The Shapeshifting Hat: Strategic Response to Fiscal and Political Incentives of Chinese National Poor Counties*

Jianhao Lin† Tingwei Luo‡ Wenbiao Sha§ Jiasong Xie¶

May 2024

Comments very welcome.

Abstract

We investigate the underlying motivations that drive local governments to manipulate GDP figures in the Chinese context of National Poor County (NPC) designation and cancellation. Using a county-year panel that combines satellite nightlight data with self-reported GDP statistics, we document that, compared to counties that were never designated, being selected as an NPC in 2011, which involves substantial fiscal transfers, is associated with a greater degree of GDP underreporting in the year of designation, suggesting a response to fiscal incentives. This underreporting effect is more pronounced for counties with worse fiscal conditions. We also find that national poor counties would overreport GDP figures to a greater degree since 2016, after the central government set poverty eradication by 2020 as a non-negotiable task, suggesting a response to political incentives. This overreporting effect appears to be persistent, exists only in provinces that lack statistical self-discipline, and can be mitigated by central statistical supervision, thus addressing the principal-agent problem. Our results are unlikely to be driven by differences in tenure years or individual characteristics of county political leaders.

Keywords: national poor counties, data manipulation, fiscal transfers, political tasks
JEL Classification: H70, H76, O20, O22, R58

*We thank Kai Li, Yupeng Lin, Ming Lu, Xianxiang Xu, Li Zhang, Jun Zhang, and seminar and conference participants at Fudan University, Sun Yat-sen University, the 5th International Conference of China Development Studies, and the China Young Economists Society Annual Meeting for many helpful comments and suggestions. Authors are listed alphabetically and first authorship is shared equally. Any remaining errors are our own.

†Lingnan College, Sun Yat-sen University. Email: linjh3@mail.sysu.edu.cn.
‡Lingnan College, Sun Yat-sen University. Email: luotw5@mail2.sysu.edu.cn.
§Lingnan College, Sun Yat-sen University. Email: wenbiao.sha@outlook.com.
¶School of Economics, Dongbei University of Finance and Economics. Email: xiejs@dufe.edu.cn.

1 Introduction

State effectiveness is central to economic development (Besley, Burgess, Khan and Xu, 2022). In developing countries with decentralized governance, local governments and bureaucrats often have a great deal of discretion over development projects that involve substantial resources in the form of foreign aid or domestic transfers.[1] By linking donors or central governments to poor households, local governments and bureaucrats managing development projects are expected to make the service work for the poor and promote economic development more broadly (World Bank, 2003). However, this may not always be the case, possibly because of misaligned incentives among the various parties involved (i.e., the poor, local governments, and the donor or central government), not to mention the pervasive corruption and the lack of accountability in developing countries (Olken and Pande, 2012). For example, one unfavorable outcome well documented in the literature is elite capture, where political leaders or their relatives, if not both, benefit more from development projects in different settings.[2] In this study, we document another type of unfavorable consequence of development projects — data manipulation by local governments — by examining the role of county-level governments in implementing one of the world's largest poverty alleviation programs in China.

There is no doubt that China has made significant progress in reducing poverty since economic reforms began in 1978 (e.g., Chen and Ravallion, 2021). As detailed in World Bank (2022), China's poverty rate fell by 2.3 percentage points per year from 1978 to 2020, and the number of poor people in rural China was reduced by an average of 18.7 million per year since 1978. However, the lessons for other developing countries, particularly the shortcomings that can arise in the implementation of anti-poverty programs and that can have efficiency implications, are less well studied. In this paper, we explore the National Poor County (NPC) designation and cancellation in China,[3] one of the poverty alleviation programs implemented by the Chinese central government from the mid-1980s to 2020, to study how local governments respond to different incentives that can lead to local behavioral distortions and thus deviations from the central policy goal. Specifically, we examine what motivates local governments to manipulate economic figures in order to qualify for NPC designation (acquiring the poverty hat) or cancellation (removing the poverty hat), and perhaps more interestingly, whether local government behavior may be biased in different directions under these two different circumstances.

The Chinese governance system is characterized as economically decentralized but politically centralized (Xu, 2011; Bardhan, 2020). Given this institutional setting, local political leaders often multitask and thus face multiple incentives that may be misaligned with the central government.[4]

[1]Even for some specific services that are best provided by for-profit companies or nongovernmental organizations, governments with regulatory authority still have an important role to play (Page and Pande, 2018).

[2]Elite capture can occur at both local and national levels (Bardhan and Mookherjee, 2000). For example, see Reinikka and Svensson (2004) and Park and Wang (2010) for local capture of government transfer programs and Andersen, Johannesen and Rijkers (2022) for elite capture of foreign aid. Scholars have also provided evidence that the welfare consequence of certain types of elite capture in Indonesian villages is not economically large (Alatas, Banerjee, Hanna, Olken, Purnamasari and Wai-Poi, 2019).

[3]Throughout this paper, "NPC" stands for "national poor county," and "NPCs" refers to "national poor counties."

[4]In addition to our empirical case, other cases of multitasking local governments in China are provided by Chen, Li and Lu (2018) and Cao, Weng, Xu and Zhou (2023), both of which study local governments' trade-off between environmental protection and GDP growth, and by Fisman, Lin, Sun, Wang and Zhao (2021), who study the trade-off between economic and public health considerations of Chinese cities' COVID-19 reopening plans.

First, local leaders are responsible for the socioeconomic functioning within their jurisdictions, ranging from paying public employees to financing development projects to providing public goods and services, and so on. In this regard, local leaders have an incentive to maximize broadly defined fiscal revenues, which we call fiscal incentives.[5] Second, the central government often sets economic or development targets for localities (e.g., growth targets, pollution controls, social stability, and poverty alleviation, to name a few), some of which are often mandatory or sometimes even politicized. Therefore, local leaders have an incentive to meet these targets,[6] which we call political incentives. Third, it is widely known that Chinese local officials are concerned with career incentives linked to their economic performance, especially at the county level (Landry, Lü and Duan, 2018; Wiebe, 2020).[7] In other words, they have an incentive to promote the local economy, which may come at the expense of not meeting the central goals or not paying street-level service providers well. Given that local budgets and resources are often limited, it is very difficult to make all incentives aligned across tasks, so local leaders have their own priorities about which ones to respond to more strongly.[8] Relatedly and importantly, the priority of local governments may differ from the goal of the central government, which in our case is to eradicate rural poverty.

In addition, accompanied and complicated by information asymmetry, local governments may also misreport their performance measures to the central government. There are two layers of government (provincial and prefectural) between the central government and county governments,[9] which leaves room for asymmetric information problems. The lack of local information makes central monitoring and evaluation difficult (e.g., Huang, Li, Ma and Xu, 2017), and therefore who collects information matters in such central-local relations (Finan, Olken and Pande, 2017). If information on the performance measure is collected by local governments and there are no independent sources of verification for the central government, county governments may have strong motives to manipulate the performance measure, which in our context is county-level GDP figures.

Conceptually, the fiscal and political incentives we study in this paper can be viewed as specific types of career incentives. After all, both can affect the promotion prospects of local leaders. Our focus is on empirically identifying the specific incentives at work in a given situation. In the empirical analysis, we are also interested in the promotion incentives of local political leaders originating from difference in tenure years or individual characteristics, mainly examining the extent to which our baseline results could be driven by such promotion incentives. Of course, an empirical challenge is to identify which of the incentives is more important in a given situation. We are able to address this challenge by using detailed county-year panel data with information on local leaders.

Applying the above logic to our empirical setting, we argue that poor Chinese counties respond strongly to fiscal incentives in the designation of NPCs and to political incentives in the cancellation of NPCs, independent of promotion incentives. In the empirical analysis, we are interested in identifying the relationship between being selected or removed as a national poor county and GDP

[5] For example, see Jin, Qian and Weingast (2005) and Han and Kung (2015).

[6] Failing to meet such targets would disqualify local political leaders (e.g., see Fang, Liu and Zhou, 2023). For pollution controls, see He, Wang and Zhang (2020). For growth targets, see Lyu, Wang, Zhang and Zhang (2018). For social stability, see Wen (2024) and King, Pan and Roberts (2013).

[7] See also Li and Zhou (2005) and Che, Chung and Qiao (2021).

[8] In other words, there are effort substitutions in such a setting. For related theoretical analysis, see Holmstrom and Milgrom (1991), Dewatripont, Jewitt and Tirole (1999), and Dewatripont, Jewitt and Tirole (2000), among others.

[9] There are five layers in China's administrative structure: the central, provincial, prefectural, county, and township governments.

misreporting by county governments. We focus on GDP figures because GDP is the main economic indicator that captures most of the relevant information in both NPC designation and cancellation. In terms of NPC designation, once a county is selected as a national poor county, it will receive numerous fiscal transfers from the central and higher-level governments, which can largely alleviate fiscal distress of poor counties. Consequently, county leaders might be tempted to underreport GDP statistics in order to qualify for and increase the chances of being designated as a national poor county. In terms of NPC cancellation, the central government mandates the elimination of rural poverty by 2020 as a critical policy objective. In response to this directive and the associated political pressure, county leaders might inflate GDP statistics to demonstrate progress toward poverty alleviation goals, as failing to meet these targets could result in political demotion.

Our focus on GDP manipulation is motivated by rich anecdotal evidence and academic research. Indeed, GDP manipulation is not uncommon in China. For example, before becoming premier, Li Keqiang reportedly expressed skepticism about the accuracy of China's official GDP figures and proposed using alternative indicators to measure the real economic situation. Many Chinese localities have been caught inflating GDP statistics, and some local leaders have been prosecuted for manipulating GDP figures for individual promotions (see Appendix Figure A2 for media coverage). As for the NPC designation and cancellation, it is also no secret that counties, including rich ones, competed for NPC status by fabricating data, and afterward the designated counties allegedly managed to get out of poor status by also manipulating economic data (see Appendix Figure A3 for media coverage). As for academic research, many have provided empirical evidence that China's official GDP statistics are subject to overreporting bias (Chen, Chen, Hsieh and Song, 2019; Chen, Qiao and Zhu, 2021; Martinez, 2022, among others).

To measure GDP manipulation, we use the log ratio of GDP per capita to the average digital number, which measures the luminosity of night lights, for a given county in a given year. We adjust GDP per capita using the provincial GDP deflator based on 2000 price levels. This ratio measure is not only simple to construct but also easy to interpret, allowing us to compare the relative degree of data manipulation across counties. Our empirical results are also robust to using measures constructed using different sources of nighttime light data or the method proposed by Henderson, Storeygard and Weil (2012), suggesting that the ratio measure of GDP manipulation is largely reliable in our empirical context.

We begin by identifying the relationship between the NPC designation and GDP misreporting. China has conducted four waves of NPC designation since the mid-1980s. Our first set of empirical analyses examines the final wave of designation in 2011. We use a difference-in-differences approach. Our treatment group is the counties that were designated as NPCs in 2011, while we use counties that were never designated as our control group. All else being equal, the never-designated counties would lack fiscal motivations stemming solely from anti-poverty funds, making these counties a clean control group. We use 2010 as our post-treatment period because the 2011 selection was based on economic statistics in 2010. The main sample is a county-year panel for the period 2003-2010, which allows us to test the parallel pre-trends assumption — the key assumption for difference-in-differences methods. We control for interactions between variables capturing initial economic conditions measured in the base year and the year dummies, allowing counties with different initial conditions to have different time trends. Our preferred specification also includes county and province-by-year fixed effects, the former mainly exploiting variation in changes before and after selection, and the latter further comparing treated and control counties within the same province-year cells.

Our empirical results show that, compared to never designated counties, being selected as a national poor county in 2011 is associated with a greater degree of underreporting GDP figures in 2010, suggesting a response to fiscal incentives. We find no evidence of different pre-trends between the treatment and control groups. One might be concerned that our results are driven by career incentives rather than the fiscal incentives we study in this paper. Our results also remain largely unchanged even when we add county party secretary tenure dummies to our preferred specification or replace the county fixed effects with party secretary fixed effects, which account for possible career incentives due to different leaders in different years in office or with different individual characteristics. In addition, we find that this underreporting effect is more pronounced for counties with more severe fiscal distress, such as greater reliance on fiscal transfers or larger budget deficits prior to the NPC designation.

Using the counties selected in 2011 as our treatment group is somewhat ad hoc, which may bias our estimates, so we also use a unique institutional arrangement to conduct instrumental variable estimation and fuzzy regression discontinuity (RD) analysis. In the 2011 NPC designation, each province was given a quota of how many counties can be designated as NPCs, which is fixed and predetermined before the selection. Thus, within a given province, counties can rank themselves in terms of GDP per capita and then use the quota to calculate a threshold below which counties would be eligible for NPC designation. We define our instrument as the distance between a county's 2009 GDP per capita and the threshold calculated using 2009 data. We choose 2009 because counties are likely to manage their 2010 GDP statistics based on 2009 data. Obviously, the distance can be positive, zero, or negative, and is negatively correlated with the probability of being designated in 2011,[10] which ensures the relevance of our instrument. Since the quotas are predetermined, the GDP manipulation behavior of county governments is unlikely to affect our instrument, which largely ensures the exclusion restriction. Our instrumented estimates obtained using panel data are larger than the baseline estimates. This could be due to the fact that the difference-in-differences estimates do not capture the impact of non-selected counties whose distance is around the zero distance point and thus also have incentives to underreport GDP figures. In addition, a fuzzy RD analysis using the 2009 distance as the running variable and cross-sectional data also yields consistent results.

Our second set of empirical analyses examines the relationship between NPC cancellation and GDP misreporting. In 2015, the Chinese central government announced a policy goal to eradicate rural poverty by 2020, which was reaffirmed in 2018. The national poor counties have been gradually removed since 2016. We again employ a difference-in-differences strategy, with the NPCs selected in 2011 as the treatment group and the counties never selected as the control group. We use county-year panel data for the period 2013-2020. In the regressions, we control for interactions between county base year characteristics and the year dummies, and include county and province-by-year fixed effects. Given the gradual elimination of NPCs over time, our empirical strategy now has

[10] On the one hand, the richest counties are certainly the ones that are never selected, and the poorest counties are the ones that are always selected. Therefore, both groups of counties have little incentive to manipulate their GDP figures. On the other hand, counties just above or just below the threshold are very likely to underreport their GDP statistics, since doing so increases their chances of being selected.

staggered variation in treatment timing, which makes the estimated coefficients difficult to interpret (Goodman-Bacon, 2021). To address this concern, we also report estimates obtained using recently proposed estimators in the literature.

The empirical results show that the national poor counties designated in 2011 would overreport GDP figures to a greater degree since 2016, suggesting a response to political incentives. As in the previous tests, we also add county party secretary tenure dummies to our preferred specification or replace county fixed effects with party secretary dummies to account for possible confounding effects stemming from promotion incentives, and find that our results remain largely unchanged. Perhaps more importantly, to address the concerns raised by the staggered variation in NPC cancellation, we also show that our results are largely robust to the new estimators proposed by de Chaisemartin and d'Haultfoeuille (2020), Callaway and Sant'Anna (2021), Sun and Abraham (2021), and Borusyak, Jaravel and Spiess (2023).

To test the assumption of parallel pre-trends and the persistence of the effects, we estimate an event study design, a dynamic variant of our difference-in-differences strategy. Again, we find no evidence of differences in pre-treatment trends between the two groups. Interestingly, we find that the overreporting effect is persistent. This result may reflect the fact that the removed counties may still be underdeveloped and unable to promote development and growth without fiscal transfers, and thus the incentive to overreport GDP figures also persists. For example, Qiu, Shi, Li and Luo (2023) show that China's national poor counties allocated more fiscal resources to poverty alleviation, which reduced fiscal support for industrial and enterprise development. Therefore, these counties are likely to be underdeveloped in terms of local development. Liu and Ma (2019) provide evidence that the NPC program failed to promote both short- and long-term economic growth, in part due to local capture. In addition, Zhu, Liu and Li (2021) show that ending the national poor county program reduced the county fiscal expenditure-to-GDP ratio. Notably, our event study estimates are also robust to the new estimators.

Finally, we provide two pieces of evidence that the overreporting effect can be mitigated. First, we conjecture that regions with a high level of statistical self-discipline, which we relate to internal motives, may have little or no incentive to inflate GDP data.[11] We rely on the National Economic Census conducted by the National Bureau of Statistics (NBS) of China to distinguish regions with better statistical self-discipline from those without. We define self-disciplined provinces as those whose self-reported GDP is lower than the census-based GDP calculated by the NBS. Our regression results consistently show that the overreporting effect is substantially reduced for counties located in provinces with statistical self-discipline. Second, we expect that central government monitoring programs, which we relate to external pressure, would deter local governments from overreporting GDP figures.[12] The NBS often conducts statistical supervision in selected provinces, which allows us to conduct a heterogeneous analysis. As expected, we find that the overreporting effect is muted for counties located in provinces supervised by the NBS. Taken together, the results suggest that the problem of misaligned incentives, which leads to local behavioral distortions and thus deviations

[11]In this case, mission incentives or intrinsic incentives more generally may be at work. For example, mission-motivated bureaucrats might behave prosocially. See Bénabou and Tirole (2006) for a theoretical analysis and Besley and Ghatak (2018) for a review.

[12]Monitoring often plays an important role in deterrence (Finan et al., 2017). For example, the existing literature has shown that government audits reduce corruption in developing countries (e.g., Olken, 2007; Avis, Ferraz and Finan, 2018).

from the central policy goal in the implementation of poverty alleviation programs, can be addressed through local self-discipline or centralized supervision.

Our paper contributes to several strands of literature. First, this paper speaks to the political economy of development literature studying the role of governance in ending poverty or promoting development more generally, which focuses on how to incentivize bureaucrats and further improve state effectiveness. This literature is relatively large; see, for example, Bardhan (2016), Finan et al. (2017), Page and Pande (2018), Olken and Pande (2019), and Besley et al. (2022) for excellent reviews.[13] We add to this literature by documenting that misaligned incentives arising in the implementation of anti-poverty programs can lead to strategic misreporting by local governments and deviations from the central policy goal, thus undermining state effectiveness. In particular, our work is related to the literature on high-powered incentives in the public sector that can lead to information manipulation or cheating behavior (Jacob and Levitt, 2003; Banerjee, Duflo and Glennerster, 2008; Fisman and Wang, 2017; Lyu et al., 2018; Acemoglu, Fergusson, Robinson, Romero and Vargas, 2020, among others), from which we differ by providing evidence that local agents would strategically manipulate the same performance measure in different directions in response to different incentives. More generally, our paper highlights the importance of information and incentives for governments and other public organizations (e.g., Banerjee, 1997; Besley and Ghatak, 2005; Acemoglu, Kremer and Mian, 2008), especially for those who are organized in M form (e.g., Maskin, Qian and Xu, 2000; Qian, Roland and Xu, 2006).

Second, this paper links to a growing literature on the manipulation of official statistics, especially in developing countries. Scholars have questioned the accuracy and reliability of government-reported GDP figures and suggested using night-time lights data (e.g., Chen and Nordhaus, 2011; Henderson et al., 2012; Donaldson and Storeygard, 2016; Hu and Yao, 2022). Some have linked the manipulation of economic data to political regimes (e.g., Hollyer, Rosendorff and Vreeland, 2011; Martinez, 2022). Specific to our empirical context, there has been considerable skepticism about China's official statistics (mainly upward bias) (e.g., Park and Wang, 2001; Young, 2003; Nakamura, Steinsson and Liu, 2016; Wallace, 2016; Chen et al., 2019),[14] while others have provided evidence that China's GDP statistics may not be systematically biased (e.g., Holz, 2004; Mehrotra and Pääkkönen, 2011; Holz, 2014), or may even be underestimated (Clark, Pinkovskiy and Sala-i Martin, 2020).[15] Our paper contributes to this literature by providing empirical evidence on the coexistence of upward and downward reporting bias in Chinese GDP statistics.

Third, our paper adds to the body of research on the role of monitoring in improving economic and governance outcomes. Previous studies have documented the importance of government audits (Olken, 2007; Ferraz and Finan, 2008; Bobonis, Cámara Fuertes and Schwabe, 2016; Avis et al., 2018), third-party audits (Duflo, Greenstone, Pande and Ryan, 2013), regulatory discretion (Duflo, Greenstone, Pande and Ryan, 2018), modern technology (Callen and Long, 2015; Dhaliwal and Hanna, 2017; Greenstone et al., 2022; Axbard and Deng, 2024), and citizen participation (Buntaine, Greenstone, He, Liu, Wang and Zhang, 2024). This literature has also pointed out that who collects

[13] See also World Bank (2017) for a comprehensive report.

[14] Some previous studies have associated overreporting bias in GDP with tournament competition in China (Xiong, 2018; Chen et al., 2021; Xu, Xu and Si, 2022); Xiao and Womack (2019) discussed China's internal information distortion patterns.

[15] Another well-known type of data manipulation by local governments in China is the misreporting of pollution data. See, for example, Ghanem and Zhang (2014) and Greenstone, He, Jia and Liu (2022).

information matters (e.g., Finan et al., 2017). We contribute to this literature by highlighting the importance of central government monitoring in improving local government performance in a principal-agent setting with multitasking agents.

Last but not least, this paper relates to the literature that studies China's poverty reduction. There is no doubt that China's growth in recent decades has been remarkable (Song, Storesletten and Zilibotti, 2011), and the poverty alleviation achieved by the Chinese government is unprecedented in human history (Chen and Ravallion, 2021; World Bank, 2022).[16] That said, our study is not intended to deny these successes. Rather, our findings suggest that local leaders may overreact to the policy goal of the central government under high-powered incentives and intensified political competition.

The rest of this paper is organized as follows. Section 2 describes the institutional context. Section 3 presents the data. Sections 4 and 5 examine the effects of NPC designation and cancellation, respectively. Section 6 concludes the paper.

2 Institutional Context

2.1 NPC Designation and Fiscal Transfers

During 1986-2011, China carried out four large-scale poverty alleviation programs that designate national poor counties. See Figure 1 for the program rollout timeline. Administered by the Leading Group for Economic Development in Poor Areas (LGEDPA), China implemented a program in 1986 aiming at eliminating rural poverty, in which 331 national poor counties were selected. In 1994, the Chinese government implemented the 8-7 Plan, aiming to lift 80 million rural people who remain poor out of poverty within 7 years (i.e., by 2000).[17] 592 counties were selected as national poor ones for this program.[18] After the 8-7 Plan ended, the list of national poor counties was adjusted in 2001.[19] The main change was removing 33 poor counties in eastern provinces which are generally richer and allocating those quotas to central or western provinces. As a result, the total number of national poor counties remains unchanged. Meanwhile, as a concentrated and contiguous impoverished area (Chinese: *Jizhong Lianpian Tekun Diqu*), all the 74 counties in Xizang (i.e., Tibet) are treated as national poor counties. In 2011, a more extensive program was carried out, with 832 national poor counties designated.[20] In this wave, 14 concentrated and contiguous impoverished areas were designated, covering 689 poor counties. It is noteworthy that the outline guiding this program made it clear that the goal is to build a moderately prosperous society (Chinese: *Xiaokang Shehui*) in 2020 by eradicating rural poverty and achieving common prosperity.

