In the Name of Love: Policy Imitation, Political Incentives, and Local Economic Growth§

Hai Hong[1], Yunfei Zhang[1], Kevin Chen[1,2*]

(1. China Academy for Rural Development, School of Public Affairs, Zhejiang University, Hangzhou, China; 2, International Food Policy Research Institute, Beijing, China)

Abstract: Policy experimentations are widely accepted instruments for promoting economic growth in many countries. However, the effects may be undermined if they're masked by top-down enforcement that ignores local conditions. Using satellite night light data, we investigate the economic impacts of the "Taobao Villages (TBVs)" program in China, a bottom-up pilot on rural e-commerce that is imitated nationwide under top-down directives. Exploiting the difference-in-differences approach, we show that the establishment of TBVs significantly *reduces* townships' night light and therefore local economic growth. We further reveal that the top-down directives in TBVs' establishment and the strategic behaviors of local officials motivated by political incentives are the potential driving forces.

Keywords: Policy experimentation, political incentives, nighttime lights, Taobao village

1. Introduction

The formulation and implementation of successful policies constitute a central issue driving economic growth (North 1990). The Chinese government, amidst its economic takeoff, has undertaken extensive policy experimentations, which have greatly propelled rapid economic growth (Rawski 1995; Cao, Qian, and Weingast 1999; Lin, Cai, and Li 2003; Xu 2011). Early policy experiments, such as the agricultural decollectivization, often adopted a bottom-up approach, in which the central government feeds the successful experiences of local pilots into national policy formation (Heilmann 2008; Xu 2011). As the state capacity strengthened, more policies, such as environmental regulations and anti-poverty campaigns, were often implemented in top-down manners, with the salience of and emphasis on central design and guidance (Chen 2017; Martinez-Bravo et al. 2022).

Although these enormous policy experimentations have played a vital role in China's economic transition and subsequent development, recent literature begins to underscore the limited and potentially biased information in the central government's policy learning, coupled with the strategic behaviors of local officials, can undermine the true policy effect (Wang and Yang 2021). Specifically, the success of locally spontaneous pilots may lead the central government to replicate it nationwide through direct interventions. But constrained by the information asymmetry (Oates 1972; 1999), when rolled out at the national level, these pilots may yield highly differentiated results due to heterogeneity across places and in the ways that policies are implemented (Allcott 2015; Vivalt 2020). Further, in a regionally decentralized system characterized by horizontal competition (Xu 2011), local officials have strong incentives to initiate and imitate policies favored by their superiors (Li and Zhou 2005; Landry 2008; Xu 2011), but likely ignore how these policies fit into

§ We are thankful to the support from Zhejiang University-International Food Policy Research Institute Center for International Development Study, the National Social Science Fund of China (No. 21&ZD091, No. 23ZDA035), the Humanities and Social Science Fund of the Ministry of Education of China (22JJD790076), and the Agricultural Science and Technology (10-IAED-05-2024; 10-IAED-RY-01-2024).

* Corresponding author: Kevin Chen; E-mail: kzchen@zju.edu.cn.

local economic conditions. Such strategic behaviors induced by political incentives may result in allocative inefficiency and undermine economic development. While extensive efforts have been made to understand the effect of either top-down or bottom-up policy experiments, yet still little is known about how the effect varies when the local information is ignored in the policy formation process and the bottom-up pilot evolves into top-down directives.

This paper seeks to understand the effects and consequences of a bottom-up pilot when it is subject to top-down directives that are imitated nationwide. We focus on the rural e-commerce programs adopted by the Chinese government to achieve rural revitalization and poverty elimination. Specifically, we study how the establishment of Taobao Villages (TBVs), the major form of rural e-commerce valued by the government, affects local economic growth.[1] Two features of TBVs make them suitable for our study. First, TBVs incorporate both bottom-up and top-down formation mechanisms,[2] which allows a discussion of the heterogeneous impacts of the two mechanisms within the same context. While both formation mechanisms play an important role in the establishment of TBVs, the logic behind each mechanism differs. In the top-down formation process, local governments tend to select villages with better prior economic foundations as policy pilots, since the probability of success is higher (Allcott 2015; Wang and Yang 2021). However, in the bottom-up formation process, some anecdotal evidence suggests that those TBVs that achieved breakthroughs tended to come from remote mountainous areas and villages with poorer economic infrastructure, as the gains from adopting new technologies are greater.[3] The two different mechanisms may result in substantial heterogeneity in policy effects.

Second, when rolled out nationally, the establishment of TBVs gradually becomes part of political competition. In response to the central government's advocates, local officials have strong incentives to engage in quantitative competitions while exerting less effort in quality improvements (see more discussion in the Background section below), since the former is more observed by their superiors (Holmstrom and Milgrom 1991; Dewatripont, Jewitt, and Tirole 1999; 2000; Alesina and Tabellini 2007; 2008). Further, under top-down enforcement, the pilots are likely to be inconsistent with the local conditions due to the information asymmetry, which may lead to substantial distortions to the local economy when the government takes direct interventions.[4]

Combining satellite night light data with detailed township-level geographic information, this paper exploits a difference-in-differences approach to firstly quantify the effects of TBV formation on local economic growth from a national scale. To account for the non-randomness of TBV establishment, we estimate an augmented two-way fixed effects model that incorporates township fixed effects, county-by-year fixed effects, and treatment-specific time trends. The inclusion of these fixed effects and time trends relaxes our identification assumption and helps to net out a wide range of time-varying unobservables. In addition, due

[1] As the name suggested, TBVs are villages that engage in online sales with certain requirement on sale scale, only villages meet the specific standards set by Alibaba (the largest e-commerce platform in China) can be certified as TBVs (see more discussion in Background section).

[2] The first one is formed bottom-up by local entrepreneurs; the second one is formed top-down, driven by local governments. Although most studies emphasize the importance of entrepreneurship in the formation of Taobao villages (Liu et al. 2020), in practice a substantial amount of TBVs are still driven by local governments (Couture et al. 2021).

[3] For example, Dongfeng Village, one of the earliest Taobao villages to be established (also heavily studied as a case study in the literature; see, e.g., (Zang et al. 2023) and the discussion in the Background section), where villagers relied on selling garbage before it establishing TBV; and Xiaying Village, the first TBV in Hubei Province, had more than half of the villagers below the poverty line set by the government, prior to the establishment of the TBV. See https://m.chinanews.com/wap/detail/zw/business/2021/02-26/9419854.shtml.

[4] Additionally, the purpose of government investment (which could be, for example, employment maximization) is often different from that of entrepreneurs who are motivated by profit maximization (Colonnelli, Li, and Liu 2024), thus may undermine the efficiency and lead to resource misallocations.

to the potential selection bias, the treatment and control group may have divergent trends even in the absence of TBV establishment. By conditioning on the above fixed effects and time trends, we exploit only within county-year variations that partialling out treatment-specific trends.

Given the above fixed effects and time trends, we find that the establishment of TBVs significantly *reduces* township-level nighttime lights. In our preferred specification and based on the elasticity estimates by Gibson et al. (2021) , the establishment of TBV reduces townships' night light by about 3%, which corresponds to a reduction in GDP of between 2.4% and 3.3% . Given the context of overall increasing night light intensity in China, this change is driven by slowing growth in TBVs. We provide several pieces of evidence in favor of the identification assumption underlying our difference-in-differences estimation. First, we exploit an event study method to detect the presence of pre-trend. Conditional on our fixed effects and time trends, we find that prior to the establishment of TBV, the disparity of night light intensity across treatment and control groups is statistically insignificant. Second, our results remain robust when we restrict control groups to some small radii of the treatment group, which largely precludes the existence of spillovers. Third, to correct for the potential biases caused by the heterogeneous treatment effects in settings with staggered treatment rollouts, we use the novel DiD estimators proposed by the literature (de Chaisemartin and D'Haultfœuille 2020; Callaway and Sant'Anna 2021; Sun and Abraham 2021) and find that negative weighting issue caused by staggered adoption is less of a concern in our setting. Additional exercises from testing for measurement error and cluster level adjustment also suggest that our results are highly robust.

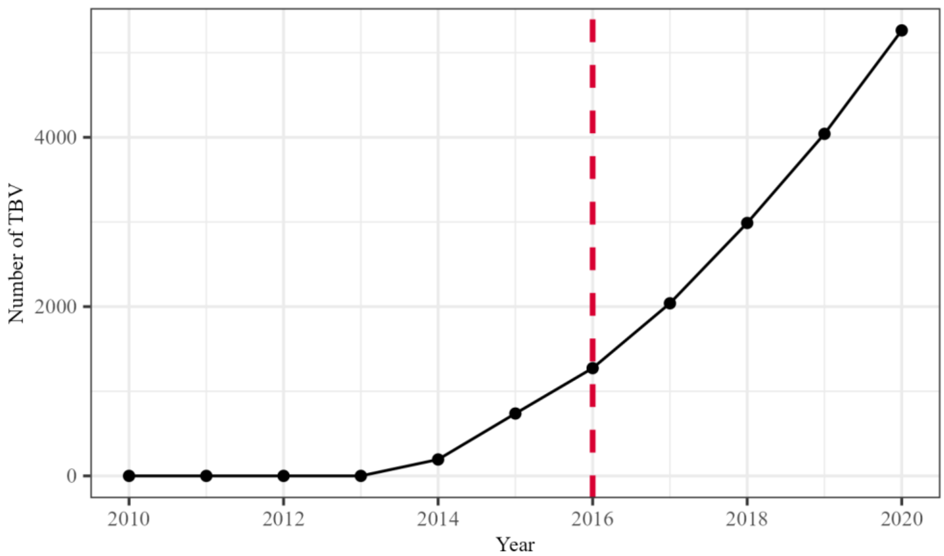


Figure 1. Yearly Number of TBVs from 2010-2020

We then discuss the potential mechanisms that drive our findings. We argue that the top-down directives and the strategic behaviors of local officials are the main driving forces. We first test whether top-down policy advocation is capable of explaining our baseline estimates. By analyzing relevant policy documents leased by the central government, we note an important turning point. At the end of 2015, the central government for the first time announced that the development of rural e-commerce should be an important means of poverty alleviation and rural revitalization. The number of TBVs started to rise rapidly since then (see Figure 1), and appeared more frequently in local government reports (see Figure 2 and the discussion in the Background section). We hypothesize that TBVs established before 2016 are more likely to be formed through bottom-up approaches, while TBVs established after 2016 are more likely to be formed through top-down manners. By estimating the treatment effect for each separate treatment cohort, we find only TBVs established before 2016 bring positive effects on townships' economic growth. In contrast, TBVs established after 2016 bring persistent negative effects to townships. This is in line with our predictions.[5]

Next, we explore the role of local governments in top-down directives. To this end, we collect the annual prefectural government reports and leverage the frequency of words related to rural e-commerce (and rural revitalization) as a proxy for the importance local governments attach to the establishment of TBVs. We then estimate an augmented version of our baseline model that incorporates the keyword frequency as an interaction term. Our results yield consistent conclusion that the negative effects are more pronounced in prefectures that attach more importance to the establishment of TBVs and the development of rural e-commerce.

In a similar vein, we find that the negative impacts of TBV establishment are more evident in counties with incumbent county officials who have been in office for longer periods, and in counties with county leader rotations, with both indicators reflecting the promotion incentives of local officials (Chen and Zhang 2021; Fang, Liu, and Zhou 2023). In China's bureaucratic selection process, economic performance most times outweighs other performance indicators in the promotion decisions made by upper-level officials (Li and Zhou 2005; Xu 2011). While devoting resources to the establishment of TBVs is conducive to serving as a vanity project, promoting the further development of TBVs contributes little to the local economy, as the growth of urban sectors is more decisive to the aggregate economy (Wang, Zhang, and Zhou 2020). Therefore, county officials with higher promotion incentives are more likely to reduce the investment in TBVs once they have been established, and lead to further distortions in the township's economy.

We conduct a series of heterogeneous analyses that help reveal the strategic behaviors of local governments. We hypothesize that, local governments have greater incentives to exploit top-down strategies to establish TBVs when the cost of investment is lower or when they have more resources at hand. Alternatively, they have greater incentives to resort to top-down approaches when townships have better economic or natural conditions. Our results provide consistent evidence. Typically, we find that the negative effects are concentrated in townships that are closer to the county government offices, as the cost of vertical control is lower (Martinez-Bravo et al. 2022), in counties with higher GDP per capita and lower fiscal pressure (likely to have more resources devoted to the establishment of TBVs), and in townships with higher initial light intensity, population, road density, and with lower geographic ruggedness, which measure townships' economic and natural conditions. Overall, the above evidence further reveals how the choices of local government led to the negative impact of TBV establishment on the local economy.

We assume the strategic behaviors of local officials on the establishment of TBVs and the resulting distortions are the main mechanisms that explain the above findings. Specifically, when the central government launches top-down advocates to develop TBVs, the principle-agent problem that local officials face incentivizes them to devote limited resources to the launch of TBVs, while disincentivizing them to allocate the resources to sustain TBVs' operation. Consistent with such prediction, we find that local governments invest more in road construction before townships establish TBVs, while the investment decreases soon after the TBVs are established. In addition, we note that such a pattern only exists in samples that establish TBVs after 2016, when top-down directives became the major formation mechanism.

As a complementary corroboration, we further examine the change in the number of firms entering before and after the establishment of TBVs. We find similar evidence to that of local government infrastructure investment, that is, the number of private firm entries increases significantly before the launch of TBVs while decreasing significantly soon after the launch. Given that the decrease in government investment may have led to a decrease in the attractiveness of firms, which further led to a depression in the local economy. This is again consistent with our narrative and the above evidence.

[5] Since the number of TBVs established after 2016 is far more than those established before 2016, the aggregate effects of TBV establishment estimated by our baseline model are still negative.

To summarize, our findings highlight the potential limitations of policy experimentations in China and reveal for the first time the substantial negative impacts of TBVs on local economies when taking local officials' strategic behaviors into account. While decentralization and horizontal political competition are conducive to promoting economic growth, they fail to accommodate the potential distortions led by local officials' political incentives and may result in unprecedented outcomes that end up hindering local economic development. It is important to note that while we document the negative effects of TBV formation on local economic growth under specific policy conditions, we are not denying the potential positive role that TBVs can play in promoting employment and alleviating poverty. As we demonstrate in the subsequent analyses, the formation of TBVs can promote local economic growth under certain conditions. What we have highlighted is the inappropriate institutional arrangements that can lead to resource misallocation, ultimately resulting in the negative effects on local economic growth we observe.

This paper contributes to several strands of literature. First, we study how the loss of local information in policy formation can have negative effects on the local economy, which is in line with Hayek's idea of the fundamental importance of local information (Huang et al. 2017). Our empirical investigation on the differential effects of TBV establishment shed new light on the potential distortions led by top-down interventions as well as the strategic behaviors of local officials (Scott 1998; Kung and Chen 2011; Suárez Serrato, Wang, and Zhang 2019; Wang and Yang 2021). Differing from the existing literature that focuses on either top-down or bottom-up pilots, we leverage the unique setting of TBVs that incorporate both top-down and bottom-up formation mechanisms, to explore how the loss of local information can substantially affect the effectiveness of policy experiments.

Second, our study provides an original political economy story about the impact of TBV development on local economies. Existing studies on TBVs have mainly focused on the formation mechanism (Qi, Zheng, and Guo 2019; Liu et al. 2020; Zang et al. 2023), while some analyze the impact of TBV formation on the local economy using either survey data or case studies (Luo and Niu 2019; Tang and Zhu 2020; Wang et al. 2021). To the best of our knowledge, this paper is the first to discuss the strategic behavior of local governments and the economic consequences in the studies of TBVs. By utilizing detailed township level data as well as comprehensive TBV establishment data, this paper is also the first to systematically test the causal impact of TBV establishment on local economic growth at the national level therefore enriching the literature related to the research on Taobao villages (Couture et al. 2021; Huang et al. 2021). Typically, we also differ from the existing literature that studies the effects of TBVs (e.g., Couture et al. 2021) in that we provide a comprehensive evaluation from a nationwide scale. The existing studies often select a subset of TBVs to evaluate its impact. Although the internal validity is guaranteed, the external validity may be weakened. In contrast, we evaluate the effects of TBVs by considering all pilots andoffer new insights from a more aggregate perspective.

Third, we add to the literature on the political economy of policy experimentation (Wang and Yang 2021). Our contribution is twofold. On the one hand, we echo a large yet growing literature related to officials' political incentives and their economic and social impacts (Qian and Weingast 1997; Li and Zhou 2005; Xi, Yao, and Zhang 2018; Li et al. 2019; Suárez Serrato, Wang, and Zhang 2019). We expand on this line of literature by linking local officials' career incentives to the formation and development of TBVs and highlight the distortionary effects in the play. On the other hand, our study also speaks to the site selection bias and heterogeneity of treatment effects in policy and program evaluation (Allcott 2015; Vivalt 2020). Research from the literature suggests that selection bias in program evaluation may overstate the true impacts of a program, thereby undermining the external validity of a program or policy (Allcott 2015). By carefully scrutinizing the TBV program, we not only confirm the conclusions of the above studies, but further suggest that such selection bias and differences in the subjects of program implementation may potentially lead to unprecedented negative effects, ultimately defeating the original purposes of the policy.

The remainder of the paper is organized as follows. Section 2 introduces the policy context in which TBVs were established and highlights the role played by local governments, with a brief summarization of the related literature on TBVs and a simple framework to guide our empirical investigations. Section 3 describes the data and main variables. Section 4 presents our empirical strategy, the primary results of the baseline regression, and the corresponding robustness checks. Section 5 considers the role of local officials and discusses the potential mechanisms. Finally, Section 6 concludes.

