ONLINE APPENDICES FOR "FIRM GROWTH AND CORRUPTION"

- Appendix A: Additional tables
- Appendix B: Validation of matching between PCI and GSO
- Appendix C: Model
- Appendix D: Description of key variables
- Appendix E: Testing alternative models

Appendix A: Tables

Appendix Table 1: Cross-Validation of PCI and GSO Data

	Firms with ≥ 10 lagged employees		All firms	
	Log VN mean in PCI	Log VN median in PCI	Log VN mean in PCI	Log VN median in PCI
	(1)	(2)	(3)	(4)
Log VN mean in GSO	0.520***		0.528***	
	(0.040)		(0.033)	
Log VN median in GSO		0.474***		0.440***
		(0.045)		(0.041)
Observations	4484	4484	5760	5760
Province–industry and year FE	✓	✓	✓	✓

Note: Each observation is a province-industry-year. The dependent variable is log mean or log median firm-level employment for each province-industry-year group in the PCI data. The independent variable is the corresponding variable computed using the GSO data. For both datasets, we keep only private firms. Since the firm employment variable in PCI is categorical, we compute the empirical mean and median for each category from the GSO data and apply these to the PCI data. All specifications control for province by industry (broad level) and year fixed effect. Robust standard errors are clustered at the province level and reported. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 2: Industry Codes and Descriptions

ISIC Rev 4 Code	Description
A	Agriculture, forestry and fishing
B	Mining and quarrying
C	Manufacturing
D	Electricity, gas, steam and air conditioning supply
E	Water supply; sewerage, waste management and remediation activities
F	Construction
G	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	Transportation and storage
I	Accommodation and food service activities
J	Information and communication
K	Financial and insurance activities
L	Real estate activities
M	Professional, scientific and technical activities
N	Administrative and support service activities
P	Education
Q	Human health and social work activities
R	Arts, entertainment and recreation
S	Other service activities

The alphabetical industry codes and descriptions are based on International Standard Industrial Classification (ISIC) of All Economic Activities, Rev.4. The list includes the 18 industries that appear in our sample.

Appendix Table 3: First Stage Results: Multiple IV

	Dep. var.: Log Vietnamese employment in industry-year (in own province)	
	Rest-of-Vietnam IV	China IV
Interaction coefficient for Agriculture	0.0532	-0.800
	(0.338)	(0.931)
Interaction coefficient for Mining and Quarrying	-1.343*	-4.052
	(0.690)	(3.398)
Interaction coefficient for Manufacturing	-0.474	-0.488
	(0.353)	(1.023)
Interaction coefficient for Electricity and Gas Supply	-0.107	0.0971
	(0.175)	(8.336)
Interaction coefficient for Water Supply	0.169	1.883
	(0.211)	(1.384)
Interaction coefficient for Construction	0.183	0.574
	(0.220)	(0.571)
Interaction coefficient for Wholesale and Retail Trade	0.216	0.611
	(0.168)	(0.415)
Interaction coefficient for Transportation and Storage	0.379*	3.180*
	(0.207)	(1.776)
Interaction coefficient for Accommodation and Food Service	0.170	0.653
	(0.174)	(0.542)
Interaction coefficient for Information and Communication	0.282*	1.303
	(0.152)	(0.868)
Interaction coefficient for Financial and Insurance Activities	0.269	1.372*
	(0.202)	(0.771)
Interaction coefficient for Real Estate Activities	0.759***	3.357***
	(0.195)	(0.992)
Interaction coefficient for Professional and Scientific Activities	0.396***	2.014**
	(0.153)	(0.872)
Interaction coefficient for Administration and Support Activities	0.625***	1.576***
	(0.176)	(0.508)
Interaction coefficient for Education	0.446***	12.37***
	(0.165)	(4.275)
Interaction coefficient for Human Health and Social Work	0.498***	3.714***
	(0.146)	(1.001)
Interaction coefficient for Arts, Entertainment and Recreation	0.166	4.249
	(0.186)	(3.937)
Interaction coefficient for Other Service Activities	-0.187	0.0148
	(0.226)	(0.522)
Observations	3367	3367
F-stats	7.986	8.369
Province–industry and year fixed effects	✓	✓

Note: Each observation is a province-industry-year. The dependent variable is log employment in industry-year in own province. The independent variable is the rest-of-Vietnam IV and China IV interacted with industry dummies. Industry refers to an ISIC alphabetical industry code. We include the interaction term for all 18 industries. The regression controls for province-industry and year fixed effects. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1. The F-test statistic for province level clustering is reported in this table and the corresponding test statistics for industry-year level clustering is higher for both the rest-of-Vietnam IV and China IV.

Appendix Table 4: Effect of Economic Performance on Bribes (All Firms)

Dependent variable: Firm's bribe payment as percentage of revenue

	Single IV	Multiple IV	RF: OLS	RF: Ordered Probit
	(1)	(2)	(3)	(4)
Panel A. Rest-of-Vietnam IV				
Log Vietnamese	-1.890**	-1.716**		
employment in industry-year (in own province)	(0.770)	(0.870)		
Log Vietnamese employment			-1.389**	-0.183**
in industry-year (excluding own province)			(0.619)	(0.0837)
Panel B. China IV				
Log Vietnamese	-2.017	-1.453**		
employment in industry-year (in own province)	(1.285)	(0.709)		
Log Chinese			-1.454	-0.221
employment in industry-year			(0.909)	(0.148)
Province–industry and year fixed effects	✓	✓	✓	✓

Note: This table repeats the same regression analysis in Table 3 on the full sample of PCI firms (without size restriction). The corresponding RHS variable and IV are also constructed on the full sample firms in the GSO survey (aggregated to province-industry-year). *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 5: Heterogeneous Effects Based on Firms' Property Rights (All Firms)

Dependent variable: Firm's bribe payment as percentage of revenue

	Single IV				Multiple IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Rest-of-Vietnam IV								
Log Vietnamese employment in industry-year (in own province)	-2.084**	0.497	-2.242**	0.464	-1.527*	-0.110	-1.474*	-0.186
	(0.922)	(2.526)	(0.977)	(2.622)	(0.853)	(1.914)	(0.788)	(1.919)
Firm owns its land			-0.442**	-0.457**			-0.410***	-0.410***
X log Vietnamese employment in industry year (in own province)			(0.188)	(0.193)			(0.130)	(0.130)
Firm owns land and has LURC	-0.326***	-0.349***	-0.0709	-0.0865	-0.301***	-0.304***	-0.0821	-0.0853
X log Vietnamese employment in industry year (in own province)	(0.118)	(0.129)	(0.132)	(0.137)	(0.0813)	(0.0829)	(0.120)	(0.121)
Observations	9,588	9,588	9,588	9,588	9,588	9,588	9,588	9,588
Panel B. China IV								
Log Vietnamese employment in industry-year (in own province)	-1.841	0.971	-1.487	0.606	-1.372*	0.546	-1.367**	0.505
	(1.667)	(3.758)	(1.554)	(3.106)	(0.721)	(1.805)	(0.681)	(1.798)
Firm owns its land			-0.602***	-0.661***			-0.295**	-0.295**
X log Vietnamese employment in industry year (in own province)			(0.170)	(0.216)			(0.126)	(0.125)
Firm owns land and has LURC	-0.412***	-0.449**	-0.0739	-0.0667	-0.300***	-0.308***	-0.131	-0.138
X log Vietnamese employment in industry year (in own province)	(0.129)	(0.188)	(0.151)	(0.145)	(0.0977)	(0.101)	(0.118)	(0.120)
Observations	9,588	9,588	9,588	9,588	9,588	9,588	9,588	9,588
Control for average firm size X log Vietnamese employment in industry-year (in own province)	NO	YES	NO	YES	NO	YES	NO	YES
Province-industry and year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Note: This table repeats the same regression analysis in Table 4 on the full sample of PCI firms (without size restriction). The corresponding RHS variable and IV are also constructed on the full sample firms in the GSO survey (aggregated to province-industry-year). *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 6: Heterogeneous Effects Based on Firms' Operation Locations (All Firms)

