European Journal of Political Economy 85 (2024) 102559

Contents lists available at ScienceDirect

European Journal of Political Economy

journal homepage: www.elsevier.com/locate/ejpe

Anti-corruption campaign in China: An empirical investigation

Li Yang [a,b,*], Branko Milanovic [c,d], Yaoqi Lin [c,e]

[a] ZEW, Leibniz Centre for European Economic Research, Germany
[b] Paris School of Economics, France
[c] Graduate Center City University of New York, USA
[d] London School of Economics, UK
[e] Washington State Department of Social and Health Services, USA

ARTICLE INFO

JEL classification:
D53
D73

Keywords:
Corruption
China
Income distribution

ABSTRACT

We create a database of officials who have been found guilty of corruption in China in the period 2012-21 with their personal characteristics and the amount of embezzled funds. We use it to investigate the correlates of corruption, estimate the effects of corruption on inequality, and find the expected increase in officials' income due to corruption and the gain in income distribution ranking. We find that the amount of corruption is positively associated with education, administrative (hierarchical) level of the official, and years of membership in the Communist Party. The sample of corrupt officials belongs to the upper income ranges of Chinese income distribution even without corruption. But corruption allows them to accede to an even higher position in income distribution. While only one-half of the corrupt officials would be in the top 5 percent of China's urban distribution without illegal incomes, practically all are in the top 5 percent when corrupt income is included.

1. Introduction

With China's economic reform in the late 1970s, and then again with a pro-market turn in 1992-93, corruption became increasingly visible. Since growth has been the top priority for China's leadership, tackling corruption took a back seat in the past four decades. Meanwhile inequality, some of it probably fueled by corruption, increased significantly (see, inter alia, Gustafsson et al. (2014), Xie and Zhou (2014). Zhuang and Shi (2016), Piketty et al. (2019), Yang et al., 2021, Targa and Yang (2023)). By 2012 corruption became the most compelling challenge confronting the ruling power of Communist Part of China (CPC).[1] Driven by such perception, a far-reaching anti-corruption campaign was started under the aegis of Xi Jinping, General Secretary of the Chinese Communist Party since the 18th Communist Party Congress. For a decade, Xi's anti-corruption campaign, sweeping across the party, state, and enterprises, targeting not only "tigers" (high-ranking corrupt officials), but also "flies" (low-ranking corrupt officials), was the largest

* Corresponding author. ZEW, Leibniz Centre for European Economic Research, Germany.
E-mail address: li.yang@zew.de (L. Yang).

[1] As concluded by both the then outgoing General Secretary Hu Jintao and by Xi Jinping in their speeches at the 18th Communist Party Congress. Xie (2016, p. 21) writes, paraphrasing CPC documents, "to govern the country, [the Party] must first run the Party well, and to run the Party, it must reinforce strict discipline."

https://doi.org/10.1016/j.ejpoleco.2024.102559
Received 2 May 2023; Received in revised form 23 February 2024; Accepted 23 May 2024
Available online 1 June 2024

Table 1
Impact of Anti-Corruption Campaign on top government officials (2012–2021).

Position	No. of Officials		No. of officials convicted of corruption		Percentage of officials convicted of corruption per year		
	Principal	Deputy	Principal	Deputy	Principal	Deputy	Total
National Leaders	12	65	1	6	0.8%	0.9%	0.9%
Departmental leaders of the State Council	41	177	5	17	1.2%	1.0%	1.0%
Provincial Leaders	124	756	10	92	0.8%	1.2%	1.2%
Prefecture Leaders	1332	–	62	–	0.5%	–	0.5%

Notes: Data of number of national leaders is from http://cpc.people.com.cn/GB/64162/394696/index.html (2023.03.03). Data of number of provincial leaders is from http://district.ce.cn/zt/rwk/index.shtml (2023.03.03). Data of number of Departments of the State Council officials is from http://www.ce.cn/ztpd/xwzt/rwk/index.shtml (2023.03.03). Principal provincial leaders include Secretary of the Provincial Party Committee, Provincial Governor, Chairman of the Provincial Political Consultative Conference, and Director of Provincial People's Congress from 31 provincial administrative units in China. Deputy provincial leaders include Deputy Secretary of the Provincial Party Committee, Deputy Governor, Vice Chairman of Provincial Political Consultative Conference, Deputy Director of Provincial People's Congress from 31 provincial administrative units in China. Departmental leaders of the State Council come from 26 Departments of the State Council, 1 Special agency directly under the State Council, and 10 Institutions directly under the State Council. Principals include secretaries of the party committee, ministers or directors, deputies include vice ministers, vice directors, director of political department, or discipline inspection team leader. Principal prefecture leaders include Secretary of the Prefecture Party Committee, Prefecture Governor, Chairman of the Prefecture Political Consultative Conference, and Director of Prefecture People's Congress from 333 prefecture administrative units in China. Number of officials convicted of corruption is calculated based on the corruption dataset which is constructed from the corruption cases published in by the Central Commission for Discipline Inspection (CCDI) and used in this paper (for further explanation see Appendix A).

organized anti-corruption effort in the history of CPC. Table 1 presents the impact of anti-corruption campaign on top government officials or "tigers". Between 2012 and 2021, the average annual conviction rate for corruption among leaders at national and provincial levels was approximately 1%, which is twice the rate observed among prefecture-level leaders.[2] By May 2021, a total of over four million cadres and officials had been investigated, with 3.7 million of them having been punished by the Central Commission of Discipline Inspection (CCDI). [3]

Utilizing data systematically compiled from individual conviction cases regularly updated on the CCDI website, we have constructed a novel database featuring senior Chinese officials, often referred to as "tigers" in anti-corruption campaigns, found guilty of corruption during the period 2012–2021. Our analysis aims to explore the correlates of corruption, assess its impact on inequality, estimate the resulting increase in officials' income, and determine the subsequent change in income distribution ranking.

Our findings indicate a positive correlation between the amount of corruption and factors such as education level, administrative hierarchy, and years of Communist Party membership among the implicated officials. Notably, the sample of corrupt officials tends to belong to the upper income echelons of the Chinese income distribution, even without factoring in corruption-related income. However, corruption emerges as a significant catalyst for climbing even higher. While only half of the corrupt officials would be in the top 5 percent of urban distribution based on legal incomes alone, virtually all of them achieve this status when their corrupt income is considered.

Corruption as a topic has been extensively examined in both China and globally, as evidenced by comprehensive surveys such as the IMF-commissioned study by Abed and Gupta (2002), along with similar investigations by Jain (2001), Zimelis (2020), Tong (2022), and Dong and Torgler (2013). Despite the considerable attention given to corruption, its inherent hidden nature makes it challenging, if not impossible, to quantify its absolute levels accurately. Existing literature has categorized the measurement of corruption into two main types. The first involves survey-based measurements, relying on respondents' perceptions or experiences of corruption (see, Liu and Peng, 2015; Jetter and Parmeter, 2018; Gründler and Potrafke, 2019; Choudhury, 2019; Gutmann et al., 2020; Pittaluga et al., 2023). While such data are widely used for measuring corruption (such as Corruption Perception Index[4]), perceptions of corruption are subjective and not necessarily reflective of factual occurrences (Olken, 2009). Meanwhile, as highlighte d by Hillman (2010), survey responses can be influenced by expressive behavior. The second is objective measurements utilizing legal statistics, predominantly at the regional or national level, focusing on the number of court cases or convictions as indicators of corruption (for example, for the U. S., Alt and Lassen, 2014; for Russia, Schulze et al., 2016; for Italy, Capasso and Santoro, 2018; Mocetti and Orlando, 2019; and for China,Dong and Torgler, 2013; Chen and Liu 2018; Jiang 2018; Demir et al., 2022; Chih et al., 2023). However, these metrics are not immune to ambiguity, as a higher number of convictions may result from more effective legal enforcement rather than a higher prevalence of corruption. Additionally, a simple count of convictions treats all acts of corruption, regardless of their scale, with equal significance.

[2] The conviction rate is obtained as the ratio of convicted officials during five years over the number of officials during the same period. It is then annualized. The five-year period is chosen because the officials' terms are normally five years and thus the overall number ("stock") of officials is normally fixed during that period.

[3] https://finance.sina.com.cn/tech/2021-06-28/doc-ikqcfnca3716443.shtml.

[4] See https://www.transparency.org/.

