University of Freiburg
Department of International Economic Policy
Discussion Paper Series
Nr. 46

The Long-term Effect of Western Customs Institution on Firm Innovation in China

Gan Jin
Günther G. Schulze

February 2024

ISSN 1866-4113

University of Freiburg
Department of International Economic Policy
Discussion Paper Series

The Discussion Papers are edited by:
Department of International Economic Policy
Institute for Economic Research
University of Freiburg
D-79098 Freiburg, Germany
Rempartstraße 16

Tel: +49 761 203 2342
Fax: +49 761 203 2414
Email: iep@vwl.uni-freiburg.de

Editor:
Prof. Dr. Günther G. Schulze

ISSN: *1866-4113*
Electronically published: 22.2.2024

The Long-term Effect of Western Customs Institution on Firm Innovation in China

Gan Jin [*]
Günther G. Schulze [†]

February 2024

Abstract

Can historical institutions affect today's firm innovation? We analyze a historical experiment in 1902, when the foreign-run Chinese Maritime Customs Service (CMC), known for its efficient and transparent governance, took over *some* of the notoriously corrupt Chinese Native Custom stations and improved their governance. Using a large data set of contemporary industrial firms in China, we show that firms in locations historically affected by the CMC rules exhibit higher innovation intensities today, which can be attributed to the persisting norms of honesty and lawfulness embedded in the CMC institution. They reduce local corruption and stimulate firms' investment in R&D and training to this day. We identify a causal effect by comparing firms in locations affected by the takeover with firms in similar but unaffected regions nearby. We also use an IV strategy that exploits the takeover criterion, which stipulated that Native Customs stations within a 25 km radius of a CMC customs station could be taken over by the Western powers.

Keywords: Innovation, Persistence, Institutions, Corruption, China

JEL Codes: N75, N45, D73, Z10, O31

[*]Corresponding author. Albert-Ludwigs-University of Freiburg, Institute of Economics, Rempartstr. 10, 79085 Freiburg, Germany. Email: gan.jin@vwl.uni-freiburg.de

[†]Albert-Ludwigs-University of Freiburg, Institute of Economics and CESifo. Email: guenther.schulze@vwl.uni-freiburg.de

1 Introduction

A growing body of literature suggests that historical institutions matter for economic prosperity today (Acemoglu et al., 2001; Dell, 2010; Michalopoulos and Papaioannou, 2013; Nunn, 2009, and others). Historical institutions have been shown to affect present urbanization, public service provision, human capital, and social capital (Chen et al., 2020; Guiso et al., 2016; Jedwab et al., 2017; Jin and Schulze, 2024), all of which are among relevant determinants of long-run economic growth. However, can historical institutions affect also today's innovation, arguably the most significant factor of growth? If so, through which channel does the effect persist?

We are the first to study this issue. Utilizing a unique policy experiment, we show that the institution established by the Chinese Maritime Customs Service (CMC) - a foreign-run customs agency in China during 1854-1949 - has a lasting causal effect on the innovation rate of Chinese firms today. We use firm-level data from the Annual Survey of Industrial Firms and find that firms located in regions historically influenced by the CMC institution exhibit an innovation intensity that is 8 ppts (or more than half of a standard deviation) larger than the average firm. The persistent effect of the CMC institution is likely explained by the long-lasting norms of honesty and lawfulness embedded in the CMC institution, which reduced local corruption and boosted firms' investment in R&D and training.

The treaty port era of China (1840-1949) offers an appropriate testing ground to identify the economic effect of historical institutions: After the defeat in the Opium War (1839-1842), the British forced China to open up "treaty ports" that allowed foreign trade. In 1854, the British established the CMC to administer trade in these port cities. Known for its efficiency and transparency, the CMC processed international trade brought to China by Western businessmen. In contrast, the original Chinese customs system at the time was organized by Native Customs with a notoriety of being corrupt and inefficient, and it only administered Chinese inland trade after the foundation of the CMC. The unique dual-customs system in late 19th and early 20th century China provides an appropriate historical experiment to compare institutions of different origins and qualities in the same country.

Our identification strategy is based on a historical event in 1902 when the CMC took over *some* of the Native Customs stations, transmitting its administrative principles to them. The affected Native Customs stations adopted clear procedural guidelines and effective anti-corruption measures, resulting in the coexistence of two types of institutions within the Chinese inland trade system. We measure the exposure to the CMC institution at the regional level, exploiting the variation in the takeover cases of Native Customs stations in each county. Our identification is further strengthened, first, by focusing on counties neighboring those CMC stations that took over Native Customs stations. This allows for comparing similar areas that only differed in the institutional quality of their customs. Second, we employ an instrumental variable (IV) strategy by utilizing the criterion determining which Native Customs stations were taken over by the CMC, i.e., those located within a 25 km radius from a CMC customs station. In the first-stage regression, we use the share of a county's area covered by a circular region with a 25 km radius centered around a CMC station to predict the takeover of a Native Customs station within that county (cf. Jin, 2023).

The pro-innovation effect of the CMC institution can be attributed to its long-lasting cultural influence, particularly in curbing corruption within the local business environment. By introducing formal rules and prohibiting corrupt practices, the transmission of the CMC institution to selected parts of the Chinese inland trade system fostered a cultural norm of honesty and lawfulness in the affected regions. Such norms have persistently reduced corruption at the local level, with significant implications for today's firm innovation. We show that the positive impact of the CMC institution on innovation is most pronounced when corruption poses a substantial barrier to innovative activities: our findings highlight that the pro-innovation effect is specifically observed in private firms, which encounter greater exposure to rent-seeking behavior and governmental expropriation in China. Additionally, our analysis using an industry-level corruption index - derived from World Bank Enterprise Survey data - demonstrates that the positive impact of the CMC institution on innovation is only significant in industries characterized by a high level of corruption. Notably, the anti-corruption channel is unlikely to be confounded by other relevant factors such as competition, foreign direct investment, and trade.

Our paper is inspired by the economic history literature on the determinants of innovation (Babina et al., 2023; Cinnirella and Streb, 2017; Danzer et al., 2023; Donges et al., 2023; Gross and Sampat, 2023; Moser et al., 2014). For instance, recent studies document how the Great Depression and public investment in R&D during World War II affected innovation in the U.S. (Babina et al., 2023; Gross and Sampat, 2023). In a related study, Donges et al. (2023) examine the history of the French occupation of German regions in the 19th century. They show that regions occupied by the French demonstrated a higher number of patents per capita than the non-occupied areas, which is largely explained by the more inclusive French legal institutions. Our paper documents a similar pro-innovation effect of good institutions but from a persistent perspective: we are the first to link *historical* institutions to *today's* innovation. Our evidence suggests that long-gone historical institutions can leave a far-reaching legacy on innovation through culture.[1]

More broadly, our paper speaks to the literature on the long-term effect of historical institutions on economic development, which is based on seminal studies such as Acemoglu et al. (2001), Acemoglu et al. (2002), and Engerman and Sokoloff (1997). By focusing on the institution-innovation nexus, we broaden the spectrum of potential channels through which history shapes economic development today, arguably by a very important one – innovations. Our paper also underscores the role of cultural persistence in explaining the long-run effect of historical institutions as shown by Voigtlander and Voth (2012); Guiso et al. (2016); Becker et al. (2016); Lowes et al. (2017); Dell et al. (2018); Chen et al. (2020); Jin (2023) and others. For instance, Becker et al. (2016) demonstrate the long-lasting cultural legacy of trust in public services and reduced corruption left by the Habsburg Empire's institutions, documenting a similar corruption-reducing effect of historical institutions. Our evidence adds to the literature by underscoring the cultural legacy of norms of honesty and lawfulness stemming from historical institutions. This sheds light on an essential yet under-researched positive cultural outcome of colonial institutions. Finally, our focus on a particular trait of historical institutions - being efficient and transparent - distinguishes this study from previous works that consider institutions as a package of state-level laws and bureaucracy.

This paper also adds to the analysis of the determinants of innovation, which plays a pivotal role in driving economic growth (Solow, 1957; Kogan et al., 2017). Corruption - the main focus of this study - is widely acknowledged as a major barrier to firm innovation (Murphy et al., 1991; Anokhin and Schulze, 2009; Paunov, 2016; Dincer, 2019; Ellis et al., 2020; Lee et al., 2020; Huang and Yuan, 2021).[2] The relationship between corruption and innovation in the above literature is further investigated through case studies in China, particularly within the context of President Xi Jinping's anti-corruption campaign initiated in 2012. For example, Xu and Yano (2017) show that intensified anti-corruption efforts led

[1] Recent studies also document the persistence of innovation patterns across space and industry (Andrews and Whalley, 2022; Danzer et al., 2023). However, our approach is different in addressing a fundamental cause of such persistence.

[2] This literature is based on a larger body of research discussing the macroeconomic impact of corruption (Shleifer and Vishny, 1993; Mauro, 1995; Aidt, 2009; Lambsdorff and Schulze, 2015; Dimant and Tosato, 2018; Zakharov, 2019).

to increased R&D investment by firms, while Fang et al. (2018) find that the campaign resulted in more effective utilization of R&D subsidies. By uncovering the effect of the CMC institution on today's innovation, our study emphasizes the role of institutional quality in fostering a pro-innovation business environment. Our findings suggest that the corruption–innovation nexus may be deeply rooted in history.