2. Background, Related Literature and Theoretical Framework

2.1 Bottom up: Early TBV formations and the main characteristics

Taobao Village is an e-commerce village with Taobao as the sale platform, where its name is derived from.[6] According to the definition, a village is recognized by Alibaba as a TBV when it meets the following conditions: annual e-commerce sales exceed 10 million yuan (approximately 1.6 million dollars), the number of active online stores in the village exceeds 100, or the number of active online stores reaches above 10% of the number of households.[7] In addition, if there are more than 3 villages within the same township developed TBVs, the township would be certified as Taobao Township.[8] Besides the above criteria, local governments also engage in TBVs' establishment and application process. Specifically, the designation of TBV does not entirely depend on the above prerequisites, Alibaba would also reference information from other channels, such as local government, media, or scholars to decide whether to certify a village as TBV.[9]

Villages identified as TBVs will receive support from Alibaba in two aspects. First, credit support. Villages designated as TBVs will be given to high-quality sellers, and a package of financial support plans will be provided for TBVs in terms of seller operations and consumption. Second, training and promotion. Villages identified as TBVs will receive professional training from Alibaba, and will be promoted to get higher exposure.

The earliest TBV dates back to 2012 and has gone through a major increase since 2014, at which time Alibaba started to formally document the establishment of TBVs with the above standards. In the early stages, TBVs were characterized by bottom-up characteristics. Typically, these villages are established by entrepreneurs and start-ups who set up local e-commerce companies and use Taobao as their platform to sell local products across the country (Qi, Zheng, and Guo 2019; Wang et al. 2021). The success of a few entrepreneurs leads other villagers to engage in similar ventures due to the "salience effect" in entrepreneurial activities (Y. Huang et al. 2021). A concrete example is Dongfeng Village in Chao County, Jiangsu Province, studied by Zang et al. (2023). After entrepreneur Sun Han's success in selling furniture using the Taobao platform, more villagers followed and registered Taobao stores to engage in e-commerce activities. As the practitioners continued to grow, Dongfeng became one of the first three TBVs in the country.

[6] It should be noted that Taobao Village is not the only form for developing rural e-commerce. Around 2015, in order to compete for the rural market, other e-commerce giants besides Taobao, such as Jingdong and Pinduoduo, were also actively developing rural e-commerce. However, due to Taobao's market size advantage, the competition for rural consumers and online retailers between different e-commerce platforms stood to end in Taobao's victory. In fact, Taobao Village is currently the only successful rural e-commerce that has been studied by academics, and the two concepts between Taobao Village and rural e-commerce are almost equivalent in the local government's annual report. Therefore, in the following, we will use Taobao Village and rural e-commerce interchangeably, as there is no substantive difference between the two in practice.

[7] The certification adopts a rolling mechanism that evaluates the eligibility of TBVs on a yearly base. Villages that fail to meet the standards will be excluded from the list. However, exit cases are rare.

[8] Alternatively, a township would be designated as a Taobao Township if its annual e-commerce sales exceed 30 million yuan, the number of active online stores in the village exceeds 300, or the number of active online stores reaches above 30% of the number of households.

[9] See https://xueqiu.com/1527849020/179758561.

In the early development of TBVs, Taobao's parent company, Alibaba, and local governments played a critical role. For example, Alibaba launched its "Thousand Counties and Ten Thousand Villages Plan" in 2014, stating that it would invest 10 billion yuan over the next three to five years to establish 1,000 county-level operation centers and 100,000 village-level service stations that facilitate rural e-commerce activities and would work with the government to establish a comprehensive rural e-commerce system. At the same time, local governments, realizing the potential benefits of e-commerce for rural economic development, have also begun to actively intervene, providing a series of support policies for rural entrepreneurs to promote the development of rural e-commerce.[10] In addition, local governments have invested in road construction and the establishment of industrial parks to further promote the expansion of TBVs (Zang et al. 2023).

2.2 Top down: The role of central and local governments

The early success of TBVs began to gradually attract the attention of the central government. For example, at the Davos Forum in 2015, the then-Premier Keqiang Li expressed his approval of the TBV development paradigm in his speech.[11] At the end of 2015, the central government issued a document for the first time pointing out that the development of rural e-commerce should be taken as an important and major means of poverty alleviation and rural revitalization.[12] Specifically, the document emphasized the critical need to expedite the integration of rural e-commerce with the agricultural industry. It highlighted the pivotal role of rural e-commerce in fostering entrepreneurship and employment opportunities among villagers, stimulating the growth of the rural consumer market, and ultimately driving the elimination of rural poverty. The document also puts forward several policy measures, including strengthening the construction of rural infrastructure (e.g., rural roads and broadband). In addition, the document underscores promoting innovation and entrepreneurship, guiding groups like rural migrant workers and youth to engage in entrepreneurial and e-commerce activities.

Local governments responded immediately. To measure the response of local governments, we collect the annual reports of prefecture governments from 2010 to 2020, and obtain keywords related to rural e-commerce as well as rural revitalization. The mention and frequency of these specific keywords reflect the importance local governments attached to these issues.[13] Figure 2 plots the yearly trend of these keywords mentioned in government reports. We find that after the central government announced the vigorous development of rural e-commerce at the end of 2015, the word frequencies related to rural e-commerce that appeared in the subsequent government reports increased significantly. At the same time, the frequency of words related to rural revitalization, as the ultimate purpose of developing rural e-commerce, has also increased over time.

Under the top-down formation mechanism, the establishment of TBVs aligns with the process of policy experiments in several aspects. First, local governments have the power to decide where to set up TBVs and provide specific subsidies to facilitate TBVs' establishment. For example, Cao county, a leading county of rural e-commerce in China, has set up specialized administrative agencies to be responsible for the planning and organization of rural e-commerce activities. The county government deeply engages in planning and selecting TBVs, and has established 168 TBVs within a couple of years.[14] Second, unlike the bottom up formation mechanism that requires the diffusion of rural e-commerce entrepreneurship (which is often time-

---

[10] This policies include enabling rural e-commerce entrepreneurs to enjoy government subsidies and preferential treatment in the registration of online stores, access to start-up loans, and housing leases, among other things.

[11] See https://www.gov.cn/zhuanti/2015-01/23/content_2808977.htm.

[12] See https://www.gov.cn/zhengce/content/2015-11/09/content_10279.htm.

[13] All the specific keywords are listed in Appendix Table B1.

[14] See report from https://www.thepaper.cn/newsDetail_forward_17293661.

consuming), the top-down formation mechanism can establish a TBV within a couple of months of intensive training and resource support. For instance, local governments often couple with professional service providers to train the villagers and help to set up online stores to meet the criterion of TBV designation. The process can be done within six months only.[15] Given these characteristics, we prefer to view the TBV designation as a means of policy experimentation.

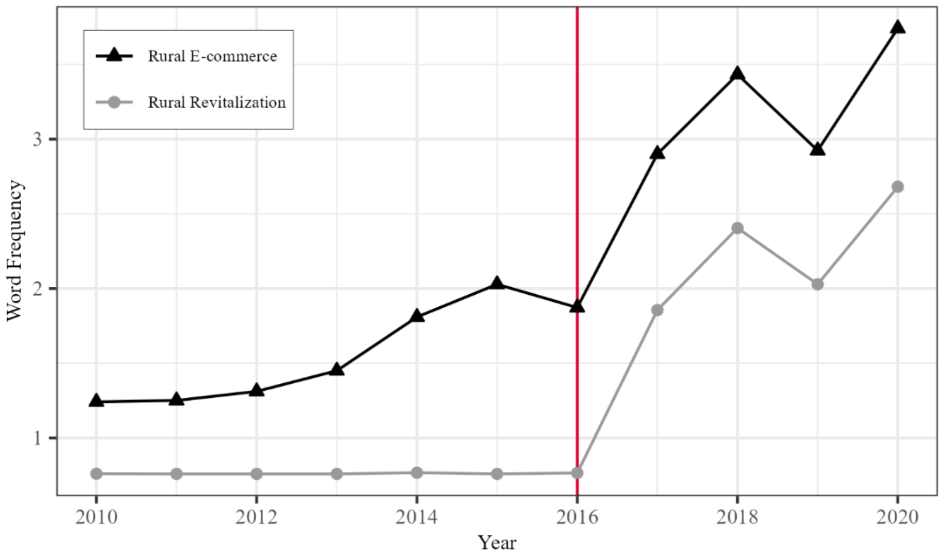


Figure 2. Keyword frequencies in local government annual reports

Specific textual information in government reports also provides qualitative evidence of the existence of strong incentives for local governments to exploit top-down interventions to establish TBVs. A typical example is the government of Heze City, Shandong Province. Due to its heavy reliance on the e-commerce industry, the Heze government includes the goal of "maintaining the leading position of Taobao villages and Taobao townships in the country" in its annual reports and has required prefectures, counties, and districts to include the development of TBVs into their performance appraisal systems. In addition, many county governments also emphasize reporting the number of Taobao villages in their annual reports. For instance, in the 2019 government report of Suining County, Jiangsu Province, the government stress the target setting for the number of TBVs to be achieved in the year. Specifically, the county government requires 58 new TBVs to be established within a single year, which equates to a 63% increase relative to the 92 TBVs that took approximately 5 years to complete. Given the limited financial resources of the county government, as well as the existence of other objectives in the appraisal system (such as promoting industrial growth and pollution reduction), the multitasking local agents may strategically exert more effort and resources in the pursuit of the quantity of TBVs. However, the formation of TBVs depends on local comparative advantages, as well as the expected revenues from engaging in e-commerce. The overemphasis on quantity in terms of the number of TBVs established may lead to selections of villages with little development advantages. Additionally, due to annual reevaluations, TBVs with poor business performance under top-down directives are likely to stay as their numbers are tied to the performance appraisal of local governments.

We also find that, in terms of the number of TBVs established, there will be a tendency for competition among local governments. For example, the government of Shantou City in Guangdong Province noted in its 2019 government report that the number of TBVs in the city ranked among the highest in the province.[16]

[15] See report from https://www.100ec.cn/detail--6487225.html.

[16] See https://www.shantou.gov.cn/cnst/zwgk/zfgzbg/content/post_1568767.html.

The government of Chao'an District, in the same province as Shantou also emphasized in its 2021 government report that the number of TBVs established in the city leads both in the country and the province.[17] Similar patterns are found in some counties' government annual reports. To name a few, in the 2019 government report of Jing County in Anhui Province,[18] the government stressed that the number of TBVs in the county ranked first in the province. Yet, while we find a disproportionate emphasis on the number of TBVs by local governments, the performance of these villages after their establishment is largely left out from the reports. Indeed, the number of TBVs is a relatively easier measure for local governments to stress, and whether the establishment of TBVs is conducive to the local economy remains questionable.

2.3 Related literature

Existing studies directly assessing the impact of TBVs are few, and there are considerable variations in their findings. Some research finds that the formation of TBVs can be beneficial to local economic development (Li and Qin 2022). In the meantime, TBVs are widely recognized in official government reports and media reports as an important means of achieving poverty elimination. They are also highly recognized by the World Bank and ADB.[19] An assessment provided by the World Bank showed that participation in TBVs can significantly improve household welfare (Luo and Niu 2019). However, the results from an RCT conducted by Couture et al. (2021) suggest that TBVs have only contributed to an increase in local consumption while having no effect in inducing firms to use e-commerce technologies, making their impact on the local economy largely small and insignificant. A few studies also extend concerns about the over-optimistic TBV development (W. Tang and Zhu 2020).

There are three main reasons for the divergent views in the literature. First, the non-randomness of TBV establishments. In either the top-down or bottom-up formation process, the decisions of individual entrepreneurs or local governments to choose where to establish a TBV are not randomized, thus making it difficult to derive reliable causal estimates from simple ex ante-post comparisons, or comparisons between treated and non-treated groups. Second, sample limitations. As of 2020, there are more than 5,000 Taobao villages across the country, and there likely exists strong heterogeneity among different villages, making it hard to generalize the results of small-scale project evaluations. Meanwhile, as noted above, there are also sheer differences in the subjects involved in project implementation (e.g., local entrepreneurs in bottom-up formation and local governments in top-down formation), both of which can lead to a large degree of heterogeneity in the results of program evaluation (Allcott 2015; Vivalt 2020). However, neither the small-scaled survey nor the RCT conducted by Couture et al. (2021) has been able to achieve national level coverage. This is in large part due to data accessibility. Existing published official statistics are only up to the county level, which usually consists of hundreds of villages, making it challenging to evaluate the economic effects of TBVs from the county level. Economic and social indicators for administrative divisions below the county level (e.g., townships or villages) are difficult to obtain, thus limiting the feasibility of assessing TBVs' effects from the national scale. Third, the formation mechanisms in the establishment of TBVs are largely not considered in the literature. While a bottom-up mechanism excels in harnessing local information and facilitating specific local conditions, top-down directives, on the other hand, are likely to fail to fully leverage local information, potentially resulting in biases and distortions in the establishment process. In the next section, we will conduct a detailed causal identification and scrutinize the economic effects of TBV establishment using the township level data.

---

[17] See http://www.chaoan.gov.cn/zwgk/qzfgzbg/content/post_3786807.html.

[18] See https://www.ahjx.gov.cn/OpennessContent/show/2007781.html.

[19] For instance, see reports https://country.cnr.cn/mantan/20181129/t20181129_524431739.shtml, and https://view.inews.qq.com/k/20200822A0JBGX00?no-redirect=1&web_channel=wap&openApp=false.

2.4 Conceptual Framework

We present a simple framework to motivate our empirical study and guide our further discussion. We first consider how different formation mechanisms affect the local economy. Due to the lack of precise local information, the probability that these TBV pilots being inconsistent with the local conditions is high in top-down formation, whereas the likelihood of mismatch is relatively low in bottom-up cases. We characterize the idea by assuming the local production function as $Y = A(p)^{\alpha}\big(1 - \tau(p)\big)^{1-\alpha}$, where p is the probability that a TBV pilot being inconsistent with the local conditions, A denote the productivity, and τ is the degree of misallocation caused by such mismatch. We assume $\partial A/\partial p < 0$ and $\partial \tau/\partial p > 0$. We can then derive that $\partial Y / \partial p < 0$, that is, the higher probability of mismatch, the lower the economy grows. Further, the above exposition also allows us to directly compare the economic outcomes between bottom-up and top-down approaches. Typically, we anticipate that the bottom-up formation approach can have positive effects on the local economy, whereas the top-down approach may turn out to have negative effects.

We then delve into modeling the strategic behaviors of local government. We consider several factors that may affect the selective establishment, that is, the importance local bureaucrats attach to the development of TBVs, local government fiscal capacity, and the establishment cost of TBVs. We also model the tradeoff between quality and quantity in local officials' political incentives. We make several assumptions for the sake of simplicity. First, we assume the local economy is primarily driven by the urban sector, while the emphasis on rural development is more focused on poverty reduction, which is less conducive to local economic growth (He, Lu, and Lee 2023).[20] Second, we assume that the local government considers two tasks that are important to their promotion, i.e., those easily be observed and measured (e.g., the number of TBVs or urban economic growth), and those hard to be observed (e.g., the quality of established TBVs). Both tasks are resource-consuming. We also simplify our analysis by assuming that the fiscal expenditure and the establishment cost are constant.[21] Consider the multitasking local government to maximize its promotion probability by considering how it relocates fiscal resources, fixed at T, over urban economic development and the development of rural e-commerce (i.e., establishing TBVs). Specifically, the objective promotion (or the utility) function of local officials is given by the following CES-CD form:

$$U = \left\{\left(\alpha v^{\frac{\sigma-1}{\sigma}} + (1-\alpha) G^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}\right\}^{\gamma} q^{1-\gamma}$$

Where v is the number of TBVs established, with each having an establishment cost constant at c. G is the resource devoted to the urban sector, which is the main driving force of economic growth. q is the quality of the established TBVs. The better the quality, the more capable the TBV program can facilitate poverty reduction. The budget constraint for local government is given by $cv + +c'q + G \leq T$. Finally, α and γ are parameters that govern the relative importance local government attached to establishing TBVs and observed tasks. The first-order conditions of the above optimization problem yield that $v^{*} = \Psi^{-1}T$, where

$$\Psi = \frac{c}{\gamma}\left[1 + \left(\frac{1-\alpha}{\alpha}\right)^{\sigma} c^{\sigma-1}\right]$$

The optimal number of TBVs is thereby a function of local government fiscal capacity T, the relative importance α, the establishment cost c, and the emphasis over quantity γ. We can also derive that the

[20] As documented in He, Lu, and Lee (2023), the recent emphasis on rural poverty reduction in China has gain growing weights in the performance appraisal system of local officials. There is thus a tradeoff between driving economic growth and reducing rural poverty, as we modelling below.

[21] Such assumptions are only made for the ease of deduction. It won't affect our main results if we allow the fiscal capacity and establishment cost to vary across counties.

quality of TBVs is determined by $q = \frac{1-\gamma}{c'}T$. By simple algebraic manipulations, we have the following proposition.

Proposition 1. *Local governments*

(a) that attach more importance to TBVs are more willing to resort top-down directives to establish TBVs, as we have

$$\frac{\partial v^*}{\partial \alpha} = \Psi^{-2}\left\{\frac{\sigma c^{\sigma-1}}{1-\gamma}\left(\frac{1-\alpha}{\alpha}\right)^{\sigma-1}\frac{1}{\alpha^2}\right\}T > 0$$

(b) *with higher fiscal capacity are more likely to establish TBVs through top-down manners as we have* $\frac{\partial v^*}{\partial T} > 0$.

(c) *are more likely to establish TBVs when the cost is relatively low, as we have*

$$\frac{\partial v^*}{\partial c} = -\Psi^{-2}\frac{1}{1-\gamma}\left\{\left[1+\left(\frac{1-\alpha}{\alpha}\right)^{\sigma}c^{\sigma-1}\right] + c(\sigma-1)\left(\frac{1-\alpha}{\alpha}\right)^{\sigma}c^{\sigma-1}\right\}T < 0$$

(d) *are less willing to devote resources in quality when they are more inclined to the observed tasks as we have* $\frac{\partial q^*}{\partial \gamma} < 0$.

