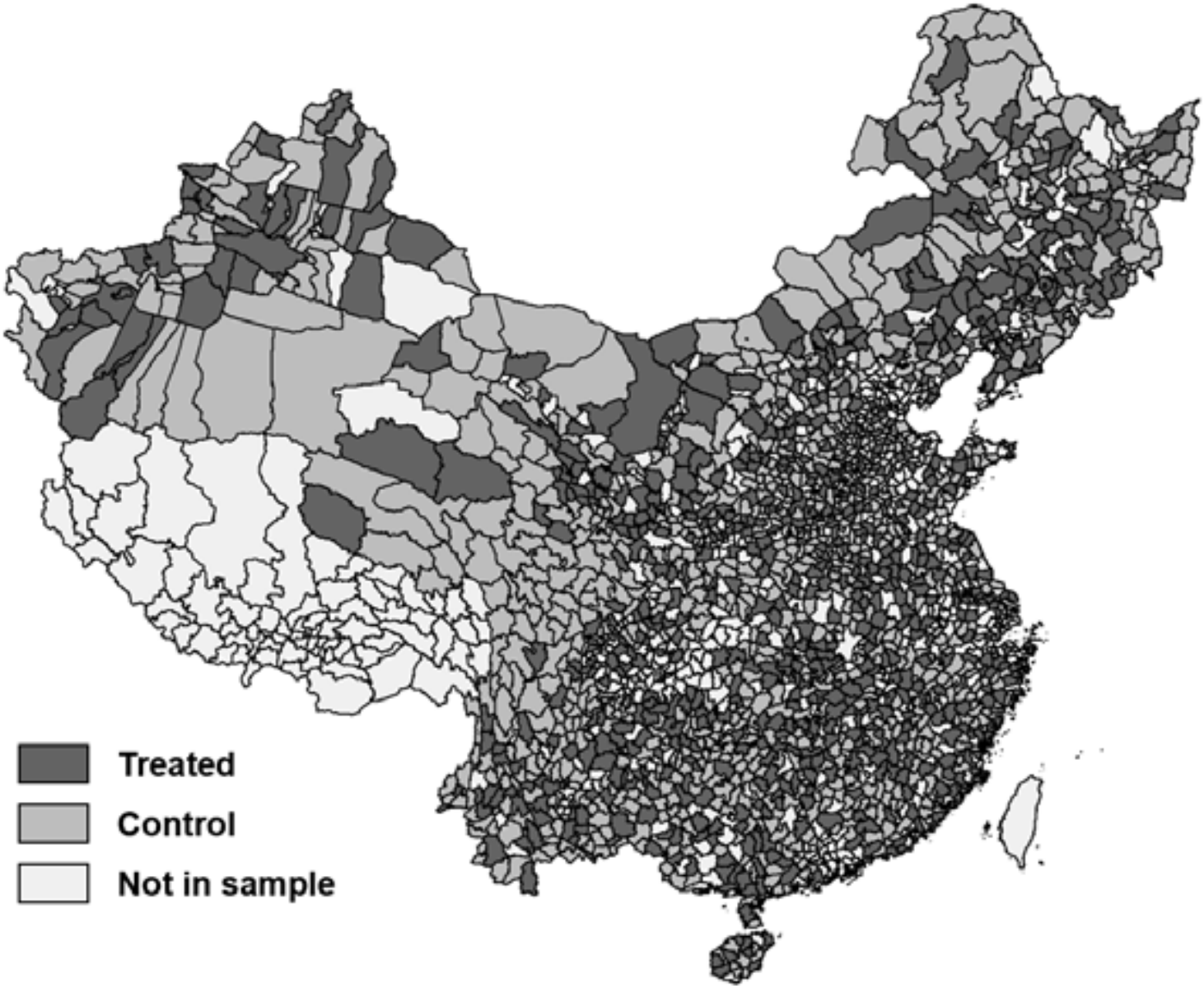
Data structure: county-level panel data from 2005-2018

- **County-level economic outcomes**
  - County-level data on GDP from statistics yearbooks
  - Harmonized nighttime light data up to 2018 (Li et al. 2020)
- **Local leader resumes**
  - Data on the top leaders manually from govt. website and online encyclopedia
- **Household surveys**
  - The China Family Panel Studies to measure citizens' trust in local officials and evaluation of local govt.
- **Other auxiliary data**
  - Government work reports, bank loans and deposits, firm registration, etc.

Sample selection

- Treated counties: counties with the survey teams
- Control counties: other counties, excluding the following types:
  - Counties in the four centrally-managed cities: Beijing, Shanghai, Tianjin, and Chongqing; these counties ranked one-level higher than other counties
  - Counties in Tibet where data are unavailable
  - Counties outside mainland China
  - Urban districts
- Final data: 1779 counties in total, 40% treated

Dist. of treated counties (those with the survey teams)



Mean comparison: no significant difference for most covariates[1]

	Treated		Control		Difference		
	Mean	SD	Mean	SD	T - C	SE	*p*-value
Panel A: Demography							
log Population (2010)	13.00	0.70	12.50	0.81	0.50	0.04	**0.00**
Share urban (%, 2010)	33.65	12.30	34.64	14.31	-0.99	0.66	**0.13**
Share 15-64 (%, 2010)	72.53	4.56	72.44	4.53	0.10	0.22	**0.67**
Years of schooling (2010)	8.23	0.73	8.19	1.05	0.04	0.05	**0.34**
Panel B: Economic development							
Unemployment rate (%, 2010)	2.08	1.39	2.16	1.52	-0.09	0.07	**0.23**
Share primary sectors (%, 2010)	64.97	18.20	64.30	17.93	0.68	0.88	**0.44**
Share secondary sectors (%, 2010)	16.10	12.79	15.49	11.54	0.61	0.59	**0.30**
log GDP (2004)	12.47	0.94	11.97	1.03	0.50	0.05	**0.00**
log GDP (2008)	12.96	0.97	12.48	1.06	0.48	0.05	**0.00**
GDP growth (%, 2002-2004 average)	11.28	6.26	11.57	6.53	-0.29	0.33	**0.38**
GDP growth (%, 2006-2008 average)	12.86	6.38	12.67	6.63	0.19	0.32	**0.54**
Light growth (%, 2002-2004 average)	18.24	12.27	18.82	16.28	-0.58	0.72	**0.42**
Light growth (%, 2006-2008 average)	6.86	10.31	7.67	11.93	-0.81	0.55	**0.14**
Distance to major roads (km, 2010)	69.18	96.10	75.21	80.62	-6.03	4.24	**0.16**
Distance to major railways (km, 2010)	70.16	102.37	74.22	90.45	-4.06	4.63	**0.38**
Panel C: Geography							
County area (km^2)	3900	7453	4128	10222	-228	446	**0.61**
Precipitation (inches, 2004)	0.04	0.09	0.03	0.10	0.00	0.00	**0.55**
Temperature (degrees, 2004)	13.91	5.23	13.57	5.41	0.34	0.26	**0.19**
Precipitation (inches, 2008)	0.05	0.11	0.05	0.11	0.01	0.01	**0.20**
Temperature (degrees, 2008)	13.73	5.08	13.43	5.25	0.30	0.25	**0.23**
Distance to major rivers (km)	59.17	61.15	57.76	59.03	1.41	2.90	**0.63**
Distance to country border (km)	346.52	251.00	345.25	251.49	1.26	12.16	**0.92**
Distance to coastline (km)	616.84	612.04	640.02	568.79	-23.18	28.46	**0.42**
Distance to prefecture center (km)	60.14	41.78	62.66	46.66	-2.52	2.17	**0.25**

[1] Except for GDP, all variables in this table are from sources that the county has no control on. The reason for using the year 2010 for the demographic data, instead of years before the reform, is that the population census was only conducted in 2000 and 2010. Similarly, the transportation data is also in 2010 due to data limitation

Link GDP growth manipulation to the reform

Notation

c, t:	county, year	l_{ct}:	nighttime light intensity growth
z_{ct}:	reported GDP growth	m_{ct}:	GDP manipulation
y_{ct}:	unobservable true GDP growth	$z_{ct} - m_{ct}$:	observable true GDP growth

According to Henderson, Storeygard, and Weil 2012, observable true GDP growth, $z_{ct} - m_{ct}$, can be written as:

$$z_{ct} - m_{ct} = y_{ct} + \epsilon^z_{ct} \tag{1}$$

nighttime light intensity growth, l_{ct}, can be written as:

$$l_{ct} = \gamma y_{ct} + \epsilon^l_{ct} \tag{2}$$

(1) & (2) $\Rightarrow$

$$m_{ct} = z_{ct} - \frac{1}{\gamma} l_{ct} + \epsilon^m_{ct} \tag{3}$$

Include m_{ct} in a DiD framework:

$$m_{ct} = \beta \, Treat_c \times Post_t + \delta_c + \lambda_t + \varepsilon^m_{ct} \tag{4}$$

where $Treat_c = 1_{\{\text{counties with survey teams}\}}$, $Post_t = 1_{\{t \geq 2009\}}$

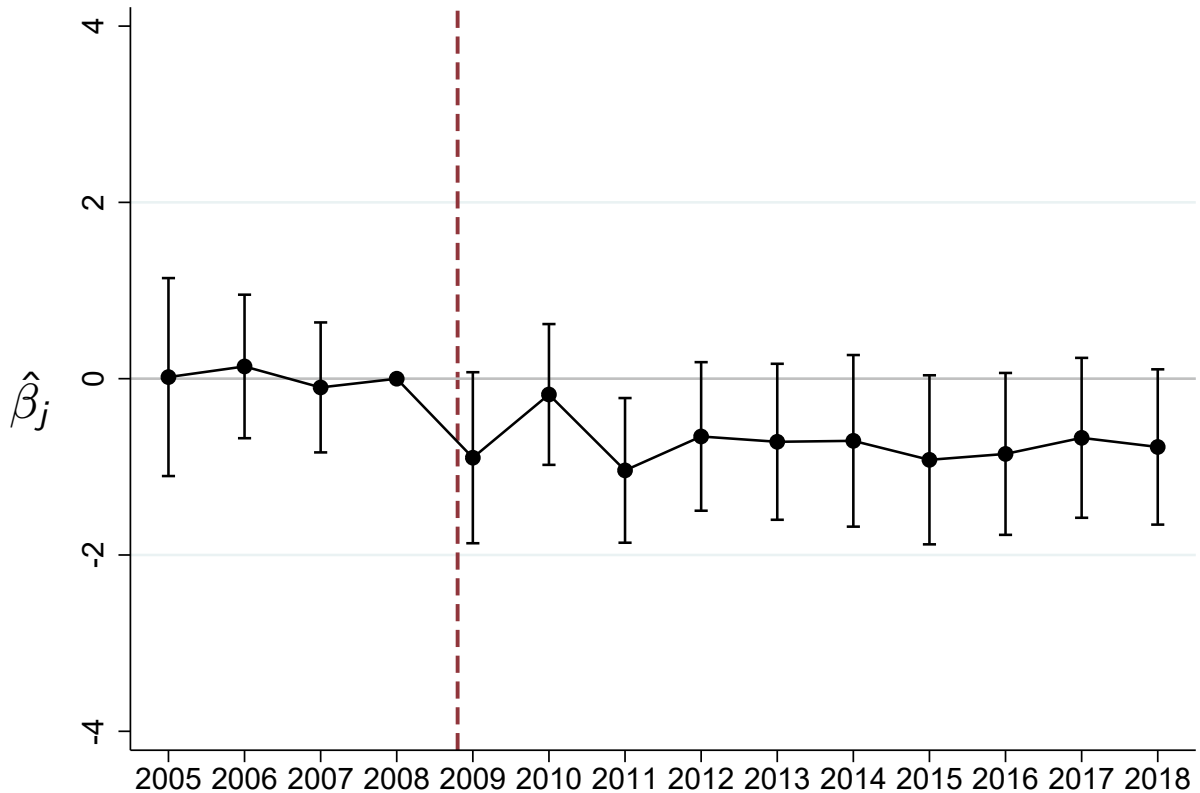
Insert (3) in (4) and rearrange to get the final regression equation:

$$\underbrace{z_{ct}}_{\substack{\text{Reported} \\ \text{GDP growth}}} = \frac{1}{\gamma} \underbrace{l_{ct}}_{\text{Light growth}} + \beta \, Treat_c \times Post_t + \delta_c + \lambda_t + \varepsilon^z_{ct} \tag{5}$$