[Figure 1 about here]

[16] For example, see Park, Wang and Wu (2002) and Meng (2013) for impact evaluations of China's earlier national poor county programs; and see Park and Wang (2010), Cai (2020), Cai, Park and Wang (2023), and Zhang, Xie and Zheng (2023) for impact evaluations of village-level programs. Almost all studies find positive impacts.

[17] See Wang, Li and Ren (2004) for more detail.

[18] The national poor counties were incrementally added up to 331 counties by 1993.

[19] See the *Outline of Rural Poverty Alleviation and Development of China (2001-2010)* issued by the State Council in 2001 for detailed guidance. Available at `https://bit.ly/3yPUsiN`, last access on October 18, 2022.

[20] See the *Outline of Rural Poverty Alleviation and Development of China (2011-2020)* issued by the CPC Central Committee and the State Council in 2011 for detailed guidance. Available at `https://bit.ly/3S8UdpT`, last access on October 18, 2022.

The poor county designation is based mainly on rural net income per capita at the county level.[21] Generally, a standard poverty line is necessary for selecting which counties would receive the "poverty hat" with numerous poverty alleviation funds. The standard may vary across counties. For example, the government allows for a higher poverty line for counties in old revolutionary base areas or minority areas,[22] indicating that such political concerns also influence the selection process. The poverty line has also increased over time. Its composition was adjusted in 2002. Specifically, the state council makes county GDP per capita account for 30% when selecting the poor counties (rural net income per capita still accounts for 60%; the remaining 10% goes to county fiscal revenue per capita). Since then, the Chinese government no longer makes the selection criteria publicly available, but this is allegedly the standard of poor county designation. In our empirical analysis, we focus on GDP figures because they largely measure the final, overall result of manipulating economic data at the local level. Our first set of empirical analyses focuses on the 2011 wave of national poor county selection.

The national poor counties would receive numerous fiscal transfers, including general and special transfers. The general fiscal transfer payment is used to fill the fiscal gaps in less developed regions, with priority given to the national poor counties. The special fiscal transfer payment is based on specific policies. In our case, it mainly includes poverty alleviation funds. The central government provides a certain amount of funds, and the provincial government (together with lower-level governments for rich regions) provides 30% of what the central government offers. In addition, the national poor counties also receive other policy preferential treatments such as tax exemption. Since those transfers constitute a large proportion of fiscal revenue for local governments, county officials have strong incentives to acquire them. Even rich counties have incentives to do so. For example, *Zhungeer Qi*,[23] a national poor county designated in 1993 (and removed in 2011), was the twelfth richest county in China in 2011. Park et al. (2002) also pointed out that it is hard to eliminate national poor counties as county leaders fought to keep access to the anti-poverty funds. More generally, local officials in China usually focus their policy choices on what could raise their revenue base (e.g., Han and Kung, 2015).

Given the strong incentives to obtain the poverty alleviation funds and other preferential treatments, it should not be surprising to observe local leaders exerting efforts to acquire or maintain the access to the funds. As noted by Zheng and Pu (2011), many counties competed for the national poor county designations by manipulating local economic data (mainly, by underreporting income or GDP per capita). Such data manipulation phenomena have been covered by Chinese media, ranging from national-level state-owned ones, such as *China Daily*[24] and *People's Daily*,[25] to local ones, such as *Guangzhou Daily*.[26] Also, see the Appendix Figure A3 for more media coverage. Importantly, Park and Wang (2001) noted that one should be cautious about county income per capita,

[21] See Park and Wang (2001) and Park et al. (2002) for detailed description and discussion on how the national poor counties were selected for the period 1986-2001.

[22] The revolutionary base areas refer to the regions where the CPC had strong support and could exercise control, establish soviets, and implement its policies during the Chinese Civil War in the early half of the 20th century. For example, Jiangxi Soviet and Yan'an Area.

[23] *Zhungeer Qi* is a county located in Inner Mongolia Autonomous Region in northwestern China.

[24] For example, see `https://bit.ly/3zWRedP`, last access on November 11, 2022.

[25] For example, see `https://bit.ly/3FSxdJo`, last access on November 11, 2022.

[26] For example, see `https://bit.ly/3TynoUh`, last access on November 11, 2022.

the main determinant of poor county designations, since *(i)* substantial discretion over how it is determined was given to local governments, who tend to manipulate the income data with substantial resources at stakes, and *(ii)* oftentimes no independent source of verification was available at that time.

2.2 NPC Cancellation and Political Tasks

Since the 18th National Congress of the Communist Party of China, the Chinese central government has begun to pay more attention to poverty alleviation. In 2013, President Xi Jinping proposed the concept of "precision poverty alleviation" (Chinese: *Jingzhun Fupin*), which emphasizes the importance of accurately identifying and helping poor households or areas through precise targeting and tailored interventions. Since then, the government has focused public investment and aggressive spending on 14 concentrated and contiguous impoverished areas located in remote, mountainous areas. See the Appendix Figure A1 for the regional distribution of these 14 poor areas. In late 2015, the Chinese central government issued the *Decision on Winning the Battle Against Poverty Alleviation* (hereafter, the 2015 Decision) and announced the gradual elimination of national poor counties by 2020.[27] In 2018, the Chinese central government issued the *Guidance on Winning the Three-Year Battle Against Poverty Alleviation* (hereafter, the 2018 Guidance),[28] with the ambitious goal of lifting all national poor counties out of poverty in three years. During 2016-2020, national poor counties were gradually removed. By the end of 2020, there were no more poor counties, and the government claimed that rural poverty had been eliminated in China. This is considered one of the greatest political achievements of the Chinese government in recent years.

The criteria for removing a county from poverty status are based primarily on the poverty incidence rate and the number of people living in poverty. If a county's poverty incidence (i.e., the share of poor population living below the poverty line) is less than 2% (3% for western provinces), it is removed as a national poor county. The poverty line is defined as a per capita income of approximately USD 330. In addition, the number of people living in poverty is also taken into account. More importantly, the central government also requires local governments to be responsible for preventing households from returning to poverty.

It is widely believed that this wave of eliminating national poor counties is a non-negotiable political task. As mentioned above, in the 2011 wave of poor county designation, the Chinese central government proposed the goal of eliminating rural poverty by 2020. This was further emphasized in the 2015 Decision. Moreover, in the 2018 Guidance, the central government clearly requires local governments to eliminate all national poor counties within three years. Given the limited time, local governments face more pressure from the center. Figure 2 shows that most counties were removed from poverty status in 2018 and 2019, while only a small number of counties were removed in 2016 and 2017. Because the political pressure facing local governments are different, our second set of empirical analyses examines the effects of these two policies separately.

Three important institutional features may induce county leaders to exert as much effort as possible to get rid of poor status. First, given China's centralized personnel control system, although local leaders have a great deal of discretion in economic matters, their promotion prospects are

[27] See `https://bit.ly/3uSFDNs` for the documentation (in Chinese), last access on December 10, 2023.
[28] See `https://bit.ly/48bKjf0` for the documentation (in Chinese), last access on December 10, 2023.

controlled by higher-level governments (Xu, 2011). Therefore, under this performance-based promotion system, failure to eliminate rural poverty would certainly affect the promotion, if not the demotion, of county leaders. Second, top-down economic target setting is common in China's hierarchical administrative system, and often the central government's preference is directly transmitted to the motives and behavior of local governments (Li, Liu, Weng and Zhou, 2019). When the eradication of rural poverty by 2020 is set as a mandatory policy goal by the central government, it gives a huge incentive to local officials to achieve the goals within the time period. Third, the central government implemented the cadre freeze (Chinese: *Ganbu Dongjie*) policy in early 2016,[29] which does not allow the party secretaries of the national poor counties to be promoted without completing the poverty alleviation goal. This may make county leaders more eager to remove the poverty hat, especially for those who have been appointed to the position for several years (i.e., in the later period of their tenure). In our research setting, local leaders of national poor counties may overstate GDP per capita or income per capita to meet the requirement of completing poverty alleviation if their true economic figures are below the required levels. Again, see the Appendix Figure A3 for related media coverage.

2.3 GDP Manipulation and Central Supervision

GDP manipulation is a common problem in developing countries (e.g., Martinez, 2022). China, of course, is no exception. In fact, GDP manipulation (mainly the overreporting of local GDP statistics) is not uncommon in China. This is mainly due to the promotion incentives that local political leaders face, given the merit-based promotion system in China (Li and Zhou, 2005; Xu, 2011), which is especially true at the county level (Landry et al., 2018). For example, Chen et al. (2021) find that county officials in the early period of their tenure would exert greater effort to overstate the GDP figure, the main observable measure of economic performance, which in turn may increase their promotion prospects. Many have provided evidence that China's growth has been overestimated (e.g., Chen et al., 2019, 2021).

Of course, the central government is well aware of GDP manipulation by local officials. To curb this problem, China's National Bureau of Statistics (NBS) conducts statistical supervision (Chinese: *Tongji Ducha*), the main purpose of which is to verify the accuracy of statistics reported by local governments. If a locality commits data manipulation and is caught, the consequences for the local leaders are severe. For example, both the party secretary and the county governor of Ying County in Shanxi Province were demoted after their data manipulation behavior was caught in 2018.[30] In late 2019 and 2020, large-scale regular supervision was conducted.[31] In our empirical analysis, we expect that this wave of NBS supervision will deter counties in the supervised provinces from overstating their GDP statistics.[32]

[29] See the *Notice on Maintaining the Stability of Party and Government Leaders in Poor Counties during the Poverty Alleviation Period* issued by the State Council and the Central Organization Department in April, 2016.

[30] For more detail, see `http://bit.ly/3TniQ2T`, last access on November 9, 2022.

[31] For more detail, see `http://bit.ly/3tfvm9P`, last access on November 9, 2022.

[32] Related to our empirical context, the Central Inspection Team (Chinese: *Zhongyang Xunshi Zu*, also known as the Central Leading Group for Inspection Work and led by the Commission for Discipline Inspection and the Organization Department of the Central Committee of the Communist Party of China) conducted investigations in February and October 2018 on whether local governments manipulate the statistical figures to meet the requirement of eliminating rural poverty, with the first (second) wave covering 13 (remaining) provinces. However, we cannot use this wave of the

In addition to central supervision, it is also important to consider local statistical self-discipline. It is less likely that all localities in China will misreport GDP statistics. Some local statistical authorities may adhere to national statistical standards and thus produce statistical data of high quality, reliability, and credibility. To distinguish regions with better statistical practices from those without, we use China's national economic census, which the NBS uses to calculate census-based GDP at the provincial level. We identify provinces with a self-reported GDP lower than the census-based GDP as regions with disciplined statistical practices. We then conduct heterogeneous analyses.

2.4 A Summary

To summarize, we are particularly interested in two sets of incentives that local governments face. The first set is the fiscal incentives that may lead county leaders to underreport GDP in order to acquire the poverty hat and receive fiscal transfers from the central government (in 2011 in our research setting). We expect the underreporting effect to be more pronounced for counties with worse fiscal conditions, both in terms of fiscal dependence and fiscal pressure.

The second set of incentives we study is the political incentives that may induce county leaders to overreport GDP in order to remove the poverty hat (since 2016 in our empirical setting), after the eradication of rural poverty by 2020 was set as a must-complete policy task by the central government in 2015. We conjecture that *(i)* the overreporting effect will persist because national poor counties are generally disadvantaged and thus unable to maintain sustainable development and growth without fiscal support from higher-level governments, *(ii)* it exists only in provinces without disciplined statistical practices, and *(iii)* it can be muted by statistical supervision from the central government, solving or mitigating the principal-agent problem.

Relatedly, we also consider the promotion incentives arising from different local leaders in different term years or with different individual characteristics. We mainly examine whether our baseline results are driven by career incentives of local officials (due to these two potential sources) rather than the fiscal and political incentives we propose above.

3 Data and Measures

In this paper, we construct county-level panel data for the period 2003-2020. We exclude municipal districts (Chinese: *Shi Xia Qu*) of municipalities, sub-provincial cities, and prefecture-level cities, which have the same administrative level as counties in China's administrative hierarchy, but are generally much richer. We also exclude counties that have undergone merger or division during the sample period. Our unit of analysis is the county-year, focusing on the administrative boundaries of Chinese counties delineated in 2020. See the Appendix Table A1 for summary statistics and data sources.

central investigation to conduct heterogeneity analysis because it covers all regions in China and was conducted in the year that removes most of the national poor counties.

3.1 Night Lights Data

We rely on the night lights data to capture the artificial GDP manipulation behavior of Chinese county officials (Chen and Nordhaus, 2011; Henderson et al., 2012). We obtain our data on the digital number (DN), which measures the luminosity of night lights, from the Global Night-time Light Database (GNLD) provided by the Chinese Research Data Services Platform (CNRDS, see https://www.cnrds.com/). See Appendix B.1 for a detailed description of data processing.

The GNLD provides the digital number for the periods 1992-2013 and 2013-present based on the Defense Meteorological Satellite Program (DMSP) and the Visible Infrared Imaging Radiometer Suite (VIIRS), respectively. We focus on two time periods (2003-2010 and 2013-2020) for our empirical analyses. Since we happen to use the two different data periods to separately examine the effects of acquiring and removing the poverty hat, the data source issue is less of an empirical concern for us. Since our data does not include counties with a DN value equal to or close to 63 after excluding municipal districts, the top-coding problem is not a concern for us. In addition, some scholars have also pointed out that the Chinese night lights data do not face severe top-coding issue (e.g., Storeygard, 2016; Baum-Snow, Brandt, Henderson, Turner and Zhang, 2017).

The digital number based on the DMSP data is available annually and at the 30 arc-second grid cell level. For each county, we first take the sum of the digital number for all cells and then divide the sum by the number of cells. Also available at the 30 arc-second grid cell level, the digital number based on the VIIRS data is a monthly series. We first calculate the mean digital number for each county, as we did for the DMPS data, and then simply average over the months to obtain the annual digital number.

3.2 GDP Manipulation Measures

Next, we use the average digital number to construct our GDP manipulation measure. For county i in year t, we use the natural logarithm of the ratio of GDP per capita ($PerCapitaGDP_{it}$) to the average digital number ($AverageDN_{it}$) to measure the GDP manipulation behavior of county governments. Since the digital numbers obtained from the DMSP for the period 2000-2013 can be zero, we add 0.01 to $AverageDN_{it}$ when using the DMSP data following Jia, Liang and Ma (2021). GDP per capita is adjusted using the provincial GDP deflator and is thus measured based on 2000 RMB *yuan*. Specifically, we measure a county's GDP manipulation tendency in a given year as follows:

$$GDPManipulation_{it} = ln(\frac{PerCapitaGDP_{it}}{AverageDN_{it}}). \tag{1}$$

This assumes that, other things being equal, there is a linear relationship between real GDP and night lights intensity. In the Chinese context, one potential concern is that local officials may manipulate GDP figures because of the promotion incentives rather than the fiscal or political incentives we study in this paper. To address this issue, we show in Appendix A that our empirical results are robust to accounting for the promotion incentives facing the county leaders.

Given the complexity of measuring GDP manipulation, our measure is by no means prefect. Nevertheless, to show the extent to which our measure is plausible, we also report the results using the digital number obtained directly from other sources in Appendix B.2 and the statistical framework proposed by Henderson et al. (2012) in Appendix B.3.

3.3 National Poor Counties

Our study period is 2003-2020. We start in 2003, right after China completed a wave of national poor county selection in 2002, which we do not study due to the availability of county-level data before 2002. We end in 2020 because that is when China claimed to have eradicated poverty. To study how the Chinese counties manage to acquire the poverty hat, we focus on the 2011 wave of national poor county selection. We can see from Figure 2 that there is an increase in the number of designated national poor counties in 2011. A total of 832 counties have been designated as national poor counties. Of these, 204 were newly selected and the rest (628) were re-selected. It is noteworthy that 38 poor counties selected in the 2002 wave were removed in 2011. See Figure 3 for the regional distribution of NPCs designated in 2011.

[Figure 2 about here]

[Figure 3 about here]

This gives us the source of variation to identify the relationship between being selected as a national poor county in 2011 and GDP manipulation behavior before 2011. In determining the 2011 national poor counties, the government uses GDP per capita in 2010 as a criterion. Thus, we are primarily interested in the GDP manipulation tendency in 2010. In our regression analysis, we define the treatment group as those counties designated as poor counties in 2011 and use counties that were never selected as poor counties as our control group. We then compare the changes in the designated counties before and after 2010 with the before-and-after changes in the counties that were never designated. Note also that we do not include the 38 poor counties that were removed in 2011 in our regressions.

We can also see from Figure 2 that national poor counties have been gradually eliminated since 2015, and there are no more national poor counties after 2020. We obtain the information on the year of being officially removed from the NPC list from the website of the National Rural Revitalization Administration of China (see `https://www.nrra.gov.cn/`). In 2015, the central government set poverty elimination by 2020 as a must-complete task for local governments. We use the 2015-2020 period to examine how county governments report their GDP figures under a non-negotiable policy target. In this set of empirical analyses, we compare the changes in the counties that were removed from the list of national poor counties before and after the year of being removed with the changes in counties that were never selected.

3.4 County Characteristics

County characteristics (e.g., level of development and local economic structure) may also influence the behavior of county leaders in manipulating economic data. Therefore, we include county-level covariates in our regressions. First, we directly include county-level nighttime light intensity (i.e., the digital number) to account for potential differences due to different levels of development. Second, we include county-level household savings per capita as an additional proxy for development level.

Third, we control for the ratio of primary sector value added to GDP and the ratio of secondary sector value added to GDP to capture differences in regional industrial structure. For each regression, we fix the value of all variables in the base year and interact them with year dummies to avoid concerns about reverse causality. This also allows us to control for potentially different time trends across counties due to different initial economic conditions. Except for the data on night lights, the data are taken from the Chinese County Statistical Yearbooks.

3.5 County Party Secretaries

In our research setting, it is also important to consider local leaders who have a great deal of discretion over policy implementation and often respond to career incentives, which may confound the fiscal and political incentives we study in this paper. To this end, we hand-collected information on the secretaries of the Communist Party of China (CPC) at the county level. We focus on party secretaries (rather than county governors) for two reasons. First, they are first-hand leaders (Chinese: *Yi Ba Shou*) in the Chinese political system. The second is a practical concern about data availability. We did our best to collect information on party leaders. We first collect the name list of county party secretaries from provincial statistical yearbooks and then search for their personal information from the Internet. For example, the *Baidu* Encyclopedia, county government websites, media coverage of county leaders, or other sources containing the biographies of top leaders who have served as county party secretaries. However, we still have missing information on some party secretaries, so the information in this dataset is not complete in the sense that we either do not know who was the party secretary for a county in a given year, or we do not have enough background information (e.g., gender, age, and education, etc.) on the party secretary. Therefore, we mainly use party secretary fixed effects and the years in office of the party secretaries in the regression analysis. When collecting the data, we have more missing information on county governors and thus focus only on county party secretaries.

4 NPC Designation in 2011

4.1 Research Design

To examine the relationship between being selected as a national poor county in 2011 and GDP manipulation behavior before 2011, we estimate the following equation:

$$GDPManipulation_{it} = \beta NPC_{i,2011} \cdot \mathbf{1}[t = 2010] + (X_{it_0} \cdot \lambda_t)'\theta + \alpha_i + \gamma_p \cdot \lambda_t + e_{it}, \quad (2)$$

where i indexes counties, p provinces, and t years (t_0 the base year). $GDPManipulation_{it}$ is the outcome of interest measured as in Equation (1), which captures county governments' manipulation behavior of GDP figures. $NPC_{i,2011}$ is a treatment dummy equal to one if county i is one of the national poor counties selected in 2011, including both the newly selected and the re-selected, and zero otherwise (i.e., our control counties that were never designated). $\mathbf{1}[\cdot]$ is a indicator function, and $\mathbf{1}[t = 2010]$ indicates the year dummy for 2010, which is our post period in this regression. X_{it_0} is a set of county-level variables measured in the base year, which we control for by interacting them with year dummies ($X_{it_0} \cdot \lambda_t$). α_i, γ_p, and λ_t are county, provincial, and year fixed effects, respectively. For time effects, we directly control for province-by-year fixed effects ($\gamma_p \cdot \lambda_t$). e_{it} is the random error term, clustered at the county level.

β is the coefficient of our interest. It is worth emphasizing again that we use a county-year panel with counties including the national poor counties selected in 2011 (treatment counties) and those never designated (control counties), and the time period spanning 2003-2010. In practice, the central government uses GDP per capita in 2010 as the selection criterion of the 2011 national poor counties, so we use 2010 as the post period. In this regard, β identifies the effect of being designated as a nation poor county in 2011 on GDP reporting behavior in 2010.

By including county fixed effects, we compare changes in outcome for the same counties before and after 2010. This means that the identification is mainly based on a before-after comparison within counties. Although the county dummies absorb any time-invariant differences that are correlated with being selected as a national poor county, this research design is concerned with differential time trends between the treatment and control groups that may further affect the outcome differently. We address this concern in two ways. First, we allow the time effect to vary by province. The province-year fixed effects account for common time-varying shocks to all counties in the same province, allowing us to compare counties in the same province-year cells. Second, we also control for interactions between base-year characteristics and year dummies to control for potential time trends due to different baseline characteristics. The characteristics include nighttime light intensity (the digital number), log household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP. These variables capture initial differences in regional levels of economic development and regional economic structure.

How to interpret β? A direct interpretation of β is that it captures the percentage change in the ratio (i.e., $\frac{PerCapitaGDP_{it}}{AverageDN_{it}}$).[33] By measuring GDP manipulation using $ln(\frac{PerCapitaGDP_{it}}{AverageDN_{it}})$ and conditioning on the control variables and fixed effects, we assume conditional linearity between GDP per capita and average DN.[34] If the linear relationship becomes less steep in the year of NPC designation, it suggests the existence of GDP underreporting, which would lead to β being negatively signed. It is noteworthy that systematic reporting bias (for all counties in the same direction) due to factors other than competition for NPC status would be differenced out by county fixed effects. In addition, we cannot completely rule out the possibility that the control counties do not compete for NPC status, a concern that is again largely alleviated by using only the never-selected counties as our control group. However, even if this were the case, our results can be viewed as lower bound estimates. In this sense, β captures the relative degree of GDP misreporting between treated and control counties, which is in line with the difference-in-elasticity interpretation of Martinez (2022) who studies the relationship between GDP and night-time light (NTL) across different political regimes by estimating a log-linear equation in levels.

To test the parallel pre-trends assumption, which is crucial for our difference-in-differences

[33] Note that $e^{\beta} - 1$ gives the exact percentage change.

[34] The linear relationship can be written as: $PerCapitaGDP = \theta \cdot AverageDN + \text{controls} + \text{fixed effects}$, where θ captures the conditional linearity.

strategy, we further estimate the following event study design:

$$GDPManipulation_{it} = \sum_{\tau=2003,\tau\neq 2009}^{2010} \beta_\tau \cdot NPC_{i,2011} \cdot \mathbf{1}[t=\tau] + (X_{it_0} \cdot \lambda_t)'\theta + \alpha_i + \gamma_p \cdot \lambda_t + e_{it}, \quad (3)$$

where τ indexes the years 2003 to 2010, and $\mathbf{1}[t=\tau]$ indicate the year dummies. The rest of the econometric settings are the same as in Equation (2). The year 2009 is omitted and used as the reference group in the regression. This equation allows us to compare the pretreatment difference in manipulation behavior between the treated counties (those selected in 2011) and the control counties (those never designated).