Our study is based on objective measurements of corruption however at the individual level, which makes it a rarity in the literature. Obtaining extensive and detailed individualized corruption data is seldom achievable despite the extensive empirical studies on corruption, and research based on such individual-level corruption data is even scarcer. Minxin Pei (2016, Appendix) presents a dozen data to illustrate the type of perpetrators and the extent of corruption. These data, with some additional information, were used by Milanovic (2019, p.p. 110-1) to show that the extent of corruption increases with the administrative level as one moves from county to prefecture to province, and that, at a given level, it tends to be higher among those working for CPC bodies than among the government and SOEs officials. The most detailed work to date, similar to ours, was done by Aidt, Hillman and Liu (2020) who conducted analysis based on the judicial documents of bribe-taking cases from 1991 to 2015 posted by the Supreme People's Court in China Judgement Online.[5] Most of the convicted individuals in the dataset are low rank officials.[6] Lorentzen and Lu (2018) and Shi (2022) are two other studies that utilized information from the CCDI website regarding convicted officials or those who were under investigation. However, these two studies made use of only a narrow range of information.[7] Compiling from individual conviction cases published by the Central Commission for Discipline Inspection, our database contains detailed information on personal characteristics (age, education, job position) of the convicted perpetrators[8] as well as the involved amount of corruption, which allows us to explore the determinants of corruption empirically and systematically, surpassing anecdotal evidence.

The contribution of out paper is threefold. First, we construct a novel dataset of senior officials convicted of corruption between 2012 and 2021, which covers the entire anti-corruption campaign (up to and including 2021) that began under Xi Jinping leadership. To the best of our knowledge, we are the first to systematically compile relevant information and construct a comprehensive corruption dataset of senior officials (as called by the official sources, the "tigers"). Such detailed information enables us to explore the determinants of the amount of corruption (which includes all kinds of monetary corruption: bribery, embezzlement, unexplained source of money etc.), which is rarely studied in the empirical literature of corruption due to the lack of individual level data on corruption. [9] We find that the amount of corruption increases with the level of education, number of years of CPC membership, and administrative level of the job. In effect, there is a strong positive relationship between the administrative level and amount of corruption.[10] We also find that age and gender are not significant predictors of the amount of corruption. Thanks to a very comprehensive database we are also able to observe that graduating from specialized Communist Party schools is inversely related to amount of corruption. We also find that the cohorts that have become CPC members after 1978, and especially after 1992, are associated with greater amounts of corruption (compared to Party members since before 1978).

Second, our unique database allows us to investigate, for the first time, inequality in the amount of corruption and compare it with the data on income distribution obtained through the regular Chinese Household Income Surveys (CHIP). We are thus able to look how corruption, considered as a "rent" income, compares both in size and distribution to other sources of income. We find that inequality of income acquired through corruption is much greater than inequality of disposable income and is almost the same as inequality of income from capital: Ginis for both are 0.69.[11] Strong concentration of corruption is also reflected in the fact that the top decile of perpetrators is responsible for 58% of total corruption, and the top 1 percent, for close to 21%.

Third, for the first time in literature, we estimate, using the known perpetrators' characteristics where in the income distribution they would be located without corrupt income, and how much they gain through corruption both in absolute amount and in income ranking. It is obvious that corruption will increase the income of those who are corrupt: the monetary gain is certainly one, and perhaps the only, reason people engage in corruption, but information on how much corrupt people gain relative to what would have been their legal income had never been available and thus was not studied. The objective of our paper is therefore to go beyond the determinants of corruption alone, and to study how much corruption increases income of individuals who engage in it and how much they gain positionally. Combining the information about the skewness of corruption with the ability to locate the position of the perpetrators in income distribution before and after corruption, we show that 80% of perpetrators (when we include only their legal income) belong to the top decile of Chinese urban income distribution. On average, corruption allows them to increase their income by between 4.7 and 7 times (depending on the method of annualization of the stock of corruption) and leapfrog many others in income distribution rankings. While without corrupt income only 6% of those convicted would be in China's urban top 1 percent, with corrupt income between 83% and 91% are (again, depending on the method of annualization). We argue that corruption, at least as revealed by the results of the

[5] Chinese Political Elite Database https://wenshu.court.gov.cn/.

[6] That is, out of 45,846 individuals they study, less than 2% hold a rank equal to, or higher than, sub-provincial ministerial level, while 8% hold a rank of prefectural-bureau level (including sub-prefectural-bureau level). Our dataset however covers predominantly the "tigers": the corresponding percentages in our dataset are 14 and 85. The databank ends in 2015 when the current campaign was still in its early stages.

[7] More distantly related, Jiang (2018) constructed Chinese Political Elite Database (https://wenshu.court.gov.cn/) containing biographical information for over 4000 municipal, provincial, and national leaders in China from late 1990s to 2015. The database also includes information about government officials investigated by the Central Commission for Discipline Inspection. However, the primary focus of the paper is not on corruption. Demir et al. (2022) developed a prefecture-level bureaucratic corruption index utilizing Jiang (2018)'s database, using it to examine the impact of local corruption on the total factor productivity of manufacturing firms. Similarly, Chih et al. (2023) investigate the effects of corruption in the home city on foreign direct investment flows in China, utilizing the same corruption index.

[8] We use the terms the perpetrator, the accused and the defendant interchangeably.

[9] Aidt et al. (2020) is one exception.

[10] These results align with those of Aidt, Hillman, and Liu (2020); however, we also find that the amount of corruption for officials at the provincial level is, on average and holding other variables constant, more than three times as high than that of prefecture-level officials. In contrast, Aidt, Hillman, and Liu (2020)find that the amounts of corruption for these two levels of officials are almost the same.

[11] Calculated only across positive values of both corruption and capital income.

Table 2
Number of cases of corrupt behavior by type of corruption and type of defendant's employment.

	1	2	3	4	5	6
	Corruption (general)	Unexplained source of money	Illegal money	Illegal earnings	Embezzlement	Total number of cases
Government	487	35	19	22	14	577
Court system	22	1		1	3	27
Public institutions	43			1	5	49
SOEs	119	6	8	5	9	147
Social organizations	15	5	1		1	22
Total	*686*	*47*	*28*	*29*	*32*	*822*

Note: The total number of individuals convicted is 686, but the total number of cases is 822 with some individuals being accused of more than one type of corrupt behavior.

current campaign is an upper-income group phenomenon, it increases dramatically income of people who engage in it, and enables them to join the very top of China's income distribution.

Despite the database's innovative features and potential applications, we want to emphasize that the "tiger" corruption database shares common limitations with previous studies that attempt to measure corruption or crimes (see Glaeser and Saks, 2006; Schulze et al., 2016; Capasso and Santoro, 2018; Aidt et al., 2020). Specifically, like all measurement approaches in the field of corruption, our dataset only includes cases where convictions have occurred, that is, the observable variable. The true extent of corruption remains unobservable. Furthermore, even within the set of convicted cases, there may be instances of false verdicts. The first type of measurement error is false negative and the second type, false positive.

There are many channels which could leads to false negative, such as the variation of the strength of anti-corruption campaign across regions and industries[12] or personal connections of corrupted officials. To (partially) address this issue, we include province fixed effect in the analysis of determinants of amount of corruption (see Table 7) to control for the variation of the anti-corruption campaign across province. Nonetheless, we admit that for most of these channels, which could lead to false negative, we are not able to control for or make correction at this stage.

Regarding false positive, the main concern is whether the current anti-corruption campaign is entirely genuine or contains some political elements used to reinforce Xi Jinping's power. Consequently, some convicted senior officials may not be guilty. The existing literature has been engaged in an ongoing debate on this issue since the commencement of Xi's anti-corruption campaign, with inconclusive findings. While certain studies suggest that political influence might play a role in the campaign (Tong, 2022), emerging empirical evidence leans toward the notion that the effort is genuine and has effectively altered the incentives for individuals, political entities, and State-Owned Enterprises (SOEs) to minimize corruption and overcome structural barriers to anti-corruption enforcement (Lorentzen and Lu, 2018; Manion, 2016). Nevertheless, akin to the false negatives, we cannot definitively rule out the presence of such measurement error.

The rest of this paper is organized as follows. Section 2 offers a concise historical background on anti-corruption measures in China. Section 3 provides an overview of the data and variables used. Sections 4 and 5 delve into the determinants of the amount of corruption. Section 6 analyzes the income ranks of corrupted officials, both with and without including income from corruption. Section 7 presents the conclusion. To complement this paper, an extensive online appendix is available, encompassing the full data file and a detailed description of the database.