▸ Discussion on identification concerns

Effect on GDP growth manipulation: event study

$$ReportedGDPGrowth_{ct} = \alpha LightGrowth_{ct} + \sum_{j=2005,\ j\neq 2008}^{j=2018} \beta_j Treat_c \times 1_{\{t=j\}} + \delta_c + \lambda_t + \varepsilon_{ct}$$



▸ Decomposition of the effect

Summarizing the effect on GDP growth manipulation: DiD

$$z_{ct} = \alpha l_{ct} + \beta Treat_c \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct}$$

- ↘ in GDP growth manipulation ≈5% of mean reported GDP growth

	(1)	(2)	(3)	(4)
Dep. var.:	Reported GDP growth (%)			
Treat x Post	-0.751**	-0.869***	-0.552***	-0.576***
	(0.316)	(0.331)	(0.162)	(0.161)
Light growth (%)	0.023***	0.022***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
Demographic controls x Post	No	Yes	Yes	Yes
Economic controls x Post	No	No	Yes	Yes
Geographic controls x Post	No	No	No	Yes
County and year FE	Yes	Yes	Yes	Yes
Cluster level	County	County	County	County
Observations	23,360	22,580	20,343	20,273
R-squared	0.269	0.269	0.362	0.362
Mean dep. var.	10.97	10.97	10.84	10.84
Sample period	2005-2018			

▸ Placebo reform ▸ Randomization inference ▸ Spillover effect ▸ Flexible α ▸ Alternative DD

IV: the randomly-assigned survey teams in 1984

Treated counties

IV: counties with the earlier survey teams

- The survey teams deployed in 2005 were mostly restructured from earlier survey teams set up in 1984 (see fig.) → IV relevant
- The earlier survey teams were selected through *systematic random sampling* of counties within provinces → IV exogenous

Discussion on exogeneity of the IV

- **Random assignment $\Rightarrow$ IV uncorrelated w/ county traits**
  - Supported by ▸ Balance test
- **Legacy effect on contemporary outcomes**
  - The earlier survey teams were led by local govt. and did not conduct any disciplining actions
  - Can have any legacy effect, as long as it did not change after 2009

Effect on GDP growth manipulation - 2SLS results

$$Treat_c \times Post_t = \theta I_{ct} + \alpha Treat_c^{1984} \times Post_t + \delta_c + \lambda_t + \epsilon_{ct}$$

$$z_{ct} = \vartheta I_{ct} + \beta \widehat{Treat_c \times Post_t} + \delta_c + \lambda_t + \varepsilon_{ct}$$

	(1)	(2)	(3)	(4)
Dep. var.:	Reported GDP growth (%)			
Treat x Post	-0.835*	-0.750*	-0.498**	-0.487**
	(0.444)	(0.448)	(0.225)	(0.225)
Light growth (%)	0.023***	0.022***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
Demographic controls x Post	No	Yes	Yes	Yes
Economic controls x Post	No	No	Yes	Yes
Geographic controls x Post	No	No	No	Yes
County and year FE	Yes	Yes	Yes	Yes
Cluster level	County	County	County	County
Observations	22,998	22,580	20,343	20,273
***F*-stat of excl. inst.**	**1,945**	**2,210**	**2,044**	**2,080**
Mean of dep. var.	10.98	10.97	10.84	10.84
Treat x Post (OLS)	-0.751**	-0.869***	-0.552***	-0.576***
	(0.316)	(0.331)	(0.162)	(0.161)

▸ Instrumented event study

Mechanisms and alternative explanations

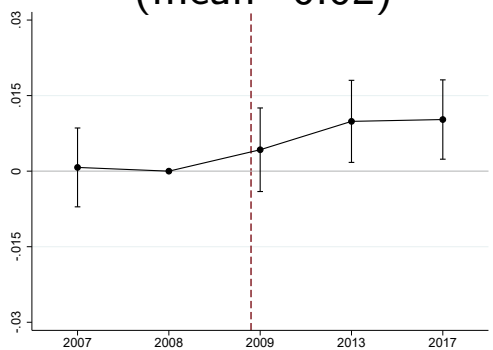
▸ Results

- **Disciplining effect on local leaders ✓**
  - Reform → higher political and legal costs of manipulation → local leaders refrain from manipulation
- **Better political selection after the reform ✗**
  - Reform → removal of bad leaders; entry of good leaders
- **Better soft information ✗**
  - Survey teams → better soft information about local leaders → less role of GDP in promotion → less manipulation
- **Improvement in local statistical capacity ✗**
  - Survey teams → improvement in local statistical capacity → less statistical errors → less gap b/w reported gdp and light
- **Concurrent reforms strengthening monitoring ✗**
  - e.g., Province-Managing-County reform, anti-corruption campaign, etc

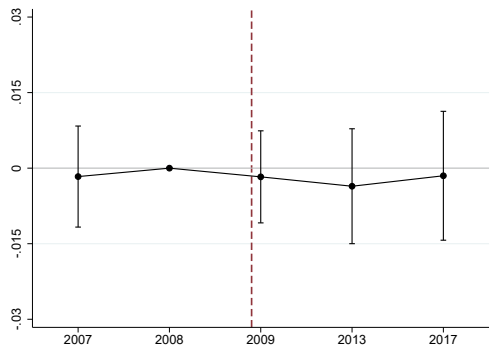
1. Govt. policies more conducive to economic growth

- Policies measured through textual analysis of government annual work reports (4 years: 2007, 2008, 2009, 2013)

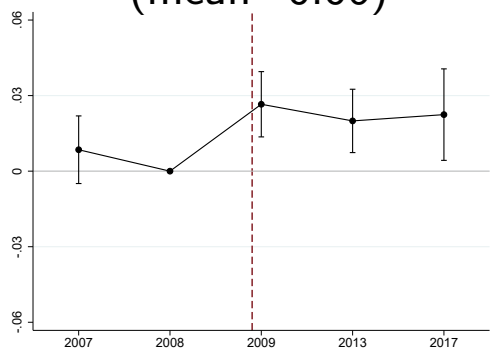
Business attraction (mean=0.02)



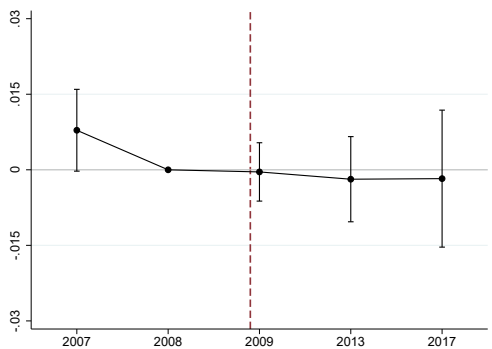
Infrastructure



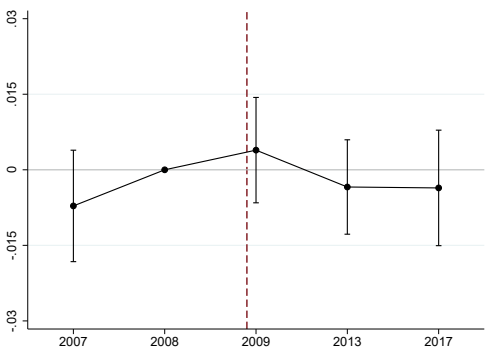
Market reform (mean=0.06)



Policy experiment



Social welfare



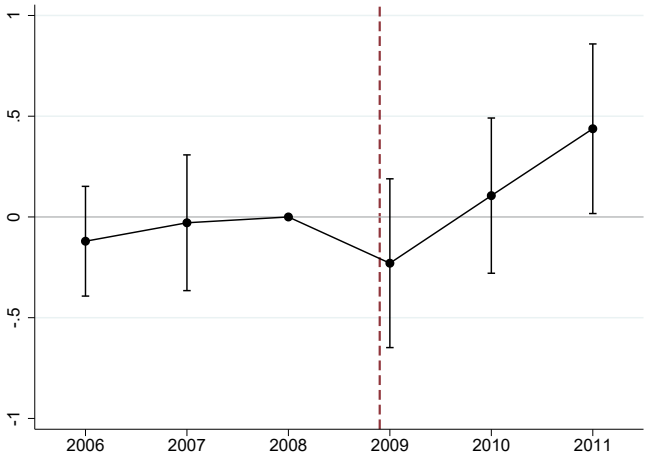
Notes: DVs: # key words in each topic / # sentences in report

2. Bank loans and deposits expanded

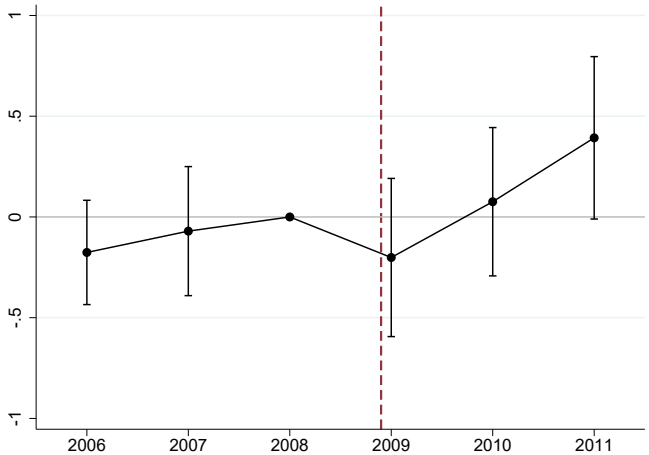
▸ Compare state banks vs other banks

- County-level banking data available from 2006-2011

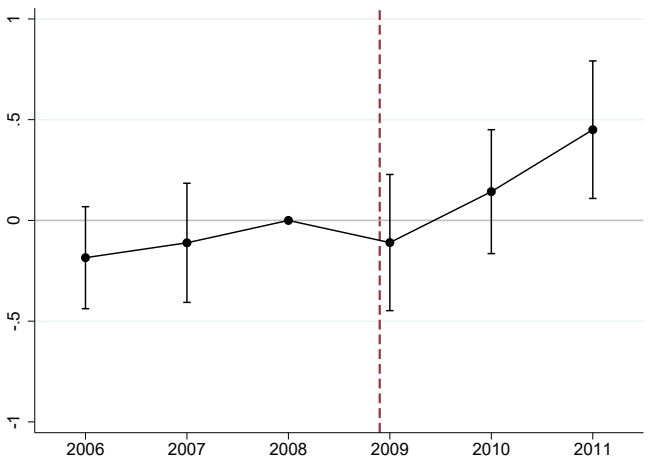
IHS (Deposits)



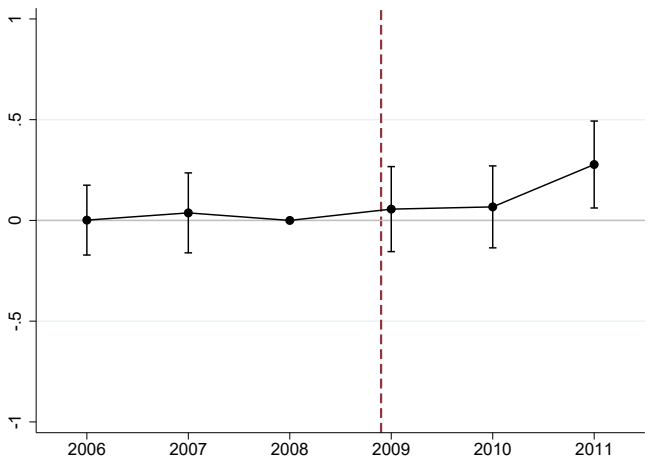
IHS (Loans)