3. Data

To empirically identify the effect of TBV establishment on local economic growth, we leverage on detailed township level shapefile, based on which we match each TBV to, and construct the dependent variables and control variables from multiple sources. Ideally, we should construct our dataset from the village level, there are several justifications why we end up choosing the township as our observation unit. First, the establishment of TBVs shows a strong salience effect (Huang et al. 2021). After a village establishes TBV, it will not only produce a radiation effect within the village but also spread to other neighboring villages (Liu et al. 2020). When the effect strengthens, other TBVs or Taobao Towns (TBT) may be established on top of the original TBVs (Zang et al. 2023). It will therefore be difficult to capture this spillover effect if only village-level data are exploited, and may lead to potential violations of our identification assumptions. Second, for the key dependent variable, nighttime light intensity, the current satellite data provided by NPP/VIIRS can only cover grids with scale of 500m×500m, while the resolution of DMSP/OLS is at a scale of 1km×1km, which results in fewer number of grids that can be covered by each village (or even none if the village is sufficiently small), thus creating additional measurement error issue.[22] Taking the above considerations into account, we choose the township as the geographic unit.

3.1 Research Sample

We use the yearly TBV information published by the Ali Research Institute, along with the national township map, to locate all TBVs established between 2014 and 2020 in the specific townships in which they are located, to determine whether there are TBVs established and the number of TBVs in the townships.[23] Figure 3 shows the geographical distributions and the timing of the establishment of TBVs in each township.[24] Throughout the sample period, a total of 1,572 townships have established TBVs, which is

[22] Another related issue is the lack of data availability. Currently, villages are geo-encoded as geometry points, and we are thus unable to map night light raster to the point-based shapefile.

[23] Although a small number of Taobao villages were also established before 2014 (numbering around 20), the criteria for determining these villages differed from the criteria after 2014. For consistent measurement, we refer to the general practice in the literature to exclude these early-established samples.

[24] If more than one TBV is established in a township, the earliest establishment will be referred as the timing of establishment for that township.

approximately 5% of all townships in China. It is observed that most TBVs are concentrated in the eastern coastal regions (e.g., Zhejiang, Jiangsu, Shandong, Fujian, etc.), with some appearing in the central and western regions.

3.2 Township level variables

Night light intensity Due to the lack of reliable statistical indicators to measure the economic activities at the township level, we adopt satellite-based nighttime lighting data as the proxy for local economic growth. A large body of literature has examined the relationship between nighttime light intensity and economic growth, and has shown its applicability (Chen and Nordhaus 2011; Henderson, Storeygard, and Weil 2012; Gibson et al. 2021). The granularity of nighttime light data also allows us to characterize the distribution of economic activity on smaller geographic units (i.e., township level). Given our sample period spans across 2010-2020, we use a combination of DMSP/OLS stabilized nighttime lighting data with NPP/VIIRS imagery from the National Oceanic and Atmospheric Administration (NOAA) NGDC data center.[25] Among them, DMSP/OLS is annual data, published from 1992 to 2013, and its image element gray value range is 0~63 with a spatial resolution of 30", while NPP/VIIRS is monthly data, which started to be published after 2012, and its spatial resolution is 15".

Based on the shapefile of the townships, we combined the two sets of data and used the DMSP/OLS data to calculate the average light intensity values of each township for 2010 and 2011, and also used the NPP/VIIRS data to calculate the average light intensity values of each township from 2012 to 2020.[26] It is important to note that the extent to which light is a desirable proxy for GDP decreases in smaller geographic units and areas with lower population densities (Gibson et al. 2021), and that light is a more suitable proxy for GDP in urban regions, where the development of industry largely depends on the consumption of electricity. It is thus important to bear in mind that one caveat of our study is that the night light luminosity may only reflect a certain aspect of the local economy in rural region, and could ignore other crucial aspects of the rural economy, such as employment. Therefore, a(n) increase/decrease in night light luminosity may not necessarily reflect the increase/decrease of overall welfare. Rather, what we are trying to emphasize is how the efficiency of the rural economy is affected by the establishment of TBVs. For this purpose, we further investigate how road investment and firm entry are affected by the establishment of TBVs (data are described below), which are both positively correlated with night light intensity and local GDP growth. Moreover, we also test the credibility of the light intensity measure in subsequent robustness tests.

---

[25] See https://ndgc.noaa.gov/eog/dmsp/dmsp.html for DMSP/OLS data and details; https://eogdata.mines.edu/download_dnb_composites.html for NPP/VIIR data and details.

[26] One potential concern is the comparability of nighttime lights between different data sources. The remote sensing literature (Ma et al. 2020) suggests using 2013 DMSP/OLS for intercalibration with NPP/VIIRS data. However, as Gibson et al. (2021) point out, DMSP data are less desirable for measuring regional economic activities than VIIRS data, and intercalibration may introduce additional measurement errors. We therefore do not correct the data in our baseline analysis, and in subsequent analyses only use the post-2012 data provided by VIIRS for robustness checks.

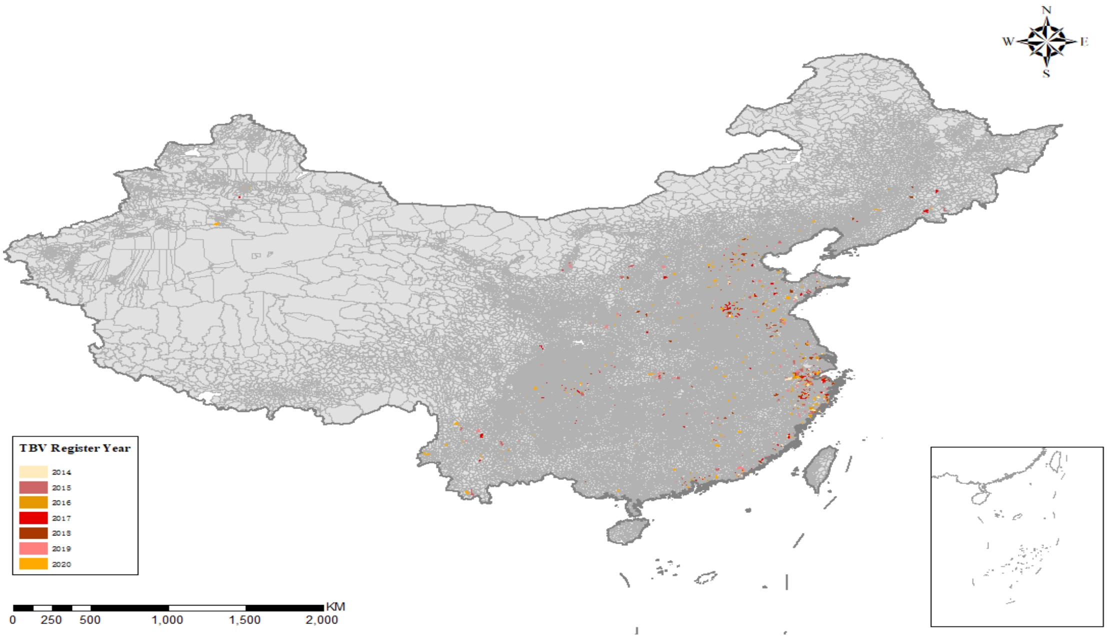


Figure 3. Temporal and Spatial Distributions of TBV Establishment

TBV establishment. We gather the list of all Taobao villages established between 2014 to 2020 from the Ali Research Institute. The dataset contains the detailed administrative code which enables us to locate each listed TBV to its specific township, to determine whether there is a TBV established, as well as the established number. The key explanatory variable is a dummy denoting whether TBV is established. The variable takes the value of 1 after the first establishment, and 0 otherwise. The simplified design makes our identification strategy analogous to the staggered difference-in-differences approach, which compares the differences in the outcome variables of the townships that have established TBV (i.e., the treatment group) before and after the treatment with the townships that have not established Taobao villages (i.e., the not-yet treated and control groups). This similar setup allows our identification to inherit both the strengths and weaknesses of the staggered DiD model, which we will discuss further in the subsequent section. As a complementary and robustness check, we use the number of Taobao villages in each township as an explanatory variable in the subsequent empirical evidence for further discussion.

Road density. The road density data used in this paper comes from the Open Street Map (OSM).[27] OSM is an open-source, user-contributed GIS-based system that allows users to view, edit, and use map data for free. Unlike proprietary mapping services (e.g., Google Maps), OSM's data is contributed by volunteers and can be used for free through a range of licensing methods.[28] The road network data provided by OSM starts from June 2013, and due to the time lag in updating the data, we consider the 2013 road data as the 2012 road inventory of the township. Using ArcGIS, we calculated the yearly road length in each township from 2012 to 2020.

Firm entry. To measure firm entry at the township level, we collected a township-level panel dataset of newly registered individual-run small businesses (IRSBs) from the National Register of Individual-Run Small Businesses (NRIRSB), maintained by the State Administration for Industry and Commerce (SAIC) of

[27] See https://www.openstreetmap.org for data and details.

[28] The open source nature of OSM leads to concerns of data accuracy. In fact, the accuracy of OSM data depends mainly on the frequency, quality and seriousness with which users edit and update the data. Existing experience suggests that for China, OSM is more accurate in non-border areas and relatively less accurate in border areas. We will test the accuracy of data measurement in subsequent robustness checks.

China. The NRIRSB is currently the most comprehensive dataset reflecting business registrations, especially small enterprises, and records the registration information of (almost) all individual industrial and commercial households on the Chinese mainland from 1949 until 2020. It covers information such as enterprise registration date, enterprise name, registered address, and industry classification. Specifically, we leverage the detailed registered address to decode the exact latitude and longitude using Gaode Map and map each registered enterprise to the specific township it is located. We then calculate the number of firm entries by industry at the township level from 2010 to 2020, and further exclude heavy industry sectors as well as non-manufactory sectors such as services based on the industry code inherited from the original dataset (the specific industries excluded are listed in Appendix Table A1). We further restrict the firm type to be an individually owned enterprise, which is more in line with our specific context since rural e-commerce businesses are more likely to be operated by individuals.

Control variables. We complement the above dependent and explanatory variables with additional control data, which we briefly outline here. First, we draw on a high resolution geo-topographical map (i.e., the 30m DEM raster file) to calculate the topography, average slope, and terrain ruggedness of each township. In particular, we calculate the ruggedness a la Nunn and Puga (2012). Second, based on the township shapefile, we measured the distance of each township to its county government, the distance to the coastline and the distance to Hangzhou (where Alibaba is headquartered). Finally, we obtained raster data on precipitation, sunshine, and agricultural production potential at 1km resolution for the year 2000 from the National Earth System Science Data Center (NESSDC).[29,30] Since the above controls are time-invariant, in the subsequent analysis, we interact the above variables with different time-trend terms (including linear, quadratic, and cubic terms) to flexibly control for the effects of these geographic characteristics. The summary statistics of the above variables are listed in Table 1.

Table 1 Summary statistics of township-level variables

Variable	Description	N	Mean	SD
log(light+1)	Logarithmic value of night light intensity	346,830	1.550	1.500
N_TBV	Number of TBVs established	346,830	0.0500	0.650
TBV	Whether a TBV is established	346,830	0.0200	0.120
log(road)	Logarithmic value of road length	346,830	3.450	1.060
log(# firm entry+1)	Number of private firm entries	346,830	0.338	0.680
topography	Average topography	346,830	161.3	184.3
slope	Average slope	346,830	2.670	3.140
ruggedness	Average terrain ruggedness	346,830	1.000	0.010
potential yield	Farmland potential yield (2000)	346,830	2963	2865
precipitation	Average annual precipitation (2000)	346,830	960.0	510.8
sunlight	Average annual sunshine (2000)	346,830	2001	575.3
distance county	Distance to the county government (KM)	346,830	19.96	18.02
distance Hangzhou	Distance to Hangzhou (KM)	346,830	1205.85	723.77
distance coast	Distance to coastline (KM)	346,830	633.8	631.4

4. Empirical Strategy and Results

4.1 Empirical Design

Baseline Model. We exploit the variations in the timing of TBV establishment and leverage the staggered difference-in-differences model to estimate the program's effect on the township economy. Specifically,

[29] The data on agricultural production potential are mainly based on the estimated annual total solar radiation of the national kilometer grid (averaged over 1950-1980) and calculated according to the production potential estimation model.

[30] See https://www.geodata.cn/ for details.

we estimate the following regression:

$$y_{it} = \beta TBV_{it} + \boldsymbol{Z}_i \times f(T) + treat_i \times T_t + \delta_i + \mu_{jt} + \epsilon_{it} \tag{1}$$

y_{it} is the measure of nighttime light intensity of township i in year t. $TBV_{it} \in \{0,1\}$ is a dummy variable that only takes 1 if the township has established TBV. $\boldsymbol{Z}_i$ is a set of township level time-invariant control variables, including the geographic location, topography and climate conditions. We interact these variables with the polynomial $f(T)$ of time trend to flexibly control for the dynamic effects of these non-time-varying factors. Individual fixed effects δ_i are also included to absorb township-specific heterogeneity. We also control for county-by-year fixed effects μ_{jt} to net out the effect of all county-level time-varying unobservables. One of the concerns in our identification is the staggered rollout of the "Comprehensive Demonstration of E-commerce in Rural Areas" program at the county level,[31] which could potentially affect both the establishment of TBV as well as the township economy. By including μ_{jt}, we are able to exploit more granular variations that only compare the outcome changes between treatment and control groups within the same county-year cell. In addition, to account for the differential trend between treatment and control groups, we control for the treatment-specific linear trend $treat_i \times T_t$ to allow the outcome trending differently. Finally, ϵ_{it} is the idiosyncratic error term clustered at the township level.

Event Study. The parameter of interest is β, which, under appropriate assumptions, identifies the causal effect of TBV establishment on the local economy. The key identification assumption in our setting is the conditional parallel trend assumption (CPTA), which states that conditional on the above controls, fixed effects, and time trends, in the absence of treatment (i.e., TBV establishment), the evolution trajectories of outcomes between treatment and control groups should follow a parallel trend. Although the CPTA is essentially untestable, we adopt an event study approach to provide a partial test on the identifying assumption, which allows us to track outcome trends before and after TBV establishment. The estimated equation is

$$y_{it} = \sum_{\tau=-5,\tau\neq-1}^{+5} \beta_\tau TBV_{it}^{\tau} + \boldsymbol{Z}_i \times f(T) + treat_i \times T_t + \delta_i + \mu_{jt} + \epsilon_{it} \tag{2}$$

In equation (2), we replace our key variable TBV_{it} with a set of leads and lags TBV_{it}^{τ} and estimate β_τ for each relative periods τ. To reduce noise and improve estimation precision, we combine all coefficients five periods before the establishment as well as five periods after the establishment. We follow the convention and omit the coefficient just one period before establishment to serve as the reference period. The flexibility of the event-study model allows us to both scrutinize the dynamic treatment effect as well as the presence of any significant pre-trend. However, this is only a partial test of the CPTA as the absence cannot be used to test what would happen had the TBV not been established. Moreover, the low statistical power in pre-trends testing may result in misleading interpretations in the causal estimates (Freyaldenhoven, Hansen, and Shapiro 2019; Roth 2022). We take three recently developed approaches to compensate for such deficiency and make more credible inferences. We first use the method proposed by Roth (2022) to explore how well-powered the pre-tests are in our context. We also adopt the "Honest DiD" estimator (Rambachan and Roth 2023) which provides bounds on the ex-post coefficients, allowing deviations in pre-period trends to be projected forward into the post-periods, rather than just assuming a parallel post trend as the common practices do. Finally, we use the synthetic DiD estimator (Arkhangelsky et al. 2021) which makes weaker assumptions than our workhorse model. Specifically, the synthetic DiD estimator combines merits from both the synthetic control method (SCM) and the DiD model that use data-driven approaches to

[31] The project aims at expanding the scope of rural markets, accelerating the flow of agricultural and industrial products between the rural and urban sectors, and promoting the wider use of e-commerce in rural areas. See http://images.mofcom.gov.cn/scjss/201507/20150727162640217.pdf for details

optimally align pre-treatment trends for each specific treatment group, leading to a more reliable comparison between treatment groups and the synthesized counterfactual groups.

Another challenge to our identification is the negative weighting issues highlighted in recent econometric literature, which point out that the standard DiD estimator by two-way fixed effects (TWFE) model may fail to recover a convex combination of the average treatment effects when there are variations in treatment timing and when treatment effects are heterogeneous across different treatment groups (de Chaisemartin and D'Haultfœuille 2020; Goodman-Bacon 2021). What's worse, the presence of heterogeneous treatment effects (HTE) may even contaminate the estimates of leads and lags in the canonical event study model (Sun and Abraham 2021). To prevent such contamination, we also use recently developed HTE-robust estimators (de Chaisemartin and D'Haultfœuille 2020; Callaway and Sant'Anna 2021; Sun and Abraham 2021) that avoid the involvement of negative weights in both DiD and event study designs.

Identification Assumption. Since the formation of TBVs may be confounded by other socio-economic conditions, it may not be considered a valid quasi-experiment. Therefore, equation (1) may not be able to identify the causal effects of TBVs. Due to the non-random selection of TBV, the estimated effects can be confounded by both omitted variables and reverse causality. In the former case (i.e., omitted variable), unobserved factors that correlate with both local night light and TBV establishment could lead to spurious effects. In the latter case (i.e., reverse causality), townships with better (or worse) economic conditions may be more likely to be selected to establish TBVs, which can either lead to overestimate or underestimate of our results, depending on the exact sign of the estimated coefficient. A careful examination of the CPTA may be able to alleviate concerns of omitted variable bias, as the disturbance of unobserved factors may be likely to break the parallel pre-trends, thus violating the CPTA. However, the presence of reverse causality can not be excluded even if no significant pre-trends are detected under the specification of equation (2).