Lastly, our paper sheds further light on the literature examining the long-run economic effect of Western presence in China. A growing body of research suggests that China's decades-long economic growth after 1978 can be attributed to factors deeply rooted in its historical context (Brandt et al., 2014; Chen et al., 2022; Keller et al., 2011). Following this argument, Jia (2014) shows regions where treaty ports were established experienced faster economic growth than unaffected areas in contemporary times. In the same vein, Jin (2023) and Jin and Schulze (2024) examine the long-run economic effect of the CMC institution and Concessions in treaty ports, respectively. In this study, we provide novel insights into the causal relationship between the CMC institution and innovation utilizing a detailed firm-level data set. By doing so, we contribute to a deeper understanding of the mechanisms through which treaty ports may have exerted a long-lasting influence on China's economy.

The remainder of the paper is organized as follows: Section 2 provides the historical background. Section 3 describes the data and the empirical strategy. In Section 4, we identify the causal effect of the CMC institution on firm innovation, Section 5 contains various checks demonstrating the robustness of our results. Section 6 examines the anti-corruption norms as the main mechanism through which CMS affects today's innovation rates. Section 7 analyzes possible confounding factors that could have invalidated the cultural channel. Section 8 concludes.

2 Historical Background

Before 1840, China's trade was exclusively administered by the Native Customs. The Qing government, which ruled China between ca. 1644 and 1912, established 40 Native Customs stations across Chinese regions, with the station in Guangzhou being the only one administering foreign trade under the "Sea Ban policy" since 1757 (Qi, 2004; Deng, 2007). The Native Customs system was characterized as highly inefficient and corrupt during the late period of Qing. Specifically, the Native Customs stations operated under informal rules and private negotiations for tax payments (Qi, 2004; Liao, 2010). Corruption thrived as the head of the station, the superintendent, devised various fees and fines to generate personal income, while positions within the customs were often occupied by family members of high officials and individuals who inherited rights without any associated responsibilities (CMC, 1907; Wright, 1950; Jin, 2023).

China's defeat in the Opium War and the signing of the Nanking Treaty in 1842 forced the opening of treaty ports for foreign trade; they also marked the end of China's self-imposed economic isolation. To address the shortcomings of the Native Customs in regulating foreign trade, the foreign consuls in Shanghai established the CMC in 1856. It operated under modern customs rules characterized by transparency, strict administrative procedures, and a merit-based personnel policy that valued honesty and discouraged corruption. The CMC introduced a dual-customs system in China, with the Native Customs administering only trade conducted by traditional Chinese vessels, primarily for inland transactions.

In the aftermath of China's defeat in 1900, the *Boxer Protocol* compelled the Chinese government to pay 450 million HK. taels of silver to foreign powers.[3] As part of the agreement, the CMC took over Native Customs stations to ensure the collection of revenues from Chinese inland trade for the payment. Through discussions and negotiations, the CMC eventually took over only Native Customs stations that were no further than 50 Chinese *li*[4] from each CMC station, resulting in 24 Native Customs stations taken over by the CMC (Huang, 1917). After the takeover, the CMC also implemented reforms in these Native Customs stations that simplified administrative procedures, issued written tax codes, and discharged incompetent personnel, resulting in increased tax collection and improved efficiency (CMC, 1907; Dai, 1989; Qi, 2004; Tsai, 2008). Historical evidence indicates the success of these reforms, with significant revenue growth observed in stations such as Santuao and two stations in Fujian province (Dai, 1989).

In light of the above historical facts, we hypothesize that the local equilibrium of business-government interactions was shifted from a corruptible one where practices such as bribery and exploiting taxpayers were common to a more predictable one where wrongdoings of businessmen and public officials were restricted. Even though the formal CMC institution ceased to exist after 1949, the cultural norms that promote honesty and lawful-

[3]HK. tael was used by customs stations to weigh silver: one HK. tael is approximately 37 g. At the exchange rate in 1900, 450 million Hk. taels of silver is worth approximately 330 million U.S. Dollars. The recipients of the payments were the United Kingdom, France, the United States, Japan, Russia, Germany, Italy, and Austria-Hungary.

[4]The distance of 50 Chinese *li* is approximately 25 km.

ness may have persisted in local business practices.

3 Data and Empirical Strategy

3.1 Data Sources and Sample

Our main data source is the Anual Survey of Industrial Firms (ASIF) in 2007, conducted by the National Bureau of Statistics of China (NBS). The survey focuses exclusively on the mining, energy, and manufacturing industries. It features a large sample size that covers more than 90% of Chinese industrial firms and a wide coverage of economic indicators. Regarding the scope of the survey, ASIF includes only firms above a certain scale (or called "above-scale" firms): from 1998 to 2006, it includes *all* state-owned firms, and private firms with sales above RMB 5 million; from 2007 to 2010, it included all firms with sales above RMB 5 million; from 2011, it includes firms with sales above RMB 10 million.

We use the 2007 survey round for the analysis because (1) the sample no longer over-represents state-owned firms because smaller state-owned firms with sales below RMB 5 million were no longer grouped with firms with sales above RMB 5 million after 2007; this ensures that the sample is more balanced and the estimated outcomes are less biased, (2) key variables that are related to innovation are still available (they become unavailable after 2007), and (3) it features a good data quality.[5] We elaborate on our choice of the 2007 survey in more detail in section 5.

Next, we follow Jin (2023) and restrict the geographic scale of the sample to firms located in neighboring counties of a *CMC county*. By using a historical Chinese county map obtained from the China Historical Geographic Information System version 6 (CHGIS, 2016), and by using the information on the location of treaty ports and CMC customs stations from CMC (1907) and Huang (1917), we define a *CMC county* as a county that contained a CMC customs station that took over at least one Native Customs station in 1902. Then, we use GIS techniques to locate directly neighboring counties of CMC counties. We compare counties neighboring the same treaty port in this research because they should share similar geographic and pre-colonial socioeconomic characteristics and differ only in the quality of the customs institution. We also exclude any county in which

[5]For example, the 2008 survey misses a substantial proportion of firms. Key variables, especially those related to innovation and R&D, are missing in surveys after 2007. In general, the literature acknowledges the reliability of the data until the year 2007 (Nie and Yang, 2012; Xiao and Xu, 2018).

a treaty port was established[6] to exclude confounding effects stemming from treaty ports.[7] Based on the address of firms provided by the ASIF, we restrict our sample to firms that are located in the aforementioned neighboring counties, resulting in 41,987 industrial firms located in 114 historical counties included in the baseline sample.

3.2 Variables

Innovation. We measure firm innovation with the value of new products generated through innovative activities. According to Acs et al. (2002), the innovative process can be measured by inputs (e.g., investment in R&D and human capital), intermediate outputs (e.g., number of patented innovations), and final outputs. Measuring innovation with the final product of innovative activities has the advantage of capturing all possible elements during that process, including those that could not be easily captured such as commercialization and the generation of ideas. Moreover, we specifically choose the value of new products over the number of patents for two key reasons. First, patenting focuses more on the legal protection of inventions, which often fails to account for the broader value and economic impact of new technologies and innovations (Acs et al., 2002; Hall et al., 2001).[8] Secondly, firms, particularly leading innovators, face a trade-off between patenting and maintaining secrecy, which may result in underreporting of patent filings firms (Hall et al., 2014; Huang and Yuan, 2021; Mukherjee et al., 2017) and introduce a bias into regression estimates. Firms may also be disinclined to patent their innovations if they find it difficult to enforce claims resulting from patent infringements.

For the above reasons, we measure firm innovation by using the value of new products produced by each firm, normalized by the value of the total output of that firm. This variable is then called *innovation intensity* for the rest of the analysis. In the robustness tests, we further examine alternative measurements of innovation.

CMC Institution. The independent variable of interest, *CMC*, varies at the county level. It equals 1 if a county contained any Native Customs station that was taken over by the CMC in 1902 and 0 otherwise. According to CMC (1907) and Huang (1917), we identified 36 Native Customs stations in 16 counties from the neighboring county sample

[6] This naturally excludes all counties where a CMC station was established.

[7] Treaty ports may grow faster than other Chinese regions(Jia, 2014; Jin, 2023), potentially facilitating firm innovation in the long run.

[8] Another concern is that the effectiveness of a patent in copyright protection is heterogeneous across industries; for example, the violation of copyrights are more difficult to prove in high-tech industries (Chen et al., 2013).

that were taken over by the CMC.

Control Variables. We include firm-level control variables that are considered important determinants of innovation. We refrain from including a full set of firm characteristics provided by the ASIF to avoid obvious problems stemming from "bad controls". Instead, we control for the log of age, log of employment, capital intensity (the capital-labor ratio), and log of sales of a firm to capture key indicators of firm size. In addition, we control for ownership types of firms; these are state ownership, foreign ownership, Hong Kong/Taiwan/Macao ownership, domestic private ownership, and collective ownership. Finally, we capture regional agglomeration of industrial activities by controlling for the number of firms in the same county-industry cell of a given firm.

Despite the inclusion of control variables that mitigates the issue of omitted variable bias, distinguishing the effect of the CMC institution from the general spillover effect of treaty ports poses a challenge: counties closer to a treaty port may facilitate innovation better than remote regions, irrespective of their exposure to the CMC institution. While our sample design addresses this concern to some extent by including only counties with similar proximity to a treaty port, the potential confounding effect of treaty ports may still exist. Thus, we control for the log distance between a county and the nearest CMC station in all regression analyses, capturing the notion that the extent of the treaty port spillover effect diminishes with increasing distance to a treaty port.

3.3 Empirical Strategy

To identify the effect of the CMC institution on firm innovation, we restrict the sample to firms located in neighboring counties of a CMC station. Moreover, we control for the distance to the closest CMC station to capture any potential confounding effect stemming from treaty ports. Nevertheless, the takeover of the Native Customs may be determined by unobserved factors such as pre-colonial economic prosperity or geographic conditions, resulting in endogeneity issues and biased estimates.