4.2 Empirical Results and Robustness Checks

Baseline Results. Table 1 presents the estimated results of the effect of being selected as a national poor county in 2011 on the degree of underreporting of GDP figures in 2010, obtained by estimating Equation (2). Panel A of Table 1 reports our baseline results using all national poor counties selected in 2011 as the treatment. Note that our control group consists of those counties that were never selected as poor counties. As noted above, while we focus on the 2003-2010 period (see columns 1-3), we also report the results using the 2008-2010 sample (see columns 4-6) to address concerns that the long time window may introduce bias and noise, especially given that we only have one post period. For both samples, we first present the results obtained using county and year fixed effects (column 1 and 4). To account for potentially differential time trends due to different initial economic conditions across counties, we then interact base-year characteristics (measured either in 2003 or 2008) with year dummies (columns 2 and 5). Finally, we further replace the year fixed effects with province-year fixed effects, which allow the time effects to be province-dependent and thus flexibly control for the time trends (columns 3 and 6). Overall, one can see that our results are robust to all specifications and all estimated coefficients are statistically significantly different from zero.

[Table 1 about here]

Focusing on the estimates reported in columns 3 and 6 with full sets of controls, we see that, compared to those never designated as national poor counties, being selected as a national poor county in 2011 is associated with an increase in the ratio of GDP per capita to average DN (before and in the year of selection) of about 5% and 6%, respectively, suggesting a greater degree of GDP underreporting by NPCs. The estimates are about only half of those reported in columns 1 and 4, respectively. This is mainly due to the fact that we flexibly control for potential time trends by including province-year dummies as well as interactions between base-year characteristics and year dummies in columns 3 and 6, while only including county and year fixed effects in columns 1 and 4. The province-year fixed effects allow us to compare counties within the same province and the same year, thereby absorbing shocks that are common to all counties in a given province. By interacting base-year characteristics with year dummies, we further allow for initial economic conditions to have different effects in different years. With these two sets of controls, this specification makes our treated and control counties more comparable, but it also reduces the variation for identification.

That said, it should not be surprising that the magnitude of the estimates reported in columns 3 and 6 are smaller than those in other columns.

Since our treatment group includes two types of national poor counties, the newly selected and the re-selected, we also report the results obtained separately using the two types of national poor counties as the treatment group (see Panels B and C of Table 1). One can see that both the newly selected counties and the re-selected counties tend to underreport their GDP figures. The estimated results also indicate that there is not much heterogeneity between the two types of selected counties. The evidence suggests that both types of counties that were eligible to be selected as national poor counties in 2011 had the same incentives to acquire the poverty hat in order to receive fiscal support from higher-level governments, including the central government.

Dynamic Estimates. A key underlying assumption of our research design is that there is no difference in pre-treatment trends between the national poor counties selected in 2011 and those never designated. To test this, we estimate Equation (3), an event study design using interactions between the dummy for national poor counties designated in 2011 and leads or lags relative to 2010, with 2010 as the post-treatment period. We plot the dynamic estimates with their corresponding 95% confidence intervals in Figure 4. The results show that prior to 2010, the difference in GDP manipulation tendency between national poor counties designated in 2011 and those never designated is not statistically significantly different from zero. That is, the two groups of counties follow similar parallel trends before the treatment. In this regard, our treatment (national poor counties designated in 2011) and control (counties never designated) groups are comparable.

[Figure 4 about here]

Permutation Test. We also conduct a permutation test to check whether our results are mainly driven by the selection of national poor counties in 2011. We randomly assign a subset of the counties in our sample as national poor counties designated in 2011. These resampled counties are coded as the treatment group (pseudo-treatment group) and the remaining counties serve as a new control group. We then re-estimate Equation (2), based on the specifications used in columns 3 and 6 of Table 1, to obtain the pseudo-treatment effect. We repeat this procedure 5,000 times. Hypothetically, if our results are driven by eligible counties competing for national poor county status, one should expect a null result that there is no association between the pseudo-assignment of national poor counties and the degree to which GDP is manipulated across counties.

[Figure 5 about here]

Figure 5 plots the distribution of placebo coefficients from the 5,000 estimated pseudo-treatment effects for the 2003-2010 sample (Panel A) and the 2008-2010 sample (Panel B). The dashed lines in the panels are the value of the coefficients reported in columns 3 and 6 of Table 1, respectively. As expected, we see that the pseudo-assignment of national poor counties produces a negligible proportion of (almost no) coefficients that are larger than the coefficient in column 3 (6) of Table 1 in absolute terms in Panel A (B). This lends strong support to the argument that our baseline results are driven by the GDP manipulation behavior of counties that were eligible to acquire national poor county status before selection.

The 38 Counties Removed in 2011. There are 38 national poor counties selected in 2002 that were removed in 2011. As mentioned above, some of these are actually rich counties. We use these 38 removed counties as the treatment group and the never designated counties as the control group, and re-estimate Equation (3) using the 2008-2010 sample. The dynamic estimates are reported in Panel A of Appendix Table A4. We do not find statistically significant results. This is also the case when we extend the sample period to 2012 (see Panel B of Appendix Table A4). This evidence suggests that more developed, and thus removed, "poor" counties may have little or even no incentives to misreport their GDP statistics.

4.3 IV and Fuzzy RD Estimates and Distance-Based Analysis

Our dynamic estimates and permutation test results presented above strongly support the causal interpretation of the results obtained using a difference-in-differences strategy. However, one may be concerned that our treatment group is somewhat ad hoc, which may bias our estimates. We now exploit a unique institutional setting to construct an instrument for being designated as a national poor county in 2011 in order to provide further causal evidence on the relationship between being selected as a national poor county in 2011 and the GDP reporting behavior in 2010. We execute this exercise using both an instrumental variable (IV) approach and a fuzzy regression discontinuity (RD) design.

The selection of national poor counties in 2011 was guided by the *Outline of Rural Poverty Alleviation and Development of China (2011-2020)*. In this wave of selection, the decision-making power over the selection is delegated to the provincial governments, but the central government's support strength, including preferential policy treatments, remains unchanged. The selection also follows the principles of maintaining the quota and replacing improved counties with undeveloped ones within provinces. Therefore, the quota of each province (i.e., the total number of counties that can be designated as national poor counties in a given province) is set and thus fixed before the 2011 selection.

[Figure 6 about here]

Under this arrangement, counties in each province can predict a rank with a threshold below which counties would be selected as national poor counties before the selection. Figure 6 shows the quota (number and ratio) of national poor counties assigned to each province used in the 2011 selection wave. Not surprisingly, provinces in the northwest (e.g., Gansu and Shaanxi) and southwest (e.g., Guizhou and Yunnan) have a higher quota than coastal provinces. In terms of ratio, the ratio has an average of about 41%, with a minimum of 23% (Heilongjiang) and a maximum of 100% (Xizang); the second highest is 68% (Guizhou). To construct our instrument, we rank the counties in a given province based on their GDP per capita in 2009. We choose 2009 because counties may strategically manipulate their 2010 GDP figures based on their relative position in the 2009 GDP per capita rankings in order to obtain the poverty hat in 2011. Using the 2009 GDP per capita rankings and the quota ratios, we then calculate a GDP per capita threshold for each province below which counties in that province are eligible for designation.

Naturally, the poorest counties are the most likely to be selected, so they may have little or no incentive to manipulate their GDP figures. On the other hand, those counties just below the threshold are highly likely to underreport their GDP figures, as doing so increases their chances of being selected, particularly if others are doing the same. In addition, those just above the threshold may also have strong incentives to manipulate their GDP figures in order to be selected as national poor counties. For the remaining rich counties, the chance of being selected will be very small (obviously zero for the richest ones).

[Figure 7 about here]

This logic is illustrated (and confirmed) in Figure 7. In Panel A, we plot the within-province distance between each county's 2009 GDP per capita and the 2009 threshold (the x-axis) against the national poor county status of the counties (the y-axis), where each point represents the proportion of national poor counties selected in 2011 in a given bin. The distance is calculated using the GDP data of the counties in each province separately, denoted as $Distcane_{ip,2009}$ (see Appendix Figure A7 for the histogram). The distant counties on the left (the poorest) are almost certain to be selected, while the distant counties on the right (the rich) have a near-zero probability of being selected. One can also see that the probability of being designated a national poor county increases as a county's ranking moves from right to left, with a jump around the zero-distance point. Panel B plots the within-province distance measured in 2009 (the x-axis) against our GDP manipulation measure (the y-axis), which is residualized. We can see a clear discontinuity at the cutoff point (i.e., the zero-distance point).

Motivated by the descriptive patterns in Figure 7, we now turn to conducting instrumental variable analysis and fuzzy RD analysis. Our instrumental variable analysis uses the panel data we used above, while our fuzzy RD analysis uses cross-sectional data constructed from our panel data.

Instrumented Results. Our instrument is the within-province distance between each county's 2009 GDP per capita and the 2009 threshold. To summarize, on the one hand, since the quota is predetermined, our instrument could not be influenced by the GDP manipulation behavior of the county governments and is therefore likely to satisfy the exclusion restriction. On the other hand, as shown above, the probability of being selected in 2011 is a monotonic function of the distance measure, thus ensuring the relevance of our instrument. In the empirical implementation, we run a two-stage least squares (2SLS) regression, re-estimating Equation (2) by instrumenting $NPC_{i,2011} \cdot \mathbf{1}[t = 2010]$ with an interaction term between the distance and the year dummy for 2010 (i.e., $Distance_{ip,2009} \cdot \mathbf{1}[t = 2010]$)

[Table 2 about here]

The results are presented in Table 2, with first-stage estimates in columns 1-2 and second-stage estimates in columns 3-4. The results show a strong first stage, with a Kleiberg-Paap F-statistic for weak identification greater than 10. Thus, according to Staiger and Stock (1997), our instrument is unlikely to suffer from weak identification problem. Our first-stage yields a negative coefficient. This is because the smaller the distance (in numerical value), the higher the probability of being selected as a national poor county. The second-stage results show that national poor counties designated in 2011 were more likely to underreport their GDP figures in 2010, which is statistically significantly different from zero (column 3). The magnitude of the estimated coefficients is larger than that of the difference-in-difference estimates reported in Table 1. This could be due to the fact that our difference-in-difference estimation uses a treatment group that includes only those counties that were designated in 2011 and a control group that was never designated, so it does not capture the tendency of the non-selected to manipulate GDP data, who are also eligible to be selected and thus have incentives to misreport GDP figures. In other words, our baseline results are likely lower bound estimates. Although our results using the 2008-2010 sample are somewhat weaker in terms of statistical significance (columns 2 and 4), the estimated coefficients have consistent signs and sizable magnitudes.

Fuzzy RD Estimates. In addition to serving as an instrument, we can also use the distance to perform a fuzzy RD analysis. Again, as we have illustrated above in Panel A of Figure 7, the likelihood of being designated an NPC in 2011 is a monotonic function of the within-province distance; and the probability is decreasing in the distance with a drop at the cutoff (the zero-distance point). In the 2011 wave, the selection of NPCs is not only based on GDP but also depends on other factors such as rural net income and county fiscal revenue. This leads us to a fuzzy RD design in which we use the distance measure as our running variable. As we discussed in Section 2, since the 2011 NPC selection was conducted in 2010 based on economic indicators measured in 2009, we construct cross-sectional data with the outcome (GDP manipulation defined as before) measured in 2010 ($GDPManipulation_{i,2010}$), the treatment status measured in 2011 ($NPC_{i,2011}$), and the running variable and covariates measured in 2009 ($Distance_{ip,2009}$ and $X_{i,2009}$, respectively).[35] We conceptualize our fuzzy RD design in an instrumental variable framework. We use "initial eligibility" as the instrument for whether or not a county was designated an NPC in 2011, and eligibility is defined as follows: $Eligible_i = \mathbf{1}[Distance_{ip,2009} \leq 0]$, which takes a value of one if the distance is smaller than or equal to zero and zero otherwise. Specifically, the first-stage and second-stage equations can be written as follows:

$$\begin{aligned} NPC_{i,2011} = {} & \alpha_1 + \beta_1 Eligible_i + \mu_1 Distance_{ip,2009} + \delta_1 Eligible_i \cdot Distance_{ip,2009} \\ & + X'_{i,2009}\theta_1 + \gamma_p + u_i, \end{aligned} \tag{4}$$

and

$$\begin{aligned} GDPManipulation_{i,2010} = {} & \alpha_2 + \beta_2 NPC_{i,2011} + \mu_2 Distance_{ip,2009} + \delta_2 NPC_{i,2011} \cdot Distance_{ip,2009} \\ & + X'_{i,2009}\theta_2 + \gamma_p + v_i, \end{aligned} \tag{5}$$

where we use first-order (linear) polynomial, as guided by Panel B of Figure 7 and also suggested by Gelman and Imbens (2019), and model different slopes on the two sides of the cutoff. In the regressions, we control for provincial fixed effects (γ_p), which allow us to compare counties within the same province, in part because the NPC quota is specified separately and thus different for each province. The random error term is clustered by province.

[Table 3 about here]

The estimated results are presented in Table 3, with first-stage estimates in columns 1-4 and second-stage estimates in columns 5-8. Again, our first-stage results indicate that the eligibility

[35] We find that there is no evidence of manipulation of the running variable around the cutoff (see the Appendix Figure A8), and the covariates are balanced at the cutoff (see the Appendix Figure A9).

indicator (i.e., the instrument) is a strong predictor of whether or not a county would be designated as poor in 2011, with a Kleiberg-Paap F-statistic for weak identification much greater than 10 in all specifications. Columns 1 and 5 report the results without the running variable, while columns 2 and 6 report the results with the running variable. In columns 3 and 7, we restrict the sample to the counties with a distance in the interval from -3 to 3 (the unit is 10,000 yuan). The interval is further reduced to [-2, 2] in columns 4 and 8. All specifications show that the counties selected as NPCs in 2011 are more likely to underreport their GDP figures. Consistent with the instrumented results, the fuzzy RD estimates are also larger than the baseline results, which again may be due to the fact that our baseline difference-in-differences estimation is likely to yield lower bound estimates. Taken together, our instrumented results and fuzzy RD estimates provide strong support for the causal interpretation of the relationship between being selected as an NPC and GDP underreporting.

Distance-Based Analysis. Finally, we can also use the distance to perform a reduced-form analysis. As we have illustrated above in Figure 7, the counties around the threshold (i.e., the zero-distance point), both on the left and on the right, should have the strongest incentives to underreport their GDP figures, while the distant counties have the least or no incentives to misreport. In this regard, we can use the absolute value of the distance to measure the extent to which counties are motivated to manipulate their GDP data. To do so, we re-estimate Equation (2) by replacing $NPC_{i,2011}$ with the absolute value of the within-province distance between each county's 2009 GDP per capita and the 2009 threshold (i.e., $|Distcane_{ip,2009}|$). This predetermined distance is likely exogenous, since the provincial-level quota is set long before the 2011 selection.

[Table 4 about here]

We report the estimated results in Table 4, again using both the 2003-2010 sample (columns 1-4) and the 2008-2010 sample (columns 5-8). Column 1 shows that counties with a shorter distance (in absolute terms) underreport their GDP figures in greater degree, while column 5 shows a statistically insignificant estimate. One can see that the shorter distance leading to a greater degree of underreporting GDP figures also holds when we focus on all national poor counties selected in 2011 (see the estimated coefficients on the triple interaction term, | Distance | $\times$ 2010 $\times$ NPC, in columns 2 and 6), or when we perform subgroup analyses using the selected counties with 2009 GDP per capita above or below the threshold separately (see columns 3 and 7 or columns 4 and 8, respectively). We can also see that the selected counties with a distance of zero (or very close to zero) in 2009 are more severe in underreporting GDP figures compared to the non-selected counties (see the estimated coefficients on the interaction term, 2010 $\times$ NPC, in columns 2-4 and 6-8). Interestingly, we find that among counties that were not selected in 2011 but had GDP per capita below the threshold in 2009 (about 100 counties), those with greater distance are also more severe in underreporting their GDP figures (see the estimated coefficients on the interaction term, | Distance | $\times$ 2010, in columns 4 and 8). This suggests that the non-selected also have incentives to underreport, which partly explains why our 2SLS estimates are larger than the baseline estimates.

4.4 The Role of Fiscal Concerns

This subsection examines the fiscal circumstances under which local governments may exert more effort to obtain the poverty hat, conditional on the national poor counties designated in 2011. We focus on two important aspects: fiscal dependence and pressure. Within their jurisdictions, Chinese county governments are responsible for providing public goods and services, paying wages to public employees (including school teachers and health workers), promoting local socioeconomic development (e.g., financing investment and development projects and subsidizing the agricultural sector), and so on. For those poor counties that do not have sufficient tax revenue, it is almost impossible to meet these requirements on their own. Therefore, many poor counties rely heavily on fiscal transfers from the the central government and higher-level governments. In theory, counties depending more on fiscal transfers or facing more fiscal pressure should be more aggressive in competing for the national poor county status.

To test this logic, we examine heterogeneous effects across counties with different levels of fiscal dependence on higher-level governments or fiscal pressure faced by county governments. In a given year, we measure fiscal dependence by a county's fiscal transfers from higher-level governments as a share of its fiscal revenues. We then average the transfer-to-revenue ratios over 2003-2009 (or 2008-2009) and use this average ratio as our final measure of fiscal dependence. To make the interpretation of the coefficients more intuitive, we demean the fiscal dependence measure in our regressions by subtracting its mean from each data point. We present the estimated effects in columns 1 and 3 of Table 5, which shows that the main effect is not statistically significantly different from zero, suggesting that average counties do not differ in terms of fiscal dependence in manipulating GDP figures. However, among the counties with higher than average fiscal dependence, the national poor counties selected in 2011 made more efforts to underreport GDP figures.

[Table 5 about here]

Our measure of fiscal pressure is constructed in a similar way. We first calculate the fiscal deficit for each county by subtracting fiscal expenditures from fiscal revenues. We then average the deficit-to-revenue ratios over 2003-2009 (or 2008-2009) and use this average ratio as our final measure of fiscal pressure. In the regressions, we also demean this variable to facilitate the interpretation of the main effect. The estimated impacts are reported in columns 2 and 4 of Table 5. We can see that for counties with average levels of fiscal pressure, those designated as poor counties in 2011 are more likely to underreport their GDP figures in 2010 (the main effect), and this effect is even larger for counties with above-average fiscal pressure (the interactive effect).

Taken together, the evidence reported above suggests that fiscal incentives play an important role in motivating less developed counties, especially those facing severer fiscal pressures, to manipulate their GDP figures in order to obtain the poverty hat, which in turn leads to numerous fiscal transfers from central and higher-level governments, thus easing their fiscal constraints and further supporting their development policies.

4.5 The Role of CPC Secretaries

We now turn to a more detailed consideration of county political leaders. The literature examining local governments in China has emphasized the important role that local political leaders play in shaping economic outcomes, particularly in terms of the promotion incentives they face (e.g., Li and Zhou, 2005; Xu, 2011; Landry et al., 2018). For example, *(i)* local leaders in different tenure years may face different promotion prospects (e.g., Chen et al., 2021); and *(ii)* the individual characteristics of local political leaders per se, such as the starting age and political rank at the start of office, may also motivate their misreporting behavior (e.g., Wang, Zhang and Zhou, 2020). Thus, it is important for us to examine whether our baseline results are confounded by career incentives stemming from these two possible sources. By accounting for the role of party secretaries, we also address the concern that our results may be driven by control counties reporting higher (due to promotion incentives) rather than designated counties reporting lower GDP statistics.

To this end, we first augment the specifications in columns 3 and 6 of Table 1 by including a set of dummy variables indicating years in office for party secretaries. In doing so, we further compare counties governed by party secretaries with the same duration in office, thereby controlling for potential variations in career incentives. The results are reported in columns 1-2 of Table 6. We can see that our baseline results remain largely unchanged for both sample periods (2008-2010 or 2003-2010).

[Table 6 about here]

Second, another way to control for the leader effects is to directly include party secretary fixed effects. We do so in columns 3-4 of Table 6. The average tenure length of the party secretaries is about 4.8 years for the period 2003-2010 (see Appendix Table A1), which may lead to colinearity between the CPC secretary and county fixed effects since our sample period is 8 and 3 years, respectively. It is also possible for a leader to serve as party secretary in two (or more) counties, and thus the party secretary fixed effects absorb the county fixed effects. Therefore, we do not include county dummies in columns 3-4. In this case, we compare the misreporting behavior of the same leader before and after the county he or she governs was designated as a national poor county. We can see that our baseline results hold, although the magnitude of the estimates shrinks about one third. Taken together, this evidence suggests that our empirical results reported in Table 1 are primarily driven by fiscal incentives, while career incentives (arising from different leaders in different years in office or with different individual characteristics) appear to play a limited role.

5 NPC Cancellation since 2016

5.1 Research Design

To examine the relationship between being cancelled as a national poor county since 2016 and GDP manipulation behavior, we estimate the following equation:

$$GDPManipulation_{it} = \beta NPCRemoved_{it} + (X_{it_0} \cdot \lambda_t)'\theta + \alpha_i + \gamma_p \cdot \lambda_t + e_{it}, \tag{6}$$

where again i indexes counties, p provinces, and t years (t_0 the base year). Instead of the redefined treatment variable, the econometric settings are the same as in Equation (2). The treatment dummy, $NPCRemoved_{it}$, equals to one if the national poor county i designated in 2011 is removed in year t, and zero otherwise. Note that we still use those counties that were never designated as national poor counties as our control group. The outcome variable, $GDPManipulation_{it}$, captures the GDP manipulation behavior of county governments, measured as in Equation (1). We control for interaction terms between county characteristics measured in the base year (X_{it_0}) with the year dummies (λ_t), accounting for potentially different time trends due to different initial economic conditions. We include county fixed effects (α_i) that control for any time-invariant confounding factors. We also allow the time effect to be province-dependent ($\gamma_p \cdot \lambda_t$), comparing counties within the same province and year. The error term, e_{it}, is clustered by county.

As noted above, the county fixed effects allow us to base our identification mainly on before-after comparisons. The coefficient to be estimated, β, identifies the effect of being removed as a national poor county on GDP manipulation behavior before and after removal. Assuming conditional linearity between GDP per capita and average DN, there may be GDP overreporting if the linear relationship becomes steeper after being removed from poor status. Again, β can be interpreted as the percentage change in the ratio of GDP per capita to average DN and captures the relative degree of GDP misreporting between removed and never designated counties.

In this set of empirical analyses, we examine the effects of the 2015 Decision and the 2018 Guidance separately. We do so for the following two reasons. First, as discussed in Section 2, the 2018 Guidance appears to have a greater deterrent effect because it has a clear timeline and a limited time period of only three years. Second, when the central government began requiring local governments to remove national poor counties, the more advantaged counties were more likely to be removed earlier. Thus, the counties removed under these two policies could be very different. Nevertheless, our results are robust to pooling all counties together.