2. Historical background of anti-corruption measures in China

Having been ruled by authoritarian regime for two thousand years, China is not unfamiliar with corruption or, relatedly, with extreme measures to eliminate it.[13] The first major anti-corruption campaign in China can be traced back to the Han Dynasty (202 BC-220 AD), when Emperor Wu of Han (156-87 BCE), a particularly active anti-corruption campaigner, established the office of regional inspector (*cishi*) to control and supervise the officials in the thirteen provinces (*zhou*). From Han to Ming to China's inter-war republican government to the People Republic of China, offices whose role was to fight corruption and ferret out the criminals have been common. The first full-time central supervisory agency, Yushi Tai (also known as the Imperial Censorate), was established in the Han Dynasty (206 BC-220 AD). In the following millennium, despite changes in organization and prescribed functions, it played a significant role in maintaining the integrity of the Chinese government and ensuring that officials were held accountable for their actions. In Ming Dynasty (1368–1644), Yushi Tai was replaced by Ducha Yuan (Chief Investigating Bureau). The so-called censorate institutions (督察院) were thus a prominent part of China's governmental structure for more than two thousand years. Censors were both an arm of government and its controllers. In modern systems like the American, their role could be seen to combine that of the Government Accounting Office with the court system.

[12] Between 2013 and 2022, the Central Committee of the Communist Party of China dispatched 19 rounds of Central Leading Groups for Inspection Work, each with distinct focuses (See http://m.ccdi.gov.cn/content/98/f8/20269.html).
[13] Chunyu Wang (2022).

Table 3
Number of cases of corruption and the amounts of corruption by type of employment.

	Total corrupt money		Total number of cases		Corruption per case	
	In 10k yuans	Share (in %)	Number	Share (in %)	In 10k yuans	Relative (average = 100)
Government	1,953,505	70	577	70	3386	99
Court system	58,567	2	27	3	2169	63
Public institutions	35,866	1	49	6	732	21
SOEs	642,056	23	147	18	4368	128
Social organizations	119,928	4	22	3	5451	159
Total	*2,809,923*	*100*	*822*	*100*	*3418*	*100*

Note: The amounts are expressed in 10,000 yuans in 2018 prices.

Table 4
Number of defendants and the amount of corruption by administrative level of the defendant's job.

	Total corrupt money		Total number of defendants		Corruption per defendant	
	In 10k yuans	Share (in %)	Number	Share (in %)	In 10k yuans	Relative (average = 100)
National level	12,096	0.4	1	0.1	12,096	295
Sub-national level	41,545	1.5	3	0.4	13,848	338
Provincial ministerial level	186,578	7	23	3	8112	198
Sub-provincial ministerial level	1,314,744	47	140	20	9391	229
Prefectoral- level	760,460	27	219	32	3472	85
Sub-prefectural-level	483,205	17	297	43	1627	40
County level	11,296	0.4	3	0.4	3765	92
Total	*2,809,923*	*100*	*686*	*100*	*4096*	*100*

Note: The amounts are expressed in 10,000 yuans in 2018 prices.

Table 5
Number of defendants and the amounts of corruption by CCDI-defined type of cadre.

	Total corrupt money		Total number of defendants		Corruption per defendant	
	In 10k yuans	Share (in %)	Number	Share (in %)	In 10k yuans	Relative (average = 100)
CMCs	1,766,165	63	182	27	9704	237
PMCs	891,363	32	453	66	1968	48
CLC	152,395	5	51	7	2988	73
Total	*2,809,923*	*100*	*686*	*100*	*4096*	*100*

Note: The amounts are expressed in 10,000 yuans in 2018 prices. CMC=Centrally-Managed Cadres; PMC=Provincially-Managed Cadres, CLC=Central-Level Cadres.

In the Republic of China (1912–1949), Ducha Yuan was replaced by the Control Yuan (监察院). It was one of the five Yuans or branches of government according to Sun Yatsen's principles, namely executive, legislative, judiciary, examination and recruitment of officials, and finally control. The Control Yuan was responsible for investigating and disciplining government officials who engaged in misconduct or abuse of power.

People's Republic of China maintained from its inception similar organs, working through two channels: the government administrative channel with the Committee of Public Supervision, and the Communist Party of China (CPC) channel with the Disciplinary Commission (which typically exists in all communist parties). Today's anti-corruption controls in People's Republic of China are carried by (i) Central Commission for Discipline Inspection (CCDI, 中央纪律检查委员会) which is an arm of the Central Committee of CPC, with CCDI ruling body selected for five-year terms to coincide with those of the Central Committee members, and by (ii) the National Supervision Commission (NSC, 国家监察委员会) which is a government body.[14] CCDI in particular plays a central role in investigating and punishing corrupt officials, who are almost all Communist Party members, from the central, provincial and county levels.

3. Data

We compiled our corruption dataset using mostly data obtained from the website of the Central Commission for Discipline Inspection.[15] CCDI has been consistently updating corruption cases involving senior officials in China since 2012 to demonstrate the

[14] Their predecessors are Central Control Commission of CPC established in 1927 (中央检查委员会) and Committee of People's Supervision (人民监察委员会), established in 1951 under the State Administration Council. See also Xie (2016).
[15] https://www.ccdi.gov.cn/scdc/. The last data point is collected on May 28, 2021.

Table 6
Concentration of total corruption compared to concentration of disposable income.

Decile	Total corruption		Disposable income	
	Average amount per recipient (in 10k yuans)	Share (in %)	Average amount per capita (in yuan)	Share of total (in %)
1	190	0.5	2927	1.3
2	421	1.0	5884	2.7
3	644	1.6	8213	3.7
4	888	2.2	10896	4.9
5	1216	3.0	13899	6.3
6	1676	4.1	17408	7.9
7	2313	5.6	21942	9.9
8	3756	9.2	28241	12.8
9	6300	15.4	38584	17.5
10	23746	58.0	72891	33.0
Mean/total	*4096*	*100*	*22085*	*100*
Gini	0.69		0.46	

Note: The amounts of corruption are cumulative (in real 2018 yuans); the amounts of income are annual (calculated from the CHIP, 2018).

progress and outcomes of anti-corruption efforts. The dataset includes senior officials categorized by CCDI, by order of importance, as Centrally-Managed Cadres (CMC),[16] Provincially-Managed Cadres (PMC),[17] and central-level cadres (CLC) from the party, state institutions, state-owned enterprises, and financial institutions (excluding CMC or PMC).[18] Therefore, our dataset can be referred to as the "Tigers' corruption" dataset. In cases where essential variables were missing from the primary data source, an extensive search of various online platforms was conducted to supplement our dataset. [19]

To be more precise, our dataset offers detailed information on convicted officials, beginning with their demographic particulars, including their name, gender, age, birthplace, and education level and major of studies.[20] As a distinctive aspect of the education of Chinese officials, we introduce a dummy variable to indicate whether the officials have graduated from the Central Party School of the Central Committee of the CPC (中央党校). Approximately one-third of the defendants in our dataset have graduated from the Central Party School, and as we shall demonstrate later, the Central Party School dummy variable is inversely associated with the amount of corruption.

Secondly, given that high-ranking officials in China are predominantly members of the CPC (99.5% of the officials in our dataset), we introduce a variable indicating the year when the defendant joined the CPC, enabling us to have information on defendants' duration of CPC membership and to generate dummy variables for various CPC cohorts (by year of membership).

Moreover, we have comprehensive employment information regarding the defendants, including geographical location of their workplaces, the year of initiation and termination of their most recent job, as well as the classification and administrative level of their respective job posts. We categorize the defendants' job assignments into five employment types[21] and seven administrative levels (see Tables 3 and 4 below).[22]

[16] Centrally Managed Cadres (中管干部) refer to the positions of leading cadres who are listed in the "List of Positions of Cadres Managed by the Central Committee of the Communist Party of China" and appointed and removed by the Central Committee of the Communist Party of China. The Organization Department of the Central Committee has the right to make suggestions on the appointment. Generally speaking, CMCs are above the Sub-provincial-ministerial level; some cadres at the Prefectural-bureau level have also been included in the cadres of the central management.

[17] Provincially Managed Cadres (省管干部) refer to the positions of provincial cadres appointed and removed directly by the Organization Department of the Provincial Party Committee. Therefore, the scope of provincial cadres generally covers the chief and deputy secretaries, municipal party committee members, and chief and deputy mayors of prefecture-level cities; chief and deputy officials of provincial departments, secretaries of universities and colleges, principals, chairmen of provincial enterprises, enterprise party secretaries, general managers etc.

[18] The officials in the CLC group come from central enterprises, universities, Ministries or Bureaus at provincial level (but the administrative rank is relatively low, i.e. it is prefectoral-level or below).

[19] Including Xinhua News Agency (http://www.xinhuanet.com), The Paper (https://m.thepaper.cn), The State Council, The People's Republic of China (http://www.gov.cn), Reuters (http://www.reuters.com), Sina (https://news.sina.com.cn, https://finance.sina.com.cn), and The Chinese Court Net (http://www.chinacourt.org).

[20] We use the terms "defendant" and "convicted official" interchangeably because all defendants included in our database have been found guilty.