Dependent variable: Firm's bribe payment as percentage of revenue

	Single IV		Multiple IV	
	(1)	(2)	(3)	(4)
Panel A. Rest-of-Vietnam IV				
Log Vietnamese employment in industry-year (in own province)	-1.658**	0.0289	-1.536*	-0.248
	(0.754)	(1.911)	(0.824)	(1.615)
Firm currently operates in more than one province	-0.295***	-0.297***	-0.246**	-0.251**
X log Vietnamese employment in industry-year (in own province)	(0.103)	(0.101)	(0.104)	(0.105)
Observations	10,963	10,963	10,963	10,963
Panel B. China IV				
Log Vietnamese employment in industry-year (in own province)	-1.654	0.554	-1.196*	0.289
	(1.262)	(2.788)	(0.652)	(1.526)
Firm currently operates in more than one province	-0.355**	-0.334**	-0.262***	-0.268***
X log Vietnamese employment in industry-year (in own province)	(0.146)	(0.138)	(0.0931)	(0.0943)
Observations	10,963	10,963	10,963	10,963
Control for average firm size X log Vietnamese employment in industry-year (in own province)	NO	YES	NO	YES
Province-industry and year fixed effects	✓	✓	✓	✓

Note: This table repeats the same regression analysis in Table 5 on the full sample of PCI firms (without size restriction). The corresponding RHS variable and IV are also constructed on the full sample firms in the GSO survey (aggregated to province-industry-year). *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 7: Robustness Check: Exclusion Restriction of the Rest-of-Vietnam IV

	Dep. var: Bribe as % of revenue
Log Vietnamese employment in industry-year	-1.163**
(excluding own region)	(0.453)
Constant	18.43***
	(5.690)
Observations	10,898
Province-industry and year fixed effects	✓

Note: This table reports the reduced form results using the alternative construction of the rest-of-Vietnam IV. In particular, we construct total employment in the same industry in other *regions* (as opposed to other provinces as in Table 3). The regression controls for province-industry and year fixed effects. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 8: Plausibly Exogenous Instruments: 95% Interval Estimates

	(1) Coefficient	(2) 95CI lower bound	(3) 95CI upper bound
Main IV estimate (Column 1 Table 3)	-1.704	-3.03	-0.38
Clustering s.e. at jt:			
δ=0		-3.4	-0.008
δ=0.05		-3.059	-0.308
δ=0.1		-2.802	-0.527
δ=0.15		-2.546	-0.743
δ=0.3		-1.792	-1.383
δ=0.5		-2.216	-0.817
Clustering s.e. at r:			
δ=0		-3.37	-0.037
δ=0.05		-3.033	-0.378
δ=0.1		-2.798	-0.616
δ=0.15		-2.565	-0.851
δ=0.3		-1.882	-1.543
δ=0.5		-2.43	-1.004

Note: This table presents 95% confidence intervals for the effect of economic performance on bribes using the "Union of CI" method proposed in Conley, Hansen, and Rossi (2012) across various prior settings. The "Union of CI" intervals impose the prior information that the support of γ, the direct impact of the instrument on the outcome, is $[-2\delta|\beta|, 2\delta|\beta|]$, where β is our original IV estimate (-1.7 in column 1 of Table 3) and δ ranges from 0 to 0.5. In our main analysis, standard errors are two-way clustered at industry-year (jt) and region (r) level. For the plausibly exogenous tests, there isn't a clear theoretical guidance to perform the inference for two-way clustered standard errors. Therefore, we report the results for both ways. For each, we also report the IV estimate for $\gamma = 0$ (i.e., $\delta = 0$).

Appendix Table 9: Heterogeneous Effects Based on Firms' Property Rights (Contemporaneous Measure of LURC)

Dependent variable: Firm's bribe payment as percentage of revenue

	Rest-of-Vietnam Singe IV				China Multiple IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Vietnamese employment in industry-year	-1.910**	1.046	-2.095**	1.055	-0.967	0.973	-1.260**	1.021
(in own province)	(0.840)	(2.707)	(0.904)	(2.784)	(0.599)	(1.976)	(0.585)	(1.944)
Firm owns its land X log Vietnamese employment			-0.326*	-0.350*			-0.218	-0.222
in industry year (in own province)			(0.188)	(0.203)			(0.151)	(0.153)
Firm owns land and has LURC X log Vietnamese	-0.309**	-0.321**	-0.0935	-0.0904	-0.263**	-0.266**	-0.114	-0.116
employment in industry year (in own province)	(0.140)	(0.145)	(0.164)	(0.170)	(0.120)	(0.121)	(0.155)	(0.157)
Observations	9,535	9,535	9,535	9,535	9,535	9,535	9,535	9,535
Control for average firm size X log Vietnamese employment in industry-year (in own province)	NO	YES	NO	YES	NO	YES	NO	YES
Province-industry and year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Note: This table shows the IV results for heterogeneous effects based on firms' property rights using contemporaneous Measure of LURC. The sample contains firms with 10 or more lagged employees reported for the year before the survey.

Appendix Table 10: Heterogeneous Effects Based on Firms' Property Rights (Ordered Probit)

Dependent variable: Firm's bribe payment as percentage of revenue

	Firms that rent their land		Firms that own land without LURC		Firms that own land with LURC	
	(1)	(2)	(3)	(4)	(5)	(6)
Log Vietnamese employment in industry-year (excluding own province)	-0.0402		0.923		-0.432*	
	(0.219)		(0.573)	(0.241)		
Log Chinese employment in industry-year		-0.148		0.0210		-0.0379
		(0.636)		(0.813)		(0.386)
Observations	2,928	2,928	1,399	1,399	4,428	4,428
Province-industry and year fixed effects	✓	✓	✓	✓	✓	✓

Note: This table reports the reduced form results using the ordered probit model for firms in three categories: (1) firms that rent their land; (2) firms that own land without an LURC; (3) firms that own land with an LURC. The sample contains firms with 10 or more lagged employees reported for the year before the survey. All regressions control for province-industry and year fixed effects. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 11: Propensity Score Regressions for LURC and Operation Locations