IHS (Loans to small firms)



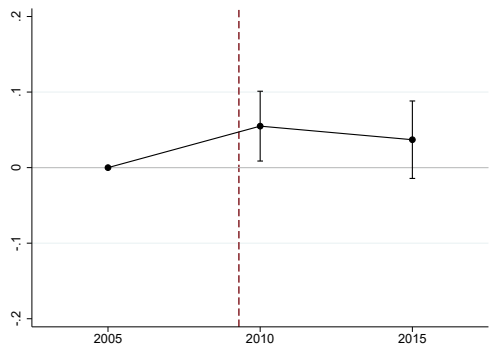
IHS (# Firms granted loans)



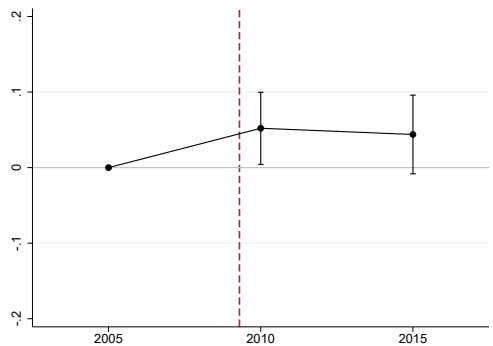
3. Firm entry increased, the more so for productive types

- Firm entry measured using the universe of firm registration data in 2005, 2010, and 2015 (only 3 years due to data limitation)

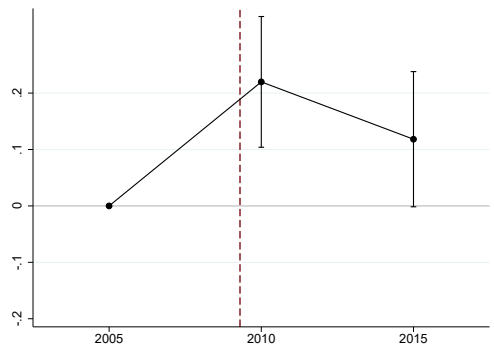
IHS (#All firms)



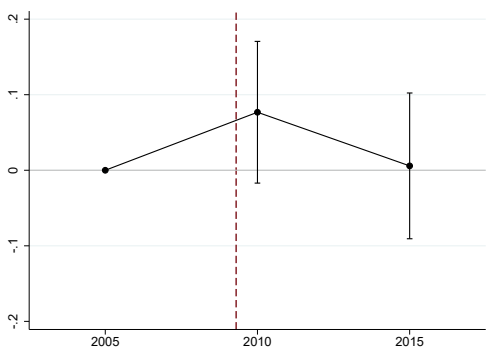
IHS (#Private)



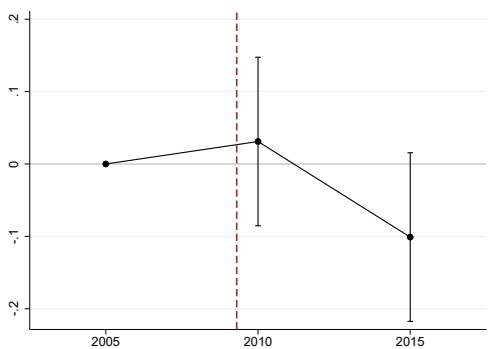
IHS (#SOEs)



IHS (#Foreign)



IHS (#Collective)



▸ IV estimation

4. Citizen attitudes towards govt.: specification

- Cohort DD due to lack of pre-reform survey data (2014, 2016)
  - Compare: (1) treatment counties vs control counties (2) most affected vs less affected cohorts
  - Keep and sort cohorts into 3 decades: 1970s, 1980s, and 1990s
  - Cohorts born in the 1990s are most affected
    - younger than 30s during the reform period
    - attitudes most permeable before 30s (Krosnick and Alwin 1989)
  - Cohorts born in the 1970s are unaffected

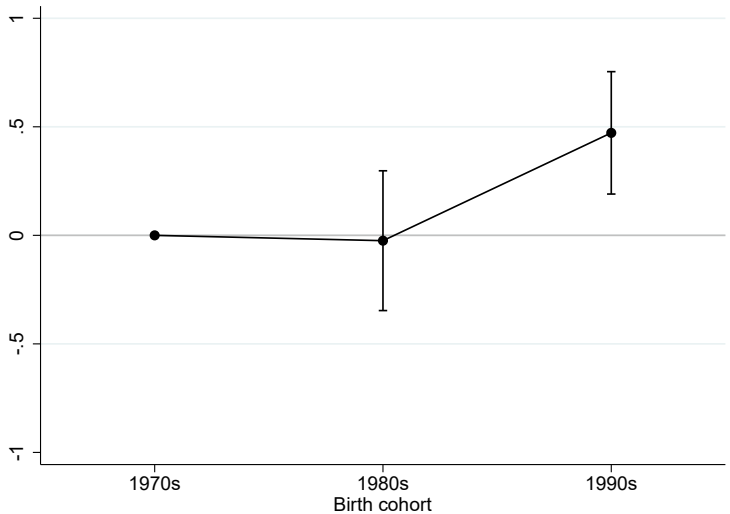
$$Y_{icg} = \sum_{k=1980s,1990s} \beta_k Treat_c \times 1_{\{g=k\}} + \delta_c + \lambda_g + W_i\Omega + X_{cg}\Psi + \varepsilon_{icg}$$

- i: individual c: county g: cohort
- omitted group: 1970s

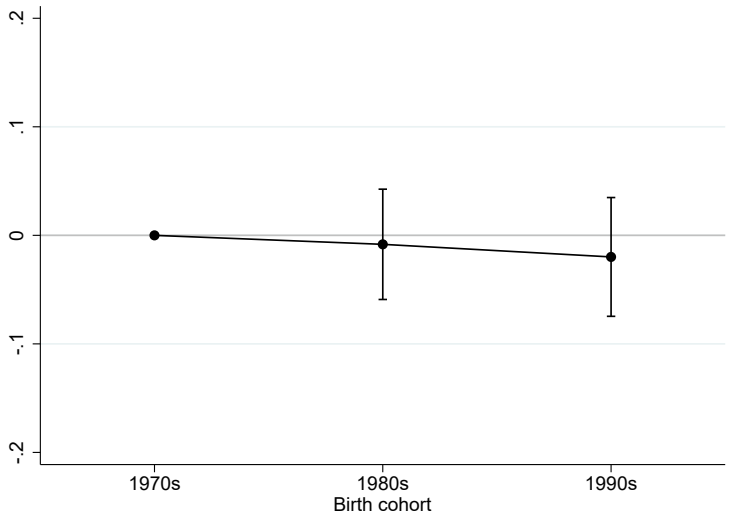
4. Citizen attitudes towards govt. improved

▸ IV estimation

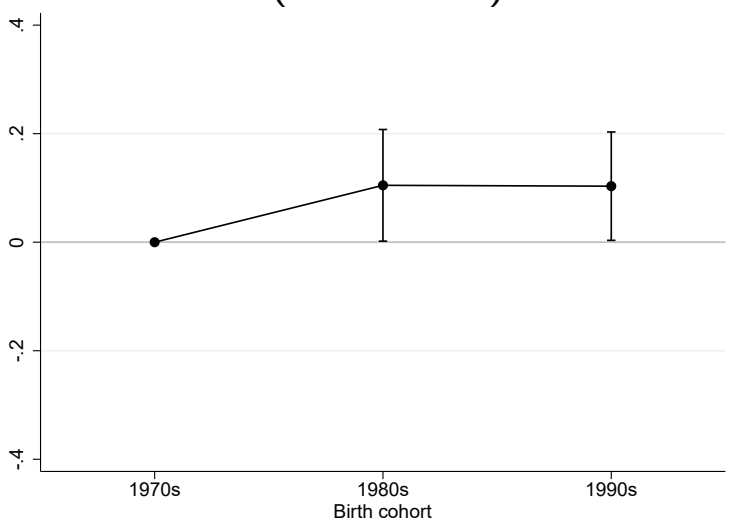
Trust in local officials
(mean=4.8)



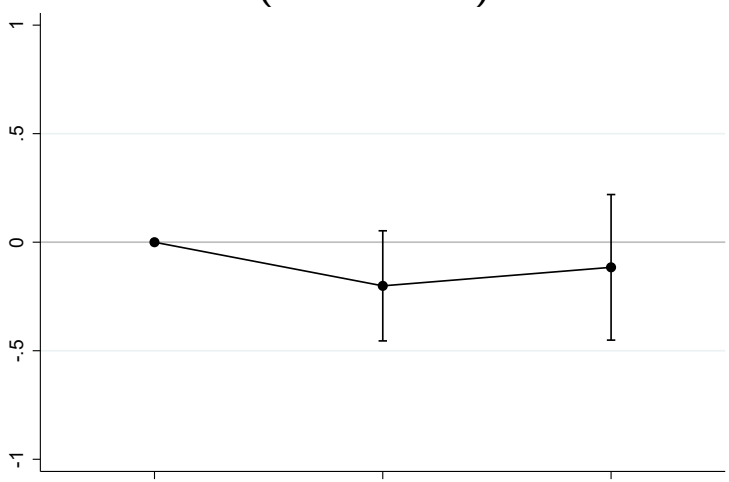
Trust in most people (0/1)
(mean=0.57)



Evaluation of govt. performance
(mean=3.4)



Perceived corruption
(mean=6.9)



▸ Objective corruption measure

Summary and Discussion

- Casual evidence on the effectiveness of top-down monitoring in reducing data manipulation
- Downstream effects: eliciting government efforts in directions conducive to the economy
- New perspective on how politician misbehaviors affect the economy:
  - Corruption: directly stifles growth
  - GDP manipulation: through unfavorable shift of efforts

Thank you!

How is regional GDP calculated and manipulated in China?

Back

Calculation of GDP

- Sum of value added in all sectors in a region. County is lowest level of region for GDP calculation
- Calculated by the local statistics bureau, which is controlled by local leaders in terms of personnel and funding

How do local leaders manipulate GDP?