Because different national poor counties were removed in different years, we have variation in the treatment timing. In such staggered settings, it is difficult to interpret the estimated coefficients, especially if there is treatment effect heterogeneity (Goodman-Bacon, 2021). To ensure our the robustness of our results, we report the estimates obtained using new estimators (e.g., Borusyak et al., 2023; Callaway and Sant'Anna, 2021; de Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2021).

To test whether our treatment and control counties follow similar pretreatment trends and whether the overreporting behavior is persistent for treated counties, we estimate the following event study equation:

$$GDPManipulation_{it} = \sum_{\tau=-k+1,\tau\neq-k}^{h} \beta_\tau \cdot \mathbf{1}[YearRemoved_i - t = \tau] + (X_{it_0} \cdot \lambda_t)'\theta + \alpha_i + \gamma_p \cdot \lambda_t + e_{it}, \tag{7}$$

where τ indexes the relative periods since being removed from poor status, with k (h) indicating the most distant pre (post) period, $YearRemoved_i$ indicates the year of being removed from poor status for a given county. $\mathbf{1}[\cdot]$ is a indicator function, and $\mathbf{1}[YearRemoved_i - t = \tau]$ indicates the relative τ's period to the year of being removed from poor status for county i in year t. The remaining econometric setups are the same as in Equation (6). We use period -k (the most distant pre period) as the reference period and thus omit it from the regression. For each regression using this equation, we always use a sample without treated counties in the earlier periods. The β_τ's estimated by this specification allow us to check for *(i)* the parallel trends assumption when $\tau < 0$, *(ii)* the immediate effect of being removed from poor status when $\tau = 0$, and *(iii)* the persistence of the effect when $\tau > 0$. Given that our empirical setting has staggered variation in treatment timing, we also report the event study estimates obtained using the newly proposed estimators as above.

5.2 Empirical Results and Robustness Checks

Baseline Results. We present the results of the effects of being removed as a national poor county on the overreporting behavior of county governments in Table 7, obtained by estimating Equation (6). See Figure 2 for the number of counties removed in different years since 2016. In the regressions, we control for county fixed effects, province-by-year fixed effects, and the interactions between base-year characteristics and year dummies. Using the 2013-2017 sample, column 1 of Table 7 examines the effect of the 2015 Decision, which was announced in November 2015. The treatment group is the counties that were removed in 2016 or 2017, while the control group is the counties that were never designated. Somewhat surprisingly, we find no statistically significant results. As discussed above, this could be due to the fact that better performing poor counties, which have neither the incentive nor the need to misreport, were removed earlier.

[Table 7 about here]

Columns 2-5 of Table 7 examine the impact of the 2018 Guidance announced in August 2018, using the 2016-2020 sample. In column 2, we use all counties that were removed in 2018-2020 as our treatment group. We then estimate the effect of being removed in different years separately, with the treatment groups in columns 3, 4, and 5 being the counties that were removed in 2018, 2019, and 2020, respectively. All columns use the counties that were never designated as the control group. Focusing attention on the estimated average effect reported in column 2, we can see that being removed from poor status since 2018 is associated with about a 6% increase in the ratio of GDP per capita to average DN, suggesting a greater degree of GDP overstatement by removed NPCs. This effect appears to be driven by those removed in 2018 and 2019 (see columns 3-5), while the effect is not statistically significantly difference from zero for counties removed in 2020. This may be due to the central statistical supervision implemented since late 2019 (see below for more analysis). Overall, our results suggest that the national poor counties designated in 2011 did have incentives to overreport GDP figures in order to respond to the central government's policy goal of eradicating rural poverty by 2020, especially after the 2018 Guidance was announced, which makes completing the task even more urgent given the limited time frame of three years. (In what follows, we focus mainly on the effect of the 2018 Guidance since the effect of the 2015 Decision is not statistically different from zero.)

Due to the fact that different counties were removed in different years, our results are concerned with the staggered variation in treatment timing, which may make our estimates difficult to interpret. To address this concern, in the Appendix Table A2, we show that our empirical results are largely robust to the new estimators proposed by Borusyak et al. (2023), Callaway and Sant'Anna (2021), de Chaisemartin and d'Haultfoeuille (2020), and Sun and Abraham (2021). Of the estimates obtained using the four new estimators, only the estimator proposed by Callaway and Sant'Anna (2021) yields an estimate about three times larger, while the others produce similar estimates. This result suggests that variation in treatment timing is less of a problem for us.

Dynamic Estimates. We now estimate the dynamic effects of being removed as a national poor county on the overreporting of GDP statistics using the event study specification of Equation (7).

We focus on counties removed since 2018, as we do not find statistically significant average effects for counties removed in 2016 and 2017. In our empirical setting, the dynamic results are important in terms of both internal validity checks and political economy implications. On the one hand, the dynamic estimates can serve as an internal validity check of the parallel pre-trends assumption, a key assumption underlying causal identification in difference-in-difference models. On the other hand, if the effect persists over time, it suggests that *(i)* the counties that were lifted out of poor status were not developed enough to promote local economic development and growth without substantial fiscal transfers from the center, and *(ii)* local leaders in charge of these counties have persistent incentives to overstate GDP figures when there are substantial political risks at stake.

[Figure 8 about here]

The event study specification to be estimated is based on the specification we use in column 2 of Table 7, where we examine the effect of the 2018 Guidance. Note that the sample period is from 2016 to 2020. Also, we have counties removed since 2018, so we have no treated units in the first two years of the sample period. We plot the estimated coefficients and their corresponding 95% confidence intervals in Figure 8. To increase the statistical power, we group periods -3 and -4 together as our reference periods. We can see that there is no difference in pretreatment trends between the national poor counties and those never designated. However, immediately after being removed from poor status, national poor counties appear to overreport GDP figures with a larger degree. More importantly, this overreporting effect is persistent and the magnitude increases slightly over time. Again, this result may reflect the fact that these removed counties were still far from well developed and therefore unable to promote development and growth without fiscal support.

We provide three pieces of evidence to support this conjecture. First, we explore the relationship between geographic factors related to economic development and the year of being removed from poor status. Specifically, we regress distance to the provincial capital and terrain ruggedness on dummies indicating different years of being removed from the NPC list. The results are reported in Appendix Table A3, showing that counties removed earlier are closer to the provincial capital and less rugged, while those removed more recently are both the most distant and the most rugged, and thus likely to be less well developed. Second, and relatedly, recall that we find no evidence of GDP overreporting for counties removed in 2011, some of which are recognized as the richest Chinese counties and most of which are indeed rich (see Appendix Figure A4). Third, this conjecture is also in part supported by the literature. For example, using detailed county-level fiscal data, Qiu et al. (2023) show that due to the poor fiscal situation in China's impoverished counties, poverty alleviation funds reduced fiscal support for industrial and enterprise development and innovation in these counties. Liu and Ma (2019) also show that the NPC program failed to promote both short- and long-term economic growth, in part due to local capture. Moreover, Zhu et al. (2021) show that the end of the NPC program reduced the county fiscal expenditure-to-GDP ratio for NPCs. Under pressure from the central government to eradicate rural poverty by 2020, local political leaders responsible for these counties may have to persistently overstate GDP figures if truthful reporting cannot meet the requirement.

Given the staggered removal of national poor counties, we check the robustness of our dynamic results presented above by re-estimating Equation (7) using new estimators recently proposed by

Borusyak et al. (2023), Callaway and Sant'Anna (2021), de Chaisemartin and d'Haultfoeuille (2020), and Sun and Abraham (2021). The estimates are plotted in the Appendix Figure A5, which do not differ much from our dynamic estimates reported in Figure 8.

The Appendix Figure A6 reports the event study estimates using the sample that combines all the removed national poor counties (i.e., examining the effects of the 2015 Decision and the 2018 Guidance together). We can see that there are no significant differences in the pre-trends and that the effect does not dissipate over time. Although the impact becomes statistically insignificant in the last post period, the magnitude is still economically remarkable.

Permutation Test. As in Section 5, we now perform permutation tests for the baseline results reported in columns 1 and 2 of Table 7. For each test, we randomly re-sample a group of counties as the treated counties (pseudo-treatment group) in each year based on the original sample structure. We then re-estimate Equation (6) to obtain the pseudo-treatment effect. We repeat this process 5,000 times. In Figure 9, we plot the distribution of placebo coefficients from the 5,000 estimated pseudo-treatment effects for the 2013-2017 sample (Panel A) and the 2016-2020 sample (Panel B), with the dashed lines indicating the value of the coefficients reported in columns 1 and 2 of Table 7, respectively.

[Figure 9 about here]

One can see from Panel A (examining the effect of the 2015 Decision) that the dashed line is in the middle of the distribution, confirming that those counties that were removed earlier had almost no incentive to misreport. In Panel B (examining the effect of the 2018 Guidance), the pseudo-assignment does not produce coefficients that are larger than the coefficient in column 2 of Table 7 in absolute terms, suggesting that the baseline results are unlikely to be driven by factors other than being removed from poor status.

5.3 The Role of Local Self-Discipline

We now turn to examine the conditions that might mitigate or even eliminate the misreporting behavior of local governments in the face of political pressure from the Chinese central government to eradicate poverty by 2020. In our empirical context, two broadly grouped factors may discourage local governments from exerting efforts to overstate GDP figures. The first is the internal motives of local governments, which we examine in this subsection. For example, if local governments have a reputation for disciplined statistical practice, these local governments are less likely to overstate GDP data. The second is external pressure from the center, which we examine in the next subsection. For example, if the central government implements monitoring programs aimed at data manipulation by local governments, the overreporting effect may be muted.

To examine the role of local statistical self-discipline in influencing the overreporting of GDP statistics by county governments, we rely on China's Fourth National Economic Census (Chinese: *Qunaguo Jingji Pucha*) to determine which regions adhere to national statistical standards. The census was conducted by China's National Bureau of Statistics (NBS) in 2018 and early 2019, with the aim of correcting and adjusting economic data published before and in 2018, including GDP figures. The NBS uses the census data to calculate census-based GDP at the provincial level. We compare the 2018 GDP figures reported by the provincial governments with the 2018 census-based GDP calculated by the NBS. We then identify those provinces whose self-reported GDP is lower than the census-based GDP as provinces with disciplined, better statistical practices.[36] Finally, we label the counties in these provinces as counties with good quality GDP statistics.

[Table 8 about here]

We conduct heterogeneous analyses by creating a dummy variable ("good quality") to indicate counties with good quality of GDP statistics. We focus on investigating the effect of the 2018 Guidance using the 2016-2020 sample. The results are reported in Table 8. Column 1 interacts the good quality dummy variable with our treatment variable ($NPCRemived_{it}$). As expected, we can see that counties with disciplined statistical practices almost completely mitigate the overreporting effect. This is confirmed in columns 2 and 3, where we run regressions using subsamples (i.e., counties with poor (good) quality GDP statistics in column 2 (3)) and find no statistically significant effects for counties with good quality. In summary, the evidence suggests that self-discipline in statistical practice plays an important positive role in the accurate reporting of GDP at the local level.

5.4 The Role of Central Supervision

Previous studies have argued that monitoring plays an important role in deterrence (Finan et al., 2017) and documented that government audits reduce corruption in developing countries (e.g., Olken, 2007; Avis et al., 2018). In our empirical context, as noted in Section 2, the Chinese central government is well aware of data manipulation by local governments and has implemented monitoring programs. In this subsection, we examine whether statistical supervision by the central government can deter local governments from overreporting GDP figures. To this end, we use the statistical supervision (Chinese: *Tongji Ducha*) implemented by the NBS in late 2019 and 2020.[37] Statistical supervision is regularly carried out by the NBS to verify the accuracy of statistical data reported by local governments. As we discussed in Section 2, several local leaders were demoted after their data fabrication behavior was caught. See the Appendix Figure A2 for the example of the Ying County, where both the party secretary and the county governor were demoted. Hence, we conjecture that statistical monitoring by the NBS has a deterrent effect on the overreporting behavior of targeted localities. Again, we focus on the 2016-2010 sample, in which national poor counties have been gradually removed since 2018. Given that the wave of statistical supervision took place in late 2019 and 2020, we expect the deterrence effect to be present only in 2020, but not in 2018 or 2019.

[Table 9 about here]

[36] According to this round of economic census, provinces with good quality GDP data include Shanghai, Yunnan, Sichuan, Anhui, Guangdong, Xinjiang, Jiangxi, Henan, Zhejiang, Hainan, Tibet, Hubei, Fujian, Jiangsu, Guizhou, and Chongqing.

[37] Related to our empirical context, the Central Inspection Team (Chinese: *Zhongyang Xunshi Zu*; also known as the Central Leading Group for Inspection Work directed by the Commission for Discipline Inspection and the Organization Department of the Central Committee of the Communist Party of China) conducted two waves of investigations in February and October 2018, respectively, on whether local governments manipulate statistical figures to meet the rural poverty alleviation requirement, with the first (second) wave covering 13 (remaining) provinces. However, we cannot take advantage of these inspections as all provinces were covered and our analysis sample includes counties that have only been treated since 2018.

In the regression analysis, we first construct a dummy variable ("supervision") that equals to one if a county is located in a province supervised by the NBS in late 2019 or 2020,[38] and zero otherwise. We then perform heterogeneous analyses. The results are presented in Table 9. In columns 1-3, we define our treated counties as the national poor counties that were removed from poor status in 2018, 2019, and 2020 separately, and interact them with the corresponding post-periods, which is further interacted with the supervision dummy variable. While we find both statistically significant and economically sizable overreporting effects for all subgroups, the deterrent effect is only present for counties removed in 2020 (see the triple interactions), consistent with our conjecture. In columns 4 and 5, we conduct a subgroup analysis for counties removed in 2020 based on the good quality dummy variable created above, which confirms that the overreporting effect (and thus the deterrent effect) exists only for counties that lack statistical self-discipline. Taken together, these results suggest that monitoring by the central government can largely solve the principal-agent problem, where county governments may overstate GDP statistics that do not match the central government's target.

5.5 The Role of CPC Secretaries

This subsection directly examines the role of local political leaders in determining their overreporting of GDP when they face the political task of rural poverty alleviation set by the central government. Our goal is twofold. First, we examine the extent to which the baseline results presented in Table 7 are driven by incentives for promotion, which again may be due to different leaders in different years in office or with different individual characteristics. Second, we examine how the political incentives we study in this section interact with promotion incentives, which may in turn influence the overreporting behavior of local governments. Table 10 presents the regression results.

[Table 10 about here]

Columns 1 and 2 of Table 10 use the 2013-2017 sample and the 2016-2020 sample, respectively, and augment Equation (6) with the dummies for years in office of different party secretaries to control for potential confounding effects arising from career incentives due to different party leaders in different years in office. Columns 3 and 4 of Table 10 replace the county fixed effects with dummies for party secretaries, addressing the concern that the baseline results may be driven by individual leaders with different characteristics competing for a higher likelihood of promotion.[39] We can see that none of the estimates in all four columns differ significantly from our baseline estimates reported in columns 1 and 2 of Table 7, suggesting that such career incentives are unlikely to drive our baseline estimates.

As mentioned in Section 2, a cadre freeze policy was implemented in early 2016. Under this policy, the party secretaries of NPCs are not allowed to be "fully" promoted if the goal of poverty elimination is not completed. Although they can be appointed to higher positions, they need to

[38] The supervised provinces in this wave include Anhui, Henan, Jilin, Jiangsu, Liaoning, Ningxia, Shanxi, Sichuan, and Yunnan.

[39] Again, since our sample period is 5 years and the average tenure of party secretaries is 5.1 years (see Appendix Table A1) and one leader can serve as party secretary for two or more counties, we cannot simply add leader fixed effects to the regressions, which in this case will highly likely be collinear with the county fixed effects.

continue to hold the same leader position as party secretary. As a result, one might expect that those who are much more eager to be promoted might exert more effort to overstate GDP figures. Importantly, given that lifting all poor areas out of poverty by 2020 is an imperative goal set by the Chinese central government, those nearing the end of their terms may be more eager not only to take off the poverty hat, but also to perform better in order to boost their promotion prospects. This would be especially true given that the cadre freeze policy implemented in 2016 makes it impossible for county leaders to be promoted without completing the political task of poverty elimination.[40] Column 5 of Table 10 uses the 2016-2020 sample and examines the interactive effect between the political incentives of our main interest and such promotion incentives, where we interact our treatment variable with a dummy variable indicating whether the party secretary is in his or her last two years in office. The estimated results confirm that the overreporting effect is much more pronounced for leaders in their last two years in office.

6 Conclusion

This study examines the relationship between the designation and cancellation of national poor counties and GDP misreporting among Chinese counties. Combining satellite nightlight data with self-reported GDP statistics and using difference-in-differences methods, we find that *(i)* being selected as an NPC in 2011, which would then receive substantial fiscal transfers, is associated with a greater degree of underreporting of 2010 GDP figures, and this underreporting effect is more pronounced for counties with worse fiscal conditions; and that *(ii)* national poor counties would overreport GDP figures to a greater degree since 2016, after poverty eradication by 2020 was set as a non-negotiable task by the central government, and this overreporting effect appears to be persistent and can be mitigated by self-discipline and central monitoring. We interpret the underreporting effect as a response to fiscal incentives and the overreporting effect as a response to political incentives. We rule out the possibility that our results are driven by the fact that local leaders are in different tenure years or have different individual characteristics.

Although we provide empirical evidence of the existence of GDP manipulation in one of China's major poverty alleviation programs — the National Poor County (NPC) Program, our study is not meant to deny China's remarkable successes in poverty reduction, as evidenced by, for example, Chen and Ravallion (2021) and World Bank (2022). Rather, our findings suggest that local leaders, who are often under high-powered incentives and intensified political competition, may overreact to the central government policy mandates. We leave the long-term sustainability of such large-scale anti-poverty programs to future research.

[40] Our sample shows that the average tenure length for the party secretaries of NPCs (about 6.4 years) is more than two years longer than that of non-NPCs (about 4.1 years) for those who stayed in office in 2016 or became party secretaries after 2016.

References

Acemoglu, Daron, Leopoldo Fergusson, James Robinson, Dario Romero, and Juan F Vargas, "The perils of high-powered incentives: Evidence from Colombia's false positives," *American Economic Journal: Economic Policy*, 2020, *12* (3), 1–43.

__ **, Michael Kremer, and Atif Mian**, "Incentives in markets, firms, and governments," *The Journal of Law, Economics, & Organization*, 2008, *24* (2), 273–306.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken, Ririn Purnamasari, and Matthew Wai-Poi, "Does elite capture matter? Local elites and targeted welfare programs in Indonesia," *AEA Papers and Proceedings*, 2019, *109*, 334–339.

Andersen, Jørgen Juel, Niels Johannesen, and Bob Rijkers, "Elite capture of foreign aid: Evidence from offshore bank accounts," *Journal of Political Economy*, 2022, *130* (2), 388–425.

Avis, Eric, Claudio Ferraz, and Frederico Finan, "Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians," *Journal of Political Economy*, 2018, *126* (5), 1912–1964.

Axbard, Sebastian and Zichen Deng, "Informed Enforcement: Lessons from Pollution Monitoring in China," *American Economic Journal: Applied Economics*, 2024, *16* (1), 213–252.

Banerjee, Abhijit V, "A theory of misgovernance," *The Quarterly journal of economics*, 1997, *112* (4), 1289–1332.

__ **, Esther Duflo, and Rachel Glennerster**, "Putting a band-aid on a corpse: Incentives for nurses in the Indian public health care system," *Journal of the European Economic Association*, 2008, *6* (2-3), 487–500.

Bardhan, Pranab, "State and development: The need for a reappraisal of the current literature," *Journal of Economic Literature*, 2016, *54* (3), 862–892.

__ , "The Chinese governance system: Its strengths and weaknesses in a comparative development perspective," *China Economic Review*, 2020, *61*, 101430.

__ **and Dilip Mookherjee**, "Capture and governance at local and national levels," *American Economic Review*, 2000, *90* (2), 135–139.

Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang, "Roads, railroads, and decentralization of Chinese cities," *Review of Economics and Statistics*, 2017, *99* (3), 435–448.

Bénabou, Roland and Jean Tirole, "Incentives and prosocial behavior," *American Economic Review*, 2006, *96* (5), 1652–1678.

Besley, Timothy and Maitreesh Ghatak, "Competition and incentives with motivated agents," *American economic review*, 2005, *95* (3), 616–636.

__ **and** __ , "Prosocial motivation and incentives," *Annual Review of Economics*, 2018, *10*, 411–438.

__ **, Robin Burgess, Adnan Khan, and Guo Xu**, "Bureaucracy and development," *Annual Review of Economics*, 2022, *14*, 397–424.

Bobonis, Gustavo J, Luis R Cámara Fuertes, and Rainer Schwabe, "Monitoring corruptible politicians," *American Economic Review*, 2016, *106* (8), 2371–2405.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event Study Designs: Robust and Efficient Estimation," 2023. Working Paper.

Buntaine, Mark T, Michael Greenstone, Guojun He, Mengdi Liu, Shaoda Wang, and Bing Zhang, "Does the Squeaky Wheel Get More Grease? The Direct and Indirect Effects of Citizen Participation on Environmental Governance in China," *American Economic Review*, 2024, *114* (3), 815–850.

Cai, Shu, "Migration under liquidity constraints: Evidence from randomized credit access in China," *Journal of Development Economics*, 2020, *142*, 102247.

__ **, Albert Park, and Sangui Wang**, "Microfinance can raise incomes: Evidence from a randomized control trial in China," 2023. Working Paper.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Callen, Michael and James D Long, "Institutional corruption and election fraud: Evidence from a field experiment in Afghanistan," *American Economic Review*, 2015, *105* (1), 354–381.

Cao, Guangyu, Xi Weng, Mingwei Xu, and Li-An Zhou, "Hybrid Contracts, Multitasking, and Incentives: Theory and Evidence from China's Air Pollution Controls," 2023. Working Paper.

Che, Jiahua, Kim-Sau Chung, and Xue Qiao, "Career concerns, Beijing style," *International Economic Review*, 2021, *62* (4), 1513–1535.

Chen, Shaohua and Martin Ravallion, "Reconciling the conflicting narratives on poverty in China," *Journal of Development Economics*, 2021, *153*, 102711.

Chen, Shuo, Xue Qiao, and Zhitao Zhu, "Chasing or cheating? Theory and evidence on China's GDP manipulation," *Journal of Economic Behavior & Organization*, 2021, *189*, 657–671.

Chen, Wei, Xilu Chen, Chang-Tai Hsieh, and Zheng Song, "A Forensic Examination of China's National Accounts," *Brookings Papers on Economic Activity*, 2019, *1*, 77–141.

Chen, Xi and William D Nordhaus, "Using luminosity data as a proxy for economic statistics," *Proceedings of the National Academy of Sciences*, 2011, *108* (21), 8589–8594.

Chen, Yvonne Jie, Pei Li, and Yi Lu, "Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in China," *Journal of Development Economics*, 2018, *133*, 84–101.

Clark, Hunter, Maxim Pinkovskiy, and Xavier Sala i Martin, "China's GDP growth may be understated," *China Economic Review*, 2020, *62*, 101243.

de Chaisemartin, Clément and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole, "The economics of career concerns, part II: Application to missions and accountability of government agencies," *Review of Economic Studies*, 1999, *66* (1), 199–217.