	Firm has LURC	Number of other provinces in which firm operates	Firm currently operates in more than one province
	(1)	(2)	(3)
Share of registration documents held	0.605***	0.511***	0.490***
	(0.098)	(0.094)	(0.091)
Former HH firm (dummy)	0.432***	-0.058	-0.051
	(0.056)	(0.055)	(0.054)
Former SOE (dummy)	0.192**	0.250***	0.275***
	(0.092)	(0.088)	(0.085)
Owner is government official (dummy)	0.440***	0.391**	0.305**
	(0.163)	(0.153)	(0.143)
Government holds positive share (dummy)	-0.354***	0.116	0.115
	(0.125)	(0.118)	(0.114)
Log of business premise size (hectare)	0.264***	0.024	0.027*
	(0.018)	(0.017)	(0.016)
Years since establishment	0.052***	0.007*	0.009**
	(0.005)	(0.004)	(0.004)
Observations	7074	7872	7872

Note: Column 1 shows the logit regression coefficients of LURC (dummy) on various predictors of obtaining the certificate, including share of registration documents held, former household firm (dummy), former SOE (dummy), owner being a government official (dummy), government holding positive share (dummy), log of business premise size, and firm's age. Column 2 and 3 show the (ordered) logit regression results for firm's operation location variables on the same set of explanatory variables. The sample contains firms with 10 or more lagged employees reported for the year before the survey. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 12: Robustness Check: Heterogeneous Effects Based on Firms' Property Rights (Single IV)

Dependent variable: Firm's bribe payment as percentage of revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Includes characteristic [. . .] and interaction with log Vietnamese employment in industry-year (in own province)							
	share of registration documents held	former HH firm	former SOE	government official	government share	premise size	years of establishment	propensity scores for having LURC and owning land
Panel A. Rest-of-Vietnam IV								
Log Vietnamese employment in industry-year (in own province)	0.00285	1.023	0.841	1.226	1.011	0.650	1.135	1.377
	(2.870)	(2.772)	(2.789)	(2.807)	(2.758)	(3.044)	(2.865)	(2.951)
Firm owns land and has LURC X log Vietnamese employment in industry year (in own province)	-0.102	-0.122	-0.116	-0.105	-0.119	-0.216	-0.0855	-0.130
	(0.142)	(0.130)	(0.133)	(0.130)	(0.132)	(0.156)	(0.127)	(0.164)
Firm owns its land X log Vietnamese employment in industry year (in own province)	-0.438**	-0.338*	-0.349*	-0.337*	-0.342*	-0.252	-0.373*	-0.385*
	(0.195)	(0.190)	(0.186)	(0.183)	(0.189)	(0.207)	(0.205)	(0.212)
Observations	7,671	9,534	9,534	9,534	9,534	8,874	9,347	7,074
Panel B. China IV								
Log Vietnamese employment in industry-year (in own province)	-0.623	0.695	0.639	1.193	0.813	1.476	0.671	1.080
	(2.218)	(2.658)	(2.637)	(2.845)	(2.646)	(2.264)	(2.440)	(2.737)
Firm owns land and has LURC X log Vietnamese employment in industry year (in own province)	-0.0825	-0.0810	-0.0898	-0.0891	-0.103	-0.190	-0.0546	-0.120
	(0.168)	(0.145)	(0.150)	(0.153)	(0.151)	(0.183)	(0.143)	(0.209)
Firm owns its land X log Vietnamese employment in industry year (in own province)	-0.682***	-0.601***	-0.570***	-0.575***	-0.566***	-0.517***	-0.592***	-0.695***
	(0.174)	(0.170)	(0.165)	(0.161)	(0.165)	(0.195)	(0.178)	(0.198)
Observations	7,671	9,534	9,534	9,534	9,534	8,874	9,347	7,074
Control for average firm size X log Vietnamese employment in industry-year (in own province)	YES	YES	YES	YES	YES	YES	YES	YES
Province-industry and year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Note: This table shows the single IV results for heterogeneous effects based on firms' property rights after controlling for other firm characteristics. Column 1 to 8 controls for the following explanatory variables and their interactions with log Vietnamese employment in industry-year in own province: share of registration documents held, former household firm (dummy), former SOE (dummy), owner being a government official (dummy), government holding positive share (dummy), log of business premise size, firm's age, and the corresponding propensity score(s) for having LURC and/or owning land. The sample contains firms with 10 or more lagged employees reported for the year before the survey. All regressions control for the main effects of the property right variables, but only the interaction coefficients are reported in this table. All regressions control for province-industry and year fixed effects. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 13: Robustness Check: Heterogeneous Effects Based on Firms' Property Rights (Multiple IV)

Dependent variable: Firm's bribe payment as percentage of revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Includes characteristic [. . .] and interaction with log Vietnamese employment in industry-year (in own province)							
	share of registration documents held	former HH firm	former SOE	government official	government share	premise size	years of establishment	propensity scores for having LURC and owning land
Panel A. Rest-of-Vietnam IV								
Log Vietnamese employment in industry-year	0.117	0.109	-0.386	0.0878	-0.0225	-0.320	0.0849	0.475
(in own province)	(2.356)	(1.994)	(1.985)	(1.999)	(2.003)	(2.610)	(2.039)	(2.362)
Firm owns land and has LURC X log Vietnamese	-0.0855	-0.0896	-0.0835	-0.0751	-0.0837	-0.165	-0.0869	-0.128
employment in industry year (in own province)	(0.133)	(0.119)	(0.119)	(0.117)	(0.119)	(0.138)	(0.114)	(0.157)
Firm owns its land X log Vietnamese employment	-0.372***	-0.323**	-0.334***	-0.306**	-0.316**	-0.243*	-0.341**	-0.351***
in industry year (in own province)	(0.111)	(0.131)	(0.127)	(0.126)	(0.128)	(0.143)	(0.136)	(0.136)
Observations	7,671	9,534	9,534	9,534	9,534	8,874	9,347	7,074
Panel B. China IV								
Log Vietnamese employment in industry-year	0.647	1.051	0.564	1.109	0.917	0.795	1.110	0.970
(in own province)	(2.149)	(1.926)	(1.929)	(1.930)	(1.911)	(2.159)	(1.937)	(2.236)
Firm owns land and has LURC X log Vietnamese	-0.115	-0.116	-0.125	-0.112	-0.121	-0.212	-0.117	-0.158
employment in industry year (in own province)	(0.134)	(0.119)	(0.118)	(0.118)	(0.118)	(0.129)	(0.113)	(0.152)
Firm owns its land X log Vietnamese employment	-0.291**	-0.243*	-0.228*	-0.214	-0.215	-0.128	-0.242*	-0.229*
in industry year (in own province)	(0.123)	(0.137)	(0.131)	(0.131)	(0.134)	(0.138)	(0.138)	(0.137)
Observations	7,671	9,534	9,534	9,534	9,534	8,874	9,347	7,074
Control for average firm size X log Vietnamese employment in industry-year (in own province)	YES	YES	YES	YES	YES	YES	YES	YES
Province-industry and year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Note: This table shows the multiple IV results for heterogeneous effects based on firms' property rights after controlling for other firm characteristics (see footnote of Appendix Table 12). The sample contains firms with 10 or more lagged employees reported for the year before the survey. All regressions control for province-industry and year fixed effects. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 14: Robustness Check: Heterogeneous Effects Based on Firms' Operation Locations (Single IV)