4.2 Baseline Results

Table 2 reports the results of the baseline specification. Column (1) reports the most parsimonious results, controlling only for township fixed effects, county-year fixed effects, and treatment-specific linear trends. We add time-invariant controls that interacted with at most third-order polynomial functions of time trends gradually through columns (2) to (4). Throughout different specifications, our estimates hold significance at least at the 5% level with little changes in magnitude, suggesting less selections on unobservables (Altonji, Elder, and Taber 2005; Oster 2019). To give a direct illustration, we follow Oster (2019) and calculate the δ value to gauge to what extent unobserved omitted variables can explain our results. Oster (2019) suggests that $|\delta| > 1$ leaves limited scope for unobservables to explain the results. Reassuringly, the calculated absolute values of δ are large in magnitude and are well above 1, which suggests that the omitted variable bias has limited effects on our results.

Contrary to the common expectation, however, we find that the establishment of TBV *decreases* the night light intensity of townships. Specifically, according to our preferred specification (column (4)), we find that the establishment reduces township light intensity by about 3%, which is approximately equivalent to a reduction in GDP by 2.4% to 3.3%, based on the elasticity estimates from Gibson et al. (2021). Note that, since our main dependent variable is defined in log value, a negative result does not necessarily mean that the establishment of TBV decreases the absolute value of light. Rather, it may reflect the fact that the night light increases slower in treatment groups compared to the control groups.

We also try out several alternative econometric settings to our baseline specifications. First, we adjust for the potential selection bias using inverse probability weighting (IPW). We first estimate the Propensity Score (PS) for the treatment status based on logistic regression using the covariate $\boldsymbol{Z}_i$, with PS at the 1st and 99th percentiles excluded to ensure that the common support assumption. We then re-estimate our baseline results using a weighted least square with the inverse of PS serving as the weights. Results are reported in column (5) of Table 2. The estimates remain robust.[32] We also adopt the Coarsened Exact Matching (CEM) approach (Iacus, King, and Porro 2012) to ensure that the treatment and control groups are comparable.[33] Specifically, we stratify all our control variables into 10 equally sized bins and restrict comparison within groups that have similar observed characteristics. The results are highly robust. Next, we replace our dependent variable with other forms of transformations of the light intensity. Column (7) replaces the log transformation with the inverse hyperbolic sine transformation (IHS) to avoid zero values in light intensity.[34] Based on the elasticity formula of IHS transformation (Bellemare and Wichman 2019), the results from column (7) suggest that the establishment would lead to a reduction in night light intensity by 4.3%.

We also examine how the establishment of TBVs affects the absolute value of night light intensity. This exercise helps to delineate whether the TBV decreases the total light intensity of treated townships or just slows the increasing rate of township night light. The results are reported in Appendix Table C1. In general, we do not find strong evidence supporting that the TBV lowers the absolute township night light intensity, the estimated results are only marginally significant and are sensitive to the inclusion of controls.

Table 2 Baseline results

Dep. Var.	log(light+1)						IHS(light)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
TBV	-0.024**	-0.025***	-0.030***	-0.030***	-0.021**	-0.032***	-0.038***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.012)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Linear Trend	-	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Quadratic Trend	-	-	Yes	Yes	Yes	Yes	Yes
Controls × Cubic Trend	-	-	-	Yes	Yes	Yes	Yes
Inverse Probability Weighting	-	-	-	-	Yes	-	-
Coarsened Exact Matching						Yes	-
Oster δ for $\beta = 0$	-	-57.189	-15.306	-15.454	-	-	-16.062
Dep. Var. Mean	1.552	1.552	1.552	1.552	1.555	2.104	1.882
# of Clusters	31530	31530	31530	31530	30890	14884	31530
Observations	346,830	346,830	346,830	346,830	339,790	163,724	346,830
Adjusted R-squared	0.974	0.974	0.974	0.974	0.977	0.967	0.972

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects, and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. In columns (1)-(6), the dependent variable is log (1+light), whereas in columns (7), the dependent variable is the IHS transformation of night light luminosity. Standard errors reported in the parentheses are clustered at the township level.

Since there may be concerns that our estimates may be biased by the presence of negative weighting issues (de Chaisemartin and D'Haultfœuille 2020; Goodman-Bacon 2021), we replicate the baseline results

[32] It should be noted that the concerns over selections are of less severity given our negative estimates. Intuitively, one may be concerned that the existence of positive selections could lead to an overestimate of the treatment effects, as the selected TBVs are largely located on the eastern coast where the economy is more prosperous (as shown in Fig. 3). However, given such selection, the negative results are therefore more likely to be underestimated and a lower bound of the true effects.

[33] Compared with other matching techniques, the CEM method matches based on the coarsened joint distribution rather than precise values of observable characteristics.

[34] The formula for Inverse Hyperbolic Sine transformation is $arcsinh(x) = ln\left(x + \sqrt{1 + x^2}\right)$, which is quite analogous to the log transformation.

with the HTE-robust estimator proposed by Callaway and Sant'Anna (2021)[35] and the synthetic DiD estimator proposed by Arkhangelsky et al. (2021), which makes comparisons in a more " local" way (Bhalotra et al. 2023), i.e., giving more weight to control units that share more similar pre-treatment trends with treated units and more weight to periods that are more similar to the post-treatment periods. The results are reported in Appendix Table C2 and C3, respectively.[36] Specifically, for the DiD_{CS} estimator, we show in Appendix Table C2 that the results are robust to either the conventional aggregation using only OLS estimation or the (improved) doubly-robust aggregation based on stabilized IPW and OLS (Sant'Anna and Zhao 2020). For the synthetic DiD estimator, we show in Appendix Table C3 that the results are robust to the inclusion of controls interacted with different orders of time-trend polynomials. In sum, the estimated coefficients using these alternative estimators are consistently negatively significant, with little variations in absolute magnitudes.

However, one potential caveat of our research design is on the treatment unit. While the township information may capture more detailed information than county level official statistics, using township as the unit of analysis may still lead to biased estimates since the unit of treatment is defined at the village level. Specifically, the economic activity at the township level may also be driven by other villages other than TBV, which could introduce additional confounding factors that may substantially bias our results. To ensure that our results are not driven by other villages within the same township, we additionally construct a village panel that covers all villages that have ever established TBVs. Appendix D illustrate the detailed steps for constructing the village panel. Note that, since the village panel only comprises TBVs, when running the DiD specification, we can only compare outcomes of treatment groups with those not-yet treated groups (de Chaisemartin and D'Haultfœuille 2020; Callaway and Sant'Anna 2021). Although such design makes sure that the results may be more comparable as all units are eventually, treated, such a design could also exacerbate the potential negative weighting issues. To ensure that the results are not biased due to negative weights, we report the results from the DiD_{CS} estimator.

The corresponding results are shown in Appendix Table C4. Reassuringly, the estimated coefficients are still negative and significant at conventional levels even if we only conduct our analysis based on a subset of all villages. Related to our subsequent discussions, we also report the estimated effects for groups that are treated before and after 2016. We find that the effects are mostly driven by cohorts that are treated after 2016. This is in line with our narratives in the Background section, that TBVs established after 2016 are mostly characterized by top-down mandates, which may lead to resource misallocation and thus distort the process of economic development.

Further, the above result excludes two alternative hypotheses that may explain our main findings. First, it largely excludes the possibility that our results are driven by the economic activity of other non-TBVs within the same township, as our main finding still holds using the TBV sample. Second, it excludes the possibility that our results are driven by the inappropriate comparison between treatment and control groups (i.e., comparison between eventually treated and never treated groups), thus alleviating concerns of self-selection and omitted variable bias. Taking together, the above exercises and alternative specifications lend additional credence to our baseline finding.

4.3 Event Study

The validity of the above results primarily hinges on the assumption of the parallel trend. We provide

---

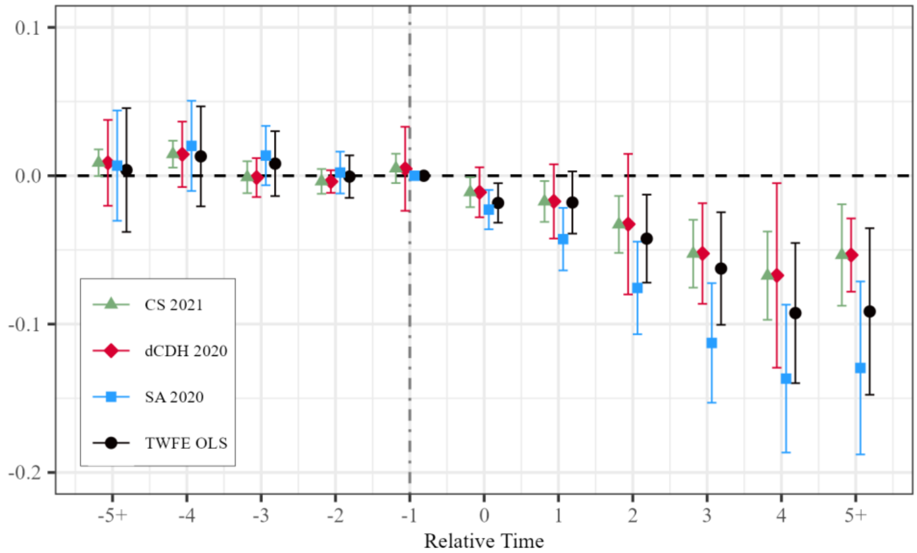
[35] We do not use the DiD_{M} estimator proposed by de Chaisemartin and D'Haultfœuille (2020) since it can only estimate the instantaneous treatment effects, while we are more interested in the aggregate treatment effects across the sample periods, which is more informative on the magnitude of both short-term and long-term effects.

[36] Both the DiD_{CS} estimator and the synthetic DiD estimator are robust to the biases inherent in standard two-way fixed effects models.

suggestive evidence for the assumption using an event-study approach. Figure 4 presents the estimated coefficients based on equation (2) as well as their corresponding 95% confidence intervals. Along with the canonical event-study estimates based on the two-way fixed effects model, we also report estimates from recent econometric literature that are robust to the presence of heterogeneous treatment effects and potential negative weights (de Chaisemartin and D'Haultfœuille 2020; Callaway and Sant'Anna 2021; Sun and Abraham 2021). To allow for a direct comparison, we plot all 4 estimates within the same figure.

We briefly summarize the findings from Figure 4. First, we detect no salient pre-trends in light intensity prior to the establishment of TBVs. Most of the leading terms are insignificant with the magnitude centered around zero. This partly alleviates our concerns over the potential violations of the identifying assumption. Given the controls, fixed effects, and time trends that are included in the regressions, we find that prior to the establishment, the outcome between treatment and control groups trending in a plausibly parallel way, which suggests that our estimated results from Table 2 are less likely driven by the ex-ante differences in light intensity.

Second, after the establishment of TBVs, the light intensity begins to decrease with a negative slope, suggesting the gradual deterioration in local economic conditions. Though there are slight differences across the four estimators, the downward trend is consistent. This alleviates our concerns about the negative weighting issues. Indeed, due to the presence of a large pure control group in our context, the variation exploited in estimation should be primarily driven by the comparison with the never-treated group, instead of comparison within treatment groups, where the latter happens to be the source of negative weighting problems (Goodman-Bacon 2021).[37] However, as our estimated effects stand in sharp contrast with the existing literature, there could be strong treatment heterogeneity across different treatment groups.[38] We discuss the issues in the subsequent section.



Note: This figure displays point estimates and the corresponding 95% confidence intervals from the event study as specified in Equation (2), with a time window ranging from 5 years before to 5 years after the establishment of TBV. Standard errors are clustered at the township level. The period just before establishment (i.e., $\tau = -1$) is set to be the reference period and thus omitted from the regression. The x-axis represents the time period relative to the establishment, while the y-axis represents the estimated coefficients. We simultaneously display four estimators. The first is the TWFE

[37] Using the Bacon decomposition proposed by Goodman-Bacon (2021), we find that comparisons between treatment groups account for only 1.9% of the total estimated weights, with the remaining 98.1% all comparisons between treatment and control groups.

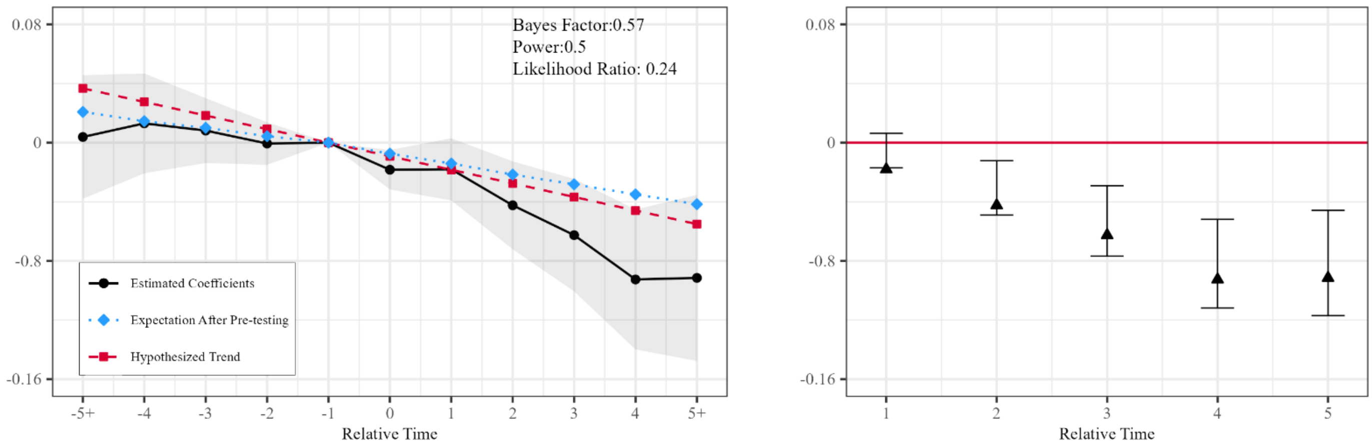
[38] We should note that the treatment heterogeneity itself does not cause much of concerns without staggered adoption, as in such cases, the DiD estimator still delivers the convex combination of all unit treatment effects, which can be interpreted as the aggregate effect. However, in the presence of staggered adoption, the negative weighting issue could make it hard to interpret the estimated coefficients, as they are no longer a sensible combination of all unit treatment effects.

OLS estimator, which includes township fixed effects and county-by-year fixed effects, as well as treatment-specific year trend and control variables interacted with a full set of polynomial functions of year trend. The second estimator is from de Chaisemartin and D'Haultfœuille (2020) that computes post-treatment effects by comparing outcome evolution in post-treatment period *t* with period *t-l*-1, where *l* represents the current period of treatment for a certain treated group and computes pre-treatment effect (or the placebo effects) by comparing the outcome of treated groups before treatment with those never treated groups. The third estimator is from Callaway and Sant'Anna (2021) which estimates the dynamic ATTs for each period relative to the first-treated period, across all cohorts. The fourth estimator is from Sun and Abraham (2021) which implements the interaction weighted approach and constructs pointwise confidence intervals for the estimation of dynamic treatment effects.

Figure 4. Event study results with multiple estimators

Nonetheless, recent studies have pointed out that the presence of no significant pre-trends is neither a sufficient nor necessary condition for the parallel trend assumption (Kahn-Lang and Lang 2020). Instead, the pre-test may be unable to detect potential deviations due to low statistical power, even when the ex-ante estimates are insignificant (Freyaldenhoven, Hansen, and Shapiro 2019; Roth 2022). To determine how well-powered our pre-trends test is, we follow the advice of Roth (2022) which allows for the presence of potential undetected pre-trends. A la Roth (2022), we hypothesize a linear pre-trend that can be detected with 50% power (i.e., we have a 50 percent probability of finding a significant pre-trend under the hypothesized pre-trend). As shown in Figure 5, Panel A, the bias generated by the hypothesized trend is substantially smaller than the true estimates, which alleviates concerns that our event-study estimates on post-treatment coefficients are driven by potential undetected pre-trend.

In addition, to further relax the parallel trend assumption and allow for potential deviations in the ex-post counterfactual trend, we adopt an "Honest DiD" approach proposed by Rambachan and Roth (2023). Essentially, their method imposes restrictions on the possible differences in trends between the treated and control groups but does not require the parallel trend to hold exactly. Our result, as shown in Panel B of Figure 5, indicates our estimates are still valid when considering potential violations of the parallel trend assumption.



Panel A Pre-trends test

Panel B Honest DiD

Note: This figure presents the pre-trends test proposed by Roth (2022), which is shown in Panel A, and the "Honest DiD" estimates proposed by Rambachan and Roth (2023), which is shown in Panel B. We plot the estimated coefficients in both panels, and their corresponding 95% confidence intervals for Panel A and the robust interval calculated from the valid inference under Rambachan and Roth (2023) methods. In panel A, we add the hypothesized trend and the expected values after pre-testing. The analysis suggests that we can detect a fairly small positive linear trend of a magnitude of 0.012 (the slope of the red line) or greater (in absolute terms) with 50% power in our event-study design. Based on the hypothesized trend, we calculate the potential bias it would generate, as shown by the blue line. In panel B, we construct the valid 95% confidence intervals by allowing differential pre-trends for treatment and control groups, permitting violations of standard parallel trend assumptions.