—, —, and —, "Multitask agency problems: Focus and task clustering," *European Economic Review*, 2000, *44* (4-6), 869–877.

Dhaliwal, Iqbal and Rema Hanna, "The devil is in the details: The successes and limitations of bureaucratic reform in India," *Journal of Development Economics*, 2017, *124*, 1–21.

Donaldson, Dave and Adam Storeygard, "The view from above: Applications of satellite data in economics," *Journal of Economic Perspectives*, 2016, *30* (4), 171–198.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India," *Quarterly Journal of Economics*, 2013, *128* (4), 1499–1545.

—, —, —, and —, "The value of regulatory discretion: Estimates from environmental inspections in India," *Econometrica*, 2018, *86* (6), 2123–2160.

Fang, Hanming, Chang Liu, and Li-An Zhou, "Window dressing in the public sector: Evidence from China's compulsory education promotion program," *Journal of Public Economics*, 2023, *222*, 104878.

Ferraz, Claudio and Frederico Finan, "Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes," *Quarterly Journal of Economics*, 2008, *123* (2), 703–745.

Finan, Frederico, Benjamin A Olken, and Rohini Pande, "The personnel economics of the developing state," *Handbook of Economic Field Experiments*, 2017, *2*, 467–514.

Fisman, Raymond and Yongxiang Wang, "The Distortionary effects of incentives in government: Evidence from China's "death ceiling" program," *American Economic Journal: Applied Economics*, 2017, *9* (2), 202–218.

—, Hui Lin, Cong Sun, Yongxiang Wang, and Daxuan Zhao, "What motivates non-democratic leadership: Evidence from COVID-19 reopenings in China," *Journal of Public Economics*, 2021, *196*, 104389.

Gelman, Andrew and Guido Imbens, “Why high-order polynomials should not be used in regression discontinuity designs,” *Journal of Business & Economic Statistics*, 2019, *37* (3), 447–456.

Ghanem, Dalia and Junjie Zhang, “"Effortless Perfection:" Do Chinese cities manipulate air pollution data?,” *Journal of Environmental Economics and Management*, 2014, *68* (2), 203–225.

Goodman-Bacon, Andrew, “Difference-in-differences with variation in treatment timing,” *Journal of Econometrics*, 2021, *225* (2), 254–277.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu, “Can technology solve the principal-agent problem? Evidence from China’s war on air pollution,” *American Economic Review: Insights*, 2022, *4* (1), 54–70.

Han, Li and James Kai-Sing Kung, “Fiscal incentives and policy choices of local governments: Evidence from China,” *Journal of Development Economics*, 2015, *116*, 89–104.

He, Guojun, Shaoda Wang, and Bing Zhang, “Watering down environmental regulation in China,” *Quarterly Journal of Economics*, 2020, *135* (4), 2135–2185.

Henderson, J Vernon, Adam Storeygard, and David N Weil, “Measuring economic growth from outer space,” *American Economic Review*, 2012, *102* (2), 994–1028.

Hollyer, James R, B Peter Rosendorff, and James Raymond Vreeland, “Democracy and transparency,” *Journal of Politics*, 2011, *73* (4), 1191–1205.

Holmstrom, Bengt and Paul Milgrom, “Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design,” *Journal of Law, Economics, and Organization*, 1991, *7* (special_issue), 24–52.

Holz, Carsten A, “Deconstructing China’s GDP statistics,” *China Economic Review*, 2004, *15* (2), 164–202.

— , “The quality of China’s GDP statistics,” *China Economic Review*, 2014, *30*, 309–338.

Hu, Yingyao and Jiaxiong Yao, “Illuminating economic growth,” *Journal of Econometrics*, 2022, *228* (2), 359–378.

Huang, Zhangkai, Lixing Li, Guangrong Ma, and Lixin Colin Xu, “Hayek, local information, and commanding heights: Decentralizing state-owned enterprises in China,” *American Economic Review*, 2017, *107* (8), 2455–2478.

Jacob, Brian A and Steven D Levitt, “Rotten apples: An investigation of the prevalence and predictors of teacher cheating,” *Quarterly Journal of Economics*, 2003, *118* (3), 843–877.

Jia, Junxue, Xuan Liang, and Guangrong Ma, “Political hierarchy and regional economic development: Evidence from a spatial discontinuity in China,” *Journal of Public Economics*, 2021, *194*, 104352.

Jin, Hehui, Yingyi Qian, and Barry R Weingast, “Regional decentralization and fiscal incentives: Federalism, Chinese style,” *Journal of Public Economics*, 2005, *89* (9-10), 1719–1742.

King, Gary, Jennifer Pan, and Margaret E Roberts, “How censorship in China allows government criticism but silences collective expression,” *American Political Science Review*, 2013, *107* (2), 326–343.

Landry, Pierre F, Xiaobo Lü, and Haiyan Duan, “Does performance matter? Evaluating political selection along the Chinese administrative ladder,” *Comparative Political Studies*, 2018, *51* (8), 1074–1105.

Li, Hongbin and Li-An Zhou, “Political turnover and economic performance: The incentive role of personnel control in China,” *Journal of Public Economics*, 2005, *89* (9-10), 1743–1762.

Li, Xing, Chong Liu, Xi Weng, and Li-An Zhou, “Target setting in tournaments: Theory and evidence from China,” *The Economic Journal*, 2019, *129* (623), 2888–2915.

Liu, Chang and Guangrong Ma, “Are place-based policies always a blessing? Evidence from China’s national poor county programme,” *Journal of Development Studies*, 2019, *55* (7), 1603–1615.

Lyu, Changjiang, Kemin Wang, Frank Zhang, and Xin Zhang, “GDP management to meet or beat growth targets,” *Journal of Accounting and Economics*, 2018, *66* (1), 318–338.

Martinez, Luis R, “How much should we trust the dictator’s GDP growth estimates?,” *Journal of Political Economy*, 2022, *130* (10), 2731–2769.

Maskin, Eric, Yingyi Qian, and Chenggang Xu, “Incentives, information, and organizational form,” *Review of Economic Studies*, 2000, *67* (2), 359–378.

Mehrotra, Aaron and Jenni Pääkkönen, "Comparing China's GDP statistics with coincident indicators," *Journal of Comparative Economics*, 2011, *39* (3), 406–411.

Meng, Lingsheng, "Evaluating China's poverty alleviation program: A regression discontinuity approach," *Journal of Public Economics*, 2013, *101*, 1–11.

Nakamura, Emi, Jón Steinsson, and Miao Liu, "Are Chinese growth and inflation too smooth? Evidence from Engel curves," *American Economic Journal: Macroeconomics*, 2016, *8* (3), 113–144.

Olken, Benjamin A, "Monitoring corruption: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 2007, *115* (2), 200–249.

__ and Rohini Pande, "Corruption in developing countries," *Annual Review of Economics*, 2012, *4* (1), 479–509.

__ and __ , "Governance initiative review paper: J-PAL governance initiative," 2019. Unpublished Manuscript.

Page, Lucy and Rohini Pande, "Ending global poverty: Why money isn't enough," *Journal of Economic Perspectives*, 2018, *32* (4), 173–200.

Park, Albert and Sangui Wang, "China's poverty statistics," *China Economic Review*, 2001, *12* (4), 384–398.

__ and __ , "Community-based development and poverty alleviation: An evaluation of China's poor village investment program," *Journal of Public Economics*, 2010, *94* (9-10), 790–799.

__ , __ , and Guobao Wu, "Regional poverty targeting in China," *Journal of Public Economics*, 2002, *86* (1), 123–153.

Qian, Yingyi, Gerard Roland, and Chenggang Xu, "Coordination and experimentation in M-form and U-form organizations," *Journal of Political Economy*, 2006, *114* (2), 366–402.

Qiu, Tongwei, Xinjie Shi, Yifei Li, and Biliang Luo, "Economic performance of fiscal anti-poverty funds in China," *Journal of Development Studies*, 2023. Forthcoming.

Reinikka, Ritva and Jakob Svensson, "Local capture: Evidence from a central government transfer program in Uganda," *Quarterly Journal of Economics*, 2004, *119* (2), 679–705.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing like China," *American Economic Eeview*, 2011, *101* (1), 196–233.

Staiger, Douglas and James H Stock, "Instrumental variables regression with weak instruments," *Econometrica*, 1997, *65* (3), 557–586.

Storeygard, Adam, "Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa," *Review of Economic Studies*, 2016, *83* (3), 1263–1295.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Wallace, Jeremy L, "Juking the stats? Authoritarian information problems in China," *British Journal of Political Science*, 2016, *46* (1), 11–29.

Wang, Sangui, Zhou Li, and Yanshun Ren, "The 8-7 National Poverty Reduction Program in China—The National Strategy and Its Impact," in "Scaling Up Poverty Reduction: A Global Learning Process and Conference" 2004.

Wang, Zhi, Qinghua Zhang, and Li-An Zhou, "Career incentives of city leaders and urban spatial expansion in China," *Review of Economics and Statistics*, 2020, *102* (5), 897–911.

Wen, Jaya, "State Employment as a Strategy of Autocratic Control in China," 2024. Working Paper.

Wiebe, Michael, "Essays in Chinese political economy." PhD dissertation, University of British Columbia 2020.

World Bank, *World Development Report 2004: Making services work for poor people*, The World Bank, 2003.

__ , *World Development Report 2017: Governance and the law*, The World Bank, 2017.

__ , *Four Decades of Poverty Reduction in China: Drivers, Insights for the World, and the Way Ahead*, The World Bank, 2022.

Xiao, Kezhou and Brantly Womack, "Distortion and credibility within China's internal information system," in "Chinese Authoritarianism in the Information Age," Routledge, 2019, pp. 139–156.

Xiong, Wei, "The mandarin model of growth," 2018. Working Paper.

Xu, Chenggang, "The fundamental institutions of China's reforms and development," *Journal of Economic Literature*, 2011, *49* (4), 1076–1151.

Xu, Gang, L Colin Xu, and Ruichao Si, "Bureaucrats, Tournament Competition, and Performance Manipulation: Evidence from Chinese Cities," 2022. Working Paper.

Young, Alwyn, "Gold into base metals: Productivity growth in the People's Republic of China during the reform period," *Journal of Political Economy*, 2003, *111* (6), 1220–1261.

Zhang, Li, Lunyu Xie, and Xinye Zheng, "Across a few prohibitive miles: The impact of the Anti-Poverty Relocation Program in China," *Journal of Development Economics*, 2023, *160*, 102945.

Zheng, Fengtian and Mizhe Pu, "Examining the negative consequences of the national poor county policy (Yixianfupin Moshi Fumian Xiaoying Tanjiu)," *People's Tribune (Renmin Luntan)*, 2011. (in Chinese).

Zhu, Jiong, Shouying Liu, and Yihao Li, "Removing the "Hats of Poverty": Effects of ending the national poverty county program on fiscal expenditures," *China Economic Review*, 2021, *69*, 101673.

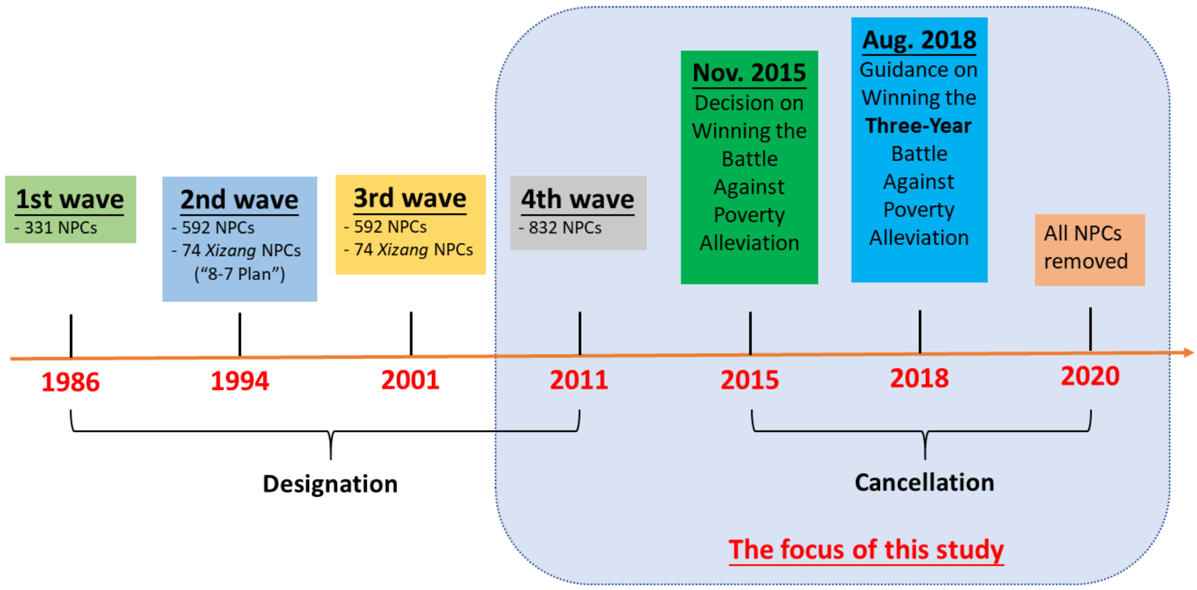

Figure 1: China's Poverty Alleviation Programs

Notes: This figures shows the timeline of the poverty alleviation programs of National Poor County (NPC) designation and cancellation in China for the period 1986-2020. For NPC designation, we focus on the 2011 wave of national poor county selection. For NPC cancellation, we focus on two policies: the 2015 Decision and the 2018 Guidance.

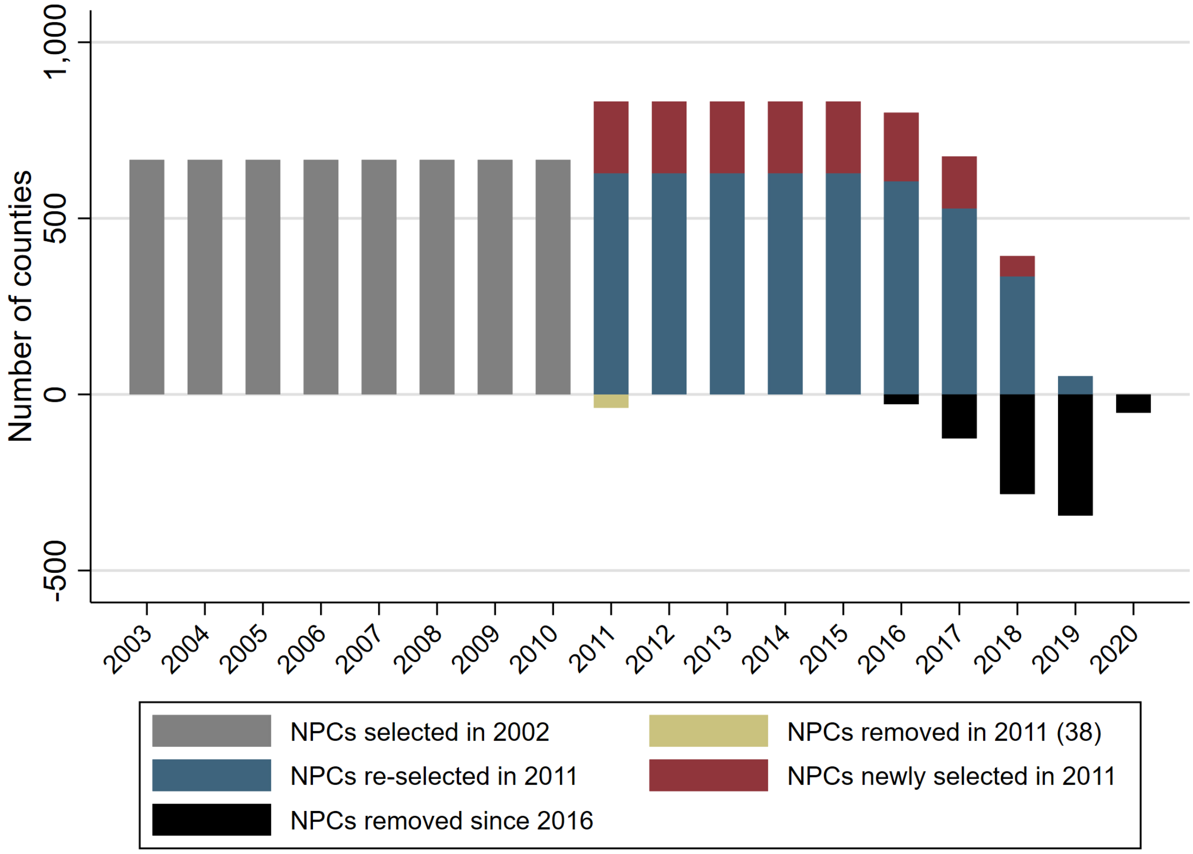

Figure 2: Number of National Poor Counties, 2003-2020

Notes: This figure shows the number of national poor counties for the period 2003-2020. When examining the 2011 designation wave, we use both the newly selected and the reselected counties.

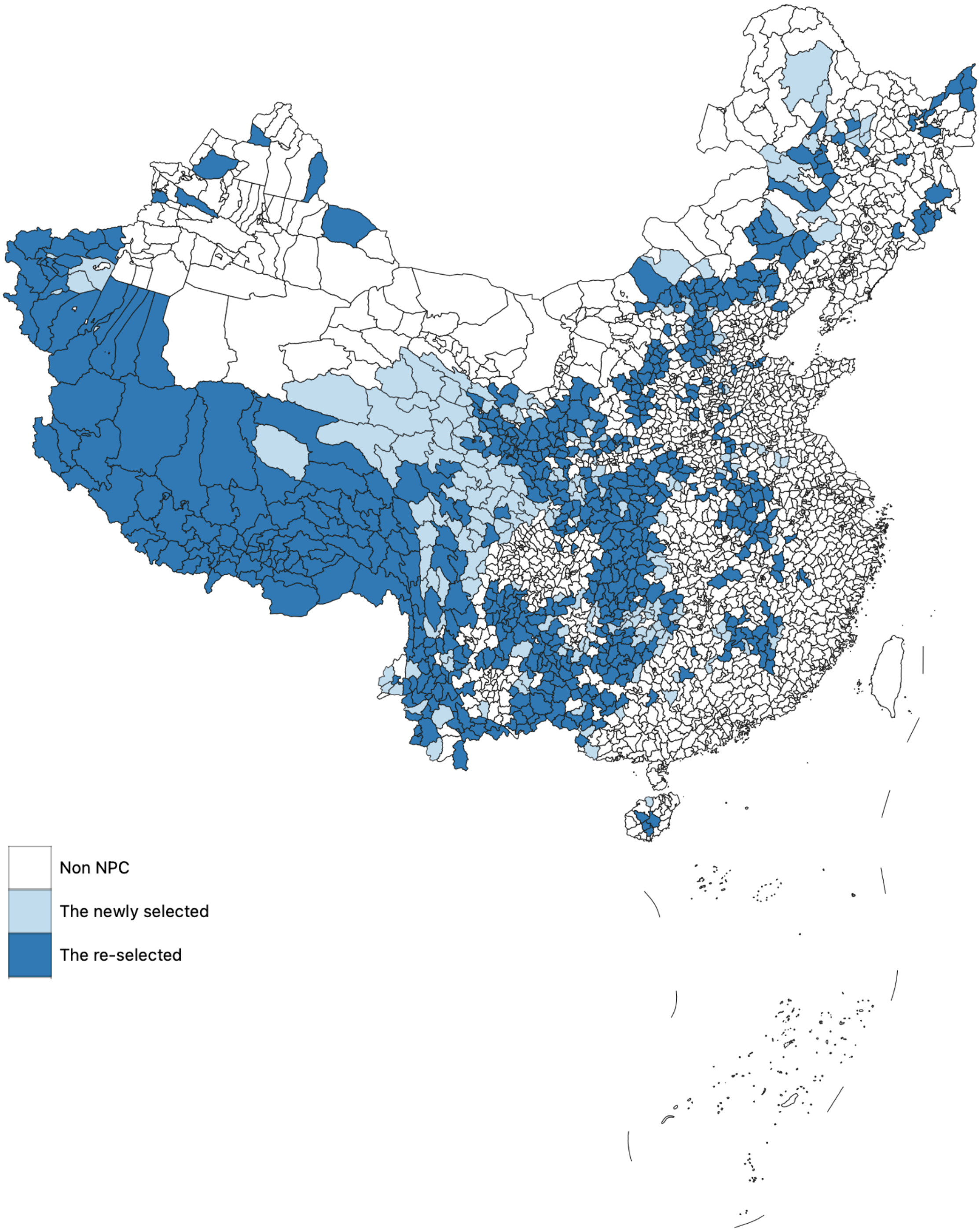

Figure 3: Regional Distribution of National Poor Counties in 2011

Notes: This figure shows the regional distribution of the 832 national poor counties selected in 2011 including both the newly selected and the re-selected counties.

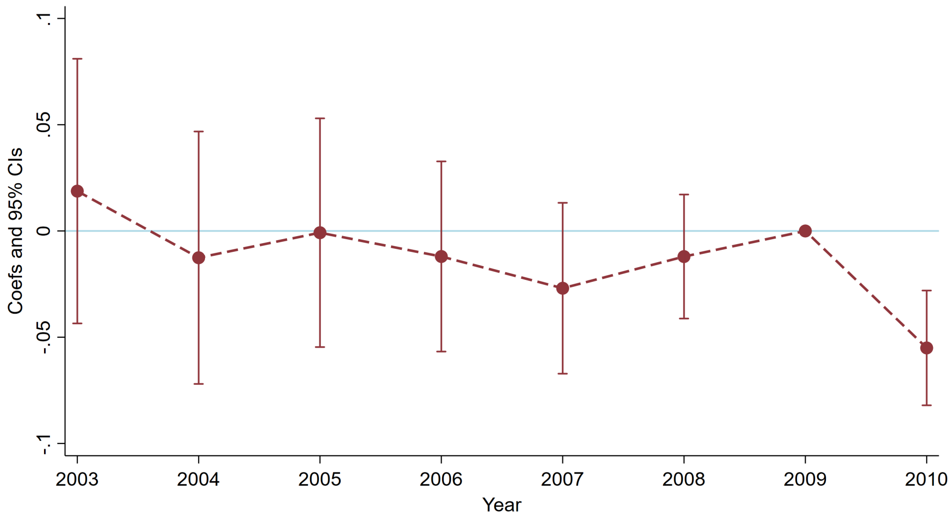

Figure 4: Dynamic Estimates of Being Selected as NPCs in 2011

Notes: The figure shows the dynamic estimates (the estimated coefficients and their corresponding 95% confidence intervals) of being selected as a national poor county on the underreporting of GDP in 2010, obtained by estimating the event study specification of Equation (3).