Dependent variable: Firm's bribe payment as percentage of revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Includes characteristic [. . .] and interaction with log Vietnamese employment in industry-year (in own province)							
	share of registration documents held	former HH firm	former SOE	government official	government share	premise size	years of establishment	propensity score for having multiple locations
Panel A. Rest-of-Vietnam IV								
Log Vietnamese employment in industry-year	-0.539	0.292	0.0605	0.447	0.245	-0.158	0.230	0.804
(in own province)	(2.186)	(1.859)	(1.889)	(1.943)	(1.872)	(1.708)	(1.951)	(2.162)
Firm currently operates in more than one province	-0.297**	-0.264***	-0.267***	-0.263***	-0.266***	-0.201**	-0.283***	-0.258*
X log Vietnamese employment in industry-year	(0.135)	(0.0921)	(0.0944)	(0.0923)	(0.0917)	(0.0912)	(0.0936)	(0.142)
(in own province)								
Observations	8,612	10,900	10,900	10,900	10,900	10,040	10,682	7,872
Panel B. China IV								
Log Vietnamese employment in industry-year	-0.554	0.767	0.623	1.090	0.825	0.810	0.805	2.027
(in own province)	(1.914)	(1.907)	(1.920)	(2.086)	(1.919)	(1.690)	(1.943)	(2.187)
Firm currently operates in more than one province	-0.383**	-0.338**	-0.343**	-0.337**	-0.345**	-0.274*	-0.367**	-0.339*
X log Vietnamese employment in industry-year	(0.190)	(0.150)	(0.152)	(0.147)	(0.148)	(0.158)	(0.155)	(0.199)
(in own province)								
Observations	8,612	10,900	10,900	10,900	10,900	10,040	10,682	7,872
Control for average firm size X log Vietnamese employment in industry-year (in own province)	YES	YES	YES	YES	YES	YES	YES	YES
Province-industry and year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Note: This table shows the single IV results for heterogeneous effects based on firms' operation locations after controlling for other firm characteristics (see footnote of Appendix Table 12). The sample contains firms with 10 or more lagged employees reported for the year before the survey. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 15: Robustness Check: Heterogeneous Effects Based on Firms' Operation Locations (Multiple IV)

Dependent variable: Firm's bribe payment as percentage of revenue

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Includes characteristic [...] and interaction with log Vietnamese employment in industry-year (in own province)							
	share of registration documents held	former HH firm	former SOE	government official	government share	premise size	years of establishment	propensity score for multiple locations
Panel A. Rest-of-Vietnam IV								
Log Vietnamese employment in industry-year (in own province)	-0.425	-0.307	-0.716	-0.0649	-0.404	-1.035	-0.405	1.391
	(2.048)	(1.574)	(1.633)	(1.738)	(1.615)	(2.184)	(1.661)	(2.133)
Firm currently operates in more than one province X log Vietnamese employment in industry-year (in own province)	-0.224*	-0.209**	-0.222**	-0.220**	-0.219**	-0.163	-0.250**	-0.203
	(0.120)	(0.102)	(0.105)	(0.101)	(0.103)	(0.104)	(0.106)	(0.130)
Observations	8,612	10,900	10,900	10,900	10,900	10,040	10,682	7,872
Panel B. China IV								
Log Vietnamese employment in industry-year (in own province)	0.0657	0.481	0.0799	0.797	0.448	0.422	0.478	1.400
	(2.010)	(1.599)	(1.659)	(1.753)	(1.629)	(1.708)	(1.645)	(2.082)
Firm currently operates in more than one province X log Vietnamese employment in industry-year (in own province)	-0.242**	-0.229***	-0.247***	-0.238***	-0.241***	-0.191**	-0.273***	-0.205
	(0.119)	(0.0865)	(0.0919)	(0.0881)	(0.0900)	(0.0928)	(0.0908)	(0.126)
Observations	8,612	10,900	10,900	10,900	10,900	10,040	10,682	7,872
Control for average firm size X log Vietnamese employment in industry-year (in own province)	YES	YES	YES	YES	YES	YES	YES	YES
Province-industry and year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Note: This table shows the multiple IV results for heterogeneous effects based on firms' operation locations after controlling for other firm characteristics (see footnote of Table 12). The sample contains firms with 10 or more lagged employees reported for the year before the survey. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix Table 16: Testing Alternative Models

	Herfindahl index for province-industry-year group	Bribes as % of revenue (excluding construction industry)	Bribes as % of revenue
	(1)	(2)	(3)
Log Vietnamese employment in industry-year	0.0421***	-1.744**	-1.746***
(in own province)	(0.00319)	(0.775)	(0.629)
Log Vietnamese province-year employment			0.884
			(3.335)
Observations	2,780	9,001	10,901
Province-industry and year fixed effects	✓	✓	✓

Note: This table shows the IV regression results for testing several alternative models. In Column 1, the dependent variable is the employment Herfindahl index constructed for each province-industry-year group. The index is normalized to be between 0 and 1, where a small index indicates a competitive industry. For this regression, we keep 1 observation per province-industry-year group. Column 2 shows the IV regression results using single rest-of-Vietnam IV where we exclude firms in the construction industry. For column 3, we include the province-year level total employment and instrument it using a predicted aggregate shock. In particular, the instrument is the weighted sum over all industries of industry-year employment in other provinces, normalized by the average over all years of total employment in the industry in other provinces. The weights are the average over all years of industry's share of employment in the given province. The sample contains firms with 10 or more lagged employees reported for the year before the survey. All regressions control for province-industry and year fixed effects. Standard errors are two-way clustered at the province and industry-year level. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Appendix B: Validation of matching between PCI and GSO

This section assesses the quality of the matching across data sets used in our analysis. Ideally, we would have constructed our endogenous regressor (employment) using the same data set that has our outcome (bribe) data. However, as discussed above, the PCI data, which has information on bribes and firm mobility, is a sample, and does not include all firms. As such, while the PCI is suitable for examining how a typical firm changes, we cannot use it for accurately calculating aggregate shocks. For example, an increase in prices for goods sold by industry j (one source of an increase in A) might lead to entry of firms, so even though A increased, average firm size might decrease. For this reason, we use the GSO data, which is a census, to construct our measure of A. However, before proceeding, it is important to make sure that the PCI firms are a reasonably representative sample of all firms in the GSO data, and that the industry codes we merge on are comparable across the data sets. If not, then the reduced form results from regressing bribes as measured in the PCI data on the GSO-based instrumental variable could be spurious, or null results could reflect poorly matched data.