- Directly ask local statistics bureau to make up numbers, ask firms to overstate income or pay additional "tax" and return later, or double counting: count firms' non-local subsidiaries

The statistics reform in 2009

- **Who started the reform?**
  - The National Bureau of Statistics (NBS), together with other central authorities, in May 2009
  - Goal: discipline misbehaviors of local governments in data processing
- **Who detects misbehaviors?**
  - Local statistics bureau: funded and controlled by local leaders $\rightarrow$ dysfunctional due to aligned incentives with local leaders
  - The survey teams: appointed and funded by the central govt. $\rightarrow$ higher probability of detecting misbehaviors
- **What is the punishment upon detection?**
  - Warning, demerit, demotion, or dismissal
- **Takeaway**
  - The survey teams, which only exist in some counties, create key source of variation in exposure to the reform

The deployment of the survey teams

- **When were they deployed?**
  - Deployed by the NBS in about 40% counties in 2005. County coverage fixed over time
- **What is their job?**
  - Conduct sampling surveys to collect information on CPI, household income, grain output, and micro firm dynamics
  - Detect misbehaviors since 2009 as previously mentioned
- **How did the NBS select these counties?**
  - The NBS did not reveal the criteria. But as the goal of the survey teams is to generate nationally representative information on indicators like CPI, these counties should not differ too much from other counties

Relation to literature

Back

- **GDP manipulation in autocracies** (Wallace 2016; Xiong 2018; Martinez 2022)
  - Aggregate analysis; correlational; only existence of manipulation
  - This paper: county-level quasi-experimental design; monitoring ↘ manipulation
- **The role of monitoring in disciplining politicians** (Olken 2007; Ferraz and Finan 2008; Avis, Ferraz, and Finan 2018; Vannutelli 2021)
  - Settings with electoral and judicial accountability
  - This paper: politicians accountable to superiors; weak legal system ⇒ different incentives
- **Real consequences of corruption (misbehaviors)** (Giannetti et al. 2021; Colonnelli and Prem 2022; Ajzenman 2021)
  - Corruption can affect govt. policies and resource allocation
  - This paper: similar findings but focus on an overlooked and seemingly harmless aspect—GDP manipulation

Setup

Back

- Two leaders $i, -i$ competing for promotion in one-period tournament
- Each leader can stimulate GDP growth with effort e_i and convex cost $g(e_i)$, and manipulate with effort m_i and convex cost $h(m_i)$. These are the direct cost of exerting efforts and subject to $g(e_i) + h(m_i) \leq \bar{C}$
- To make manipulation secret, assume the manipulation happens simultaneously with stimulating GDP growth
- The reported GDP growth is given by

$$G_i = e_i + m_i + \varepsilon_i, \qquad \varepsilon_i - \varepsilon_{-i} \sim [-\frac{1}{2\phi}, \frac{1}{2\phi}]$$

- The manipulation is detected with prob. θp , incurring an indirect cost: the punishment λm_i. λ marginal punishment, p exogenous rate of detection, and θ institutional quality.
- The overall cost for leader i is:

$$C_i = g(e_i) + h(m_i) + 1_{\{i \text{ detected}\}} \lambda m_i$$

- Leader i receives reward R if promoted and r if not, $R \gg r > 0$:

$$U_i = \delta[1_{\{i \text{ promoted}\}} u(R) + (1 - 1_{\{i \text{ promoted}\}}) u(r)] + (1 - \delta) u(\Omega) - C_i$$

where δ captures career concerns; Ω exogenous rents from current office

Promotion rule

- The principal (upper-level govt.) mainly promotes the leader with the highest reported GDP growth (Maskin, Qian, and Xu 2000; Li and Zhou 2005)

- Upon detection of manipulation, the principal subtracts the degree of manipulation from a leader's reported GDP growth

- Formally, leader i is promoted if

$$G_i - 1_{\{i \text{ detected}\}} m_i > G_{-i} - 1_{\{-i \text{ detected}\}} m_{-i}$$

Timing and solution

- **Timing:** m_i, $e_i \longrightarrow \varepsilon_i \longrightarrow$ detection, punishment, promotion
- **Equilibrium:** pure strategy Nash equilibrium, where leader i maximizes his expected payoff taking leader $-i$'s choice as given:

$$m^* = K[\frac{\theta p(\lambda + \delta V \phi)}{\delta V \phi} - 1]$$
$$e^* = Q(m^*)$$

  where $K(\cdot)' < 0, Q(\cdot)' < 0$. $V = [u(R) - u(r)]$: utility gains from promotion

- **Proposition 1:** Equilibrium GDP manipulation m^* is decreasing in institutional quality θ, prob. of detection p, marginal punishment λ, and increasing in promotion gains V, career concerns δ, and noise density ϕ
- **Proposition 2:** Compared to m^*, equilibrium effort in stimulating the economy changes in opposite directions w.r.t. the parameters

Testable predictions on the effect of the reform

- **GDP manipulation:** Treated counties (T) should see a larger drop in GDP manipulation than control counties (C):

$$\Delta m_T - \Delta m_c = \underbrace{(m'_T - m_T)}_{<0} - \underbrace{(m'_C - m_C)}_{=0} < 0$$

- **Effort in stimulating the economy:** Treated counties should see a larger increase in effort in stimulating the economy than control counties:

$$\Delta e_T - \Delta e_c = \underbrace{(e'_T - e_T)}_{>0} - \underbrace{(e'_C - e_C)}_{=0} > 0$$

- **Real economic growth:** Let $y = e + \varepsilon$. The effect of the reform on real economic growth is indeterminate:

$$\Delta y_T - \Delta y_c = \underbrace{(\Delta e_T - \Delta e_c)}_{>0} + \underbrace{(\Delta \varepsilon_T - \Delta \varepsilon_C)}_{\gtreqless 0} \gtreqless 0$$

Testable predictions on the effect of the reform (cont.)

- **Heterogeneity:** The magnitude of the treatment effect on GDP manipulation $|\Delta m_T - \Delta m_c|$ is decreasing in career concern δ but increasing in institutional quality θ

$$\frac{\partial |\Delta m_T - \Delta m_c|}{\partial \delta} < 0, \quad \frac{\partial |\Delta m_T - \Delta m_c|}{\partial \theta} > 0$$

Identification concerns

Back

$$\underbrace{z_{ct}}_{\text{Reported GDP growth}} = \frac{1}{\gamma}\underbrace{l_{ct}}_{\text{Light growth}} + \beta\, Treat_c \times Post_t + \delta_c + \lambda_t + \varepsilon^z_{ct}$$

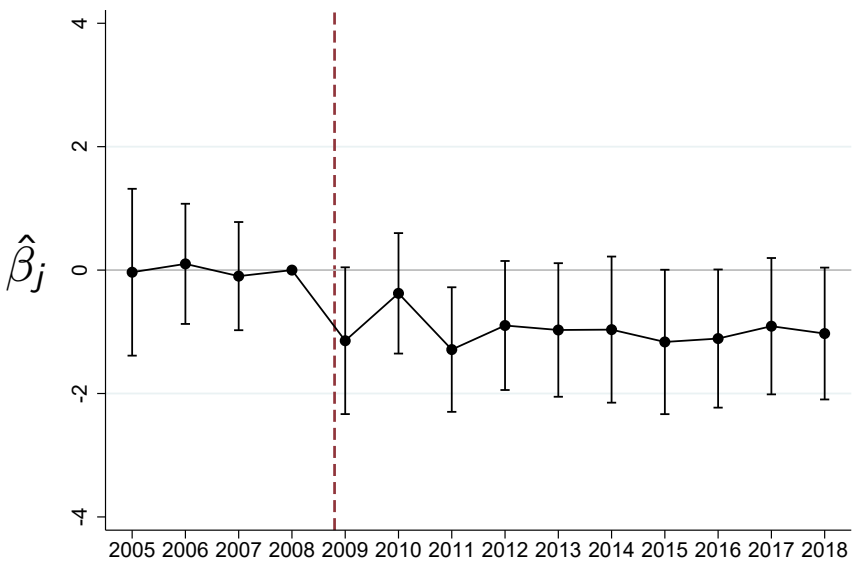
- Identifying assumption: z_{ct}, after adjusting for l_{ct}, evolve in parallel between treated and control counties in the absence of the reform
- Event study can lend support but unable to fully verify it
- Two remaining concerns:
  1. γ, relationship b/w light and economic activities, may differ across counties or years
  2. (Unobserved) heterogeneity between treated and control counties
- Solution to concern 1:
  - Allow the effect of light growth (γ) to vary by a large number of county characteristics and also by years
- Solution to concern 2:
  - Balance the covariates through entropy balancing (Hainmueller 2012)
  - Include county-specific controls or time trends
  - Placebo event study around 2005 (teams launched with no monitoring)
  - Instrumental variable for $Treat_c$

Decomposing the effect

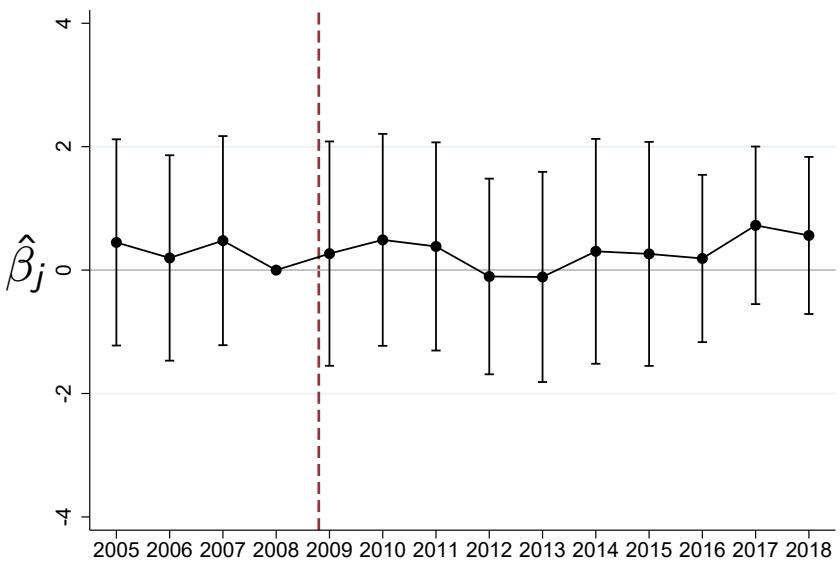
Back

$$Y_{ct} = \sum_{j=2005,\ j\neq 2008}^{j=2018} \beta_j \, Treat_c \times 1_{\{t=j\}} + \delta_c + \lambda_t + \varepsilon_{ct}$$