To address the above concern, we follow the strategy in Jin (2023) and construct an instrumental variable that predicts the takeover of Native Customs stations at the county level.[9] As specified in section 2, a Native Customs station was taken over by the CMC if it

[9]It is important to state that a sharp regression discontinuity design (RDD) based on the boundary of CMC circles is not a suitable identification strategy here. A CMC circle determined only the takeover of Native Customs stations and its boundary did not necessarily generate a spatial discontinuity of economic prosperity.

was located within a 25 km radius of a CMC customs station. Thus, the more a county's area was covered by a circular area with a radius of 25 km around a CMC station, the more likely a Native Customs station was taken over in that county. Importantly, a prerequisite for such a positive relationship is that Native Customs stations should geographically spread out throughout a county, which was, in effect, the case and supported by historical documents.[10]

Following this logic, we apply GIS techniques and draw imaginary circular areas, referred to as *CMC circles*, with a radius of 25 km, each centered around a CMC station that took over Native Customs stations. Then, we calculate *CMC Coverage* for each county, which is the share of a county's area covered by CMC circles, as our instrumental variable for *CMC*. Using the 114 neighboring counties under investigation as a testing sample, we indeed find a positive and significant relationship between *CMC Coverage* and *CMC* with the coefficient being 0.91 and significant at the one percent level.[11]

Our baseline regressions include the following models:

$$CMC_c = \beta_0 + \beta_1 CMC\,coverage_c + \mathbf{X}'\beta + \delta_j + \delta_k + \epsilon_i \tag{1}$$

$$Y_i = \mu_0 + \mu_1 CMC_c + \mathbf{X}'\mu + \delta_j + \delta_k + \epsilon_i \tag{2}$$

where subscripts i, c, j, and k stand for a firm, a county, an industry, and a CMC station. Equation 1 is the first-stage regression where the variable of interest CMC is predicted by the instrumental variable $CMC\,coverage$.[12] In the second-stage regression described by Equation 2, the outcome variable Y_i is the innovation intensity of a firm i. $\mathbf{X}$ contains firm-, industry-, and county-level control variables listed in Table A1. δ_j represents industry fixed effects, and δ_k refers to CMC fixed effects capturing common characteristics in the area around a CMC station.

One may be concerned that the radius of 25 km is the result of a deliberate individual

[10]For example, Qi (2004) documents such spatial distribution of customs in the Fujian province.

[11]Even if we control for the log distance to the closest CMC station, the coefficient drops only to 0.800 and remains highly significant.

[12]In a recent study, Casey and Klemp (2021) discover an overestimation of regression coefficients in the economic history literature when the endogenous and key exogenous variables are both measured in the contemporary period, and the instrument is measured in the historical period. Our paper, however, does not present such an issue because both *CMC* and *CMC Coverage* are measured in the historical period. While our baseline IV estimation constitutes a "reduced form" analysis connecting historical institutions and contemporary outcomes, the mechanisms will be discussed in detail in section 6.

decision, and as such *CMC Coverage* may therefore be endogenous. For example, if the CMC set up the 25 km radius to target certain Native Customs stations with better performance, our instrument may no longer be valid. Historical evidence, however, shows that the radius is more likely an arbitrary decision drawn solely by Chinese officials who had limited information about the distribution of customs stations (General Administration of Customs, 2003; Tsai, 2008). The CMC officials also complained about the arbitrariness of the radius (CMC, 1907; General Administration of Customs, 2003). Finally, we argue that *CMC Coverage* is an exogenous measurement because it primarily depends on the county's shape and the relative position between a county and a CMC station.[13]

4 Baseline Results

The baseline results are presented in Table 1. Controlling for CMC fixed effects, the estimation in column (1) shows that firms located in counties affected by the CMC institution demonstrate a higher level of innovation intensity than other firms: the coefficient is 0.088 and is significant at the one percent level. When we control for industry fixed effects and firm-level controls in columns (2) and (3), the estimated coefficient reduces slightly to 0.085 and 0.080, respectively. This result suggests that firm and industrial characteristics may capture a small proportion of the CMC institutional effect on innovation; however, the coefficient of *CMC* remains highly significant and sizeable. Finally, we control for agglomeration in column (4) and the estimated effect reduces further to 0.078 with the statistical significance remaining at the one percent level. Regarding the magnitude of the effect, our preferred specification in column (4) suggests that the innovation intensity of firms in places affected by the CMC institution is on average higher than that in other firms by 8 ppt. This difference accounts for almost three times the sample mean (0.03) and more than half of the standard deviation of firm innovation intensity (0.14), suggesting a very substantial effect of the CMC institution on innovation in the long run.

Table A2 reports the full coefficients. We find that larger firms, as measured by sales and number of employment, demonstrate a higher level of innovation intensity than smaller firms, which is not surprising. Younger firms are more innovative than older firms, as presented by the negative and significant coefficient of *Age*. Finally, the coefficients of ownership dummies suggest that state-owned firms are on average more innovative than

[13]To further support the validity of the IV, Jin (2023) shows that *CMC Coverage* is unrelated to a large set of geographic and pre-colonial socioeconomic factors.

firms of all other types of ownership; this may be explained by the fact that state-owned firms are generally larger and more resourceful.

[Table 1 about here.]

5 Extensions and Robustness Tests

5.1 Alternative Innovation Variables

First, we show that our baseline results are robust to alternative innovation variables. In column (1) of Table 2, the outcome variable is a dummy variable that switches on when a firm's innovation intensity is above the industrial average, indicating that a firm is leading the technological progress of an industry. In column (2), we replace innovation intensity with a dummy variable indicating whether a firm produces any new product at all. In columns (3) and (4), innovation intensity is normalized by that firm's total sales and total revenue, respectively. The results presented in Table 2 show that the coefficient of *CMC* remains positive and significant regardless of how innovation intensity is defined.

[Table 2 about here.]

5.2 Intermediate Outcome Variables

In this subsection, we show that the historic CMC institution has a positive impact on present firms' innovation rates because it influenced firm behavior. Specifically, we expect firms in places affected by the CMC institution to invest more resources into R&D, and employee training, which eventually translate into more new products (Acemoglu, 1997; Parisi et al., 2006; Bauernschuster et al., 2009; Chen et al., 2013; Audretsch and Belitski, 2020).

To examine the long-run impact of the CMC institution on R&D, we analyze firm-level data on R&D expenditure and normalize it by each firm's revenue, referred to as *R&D intensity*. We also measure R&D spillover effects by calculating the average R&D expenditure of all other firms within the same county-industry cell. The results in column (1) of Table 3 demonstrate that firms in regions affected by the CMC institution invest more in R&D compared to firms in other areas. With an estimated coefficient of 0.004, the effect of the CMC institution on R&D expenditure is large, considering that the mean *R&D intensity* in the sample is only 0.002. In column (2), we include *R&D intensity* as an additional control variable in Equation 2. The results confirm a positive and significant association between R&D expenditure and firm innovation, which is expected. The estimated coefficient of *CMC* reduces to 0.071 compared to the baseline estimate in Table 1, suggesting that part of the CMC institutional effect on innovation is mediated by its influence on R&D expenditure. In column (3), we include R&D spillover instead of R&D intensity and find very similar results. Finally, we add both R&D intensity and R&D spillover in the regression, resulting in the coefficient of *CMC* decreasing to 0.066.

[Table 3 about here.]

Apart from R&D, investment in human capital is also crucial for firm innovation. The majority of human capital investment within firms is made in the form of training (Acemoglu, 1997): continuous training ensures that employees get access to updated knowledge, facilitating the possibility to innovate (Bauernschuster et al., 2009).[14] It is thus likely that firms in areas affected by the CMC institution spend more on employee training, resulting in better firm innovation. To test this hypothesis, we regress expenditure on training (normalized by revenue) on *CMC* in column (1) of Table 4. The estimated coefficient is very small in size and statistically insignificant, suggesting that the CMC is unrelated to the expenditure on training in relative terms. When the training expenditure is measured by its absolute term (in log) in column (2), however, the coefficient of *CMC* becomes highly significant.

The difference between columns (1) and (2) implies that *CMC* affected firm innovation through its influence on firm size. Specifically, firms in areas affected by the CMC institution tend to be larger, resulting in higher expenditure on training. This is supported by the evidence in column (3) when we regress firm revenue (in log) on *CMC*: the estimated coefficient is positive and significant, suggesting that *CMC* indeed affected firm size. Combining all evidence, we argue that the CMC institution fostered firms to grow larger. As a consequence, these firms spend more on training and demonstrate a higher innovation intensity. Notably, columns (4) and (5) show that both relative and absolute volumes of training expenditure are positively and significantly related to innovation intensity.

[Table 4 about here.]

[14] More studies in the business literature document a positive effect of training on firm innovation such as Shipton et al. (2006), Guisado-González et al. (2016), and Demirkan et al. (2022)

5.3 County-Level Control Variables

Finally, we address a possible concern that the transmission of the CMC institution might be associated with unobserved socioeconomic and geographic characteristics, resulting in biased estimates. In particular, neglecting co-determinants of innovation that are potentially related to *CMC Coverage* may lead to the violation of the exclusion restriction, invalidating the IV estimation. To address this issue, we control for important pre-colonial socioeconomic characteristics and geographic conditions in the baseline estimation. In column (1) of Table 5, we control for prefecture-level population density in 1820[15] and the distance to the closest prefecture capital. The estimation in column (2) adds basic geographic information including latitude, longitude, and the size of the county (in log). Column (3) includes location indicators, which capture the distance between a county's centroid and key geographic features such as the Yangtze River, the coastline, and the Grand Canal.[16] Finally, we include additional geographic characteristics that might be related to *CMC Coverage*, including elevation, terrain ruggedness, small river density, and a second order polynomial of geographic coordinates.[17] In Table 5, we progressively include the above variables from column (1) to column (4), and the coefficients of *CMC* remain highly significant.