To cross-validate the two data sets and ensure that we are matching them appropriately, we compare mean and median firm employment among private firms for each pjt group. One issue with the PCI data is that employment is coded as a categorical variable: 10 to 50, 50 to 100, etc. To assign cardinal values to these bins, we compute the empirical mean and median employment for all firms in GSO for each of these PCI bins, and use these values to create the cardinal employment measure for the PCI firms. We then run the following regression, with province-industry and year fixed effects:

$$\log(Employ_{pjt}^{\text{PCI}}) = \alpha + \beta \log(Employ_{pjt}^{\text{GSO}}) + \nu_{pj} + \mu_t + \epsilon_{pjt} \tag{7}$$

If the PCI firms are a perfect random sample of GSO firms, stratified by province, industry and year, we should have $\beta = 1$. We report the estimates in Appendix Table 1. We can see that the changes in mean employment in PCI and mean employment in GSO are positively correlated: β is 0.520 and significant at the 1% level. Similarly, the median employment in PCI and median employment in GSO are positively correlated and the coefficient is 0.474 and highly significant. The correlations are equally robust when we computed the mean and median on the full sample of firms. These results confirm that, while the match between the two data sets is not perfect, they are indeed comparable, even looking just over time at changes within a given province-industry cell.

Appendix C: Model

In this appendix, we formally present a model, which is verbally summarized in the paper, in which governments choose how much to extract from firms to maximize their bribe revenue. Governments balance the revenues they raise from extracting higher bribes from firms with an increased risk that by extracting too much from firms in their jurisdictions, some firms may choose to relocate to other jurisdictions with lower corruption. The model generates the prediction that bribes as a fraction of revenues should decrease with firm growth, which is the main empirical fact we present in our paper. The model also generates the prediction, confirmed in the data, that this relationship is especially strong for firms that can more readily move to another jurisdiction. Other mechanisms could also generate the prediction about the average relationship between firm growth and bribes, but do not easily explain the heterogeneity pattern we find.

For firms, a bribe is an additional payment to government, analogous to a tax. The model is therefore similar to models of inter-regional tax competition, where we think of a bribe payment as a type of tax. The key distinction of our model compared to the previous literature is that we focus not just on the equilibrium level of taxes/bribes, but also examine how the level of bribes changes with productivity shocks. It is this comparative static that generates predictions about how growth affects the amount of corruption in the economy. We also derive how the relationship between productivity shocks and the equilibrium bribe rate varies based on the firm's ease of relocating to another jurisdiction.

We assume that there are two provinces, denoted 1 and 2.[16] Each province is endowed with a unit mass of incumbent firms. Note that this assumption of no entry, which we adopt for tractability, implies that firm growth is equivalent to industry growth. In our empirical work, we examine the effects of industry growth, which combines both growth in firm size and growth in the number of firms.

Government and firms play a static game and move sequentially. First, the government in each province p sets a bribe rate b_p, which is the percent of a firm's revenues that it must pay in bribes. Next, firms in each province choose whether to stay in the province or relocate to the other province. Finally, firms choose their factors of production, they produce, and the government collects bribes.

We begin by specifying the firm's problem, then the problem for local governments, and lastly characterize the equilibrium. Suppose all firms have the same two-factor Cobb-Douglas production function with diminishing returns to scale. We assume diminishing returns to scale in order to pin down firm size and generate profits in equilibrium. Capital and labor are perfectly elastically supplied at the same wage rate w and interest rate r in both provinces.[17] Denote the bribe rate set in period 1 in province p as b_p. We focus on the problem for firms in province 1 (naturally the analysis is symmetric for firms in province 2). A typical firm in

[16] The same results apply in a context where we have a large number of jurisdictions, and firms everywhere face some fixed outside option.

[17] One can think of this as there being a homogeneous outside-good sector that is perfectly traded with regions having the same productivity in this sector. The model assumes symmetric regions and abstracts away from heterogeneous factor prices. The qualitative predictions we discuss in Proposition 1 and 2 do not hinge on this assumption.

province 1 solves

$$\max_{K\geq 0, L\geq 0} (1-b_1)AK^{\alpha}L^{\beta} - wL - rK \tag{1}$$

where A is the total factor productivity of the firm. We can also think of A as encompassing the price of the products in the firm's industry. This maximization problem yields the following familiar results:

$$\frac{L^*}{K^*} = \frac{r}{w}\frac{\beta}{\alpha} \tag{2}$$

$$K^* = \left(\frac{r}{(1-b_1)A\alpha}\left(\frac{r}{w}\frac{\beta}{\alpha}\right)^{-\beta}\right)^{\frac{1}{\alpha+\beta-1}} \tag{3}$$

$$\pi^* = (1-b_1)AK^{*\alpha}L^{*\beta} - wL^* - rK^* \tag{4}$$

In addition to affecting the firm's decision of whether to move as described below, the bribe rate also affects the firm's optimal choice of capital and its profits: the higher the rate of bribe extraction b_1, the smaller the firm's capital stock and profits will be.

The firm will choose to stay in province 1 if and only if profits in province 1 are greater than profits in province 2 less moving costs, i.e. if $\pi^*_{f1} \geq \pi^*_{f2} - m$, where m is the firm's moving costs. To proceed, we need to impose some structure on the moving costs m. We specify the moving costs for firm i as

$$m_i = \theta A^{\eta}\epsilon_i. \tag{5}$$

The term A^{η} captures the fact that the moving costs should be increasing in firm size. For example, the firm's capital stock would need to be moved or sold and repurchased with transaction costs, and larger firms have a larger capital stock. Similarly, new employees would have to be recruited, hired, and trained, and larger firms have more employees. In the context of the model, A is directly related to firm size; higher TFP firms have a larger capital stock and more employees, and therefore larger moving costs.[18] The exponent $\eta \geq 0$ captures the degree to which moving costs are increasing in the size of the firm.

Conditional on firm size, moving costs vary across firms in two ways. First, the θ term captures the part of the firm's moving costs that is observable to the government, with higher θ corresponding to higher moving costs. In our empirical analysis, we focus on a firm's property rights status and whether it has operations in multiple provinces as proxies for the observable components of its moving costs. Second, moving costs include a stochastic term ϵ that varies across firms. While θ will be observable to the government in determining bribe rates, the idiosyncratic part of the moving costs ϵ is unobserved.

[18]Note that all of our key results are robust to instead parameterizing the moving costs in terms of the capital stock K^*, rather than in terms of A, but this is more complicated because K^* is endogenously determined, whereas A is an exogenous parameter of the model. Details for this alternative model are available upon request.

Putting the pieces together, a firm in province 1 chooses to stay if and only if

$$\begin{aligned} \pi_1^* &\geq \pi_2^* - \theta A^\eta \epsilon, \quad \text{or} \\ \epsilon &\geq \frac{\pi_2^* - \pi_1^*}{\theta A^\eta} \end{aligned} \tag{6}$$

To simplify the algebra, we further assume that ϵ is uniformly distributed over $[0, 1]$.[19] The equilibrium number of firms for a given θ in province 1 is therefore simply $1 - \frac{\pi_2^* - \pi_1^*}{\theta A^\eta}$.[20] Since the problem is symmetric for both provinces, this expression will be greater than 1 if $b_1 < b_2$ (firms are moving into province 1 from province 2), and less than 1 if $b_1 > b_2$ (firms are moving out of province 1 to province 2).