Y: Reported GDP growth, %



Y: Light growth, %

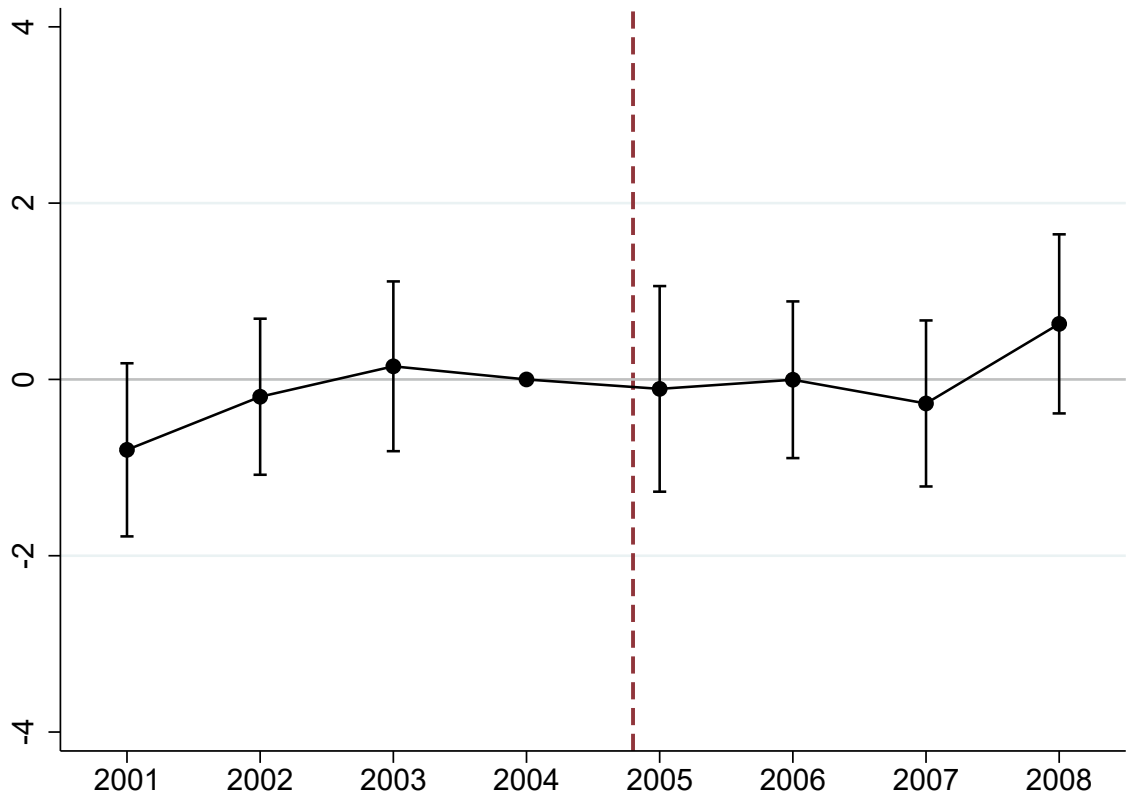


Placebo reform: the launch of survey teams in 2005

Back

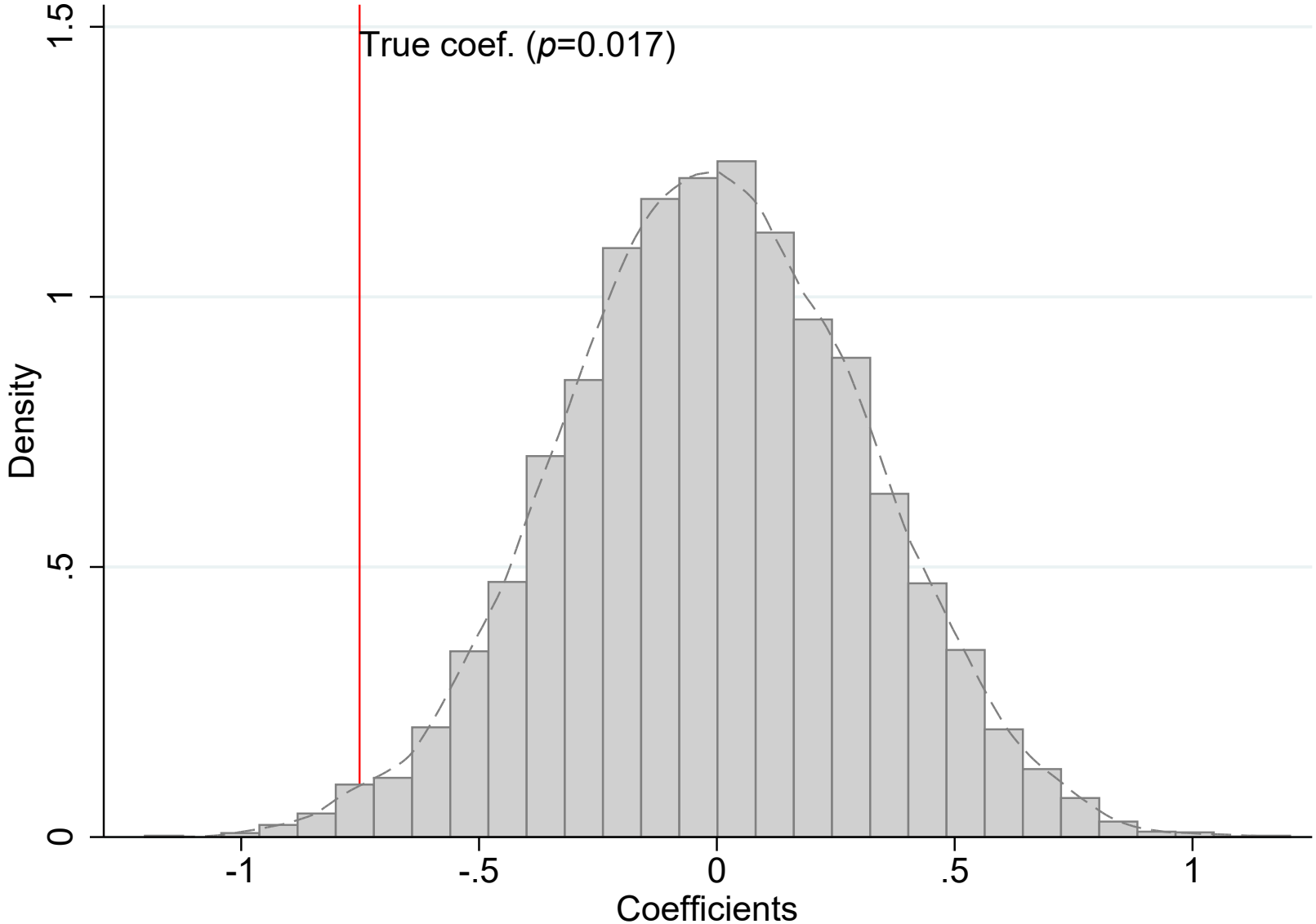
$$ReportedGDPGrowth_{ct} = \alpha LightGrowth_{ct} + \sum_{j=2001,\ j\neq 2004}^{j=2008} \beta_j Treat_c \times 1_{\{t=j\}} + \delta_c + \lambda_t + \varepsilon_{ct}$$

$$\text{s.t. } 2001 \leq t \leq 2008$$



Randomization inference

Back



Estimating spillover effect

Back

$$z_{ct} = \alpha l_{ct} + \beta^{Direct} Treat_c \times Post_t + \beta^{Spillover} Spillover_c^{Neighbors \to c} \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct}$$

	(1)	(2)	(3)	(4)
Dep. var.	Reported GDP growth (%)			
Treat x Post	-0.576***	-0.580***	-0.565***	-0.585***
	(0.161)	(0.164)	(0.164)	(0.162)
#Treated neighbors x Post		-0.010		
		(0.071)		
1(#Treated neighbors>0) x Post			0.108	
			(0.263)	
1(#Treated neighbors>Median=2) x Post				-0.133
				(0.178)
Light growth (%)	0.017***	0.017***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
County controls x Post	Yes	Yes	Yes	Yes
Neighbor number FE x Post	No	Yes	Yes	Yes
Cluster level	County	County	County	County
Mean dep. var.	10.84	10.84	10.84	10.84
Mean number of neighbors	5.91	5.91	5.91	5.91
Mean number of treated neighbors	1.97	1.97	1.97	1.97

Estimating spillover effect: robustness

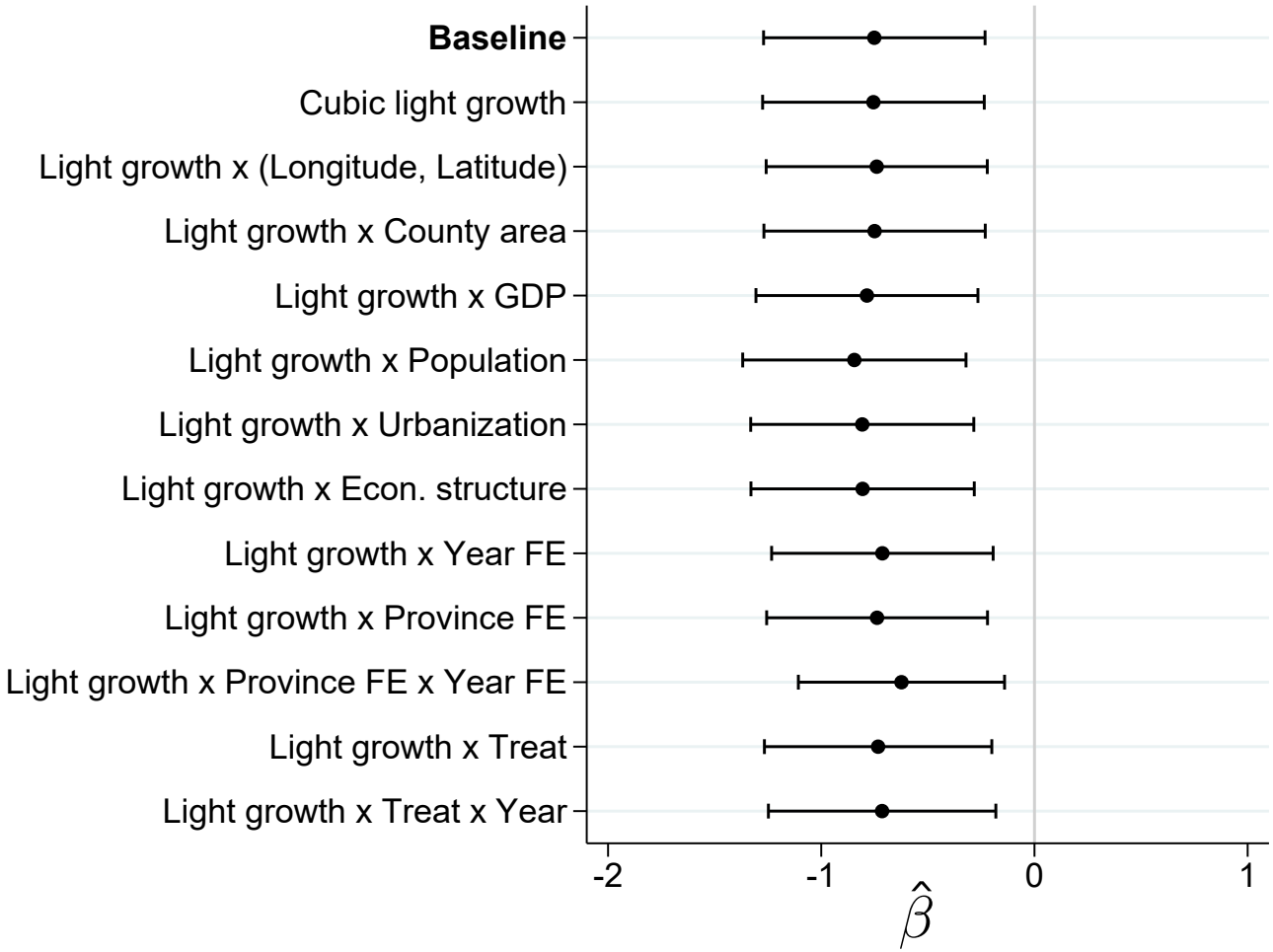
$$z_{ct} = \alpha l_{ct} + \beta^{Direct} Treat_c \times Post_t + \beta^{Spillover} Spillover_c^{Neighbors \to c} \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct}$$

	(1)	(2)	(3)	(4)
Dep. var.	Reported GDP growth (%)			
Treat x Post	-0.610***	-0.559***	-0.581***	-0.576***
	(0.162)	(0.164)	(0.164)	(0.164)
#Treated neighbors within 50km x Post	-0.094			
	(0.091)			
#Treated neighbors within 100km x Post		0.027		
		(0.036)		
#Treated neighbors (GDP-weighted) x Post			-0.011	
			(0.068)	
#Treated neighbors (population-weighted) x Post				0.003
				(0.069)
Light growth (%)	0.017***	0.017***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
County controls x Post	Yes	Yes	Yes	Yes
Neighbor number FE x Post	Yes	Yes	Yes	Yes
Cluster level	County	County	County	County
Mean dep. var.	10.84	10.84	10.84	10.84
Mean number of neighbors	2.40	11.45	5.91	5.91
Mean number of treated neighbors	0.88	4.47	1.97	1.97