[21] Namely, government, court system (judiciary), public institutions (including junior colleges, universities and hospitals), state-owned enterprises, and social organizations (Including Union of Supply and Marketing Cooperatives(供销社), Credit Union(信用社), Federation of Industry and Commerce(工商联), and other social associations with government backing.). Additionally, two finer classifications of job posts of the defendants are available in our dataset, with 14 and 84 categories of posts respectively.

[22] The administrative level is coded according to the level and ranking system stipulated in "Civil Servant Law of the People's Republic of China (2018 Revision)". In total there are 12 administrative levels i.e., 1. National level, 2. Sub-national level, 3. Provincial ministerial level, 4. Sub-provincial ministerial level, 5. Prefectural-bureau level, 6. Sub-prefectural-bureau level, 7. County-division level, 8. Deputy-county-division-head level, 9. Section-head level, and 10. Deputy-Section-head level, 11. Section member and 12. Ordinary staff. In our dataset, the defendants are only from the top 7 levels.

Table 7
Determinants of amount of corruption.

	Formulation (1)	Formulation (2)
Male	−0.084	−0.032
	(0.221)	(0.238)
Age	−0.175	−0.213
	(0.195)	(0.193)
Age^2	0.002	0.002
	(0.002)	(0.002)
Education level (baseline: bachelor or less)		
Graduate education dummy	0.307[a]	0.293[b]
	(0.101)	(0.108)
Central Party School dummy	−0.243[b]	−0.208[c]
	(0.114)	(0.115)
Years of CPC membership	0.060[a]	0.051[b]
	(0.020)	(0.020)
CPC Cohort (baseline: joined CPC before 1978)		
1979 and 1992	0.712[a]	0.706[b]
	(0.259)	(0.263)
After 1992	1.243[a]	1.113[b]
	(0.398)	(0.405)
Employment type (baseline: government)		
Court system	−0.094	−0.170
	(0.252)	(0.270)
Public institution	−0.59[a]	−0.62[a]
	(0.118)	(0.134)
State Owned Enterprises	0.243[b]	0.372[b]
	(0.119)	(0.138)
Social Organizations	0.973[a]	0.965[a]
	(0.265)	(0.230)
Administrative level of the defendant (Baseline: Sub-prefectural level or lower)		
National level and sub-national level	2.396[a]	2.758[a]
	(0.194)	(0.213)
Provincial ministerial level	1.466[a]	1.698[a]
	(0.218)	(0.263)
Sub-provincial ministerial level	1.381[a]	1.379[a]
	(0.180)	(0.196)
Prefectural level	0.283[b]	0.259[b]
	(0.103)	(0.111)
Job Region (baseline: central)		
East	0.284	
	(0.180)	
Northeast	0.163	
	(0.151)	
West	0.378	
	(0.235)	
Constant	8.650	9.819*
	(5.673)	(5.613)
Province Fixed effect	No	Yes
Cluster	Province level	Province level
Observations	567	567
R-squared	0.297	0.376

Notes: OLS regression. Standard errors are robust, clustered at the provincial level.
Our analysis is based CCDI dataset. We restrict our sample to only CPC members. The dependent variable is natural log of the amount of corruption in 2018 yuans. The regression is run across all convicted individuals but since for some of them not full information is available the sample is reduced from 686 to 567.
[a] $p < 0.01$.
[b] $p < 0.05$.
[c] $p < 0.1$.

The crimes in CCDI database are divided into seven types: (1) crime of acceptance of bribes[23] and crime of dereliction of duty,[24] (2) organized crime, (3) drug- and sex-related crimes, (4) bribes given, (5) homicide, (6) unauthorized access to national secrets, and (7) malfeasance. Given that the defendants may have committed more than one type of crime, there are in total 1728 cases of crimes for 1451 defendants. From among all the cases of crimes, we are only interested in (1), that is, crime of acceptance of bribes and crime of dereliction of duty, which accounts more than 80 percent of all cases. We were able to obtain data on 686 defendants who are guilty of (1) using, as mentioned above, an extensive search through various media platform. We provide the sources for each data point in our online appendix.

The illicit gains in (1), which are always expressed in monetary units, are in turn classified into five types: general corruption (i.e. acceptance of bribes), unexplained sources of income, illegally obtained money, illicit earnings, and misappropriation of public funds (see Table 2). There are a total of 822 instances of individual illicit gains for the 686 convicted individuals, as some defendants have multiple types of illicit gains in certain cases. We refer to these illegal gains as "corruption" or the "amount of embezzlement".

The amount of corruption in our dataset measures the stock of illicit gains accumulated over years. That stock however is estimated at the time the corrupt official is arrested, and is expressed in the values of that year. (From the readings of individual cases, we note that the stock often consists of foreign currencies, gold, works of art and jewelry.) We convert the nominal value of the stock into 2018 prices using the ratio between the price level in the year when the official was arrested and 2018. For example, if the official was arrested in 2012, the estimated stock of corruption will be increased by 12 percent, reflecting the inflation between 2012 and 2018. We thus obtain all the corrupt amounts expressed in 2018 prices and these are the amounts we use in the entire analysis.

Finally, to allow a comparison between the amount of corruption and the likely legal earnings of the convicted officials, we combined our corruption dataset with the China Household Income Project 2018 (CHIP18), which is the latest wave of a nationally representative household income survey. This survey contains comprehensive data on earnings (including annual wage and business income), occupation, and demographic characteristics of 70,431 individuals living across 16 provinces. [25] We estimated the earning function using the data from CHIP18 and to do so selected the same variables regarding individual characteristics as available in our corruption database. They are gender, age, CPC membership status, educational level and major of studies, type of contract, sector of industry of work unit, ownership and region of the work unit, and the administrative level of the job post of the respondent. We were thus able to "locate" where the defendants would be in China's urban income distribution if they had only legal earnings.

4. Describing and analyzing income from corruption

4.1. Corruption by type of employment

As already mentioned, we have full data in the sense that the case has been carried to the end and information about perpetrators is complete for 686 defendants (with 822 instances of individual illicit gains; some defendants are found guilty of more than one crime). These are the closed cases on which we shall focus in the rest of the paper.

Table 2 summarizes the five types of corruption according to the type of employment held by the officials when they were arrested. Almost 500 of the convicted 686 individuals were employed in the government administration (see Table 2, column 1).[26] Or differently, 577 out of 822 cases (or 70%) are related to the individuals who were working in the government apparatus (Table 2, column 6). About 20% of the cases are linked to the individuals working in SOEs. Thus, these two thus groups account for 90% of either individuals convicted or cases investigated. The remaining three types of employment (the court system including prosecution office, public institutions, and social organizations) are of marginal importance.

Table 3 breaks down both the total number of cases and the total amounts of corrupt money by the five employment types. The two key types of employment connected with corruption (namely, work in the government apparatus and in SOEs) have a slightly higher share of embezzled money than is their share of cases. For example, 18% of cases concern government officials, but they are convicted of having embezzled 23% of the total amount. The basic picture, namely of SOEs employees and those working in the government apparatus being both the most frequent culprits remains. It is also revealing to look at the amount embezzled per case.[27] Here, SOEs employees are placed at the top (after the relatively insignificant number of those working in social organizations) as most corrupt with

[23] According to the provisions of the Criminal Law of the People's Republic of China, if a state functionary takes advantage of his position to extort other people's property, or illegally accepts other people's property to seek benefits for others, he is guilty of accepting bribes.

[24] According to the provisions of the Criminal Law of the People's Republic of China, the crime of dereliction of duty refers to the crime of dereliction of duty when state agency staff take advantage of their position or engage in malpractice for personal gain, abuse their power, or neglect their duties to obstruct the legal, fair and effective execution of official activities of state agencies, thereby damaging the public's trust in the objectivity and impartiality of the official activities of state agency staff and causing heavy losses to the interests of the country and the people.

[25] The sample of CHIP 18 is coming from the big sample of the annual integrated household survey conducted by the National Bureau of Statistics in 2018. The latter contains 160 thousands households in 31 provinces. The CHIP sample was selected by systematic sampling method in three layers of east, center and west and contains 15 provinces. For more details, please see http://www.ciidbnu.org/chip/chips.asp?year=2018&lang=EN.

[26] It includes all those working at different types and different levels of governmental organs (CPC apparatus, People's Congress, Political Consultative Conference, Commission for Discipline Inspection etc.).

[27] These are cumulative amounts (stock of corruption) which the defendants are accused of having stolen over a number of years, extending to several decades and which we have converted into constant 2018 amounts. We shall later annualize these amounts in order to compare them with yearly earnings.

44 million yuan per case (or about $6.6 million at the 2018 average exchange rate[28]). They are followed by government officials (34 million yuan, or about $5.1 million at the 2018 average exchange rates). For those working in health and education (public institutions), amount of corruption per case was significantly less.