The two governments in period 1 set bribe rates, taking firms' response and the other province's bribe rate as given. To solve this, we consider the government in province 1. It takes b_2 as given and solves,

$$\max_{b_1 \geq 0} b_1 A K^{*\alpha} L^{*\beta} \left(1 - \frac{\pi_2^* - \pi_1^*}{\theta A^\eta}\right) \tag{7}$$

Assuming a symmetric equilibrium, the first-order condition can be simplified to:

$$K^* + b_1^*(\alpha + \beta)\frac{dK^*}{db_1} + \frac{b_1^* K^*}{\theta A^\eta}\frac{d\pi_1^*}{db_1} = 0 \tag{8}$$

After some algebra, we get:

$$\left(\frac{1}{\theta} A^{1-\eta} \left(\frac{r\beta}{w\alpha}\right)^\beta K^{*\alpha+\beta} + \frac{\alpha + \beta}{1 - \alpha - \beta}\frac{1}{1 - b^*}\right) b^* = 1 \tag{9}$$

Note that we have suppressed the province subscript since $b_1^* = b_2^*$ in equilibrium.

Several aspects of the equilibrium condition in Equation (9) are worth noting. First, as θ goes to $+\infty$, or firms are completely immobile, the expression simplifies such that $b^* = 1 - \alpha - \beta$. This implies that the greater the diminishing returns to scale, the higher the bribe rate. Intuitively, if output is highly concave in capital, even when the bribe rate is reduced, firms will not expand their capital stock much due to diminishing returns. Thus, the elasticity of capital with respect to the bribe rate is low. The same applies to labor. Therefore, when the government increases the bribe rate, it can extract more revenue from firms without discouraging production. Hence, the optimal bribe rate is higher.

The second observation is that as θ decreases, so that moving costs decrease, inter-regional competition increases and the equilibrium bribe rate decreases. Thus far, the model captures the idea that increasing competition between political jurisdictions can drive down corruption, as in Shleifer and Vishny (1993) and Burgess et al. (2012).

[19] This assumption simplifies the algebra but is not essential; all of the key results go through for arbitrary distributional forms of the error term.

[20] Even though we have in mind a world of many firms with heterogenous θ, we are solving the model for a particular θ. After we obtain the equilibrium bribe rate, which is a function of θ, we will examine how bribes and the effect of firm growth on bribes vary with θ.

Next, we examine how the equilibrium bribe rate responds to increases in the productivity of firms, i.e. increases in A. Taking the derivative with respect to $\log A$ on both sides of Equation (9) and re-arranging terms, we get our first result:

Proposition 1. $\frac{db^*}{d\log A} < 0$ *if* $0 \leq \eta < \frac{1}{1-\alpha-\beta}$*;* $= 0$ *if* $\eta = \frac{1}{1-\alpha-\beta}$*; and* > 0 *if* $\eta > \frac{1}{1-\alpha-\beta}$.

Proof. : One way to prove the results is using standard first order condition and taking derivative by invoking implicit function theorem. Here we first make a transformation of variables to simplify the problem. First notice that by replacing K, L, A, θ by $rK, wL, \frac{A}{r^\alpha w^\beta}, \theta(r^\alpha w^\beta)^\eta$, we can assume $w = r = 1$ without changing the problem. This linear change of variable also preserves the signs of the derivatives in question.

We can rewrite the firm's objective function:

$$\pi_1^* = max_{K,L\geq 0}(1-b_1)AK^\alpha L^\beta - K - L$$

which gives

$$\begin{aligned} K_1^* &= \alpha((1-b_1)A\alpha^\alpha\beta^\beta)^{\frac{1}{1-\alpha-\beta}} \\ L_1^* &= \beta((1-b_1)A\alpha^\alpha\beta^\beta)^{\frac{1}{1-\alpha-\beta}} \\ \pi_1^* &= (1-\alpha-\beta)((1-b_1)A\alpha^\alpha\beta^\beta)^{\frac{1}{1-\alpha-\beta}} \end{aligned}$$

Substituting these into the government's problem, we are able to calculate the total bribe received by government in province 1:

$$B_1 = t(1-\lambda(R_2-R_1))(R_1^{\alpha+\beta} - R_1) \tag{10}$$

where

$$\begin{aligned} R_1 &= (1-b_1)^{\frac{1}{1-\alpha-\beta}} && (11) \\ R_2 &= (1-b_2)^{\frac{1}{1-\alpha-\beta}} && (12) \\ \lambda &= \frac{A^{\frac{1}{1-\alpha-\beta}}}{\theta A^\eta}(1-\alpha-\beta)(\alpha^\alpha\beta^\beta)^{\frac{1}{1-\alpha-\beta}} && (13) \\ t &= (A\alpha^\alpha\beta^\beta)^{\frac{1}{1-\alpha-\beta}} && (14) \end{aligned}$$

Instead of choosing $b_1, b_2 \in (0,1)$, the two provinces could choose $R_1, R_2 \in (0,1)$ as well. The FOC for (10) then yields

$$\lambda(R_1^{\alpha+\beta} - R_1) = (1-\lambda(R_2-R_1)) * (1-(\alpha+\beta)R_1^{\alpha+\beta-1}) \tag{15}$$

The symmetric equilibrium $R_1 = R_2 = R^*$ must then satisfy

$$\lambda(R_1^{\alpha+\beta} - R_1) = 1-(\alpha+\beta)R_1^{\alpha+\beta-1} \tag{16}$$

This gives the equilibrium bribe rate b^* via (11).

With that preliminary math completed, we can complete the proof. Re-arrange (16) and

notice that the ratio

$$\phi := \frac{1-(\alpha+\beta)R_1^{\alpha+\beta-1}}{(R_1^{\alpha+\beta}-R_1)}$$

is strictly increasing in R, because $R_1^{\alpha+\beta}-R_1$ is concave and hence log-concave in R. Therefore, (16) has a unique solution for fixed parameters (so fixed λ), and this solution is strictly increasing in λ. In other words,

$$\frac{dR^*}{d\lambda} > 0 \tag{17}$$

From (11), $b^* = 1-(R^*)^{1-\alpha-\beta}$, and so $\frac{db^*}{dR^*} < 0$. Also from (13), $\frac{d\lambda}{dA}$ is positive when $\eta < \frac{1}{1-\alpha-\beta}$ and negative otherwise. Multiplying the three inequalities, we obtain

$$\begin{aligned} \frac{db^*}{dA} &> 0, \text{when } \eta > \frac{1}{1-\alpha-\beta} \\ \frac{db^*}{dA} &< 0, \text{when } \eta < \frac{1}{1-\alpha-\beta} \end{aligned}$$

■

The critical factor that determines the sign of $db^*/d\log A$ is η, which characterizes the concavity of the moving costs with respect to the capital stock. The intuition is that with a positive shock to A, for a given size, firms enjoy higher revenues and hence care more about the bribes they will pay and less about the moving costs. This tends to drive down the equilibrium bribe rate due to inter-regional competition. However, at the same time, the cost of moving rises as firms expand in size to take advantage of the higher productivity. This instead tends to drive up the equilibrium bribe rate. The two effects exactly cancel at $\eta = \frac{1}{1-\alpha-\beta}$.[21] If $\eta < \frac{1}{1-\alpha-\beta}$, then the first effect (inter-regional competition effect) dominates the second effect (moving cost effect), and the equilibrium bribe rate falls. Given that $1-\alpha-\beta < 1$, a sufficient condition for $\frac{db^*}{d\log A} < 0$ is that moving costs scale up less than linearly with firm size, as proxied by A. Moving costs seem likely to fulfill this assumption in practice and, moreover, because $1-\alpha-\beta$ can in fact be much less than 1, it seems plausible that $\eta < \frac{1}{1-\alpha-\beta}$ and therefore $db^*/d\log A < 0$ in most settings. The empirical evidence presented in the previous section matches this prediction of the model.