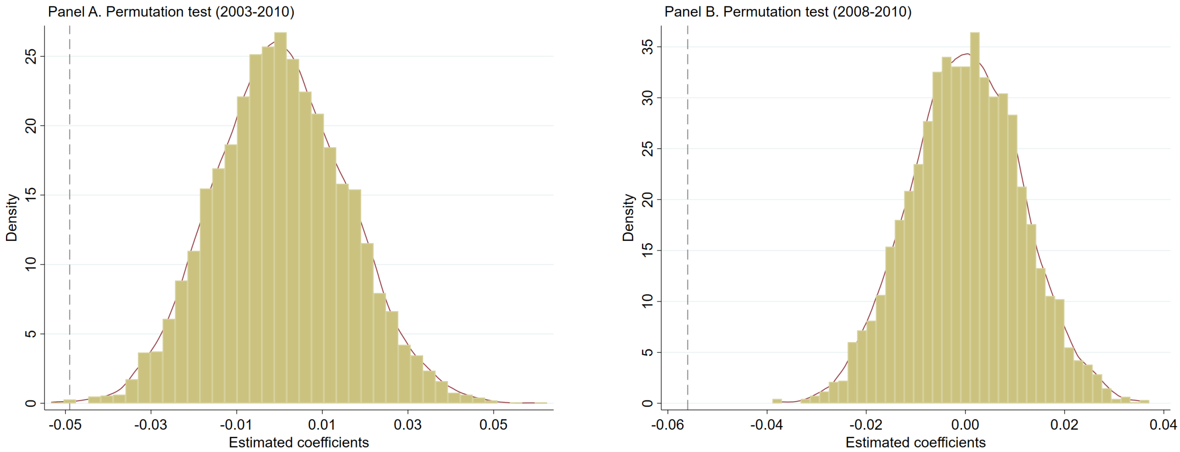

Figure 5: Permutation Tests for Being Selected as NPCs in 2011

Notes: This figure shows the distribution of placebo coefficients from the 5,000 estimated pseudo-treatment effects on GDP underreporting using the baseline model by randomly assigning pseudo-exposure to being selected as a national poor county for the 2003-2010 sample (Panel A) and the 2008-2010 sample (Panel B). The dashed lines in the panels are the value of the coefficients reported in columns 3 and 6 of Table 1, respectively.

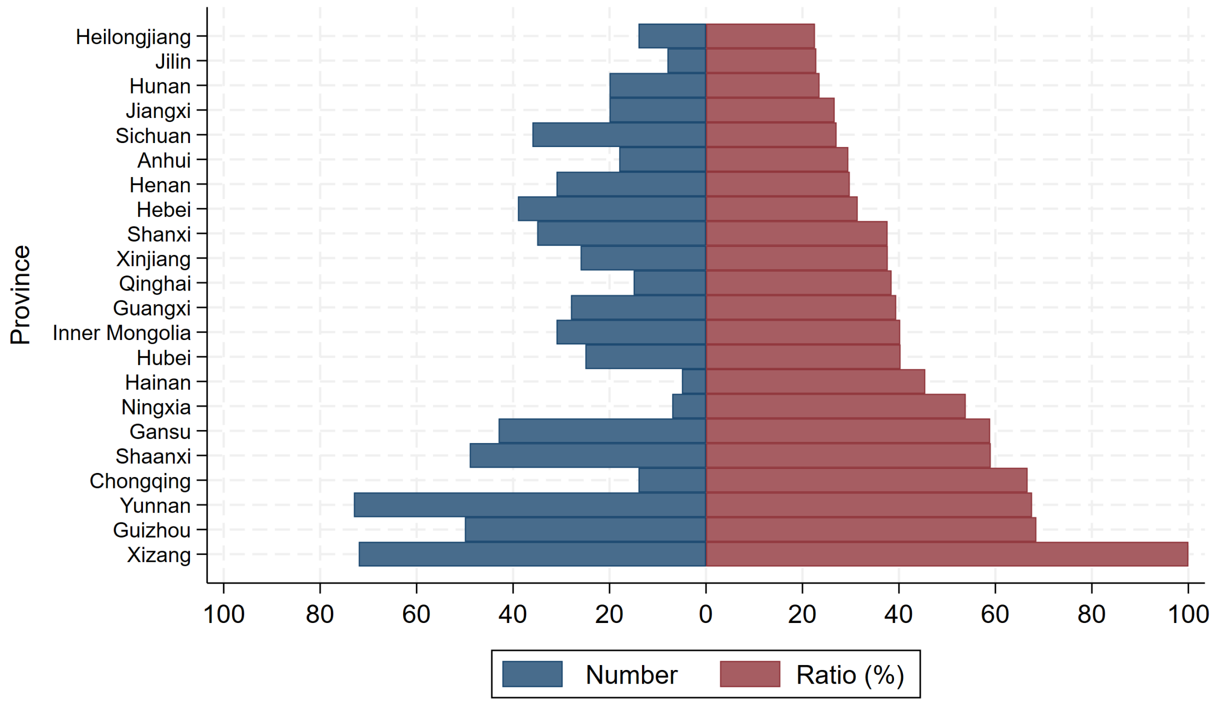

Figure 6: Number and Ratio of NPC Quota across Provinces

Notes: This figure shows the number (the left part) and ratio (the right part) of national poor counties assigned to each province (the quota of national poor counties that one province can designate in the 2011 selection wave). Note that provinces in the coastal region were not assigned the NPC quota, so they are not included in this figure.

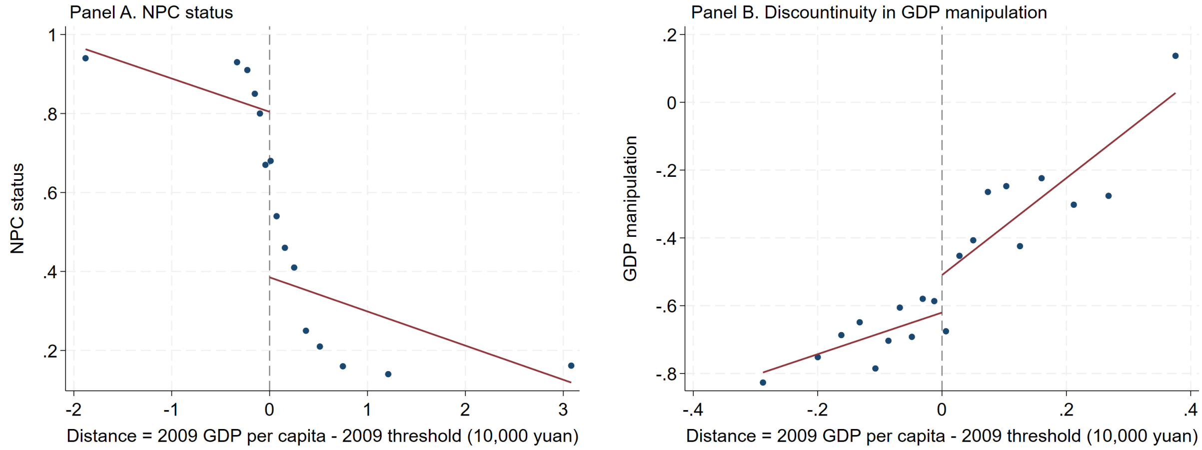

Figure 7: NPC Status and Discontinuity in GDP Manipulation

Notes: Panel A plots the within-province distance between each county's 2009 GDP per capita and the 2009 threshold (the x-axis) against the proportion of poor counties selected in 2011 in a given bin (the y-axis). Panel B plots the within-province distance between each county's 2009 GDP per capita and the 2009 threshold (the x-axis) against the residualized GDP manipulation variable, obtained by regressing the 2010 GDP manipulation measure on provincial fixed effects and baseline control variables measured in 2009. The 2009 threshold in a given province is calculated based on the quota of national poor counties assigned to that province and the GDP per capita ranking of the counties in that province, so that counties with GDP per capita below the threshold would be eligible to be selected as national poor counties. Panel B uses a restricted sample ($Distcane_{ip,2009} \in [-0.4,\ 0.4]$).

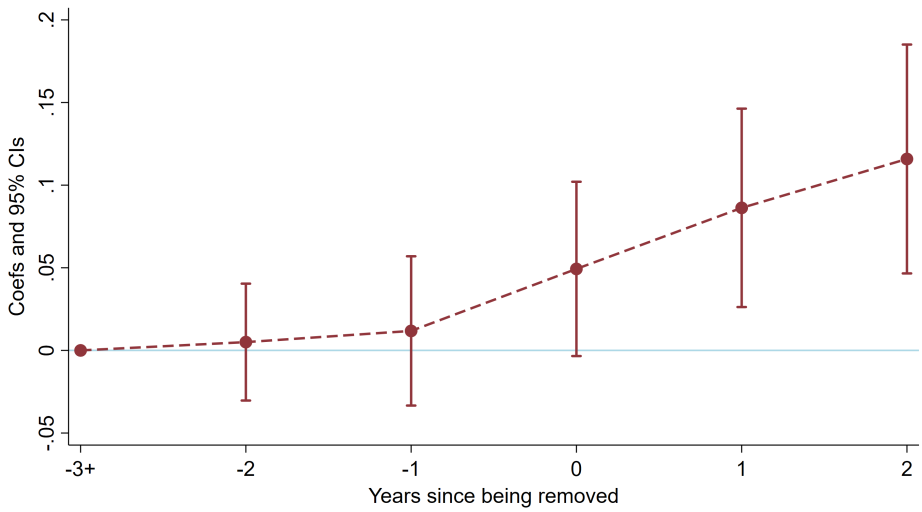

Figure 8: Dynamic Estimates of Being Removed from Poor Status since 2018

Notes: The figure shows the dynamic estimates (the estimated coefficients and their corresponding 95% confidence intervals) of being removed from poor status on the overreporting of GDP for the 2016-2020 sample, obtained by estimating the event study specification of Equation (7).

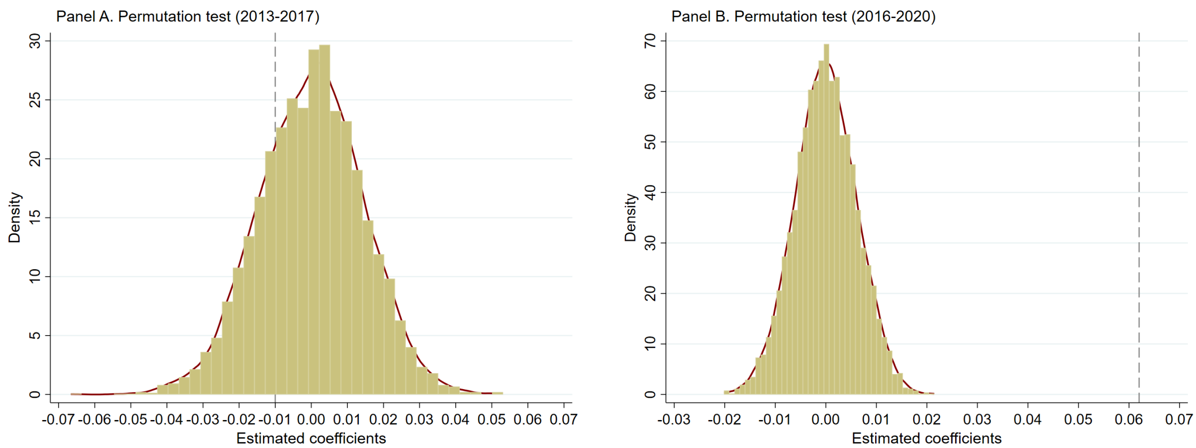

Figure 9: Permutation Tests for Being Removed from Poor Status

Notes: This figure shows the distribution of placebo coefficients from the 5,000 estimated pseudo-treatment effects on GDP overreporting using the baseline model by randomly assigning pseudo-exposure to being removed from poor status for the 2013-2017 sample (Panel A) and the 2016-2020 sample (Panel B). The dashed lines in the panels are the value of the coefficients reported in columns 1 and 2 of Table 7, respectively.

Table 1: NPC Designation and Underreporting

	GDP manipulation					
	2003-2010			2008-2010		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All NPCs selected in 2011						
NPC × 2010	-0.130***	-0.101***	-0.049**	-0.111***	-0.074***	-0.056***
	(0.016)	(0.019)	(0.019)	(0.013)	(0.015)	(0.014)
Num. obs.	14591	12558	12558	5465	5410	5410
Num. clu.	1835	1573	1573	1828	1806	1806
Adj. R-sq.	0.970	0.969	0.973	0.985	0.985	0.987
Panel B. NPCs newly selected in 2011						
NPC × 2010	-0.122***	-0.127***	-0.095***	-0.075***	-0.051**	-0.050**
	(0.027)	(0.030)	(0.033)	(0.020)	(0.022)	(0.022)
Num. obs.	9627	8311	8311	3612	3593	3593
Num. clu.	1207	1040	1040	1205	1198	1198
Adj. R-sq.	0.970	0.968	0.973	0.983	0.983	0.985
Panel C. NPCs re-selected in 2011						
NPC × 2010	-0.133***	-0.091***	-0.039*	-0.122***	-0.083***	-0.063***
	(0.017)	(0.022)	(0.021)	(0.014)	(0.016)	(0.016)
Num. obs.	13037	11161	11161	4881	4834	4834
Num. clu.	1640	1398	1398	1633	1614	1614
Adj. R-sq.	0.969	0.968	0.972	0.984	0.985	0.987
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	No	Yes	Yes	No
Province-year FEs	No	No	Yes	No	No	Yes
2003 controls × year FEs	No	Yes	Yes	No	No	No
2008 controls × year FEs	No	No	No	No	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2003-2010 and 2008-2010 for columns 1-3 and 4-6 respectively, and counties including national poor counties designed in 2011 and counties never designed. Panels A, B, and C use all selected counties, the newly selected counties, and the re-selected counties in 2011 as the treatment group, respectively. The control group is the counties that were never selected. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2003 for columns 1-3 and 2008 for columns 4-6). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 2: NPC Designation and Underreporting: IV Estimation

	NPC X 2010		GDP manipulation	
	First-stage		Second-stage	
	2003-2010	2008-2010	2003-2010	2008-2010
	(1)	(2)	(3)	(4)
Distance × 2010	-0.069***	-0.060***		
	(0.015)	(0.016)		
NPC × 2010			-1.498***	-0.269
			(0.482)	(0.372)
Kleiberg-Paap Wald F-statistic	19.585	13.854		
Num. obs.	10582	4438	10582	4438
Num. clu.	1324	1480	1324	1480
Adj. R-sq.	0.702	0.658	-0.618	-0.056
County FEs	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes
2003 controls × year FEs	Yes	No	Yes	No
2008 controls × year FEs	No	Yes	No	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2003-2010 and 2008-2010 for columns 1 and 3 and 2 and 4 respectively, and counties including national poor counties designed in 2011 and counties never designed. The second-stage dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The instrumental variable (Distance) is the within-province distance between each county's 2009 GDP per capita and the 2009 threshold. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2003 for columns 1 and 3 and 2008 for columns 2 and 4). The Kleiberg-Paap Wald F-statistic is reported. Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 3: NPC Designation and Underreporting: Fuzzy RD Analysis

	NPC				GDP manipulation			
	First-stage				Second-stage			
	Full sample		[-3, 3]	[-2, 2]	Full sample		[-3, 3]	[-2, 2]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eligible	0.287***	0.275***	0.253***	0.216***				
	(0.036)	(0.039)	(0.037)	(0.036)				
NPC					-1.467***	-1.100***	-0.846**	-0.623*
					(0.432)	(0.336)	(0.333)	(0.361)
Kleiberg-Paap Wald F-statistic	62.169	50.520	47.053	35.583				
Num. obs.	1482	1482	1447	1350	1482	1482	1447	1350
Adj. R-sq.	0.435	0.436	0.440	0.417	-0.102	0.173	0.239	0.307
2009 controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Provincial FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Running variable	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Notes: The unit of analysis is the county. The sample is cross-sectional and excludes provinces with no counties selected as national poor counties in 2011. Columns 3-4 and 7-8 report the results obtained using restricted samples by reducing the distance (the unit is 10,000 yuan). The second-stage dependent variable (GDP manipulation measure in 2010) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The instrumental variable (Eligible) is a dummy variable that is one if a county's within-province distance (i.e., the running variable) is zero or negative, and zero if it is positive. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured in 2009. The Kleiberg-Paap Wald F-statistic is reported. Robust standard errors in parentheses are clustered by province. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 4: NPC Designation and Underreporting: Distance-Based Analysis

	GDP manipulation							
	2003-2010				2008-2010			
			Distance>0	Distance≤0			Distance>0	Distance≤0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mid$ Distance $\mid$ $\times$ 2010	0.067**	0.043	0.031	-1.131***	-0.010	-0.025	-0.023	-0.447***
	(0.026)	(0.030)	(0.032)	(0.165)	(0.024)	(0.027)	(0.028)	(0.145)
$\mid$ Distance $\mid$ $\times$ 2010 $\times$ NPC		0.084*	0.097***	0.637***		0.059*	0.066**	0.277*
		(0.044)	(0.037)	(0.192)		(0.035)	(0.031)	(0.144)
NPC $\times$ 2010		-0.067**	-0.026	-0.031		-0.074***	-0.084***	-0.022
		(0.028)	(0.032)	(0.042)		(0.021)	(0.028)	(0.037)
Num. obs.	10037	10037	6077	3960	4182	4182	2490	1692
Num. clu.	1255	1255	760	495	1394	1394	830	564
Adj. R-sq.	0.965	0.965	0.974	0.956	0.981	0.981	0.986	0.978
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2003 controls $\times$ year FEs	Yes	Yes	Yes	Yes	No	No	No	No
2008 controls $\times$ year FEs	No	No	No	No	Yes	Yes	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2003-2010 and 2008-2010 for columns 1-4 and 5-8 respectively, and counties including national poor counties designed in 2011 and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The | Distance | is the absolute value of the within-province distance between each county's 2009 GDP per capita and the 2009 threshold. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2003 for columns 1-4 and 2008 for columns 5-8). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 5: NPC Designation and Underreporting: Fiscal Concerns

	GDP manipulation			
	2003-2010		2008-2010	
	(1)	(2)	(3)	(4)
NPC × 2010	0.014	-0.047**	-0.011	-0.054***
	(0.022)	(0.019)	(0.018)	(0.017)
NPC × 2010 × fiscal dependence (demeaned)	-0.681***		-0.547***	
	(0.173)		(0.137)	
NPC × 2010 × fiscal pressure (demeaned)		-0.531*		-0.155
		(0.305)		(0.261)
2010 × fiscal dependence (demeaned)	-0.076		0.081	
	(0.095)		(0.080)	
2010 × fiscal pressure (demeaned)		0.477***		0.227
		(0.173)		(0.166)
Num. obs.	9438	9438	4201	4201
Num. clu.	1181	1181	1401	1401
Adj. R-sq.	0.969	0.969	0.985	0.985
County FEs	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes
2003 controls × year FEs	Yes	Yes	No	No
2008 controls × year FEs	No	No	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2003-2010 and 2008-2010 for columns 1-2 and 3-4 respectively, and counties including national poor counties designed in 2011 and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. Fiscal dependence is measured as the annual average transfer-to-revenue ratio (a county's fiscal transfers from higher-level governments as a share of its fiscal revenues) over the sample period. Fiscal pressure is the annual average deficit-to-revenue ratio (subtracting fiscal expenditures from fiscal revenues as a share of its fiscal revenues) over the sample period. Both fiscal dependence and fiscal pressure are time invariant and demeaned in the regressions. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2003 for columns 1-2 and 2008 for columns 3-4). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 6: NPC Designation and Underreporting: CPC Secretaries

	GDP manipulation			
	2003-2010	2008-2010	2003-2010	2008-2010
	(1)	(2)	(3)	(4)
NPC × 2010	-0.058***	-0.060***	-0.037**	-0.036**
	(0.020)	(0.016)	(0.018)	(0.016)
Num. obs.	10193	4423	9594	3879
Num. clu.	1401	1495	1398	1481
Adj. R-sq.	0.968	0.983	0.979	0.984
County FEs	Yes	Yes	No	No
Tenure FEs	Yes	Yes	No	No
CPC Secretary FEs	No	No	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes
2003 controls × year FEs	Yes	No	Yes	No
2008 controls × year FEs	No	Yes	No	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2003-2010 and 2008-2010 for columns 1 and 3 and 2 and 4 respectively, and counties including national poor counties designed in 2011 and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. First two years in office is a dummy variable that equals one if the county party secretary is in his or her first two years in office. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2003 for columns 1 and 3 and 2008 for columns 2 and 4). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 7: NPC Cancellation and Overreporting

	GDP manipulation				
	2013-2017	2016-2020			
	Removed 16-17	Removed 18-20			
	(1)	(2)	(3)	(4)	(5)
NPC removed	-0.016	0.056^{***}			
	(0.026)	(0.014)			
NPC removed 2018 $\times \geq$2018			0.060^{***}		
			(0.021)		
NPC removed 2019 $\times \geq$2019				0.084^{***}	
				(0.023)	
NPC removed 2020 $\times$ =2020					-0.025
					(0.049)
Num. obs.	5754	7117	5606	5868	4863
Num. clu.	1153	1444	1142	1198	1000
Adj. R-sq.	0.924	0.941	0.936	0.940	0.941
County FEs	Yes	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes	Yes
2013 controls $\times$ year FEs	Yes	No	No	No	No
2016 controls $\times$ year FEs	No	Yes	Yes	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2013-2017 and 2016-2020 for columns 1 and 2-5 respectively, and counties including national poor counties removed since 2016 (those removed in 2016-2017 for column 1 and those removed in 2018-2020 for columns 2-5) and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The variable NPC Removed equals one if a county has been removed from poor status in a given year and zero otherwise. We disaggregate the counties removed in 2018-2020 into different single years in columns 3-5 (corresponding to 2018, 2019, and 2020, respectively). Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2013 for column 1 and 2016 for columns 2-5). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 8: NPC Cancellation and Overreporting: Local Self-Discipline

	GDP manipulation		
	2016-2020		
	Removed 18-20		
		Poor quality	Good quality
	(1)	(2)	(3)
NPC removed	0.094***	0.082***	0.024
	(0.019)	(0.020)	(0.019)
NPC removed × good quality	-0.071***		
	(0.025)		
Num. obs.	7117	3027	3961
Num. clu.	1444	643	819
Adj. R-sq.	0.941	0.938	0.938
County FEs	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes
2016 controls × year FEs	Yes	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2016-2020, and counties including national poor counties removed since 2018 and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The variable NPC Removed equals one if a county has been removed from poor status in a given year and zero otherwise. Good (poor) quality is a dummy variable that equals one if a county is located in a province with (without) disciplined statistical practice, calculated based on the 4th National Economic Census of China, and zero otherwise. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2016). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 9: NPC Cancellation and Overreporting: Central Supervision

	GDP manipulation				
	2016-2020				
				Poor quality	Good quality
	(1)	(2)	(3)	(4)	(5)
NPC removed 2018 × ≥2018	0.065**				
	(0.026)				
NPC removed 2019 × ≥2019		0.085***			
		(0.025)			
NPC removed 2020 × =2020			0.096*	0.200***	-0.037
			(0.051)	(0.060)	(0.087)
NPC removed 2018 × ≥2018 × supervision	-0.011				
	(0.036)				
NPC removed 2019 × ≥2019 × supervision		-0.003			
		(0.041)			
NPC removed 2020 × =2020 × supervision			-0.309***	-0.147*	-0.156
			(0.087)	(0.076)	(0.114)
Num. obs.	5606	5868	4863	1948	2804
Num. clu.	1142	1198	1000	418	595
Adj. R-sq.	0.936	0.940	0.941	0.939	0.933
County FEs	Yes	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes	Yes
2016 controls × year FEs	Yes	Yes	Yes	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2016-2020, and counties including national poor counties removed since 2018 and counties never designed. The counties removed in 2018-2020 are disaggregated into different single years. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. Supervision is a dummy variable that takes the value of one if a county is located in a province that was supervised by the National Bureau of Statistics of China in late 2019 or 2020 and zero otherwise. Good (poor) quality is a dummy variable that equals one if a county is located in a province with (without) disciplined statistical practice, calculated based on the 4th National Economic Census of China, and zero otherwise. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2016). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table 10: NPC Cancellation and Overreporting: CPC Secretaries