4.2. Corruption by administrative level

People accused and convicted of corruption can work at different administrative levels. For example, a worker in the government apparatus can work at the national level (the highest level in our database) or at the country level (the lowest level in the database). Here we are concerned with the hierarchical level at which corruption takes place. Table 4 shows a positive relationship between the average amount of corruption per case and hierarchical level. At country, sub-prefectoral and prefectoral levels, corruption per case is below the average; at sub-provincial and provincial level, it is about twice the average; at subnational and national level, it is about three times the average. The ratio between the average amount embezzled at national or sub-national vs. sub-prefectoral level is around 8-to-1. If we assume that power is proportional to the hierarchical level, it is not surprising that corruption (per case) will be proportional too, as the value of favors given by higher-level officials outstrips by far the amount of favors that can be provided by low-level officials.

But when we look at what level the bulk of corruption takes place, the situation changes. The most important levels are middling levels. Officials at the sub-provincial and prefectural levels account for about one-half of all cases and almost three-quarters of embezzled money. Combining this finding with the previous we note that the "heart" of corruption lies at the government apparatus and SOEs at just below the provincial level.

4.3. Corruption by type of cadre

The data provide also the information on the type of cadres who were convicted of corruption. This refers to the classification, mentioned above, used by CCDI that distinguishes Centrally-Managed Cadres (CMC), Provincially-Managed Cadres (PMC), and Central Level Cadres of the party and state institutions, state-owned enterprises and financial institutions (CLC).

As shown in Table 5, the amount of corruption increases with the cadre level: CMCs, on average, have been convicted of embezzling more than 4½ times as much per case as the PMCs. In fact, the corruption-amount gap between the CMCs and the other two categories is very high. While the number of cases of corruption is the highest at middling levels (two-thirds of corruption cases are due to PMCs), the per-case amounts of corruption are much greater for CMCs, and consequently, the share of total corrupt money that is attributable to the higher-level cadres is very large (63 percent).

4.4. Corruption as an income source and its distribution

Fig. 1 shows a histogram of corruption amounts. Its shape is similar to lognormal, but with a thicker upper tail, and a somewhat greater kurtosis. As discussed below, the top of corruption follows a very clear Pareto distribution.

Table 6 gives data for 686 defendants divided into ten deciles by the amount of corruption. It compares it then with the similar partitioning into ten deciles according to per capita disposable income in urban China in 2018. The objective is to contrast the concentration of corruption to that of disposable (after-tax and after-transfer) income. Obviously, the deciles are composed of very different individuals, and the "horizontal" comparison of corruption amount of (say) first corruption decile with disposable income of the first income decile is meaningful only after the nominal amounts of corruption which are stocks are annualized (and thus converted into flows). We proceed to do that in the next section.

Here, however, we simply compare the concentration of corruption with that of disposable income. The former is much more concentrated with the top decile receiving almost 58 percent of all corruption whereas the top decile by income receives about a third of total urban China's income. The difference is also reflected in the Gini coefficient: it's 0.69 for corruption and 0.46 for total income. The gap is even more striking at the level of the top 1 percent: the top 7 cases of corruption, which is approximately equal to the top 1 percent of the defendants from our sample, account for 20.6% of total corruption, while the top 1 percent income share for disposable income is 7%. Much greater concentration of corruptions is also reflected in the Lorenz curves (see Fig. 2) with the one for corruption clearly much further away from the 45-degree line.

Concentration of corruption however has a strong similarity with the concentration of (legal) capital income as reported in household surveys. Fig. 3 shows the Lorenz curves for the two. Corruption however is more concentrated above the median, with the top decile taking (as we have seen) almost 58% of all corrupt income against 53% of capital income. The Gini coefficients for capital income, and for corruption are both 0.69.

Given the high concentration of corruption income, we can focus more on the top end. As we expect, top-end corruption follows a Pareto distribution. In Fig. 4, this is shown by looking at the top decile of corruption income. The Pareto line matches the actual data quite well with an R-square of 0.97 and a highly significant Pareto coefficient of −1.43. A glance at the graph allows to notice the likely truncation of corruption income at the very top, where the two highest recipients both show the same corruption income, and to surmise that, by prolonging the line, one might find some people with even higher corruption incomes. In other words, it seems that the full upper right-end tail of corruption is not shown in the actual data.

[28] 1 USD = 6.88 Yuan.

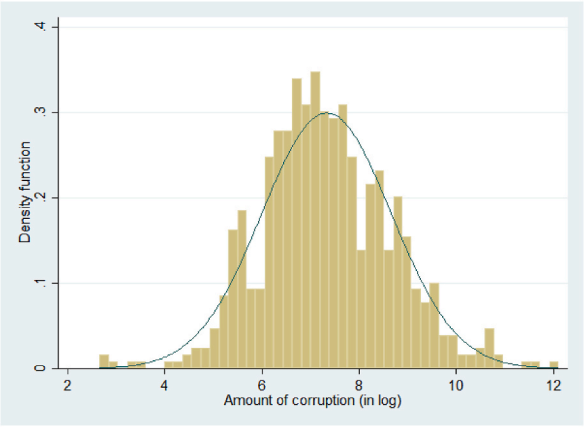

Fig. 1. Distribution of corruption (amounts in yuan)
(log-normal distribution superimposed)
Note: Bin = 50. Blue curve is normal distribution. Corruption amounts expressed in 2018 yuans. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

4.5. CPC membership and corruption

All but seven individuals accused of corruption in our dataset are CPC members. Therefore, in the analysis we do not discuss CPC membership as a binary variable (member/no member), but take membership as a given. The correspondence between membership in the Party and corruption is not only due to the presumed likelihood that top positions at which corruption takes place are mostly filled by CPC members, but also to the fact that the campaign to uproot corruption is conducted by the Party and its Disciplinary Commission and consequently the focus is almost solely on Party members.

The median age of the person convicted of corruption is 58, the median number of years of CPC membership is 34, and the median year when they joined the party is 1985. There is a noticeable absence of old people (many of them retired, and less likely to be troubled by investigation even if they might have been involved in some corruption in the past) and younger people whose positions are probably not high enough to "deserve" (or attract) much corruption. The distribution by age is also reflected in the year when defendants have jointed CPC: 75 percent of defendants have joined the Party before 1988, and 87 percent before, and including, 1992. Practically nobody among the defendants has joined the Party after 2012 when Xi Jinping became the President and the head of the Party.

5. correlates of the amount of corruption

Table 7 looks at the amount of corruption in function of the various characteristics whose bilateral relations with corruption we have just explored. We report summary statistics in appendix table AP10 in Appendix B. We show two specifications: OLS with and without provincial fixed effects. The results are practically the same and we shall discuss them together. Age and gender are not statistically significant. Variables that are significantly positively associated with the amount of corruption (at least at 5% level) are education, years of CPC membership, having joined the Party after 1978 (as opposed to the control of having joined before 1978), administrative level, and working in SOEs or social organizations (as opposed to working in the government apparatus). The fact that age, either linearly nor in a life-cycle formulation (with age squared), is not a significant predictor of the amount of corruption is important. It shows that it is other factors, like the length of CPC membership or level of education that exert their own independent effect. It is remarkable that cohorts that joined CPC after 1978, and especially so the cohort that joined the Party after 1992, tend to engage in greater corruption. Education is a significant predictor of greater corruption, but the attendance of the Central Party School has the opposite effect, almost fully offsetting the effect of graduate education on corruption.

When we turn to the type of employment (with employment in government being the control variable), SOE employment adds, statistically significantly, to corruption between 27% and 45% (between $e^{0.24}$ *and* $e^{0.37}$), and even more so employment in government-backed social organizations. Employment in public institutions (health, education etc.) is, compared to the control, negatively correlated with the amount of corruption. The most likely interpretation is that the opportunities for large scale corruption are much greater in SOEs than in health or education.

Administrative level of the defendant is very strongly positively correlated with the amount of corruption. It is notable that in both formulations (with and without provincial fixed effects), the size of the coefficient monotonically increases with the administrative level. While at the prefectural level, the amount of corruption is likely to be only 30% higher than at the sub-prefectural-or-lower (control variable) level, at the provincial or sub-provincial level, it is about 4–5.5 times greater, and at national or sub-national level, it is estimated to be 9 to 16 times greater. Finally, the four standard Chinese regions do not appear to be different from each other in respect of corruption. None of the coefficients is statistically significant.