Figure 5. Testing for the Pre-trends

4.4 Other Robustness

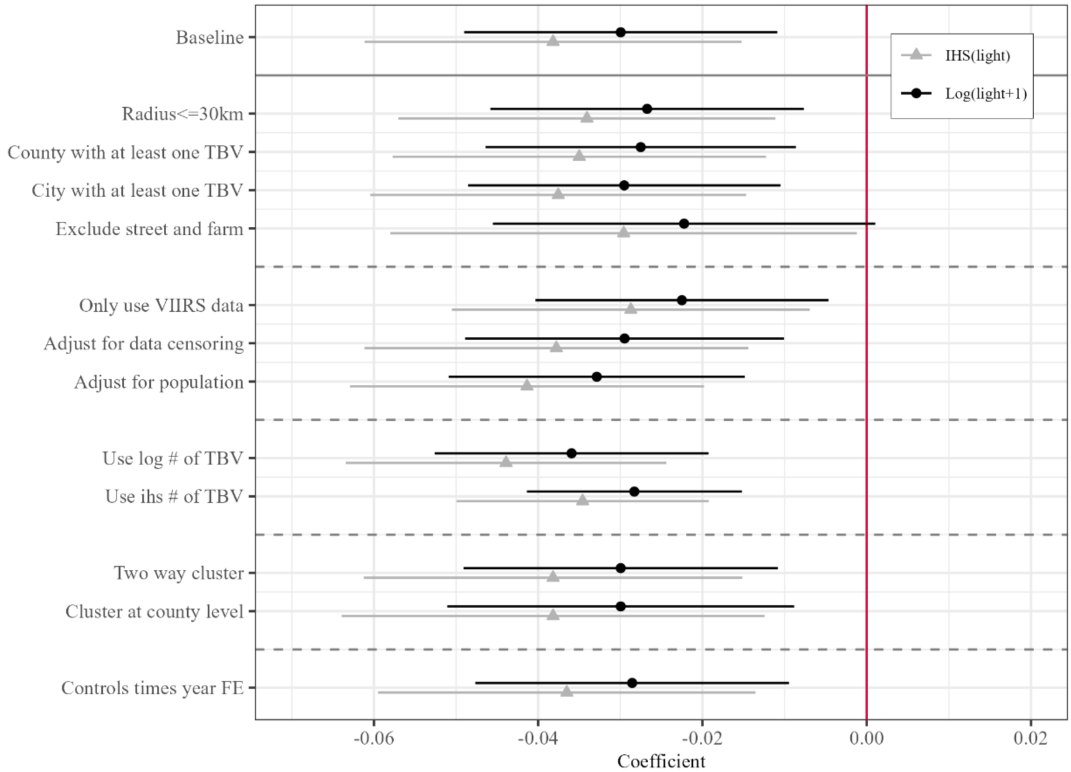
In this section, we briefly summarize some of the additional exercises that ensure the robustness of our baseline estimates. We summarize the results in Figure 6 and relegate other additional exercises in Appendix

C. In particular, the first row of Figure 6 replicates the results of columns (4) and (6) of the baseline regression in Table 2 to serve as a comparison. The remaining rows report the results across different robustness specifications. We summarize them in order.

Comparability of treatment and control groups. One potential concern with our identification is the comparability between treatment and control townships. Although we exploit inverse probability weighting in our baseline regression to adjust for the probability of being treated, the differences in geographic distribution may nonetheless introduce unobserved heterogeneity that confound our results. As shown in Figure 2, most of the treated townships are concentrated on the eastern coast, with only a small number of townships located in the central and western regions. We therefore replicate our baseline estimate within a restricted sample to alleviate concerns arising from such an issue. First, we restrict the control group to townships within a radius of 30 KM of their nearest treatment township.[39] Second, we restrict the sample to counties that contain at least one treatment township. Third, we restrict the sample to prefectures with at least one treatment township. The key idea of the above exercises is to limit the control group to a geographically adjacent range from the treatment group, thus reducing the potential confound due to differences in geographic distribution. The corresponding estimation results are reported in rows 2, 3, and 4 of Figure 6. As can be seen, the estimates remain largely unchanged.

The comparability is not only in terms of geographic distribution, but also in terms of different township types. Traditionally, China divides sub-county administrative divisions into four categories: streets, towns, townships, and other forests, farms, etc. The streets represent the more developed and modern urban areas, while the townships are mainly rural areas. The definition of a town is more complex, as it has both urban and rural features and is engaged in both agricultural and industrial production. Since TBVs are mainly formed in townships or towns, we therefore exclude the samples of streets and farms and re-estimate the baseline regressions, the results of which are plotted in row 5 of Figure 6. This test also addresses a potential concern that the slowdown in township economic growth following the establishment of TBVs may have been primarily due to the faster development of the urban sector. Our results suggest that excluding the street sample does not significantly affect the negative effect.

[39] Appendix Table C5 further examines the samples within different radii. Specifically, we re-estimate the baseline regressions by restricting control townships within the radii of 10KM-50KM (at 10 KM intervals) from the nearest treatment group. The results remain robust.



Note: This figure displays point estimates and the corresponding 95% confidence intervals from various robustness checks. The x-axis represents the value of coefficients, while the y-axis denotes the specific robustness checks. All regressions include township fixed effects and county-by-year fixed effects, as well as treatment-specific year trend and control variables interacted with either a full set of polynomial functions of year trend or year fixed effects. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast.

Figure 6. Summary of robustness checks

Measurement error and proxy issues of night light data. Another challenge to our identification originates from the appropriateness of exploiting nighttime light as a proxy variable for township economic development. This challenge is twofold. First, there is measurement error in the satellite data itself. Since we cross-utilize two data sources for lights, DMSP and VIIRS, the specific level of GDP represented by each value may change, despite the fact that both measure lights between 0 and 63. In addition, there is considerable data censoring in the night light data. Due to problems with satellite measurements, a large number of observations may actually be at 0 or at 63, while the actual level of represented GDP may be different. Second, there may be the problem of proxying GDP in rural areas. Gibson et al. (2021) note that nighttime lights may not fully reflect the level of economic development in rural areas, as much of the agricultural production does not require light input in practice.

Rows 6, 7, and 8 of Figure 6 respond to the above challenges. First, since the DMSP is relatively less precise and nosier than the VIIRS in predicting GDP growth, in row 6 of Figure 6 we exploit only the VIIRS data, i.e., we use only the post-2012 sample for the estimation. Second, to address the problem of data truncation, mainly due to the presence of extremely backward and developed regions, we exclude samples with either minimum or maximum values (i.e., 0 and 63) from the regression in row 7 of Figure 6. Finally, while we are unable to directly address the proxy issues in rural areas, we argue that the bias induced by such problem is unlikely to explain the pattern in our data, as the proxy power has to mimic the exact decline upon the establishment of TBV. This is less likely given that the number of established TBVs is large in absolute amount (i.e., more than 5000 in total). Nonetheless, we still attempt to partially correct the proxy issues by using the population density as weights and run a weighted regression. The population data comes from the population raster provided by Landscan.[40] Since night lights proxy for GDP better at higher population densities (Gibson et al. 2021), the weighted least squares estimation essentially gives more weight to areas that are more densely populated, thus mitigating the proxy issue to some extent. The population-adjusted estimates are reported in row 8 of Figure 6. The estimates remain robust.

Alternative measures. For the ease of interpretation, in our baseline estimation, we only use a parsimonious measure of the establishment of TBV, while ignoring how the number of TBVs would affect our results. We consider the impact of changes in the number of TBVs on township light intensity in rows 9 and 10 of Figure 6. The estimated coefficients we obtain using the log-log (ihs-ihs) form are not significantly different from the estimated coefficients from the baseline regression and suggest the elasticity ranging from 3-4.

We also considered the potential problems associated with using logarithmic forms for nighttime lights (Chen and Roth 2023). It is important to note that nighttime lights do not have a unit of measurement like income (e.g., dollar or cent), but are more akin to discrete variables. We therefore follow the recommendation of Chen and Roth (2023) and use the PPML model for estimation. The estimation results are reported in Table C6 in Appendix C, which replicates the identification of columns (1)-(5) of Table 2 of the benchmark regression. Reassuringly, we find no substantial differences when exploiting the PPML model and the estimated coefficients are all significant at 1% level.

Adjusting clustered standard errors and control variables. Our baseline results are robust to adjustments in both standard errors and control variables. Specifically, row 11 of Figure 6 uses two-way clustering (Cameron, Gelbach, and Miller 2008) to cluster standard errors at both the township level as well as the county-year level to control for both serial correlation among townships as well as yearly cross-sectional correlation within the same county. Row 12 then clusters standard errors at the county level to control for any arbitrary correlations between townships within the same county. The estimates are still significant at the 1% level. Finally, row 13 of Figure 6 interacts the control variables with the full set of time dummies to allow the flexible effect of control variables. The estimates remain robust.

5. Further Analysis

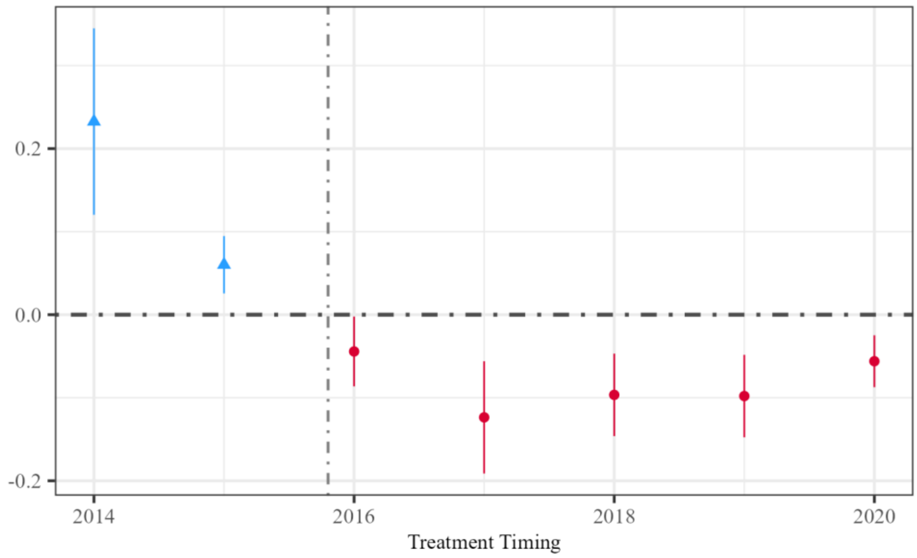
We now take one step further and provide a political economy explanation for the above counter-intuition results. Our analyses are mainly focused on the verification of our theoretical predictions in section 2.4 and explore the role of local officials' political incentives. We provide several pieces of evidence to support the above argument. We first show that the negative effects are salient only after the central government advocates local officials to exploit top-down directives to establish TBVs. Consequently and correspondingly, the negative effects are concentrated in places where local governments put more emphasis on developing rural e-commerce, and the effects are more salient when local governments are motivated by their promotion incentives. Second, we exploit a battery of heterogeneous analysis to lend additional credence to our hypothesis and provide indirect inference on the selection criteria in the top-down formation mechanism. We suggest that townships that are closer to the county government, and townships with better economic and natural conditions are more likely to be chosen as TBVs. Alternatively, counties with more resources are more likely to use top-down directives to establish TBVs. Finally, we present additional evidence on changes in road investment that help to reveal the strategic behaviors of local officials. We also document changes in firm entry, as one of the potential mechanisms that undermine the development of the local economy.

[40] The data is at a spatial resolution of 1 kilometer, with yearly updates starting in 2000. The LandScan Global is the highest-resolution global population distribution data available that represents the ambient (24-hour average) population, using an innovative methodology that combines geospatial sciences, remote sensing and machine learning algorithms. The data are available at https://landscan.ornl.gov/.

5.1 The Effects of Top-down Enforcement

Central government's advocates. Since the motivation of local governments to establish TBVs depends to a large extent on the policy orientation of the central government, if our negative effects are mainly driven by government intervention, a predictable result is that a change in the central government's policy guidance will significantly affect our estimated effects. We note a highly representative example occurred at the end of 2015. Specifically, in October, the then government premier Keqiang Li hosted a State Council executive meeting and proposed for the first time to accelerate the development of rural e-commerce.[41] Therefore, local governments are more likely to respond to the State's call to develop rural e-commerce (as represented by TBVs) after 2016, resulting in a top-down formation mechanism in TBVs' establishment. In contrast, prior to 2016 (e.g., 2014 and 2015, the first two years when TBVs began to emerge), there is limited incentives for local officials to develop rural as much of the priorities are given to urban development (Wang, Zhang, and Zhou 2020), and thus TBVs established before 2016 are more likely follow a bottom-up formation mechanism by local entrepreneurs who seek to maximize their expected profit.

We verify our argument in Figure 7, where we divide the treatment group into different cohorts based on their treatment timing, and estimate the treatment effects for each specific cohort. We find strong treatment effects heterogeneity across different cohorts. Specifically, the effects of TBV establishment are positive in the first two years, and the estimated coefficients are both statistically and economically significant. For instance, we find the establishment of TBVs in the first treatment period (i.e., 2014) increases township light intensity by over 20%, suggesting the early bottom-up establishment of TBVs can significantly drive local economic growth. In stark contrast, we find after 2016, possibly when the central government started emphasizing the need to vigorously develop rural e-commerce, the establishment brought consistent and persistent negative effects on the local economy. The pattern in the figure suggests a decrease in light intensity by about 10%. Since most of the villages are established after 2016, as shown in Figure 1, the aggregated effects are still negatively significant, as we show in our baseline and event-study estimates.



Note: This figure displays point estimates and the corresponding 95% confidence intervals for each treatment cohort. Standard errors are clustered at the township level. The x-axis represents the treatment timing, while the y-axis denotes the corresponding coefficients. All regressions include township fixed effects and county-by-year fixed effects, as well as treatment-specific year trend and control variables interacted with a full set of polynomial functions of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000),

[41] See https://www.gov.cn/guowuyuan/2015-10/14/content_2946877.htm. Shortly thereafter, the State Council issued a guiding opinion on promoting the development of rural e-commerce, proposing that the development of rural e-commerce should lead to rural poverty alleviation and development, among other things. See https://www.gov.cn/zhengce/content/2015-11/09/content_10279.htm for more detail.

sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast.

Figure 7. Estimation for treatment-specific cohort

The above evidence explains at the same time why most of the existing studies on TBVs estimate a positive impact. A highly likely reason lies in the sample selection. Since these studies are mainly based on survey data or case studies, there may be positive selections in the villages being surveyed, as these successful villages are more likely to be noticed, either through media coverage or government publicity. Our study further suggests that although Taobao villages established in the early stages do have a positive impact on the local economy, with the strengthening of government intervention, TBVs may instead undermine the local economy. Given that the development paradigms of TBVs are highly respected, our findings therefore emphasize the importance of a systematic evaluation of the aggregate and overall effects of TBVs.

Local governments' emphasis. The above evidence only exploits the temporal variations and may be obscured by unobserved time-varying confounders. To provide additional support and reveal the role of local governments in top-down enforcement, we test whether the government's emphasis on developing rural e-commerce can explain our negative results. To this end, we manually collected the government annual reports of 284 prefectures across China from 2010 to 2020 from either government websites or news reports. We use the word frequencies of relevant keywords mentioned in the government reports to reflect the relative importance that local governments attach to the development of rural e-commerce. Intuitively, more mentions of these keywords represent higher emphasis on the corresponding issue (Chen, Li, and Lu 2018). The keywords we extracted can be mainly divided into four categories. The first category is directly related to rural e-commerce, such as "rural e-commerce (Nongcun Dianshang)", "Taobao village (Taobao Cun)" and "digital villages (Shuzi Xiangcun)". The second category is related to rural development, such as "rural revitalization" and "poverty alleviation". The main reason for our focus on rural revitalization is that the development of TBVs and rural e-commerce has been closely integrated with poverty alleviation since 2016, especially after the central government included "rural revitalization" in the appraisal objectives of local officials in 22 provinces across the country in 2018 (He, Lu, and Lee 2023). The third and fourth categories are not directly related to rural e-commerce and are mainly used to serve as placebo tests. Specifically, the third category only relates to the development of e-commerce in the urban sector, such as "e-commerce (Dianzi Shangwu)" and "digital economy (Shuzi Jingji) ". The fourth category involves agricultural industrial policies, such as "rural industry (Nongcun Chanye)" and "agricultural industry (Nongye Chanye)". Overall, the first two categories are more relevant to rural e-commerce, while the last two categories are relatively less relevant.

For each of the categories, we generate a dummy variable High KW denoting above-average frequencies and estimate an augmented version of our baseline model with the keyword dummy included as an interaction term. Columns (1) and (2) of Table 3 firstly report the results for the rural e-commerce category. The estimated coefficient of the interaction term, TBV×High KW, is significantly negative, while the coefficient of the main term, TBV, is no longer significant and close to zero. This suggests that the local government's emphasis indeed drives our baseline results. We find consistent evidence when we use the rural revitalization category as the interaction in columns (3) and (4), that the local government's emphasis on rural revitalization is equally likely to lead to negative effects. In columns (5) and (6), we sum the frequencies of the rural e-commerce and the rural revitalization categories, and the results remain robust and significantly negative.

In the remaining four columns, we conduct a placebo test using word frequencies that are not directly related to rural e-commerce. In particular, the word frequencies in columns (7) and (8) contain only general e-commerce and do not directly target rural areas. We find that the development of general e-commerce by local governments does not significantly affect the rural economy. The coefficients of the interaction term, although negative, are not significant at the conventional level. Columns (9) and (10) replicate the previous results with the word frequencies replaced with categories related to agricultural industrial policies. Again, we do not find that the government's emphasis on these efforts significantly affects our baseline results. Overall, the above exercises suggest that the township economy is negatively affected only when the local government directly emphasizes the development of rural e-commerce or rural revitalization.

Nonetheless, since we do not directly observe which TBVs are established through top-down directives and which are not, to prove that the negative results are indeed driven by the local governments' emphasis, it is crucial to show that there exists a "first stage" channel, that is, the emphasis of local governments promote the establishment of TBVs, which eventually leads to the observed negative results. In Appendix Table C7, we perform a simple test by regressing the independent variable (either dummy of TBV establishment of the number of TBVs established) on three set of keyword frequencies (rural e-commerce, rural revitalization, and the total frequency by summing all 4 categories in Table 4 together). We find that only the keyword frequency of rural e-commerce has significant positive effects on the probability of TBV establishment (or the number of TBVs established). Importantly, we note that the effects are only significant for TBVs established after 2016, aligning with the pattern we show in Figure 7. This evidence suggests that local government indeed response to central government's initiatives by establishing more TBVs.