	GDP manipulation				
	2013-2017	2016-2020	2013-2017	2016-2020	
	Removed 16-17	Removed 18-20	Removed 16-17	Removed 18-20	
	(1)	(2)	(3)	(4)	(5)
NPC removed	0.006	0.052***	0.035	0.058***	0.013
	(0.029)	(0.015)	(0.032)	(0.014)	(0.016)
NPC removed × last two years in office					0.069***
					(0.019)
Last two years in office					-0.010
					(0.013)
Num. obs.	5217	6259	4834	5883	5883
Num. clu.	1057	1286	1056	1279	1279
Adj. R-sq.	0.922	0.938	0.927	0.943	0.943
County FEs	Yes	Yes	No	No	No
Tenure FEs	Yes	Yes	No	No	No
Province-year FEs	Yes	Yes	Yes	Yes	Yes
CPC Secretary FEs	No	No	Yes	Yes	Yes
2013 controls × year FEs	Yes	No	Yes	No	No
2016 controls × year FEs	No	Yes	No	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2013-2017 and 2016-2020 for columns 1 and 3 and 2, 4, and 5 respectively, and counties including national poor counties removed since 2016 (those removed in 2016-2017 for columns 1 and 3 and those removed in 2018-2020 for columns 2, 4, and 5) and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The variable NPC Removed equals one if a county has been removed from poor status in a given year and zero otherwise. Last two years in office is a dummy variable that equals one if the county party secretary is in his or her last two years in office. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2013 for column 1 and 2016 for columns 2-5). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Online Appendix

The Shapeshifting Hat: Strategic Response to Fiscal and Political Incentives of Chinese National Poor Counties

Jianhao Lin, Tingwei Luo, Wenbiao Sha, and Jiasong Xie

A Appendix Figures and Tables

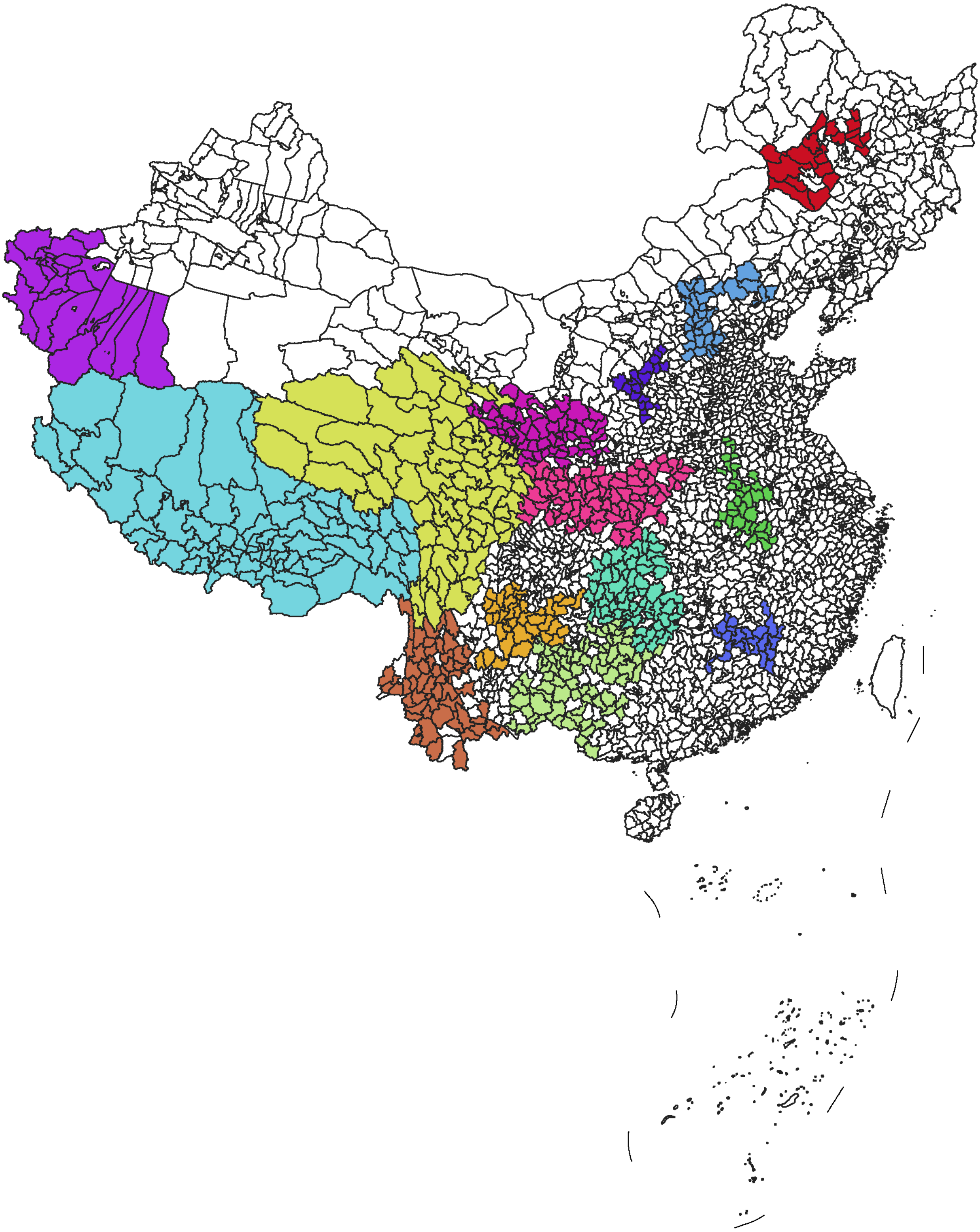

Figure A1: National Poor Counties in the 14 Concentrated and Contiguous Impoverished Areas

Notes: This figure shows the regional distribution of the 14 concentrated and contiguous impoverished areas (Chinese: *14 Ge Jizhong Lianpian Tekun Diqu*), designated in 2011.

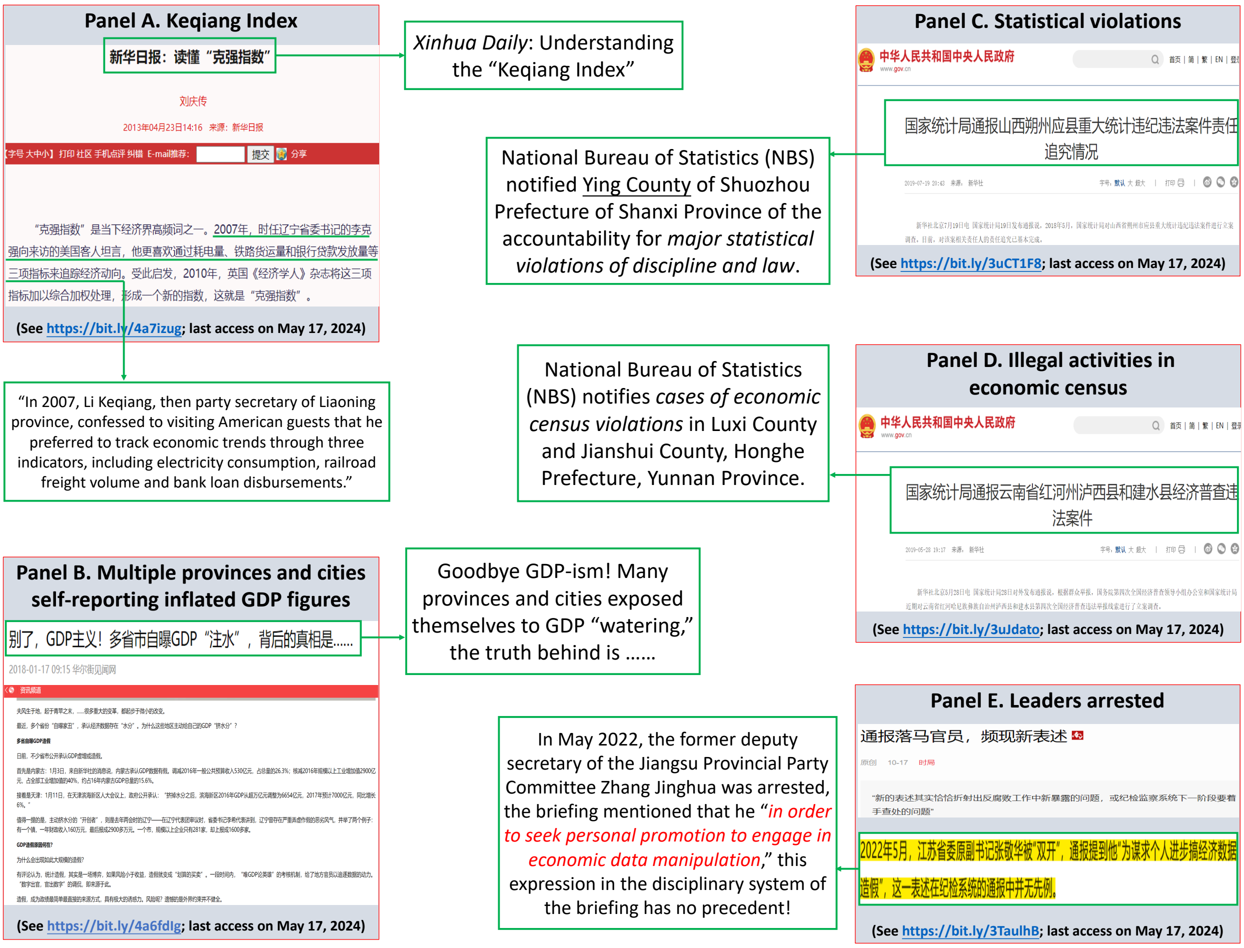

Figure A2: Media Coverage on Manipulation about Economic Data

Notes: This figure shows the media coverage of GDP manipulation by local governments.

Figure A3: Media Coverage on Anti-poverty-related Data Manipulation

Notes: This figure shows the media coverage of anti-poverty-related data fabrication.

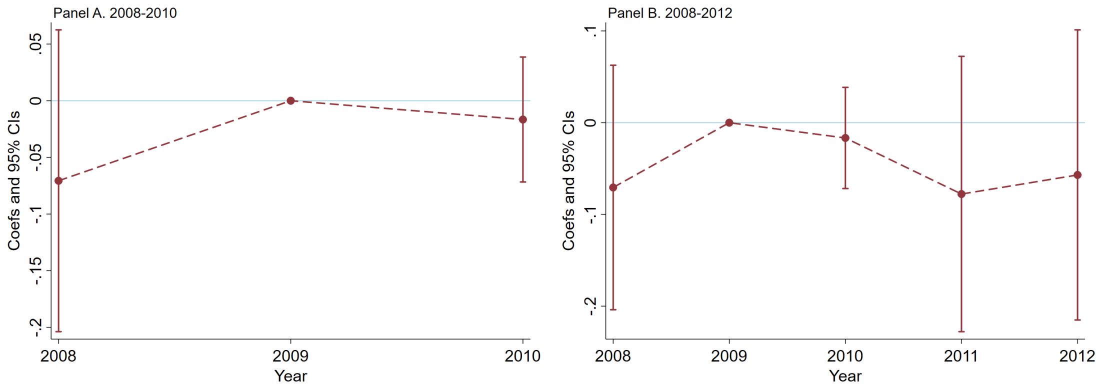

Figure A4: Dynamic Estimates of Being Removed from Poor Status in 2011

Notes: The figure shows the dynamic estimates (the estimated coefficients and their corresponding 95% confidence intervals) of being removed from poor status in 2011 on GDP reporting using the 38 counties removed in 2011.

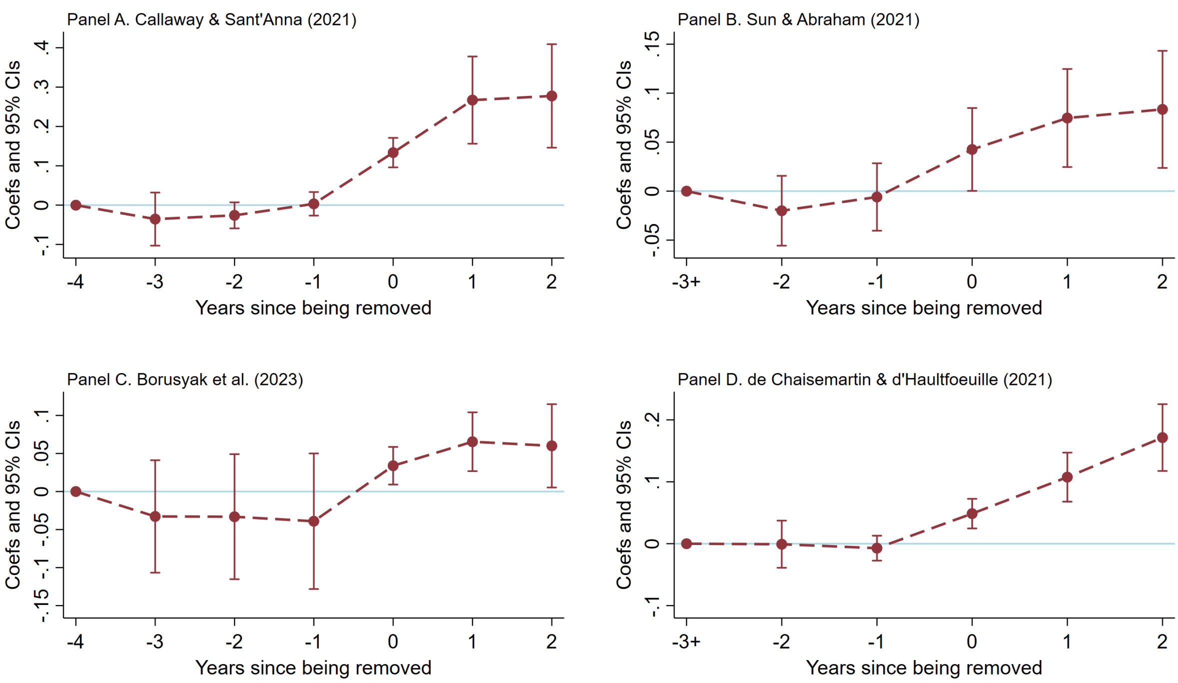

Figure A5: Dynamic Estimates of Being Removed from Poor Status since 2018: New Estimators

Notes: The figure shows the dynamic estimates (the estimated coefficients and their corresponding 95% confidence intervals) of being removed from poor status on the overreporting of GDP for the 2016-2020 sample, obtained by estimating the event study specification of Equation (7) using new estimators: Callaway and Sant'Anna (2021) for Panel A, Sun and Abraham (2021) for Panel B, Borusyak et al. (2023) for Panel C, and de Chaisemartin and d'Haultfoeuille (2020) for Panel D.

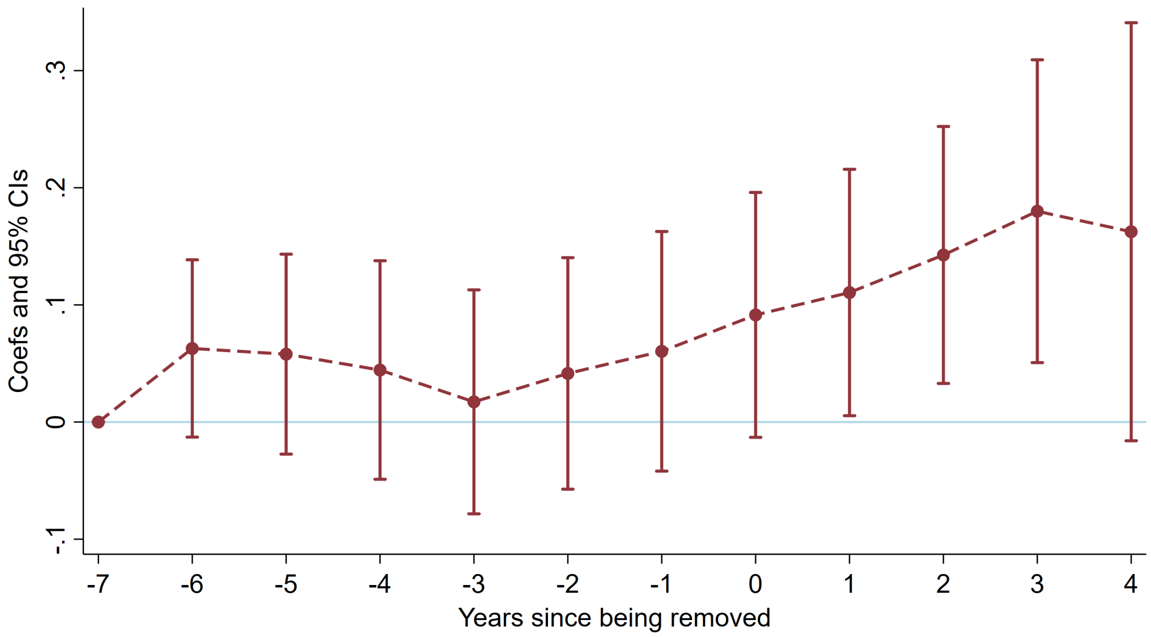

Figure A6: Dynamic Estimates of Being Removed from Poor Status since 2016

Notes: The figure shows the dynamic estimates (the estimated coefficients and their corresponding 95% confidence intervals) of being removed from poor status on the overreporting of GDP for the *pooled* 2013-2020 sample, obtained by estimating the event study specification of Equation (7).

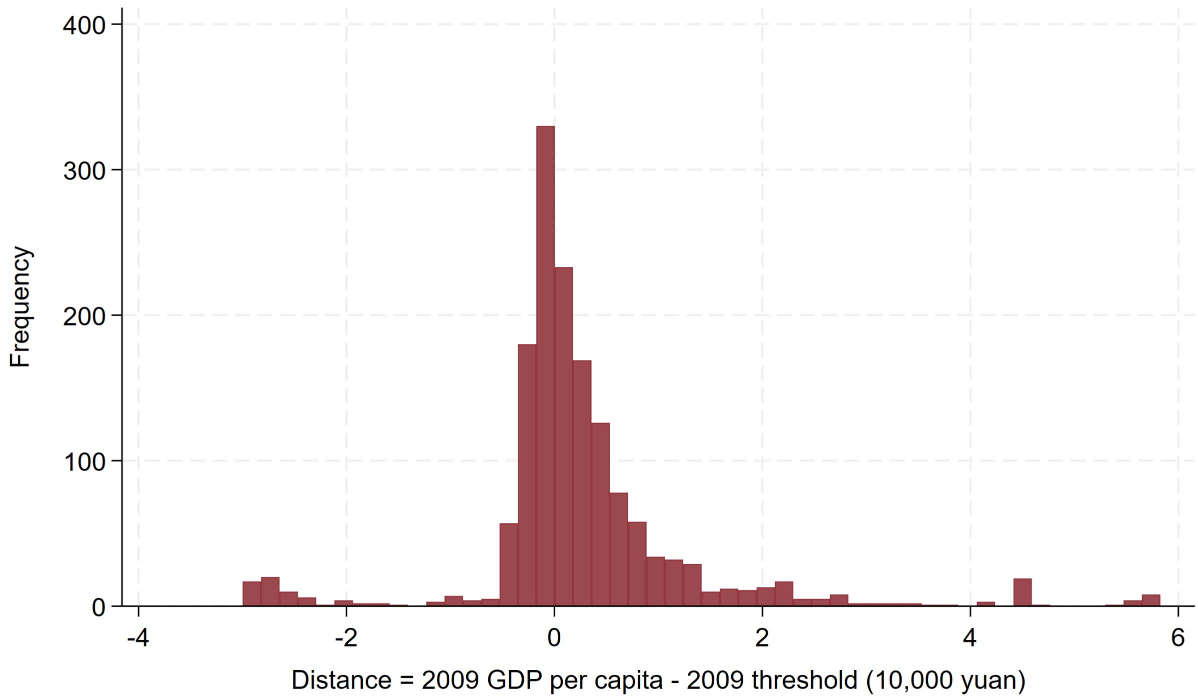

Figure A7: Histogram of the 2009 Distance Measure

Notes: This figure plots histogram of the 2009 distance measure (mean = 0.275, max = 5.826, min = -2.998).

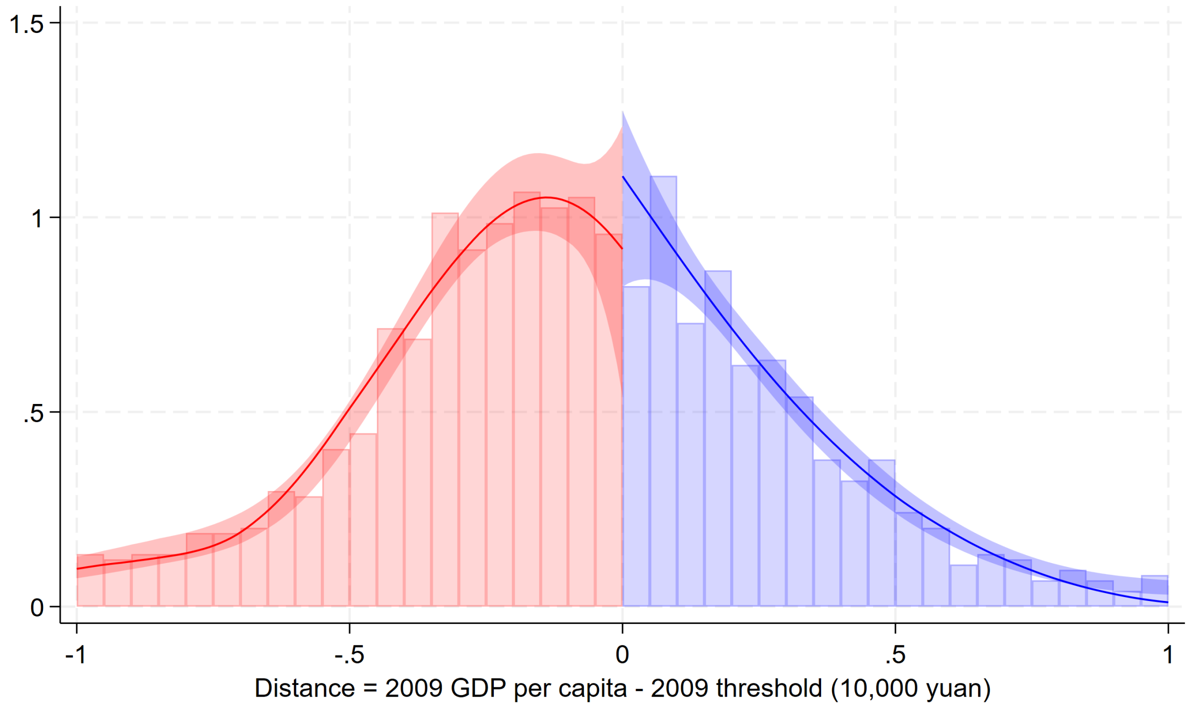

Figure A8: Testing of Non-manipulation of Running Variable

Notes: This figure plots the density of the residualized running variable (i.e., within-province distance between each county's 2009 GDP per capita and the 2009 threshold).