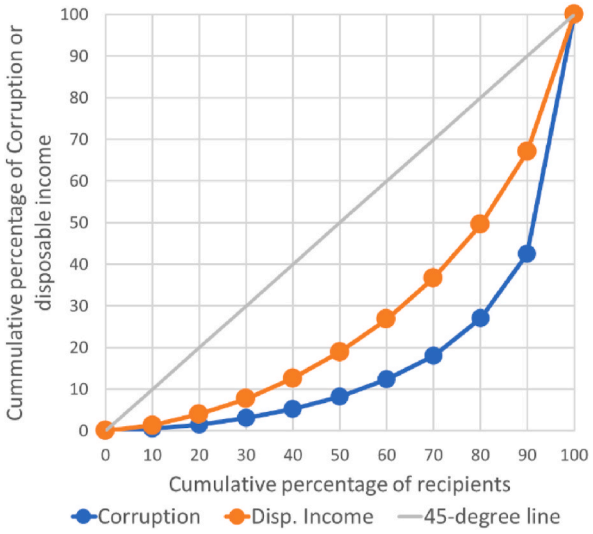

Fig. 2. Lorenz curves for corruption and disposable income
Note: The Lorenz curve for corruption refers to total amount of corrupt money over 686 individuals divided into ten deciles. The Lorenz curve for per capita disposable income refers to the 2018 data obtained from CHIP 2018as standardized by LIS, divided into ten deciles. The point x = 90% at the horizontal axis for corruption corresponds to y = 42% on y axis. This means that the highest decile of corruption recipients has received (100-42) = 58% of all corruption-related income. The richest 10 percent of the urban Chinese have in 2018 received 33% of all disposable urban income.

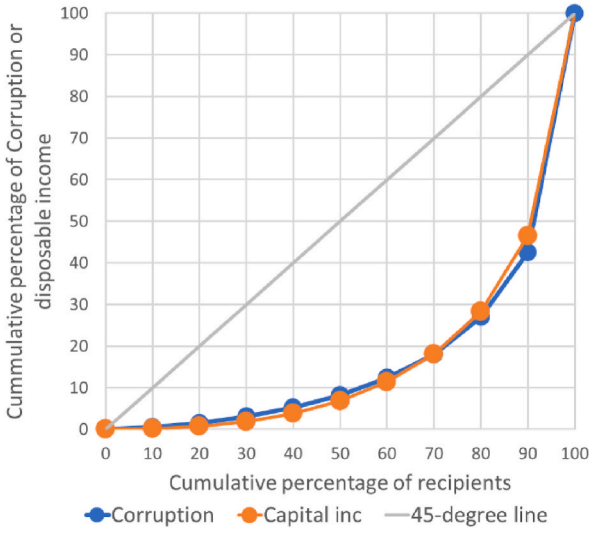

Fig. 3. Lorenz curves for corruption and income from capital
Note: The Lorenz curve for corruption refers to total amount of corrupted money over 686 individuals divided into ten deciles. The Lorenz curve for per capita income from property (or capital income) refers to the 2018 micro data obtained from CHIP 2018, with recipients divided into ten deciles. The point x = 90% at the horizontal axis for corruption corresponds to y = 42% on y axis. This means that the highest corrupt decile of recipients has received (100-42) = 58% of all corruption-related income. The richest 10 percent of the urban Chinese by capital income have in 2018 received 53% of all urban capital income.

6. Where in the distribution are the corrupt officials and how much do they gain from it?

The location of corrupt officials within the distribution of income, specifically in terms of earnings (i.e., annual wage and business income), is unstudied in the literature due to data limitations. However, leveraging the comprehensive data available in our dataset and access to separate data on China's income distribution, we can offer the initial estimates on how corruption assists officials in advancing up the income ladder.

By combining our corruption dataset with the China Household Income Project 2018(CHIP18) we are able to estimate the positions of corrupt officials within the income distribution. Using CHIP18 dataset, we estimate the earnings regression for employed adults

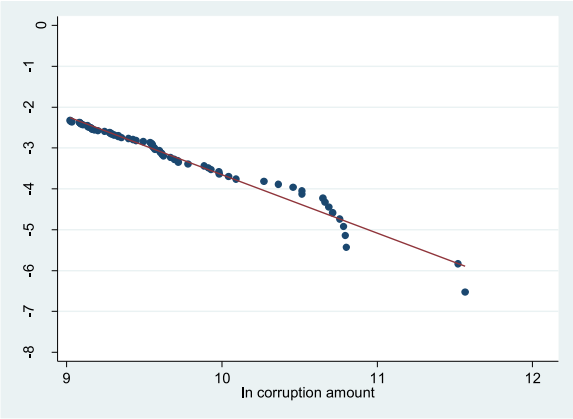

Fig. 4. Pareto relationship for the top decile of corrupt officials
Note: The horizontal axis gives the natural logarithm of amounts of corruption (in real 2018 yuans) among the top decile of recipients of corruption. The vertical axis shows the logarithm of the inverse cumulative distribution of recipients. The relationship between the two shows by how much a given increase in (log of) corruption amount reduces the percentage of recipients of such (high) corruption. The coefficient linking the two is the so-called "Pareto constant". The regression line in this figure shows that its value for the top decile is −1.43.

residing in urban areas.[29] We limit the CHIP18 sample to employed adults aged 18 to 65 who possess an urban residence permit (*hukou*). [30] The variables used in the earnings regression include gender, age, CPC membership dummy variable, education level, type of education (social or natural science), type of work contract, industry (manufacturing, energy, finance, etc.), ownership of the work unit, and leadership position variables that indicate the administrative level (prefecture, country, province) or executive status. We also control for region and apply sample weights. Table 8 displays the results of the regression. We report summary statistics in appendix table AP11 in Appendix B.

Unsurprisingly, education level, type of education, and industry sector are influential factors in the earnings regression analysis. However, the variables that are of particular interest to us, given the profile of individuals who engage in corrupt activities, are those related to leadership/executive positions and the corresponding administrative level. These variables are highly significant in the earnings regression, which aligns with our expectations that higher-level positions have a significant impact on earnings.

We proceed to predict the earnings of corrupt officials in the absence of corrupt activities, utilizing the estimated coefficients from the earnings regression and the characteristics of the corrupt officials from our corruption dataset. The median of the predicted annual earnings of corrupt officials is about 157,000 yuans per year,[31] which is more than three times the median of the earnings in the CHIP18 urban subsample (48,000 yuans), and would also place such an official at the 95th percentile of the urban income distribution. The results regarding the mean are similar: the ratio mean-to-mean is in excess of 3 (see Table 9, column 3). This indicates that the corrupt officials are not just a random sample of the urban working population. They are significantly better off (excluding their corrupt earrings) than the average urban worker and belong to the top portion of the distribution.

To estimate the officials' total earnings, we add the amount of corrupt earnings to their predicted legal earnings. Corruption may be considered as a "bonus" or "rent". The challenge is to determine the size of the "bonus" accurately. As explained before, the stock of corruption is expressed in 2018 prices. That stock however was acquired over the years. We annualize it by estimating the number of years of corruption (N) in two different ways.[32] First, we estimate N_i for each corrupt official using the number of years they have been a member of the CPC.[33] This is an individual-based annualization, and is justified by the likelihood that the membership of the party was a facilitating, or perhaps often the indispensable, condition for getting involved in corruption. The second assumption is simpler: we use a given number of years of corruption across the board, that is the same for all individuals. We present the results for individual Ns and for $N = 20$ in the main text, and the results for other Ns in the Appendix (see Appendix B, Table AP2).[34] In all cases, the corruption "bonus" gives a dramatic increase to the annual income of the corrupt officials.

When we annualize the overall amount of corruption by the number of years of CPC membership, the new mean income of corrupt

[29] The earnings include after-tax wages plus fringe benefits, and net business income. The regression is run across individuals.
[30] We thus exclude rural-urban migrants.
[31] This is about $23,000 at 2018 average exchange rate.
[32] Under both scenarios, the implicit assumption is that yearly corruption is constant. That, of course, is unlikely to be true in real life but we do not have information on the dynamics of corruption. Furthermore, our objective is to "locate" individuals in the "average" income distribution of the urban China over a number of years, and consequently the variability of his or her position between the years Is of little importance.
[33] The corruption amount could be interpreted as the "compensation" the corrupt officials took illegally for themselves since in Chinese public administration, individuals are often overly qualified for the tasks they are assigned and are underpaid relative to their qualifications (Wu and Wang, 2018).
[34] While N is equal to 30, 40 or 50.

Table 8
Earnings regression estimates (based on CHIP18).