Proposition 1 specifies conditions when the *rate* of bribe extraction falls as A increases; the rate b is the size of the distortion to production. It is worth noting that another (testable) prediction is that the *total amount of bribes* extracted from the firm will increase when A increases. To see this note that the firm's moving decision is a tradeoff between its total moving costs and its total bribes. Since when A increases, the firm's moving costs increase, the government can retain the same firms even with a higher total bribe extraction.

Also worth noting is that to the extent that taxes follow similar patterns to bribes, another implication of the model is that taxes on firms should also be lower in rich countries than in poor countries. There is suggestive evidence along these lines: Gordon and Li

[21] Though the specific cut-off value for η at which the sign of the comparative statics switches depends on the Cobb-Douglas functional form for production and the particular parameterization of the moving costs, the general intuition is robust: The more concave the moving costs with respect to firm size are, the more likely that equilibrium bribe rate decreases with A.

(2009) show that for poor countries (with per-capita GDP below $745), corporate income taxes represent 7.5 percent of GDP, whereas for rich countries (with per-capita GDP above $9,200), corporate income taxes represent only 4.5 percent of GDP, although they suggest a different explanation than the one proposed here.

Next, we examine how the effect of a productivity shock on bribes varies across firms with different θ. We will focus on the firm's property right status or multi-province operations as the empirical analogue of θ, where higher θ corresponds to less transferable property rights or concentration of operations in one province and thus a higher cost of moving. The next proposition derives how the elasticity of bribes with respect to productivity varies with θ.

Proposition 2. *The elasticity* $-\frac{d\log b^*}{d\log A}$ *is monotonically decreasing in* θ *if and only if* $0 \leq \eta < \frac{1}{1-\alpha-\beta}$. *That is,* $\frac{d^2\log b^*}{d\log A d\theta} > 0$ *iff* $0 \leq \eta < \frac{1}{1-\alpha-\beta}$.

Proof. : From (11),

$$\epsilon_{R^*} = \frac{1}{1-\alpha-\beta}\frac{b^*}{1-b^*}\epsilon_{b^*} \tag{18}$$

where $\epsilon_R = \frac{dR}{dA}\frac{A}{R}$ and $\epsilon_b = -\frac{db}{dA}\frac{A}{b}$. Taking the derivative of (16) with respect to A, we get (omitting stars)

$$\epsilon_\lambda := \frac{1}{1-\alpha-\beta} - \eta = \left(\frac{(\alpha+\beta)(1-\alpha-\beta)R^{\alpha+\beta-1}}{1-(\alpha+\beta)R^{\alpha+\beta-1}} + \frac{R-(\alpha+\beta)R^{\alpha+\beta}}{R^{\alpha+\beta}-R}\right)\epsilon_R \tag{19}$$

Substitute (18) and (11) into (19), we have

$$\frac{\epsilon_\lambda}{\epsilon_b} = \left((\alpha+\beta)R^{\alpha+\beta-2} + \frac{R^{\alpha+\beta-1}-1}{(1-\alpha-\beta)R}\right)\phi^{-1} \tag{20}$$

It is easy to see that b^* (R^*) increases (decreases) with θ, and thus the right hand side increases with θ. Therefore,

$$\frac{d\epsilon_{b^*}}{d\theta} < 0, \text{when } \epsilon_\lambda > 0 \Leftrightarrow \eta < \frac{1}{1-\alpha-\beta} \tag{21}$$

■

Intuitively, Proposition 1 implied that bribes fall when there is an increase in A, because more profitable firms are more willing to pay moving costs and escape from high bribe rates. Proposition 2 states that the bribe rate falls more after such a shock for firms with lower observable moving costs because the fraction of firms who are on the margin of moving is larger, so a given change in bribes will induce a larger number of them to leave.[22]

[22]Note that Proposition 2 is stated in terms of elasticity or percentage change in the bribe rate (i.e. the change in $\log b$). The sign for the cross-partial of the level change (ie. $\frac{d^2b^*}{dAd\theta}$) is in general indeterminant because though the elasticity falls with θ (under the condition in Proposition 2), the level of bribe also falls with θ. In particular, we can show that the relationship between $\frac{db^*}{dA}$ and θ is U-shaped and increasing as $\theta \to \infty$, and $\frac{d^2b^*}{dAd\theta} > 0$ for a reasonable range of θ as assessed by the ratio of moving costs to revenue.

Appendix D: Description of key variables

The key variables described in Table 1 are below. Note that each observation is a firm; the sample pools firms for the years 2006 to 2010. The PCI firms in our sample are firms with at least 10 employees, valid industry and bribe payment data and whose province-industry-year is represented in the GSO data.

- Bribes as percentage of revenue (PCI): This is a 7-point categorical variable drawn from question D6 of the annual PCI survey. Respondents answered within ranges: 1) 0%; 2) less than 1%; 3) 1-2%; 4) 2-10%; 5) 10-20%; 6) 20-30%; 7) over 30%. We recode each category with the corresponding cell mean with over 30% recoded as 35%.

- Years since establishment (PCI): Continuous variable that subtracts year of establishment from the year the firm completed the survey (2006 to 2010). Establishment only captures when the firm began doing business and has no legal connotation. A follow-up question asks when a firm registered as a formal business.

- Number of employees (PCI): This variable is categorical in PCI: for example, 10-49, 50-199, etc. We recode each category with the corresponding empirical cell mean in GSO.

- Mean employment (GSO, mean for industry-year-province level): Continuous measure collected in the GSO Enterprise Census.

- Log employment (GSO, aggregate for industry-year-province): Continuous measure collected in the GSO Enterprise Census.

- Log of business premise size (hectares) (PCI): Continuous measure collected in only the 2009 to 2011 surveys.

- Land ownership (dummy) (PCI): Dichotomous variable measuring whether a firm purchased (=1) or leases (=0) its main business premises.

- Land use right certificate (LURC, dummy) (PCI): Dichotomous variable measuring whether a firm possesses an LURC for its main business premises.

- Land ownership without land use right certificate (dummy) (PCI): Dichotomous variable, calculated from questions about whether a firm owns land and whether it has an LURC for that land.

- Number of other provinces in which firm operates (PCI): Count variable calculated from firm's response to a question asking it to record all provinces and national-level cities in which it has operations or branch offices, outside of its headquarters location.

- Firm currently operates in more than one province (dummy) (PCI): Dichotomous variable for whether firm listed operations in at least one province outside the province with its main headquarters.

- Share of registration documents held: Based on the following question: "In addition to the certificate of business registration, your business may need additional permits or business to be fully legal (e.g. mining licenses....). How many of these documents were required for your firm?' "Y" Variable defined as the number of registration documents held by the firm divided by the 95% percentile of the firm's industry (a proxy for formality).