Flexible mapping b/w light growth and economic growth

Back

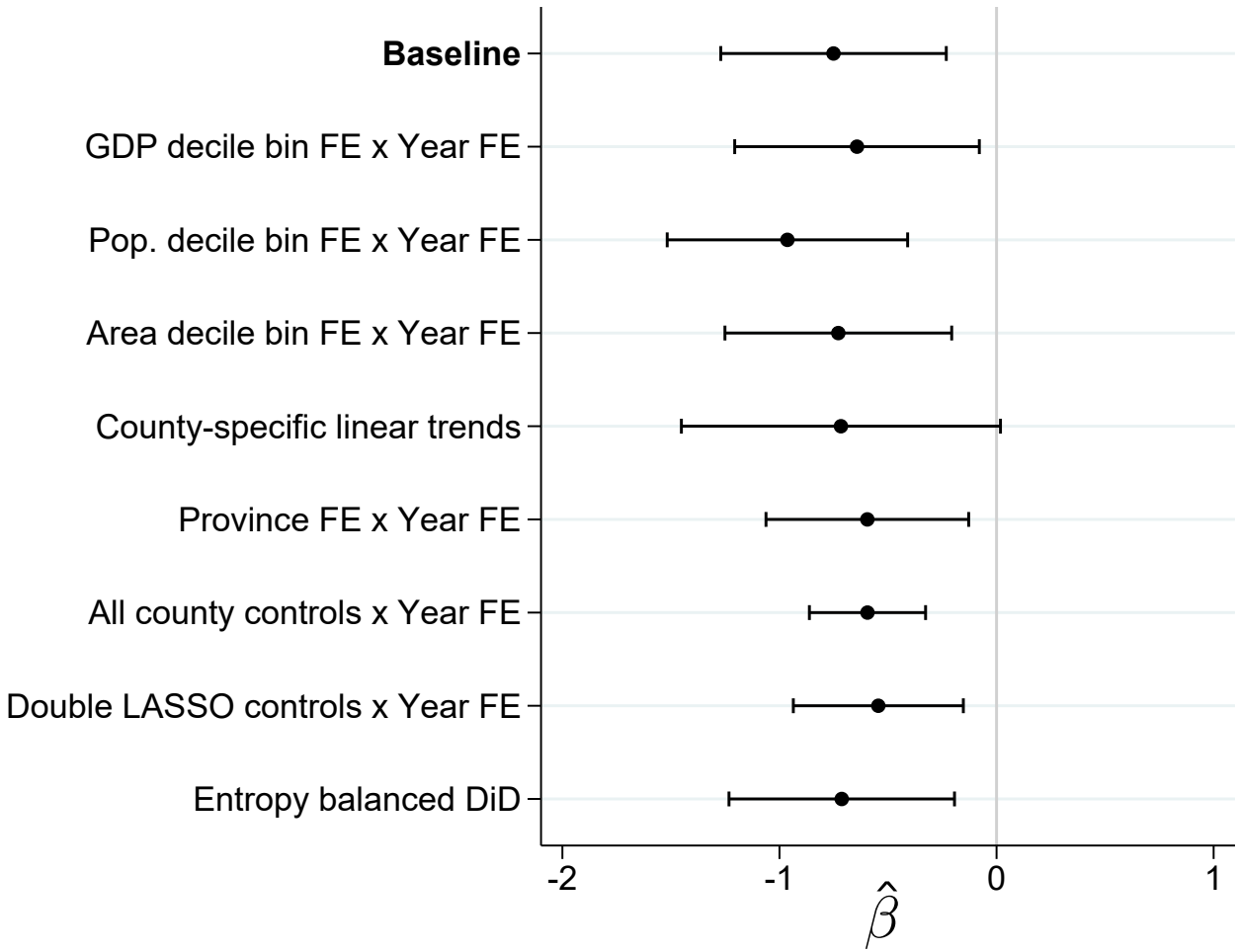
$$z_{ct} = \alpha l_{ct} + f(l_{ct}) + \beta Treat_c \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct}$$



Addressing unbalancedness b/w treated and control groups

Back

$$z_{ct} = \alpha I_{ct} + \beta \mathit{Treat}_c \times \mathit{Post}_t + \delta_c + \lambda_t + \text{Alter. FEs} + \varepsilon_{ct}$$



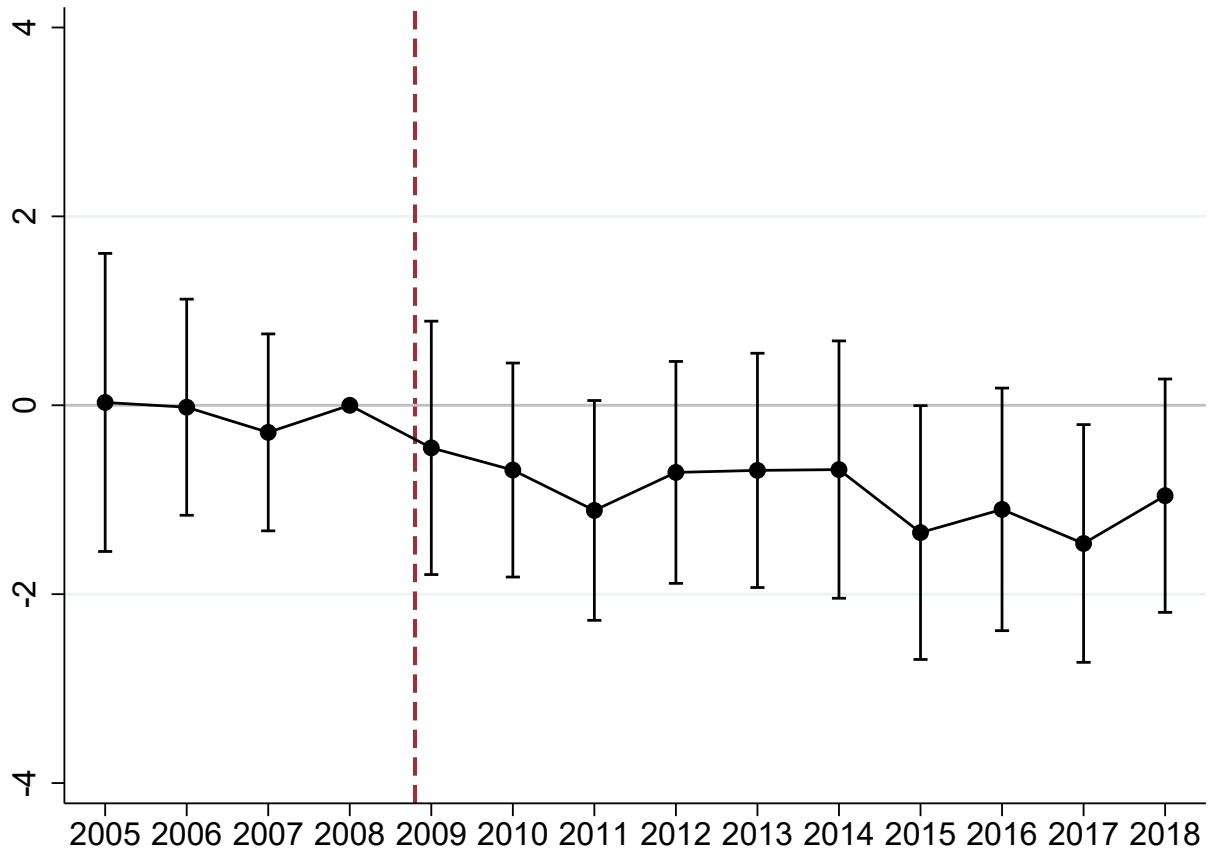
IV balance test [2] Back

	W/ rural survey teams		W/O rural survey teams		Difference		
	Mean	SD	Mean	SD	Diff	SE	*p*-value
Panel A: Demography							
log Population (2010)	12.73	0.82	12.69	0.80	0.05	0.04	**0.25**
Share urban (%, 2010)	34.09	12.69	34.33	14.00	-0.24	0.68	**0.72**
Share 15-64 (%, 2010)	72.62	4.35	72.40	4.65	0.22	0.23	**0.32**
Years of schooling (2010)	8.24	0.82	8.19	0.99	0.05	0.05	**0.30**
Panel B: Economic development							
Unemployment rate (%, 2010)	2.13	1.38	2.16	1.51	-0.03	0.07	**0.64**
Share primary sectors (%, 2010)	63.98	17.66	64.89	18.24	-0.91	0.90	**0.31**
Share secondary sectors (%, 2010)	15.92	11.90	15.64	12.14	0.28	0.60	**0.64**
log GDP (2004)	12.20	0.99	12.13	1.04	0.07	0.05	**0.21**
log GDP (2008)	12.71	1.01	12.63	1.08	0.07	0.05	**0.16**
GDP growth (%, 2002-2004 average)	11.26	6.32	11.54	6.50	-0.28	0.34	**0.42**
GDP growth (%, 2006-2008 average)	12.96	6.56	12.62	6.51	0.34	0.33	**0.29**
Light growth (%, 2002-2004 average)	18.67	14.60	18.67	15.04	0.00	0.74	**1.00**
Light growth (%, 2006-2008 average)	7.24	10.92	7.48	11.60	-0.24	0.57	**0.67**
Distance to major roads (km, 2010)	74.78	99.81	72.54	80.07	2.23	4.40	**0.61**
Distance to major railways (km, 2010)	77.02	105.61	71.13	90.01	5.89	4.80	**0.22**
Panel C: Geography							
County area (km^2)	3883	7003	4160	10298	-277	463	**0.55**
Precipitation (inches, 2004)	0.030	0.083	0.035	0.095	-0.005	0.005	**0.27**
Temperature (degrees, 2004)	13.83	5.22	13.60	5.41	0.23	0.27	**0.39**
Precipitation (inches, 2008)	0.04	0.10	0.05	0.12	-0.01	0.01	**0.14**
Temperature (degrees, 2008)	13.66	5.07	13.45	5.26	0.21	0.26	**0.43**
Distance to major rivers (km)	58.51	59.50	58.40	60.42	0.11	3.00	**0.97**
Distance to country border (km)	340.46	252.42	348.81	250.34	-8.36	12.55	**0.51**
Distance to coastline (km)	632.09	618.89	631.84	571.17	0.25	29.48	**0.99**
Distance to prefecture center (km)	62.20	44.79	61.42	44.92	0.78	2.25	**0.73**

[2] Except for GDP, all variables in this table are from sources that the county has no control on. The reason for using the year 2010 for the demographic data, instead of years before the reform, is that the population census was only conducted in 2000 and 2010. Similarly, the transportation data is also in 2010 due to data limitation

Instrumented event study

Back study



Key mechanism: disciplining effect of the reform

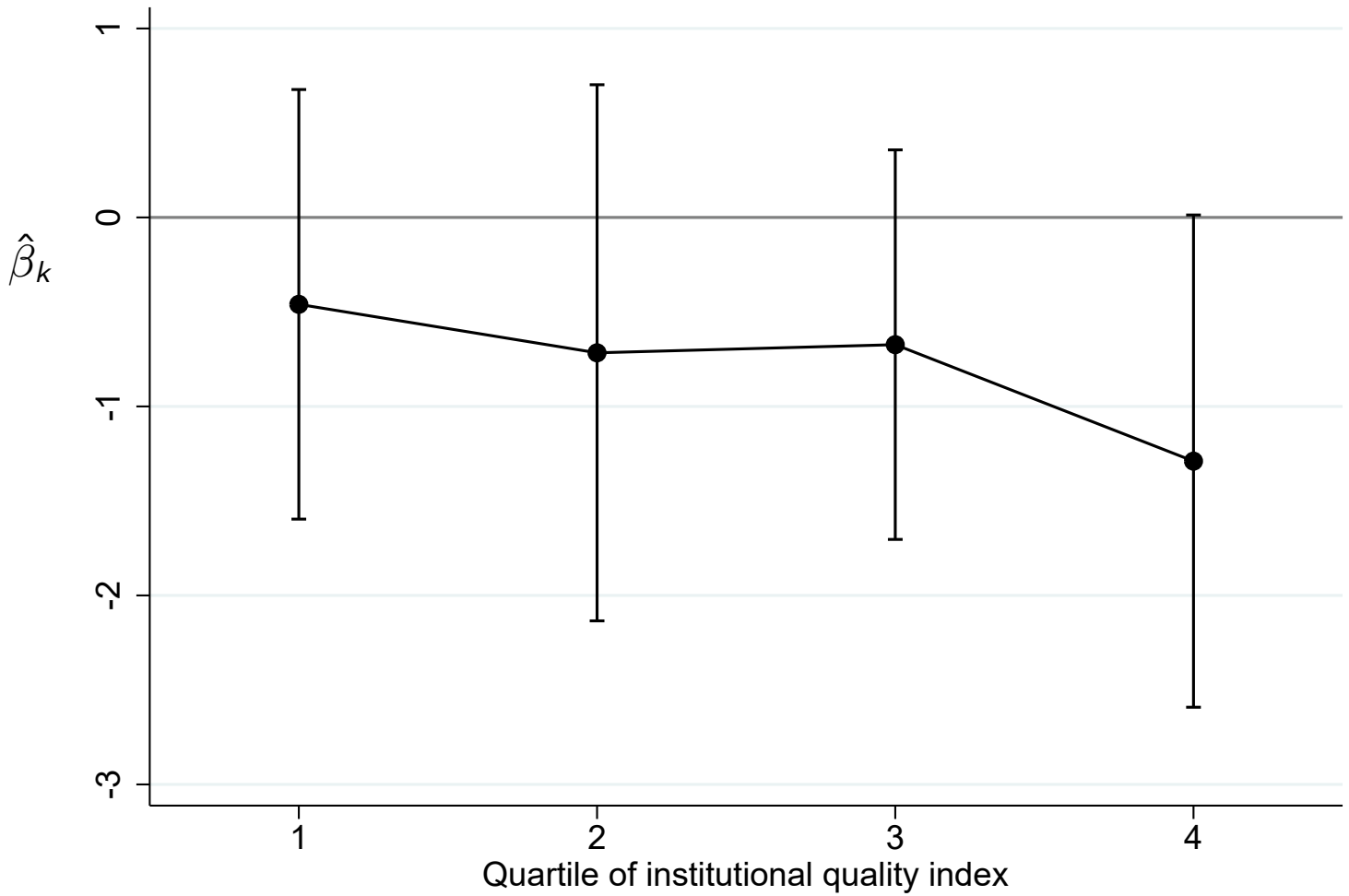
Back

	(1)	(2)	(3)
	Baseline	With leader FE	No turnover
Dep. var.	Reported GDP growth (%)		
Treat x Post	-0.751**	-0.797*	-0.789**
	(0.316)	(0.417)	(0.389)
Light growth (%)	0.023***	0.016***	0.017***
	(0.005)	(0.005)	(0.006)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Leader FE	No	Yes	No
Cluster level	County	County	County
Observations	23,360	21,101	15,021
R-squared	0.269	0.343	0.297
Mean of dep. var.	10.97	10.87	11.50