	Coefficient	Std. error
Male	0.263[a]	(0.021)
Age	0.089[a]	(0.014)
Age^2	−0.001[a]	(0.000)
CCP membership dummy	0.038[c]	(0.012)
Education level (baseline: secondary education or below)		
Master's degree or above	0.755[a]	(0.048)
Bachelor's degree (*benke*)	0.421[a]	(0.037)
Junior college (*zhuanke*)	0.190[a]	(0.029)
Major of education (baseline: others)		
Natural Sciences	0.164[a]	(0.015)
Social Sciences	0.107[b]	(0.023)
Permanent contract dummy	0.207[b]	(0.045)
Industry of work unit (baseline: others)		
Finance	0.073	(0.034)
Energy	0.067[b]	(0.018)
Transportation	0.058	(0.037)
Manufacturing	0.163[b]	(0.030)
Media, Culture, and Tourism	0.052	(0.038)
Public Sector and Social Organization	−0.090[c]	(0.034)
Ownership of work unit (baseline: others)		
Government and Party Agencies	0.105[a]	(0.009)
Public institutions	0.071[b]	(0.014)
SOEs	0.072[a]	(0.009)
Dummy for principal at work unit	0.261[a]	(0.006)
Dummy for prefectural or higher level	0.407[c]	(0.169)
Dummy for county level	0.213[b]	(0.066)
Dummy for executive in enterprise	0.499[a]	(0.068)
Regions (baseline: central)		
East	0.267[a]	(0.003)
Northeast	−0.075[a]	(0.007)
West	0.074[a]	(0.004)
Constant	8.380[a]	(0.265)
Cluster	**Regional level**	
Observations	9229	
R-squared	0.272	

Notes: OLS regression. Standard errors are robust, clustered at the provincial level.
Our analysis is based CHIP 2018. We restrict the sample to employed adults aged 18 to 65, who live in urban area of China with urban *hukou*.
[a] $p < 0.01$.
[b] $p < 0.05$.
[c] $p < 0.1$.

officials (legal plus corrupt income) is estimated at 7.6 times their legal earnings, and when the annualization is uniform (20 years of corruptions), the new mean income is 13.6 times greater than the legal income (see Table 9, column 5). It is clear that, for most corrupt officials, the amount of corrupt income is several fold greater that the amount of legal income. Corruption therefore advances the corrupt officials' income position significantly up. Moreover, given that corruption that we observe here is an upper-income phenomenon, and that it is itself very unequally distributed (Gini of 0.69 as mentioned in Section 4.4), it is clear that corruption increases overall urban inequality in China.

Column 1 in Table 10 shows at what percentile of income distribution in urban China the corrupt officials would be when using their predicted legal income alone. A very large majority of the corrupt officials (80%) would be in top ten percent of the distribution; 50% of the corrupt officials would be in top five percent, 6.5% would be in top one percent, and 3.4% would be in the top 0.5 percent. As we have already mentioned, the corrupt officials –even without corrupt earnings—are placed fairly high in the Chinese urban income distribution.

But corruption makes them move up higher. While one-half of corrupt officials were in the top 5 percent when corruption was not taken into consideration, with corruption, between 98 and 99 percent are in the top 5 percent, and 43%–60% of them (depending on the assumption of N) became part of 0.1 percent.[35]

[35] There is an issue of endogeneity here. When we add the corrupt income, these new values do not affect the distribution of legal income from CHIP18. In other words, if we had the distribution of earnings that would include both legal and illicit earnings, the position of corrupt officials would be somewhat lower than as shown here. The income distribution that we use to find out where the corrupt officials are is in principle an income distribution of legal incomes. If there are, in addition to the illegal incomes considered here, other illegal incomes that have not been found out, the "true" income distribution has higher incomes throughout and hence the position of the corrupt officials may be somewhat lower than estimated.

Table 9
Annual earnings of the overall urban sample of workers and of the corrupt officials.

	1	2	3	4		5	
	Annual earnings from CHIP2018 (all urban residents)	Estimated annual earnings of corrupt officials (before corruption)	Ratio (2): (1)	Predicted annual income of corrupt officials (including the corruption bonus)		Increase in income due to corruption (times)	
				Using years of CPC membership	Assuming 20 years of corruption	Using years of CPC membership	Assuming 20 years of corruption
Mean	46,019	169,031	3.7	1,292,409	2,305,980	7.6	13.6
Median	48,370	156,888	3.2	634,706	906,843	4.0	5.8
No of observations	9229	1322		567	642		

Notes: The complete corruption dataset contains 1451 observations (crimes 1 to 7). However, due to missing values of some explanatory variables used for the earnings regression in Table 8, we could estimate the annual earnings of only 1322 convicted officials as shown in column 2. The difference in the number of observations in Column 4 for the total number of officials convicted for acceptance bribes and dereliction of duties is caused by the lack of data on the years of CPC membership for some of them.

Table 10
Change in the rankings in income distribution of officials before and after corruption (all calculated in 2018 yuans).

	Percentage of officials ranked above the threshold (without corrupt income)	Percentage of officials ranked above the threshold (including corrupt income)	
		Using years of CPC membership*	Assuming 20 years of corruption
Top 40 percent	99.6	100	100
Top 10 percent	79.3	99.5	99.7
Top 5 percent	54.5	97.98	99.5
Top 1 percent	6.5	82.0	91.1
Top 0.5 percent	3.4	75.0	86.9
Top 0.1 percent	0	43.0	60.3
Number of observations	1322	567	642

Notes: The complete corruption dataset contains 1451 observations (crimes 1 to 7). However, due to missing values of some explanatory variables used for the earnings regression in Table 8, we could estimate only the annual earnings of 1322 convicted officials as shown in column 2. Appendix B, Table AP1 presents comparable results obtained from a restricted sample that includes only 567 convicted officials.

7. Conclusions

The government- and Communist Party of China-led anti-corruption campaign has probably for the first time in history allowed researchers to access a consistent dataset on corrupt officials. The Information includes name, gender, age, education, duration of membership in the ruling party, position of authority, the amount of embezzled money and several other characteristics.

The studies of corruption have so far been hampered by the lack of similar data for those accused or convicted of corruption. Thus they were limited to individual or ethnographic case studies, general discussions of corruption, or use of expert opinions to gauge the extent of corruption. This was due to the fact that corruption was almost never prosecuted as a matter of specific policy; the cases were dealt with sporadically, at different courts, and information was neither uniform in its form (i.e. the same information was not available for each defendant) nor centralized.

Our empirical knowledge of characteristics of people who engage in corruption, where in the income distribution they are, and the amounts of money embezzled has therefore been limited. The database on 686 officials convicted of corruption in the period 2012-21 that we have constructed allows us, in conjunction with China-wide 2018 urban income survey, to predict corrupt officials' legal income and to estimate where in China's urban distribution they would be in the absence of corruption. We are thus able for the first time to calculate gains from corruption relative to corrupt officials' legal income and to estimate their income distribution gains (i.e. their pre- and post-corruption income distribution positions).

It should be noted that our database, given the objectives of the campaign, covers (a) almost exclusively CPC members, (b) people with high level of education and reasonably high executive power whether in the Party or government apparatus, SOEs or elsewhere, and therefore (c) people whose legal income would have been rather high.

We consider three related, but separate, questions. First, what are the correlates of corruption (measured by the amount of money individuals are accused of having embezzled); second, what are the features of corruption as a source of income (treating it as any other source of income); and third, what are the gains, both in relative terms and in income position, that the individuals accused of corruption realize.

We find that that age and gender are not correlated with the amount of corruption; neither do different regions of China display differences in the determinants of corruption. The use of provincial dummies does not make much of a difference either. The variables that are positively associated with the amount of corruption are education, administrative level at which the corrupt official is, number of years of CPC membership, and having joined the party after 1978, and even more so after 1992. We also use information on whether the defendant has graduated from the CPC party school and find that it is negatively correlated with the amount of corruption.

When we consider corruption as an alternative source of income and compare its distribution with the distribution of disposable income from the 2018 income survey of urban population, we find that corruption is much more unequally distributed than disposable income (Gini of 0.69 vs. 0.47), and that its distribution is very similar to the distribution of income from capital.

The estimated median legal earnings of corrupt officials are 3.2 times greater than the overall urban median. Consequently, about 80% of them would belong to the top urban decile and 6% to the top percentile, even without corrupt earnings. However, corrupt earnings are huge and allow them to climb much higher in income distribution hierarchy. Thanks to corruption, the median earnings of corrupt officials surge by between 4 and 5.8 times (depending on the assumption used to annualize the data), and between 82% and 91% percent of corrupt officials "end up" in the urban top one percent.

CRediT authorship contribution statement

Li Yang: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Branko Milanovic:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Yaoqi Lin:** Software, Investigation, Data curation.

Declaration of competing interest

There is no confict of interest.

Data availability

Data available within the article or its supplementary materials.

Appendix A. Construction of "Tigers Corruption" Dataset

We name our corruption dataset "Tigers' Corruption Dataset", since it includes only the corruption cases of the high-ranking officials. Starting in 2012, the Central Commission for Discipline Inspection of the Communist Party and the National Supervisory Commission of China (CCDI) has been regularly updating corruption cases involving high-ranking officials in its website (https://www.ccdi.gov.cn/scdc/). This website serves as a platform to showcase the progress and results of anti-corruption efforts. The website reports on two types of cases: officials who have been investigated and officials who have received administrative punishment. Our dataset only includes cases in which officials received administrative punishment.