[Table 5 about here.]

5.4 Sample Choice

This paper uses the cross-section data set of the year 2007 because it offers the best sample for our analysis for the following reasons. First, data sets after the survey round of 2007 are not considered because variables related to innovation are no longer available.[18] Second, the NBS altered the scope of the survey in 2007, switching it from *all* state-owned firms and above-scale (i.e. firms with sales above RMB 5 million) private firms to all above-scale firms, regardless of the ownership type. Thus, using a sample from surveys before 2007 may result

[15] Chinese population data before 1840 at the prefecture level is documented by Cao (2000).

[16] Yangtze River is the most navigable river in China. The Grand Canal was the most important grain transportation route, which is closely related to political stability in historical China (Cao and Chen, 2022).

[17] Elevation data is obtained from NASA's Shuttle Radar Topography Mission (SRTM) and terrain ruggedness data set is from Nunn and Puga (2012). We calculate the average elevation and terrain ruggedness of a county using all grids within that county.

[18] The 2004 survey also misses innovation variables and is thus not considered by this study. Moreover, R&D and training expenditures are only available for the period from 2005 to 2007.

in an over-representation of state-owned firms, potentially biasing the estimates. The 2007 survey, on the other hand, offers a unique sample where state-owned and private firms are selected by the same standard while innovation variables are still available. Third, the 2007 survey represents the best data quality. For example, when considering the mismatch of a firm's registration type and its ownership type as an indication of low-quality observation (cf. Nie and Yang, 2012),[19] Table A3 shows that such mismatch cases are the lowest in the 2007 survey. To sum up, the use of the 2007 survey as a cross-sectional data set apparently serves best our study.

Nevertheless, we estimate the effect of the CMC institution on innovation using a larger data set for the period 2000-2007 as a robustness check. Due to the inclusion of multiple years in the regression analysis, we normalize the value of the fixed asset and the total sales by the fixed asset investment price index and production price index, respectively; both indexes are obtained from the National Bureau of Statistics of China and are constructed to normalize the price at the 2007 level. It is important to notice that this dataset, albeit larger, suffers not only the aforementioned quality issues but also inconsistent measurement of key variables. For example, employment is measured by end-of-year figures for 2000, 2005, and 2007, and by the yearly average for the rest of the period. With these limitations in mind, the results reported in Table 6 show that firms in places affected by the CMC institution are more innovative than firms in other places. Notably, the smaller magnitude of the effect reported here may very likely be the result of low data quality and sample selection bias. Overall, our results are confirmed in the larger sample, data quality issues notwithstanding.

[Table 6 about here.]

6 Reducing Corruption as the Main Channel

6.1 CMC, Corruption, and Innovation

Why does the CMC have an impact on today's firm innovation even though its institution was long abolished? We argue that the norms of honesty and lawfulness embedded in the business practice under the CMC rule is the most plausible explanation. The reform in

[19] This refers to the issue that a firm with less than 50% of its capital owned by the state reports its registration type as state-owned. We also examine similar mismatch cases for foreign and Hong Kong/Macau/Taiwan firms.

1902 in Native Customs stations supervised by the CMC prohibited corrupt practices while promoting honesty and compliance. This drastic institutional shift reshaped the interaction between businesses and government in the affected areas, establishing a self-enforcing local equilibrium that continues to affect today's firm behavior. For example, using data from the 2005 Investment Climate Survey published by the World Bank, Jin (2023) shows that firms affected by the CMC institution are less corruptible than unaffected firms: they spend less time with governmental assignments, and are less likely to bribe for loans or sign informal contracts.

There is little doubt that corruption impedes innovation. Corruption leads to an inefficient allocation of talent as more talent is invested in dealings with the government than in a non-corrupt environment. This leads to a reduction in the available talent for innovation and thereby hindering technological progress (Murphy et al., 1991). Moreover, in a corrupt environment, government officials pose a heightened risk of expropriating innovators through actions like increasing the costs and uncertainty of granting necessary permits and licenses. Innovators, who usually have inelastic demand for these governmental services, are thus more susceptible to expropriation (Xu and Yano, 2017). Corrupt bureaucrats may increase bureaucratic complexity and regulatory uncertainty in order to extract more resources from the business sector. Returns to investment in innovation may diminish and become even more uncertain. Managers may prioritize rent-seeking activities over investing in R&D when the relative payoff of the former is higher than the latter, i.e., when the local business environment is more corruptible (Baumol, 1996; Murphy et al., 1991; Xu and Yano, 2017).

The presence of an anti-corruption culture fostered by the CMC institution is expected to discourage rent-seeking and expropriation, thereby reducing the uncertainty associated with innovation. As a result, firms operating in areas with stronger anti-corruption norms are likely to experience higher returns to innovation and consequently increase their investment in innovation-related activities such as R&D and training. This, in turn, leads to a higher level of innovation in those firms (and ultimately to stronger firm growth).

Ideally, we would have tested the aforementioned hypothesis directly with a geo-referenced firm-level dataset that contained information on corruption-related behavior or perceptions of corruption. Unfortunately, such data is unavailable and the ASIF does not contain information on corruption. Therefore, we provide three pieces of evidence - each utilizing a unique measurement of the corruption environment - that indirectly explain why the anti-corruption culture is very likely the key mechanism here.

6.2 CMC and the Anti-Corruption Campaign

We show that the CMC institution is related to a reduced level of regional corruption. This first step is necessary to build a solid connection between historical and contemporary variables that are conceptually close (Voth, 2021), offering more compelling evidence of the persistence effect of the CMC institution. To that end, we use the data set from Wang and Dickson (2022) on corruption investigation cases from 2012 to 2015.[20] The period under investigation here is at the height of China's anti-corruption campaign, which was initiated by President Xi after he assumed office.[21] Given the fact that reliable data on county-level corruption cases around the year 2007 do not exist, we assume that the corruption pattern revealed during the anti-corruption campaign is not systematically different from the one in 2007 (the year for our baseline analysis). This seems a reasonable assumption.

We aggregate the total number of corruption cases from the data set at the county-year level, restricting our analysis to those counties that can be geographically matched to the neighboring counties in our baseline sample; this leads to 147 contemporary counties in the data set. Next, we obtain the grid-level population data from *WorldPop*[22] to calculate a county-year panel data set on the Chinese population, which is used for calculating the per capita corruption cases. Finally, we match each contemporary county in the corruption data set to a historical county, and hence the associated key variables, in our baseline data set.

[Figure 1 about here.]

As summarized in Figure 1, we compare the yearly total and per capita corruption cases of counties affected by the CMC institutions with those that were not, from 2012 to 2015. While apparently the total and per capita number of corruption cases increased sharply since the anti-corruption campaign was launched, the increase was much steeper for non-CMC counties than for CMC counties.

[20] The original data source was at *http://news.qq.com/zt2016/fanfu˙ccdi/index.html* (accessed 29. January 2024), which is no longer available (as of February 2024). According to Wang and Dickson (2022), the database was organized by Tencent, the biggest Chinese internet company, that allowed visitors to search every corruption investigation since 2011. It collected data from the Party and governmental departments at every administrative level. We dropped the year 2016 from the data set because the collection ended in August 2016, leading to fewer observations than other years.

[21] This campaign is widely acknowledged as the most intensive one so far (Wedeman, 2016). More than 600,000 officials had been removed as of 2021, and it affects officials at all governmental levels.

[22] Available at *https://www.worldpop.org*.

To investigate this difference more concretely, we regress the county-year per capita corruption cases on *CMC*, and report the results in Table 7. In column (1), the negative and significant coefficient confirms the pattern shown in Figure 1: counties affected by the CMC institution suffered less from corruption. Moreover, the coefficient remains stable even if we control progressively for the distance to the closest CMC station (columns (2)-(4)), year fixed effects (columns (3) and (4)), and province fixed effects (column (4)). Our preferred specification in column (4) suggests that the CMC institution explains a sizable decrease in corruption cases per capital: the coefficient -0.01 accounts for almost half of the sample mean (0.024) and a quarter of the standard deviation (0.04). Our evidence thus suggests that the CMC institution has led to a significant and substantial overall decrease in corruption at the local level.

[Table 7 about here.]

6.3 Corruption and Ownership Type

We investigate next how corruption is related to the pro-innovation effect of the CMC institution. Specifically, if the pro-innovation effect of the CMC institution is due to a reduction in corruption, it should be more pronounced among those firm types that are typically - outside of the area of CMC influence - more exposed to corruption. One distinct determinant of exposure to corruption is the ownership type. There is ample evidence that private firms in China often face stricter regulations and financial constraints, making especially domestically owned private firms more susceptible to engaging in bribery and seeking political favors (Jiang and Nie, 2014; Poncet et al., 2010).[23] Moreover, private firms are more vulnerable to expropriation, as they lack the protection and resources that state-owned firms may possess. To test this hypothesis, we analyze the heterogeneous impact of the CMC institution on innovation based on the ownership types. We divide the sample into three subgroups: state-owned enterprises (SOEs), domestic private firms, and foreign-owned firms.