However, we note that this evidence is only suggestive as it only shows that higher emphases on rural e-commerce are positively correlated with more established TBVs. There could be issues like reverse causality that may obscure the interpretation of the results, that is, the rapid growing of TBVs leads the local government to put more weights on the development of rural e-commerce. We partially mitigate this concern by replacing the independent variables in Table C7 (i.e., keywords frequencies) with their lagged terms. Results are reported in Appendix Table C8, which again provide strong support to our findings.

Table 3 Heterogeneity in local government emphasis

Dep. Var.	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)
Key Words:	Rural E-commerce		Rural Revitalization		Both		E-commerce		Rural Industry	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
TBV×High KW	-0.044***	-0.056***	-0.031**	-0.046**	-0.025*	-0.039**	-0.015	-0.026	0.004	0.002
	(0.014)	(0.017)	(0.015)	(0.018)	(0.014)	(0.017)	(0.013)	(0.016)	(0.013)	(0.015)
TBV	-0.002	-0.002	-0.000	0.005	-0.005	-0.002	-0.010	-0.009	-0.032***	-0.039***
	(0.012)	(0.015)	(0.014)	(0.017)	(0.013)	(0.016)	(0.012)	(0.015)	(0.011)	(0.013)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	1.552	1.882	1.552	1.882	1.552	1.882	1.552	1.882	1.552	1.882
# of Clusters	31530	31530	31530	31530	31530	31530	31530	31530	31530	31530
Observations	346,830	346,830	346,830	346,830	346,830	346,830	346,830	346,830	346,830	346,830
Adjusted R-squared	0.974	0.972	0.974	0.972	0.974	0.972	0.974	0.972	0.974	0.972

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects, treatment-specific year trend and control variables interacted with a 3-order polynomial function of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. In odd columns, the dependent variable is log (1+light), whereas in even columns, the dependent variable is the IHS transformation of night light luminosity. The keyword frequencies interact with the TBV dummy and treatment dummy. The variable key word frequencies itself is absorbed by the county-year FE. Standard errors reported in the parentheses are clustered at the township level.

5.2 Political incentive and strategic behaviors of local officials

We further examine how local officials' political incentives may distort the local economy. The political economy literature suggests the strong presence of political cycles in China (Chen and Zhang 2021), where local officials facing the pressure of evaluation and promotion may have incentives to devote most of their resources to economic growth in urban sectors while resorting limited effort to improve the development of TBVs. As in China's meritocracy-based political selection process, the economic performance within the jurisdictions led by local officials serves as a pivotal factor in the evaluation of their promotions by upper-level authorities (Fang, Liu, and Zhou 2023). In regular times, when confronted with top-down directives to promote TBVs, local officials are inclined to allocate greater fiscal resources towards TBV development, aligning with the political competition they face. However, during critical evaluation periods, they prioritize directing resources towards enhancing economic performance in urban areas, as such achievements are more readily observable and evaluated by their superiors. These strategic responses arising from the political cycle may introduce distortions in TBV development, thereby negatively impacting the local rural economy.

To verify the distortion effects of local cadres' political incentives, we manually collect the biographies of the county's party secretaries (PS), the top leader of the county, from various sources (including the official websites, Baidu Baike, and news reports), and cross-validate the data sources whenever possible.[42] We calculate the term length of each PS, as the probability of promotion increases with the term length (Guo 2009), and local cadres with longer term lengths are more motivated to flatter their superiors by establishing more TBVs to ensure promotion. We also calculate whether there are leadership turnovers during the sample periods, since the political incentives when the turnover is approaching may also bring substantial distortions. Following Table 3, we estimate the baseline regression with PS's term length or leadership turnovers incorporated as an interaction term. The estimated results are reported in Table 4, columns (1) to (4). We find that the political incentives indeed account for the negative effects we estimate. As shown in columns (1) and (2), PS with an additional year in office would enlarge the negative effects of TBV establishment by 1.7%-2.2%. Columns (3) and (4) also suggest that the negative effects are substantially explained by counties with leadership turnovers, which again coincides with our hypothesis. We also examine the role of county governors' (CG) political incentives. We replicate the results in columns (5) to (8), where we find no significant effects driven by the promotion incentives of county governors. This is not surprising since party sectaries hold greater and more arbitrary power over county governors, and are the ones really in charge of the Chinese government.

Table 4 Heterogeneity in leadership term length and turnover

Dep. Var.	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PS Term×TBV	-0.017**	-0.022***						
	(0.007)	(0.008)						
PS Turnover×TBV			-0.048*	-0.062**				
			(0.025)	(0.029)				
CG Term×TBV					0.000	-0.001		
					(0.007)	(0.008)		
CG Turnover×TBV							-0.012	-0.023

[42] Since not all county leaders' information are publicly available, we have only collected information for 1588 county leaders.

							(0.026)	(0.030)
TBV	0.014	0.019	-0.008	-0.008	-0.018	-0.018	-0.008	-0.005
	(0.025)	(0.029)	(0.020)	(0.024)	(0.025)	(0.030)	(0.022)	(0.026)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	1.471	1.791	1.471	1.791	1.424	1.736	1.424	1.736
# of Clusters	15088	15088	15088	15088	17582	17582	17582	17582
Observations	165,968	165,968	165,968	165,968	193,402	193,402	193,402	193,402
Adjusted R-squared	0.970	0.967	0.970	0.967	0.970	0.968	0.970	0.968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects, treatment-specific year trend and control variables interacted with a 3-order polynomial function of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. In odd columns, the dependent variable is log (1+light), whereas in even columns, the dependent variable is the IHS transformation of night light luminosity. County leaders' term length and rotation status interacted with the TBV dummy. The variables on county leaders' term length and rotation status are absorbed by the county-year FE. Standard errors reported in the parentheses are clustered at the township level.

5.3 Alternative Explanations

As the above analyses only provide indirect validation to our preferred explanation (i.e., distortion driven by local officials' strategic responses), we further explore other channels and the presence of reverse causality that may explain the results. In this subsection, we discuss several alternative explanations and provide evidence that they can not explain the negative effects of TBV establishment.

Non-random targeting of TBVs. A leading competing explanation for the above results is the strategic selection of TBVs. As the central government emphasizes the role of TBVs in alleviating rural poverty, local officials may select townships with worse economic performance to promote the establishment of TBVs. Such a strategic selection may introduce negative correlations between the night light intensity and TBV establishment (which leads to the concern of reverse causality), coinciding with our baseline results. To show that this is not the case, we manually collect information on whether a county is a poverty county designated by the central government, and estimate an augmented regression with this dummy interacted with our key explanatory variable. The results, presented in columns (1) and (2) of Table 5, indicate no significant difference in treatment effects between poverty and non-poverty counties. If any, the negative effects are even weaker in poverty county, which contrasts the hypothesis that the negative results are driven by targeted selection of TBVs.[43]

Constrained structure transformation. Although TBVs are located in rural regions, the production and consumption of goods produced in TBVs require tight interconnection with the urban sector. The development of TBVs *per se* also requires the renovation of traditional rural communities. Therefore, the negative results may be driven by TBVs undergoing unsuccessful structural transformation. In China, around 800 counties are chosen to specialize in agricultural production to ensure food security (known as Major Crop (MC) County). These counties are constrained by institutional factors and are in a slow pace of structural transformation. If the negative results are driven by constrained structural transformation, then we should find stronger negative effects in these counties. However, the results from columns (3) and (4) of Table 5 suggest no differential effects between MC county and non-MC county, which is inconsistent with the above explanation.

[43] The results from Figure 8, which we present in the next subsection also find that the negative results are concentrated in townships with better economic conditions, thus highlighting the targeting selection of TBVs in the opposite direction.

Cultural and informal institutional factors. One reason for the success of TBVs in rural regions is the presence of informal institutions that facilitates the cooperation within villages. An important component is the long-lasting clan cultural (Chen, Ma, and Sinclair 2022; Tang and Zhao 2023), which promotes cooperation through risk sharing and easing credit constraints. In China, clan culture is geographically concentrated in the south, while exerting limited role in the north. This pattern aligns with the spatial diffusion of TBV formation: TBVs are first established in south-east China, where the clan culture flourishes, and then gradually diffuse to the northern and western regions, where the clan culture is less prevalence. It is thus possible that the underdevelopment of TBVs is the result of lacking support of clan culture. We test this alternative hypothesis in columns (5) and (6) of Table 5, where we measure the clan culture by the number of genealogies compiled before the founding of the People's Republic of China.[44] We do not find any significant effect of clan culture in explaining the negative results.

Table 5 Alternative explanations

Dep. Var.	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TBV×Poverty County	0.036	0.058						
	(0.034)	(0.041)						
TBV×MC County			0.050	0.056				
			(0.033)	(0.040)				
TBV×Clan Culture					0.009	0.009		
					(0.007)	(0.007)		
TBV×Internet Access							-0.017	-0.024
							(0.014)	(0.016)
TBV	-0.033***	-0.043***	-0.035***	-0.044***	-0.033***	-0.040***	-0.022*	-0.027**
	(0.010)	(0.012)	(0.010)	(0.012)	(0.011)	(0.013)	(0.011)	(0.014)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	1.471	1.791	1.471	1.791	1.424	1.736	1.424	1.736
# of Clusters	15088	15088	15088	15088	17582	17582	17582	17582
Observations	165,968	165,968	165,968	165,968	193,402	193,402	193,402	193,402
Adjusted R-squared	0.970	0.967	0.970	0.967	0.970	0.968	0.970	0.968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects, treatment-specific year trend and control variables interacted with a 3-order polynomial function of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. In odd columns, the dependent variable is log (1+light), whereas in even columns, the dependent variable is the IHS transformation of night light luminosity. County leaders' term length and rotation status interacted with the TBV dummy. The variables on county leaders' term length and rotation status are absorbed by the county-year FE. Standard errors reported in the parentheses are clustered at the township level.

Access to the Internet. The well-functioning of TBVs relies highly on the construction of ICT infrastructure. One reasonable explanation for the negative results of TBV establishment may due to the insufficient access to the ICT infrastructure. If this were the case, then we should find negative results to be concentrated in locations with worse access to the internet. To measure the extent of internet access, we rely on exogeneous variations stemming from the construction of a large optical fiber network across major cities in the beginning of 2000s.[45] The network facilitates the rapid

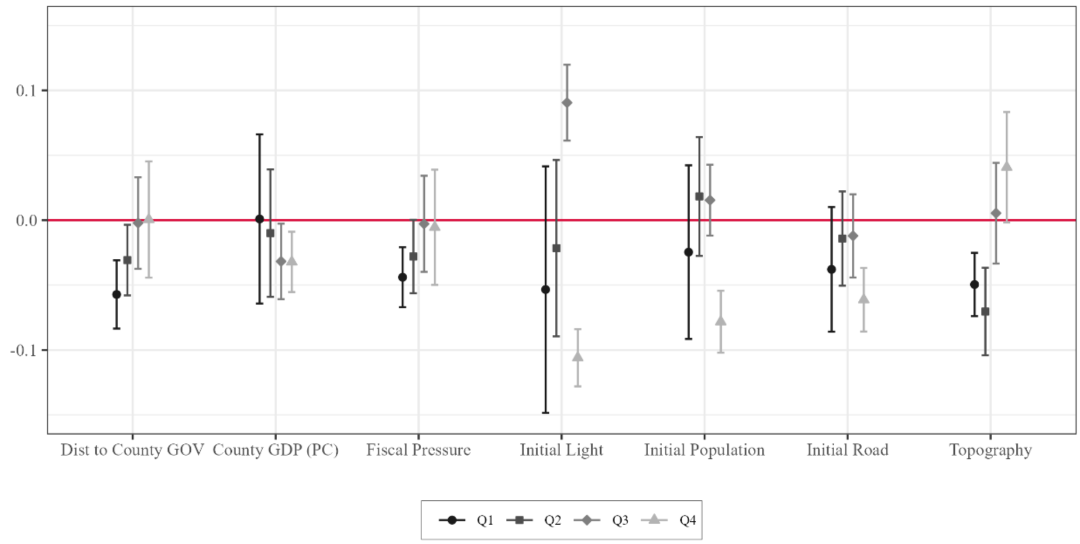
[44] Data on the location of each genealogy is from https://jiapu.library.sh.cn/#/. We calculate the number of genealogies for each county and normalize it by county's population in 2000.

[45] See report from https://www.c114.com.cn/market/37/a368093.html.

development of the communications industry and serves a good proxy for internet access (Hjort and Poulsen 2019). To measure the extent of internet access, we draw on the shapefile of the network and calculate the distance of each township to its nearest network. The estimated results using the distance to the nearest network as interaction term are reported in columns (7) and (8) of Table 5. Again, we do not find any significant evidence showing that the access to the internet can explain our baseline findings.

5.4 Heterogeneous Analysis

This subsection attempts to provide additional support and reveal the selection criteria of TBVs by exploiting a heterogeneous analysis. To this end, we examine how the following factors affect our negative effects: the distance to the county government (a measure of vertical control cost of county government); the GDP per capita, and the fiscal pressure at the county level (measures of the capacity and resources can be used to establish TBVs);[46] the initial light intensity initial population as well as the road density (measures of township's initial economic conditions); and township's topography (a measure of township's natural conditions). We estimate the quantile treatment effects at the quartile level for each of the above variables and summarize the results in Figure 8.



Note: This figure displays point estimates and the corresponding 95% confidence intervals from the estimation of the quantile treatment effects. Standard errors are clustered at the township level. The x-axis represents the specific variable that is exploited in heterogeneous analysis, while the y-axis denotes the corresponding coefficients. All regressions include township fixed effects and county-by-year fixed effects, as well as treatment-specific year trend and control variables interacted with either a full set of polynomial functions of year trend or year fixed effects. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast.

Figure 8. Heterogeneous Analysis

Three primary patterns emerge. First, we find that the negative effects are concentrated in townships that are closer to the county governments. A plausible explanation is that the county governments have better information about these townships and are more likely to resort to a top-down approach to establishing TBVs. This result again confirms that the establishment of TBVs is largely influenced by local government intervention, otherwise the distance to the county government should not have a significant mediation effect on the estimated results. Second, we find that counties with better economic and fiscal capacities are more likely to adopt a top-down approach to TBV

[46] Specifically, we measure the fiscal pressure by using the gap between fiscal expenditure and fiscal revenue, normalized by county's GDP.

establishment. Our results show that the negative effects are mainly driven by counties with higher GDP per capita and lower fiscal pressure. These counties have relatively more resources at their disposal and are more capable of exploiting these resources to invest in the establishment of TBVs. Finally, we discover that townships with better conditions (e.g., more population, road, and less roughness) are more likely to experience a negative impact, which implicitly suggests that these townships are more likely to be chosen to establish TBVs.

Interestingly, we find that TBV establishment can have a positive impact when the initial light luminosity is just moderately above the median (i.e., lies between the median and the upper quartile), as well as when the terrain is more rugged. This may be because TBVs established in these townships are more likely to follow a bottom-up formation mechanism. Some media reports also support our hypothesis. These reports indicate that a large number of Taobao villages established in mountainous areas have achieved market expansion through the use of e-commerce platforms, which has led to the rapid development of the local economy.[47]

The above results also echo some of the potential competing hypotheses. For instance, one might argue that TBVs experienced a significant expansion after 2016, which led to increased competition for online sales, and thus the negative effect we estimate might simply be due to the poor operation of these Taobao villages. A priori, we would assume that these poorly run TBVs are more likely to be those with poorer initial conditions, but based on the results in Figure 8, we are forced to assume that these poorly run Taobao villages are more likely to be located in townships with better initial conditions. Therefore, our results are more in favor of the hypothesis that local governments exploit top-down directives to establish TBVs in townships with better conditions.

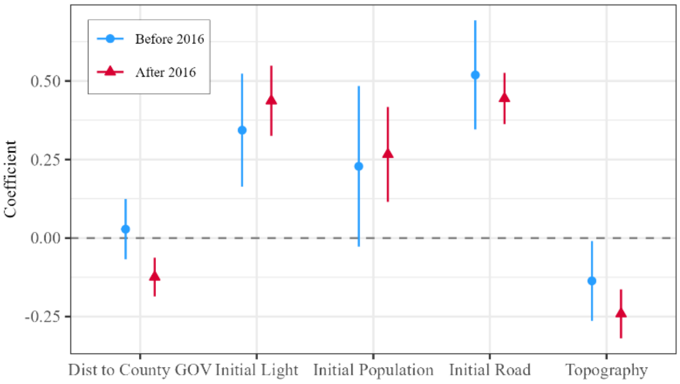
Combining the above results, we provide a further validation to our top-down hypothesis by examining how the characteristics of townships that establish TBVs change in response to the top-down directives. Specifically, as shown in Fig 8, townships that are closer to the county government and have better baseline conditions are more likely to be affected negatively by the establishment of TBVs. If the top-down channel indeed exists, then we should find that townships with these characteristics are more likely to establish TBVs after 2016, and in prefectures with higher emphasis on rural e-commerce and rural revitalization.

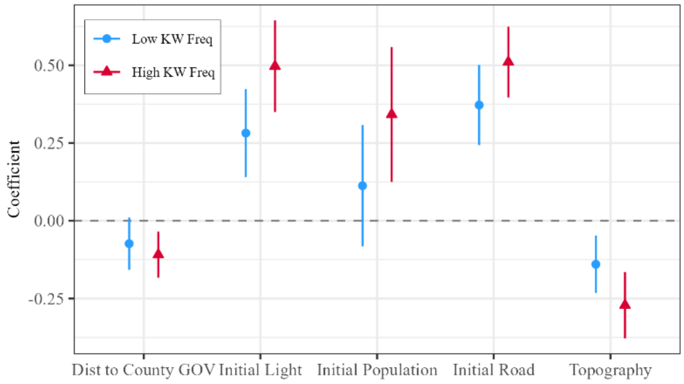
We verify this hypothesis in Figure 9, where we regress the characteristics of townships on the treatment dummy, and show how the differences in baseline characteristics between treatment and control groups are shaped by the top-down directives. As above, we measure the strength of top-down directives in two ways. In the left panel of Figure 9, we find that TBVs established after 2016 are significantly closer to the county government, compared with those established before 2016. TBVs established after 2016 also have better economic conditions (e.g., higher light intensity, more population, and less rugged terrain at baseline), though the difference are generally not statistically significant.