Notes: Col 3 focuses on a trimmed sample with only leaders whose terms straddle 2009

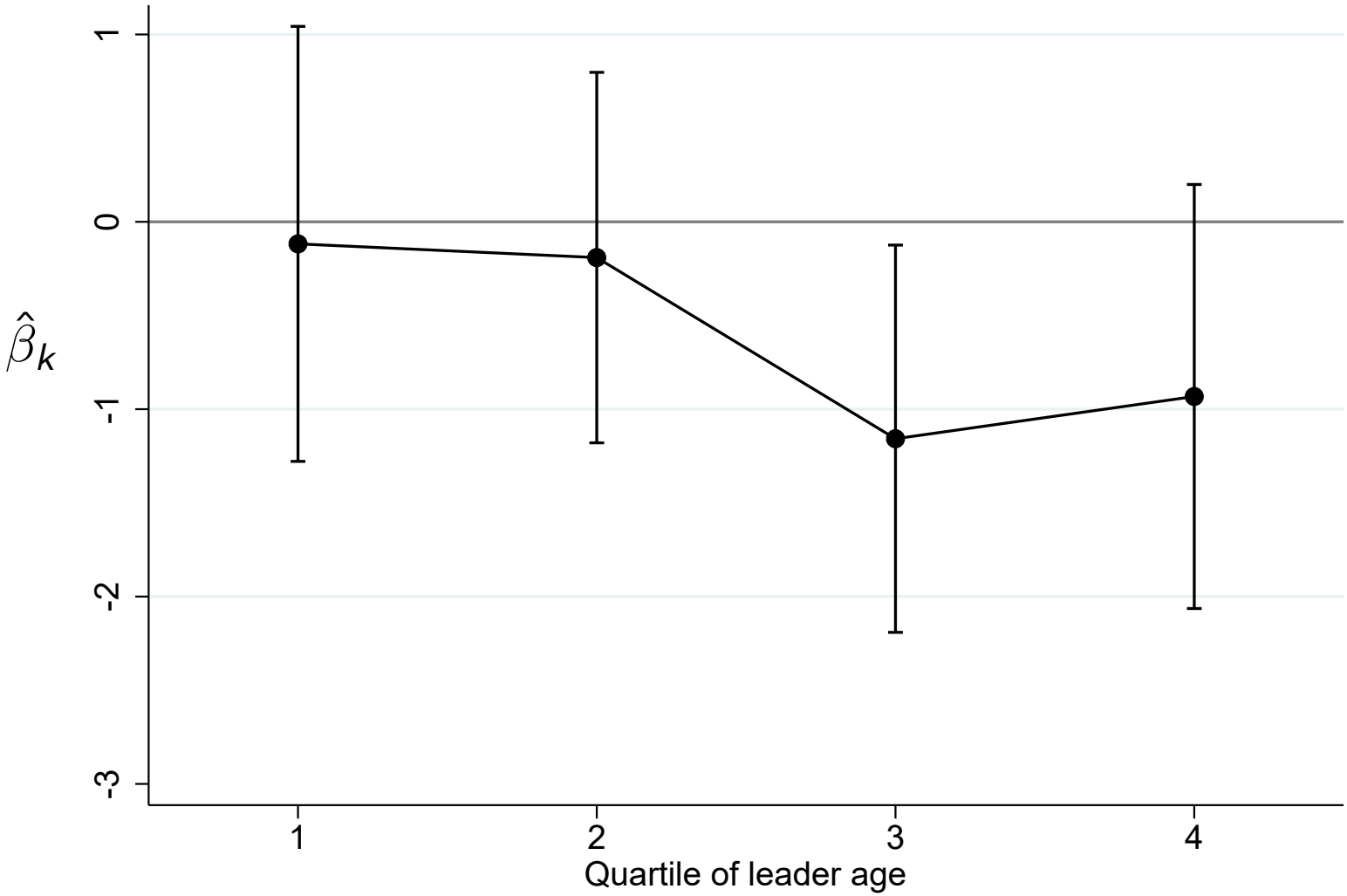
Heterogeneity by institutional quality: $\frac{\partial|\Delta m_T - \Delta m_c|}{\partial\theta} > 0$

- Provincial level index describing market intermediary and the legal environment (Fan, Wang, and Zhu 2011)



Low institutional quality High institutional quality

Heterogeneity by career concern: $\frac{\partial|\Delta m_T - \Delta m_c|}{\partial \delta} < 0$



Alternative explanation #1: better political selection

Dep. var.	(1) 1(Turnover)	(2) Years in office	(3) Age	(4) 1(Local)	(5) Schooling	(6) 1(<Master)	(7) 1(Master)	(8) 1(PhD)
Treat x Post	-0.001	0.113	-0.075	-0.009	-0.157	0.006	0.003	-0.009
	(0.006)	(0.163)	(0.294)	(0.016)	(0.104)	(0.023)	(0.024)	(0.012)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	County	County	County	County	County	County	County	County
Observations	22,695	22,695	10,674	8,243	7,994	7,994	7,994	7,994
R-squared	0.151	0.879	0.769	0.857	0.829	0.826	0.817	0.736
Mean dep. var.	0.05	8.07	47.68	0.16	16.40	0.57	0.39	0.04

Alternative explanation #1: better political selection(cont)

Dep. var.	(1) Reported GDP growth (%)	(2) Reported GDP growth (%)	(3) Reported GDP growth (%)
Treat x Post	-0.751**	-0.755**	-0.731**
	(0.316)	(0.315)	(0.315)
Treat x Post x Post-reform turnover number		0.192	-0.204
		(0.246)	(0.319)
Light growth (%)	0.023***	0.023***	0.022***
	(0.005)	(0.005)	(0.005)
Observations	23,360	23,360	23,346
R-squared	0.269	0.269	0.276
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Post-reform turnover number FE x Year FE	No	No	Yes
Cluster level	County	County	County
Mean of dep. var.	10.97	10.97	10.98
Mean of post-reform turnover number	0.50	0.50	0.49

Alternative explanation #2: soft information

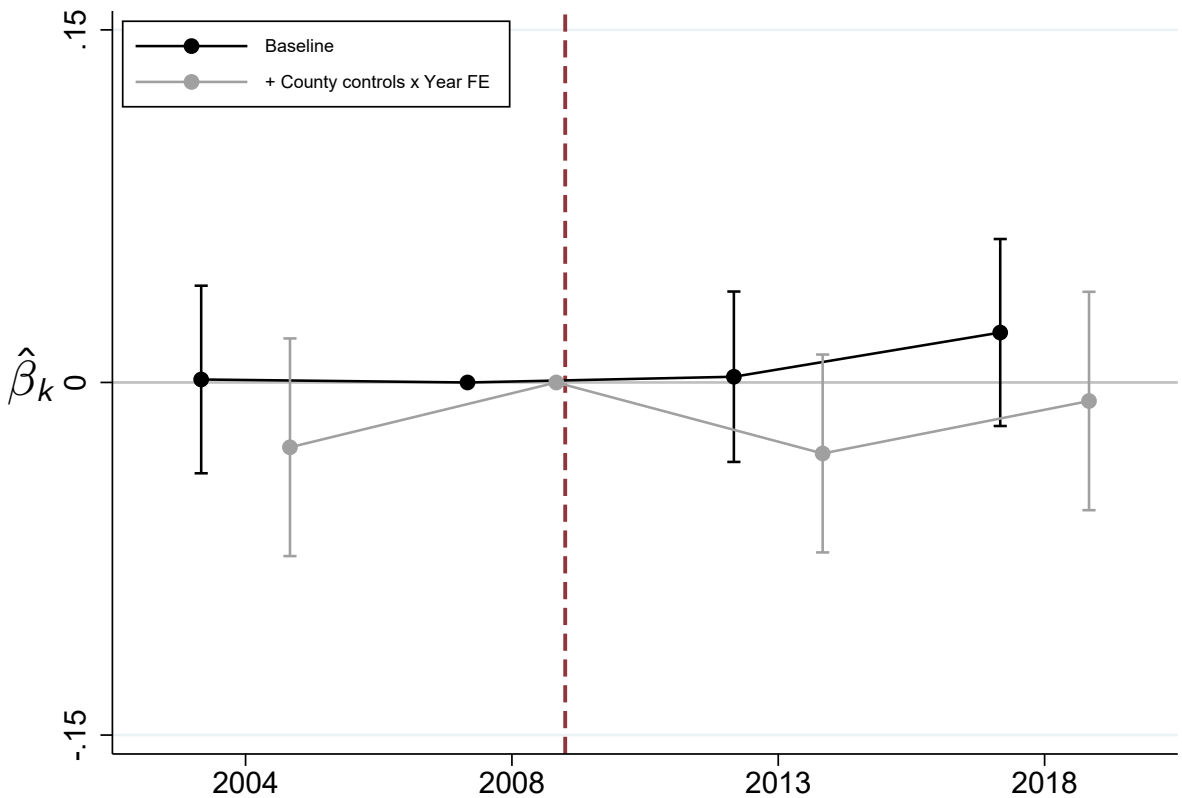
- Arrival of survey teams → better soft information → less weight of GDP in promotion → less manipulation

Dep. var.	(1) Reported GDP growth (%)	(2) Reported GDP growth (%)	(3) Reported GDP growth (%)
Treat x Post	-0.751**	-0.749**	-0.751**
	(0.316)	(0.317)	(0.317)
Treat x Post x Distance to upper-level govt.		-0.138	
		(0.292)	
Treat x Post x 1(No connection → connection)			0.082
			(3.117)
Light growth (%)	0.023***	0.023***	0.023***
	(0.005)	(0.005)	(0.005)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Distance to upper-level govt. x Year FE	No	Yes	No
1(No connection → connection) x Year FE	No	No	Yes
Cluster level	County	County	County
Observations	23,360	23,239	23,360
R-squared	0.269	0.269	0.269
Mean of dep. var.	10.97	10.97	10.97

Alter. exp. #3: improvement in local statistical capacity

- Measured by county performance in coordinating and conducting centrally launched economic census in census years 2004, 2008, 2013, 2018
- Counties with excellent performance will receive a reputational award

$$Award_{ct} = \sum_{k=2004,2013,2018} \beta_k Treat_c \times 1_{\{t=k\}} + \delta_c + \lambda_g + \varepsilon_{ct}$$



Alter. exp. #4: concurrent reforms

- Other concurrent reforms may also increase monitoring of county govt.:
  - Fiscal PMC: province manages county's fiscal revenue
  - Full PMC: province manages county's all affairs
  - Anticorruption inspection since 2013