CCDI has categorized theses convicted officials into three types: Centrally-Managed Cadres (CMC), Provincially-Managed Cadres (PMC), and central-level cadres (CLC) from the Party, state institutions, state-owned enterprises, and financial institutions (excluding CMC or CPC). We have collected data for each type within the timeframe specified below.

Type of Cadres	The last data access date	No. of cases
CMC	04/30/2021	227
PMC	05/28/2021	1105
CLC	04/23/2021	119

We use RStudio and the R package rvest to collect data from the main data sources (websites). The detail information about this package can be found here or here.

Once we collected the information of the corruption cases of the convicted officials, we then collect the demographic and employment information of each convicted officials from Baidu Baike (https://baike.baidu.com)

Variables provided in the main data sources include.

- Personal Information: Name, Gender, Birth Year, Birth Province, Birth City, Year join CCP, Highest education degree (other than CCP school), Field of Study (Other than CCP school), CCP School, CCP School education level, CCP School Majorschool), CCP School, CCP School education level, CCP School Major
- Job Information: CCDI Classification, Type of position (Government, Public institution or Enterprise), Job Title, Position Starting, Position Ending, Administration Level, Job Province, Job City
- Case Information: Date starting the investigation, Date of the Administration and Party punishment, Type of crime (Crime of acceptance of bribes and crime of dereliction of duty, Crime-Ganster, Crime-Drugs/Sex/Others, Bribery, Crime of intentional homicide, Illegal access to national secrets, Malfeasance), Amount of illegal money involved (Amount of Money Corrupted, Currency, Money cannot explain the source, Money-Illegal Possession, Illegal Earnings, Illegal Use Public Money, Bribery, Time of Judgement), Sentence (Term of Imprisonment, Penalty Amount, Other Punishment, Death Penalty, Death Sentence with Reprieve)

In cases where essential variables were missing from the primary data source (such as Birth Province, Birth City, Year join CCP, Administration Level, Amount of Money Corrupted), we conducted an extensive search to supplement our dataset using various online platforms listed below.

- Xinhua News Agency (http://www.xinhuanet.com)
- The Paper (https://m.thepaper.cn)
- The State Council, The People's Republic of China (http://www.gov.cn)
- Reuters (http://www.reuters.com)
- Sina Corporation (https://news.sina.com.cn; https://finance.sina.com.cn)
- The Chinese Court Net (http://www.chinacourt.org/)

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejpoleco.2024.102559.

References

Abed, George T., Gupta, Sanjeev, 2002. "The Economics of Corruption: an overview." Governance, Corruption and Economic Performance. International Monetary Fund, Washington, pp. 1–16.

Aidt, Toke S., Hillman, Arye L., Qijun, Liu, 2020. Who takes bribes and how much? Evidence from the China Corruption Conviction Databank. World Dev. 133, 104985.

Alt, James E., Lassen, David Dreyer, 2014. Enforcement and public corruption: evidence from the American states. J. Law Econ. Organ. 30 (2), 306–338.

Capasso, Salvatore, Santoro, Lodovico, 2018. Active and passive corruption: Theory and evidence. Eur. J. Polit. Econ. 52, 103–119.

Chen, Yanting, Liu, Qijun, 2018. Public-sector wages and corruption: an empirical study. Eur. J. Polit. Econ. 54, 189–197.

Chih, Yao-Yu, Demir, Firat, Hu, Chenghao, Liu, Junyi, Shen, Hewei, 2023. A spatial analysis of local corruption on foreign direct investment: evidence from Chinese cities. Eur. J. Polit. Econ. 79, 102443.

Choudhury, Sanchari, 2019. WTO membership and corruption. Eur. J. Polit. Econ. 60, 101806.

Demir, Firat, Hu, Chenghao, Liu, Junyi, Shen, Hewei, 2022. Local corruption, total factor productivity and firm heterogeneity: empirical evidence from Chinese manufacturing firms. World Dev. 151, 105770.

Dong, Bin, Torgler, Benno, 2013. Causes of corruption: evidence from China. China Econ. Rev. 26, 152–169.

Glaeser, Edward L., Saks, Raven E., 2006. Corruption in America. J. Publ. Econ. 90 (6–7), 1053–1072.

Gutmann, Jerg, Padovano, Fabio, Voigt, Stefan, 2020. Perception vs. experience: Explaining differences in corruption measures using microdata. Eur. J. Polit. Econ. 65, 101925.

Gründler, Klaus, Potrafke, Niklas, 2019. Corruption and economic growth: new empirical evidence. Eur. J. Polit. Econ. 60, 101810.

Gustafsson, Björn, Shi, L.I., Hiroshi, Sato, 2014. Data for studying earnings, the distribution of household income and poverty in China. China Econ. Rev. 30, 419–431.

Hillman, Arye, 2010. Expressive behavior in economics and politics. Eur. J. Polit. Econ. 26, 403–418.

Jain, Arvind K., 2001. Corruption: a review. J. Econ. Surv. 15 (1), 71–121.

Jetter, Michael, Parmeter, Christopher F., 2018. Sorting through global corruption determinants: institutions and education matter–Not culture. World Dev. 109, 279–294.

Jiang, Junyan, 2018. Making bureaucracy work: patronage networks, performance incentives, and economic development in China. Am. J. Polit. Sci. 62 (4), 982–999.

Liu, Qijun, Peng, Yaping, 2015. Corruption in college admissions examinations in China. Int. J. Educ. Dev. 41, 104–111.

Lorentzen, Peter L., Lu, Xi, 2018. Personal ties, meritocracy, and China's anti-corruption campaign. Meritocracy, and China's Anti-Corruption Campaign (November 21, 2018).

Manion, Melanie, 2016. Taking China's anticorruption campaign seriously. Economic and Political Studies 4 (1), 3–18.

Milanovic, Branko, 2019. Capitalism, Alone: the Future of the System that Rules the World. Harvard University Press.

Mocetti, Sauro, Orlando, Tommaso, 2019. Corruption, workforce selection and mismatch in the public sector. Eur. J. Polit. Econ. 60, 101809.

Olken, Benjamin A., 2009. Corruption perceptions vs. corruption reality. J. Publ. Econ. 93 (7–8), 950–964.

Pei, Minxin, 2016. China's Crony Capitalism: the Dynamics of Regime Decay. Harvard University Press.

Piketty, Thomas, Yang, Li, Zucman, Gabriel, 2019. Capital accumulation, private property, and rising inequality in China, 1978–2015. Am. Econ. Rev. 109 (7), 2469–2496.

Pittaluga, Giovanni B., Seghezza, Elena, Morelli, Pierluigi, 2023. Media fabrication of corruption and the quality of the political class: the case of Italy. Eur. J. Polit. Econ., 102461

Schulze, Günther G., Sjahrir, Bambang Suharnoko, Zakharov, Nikita, 2016. Corruption in Russia. J. Law Econ. 59 (1), 135–171.

Shi, Xiangyu, 2022. Anti-Corruption Campaign and Economic Growth in Chinese Cities: the Dichotomous Mechanism of Network-Based Political Competition. Available at: SSRN 4644117.

Targa, Matteo, Yang, Li, 2023. The Impact of Communist Party Membership on Wealth Distribution and Accumulation in Urban China. Available at: SSRN:. ZEW - Centre for European Economic Research Discussion Paper No. 23-057 https://ssrn.com/abstract=4662171.

Tong, Shenghui, 2022. Corruption and anti-corruption in China: a review and future research agenda. Asian Pac. Econ. Lit. 36 (1), 3–16.

Wu, Na, Wang, Qunyong, 2018. Wage penalty of overeducation: new micro-evidence from China. China Econ. Rev. 50, 206–217.

Xie, Chuntao, 2016. Fighting Corruption: How the CPC Works. New World Press, Beijing.

Xie, Yu, Zhou, Xiang, 2014. Income inequality in today's China. Proc. Natl. Acad. Sci. USA 111 (19), 6928–6933.

Yang, Li, Novokmet, Filip, Milanovic, Branko, 2021. From workers to capitalists in less than two generations: a study of Chinese Urban top group transformation between 1988 and 2013. Br. J. Sociol. 72 (3), 478–513.

Zhuang, Juzhong, Shi, Li, 2016. Understanding Recent Trends in Income Inequality in the People's Republic of China, vol. 489. Asian Development Bank Economics Working Paper Series.

Zimelis, Andris, 2020. Corruption research: a need for an integrated approach. International Area Studies Review 23 (3), 288–306.