- Former household firm (dummy) (PCI): Dichotomous variable for whether the firm operated as an informal business based on household premises before formal registration.

- Former SOE (dummy) (PCI): Based on questions asking whether the firm is a privatized former State Owned Enterprise (SOE). Equals 1 if the firm is either a locally or centrally managed SOE.

- Owner is a former government official (dummy) (PCI): Based on question asking whether the owner of the private enterprise is a former government official, army office, or SOE manager.

- Government holds a positive share (dummy) (PCI): Based on question asking whether shares of the firm are held by local or central government officials.

Appendix E: Testing alternative models

Empirical confirmation of several predictions of our model supports the idea that inter-jurisdictional competition is a mechanism through which economic growth can reduce bribery. However, there are also other potential models that predict a negative correlation of economic growth and the bribe rate. The first and most direct way to distinguish between the inter-jurisdictional model and these other models is that we find that the relationship between growth and bribery is diminished for firms that are less likely to relocate outside their province. This is a direct prediction of inter-jurisdictional competition, but is not predicted by most other models.

Nonetheless, it is possible that these heterogenous effects are picking up other firm characteristics besides property rights or multiple locations. Thus, it is important to consider several other possible explanations for the general pattern that economic growth reduces bribes, and to discuss the degree to which our evidence is, or is not, consistent with them.

Product-market competition

Economic growth could increase competition among firms, and this product market competition affects the amount of rents bureaucrats can capture. If firms have less market power and smaller rents, then bureaucrats may be less able to extract bribes from them. Ades and Di Tella (1999) present empirical evidence that product market competition reduces corruption, for example. To probe the possibility of this mechanism, we test the starting premise that the variation in economic growth that we analyze increases market competition. We regress the Herfindahl index, constructed using employment (our most accurate measure of firm size) from the GSO data, on employment, instrumented with industry employment in the rest of Vietnam ($\log(Employ_{p^-jt})$). We find that higher predicted employment leads to less, not more, competition, suggesting that the main mechanism through which growth reduces bribery in our context is not increased firm competition (Appendix Table 16, column 1).[23] However, Bliss and Di Tella (1997) present a model in which, counterintuitively, less competition among firms can lead to less bribery; it is possible that this mechanism of reduced competition among firms (higher rents for firms) leading to a reduction in bribe extraction is at play in our setting.

Spurious effect of industry-specific bribe crackdowns

A second possibility we consider is that there are industry-specific crackdowns on bribes. As discussed earlier, this represents the fundamental identification assumption of the within-Vietnam analysis: There are no industry-specific crackdowns on bribes. The strongest evidence in support of our assumption is that when we use industry size in China to instrument for industry size in Vietnam, we find similar results, suggesting that internal reverse causality within Vietnam is not driving the results.

[23] Another option would be to test for changes in profit margins directly. However, the profit margin data in the GSO is known to be much less reliable than employment (Tran and Dao, 2013), as firms routinely underreport profits to avoid taxes. For example, in the GSO data, 38 percent of firms report a profit margin of less than 1 percent of revenues, with 23 percent of firms reporting 0 profits. Given these reporting issues, the PCI dataset does not ask about profits.

Moreover, the institutional structure of Vietnam, which we described in section 2.1 is such that bribery is decentralized to the province level, suggesting that national industry-level crackdowns are unlikely (recall that our identifying variation is essentially Vietnam-wide growth for an industry). In addition, we undertook a systematic review of the national anti-corruption website, which documents major anti-corruption efforts of the government. Over the study period, only one industry-specific anti-corruption campaign is documented, a crackdown in the construction industry in 2008. Appendix Table 16, column 2 shows that the main results are essentially unchanged when we re-run our main specification from column 1 of Table 3, but excluding the construction industry. As the fundamental identification assumption, it is difficult to establish empirically that there are no industry-specific shocks to bribes, but the qualitative evidence points against such an explanation for the patterns we find.

Fixed cost of anti-corruption enforcement

Another possibility is that there is a layer of oversight over bureaucrats aimed at rooting out corruption, such as an anti-corruption agency. The overseers face a fixed cost of enforcement, so as the total scale of bribery (in levels) goes up, it is easier to detect and punish bribery. Or said differently, it may be easier to detect a larger bribe. If so, then as firms grow, bureaucrats will adjust the bribe rate down.

While this explanation may be at work at the cross-country level, it does not seem to be a key factor explaining the results in this paper. In particular, since most regulatory activities are at the province level, if fixed costs were the main explanation, then the key factor determining bribes would be the overall size of economic activity in the province as a whole, not the size of particular industries.

To test this, column 3 of Appendix Table 16 reruns the main IV regression, but in addition to including log employment in a province-year in the particular industry, we also include the aggregate endogenous variable, which is log employment in the province-year in all industries. Because we have two endogenous regressors, we need two instruments. For the first, we use our standard single IV that is industry-year employment in the rest of Vietnam. The second IV aggregates employment across all industries in the rest of Vietnam. To construct this aggregate instrument, we calculate the log of the weighted sum over all industries of employment in the industry in a given year in all provinces excluding province p itself (relative to the average over all years), where the weights represent the share of total employment in province p that comes from industry j. Specifically, define the weights for each industry j in province p as:

$$w_{pj} = \frac{1}{T} \sum_{t} \frac{Employ_{pjt}}{Employ_{pt}}. \tag{22}$$

Then the predicted log employment in province p at time t is:

$$predicted \log(Employ_{pt}) = \log\left(\sum_j w_{pj} \frac{Employ_{p^-jt}}{\left(\frac{1}{T}\sum_t Employ_{p^-jt}\right)}\right) \tag{23}$$

where the $\frac{1}{T}\sum_t Employ_{p^-jt}$ term is a normalization such that the weights to each industry are given exactly by w_{pj}. Note that with a single industry j in a province, $predicted \log(Employ_{pt})$ reduces exactly to $\log(Employ_{p^-jt})$.

Using this approach, the evidence in column 3 suggests that, the IV coefficient estimate on overall province employment, $\log(Employ_{pt})$, is insignificant, and the negative coefficient on industry-specific $\log(Employ_{pjt})$ is unchanged from our main specification. This suggests that the results are not being driven by aggregate changes in enforcement practices at the province level.

Diminishing returns to bribes

A final alternative explanation for the main effect we find is that bureaucrats have diminishing marginal utility of income relative to the risk of being caught and going to jail. Thus, as it becomes easier to extract a given amount of bribe revenues, they reduce the rate. The reduced form effect of this mechanism would be similar to a fixed cost of anti-corruption enforcement within each province: A decline in bribes would be driven by aggregate employment in the province, not industry-specific employment. The results in column 3 of Appendix Table 16 suggest that this is not the entire explanation for our results.

Summary

In summary, to the extent we can examine quantitative and qualitative predictions of these alternative models, we do not find that other mechanisms can explain all the facts in the data. It still may well be that these other mechanisms are in operation and explain some of the overall effect of growth on bribery. But, the positive evidence in support of inter-jurisdictional competition and the limited evidence in support of other models suggests that the mechanism we highlight is at least one important factor in why economic growth reduces corruption in Vietnam.