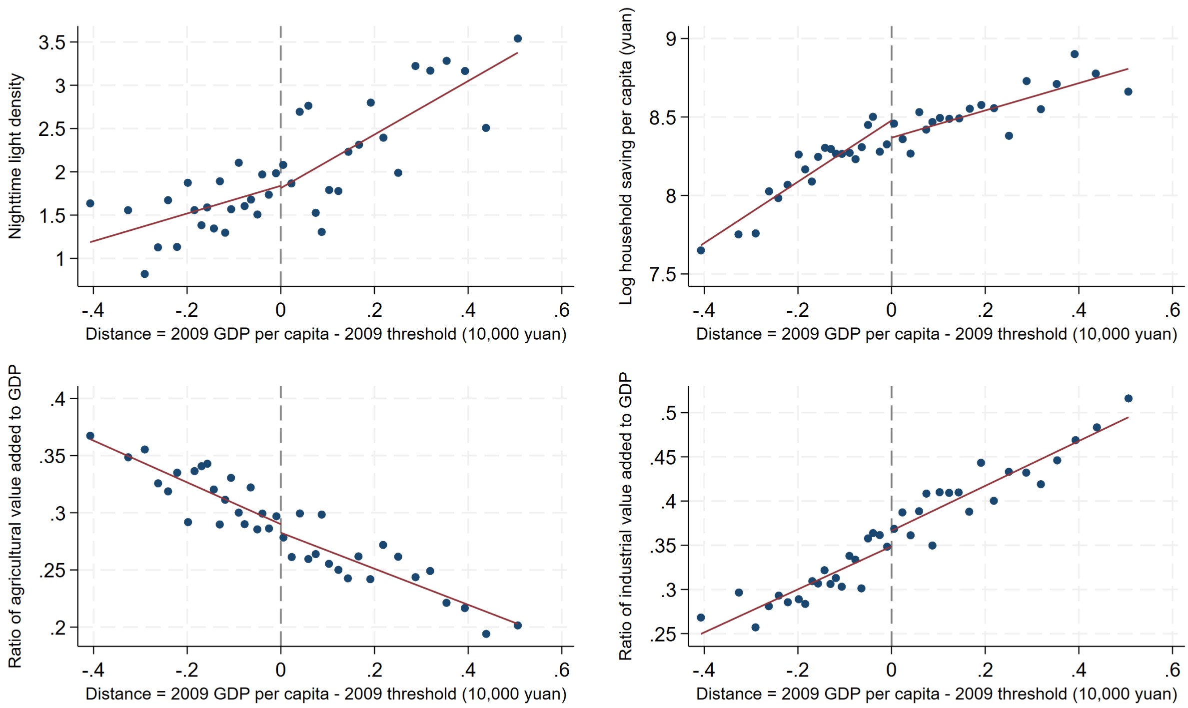

Figure A9: Testing of Balancing of Covariates

Notes: This figure plots the within-province distance between each county's 2009 GDP per capita and the 2009 threshold (the x-axis) against the residualized covariates (the y-axis) in different panels.

Table A1: Summary Statistics for Key Variables

	Mean	SD	Min	p25	p50	p75	Max
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. The 2003-2010 sample (obs. = 14,864)							
GDP manipulation	-0.44	1.52	-3.02	-1.53	-0.67	0.27	4.10
All NPCs selected in 2011	0.44	0.50	0.00	0.00	0.00	1.00	1.00
NPCs newly selected in 2011	0.10	0.31	0.00	0.00	0.00	0.00	1.00
NPCs re-selected in 2011	0.33	0.47	0.00	0.00	0.00	1.00	1.00
Nighttime light density (DMSP)	3.25	4.91	0.00	0.42	1.28	4.22	49.25
Log household saving per capita (yuan)	8.28	0.78	5.91	7.84	8.31	8.75	10.80
Ratio of agricultural value added to GDP (%)	27.43	13.29	2.32	17.51	27.22	36.42	63.23
Ratio of industrial value added to GDP (%)	39.42	16.00	8.46	27.28	38.20	50.71	77.83
Ratio of fiscal gap to fiscal revenue (%)	-3.96	8.31	-30.82	-7.99	-2.86	-0.36	27.17
Ratio of transfer payment to fiscal revenue (%)	44.21	25.40	-0.97	23.97	38.53	66.16	95.36
Years in office of county party secretary	4.60	1.84	1.00	3.00	5.00	6.00	14.00
Panel B. The 2013-2020 sample (obs. = 14,764)							
GDP manipulation	0.68	0.76	-1.08	0.66	0.20	1.13	2.88
NPCs removed in 2016 and 2017	0.08	0.27	0.00	0.00	0.00	0.00	1.00
NPCs removed in 2018	0.15	0.36	0.00	0.00	0.00	0.00	1.00
NPCs removed in 2019	0.18	0.38	0.00	0.00	0.00	0.00	1.00
NPCs removed in 2020	0.03	0.16	0.00	0.00	0.00	0.00	1.00
Nighttime light density (VIIRS)	1.21	1.14	0.14	0.93	0.63	1.42	22.20
Log household saving per capita (yuan)	9.48	0.64	5.91	9.50	9.12	9.88	10.80
Ratio of agricultural value added to GDP (%)	20.10	11.22	2.32	18.74	11.99	26.37	63.23
Ratio of industrial value added to GDP (%)	41.12	15.08	8.46	41.63	30.37	51.59	77.83
Years in office of county party secretary	5.10	2.18	1.00	5.00	3.00	6.00	13.00

Notes: GDP manipulation is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The GDP data are from various years of the County Statistical Yearbook of China, and the night light intensity data are from the Global Night-time Light Database (GNLD) provided by the Chinese Research Data Services Platform (CNRDS). Data on national poor counties are from the State Council of China. Data on county party secretaries are collected manually from the Internet. Data on the other variables are also taken from various years of the County Statistical Yearbook of China.

Table A2: NPC Cancellation and Overreporting: New Estimators

	GDP manipulation							
	Callaway & Sant'Anna (2021)		Sun & Abraham (2021)		Borusyak et al. (2023)		de Chaisemartin & d'Haultfoeuille (2021)	
	2013-2017	2016-2020	2013-2017	2016-2020	2013-2017	2016-2020	2013-2017	2016-2020
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NPC Removed	-0.003	0.210***	-0.018	0.054***	-0.029	0.034***	0.004	0.049***
	(0.034)	(0.039)	(0.026)	(0.013)	(0.028)	(0.013)	(0.033)	(0.011)
Num. obs.	5751	7095	5754	7117	5733	7437	2270	3587
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2013 controls × year FEs	Yes	No	Yes	No	Yes	No	Yes	No
2016 controls × year FEs	No	Yes	No	Yes	No	Yes	No	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2013-2017 and 2016-2020 for columns 1, 3, 5, and 7 and 2, 4, 6, and 8 respectively, and counties including national poor counties removed since 2016 (those removed in 2016-2017 for columns 1, 3, 5, and 7 and those removed in 2018-2020 for columns 2, 4, 6, and 8) and counties never designed. The variable NPC Removed equals one if a county has been removed from poor status in a given year and zero otherwise. The coefficients and standard errors are estimated using new estimators: Callaway and Sant'Anna (2021) for columns 1-2, Sun and Abraham (2021) for columns 3-4, Borusyak et al. (2023) for columns 5-6, and de Chaisemartin and d'Haultfoeuille (2020) for columns 7-8. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2013 for columns 1, 3, 5, and 7 and 2016 for columns 2, 4, 6, and 8). *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table A3: Geographic Characteristics and Years of Being Removed

	Geographic characteristics							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Dependent variable: distance to the provincial capital (km)								
Removed in 2011	21.801			31.974	-121.913			-85.420
	(34.499)			(32.528)	(141.709)			(127.889)
Removed in 2016-17		21.753		32.726*		28.856**		13.664
		(21.960)		(19.197)		(12.927)		(22.203)
Removed in 2018-20			84.536***	84.088***			67.134**	71.030**
			(22.661)	(22.867)			(30.160)	(30.850)
Num. obs.	1203	1252	1657	1768	1203	1252	1657	1768
Num. clu.	28	29	29	29	28	29	29	29
Adj. R-sq.	-0.000	0.001	0.047	0.043	0.338	0.383	0.374	0.334
Panel B. Dependent variable: terrain ruggedness (average slope)								
Removed in 2011	-0.936			-0.195	1.308			2.197***
	(1.718)			(1.718)	(0.904)			(0.790)
Removed in 2016-17		5.410***		5.954***		2.534***		3.497***
		(1.575)		(1.624)		(0.797)		(1.008)
Removed in 2018-20			6.003***	6.247***			3.900***	4.085***
			(0.925)	(0.983)			(0.662)	(0.765)
Num. obs.	1203	1252	1657	1768	1203	1252	1657	1768
Num. clu.	28	29	29	29	28	29	29	29
Adj. R-sq.	-0.000	0.001	0.047	0.043	0.338	0.383	0.374	0.334
Provincial FEs	No	No	No	No	Yes	Yes	Yes	Yes

Notes: The unit of analysis is the county. The sample is a cross-section of counties in 2020. Panels A and B use distance to provincial capital (km) and terrain ruggedness (average slope) as dependent variable, respectively. The data are obtained from China's National Earth System Science Data Center (https://www.geodata.cn/). Columns 1 and 5 use never designated counties and NPCs removed in 2011; columns 2 and 6 use never designated counties and NPCs removed in 2016-2017; columns 3 and 7 use never designated counties and NPCs removed in 2018-2020; and columns 4 and 8 use never designated counties and NPCs removed in 2011 and 2016-2020. "Removed in 2011" is a dummy variable that is one if the county was removed from the NPC list in 2011 and zero otherwise. The other two regressors are defined in the same way. The omitted group is the never-designated counties. A constant is included in the regressions but not reported. Robust standard errors in parentheses are clustered by province. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

B Measuring GDP Manipulation

B.1 The CNRDS Data

We directly use the digital number (DN) data, which measures the luminosity of night lights, from the Global Night-time Light Database (GNLD) provided by the Chinese Research Data Services Platform (CNRDS). A document describing how the CNRDS processes the data is available on the CNRDS website (see https://www.cnrds.com/). Here is a brief description of the document.

The DMSP/OLS Data for the 2003-2013 Period. The CNRDS constructs the night light data for the period 2003-2013 as follows. The CNRDS first obtains the image data from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP/OLS), which is accessible through the National Oceanic and Atmospheric Administration (NOAA) portal (see http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html). The relevant DMSP/OLS image data are collected by a constellation of six satellites, namely F10, F12, F14, F15, F16, and F18 (the data used in our paper is obtained from the latter four). A "stable light" series is then constructed, which is suitable for us to construct the GDP manipulation measure due to its refinement process that mitigates transient light disturbances such as those caused by solar and lunar illumination, atmospheric conditions, and transient anomalies such as fires or explosions.

However, the data are still subject to some limitations, such as saturation (top-coding) and inconsistencies in temporal and cross-satellite comparability. While saturation is a relatively minor concern in the Chinese context and also in our empirical context, the data discrepancies across time and satellite need to be addressed. The CNRDS employs a three-step correction approach to ensure comparability of the night lights along four different dimensions: inter-satellite DN values within the same year, intra-satellite DN values across different years, inter-satellite number of lit pixels within the same year, and intra-satellite number of lit pixels across different years. The three steps include inter-calibration, intra-annual and multi-sensor correction, and inter-annual and multi-sensor correction. See the CNRDS website for more details.

The VIIRS/DNB Data for the 2013-2020 Period. The CNRDS constructs the night lights data for the period 2013-2020 as follows. The CNRDS first obtains the image data from the Visible Infrared Imaging Radiometer Suite's Day/Night Band (VIIRS/DNB), which is accessible via the NOAA portal (see https://www.ngdc.noaa.gov/eog/viirs/download_dnb_composites.html). While the data available on the NOAA platform have been stripped of atmospheric interferences such as cloud cover, lightning, and lunar illumination, they retain anomalies due to ephemeral luminosity events such as fires and maritime navigation beacons, as well as persistent background noise. These anomalies manifest themselves as statistical outliers, negative values, and pervasive background noise within the original dataset. This is corrected by the CNRDS. See the CNRDS website for more details.

B.2 Alternative Data Sources

To check the robustness of our baseline empirical results, we also use two alternative sources of nighttime light data: *(i)* the global 500-meter resolution "NPP-VIIRS like" night light dataset (2000-2022) and *(ii)* China's "DMSP-OLS like" 1-km night light remote sensing dataset (1992-2022). Both can be downloaded from the National Earth System Science Data Center of the National Science and Technology Infrastructure of China (see http://www.geodata.cn/). Both provide the DN values at the 30 arc-second grid cell level. We average the DN values at the grid-cell level to the county level using ArcGIS. We then calculate the GDP manipulation variable using Equation (1). Finally, we perform regression analyses.

The estimated results of the effects of NPC designation and cancellation are reported in the Appendix Tables B1 and B2, respectively. Overall, the results are largely similar to our baseline estimates. Appendix Table B1 shows that the national poor counties designated in 2011 are associated with GDP underreporting in 2010; and Appendix Table B2 shows that the national poor counties removed from the poor status since 2018 are associated with GDP overreporting.

B.3 Henderson et al. (2012)

We now re-estimate our baseline results by constructing the GDP manipulation measure using the statistical framework proposed by Henderson et al. (2012). Following Henderson et al. (2012), we first use lights growth and official GDP growth to obtain an improved estimate of true GDP growth, and then measure GDP manipulation as the difference between official GDP growth and the improved GDP growth. We assume that the improved GDP growth can be written as the weighted average of official GDP growth and the fitted value of official GDP growth based on night lights growth:

$$ImprovedGDP = \lambda \cdot GDP + (1 - \lambda) \cdot \widehat{GDP}, \tag{B1}$$

where GDP is the growth rate of official GDP per capita, $\widehat{GDP}$ is the fitted value of official GDP growth. λ is the weight of GDP growth in the calculation of improved GDP growth. Specifically, λ is the optimal weight that minimizes the variance of improved GDP growth relative to true GDP growth:

$$\lambda = arg\ \min\{var(ImprovedGDP - TrueGDP)\}, \tag{B2}$$

where $TrueGDP$ is the true GDP growth that is unknown.

Using the relationship between official GDP growth, true GDP growth, and night lights growth defined in Henderson et al. (2012), we can solve for the optimal weight λ^*:

$$\lambda^* = \frac{\phi var(GDP)var(Light) - cov(GDP - Light)^2}{var(GDP)var(Light) - cov(GDP - Light)^2} = \frac{\phi - \rho^2}{1 - \rho^2}, \tag{B3}$$

where $Light$ is the night lights growth, $\phi \equiv var(TrueGDP)/var(GDP)$, and ρ is the correlation between official GDP growth and night lights growth. The values of ϕ reflect the reliability of the GDP data. To find the specific values of ϕ, we need to divide the counties into two groups based on the level of official GDP quality. According to Chen et al. (2019), we choose the counties in Guangdong and Zhejiang provinces as having high GDP quality and the rest of counties as having low GDP quality. By assuming a specific value of ϕ_h for the counties with good GDP quality (e.g., 0.9), we can derive the value of ϕ_l for the counties with low GDP quality ($\phi_l = \phi_h var(GDP_h)/var(GDP_l)$). After obtaining the values of ϕ, we can calculate the optimal weight λ^*.

In this framework, the GDP manipulation of county c in year t can be naturally written as the difference between official GDP growth and the improved GDP growth:

$$GDPManipulation_{c,t} = GDP_{c,t} - ImprovedGDP_{c,t} = (1-\lambda)(GDP_{c,t} - \widehat{GDP}_{c,t}), \tag{B4}$$

where $\widehat{GDP}_{c,t}$ is the fitted value of official GDP growth of county c and year t, which is obtained by running the following regression:

$$GDP_{c,t} = \alpha + \beta \cdot Light_{c,t} + \mu_c + \delta_{p,t} + e_{c,t}, \tag{B5}$$

where μ_c are the county fixed effects, and $\delta_{p,t}$ are province-year fixed effects. To get a more precise $\widehat{GDP}_{c,t}$, we follow Xu et al. (2022) to run this regression separately for the counties with high GDP quality and those with low GDP quality, taking into account potential heterogeneity across regions. In addition, we also consider different values of ϕ_h (e.g., 0.9 and 0.6) in our robustness checks.

Using the above procedure, we construct a new measure of GDP manipulation and then perform regression analyses using our baseline empirical strategies, replacing the outcome variable with the new measure. The estimated results of the effects of NPC designation and cancellation are reported in the Appendix Tables B3 and B4, respectively. Overall, the results are qualitatively very similar to our baseline findings. The magnitudes are different because the outcome in the regressions is measured as the difference between two growth variables (i.e., official GDP growth and improved GDP growth), while the outcome in our baseline regressions is measured as the log ratio of two level variables (i.e., GDP per capita and average DN value).

Table B1: NPC Designation and Underreporting: Alternative Data Sources

	GDP manipulation			
	"DMSP-OLS like" 1-km night light		"NPP-VIIRS like" 500-meter night light	
	2003-2010	2008-2010	2003-2010	2008-2010
	(1)	(2)	(3)	(4)
NPC × 2010	-0.107***	-0.083***	-0.042*	-0.048**
	(0.020)	(0.015)	(0.025)	(0.021)
Num. obs.	12526	5392	12526	5392
Num. clu.	1569	1800	1569	1800
Adj. R-sq.	0.977	0.986	0.897	0.928
County FEs	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes
2003 controls × year FEs	Yes	No	Yes	No
2008 controls × year FEs	No	Yes	No	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2003-2010 and 2008-2010 for columns 1 and 3 and 2 and 4 respectively, and counties including national poor counties designed in 2011 and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2003 for columns 1 and 3 and 2008 for columns 2 and 4). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table B2: NPC Cancellation and Overreporting: Alternative Data Sources

	GDP manipulation									
	"DMSP-OLS like" 1-km night light					"NPP-VIIRS like" 500-meter night light				
	2013-2017	2016-2020				2013-2017	2016-2020			
	Removed 16-17	Removed 18-20				Removed 16-17	Removed 18-20			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NPC removed	-0.034	0.047***				-0.006	0.090***			
	(0.030)	(0.015)				(0.032)	(0.015)			
NPC removed 2018 $\times \geq$2018			0.048**					0.089***		
			(0.023)					(0.025)		
NPC removed 2019 $\times \geq$2019				0.064***					0.131***	
				(0.024)					(0.025)	
NPC removed 2020 $\times$ =2020					-0.016					0.081
					(0.060)					(0.050)
Num. obs.	5729	7101	5590	5852	4847	5729	7101	5590	5852	4847
Num. clu.	1148	1440	1138	1194	996	1148	1440	1138	1194	996
Adj. R-sq.	0.979	0.980	0.982	0.980	0.982	0.973	0.976	0.978	0.977	0.979
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2013 controls $\times$ year FEs	Yes	No	No	No	No	Yes	No	No	No	No
2016 controls $\times$ year FEs	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2013-2017 and 2016-2020 for columns 1 and 6 and for columns 2-5 and 7-10 respectively, and counties including national poor counties removed since 2016 (those removed in 2016-2017 for column 1 and 6 and those removed in 2018-2020 for columns 2-5 and 7-10) and counties never designed. The dependent variable (GDP manipulation) is the log ratio of a county's reported GDP per capita to its nighttime light intensity, measured as average digital number. The variable NPC Removed equals one if a county has been removed from poor status in a given year and zero otherwise. We disaggregate the counties removed in 2018-2020 into different single years in columns 3-5 and 8-10 (corresponding to 2018, 2019, and 2020, respectively). Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2013 for columns 1 and 6 and 2016 for columns 2-5 and 7-10). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table B3: NPC Designation and Underreporting: Henderson et al. (2012)

	GDP manipulation			
	$\phi = 0.9$		$\phi = 0.6$	
	2003-2010	2008-2010	2003-2010	2008-2010
	(1)	(2)	(3)	(4)
NPC × 2010	-1.329**	-1.417*	-1.473**	-1.610*
	(0.618)	(0.756)	(0.716)	(0.875)
Num. obs.	11052	4346	11052	4346
Num. clu.	1555	1584	1555	1584
Adj. R-sq.	0.277	0.220	0.269	0.208
County FEs	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes
2003 controls × year FEs	Yes	No	Yes	No
2008 controls × year FEs	No	Yes	No	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2003-2010 and 2008-2010 for columns 1 and 3 and 2 and 4 respectively, and counties including national poor counties designed in 2011 and counties never designed. The dependent variable (GDP manipulation) is the difference between official GDP growth and the improved GDP growth. NPC equals to one if the county was selected as a national poor county in 2011 and zero if the county was never selected. Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2003 for columns 1 and 3 and 2008 for columns 2 and 4). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

Table B4: NPC Cancellation and Overreporting: Henderson et al. (2012)

	GDP manipulation									
	$\phi = 0.9$					$\phi = 0.6$				
	2013-2017	2016-2020				2013-2017	2016-2020			
	Removed 16-17	Removed 18-20				Removed 16-17	Removed 18-20			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NPC removed	1.106	2.387***				1.192	2.561***			
	(1.893)	(0.602)				(2.034)	(0.646)			
NPC removed 2018 $\times \geq$2018			2.284**					2.438**		
			(0.936)					(1.006)		
NPC removed 2019 $\times \geq$2019				3.437***					3.684***	
				(0.824)					(0.886)	
NPC removed 2020 $\times$ =2020					1.022					1.187
					(1.650)					(1.777)
Num. obs.	2995	6329	4995	5245	4349	2995	6329	4995	5245	4349
Num. clu.	1077	1432	1129	1189	990	1077	1432	1129	1189	990
Adj. R-sq.	-0.050	0.284	0.278	0.277	0.292	-0.052	0.283	0.277	0.276	0.290
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2013 controls $\times$ year FEs	Yes	No	No	No	No	Yes	No	No	No	No
2016 controls $\times$ year FEs	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes

Notes: The unit of analysis is the county-year. The sample is a county-year panel, with the sample period spanning 2013-2017 and 2016-2020 for columns 1 and 6 and for columns 2-5 and 7-10 respectively, and counties including national poor counties removed since 2016 (those removed in 2016-2017 for column 1 and 6 and those removed in 2018-2020 for columns 2-5 and 7-10) and counties never designed. The dependent variable (GDP manipulation) is the difference between official GDP growth and the improved GDP growth. The variable NPC Removed equals one if a county has been removed from poor status in a given year and zero otherwise. We disaggregate the counties removed in 2018-2020 into different single years in columns 3-5 and 8-10 (corresponding to 2018, 2019, and 2020, respectively). Control variables include nighttime light intensity, household savings per capita, the ratio of primary sector value added to GDP, and the ratio of secondary sector value added to GDP, all measured at the county level and in the base year (2013 for columns 1 and 6 and 2016 for columns 2-5 and 7-10). Robust standard errors in parentheses are clustered by county. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.