	(1)	(2)	(3)	(4)	(5)
	Baseline	PMC reforms		Anti-corruption	Both
Dep. var.	Reported GDP growth (%)				
Treat x Post	-0.751**	-0.814***	-0.744**	-0.752**	-0.806**
	(0.316)	(0.313)	(0.316)	(0.316)	(0.313)
Light growth (%)	0.023***	0.021***	0.023***	0.023***	0.022***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Fiscal PMC	No	Yes	No	No	Yes
Full PMC	No	No	Yes	No	Yes
Anticorruption inspection	No	No	No	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Cluster level	County	County	County	County	County
Observations	23,360	23,360	23,360	23,360	23,360
R-squared	0.269	0.271	0.269	0.269	0.271
Mean of dep. var.	10.97	10.97	10.97	10.97	10.97

Effect on bank loans and deposits

Back

Top panel: all banks. Bottom panel: state banks

	(1)	(2)	(3)	(4)
			IHS transformation of	
Dep. var.:	Deposits	Loans	Loans to small firms	Number of firms granted loans
Treat x Post	0.158	0.174	0.263**	0.123
	(0.129)	(0.123)	(0.116)	(0.076)
County controls x Post	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Cluster level	County	County	County	County
Observations	8,922	8,922	8,922	8,922
R-squared	0.311	0.329	0.424	0.520

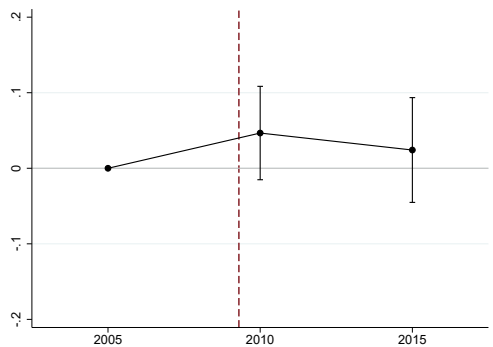
	(1)	(2)	(3)	(4)
			IHS transformation of	
Dep. var.:	Deposits	Loans	Loans to small firms	Number of firms granted loans
Treat x Post	0.203*	0.205*	0.292**	0.167**
	(0.123)	(0.119)	(0.126)	(0.085)
County controls x Post	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Cluster level	County	County	County	County
Observations	8,922	8,922	8,922	8,922
R-squared	0.323	0.328	0.512	0.574

Firm entry: IV estimation

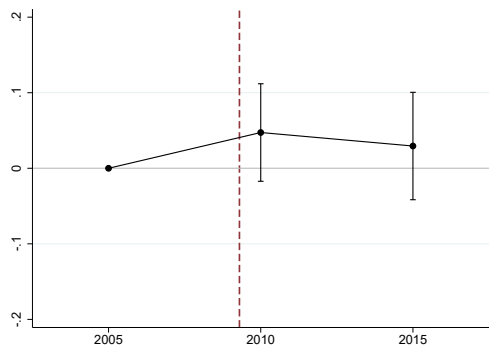
Back

- Firm entry measured using the universe of firm registration data in 2005, 2010, and 2015 (only 3 years due to data limitation)

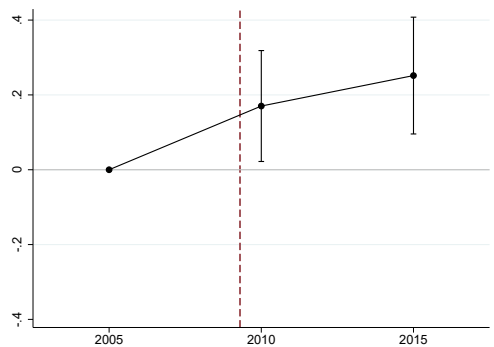
IHS(#All firms)



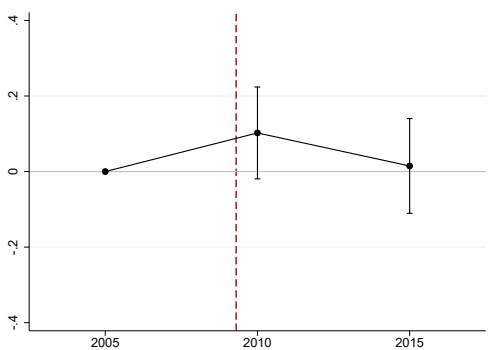
IHS(#Private)



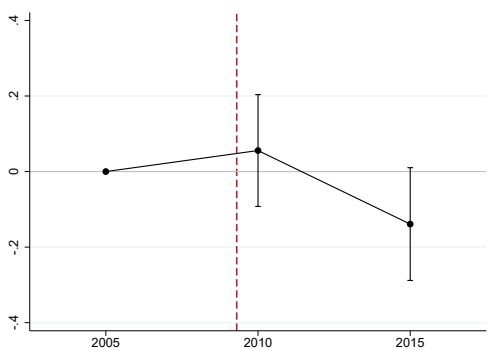
IHS(#SOEs)



IHS(#Foreign)



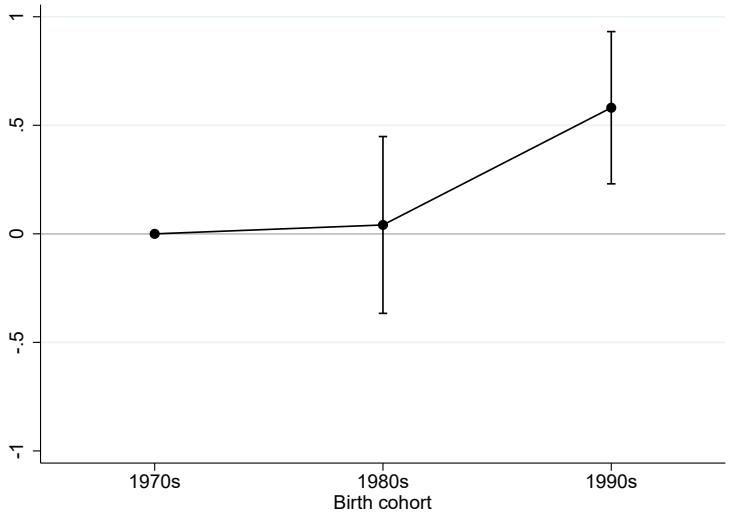
IHS(#Collective)



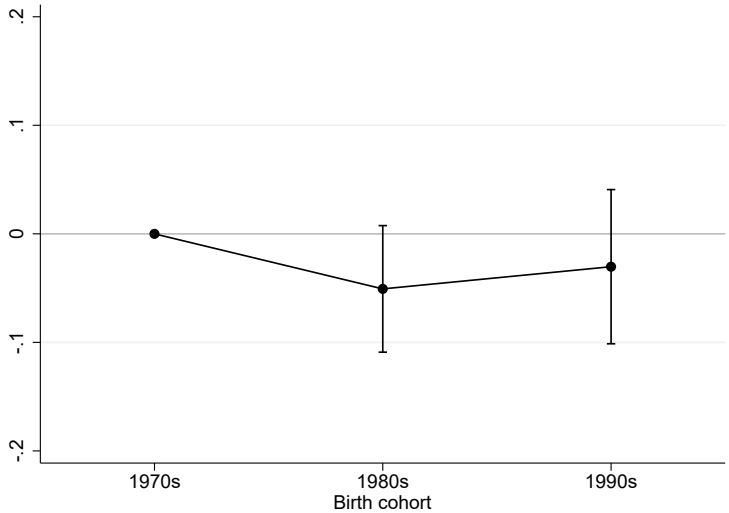
Citizen attitudes: IV estimation

Back

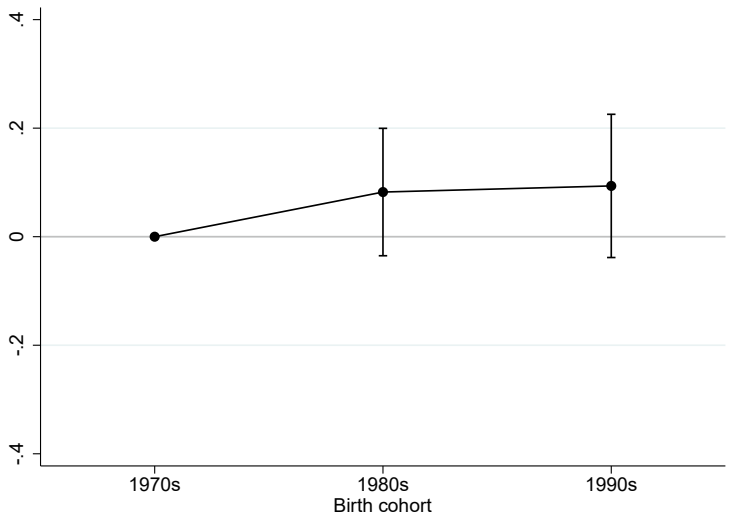
Trust in local officials
(mean=4.8)



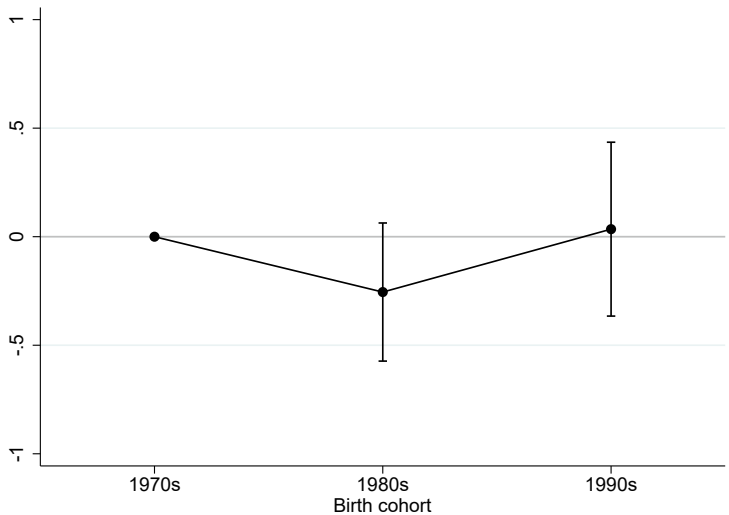
Trust in most people (0/1)



Evaluation of govt. performance
(mean=3.4)



Perceived corruption



Effect on corruption convictions - 2SLS estimates

Back

Dep. var.:	(1)	(2)	(3)	(4)	(5)
	IHS (Number of convictions)				
	Panel A: Bribery				
Treat	0.055	0.022	-0.006	0.020	-0.008
	(0.071)	(0.066)	(0.061)	(0.066)	(0.061)
#Anti-corruption inspections				-0.023*	-0.015
				(0.013)	(0.013)
Mean of dep. var.	2.603	2.718	2.718	2.718	2.718
SD of dep. var.	3.287	3.304	3.304	3.304	3.304
	Panel B: Appropriation				
Treat	0.041	0.039	0.020	0.039	0.020
	(0.052)	(0.054)	(0.054)	(0.054)	(0.054)
#Anti-corruption inspections				0.001	-0.006
				(0.011)	(0.012)
Mean of dep. var.	0.990	0.974	0.974	0.974	0.974
SD of dep. var.	1.494	1.443	1.443	1.443	1.443
	Panel C: Bribery ∪ Appropriation				
Treat	0.075	0.048	0.020	0.047	0.018
	(0.070)	(0.066)	(0.062)	(0.066)	(0.062)
#Anti-corruption inspections				-0.018	-0.018
				(0.014)	(0.014)
Mean of dep. var.	3.191	3.303	3.303	3.303	3.303
SD of dep. var.	3.594	3.604	3.604	3.604	3.604
County controls	No	Yes	Yes	Yes	Yes
Province FE	No	No	Yes	No	Yes
Observations	1,752	1,498	1,498	1,498	1,498
F-stat of excl. inst.	1942	2041	2214	2041	2204