In the right panel of Figure 9, we measure the strength of top-down directives by the average keyword frequencies on rural e-commerce and rural revitalization. We again find consistent evidence showing that TBVs established in prefectures with higher keyword frequencies are generally have better initial conditions. This lends further support to our hypothesis that under top-down directives, local officials are more incentivized to seek townships with better economic conditions for policy experiments (Figure 9), but such a strategic selection may fail to account for the suitability in local conditions for conducting policy experiments, which may create substantial distortion and

[47] See, for example, https://view.inews.qq.com/k/20200822A0JBGX00?no-redirect=1&web_channel=wap&openApp=false.

lead to negative economic outcomes (Figure 8).





Panel A Panel B

Note: This figure presents the effects of top-down directives on the selection of TBVs. We measure the strength of top-down directives in two ways. In panel A, we use time-series variation (i.e., before and after 2016) to measure the strength of top-down directives. In Panel B, we use cross-sectional variation (i.e., high versus low keyword frequencies) to measure the strength of top-down directives.

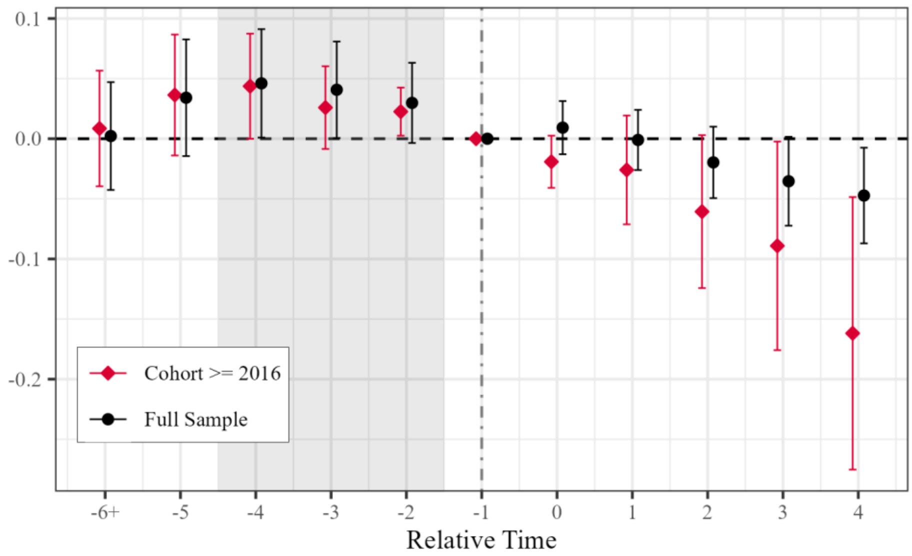
Figure 9. Top-down directives and change in TBVs' characteristics

5.5 Road Investment & Firm Entry

Although the above exercises have verified our argument that top-down enforcement as well as the political incentives of local governments can substantially moderate the negative impacts we estimate, it is still not clear about the exact mechanism through which local officials distort the local economy. By carefully scrutinizing the government reports and related documents, we find that local governments intervene in the formation of TBVs mainly through investment in infrastructure. For example, one of the key tasks emphasized by the central government is to strengthen the construction of rural roads to improve the capacity of logistics. We thus assume that the construction of roads in a given township can largely reflect the local government's investment behaviors. Several reasons justify the exploitation of road investment to proxy for the investment behaviors of local government (especially county government, as they are the final policy executors). First, local officials have long been relying on infrastructure investment to promote economic growth. Second, under China's financial system, the financial revenue of township governments is mostly governed by county governments, the so-called "county management of township finances (Xiang Cai Xian Guan)" institution. Correspondingly, the decision on road investment is largely determined by county governments and is therefore largely exogenous to the township governments, which is more of a reflection of how the county governments, as the agents, respond to their principal's advocates.

To depict how road investment changes upon the establishment of TBVs, we adopt the event-study approach and visually present the results in Figure 10, where the dependent variable is the logarithm of road length. Since local governments' responses may be more pronounced after 2016, we plot results from both full sample and a restricted sample that only incorporates the treatment cohort treated after 2016 and the control group.

Figure 10 delivers some interesting patterns. First, we find that after the establishment of TBVs, road construction in the township decreased significantly, and this downward trend is mainly concentrated in the treatment townships where TBVs were established after 2016. Second, we find a moderate increase in road construction before the township establishes TBVs. The estimated coefficients in periods -4, -3, and -2 are jointly significant at least at the 10% level. This is in line with our hypothesis that local governments would increase investment to establish TBVs. However, the significant drop in estimates after the TBVs are established suggests that local governments reduce their investment once TBVs are established (i.e., quantity requirements are met), which reveals how changes in road investment distort the local economy. Meanwhile, given the positive relationship between road infrastructure and economic growth, the results in Figure 10 shed some light on the potential mechanisms underlying our baseline results.[48]



Note: This figure displays point estimates and the corresponding 95% confidence intervals from the event study as specified in Equation (2), with a time window ranging from 6 years before to 4 years after the establishment of TBV. Standard errors are clustered at the township level. The period just before establishment (i.e., $\tau = -1$) is set to be the reference period and thus omitted from the regression. The x-axis represents the period relative to the establishment, while the y-axis represents the estimated coefficients. We simultaneously display two estimated results from the full sample as well as a restricted sample that only incorporates the control group and treatment cohort treated after 2016. All regressions include township fixed effects and county-by-year fixed effects, as well as treatment-specific year trend and control variables interacted with a full set of polynomial functions of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast.

Figure 10. Event study estimates for road investment

A potential concern is that when roads are leveraged to measure infrastructure investment, the investment decision in the current period is determined by the stock of past periods. Thus the downward trend in ex-post investment that we observe may simply be because ex-ante road levels have reached a more consensual level. However, by comparing the estimated coefficients ex-ante and ex-post (i.e., summing the estimated coefficients over all relative periods), we find the net investment is negative,[49] that is, it falls more ex-post than it rises ex-ante. This suggests that net investment by local governments indeed falls, thus further supporting our conjecture.

Following Fang, Liu, and Zhou (2023), we summarize the estimated results from Figure 10 by separately including the preparation periods to establish TBVs (i.e., within 4 periods before the establishment of TBVs). The regression specification is given by:

$$y_{it} = \beta_1 PreTBV_{it} + \beta_2\, TBV_{it} + \boldsymbol{Z}_i \times f(T) + treat_i \times T_t + \delta_i + \mu_{jt} + \epsilon_{it} \qquad (3)$$

Where $PreTBV_{it}$ denotes periods just 4 periods before the establishment of TBVs (the shaded periods in Figure 10). The other variables are defined similarly as those in Equation (1). The results

48 In our sample, the simple log-log correlation between nighttime light and road length is 0.227, which is significant at the 1% level. When we estimate using equation (1), i.e., controlling for treatment-specific time trends, township fixed effects, and county-year fixed effects, the elasticity coefficient is still significant at the 1% level.

49 This result is not significant for the full sample. However, for the restricted sample that establish TBVs after 2016, the estimated coefficient is -0.220 and the standard error is 0.120 (significant at the 10% level).

from equation (3) are reported in Table 6, where in columns (1) to (2), we report the estimates using the full sample, as well as estimates for separate treatment groups (i.e., cohorts treated before and after 2016) in the subsequent 4 columns. Consistent with the graphical results presented in Figure 10, we find a significant increase in road construction in periods that is close to the treatment, accompanied by a salient decrease after the establishment of TBVs. More interestingly, such fluctuation effect in road construction is entirely driven by cohorts treated after 2016, whereas for cohorts treated before 2016, we do not observe any differences in road investment between treatment and control groups. This is in line with our expectation since most TBVs established before 2016 were through a bottom-up approach.[50]

Table 6 Changes in road construction upon the establishment of TBVs

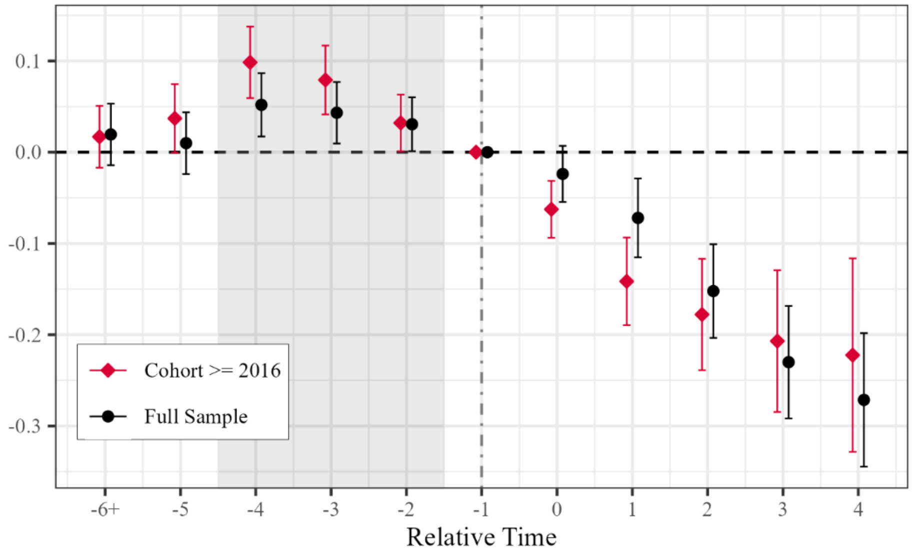
Dep. Var.: log (road)	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample		Cohort >= 2016		Cohort < 2016	
PreTBV	0.019**	0.019**	0.015*	0.015*	0.002	0.004
	(0.009)	(0.009)	(0.008)	(0.008)	(0.032)	(0.032)
TBV	-0.021*	-0.021*	-0.030**	-0.030**	-0.006	-0.008
	(0.012)	(0.012)	(0.014)	(0.014)	(0.028)	(0.028)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	No	Yes	No	Yes	No	Yes
Dep. Var. Mean	3.309	3.309	3.301	3.301	3.284	3.284
# of Clusters	31530	31530	31183	31183	30302	30302
Observations	283,770	283,770	280,647	280,647	272,718	272,718
Adjusted R-squared	0.865	0.865	0.864	0.864	0.863	0.863

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is the logarithm value of road length. The sample period is 2012-2020. Specification is listed in equation (3). All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

We now discuss how firm entries respond to different TBV formation mechanisms. Rational entrepreneurs make entry decisions based on the potential profitability, which may be affected by local economic activities as well as government support (e.g., subsidies, grants, tax deductions, etc.). By examining how firm entries are affected by TBV establishment, we resort to providing further evidence on how local governments' strategic behaviors cast distortions to the local economy. Following the same event-study approach in Figure 10, we plot the change in the number of firm entries before and after the policy for the full sample and the restricted sample in Figure 11. The results in Figure 11 exhibit almost the same pattern as in Figure 10: the number of firm entries rises before the establishment (the estimated coefficients for periods -4,-3, and -2, are all jointly significant at the 5% level), while the it declines dramatically after TBV's establishment. Meanwhile, this pattern is more prominent in the restricted sample where TBVs were established after 2016. Taken together, the above two exercises (i.e., Figure 10 and Figure 11) also provide additional support to our baseline findings that the decrease in night light intensity of treated townships is indeed caused by

[50] As an additional robustness check, we exclude the sample of townships located at country borders and replicate the regressions in Table 6 in Appendix Table C9. As noted before, the OSM data may have lower precision in border areas, and may therefore introduce additional measurement errors within these samples. Our results remain robust to sample exclusion. We also provide several exercise similar to our baseline robustness checks in Appendix Table C11 to ensure that our results on road construction is robust.

reduced economic activities.



Note: This figure displays point estimates and the corresponding 95% confidence intervals from the event study as specified in Equation (2), with a time window ranging from 6 years before to 4 years after the establishment of TBV. Standard errors are clustered at the township level. The period just before establishment (i.e., $\tau = -1$) is set to be the reference period and thus omitted from the regression. The x-axis represents the period relative to the establishment, while the y-axis represents the estimated coefficients. We simultaneously display two estimated results from the full sample as well as a restricted sample that only incorporates the control group and treatment cohort treated after 2016. All regressions include township fixed effects and county-by-year fixed effects, as well as treatment-specific year trend and control variables interacted with a full set of polynomial functions of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast.

Figure 11. Event study estimates for firm entry

Table 7 reports the results of DiD estimation using equation (3). Same as Figure 11, we find that the negative effects of firm entry in posttreatment periods, as well as the positive effects in pretreatment periods are completely driven sample of TBVs established after 2016, and we also document a significant increase in the number of firm entries using cohorts treated before 2016, which may also provide a potential explanation for the positive results estimated in Fig 7. Analogously, the decrease in firm entries also partially explains why the establishment eventually hinders local economic growth.[51]

In Appendix Table C12, we examine the effects on firm entries on a subset of industries, which is closely related to the development of TBVs. Since most of the TBVs engage in production and processing in light industry and agriculture, we focus our attention to scrutinize how the establishment of TBVs affect the entry of firms in these specific industries.[52] The results shown in Table C12 generally align with the pattern shown in Table 7. Typically, for TBVs established after 2016, we find a significant increase of firm entries prior to the TBV establishment, the pattern reversed drastically soon after the TBV establishment. This further strengthens our confidence that the entry

[51] To avoid potential bias that arises when taking logs with respect to variables with zero values (Chen and Roth 2023), in Appendix Table C10 we report PPML estimates. Although the results for the full sample are no longer significant under the PPML regression, the post-2016 cohort-estimations remain significantly negative, while the results for pre-2016 cohorts are significantly positive. We also provide several exercise similar to our baseline robustness checks in Appendix Table C11 to ensure that our results on road construction is robust.

[52] These industries include Agro-food Processing Industry, Wine, Beverage and Refined Tea Manufacturing Industry, Textile and Clothing Industry, Wood Processing and Products Manufacturing Industry, Furniture Manufacturing Industry, Leather Products and Footwear Industry'.

dynamics of firms are indeed induced by the establishment of TBVs.

Table 7 Changes in firm entry upon the establishment of TBVs

Dep. Var.: log (# firm entry+1)	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample		Cohort >= 2016		Cohort < 2016	
PreTBV	0.069***	0.066***	0.074***	0.071***	-0.010	-0.012
	(0.012)	(0.012)	(0.013)	(0.013)	(0.031)	(0.031)
TBV	-0.048***	-0.043**	-0.110***	-0.103***	0.155***	0.151***
	(0.017)	(0.017)	(0.019)	(0.019)	(0.044)	(0.044)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	No	Yes	No	Yes	No	Yes
Dep. Var. Mean	0.338	0.338	0.332	0.332	0.322	0.322
# of Clusters	31530	31530	31183	31183	30302	30302
Observations	346,830	346,830	343,013	343,013	333,322	333,322
Adjusted R-squared	0.662	0.666	0.659	0.662	0.653	0.656

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is the logarithm value of the number of private firm entries plus 1. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

5.6 Discussion

So far, we have documented that the reducing in road construction and firm entry could potentially drive the negative impacts of TBV establishment on the local economy. In this subsection, we combine the above findings to delve deeper into the potential mechanisms that drive the unprecedented negative effects of TBV establishment. Specifically, we consider how top-down interventions, combined with the strategic behaviors caused by the political incentives of local officials, contribute to the reduction in economic activities.

Table C13 summarizes our results, where we augment the specification in equation (3) by interacting the main terms (i.e., PreTBV and TBV) with variables that measure the top-down directives (word frequencies in government report) and local officials' strategic behaviors (county leader rotations).[53] Such more generalized specification allows us to scrutinize how our results are affected by these factors. In columns (1) and (2), we find that when prefectural governments attach more weights to the development of TBVs, there will be more firm entries in the pretreatment periods, while there is no amplified effect in subsequent declines. We do not find any significant changes in road construction relative to the top-down directives. This evidence suggests that the top-down directives indeed facilitates the establishment of TBVs by attracting the entry of more businesses. In columns (3) and (4), we find that the rotation of county leader significantly reduces the road constructions and firm entries after the TBVs are established. And we do not find evidence indicating that the rotation leads to more pretreatment road construction or firm entry. This is in line with our hypothesis that the local leaders are less willing to allocate resources to support the further development of TBVs (i.e., improving the quality) when they prioritize promoting economic growth in urban sectors, particularly when they have stronger promotion incentives.

Taken together, the results from Table C13 suggest that the top-down interventions along do

[53] Specifically, we consider the word frequencies of both rural e-commerce and rural revitalization (as in columns (5) and (6) of Table 3).

not necessarily induce negative effects or distortions to the local economy, but such pervasive effect could emerge when it is combined with the strategic behaviors of local officials, which originate from the meritocracy-based political selection process that prioritizes the economic performance.

6. Concluding Remarks

In this paper, we assess for the first time the impact of Taobao village establishment on the township economy at a national scale using detailed township data. We find that, in general, the establishment of Taobao villages has negative economic impacts on townships (though a small subset of TBVs indeed leads to positive effects). Our results are particularly noteworthy since they stand in sharp contrast with the findings of earlier papers. We believe that the main reason for this result is the incomplete sample and the neglect to consider the strategic behavior of local governments. In the presence of political competition and political incentives, local governments resort to top-down directives to establish Taobao villages. We test the heterogeneity of the estimation results and find that in the temporal dimension, the negative effect emerges only after the central government's advocation, while in the cross-sectional dimension, the negative effect is mainly concentrated in the places where local governments pay more attention to the development of rural e-commerce. Further analysis shows that local governments are more inclined to select townships with better initial conditions to be Taobao village pilots, and local governments with better economic and fiscal capacities are more inclined to use top-down intervention. Since the number of Taobao villages is more easily quantifiable, we argue that local governments have more incentives to use top-down directives to promote the formation of Taobao villages and less incentives to further maintain the development of Taobao villages. Using additional road data as well as business registration data, we further verify the hypothesis.