Results are reported in Table 8. We find that *CMC* has a significantly positive effect only for domestic private firms (column 2), but not for the other ownership types. This finding supports our notion that the CMC institution promotes innovation effectively for those firms that regard corruption as a major barrier to investing in innovation, i.e. domestic private firms as opposed to firms of other ownership types. Empirical evidence from

[23] Foreign-owned firms may be more mobile ad react to extortion by relocating.

similar studies aligns with the above argument. For example, Paunov (2016) demonstrates that corruption reduces innovation for domestic private firms but not for public and foreign ones. Zakharov (2019) shows, for Russia, that corruption reduces investment in private firms while not significantly impacting firms owned by the state.

[Table 8 about here.]

6.4 Corruption and Industry

To further validate the anti-corruption channel of the CMC institution, we make use of the sector as a second source of variation in corruption levels experienced by firms. To that end we create an industrial corruption index from the 2005 World Bank Investment Climate Survey. We use the following information from this survey as components of the index: (1) the tendency of bribery, which is derived from the question *"Is there a need for informal payment to staff from the banks or loan-providing institutions?"*. (2) the frequency of building connections with public officials, which is derived from the question *"How many days does the GM or Vice GM spend on the government assignments and communications per month?"*.[24] (3) Experience of signing informal contracts, which is derived from the question *"Does your company usually sign formal contracts with the client/supplier?"*. (4) the extent of red tape and legal restrictions, which is taken from the question *"How many licenses and registrations (permanent and renewable annually) are required for your company?"*. (5) expenditure on entertainment and travel as the share of total sales, which is considered a relevant indicator of corruption in Chinese firms and is negatively related to firm performance (Cai et al., 2011).

We calculate the corruption index for each industry by summing up the average value for each of the aforementioned corruption indicators across firms in that industry. Notably, we exclude firms in areas affected by the CMC institution when calculating the corruption index; this ensures that our industrial measurement of corruption is orthogonal to *CMC* and hence exogenous.[25] After matching industries between the ASIF and the World Bank Enterprise Survey, each firm in our baseline sample is assigned its industrial corruption

[24] We define "frequently" by creating a dummy indicating more than a week per month.

[25] Another concern is that the World Bank Enterprise Survey may have surveyed more firms in places affected by the CMC than firms in other regions because these places are more developed, resulting in a sample selection bias when constructing the corruption index at the industrial level. However, we show that this is not the case. As presented in Table A4, we find no difference in the number of firms between CMC- and non-CMC-affected counties.

index. This enables us to re-estimate the baseline regression model using sub-samples of firms categorized by the quintile of the corruption index: the first quintile represents the lowest corruption level whereas the fifth quintile indicates the highest corruption level.

As reported in Table 9, the results show that the effect of the CMC institution on firm innovation is insignificant for firms operating in less corruptible industries, as indicated by insignificant coefficients for CMC in columns (1) to (4).[26] Moreover, point estimates increase from (1) to (5) almost monotonically. For firms operating in the highest quintile of corrupt environments, we find a positive and significant coefficient for *CMC*, suggesting that the CMC institution facilitates firm innovation especially in industries characterized by a high level of corruption (column 5).[27] This again suggests that the innovation stimulus of the CMC institution is brought about by its corruption-reducing effect.

[Table 9 about here.]

7 Possible Confounding Factors

7.1 Competition

Could the CMC institution affect innovation through channels other than persistent anti-corruption norms? One potential argument is that the CMC institution may have persistently promoted competition in local industries and hence affect today's innovation. Existing studies on competition and innovation offer mixed evidence: while Aghion et al. (2005) document an inverted-U shape relationship of the two in the U.K., others find such a relationship sensitive to the period, the country, and the types of firms being studied (Correa, 2012; Hashmi, 2013; Askenazy et al., 2013; Mulkay, 2019). To measure competition, we generate the Herfindahl-Hirschman index (HHI) for each county-industry cell and include it as an additional industry-level control in the baseline regression. The results presented in column (1) of panel A in Table 10 show that the coefficient of the HHI is statistically insignificant, suggesting that competition does not significantly affect firm innovation in this context.

Additionally, we calculate the concentration ratio (CR) based on the sales of the top 10, 30, and 50 firms within a county-industry cell. The results in columns (2)-(3) of panel

[26]Note that the number of observations is somewhat different across columns. This is because the corruption index varies at the industry level.

[27]Alternatively we categorized the corruption index into deciles, which did not alter our results in any significant way.

A indicate that CR is negatively but insignificantly associated with innovation. To further test for a possible inverted-U shape relationship between competition and innovation, panel B replicates the estimations in Panel A including also a quadratic form of competition variables. The results, however, do not support the hypothesis: the coefficients of the competition variables and their squared terms are insignificant, and the coefficient of *CMC* remains largely unchanged. These findings suggest a limited role played by competition in explaining the relationship between the CMC institution and firm innovation in our empirical setup.

[Table 10 about here.]

7.2 Foreign Direct Investment

Next, we test whether FDI could be an alternative channel. FDI facilitates the transfer of technology and know-how from foreign investors to domestic firms, which may improve innovation capabilities and eventually enhance innovation rates (Hu and Jefferson, 2009; Liu and Buck, 2007; Chen et al., 2022). In addition, FDI tends to affect innovation through spillover channels (Blomström and Sjöholm, 1999; Cheung and Ping, 2004; Kugler, 2006). Reported in column (1) of Table 11, we find that firms in counties affected by the CMC institution receive a significantly larger amount of FDI (in logs) than other firms. However, when we explore the impact of FDI on firm innovation, as shown in column (2), we do not find any significant relationship. Similarly, the inclusion of FDI spillover effects, measured by the average FDI received by all other firms within the same county-industry cell, shows no significant result (column 3). Also including FDI and FDI spillover effects simultaneously yields no significant results for these variables and leaves the coefficient for CMC unaffected (Column 4).

Thus, while firms in CMC-affected regions may attract more FDI, this additional foreign investment does not contribute significantly to firm innovation. FDI is therefore unlikely to confound the anti-corruption channel of the CMC effect.

[Table 11 about here.]

7.3 Trade

The transfer of knowledge induced by an open economy may also take place through trade (Lileeva and Trefler, 2010; Aw et al., 2011; Shu and Steinwender, 2019; Coelli et al., 2022).

Thus, we next examine whether trade openness can be a potential further channel through which CMC has led to higher innovation rates. We measure a firm's trade by the value of its reported exports normalized by its revenue – firms with a higher value of *Trade* are thus more export-oriented. As reported in column (1) of Table 12, firms located in CMC-affected regions do not exhibit significantly higher levels of trade, suggesting that trade is unlikely a channel that explains the impact of the CMC institution on firm innovation. In columns (3) and (4), we estimate the effect of the CMC institution on firm innovation after controlling for trade, and the coefficient of *CMC* remains significant and quantitatively similar to the baseline estimation. We also find that trade is positively related to innovation (column 2); trade conducted by other firms in the same county-industry also affects the innovation of a particular firm under consideration (columns 3 and 4). The overall evidence suggests that any potential effect of trade on innovation is orthogonal to the influence of the CMC institution, and the anti-corruption culture stemming from the CMC institution hence remains a robust mechanism.

[Table 12 about here.]

7.4 Heterogeneous Effects of Potential Confounders?

Finally, we show that the potential confounding factors – competition, FDI, and trade – do not respond heterogeneously to different corruption levels in their effect on innovation. This finding lends further support to our conclusion that the anti-corruption channel matters for the effect of CMC on innovation.

If, for example, competition affected innovation positively only in a low corruption environment, our corruption index would capture not only corruption as such but also the effectiveness of competition. To investigate such a possibility, we include each of the above factors as an additional control variable individually and estimate each model for the sub-samples of the quintile of firms with the lowest and highest levels of industrial corruption as defined in subsection 6.4, respectively.

The results reported in Table 13 align with our expectation: the coefficient of *HHI*, *FDI*, and *Trade* do not exhibit differential responses for the different levels of corruption. Specifically, competition and FDI do not have a significant effect on innovation, regardless of the industrial corruption level is high or low (columns 1-2 and 5-6). Similarly, export is positively and significantly associated with innovation in both high- and low-corruption environments (columns 3 an 4). Moreover, the coefficient of *CMC* remains insignificant for low-corruption estimations (columns 1, 3, and 5) and significant for high-corruption environments (columns 2, 4 and 6). These results suggest that competition, FDI, and trade do not exhibit a heterogeneous effect on innovation based on different corruption levels, and that the anti-corruption effect of the CMC institution is likely to be the central mechanism.

[Table 13 about here.]

8 Conclusion

This paper suggests that firm innovation in China is likely affected by historical institutions that promote efficiency, transparency, and honesty. In particular, we show that firms in places affected by the Chinese Maritime Customs (CMC), a foreign-run institution known for its transparency, efficiency and rule-based approach, exhibit a higher innovation intensity than firms in similar but unaffected places. To identify causality, we exploit the historical event that the CMC transmitted its administrative guidelines to part of Native Customs stations. This allows us, also with the help of an instrumental variable strategy, to compare counties affected with those unaffected by the take over of Native Customs stations.

Our results further suggest that the embedded cultural norm of honesty and lawfulness is a plausible channel to explain the long-term effect of the CMC institution: we show that the pro-innovation effect of the CMC institution is strong and significant for firms exposed to a highly corruptible environment and becomes insignificant in low-corruption situations. This finding remains robust after taking alternative channels into account such as competition, FDI, and trade.

Innovation plays a pivotal role in driving economic growth, as reflected in its ability to spur productivity gains, enhance market competitiveness, and foster technological advancements. Our analysis, which accurately identifies the impact of historical institutions on innovation through cultural channels, provides valuable insights for policy considerations: we highlight the significance of governance quality and the rule of law in combating corruption and promoting innovation. Additionally, we emphasize the equally important role of social norms fostered by these institutions, which amplify the positive effect of institutions on innovation. These findings underscore the multifaceted nature of institutions that contribute to a favorable environment for economic development.

References

Acemoglu, Daron (1997). Training and innovation in an imperfect labour market. *The Review of Economic Studies 64*(3), 445–464.

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics 117*(4), 1234–1294.

Acs, Zoltan J., Luc Anselin, and Attila Varga (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy 31*(7), 1069–1085.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt (2005). Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Aidt, Toke S. (2009). Corruption, institutions, and economic development. *Oxford Review of Economic Policy 25*(2), 271–291.

Andrews, Michael J. and Alexander Whalley (2022). 150 years of the geography of innovation. *Regional Science and Urban Economics 94*, 103627.

Anokhin, Sergey and William S. Schulze (2009). Entrepreneurship, innovation, and corruption. *Journal of Business Venturing 24*(5), 465–476.

Askenazy, Philippe, Christophe Cahn, and Delphine Irac (2013). Competition, R&D, and the cost of innovation: Evidence for France. *Oxford Economic Papers 65*(2), 293–311.

Audretsch, David B. and Maksim Belitski (2020). The role of R&D and knowledge spillovers in innovation and productivity. *European Economic Review 123*, 103391.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu (2011). R&D investment, exporting, and productivity dynamics. *American Economic Review 101*(4), 1312–1344.

Babina, Tania, Asaf Bernstein, and Filippo Mezzanotti (2023). Financial disruptions and the organization of innovation: Evidence from the great depression. *The Review of Financial Studies 36*(11), 4271–4317.

Bauernschuster, Stefan, Oliver Falck, and Stephan Heblich (2009). Training and innovation. *Journal of Human Capital 3*(4), 323–353.

Baumol, William J. (1996). Entrepreneurship: Productive, unproductive, and destructive. *Journal of Business Venturing 11*(1), 3–22.

Becker, Sascha O., Katrin Boeckh, Christa Hainz, and Ludger Woessmann (2016). The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy. *Economic Journal 126*(590), 40–74.

Blomström, Magnus and Fredrik Sjöholm (1999). Technology transfer and spillovers: Does local participation with multinationals matter? *European Economic Review 43*(4-6), 915–923.

Brandt, Loren, Debin Ma, and Thomas G. Rawski (2014). From divergence to convergence: Reevaluating the history behind China's economic boom. *Journal of Economic Literature 52*(1), 45–123.

Cai, Hongbin, Hanming Fang, and Lixin Colin Xu (2011). Eat, drink, firms, government: An investigation of corruption from the entertainment and travel costs of Chinese firms. *The Journal of Law and Economics 54*(1), 55–78.

Cao, Shuji (2000). *History of population in China, volume 5*. Shanghai: Fudan University Press.

Cao, Yiming and Shuo Chen (2022). Rebel on the canal: Disrupted trade access and social conflict in China, 1650–1911. *American Economic Review 112*(5), 1555–1590.

Casey, Gregory and Marc Klemp (2021). Historical instruments and contemporary endogenous regressors. *Journal of Development Economics 149*, 102586.

Chen, Sheng-Syan, Yan-Shing Chen, Woan-lih Liang, and Yanzhi Wang (2013). R&D spillover effects and firm performance following R&D increases. *Journal of Financial and Quantitative Analysis 48*(5), 1607–1634.

Chen, Ting, James Kai-sing Kung, and Chicheng Ma (2020). Long live Keju! The persistent effects of China's civil examination system. *Economic Journal 130*(631), 2030–2064.

Chen, Yongmin, Haiwei Jiang, Yousha Liang, and Shiyuan Pan (2022). The impact of foreign direct investment on innovation: Evidence from patent filings and citations in China. *Journal of Comparative Economics 50*(4), 917–945.

Chen, Yuyu, Hui Wang, and Se Yan (2022). The long-term effects of Protestant activities in China. *Journal of Comparative Economics 50*(2), 394–414.

Cheung, Kui-yin and Lin Ping (2004). Spillover effects of FDI on innovation in China: Evidence from the provincial data. *China Economic Review 15*(1), 25–44.

CHGIS (2016). *China Historical Geographic Information System.* Cambridge: Harvard Yenching Institute.

Cinnirella, Francesco and Jochen Streb (2017). The role of human capital and innovation in economic development: evidence from post-Malthusian Prussia. *Journal of Economic Growth 22*, 193–227.

CMC (1907). *Native Customs Trade Returns.* Shanghai: Chinese Maritime Customs, Statistical Department.

Coelli, Federica, Andreas Moxnes, and Karen Helene Ulltveit-Moe (2022). Better, faster, stronger: Global innovation and trade liberalization. *Review of Economics and Statistics 104*(2), 205–216.

Correa, Juan A. (2012). Innovation and competition: An unstable relationship. *Journal of Applied Econometrics 27*(1), 160–166.

Dai, Yifeng (1989). On the takeover of Chinese Native Customs in the late Qing period. *History Study 6*, 95–108.

Danzer, Alexander, Natalia Danzer, and Carsten Feuerbaum (2023). Military spending and innovation: Learning from 19th century world fair exhibition data. *CESifo Working Paper No. 10347*.

Dell, Melissa (2010). The persistent effects of Peru's mining mita. *Econometrica 78*(6), 1863–1903.

Dell, Melissa, Nathan Lane, and Pablo Querubin (2018). The historical state, local collective action, and economic development in Vietnam. *Econometrica 86*(6), 2083–2121.

Demirkan, Irem, Ravi Srinivasan, and Alka Nand (2022). Innovation in SMEs: The role of employee training in German SMEs. *Journal of Small Business and Enterprise Development 29*(3), 421–440.

Deng, Yibing (2007). Chinese Native Customs in the early Qing period of China. *Studies in Qing History 2*, 54–64.

Dimant, Eugen and Guglielmo Tosato (2018). Causes and effects of corruption: What has past decade's empirical research taught us? A survey. *Journal of Economic Surveys 32*(2), 335–356.

Dincer, Oguzhan (2019). Does corruption slow down innovation? Evidence from a cointegrated panel of US states. *European Journal of Political Economy 56*, 1–10.

Donges, Alexander, Jean-Marie Meier, and Rui C. Silva (2023). The impact of institutions on innovation. *Management Science 69*(4), 1951–1974.

Ellis, Jesse, Jared Smith, and Roger White (2020). Corruption and corporate innovation. *Journal of Financial and Quantitative Analysis 55*(7), 2124–2149.

Engerman, Stanley L. and Kenneth L. Sokoloff (1997). Factor endowments, institutions, and differential paths of growth among new world economies: A view from economic historians of the United States. In Stephen Haber (Ed.), *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914*, pp. 260–304. Stanford: Stanford University Press.

Fang, Lily, Josh Lerner, Chaopeng Wu, and Qi Zhang (2018). Corruption, government subsidies, and innovation: Evidence from China. *NBER Working Paper* (25098).

General Administration of Customs (2003). *The Selection of Inspector-General's Circluars in the Chinese Maritime Customs*. Beijing: The Chinese Customs Press.

Gross, Daniel P. and Bhaven N. Sampat (2023). America, jump-started: World War II R&D and the takeoff of the US innovation system. *American Economic Review 113*(12), 3323–3356.

Guisado-González, Manuel, Mercedes Vila-Alonso, and Manuel Guisado-Tato (2016). Radical innovation, incremental innovation and training: Analysis of complementarity. *Technology in Society 44*, 48–54.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2016). Long-term persistence. *Journal of the European Economic Association 14*(6), 1401–1436.

Hall, Bronwyn, Christian Helmers, Mark Rogers, and Vania Sena (2014). The choice between formal and informal intellectual property: A review. *Journal of Economic Literature 52*(2), 375–423.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg (2001). The NBER patent citation data file: Lessons, insights and methodological tools.

Hashmi, Aamir R. (2013). Competition and innovation: The inverted-U relationship revisited. *Review of Economics and Statistics 95*(5), 1653–1668.

Hu, Albert Guangzhou and Gary H. Jefferson (2009). A great wall of patents: What is behind China's recent patent explosion? *Journal of Development Economics 90*(1), 57–68.

Huang, Qianqian and Tao Yuan (2021). Does political corruption impede firm innovation? Evidence from the United States. *Journal of Financial and Quantitative Analysis 56*(1), 213–248.

Huang, Xuyuan (1917). *The History of Chinese Customs*. Beijing: Gonghe Press.

Jedwab, Remi, Edward Kerby, and Alexander Moradi (2017). History, path dependence and development: Evidence from colonial railways, settlers and cities in Kenya. *The Economic Journal 127*(603), 1467–1494.

Jia, Ruixue (2014). The legacies of forced freedom: China's treaty ports. *The Review of Economics and Statistics 96*(4), 596–608.

Jiang, Ting and Huihua Nie (2014). The stained China miracle: Corruption, regulation, and firm performance. *Economics Letters 123*(3), 366–369.

Jin, Gan (2023). Circle of fortune: The long-term impact of Western customs institution in China. *Journal of Development Economics 163*, 103075.

Jin, Gan and Günther Schulze (2024). Historical Legacies and Urbanization: Evidence from Chinese Concessions. *Working paper, University of Freiburg*.

Keller, Wolfgang, Ben Li, and Carol H. Shiue (2011). China's foreign trade: Perspectives from the past 150 years. *The World Economy 34*(6), 853–892.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Kugler, Maurice (2006). Spillovers from foreign direct investment: Within or between industries? *Journal of Development Economics 80*(2), 444–477.

Lambsdorff, Johann G. and Günther G. Schulze (2015). What can we know about corruption? A very short history of corruption research and a list of what we should aim for. *Jahrbucher fur Nationalokonomie und Statistik 235*(2), 100–114.

Lee, Chien-Chiang, Chih-Wei Wang, and Shan-Ju Ho (2020). Country governance, corruption, and the likelihood of firms' innovation. *Economic Modelling 92*, 326–338.

Liao, Shengfeng (2010). *Chinese Native Customs and Regional Economy*. Beijing: Renmin Press.

Lileeva, Alla and Daniel Trefler (2010). Improved access to foreign markets raises plant-level productivity... for some plants. *The Quarterly Journal of Economics 125*(3), 1051–1099.

Liu, Xiaohui and Trevor Buck (2007). Innovation performance and channels for international technology spillovers: Evidence from Chinese high-tech industries. *Research Policy 36*(3), 355–366.

Lowes, Sara, Nathan Nunn, James A. Robinson, and Jonathan L. Weigel (2017). The evolution of culture and institutions: Evidence from the Kuba Kingdom. *Econometrica 85*(4), 1065–1091.

Mauro, Paolo (1995). Corruption and growth. *The Quarterly Journal of Economics 110*(3), 681–712.

Michalopoulos, Stelios and Elias Papaioannou (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica 81*(1), 113–152.

Moser, Petra, Alessandra Voena, and Fabian Waldinger (2014). German Jewish émigrés and US invention. *American Economic Review 104*(10), 3222–3255.

Mukherjee, Abhiroop, Manpreet Singh, and Alminas Žaldokas (2017). Do corporate taxes hinder innovation? *Journal of Financial Economics 124*(1), 195–221.

Mulkay, Benoît (2019). How does competition affect innovation behaviour in French firms? *Structural Change and Economic Dynamics 51*, 237–251.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny (1991). The allocation of talent: Implications for growth. *The Quarterly Journal of Economics 106*(2), 503–530.

Nie, Huihua, Ting Jiang and Rudai Yang (2012). The current stage and potential issues of the Chinese Annual Survey of Industrial Firms. *Journal of World Economy* (5), 142–158.

Nunn, Nathan (2009). The importance of history for economic development. *Annual Review of Economics 1*(1), 65–92.

Nunn, Nathan and Diego Puga (2012). Ruggedness: The blessing of bad geography in Africa. *The Review of Economics and Statistics 94*(1), 20–36.

Parisi, Maria Laura, Fabio Schiantarelli, and Alessandro Sembenelli (2006). Productivity, innovation and R&D: Micro evidence for Italy. *European Economic Review 50*(8), 2037–2061.

Paunov, Caroline (2016). Corruption's asymmetric impacts on firm innovation. *Journal of Development Economics 118*, 216–231.

Poncet, Sandra, Walter Steingress, and Hylke Vandenbussche (2010). Financial constraints in China: Firm-level evidence. *China Economic Review 21*(3), 411–422.

Qi, Meiqin (2004). *A Study in Qing's Native Customs.* Huhehaote: University of Inner Mongolia Press.

Shipton, Helen, Michael A. West, Jeremy Dawson, Kamal Birdi, and Malcolm Patterson (2006). HRM as a predictor of innovation. *Human Resource Management Journal 16*(1), 3–27.

Shleifer, Andrei and Robert W. Vishny (1993). Corruption. *The Quarterly Journal of Economics 108*(3), 599–617.

Shu, Pian and Claudia Steinwender (2019). The impact of trade liberalization on firm productivity and innovation. *Innovation Policy and the Economy 19*(1), 39–68.

Solow, Robert M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics 39*(3), 312–320.

Tsai, Weipin (2008). The Inspector General's last prize: The Chinese Native Customs Service, 1901–31. *The Journal of Imperial and Commonwealth History 36*(2), 243–258.

Voigtlander, Nico and Hans-Joachim Voth (2012). Persecution perpetuated: The medieval origins of anti-Semitic violence in Nazi Germany. *Quarterly Journal of Economics 127*(3), 1339–1392.

Voth, Hans-Joachim (2021). Persistence–myth and mystery. In Alberto Bisin and Giovanni Federico (Eds.), *The Handbook of Historical Economics*, pp. 243–267. Amsterdam: Elsevier.

Wang, Yuhua and Bruce J. Dickson (2022). How corruption investigations undermine regime support: Evidence from China. *Political Science Research and Methods 10*(1), 33–48.

Wedeman, Andrew (2016). Four years on: where is Xi Jinping's anti-corruption drive headed. *China Policy Institute: Analysis*.

Wright, Stanley F. (1950). *Hart and the Chinese customs*. Belfast: Mullan.

Xiao, Liping and Weifeng Xu (2018). The application of the database of Chinese Annual Survey of Industrial Firms to enterprise economic behavior research: Perspectives, integration and extension. *Foreign Economics & Management 40*(3), 137–152.

Xu, Gang and Go Yano (2017). How does anti-corruption affect corporate innovation? Evidence from recent anti-corruption efforts in China. *Journal of Comparative Economics 45*(3), 498–519.

Zakharov, Nikita (2019). Does corruption hinder investment? Evidence from Russian regions. *European Journal of Political Economy 56*, 39–61.

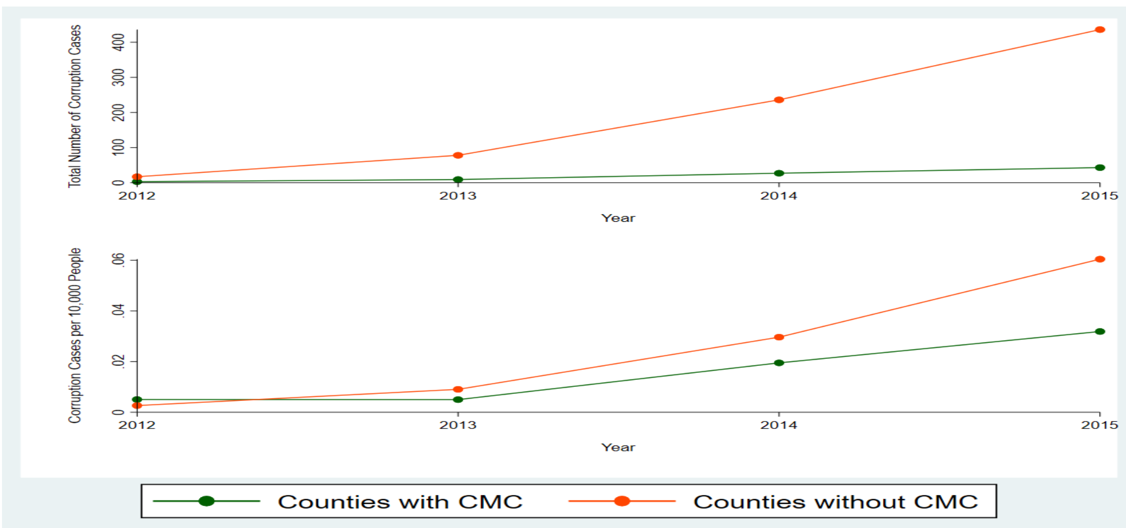

Figure 1: Comparing Corruption Cases, 2012-2015

Table 1: Baseline Results

DV: Innovation Intensity (New product output/total output)				
	(1)	(2)	(3)	(4)
CMC	0.088***	0.085***	0.080***	0.078***
	(0.034)	(0.030)	(0.028)	(0.030)
Dis_CMC	0.038***	0.037***	0.036***	0.036***
	(0.009)	(0.008)	(0.008)	(0.008)
N	41987	41987	41987	41987
R-squared	0.021	0.033	0.050	0.050
Fstat	14.674	14.617	14.780	14.531
Firm controls	No	No	Yes	Yes
Agglomeration	No	No	No	Yes
CMC FE	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes

Notes: The table reports IV estimates. *CMC* is instrumented by *CMC Coverage*. *Fstat* is the Kleibergen-Paap rk Wald F statistic. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), and ownership type. *Agglomeration* is the number of firms in a county-industry cell. Robust errors adjusted for clustering at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 2: Alternative Innovation Variables

	Innovation (abv. mean, dmy)	Innovation (dummy)	Innovation intensity over sales	over income
	(1)	(2)	(3)	(4)
CMC	0.118**	0.146**	0.083***	0.078***
	(0.057)	(0.062)	(0.031)	(0.030)
N	41987	41987	41979	41984
R-squared	0.094	0.104	0.046	0.046
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 3: CMC Institution and R&D

DV	R&D intensity	Innovation intensity		
	(1)	(2)	(3)	(4)
CMC	0.004*	0.071***	0.072**	0.066**
	(0.002)	(0.027)	(0.028)	(0.026)
R&D intensity		1.876***		1.866***
		(0.718)		(0.712)
R&D spillover			1.435**	1.143**
			(0.661)	(0.557)
N	41984	41984	41987	41984
R-squared	0.025	0.074	0.054	0.077
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates.*Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.

***, ** and * indicate significance at 1%, 5% and 10%

Table 4: CMC Institution, Employee Training, and Firm Size

DV:	Training exp. over revenue	Training exp. in log	Revenue in log	Innovation intensity	
	(1)	(2)	(5)	(3)	(4)
CMC	-0.000	1.020*	0.099**	0.080***	0.078***
	(0.000)	(0.554)	(0.038)	(0.028)	(0.027)
Training exp. over revenue				0.550**	
				(0.230)	
Training exp. (in log)					0.002***
					(0.000)
N	41984	41987	41987	41984	41987
R-squared	0.038	0.174	0.946	0.050	0.052
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 5: Additional Control Variables

DV: Innovation Intensity				
	(1)	(2)	(3)	(4)
CMC	0.119***	0.080***	0.071**	0.103***
	(0.026)	(0.022)	(0.030)	(0.027)
N	41987	41987	41987	41987
R-squared	0.027	0.055	0.060	0.054
Pre-condition	Yes	Yes	Yes	Yes
Geography	No	Yes	Yes	Yes
Location	No	No	Yes	Yes
Add. geography	No	No	No	Yes
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates. *Pre-condition* includes prefecture-level population density in 1820 and the distance to the closest prefecture capital. *Geography* includes latitude, longitude, and the size of the county (in log). *Location* includes the distance to the Yangtze River, the Grand Canal, and the coastline. *Additional Geography* includes elevation, terrain ruggedness, small river density, and a second order polynomial of geographic coordinates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 6: Estimations Using Pooled Sample, 2000-2007

DV: Innovation Intensity (New product output/total output)				
	(1)	(2)	(3)	(4)
CMC	0.057***	0.051***	0.041***	0.032**
	(0.015)	(0.014)	(0.013)	(0.015)
N	179403	179403	179325	179325
R-squared	0.020	0.033	0.047	0.050
Year FE	Yes	Yes	Yes	Yes
CMC FE	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes
Agglomeration	No	No	No	Yes

Notes: The table reports IV estimates with the ASIF data set 2000-2003, and 2005-2007. All regression include year fixed effects. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), and ownership type. *Agglomeration* is the number of firms in a county-industry cell. Robust errors adjusted for clustering at the province-industry level are reported in parentheses.

***, ** and * indicate significance at 1%, 5% and 10%

Table 7: CMC Institution and Corruption Cases

DV: Corruption cases per 10,000 people				
	(1)	(2)	(3)	(4)
CMC	-0.010***	-0.007**	-0.007**	-0.011***
	(0.003)	(0.003)	(0.003)	(0.004)
N	588	588	588	588
R-squared	0.009	0.018	0.264	0.301
Distance˙CMC	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes
Province FE	No	No	No	Yes

Notes: OLS regression at the (contemporary)county-year level, from 2012 to 2015. The dependent variable is the number of corruption cases per 10,000 people. Dis_CMC refers to the distance to the closest CMC station (in log). Robust standard errors clustered at the county level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 8: CMC Institution and Firm Ownership Type

DV: Innovation intensity			
	SOE	Private firms	Foreign firms
	(1)	(2)	(3)
CMC	-1.021	0.074***	0.052
	(1.447)	(0.023)	(0.046)
N	1120	26452	12039
R-squared	-1.051	0.061	0.070
Controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes

Notes: The table reports IV estimates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 9: CMC Institution and Industry-Level Corruption

DV: Innovation Intensity					
	Corruption index quintile (low to high)				
	1	2	3	4	5
CMC	0.028	0.070	0.128	0.097	0.147**
	(0.032)	(0.045)	(0.082)	(0.074)	(0.068)
N	10482	7251	7855	8690	7709
R-squared	0.121	0.041	0.029	0.015	0.032
Controls	Yes	Yes	Yes	Yes	Yes
CMC FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 10: CMC Institution and Competition

DV: Innovation Intensity (New product output/total sales)				
Panel A: Testing linear relationship				
	(1)	(2)	(3)	(4)
CMC	0.078***	0.078***	0.078***	0.076***
	(0.030)	(0.029)	(0.029)	(0.029)
HHI	0.002			
	(0.003)			
CR10		0.001		
		(0.015)		
CR 20			-0.011	
			(0.024)	
CR 50				-0.056
				(0.057)
N	41987	41987	41987	41987
R-squared	0.051	0.050	0.051	0.052
Panel A: Testing non-linear relationship				
CMC	0.079***	0.078***	0.078***	0.079***
	(0.030)	(0.029)	(0.029)	(0.030)
HH index	-0.027			
	(0.038)			
HH index sqr	0.002			
	(0.003)			
CR10		-0.018		
		(0.101)		
CR 10 sqr		0.014		
		(0.066)		
CR 20			-0.052	
			(0.161)	
CR 20 sqr			0.027	
			(0.093)	
CR 50				-0.247
				(0.240)
CR 50 sqr				0.109
				(0.119)
N	41987	41987	41987	41987
R-squared	0.050	0.050	0.051	0.051
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 11: CMC Institution and FDI

DV	FDI	Innovation intensity		
	(1)	(2)	(3)	(4)
CMC	0.841***	0.078***	0.078***	0.078***
	(0.306)	(0.029)	(0.029)	(0.029)
FDI		0.000		0.000
		(0.001)		(0.001)
FDI spillover			-0.001	-0.001
			(0.001)	(0.001)
N	41987	41987	41987	41987
R-squared	0.790	0.051	0.051	0.051
Fstat	14.562	14.534	15.225	15.197
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 12: CMC Institution and Trade

DV	Trade	Innovation intensity		
	(1)	(2)	(3)	(4)
CMC	0.051	0.077***	0.070**	0.076***
	(0.051)	(0.029)	(0.028)	(0.029)
Trade		0.028***		0.022***
		(0.007)		(0.006)
Trade spillover			0.067***	0.062***
			(0.019)	(0.016)
N	41984	41984	41987	41984
R-squared	0.312	0.055	0.059	0.059
Fstat	14.562	14.543	14.454	14.477
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates. *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table 13: Confounding Factors

DV: Innovation Intensity						
Corruption:	Low	High	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
CMC	0.030	0.149**	0.031	0.137**	0.027	0.148**
	(0.035)	(0.069)	(0.031)	(0.066)	(0.033)	(0.069)
HHI	-0.003	0.003				
	(0.007)	(0.007)				
Trade			0.047***	0.047*		
			(0.001)	(0.001)		
FDI					0.000	-0.001
					(0.001)	(0.001)
N	10482	7709	10481	7709	10482	7709
R-squared	0.121	0.031	0.129	0.045	0.121	0.031
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports IV estimates using samples with the lowest corruption index quintile (indicated by columns under *Low*) and samples with the highest corruption index quintile (indicated by columns under *High*). *Controls* refer to firm age (in log), employment (in log), capital intensity, sales (in log), ownership type, agglomeration, and the distance to the closest treaty port (in log). *Fixed effects* refer to CMC fixed effects and industry fixed effects. Robust standard errors clustered at the province-industry level are reported in parentheses.

***, ** and * indicate significance at 1%, 5% and 10%

Table A1: Descriptive Statistics

	Mean	S.D.	Min	Max
Innovation intensity	0.032	0.143	0	1
CMC	0.177	0.382	0	1
CMC Coverage	0.191	0.236	0	0.973
Dis_CMC (in log)	3.680	0.462	2.722	5.066
Sale (in log)	10.282	1.174	-4.605	14.753
Employment (in log)	4.815	0.908	3.401	7.903
Capital intensity	90.430	245.604	0.001	18277.291
Age (in log)	1.997	0.696	0	4.615
Agglomeration	127.469	151.495	1	739.000
Ownership dummies:				
SOE	0.027	0.161	0	1
Foreign firms	0.128	0.334	0	1
HK/TW/MC holding	0.159	0.366	0	1
Private firs	0.630	0.483	0	1
Collective ownership	0.057	0.231	0	1

Table A2: Baseline Results, Full Results

DV: Innovation Intensity				
	(1)	(2)	(3)	(4)
CMC	0.088***	0.085***	0.080***	0.078***
	(0.034)	(0.030)	(0.028)	(0.030)
Dis_CMC	0.038***	0.037***	0.036***	0.036***
	(0.009)	(0.008)	(0.008)	(0.008)
Sale			0.008***	0.008***
			(0.002)	(0.002)
Employment			0.009***	0.009***
			(0.002)	(0.002)
Capital intensity			0.000	0.000
			(0.000)	(0.000)
Age			-0.004***	-0.004***
			(0.002)	(0.002)
Agglomeration				0.000
				(0.000)
Hoding: compared to state owned firms				
Foreign			-0.043***	-0.043***
			(0.013)	(0.013)
HK/ /TW/MAC			-0.042***	-0.042***
			(0.013)	(0.013)
Private			-0.039***	-0.039***
			(0.011)	(0.011)
Collective			-0.038***	-0.038***
			(0.011)	(0.011)
N	41987	41987	41987	41987
R-squared	0.021	0.033	0.050	0.050
Fstat	14.674	14.617	14.780	14.531
CMC FE	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes

Notes: The table reports IV estimates. Robust standard errors clustered at the province-industry level are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%

Table A3: Share of Ownership Mismatched Observations

Share of mismatched observations							
Year	2000	2001	2002	2003	2005	2006	2007
State-owned	8.42 %	7.74 %	6.87 %	5.65 %	3.56 %	3.19 %	2.35 %
Foreign	6.15%	5.66%	5.74%	5.17%	4.19%	3.92%	3.06%
HK/MO/TW	6.61%	6.15%	6.05%	5.50%	4.03%	3.61%	2.77%

Notes: The table shows the share of mismatched observations across the period 2000-2007. A mismatched observation is defined as a firm's registration type inconsistent with its ownership type.

Table A4: Sampling Issue in World Bank Survey

DV: Number of firms				
	(1)	(2)	(3)	(4)
CMC	2.574	2.802	1.970	3.924
	(2.974)	(3.045)	(2.969)	(5.776)
N	1137	1137	1137	74
R-squared	0.001	0.235	0.084	0.181
Fixed Effects	No	Prefecture	Province	CMC

Notes: The table reports OLS estimates. Standard errors are reported in parentheses.
***, ** and * indicate significance at 1%, 5% and 10%