From a broader perspective, our analyses and findings also emphasize the heterogeneity of effects between top-down and bottom-up policy implementation. As an authoritarian state, China's top-down administrative system is conducive to strengthening the state's control over society, but when this institution is used to promote local economic development, the inability to adequately consider the local heterogeneity and the difficulty of effectively utilizing local information may inevitably result in allocation inefficiency. What's worse, as our study shows, it may cause severe distortions due to the strategic behavior of local governments. Failure to distinguish between top-down and bottom-up implementations may have serious economic consequences. In addition, our study also shows that although the analysis of Taobao-village-specific case and survey data can reveal the formation mechanism of Taobao villages and the effects on household welfare, the heterogeneity in treatment effects makes it hard to clearly describe the aggregate economic effects of Taobao villages. Unfortunately, due to data limitations, we are unable to discuss in depth the specific sources and channels of distortions caused by local governments to the local economy. Further studies can extend the analysis based on our paper.

References

Alesina, Alberto, and Guido Tabellini. 2007. "Bureaucrats or Politicians? Part I: A Single Policy Task." *American Economic Review* 97 (1): 169–79. https://doi.org/10.1257/aer.97.1.169.

———. 2008. "Bureaucrats or Politicians? Part II: Multiple Policy Tasks." *Journal of Public Economics* 92 (3–4): 426–47. https://doi.org/10.1016/j.jpubeco.2007.06.004.

Allcott, Hunt. 2015. "Site Selection Bias in Program Evaluation*." *The Quarterly Journal of Economics* 130 (3): 1117–65. https://doi.org/10.1093/qje/qjv015.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–84. https://doi.org/10.1086/426036.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. 2021. "Synthetic Difference-in-Differences." *American Economic Review* 111 (12): 4088–4118. https://doi.org/10.1257/aer.20190159.

Bellemare, Marc F., and Casey J. Wichman. 2019. "Elasticities and the Inverse Hyperbolic Sine Transformation." *Oxford Bulletin of Economics and Statistics* 82 (1): 50–61. https://doi.org/10.1111/obes.12325.

Bhalotra, Sonia, Damian Clarke, Joseph Flavian Gomes, and Atheendar Venkataramani. 2023. "Maternal Mortality and Women's Political Power." *Journal of the European Economic Association* 21 (5): 2172–2208. https://doi.org/10.1093/jeea/jvad012.

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–27. https://doi.org/10.1162/rest.90.3.414.

Cao, Yuanzheng, Yingyi Qian, and Barry R. Weingast. 1999. "From Federalism, Chinese Style to Privatization, Chinese Style." *Economics of Transition* 7 (1): 103–31. https://doi.org/10.1111/1468-0351.00006.

Chaisemartin, Clément de, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–96. https://doi.org/10.1257/aer.20181169.

Chen, Jiafeng, and Jonathan Roth. 2023. "Logs with Zeros? Some Problems and Solutions." *The Quarterly Journal of Economics*, December, qjad054. https://doi.org/10.1093/qje/qjad054.

Chen, Ling, and Hao Zhang. 2021. "Strategic Authoritarianism: The Political Cycles and Selectivity of China's Tax-Break Policy." *American Journal of Political Science* 65 (4): 845–61. https://doi.org/10.1111/ajps.12648.

Chen, Xi, and William D. Nordhaus. 2011. "Using Luminosity Data as a Proxy for Economic Statistics." *Proceedings of the National Academy of Sciences* 108 (21): 8589–94. https://doi.org/10.1073/pnas.1017031108.

Chen, Xuelian. 2017. "A U-Turn or Just Pendulum Swing? Tides of Bottom-Up and Top-Down Reforms in Contemporary China." *Journal of Chinese Political Science* 22 (4): 651–73. https://doi.org/10.1007/s11366-017-9515-6.

Chen, Yvonne Jie, Pei Li, and Yi Lu. 2018. "Career Concerns and Multitasking Local Bureaucrats: Evidence of a Target-Based Performance Evaluation System in China." *Journal of Development Economics* 133 (July):84–101. https://doi.org/10.1016/j.jdeveco.2018.02.001.

Chen, Zhiwu, Chicheng Ma, and Andrew J Sinclair. 2022. "Banking on the Confucian Clan: Why China Developed Financial Markets so Late." *The Economic Journal* 132 (644): 1378–1413. https://doi.org/10.1093/ej/ueab082.

Couture, Victor, Benjamin Faber, Yizhen Gu, and Lizhi Liu. 2021. "Connecting the Countryside via E-Commerce: Evidence from China." *American Economic Review: Insights* 3 (1): 35–50. https://doi.org/10.1257/aeri.20190382.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999. "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies." *Review of Economic Studies* 66 (1): 199–217. https://doi.org/10.1111/1467-937X.00085.

———. 2000. "Multitask Agency Problems: Focus and Task Clustering." *European Economic Review* 44 (4–6): 869–77. https://doi.org/10.1016/S0014-2921(00)00059-3.

Fang, Hanming, Chang Liu, and Li-An Zhou. 2023. "Window Dressing in the Public Sector: Evidence from

China's Compulsory Education Promotion Program." *Journal of Public Economics* 222 (June):104878. https://doi.org/10.1016/j.jpubeco.2023.104878.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M. Shapiro. 2019. "Pre-Event Trends in the Panel Event-Study Design." *American Economic Review* 109 (9): 3307–38. https://doi.org/10.1257/aer.20180609.

Gibson, John, Susan Olivia, Geua Boe-Gibson, and Chao Li. 2021. "Which Night Lights Data Should We Use in Economics, and Where?" *Journal of Development Economics* 149 (March):102602. https://doi.org/10.1016/j.jdeveco.2020.102602.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77. https://doi.org/10.1016/j.jeconom.2021.03.014.

Guo, Gang. 2009. "China's Local Political Budget Cycles." *American Journal of Political Science* 53 (3): 621–32. https://doi.org/10.1111/j.1540-5907.2009.00390.x.

He, Liuyang, Xi Lu, and Chyen Yee Lee. 2023. "Last Mile in Anti-Poverty Drive: Impact of Cadres' Appraisals on Growth and Poverty Reduction." *World Development* 167 (July):106255. https://doi.org/10.1016/j.worlddev.2023.106255.

Heilmann, Sebastian. 2008. "From Local Experiments to National Policy: The Origins of China's Distinctive Policy Process." *The China Journal* 59 (January):1–30. https://doi.org/10.1086/tcj.59.20066378.

Henderson, J. Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028. https://doi.org/10.1257/aer.102.2.994.

Hjort, Jonas, and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review* 109 (3): 1032–79. https://doi.org/10.1257/aer.20161385.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *The Journal of Law, Economics, and Organization* 7 (special_issue): 24–52. https://doi.org/10.1093/jleo/7.special_issue.24.

Huang, Yi, Chen Lin, Sibo Liu, and Gustavo Manso. 2021. "The Salience of Entrepreneurship: Evidence from Online Business." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3843524.

Huang, Zhangkai, Lixing Li, Guangrong Ma, and Lixin Colin Xu. 2017. "Hayek, Local Information, and Commanding Heights: Decentralizing State-Owned Enterprises in China." *American Economic Review* 107 (8): 2455–78. https://doi.org/10.1257/aer.20150592.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20 (1): 1–24. https://doi.org/10.1093/pan/mpr013.

Kahn-Lang, Ariella, and Kevin Lang. 2020. "The Promise and Pitfalls of Differences-in-Differences: Reflections on *16 and Pregnant* and Other Applications." *Journal of Business & Economic Statistics* 38 (3): 613–20. https://doi.org/10.1080/07350015.2018.1546591.

Kung, James Kai-sing, and Shuo Chen. 2011. "The Tragedy of the Nomenklatura: Career Incentives and Political Radicalism during China's Great Leap Famine." *The American Political Science Review* 105 (1): 27–45.

Landry, Pierre F. 2008. *Decentralized Authoritarianism in China: The Communist Party's Control of Local Elites in the Post-Mao Era*. Cambridge ; New York: Cambridge University Press.

Li, Guangqin, and Jiahong Qin. 2022. "Income Effect of Rural E-Commerce: Empirical Evidence from Taobao Villages in China." *Journal of Rural Studies* 96 (December):129–40. https://doi.org/10.1016/j.jrurstud.2022.10.019.

Li, Hongbin, and Li-An Zhou. 2005. "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China." *Journal of Public Economics* 89 (9–10): 1743–62. https://doi.org/10.1016/j.jpubeco.2004.06.009.

Li, Xing, Chong Liu, Xi Weng, and Li-An Zhou. 2019. "Target Setting in Tournaments: Theory and Evidence from China." *The Economic Journal* 129 (623): 2888–2915. https://doi.org/10.1093/ej/uez018.

Lin, Justin Yifu, Fang Cai, and Zhou Li. 2003. *The China Miracle*. The Chinese University of Hong Kong Press. https://doi.org/10.2307/j.ctv1fj84hd.

Liu, Min, Qian Zhang, Song Gao, and Jikun Huang. 2020. "The Spatial Aggregation of Rural E-Commerce in China: An Empirical Investigation into Taobao Villages." *Journal of Rural Studies* 80 (December):403–17. https://doi.org/10.1016/j.jrurstud.2020.10.016.

Luo, Xubei, and Chiyu Niu. 2019. *E-Commerce Participation and Household Income Growth in Taobao Villages*. World Bank, Washington, DC. https://doi.org/10.1596/1813-9450-8811.

Ma, Jinji, Jinyu Guo, Safura Ahmad, Zhengqiang Li, and Jin Hong. 2020. "Constructing a New Inter-Calibration Method for DMSP-OLS and NPP-VIIRS Nighttime Light." *Remote Sensing* 12 (6): 937.

https://doi.org/10.3390/rs12060937.

Martinez-Bravo, Monica, Gerard Padró I Miquel, Nancy Qian, and Yang Yao. 2022. “The Rise and Fall of Local Elections in China.” *American Economic Review* 112 (9): 2921–58. https://doi.org/10.1257/aer.20181249.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9780511808678.

Nunn, Nathan, and Diego Puga. 2012. “Ruggedness: The Blessing of Bad Geography in Africa.” *Review of Economics and Statistics* 94 (1): 20–36. https://doi.org/10.1162/REST_a_00161.

Oates, Wallace E. 1972. *Fiscal Federalism*. The Harbrace Series in Business and Economics. New York: Harcourt Brace Jovanovich.

———. 1999. “An Essay on Fiscal Federalism.” *Journal of Economic Literature* 37 (3): 1120–49.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 37 (2): 187–204. https://doi.org/10.1080/07350015.2016.1227711.

Qi, Jiaqi, Xiaoyong Zheng, and Hongdong Guo. 2019. “The Formation of Taobao Villages in China.” *China Economic Review* 53 (February):106–27. https://doi.org/10.1016/j.chieco.2018.08.010.

Qian, Yingyi, and Barry R. Weingast. 1997. “Federalism as a Commitment to Perserving Market Incentives.” *The Journal of Economic Perspectives* 11 (4): 83–92.

Rambachan, Ashesh, and Jonathan Roth. 2023. “A More Credible Approach to Parallel Trends.” *The Review of Economic Studies*, February, rdad018. https://doi.org/10.1093/restud/rdad018.

Rawski, Thomas G. 1995. “Implications of China’ Reform Experience.” *The China Quarterly* 144:1150–73. https://doi.org/10.1017/S030574100000477X.

Roth, Jonathan. 2022. “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends.” *American Economic Review: Insights* 4 (3): 305–22. https://doi.org/10.1257/aeri.20210236.

Sant’Anna, Pedro H.C., and Jun Zhao. 2020. “Doubly Robust Difference-in-Differences Estimators.” *Journal of Econometrics* 219 (1): 101–22. https://doi.org/10.1016/j.jeconom.2020.06.003.

Scott, James C. 1998. *Seeing Like a State*. Yale University Press. http://www.jstor.org/stable/j.ctt1nq3vk.

Suárez Serrato, Juan Carlos, Xiao Yu Wang, and Shuang Zhang. 2019. “The Limits of Meritocracy: Screening Bureaucrats under Imperfect Verifiability.” *Journal of Development Economics* 140 (September):223–41. https://doi.org/10.1016/j.jdeveco.2019.06.003.

Sun, Liyang, and Sarah Abraham. 2021. “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects.” *Journal of Econometrics* 225 (2): 175–99. https://doi.org/10.1016/j.jeconom.2020.09.006.

Tang, Can, and Zhong Zhao. 2023. “Informal Institution Meets Child Development: Clan Culture and Child Labor in China.” *Journal of Comparative Economics* 51 (1): 277–94. https://doi.org/10.1016/j.jce.2022.09.006.

Tang, Weicheng, and Jin Zhu. 2020. “Informality and Rural Industry: Rethinking the Impacts of E-Commerce on Rural Development in China.” *Journal of Rural Studies* 75 (April):20–29. https://doi.org/10.1016/j.jrurstud.2020.02.010.

Vivalt, Eva. 2020. “How Much Can We Generalize From Impact Evaluations?” *Journal of the European Economic Association* 18 (6): 3045–89. https://doi.org/10.1093/jeea/jvaa019.

Wang, Cassandra C., Julie T. Miao, Nicholas A. Phelps, and Jia Zhang. 2021. “E-Commerce and the Transformation of the Rural: The Taobao Village Phenomenon in Zhejiang Province, China.” *Journal of Rural Studies* 81 (January):159–69. https://doi.org/10.1016/j.jrurstud.2020.10.017.

Wang, Shaoda, and David Yang. 2021. “Policy Experimentation in China: The Political Economy of Policy Learning.” w29402. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w29402.

Wang, Zhi, Qinghua Zhang, and Li-An Zhou. 2020. “Career Incentives of City Leaders and Urban Spatial Expansion in China.” *The Review of Economics and Statistics* 102 (5): 897–911. https://doi.org/10.1162/rest_a_00862.

Xi, Tianyang, Yang Yao, and Muyang Zhang. 2018. “Capability and Opportunism: Evidence from City Officials in China.” *Journal of Comparative Economics* 46 (4): 1046–61. https://doi.org/10.1016/j.jce.2018.04.001.

Xu, Chenggang. 2011. “The Fundamental Institutions of China’s Reforms and Development.” *Journal of Economic Literature* 49 (4): 1076–1151. https://doi.org/10.1257/jel.49.4.1076.

Zang, Yuzhu, Shougeng Hu, Bing-bing Zhou, Ligang Lv, and Xueyan Sui. 2023. “Entrepreneurship and the Formation Mechanism of Taobao Villages: Implications for Sustainable Development in Rural Areas.” *Journal of Rural Studies* 100 (May):103030. https://doi.org/10.1016/j.jrurstud.2023.103030.

Appendix A

Table A1 List of industries excluded from IRSB data

Industry	Industry Code	Industry	Industry Code
Coal mining and washing	06	Real Estate	70
Oil and gas extraction	07	Leasing	71
Ferrous Metal Mining and Processing	08	Business Services	72
Nonferrous Metal Mining and Processing	09	Research and Experimental Development	73
Non-metallic Mining and Processing	10	Professional and Technical Services	74
Mining specialties and auxiliary activities	11	Science and Technology Promotion and Application Services	75
Other Mining	12	Water Resources Management	76
Petroleum, coal and other fuel processing industry	25	Ecological Protection and Environmental Management	77
Chemical raw materials and chemical products manufacturing	26	Public Facilities Management	78
Pharmaceutical Manufacturing	27	Land Management	79
Chemical Fiber Manufacturing	28	Residential Services	80
Rubber and Plastic Products	29	Repair of Motor Vehicles, Electronic Products and Household Products	81
Non-metallic mineral products industry	30	Other Services	82
Ferrous metal smelting and rolling processing industry	31	Education	83
Non-ferrous metal smelting and rolling processing industry	32	Health	84
Metal Products Industry	33	Social work	85
General Equipment Manufacturing	34	News and Publishing	86
Specialty Equipment Manufacturing	35	Radio, television, film and sound recording production industry	87
Automobile Manufacturing	36	Culture and Arts	88
Railroad, ship, aerospace and other transportation equipment manufacturing industry	37	Sports	89
Electrical machinery and equipment manufacturing	38	Entertainment Industry	90
Computer, communications and other electronic equipment manufacturing	39	Communist Party of China (CPC) Organs	91
Instrumentation Manufacturing	40	State Institutions	92
Other manufacturing industries	41	People's Political Consultative Conference, Democratic Party	93
Comprehensive utilization of waste resources	42	Social Security	94
Metal products, machinery and equipment repair industry	43	Mass organizations, social groups and other member organizations	95
Electricity, heat production and supply industry	44	Grassroots Mass Self-Governance Organizations	96
Gas Production and Supply	45	International Organizations	97
Water production and supply	46		

Appendix B

Table B1 Frequency of specific keywords in government annual reports

Keywords related to rural e-commerce	Keywords related to rural revitalization	Keywords related to general e-commerce	Keywords related to agricultural industrial policies
Rural e-commerce development	Rural revitalization	E-commerce development	Rural Finance
Rural network sales	Rural development	digital economy	Agricultural Insurance
Rural e-commerce support policy	Rural infrastructure construction	E-commerce platform	Agricultural science and technology innovation
Agricultural products e-commerce	Rural poverty alleviation	Online Retail	Food security
Rural e-commerce platform construction	Rural Income Increase	E-payment	Agricultural Product Processing
Rural e-commerce industry chain	Rural Habitat Environment	Cross-border E-commerce	Agricultural Modernization
Rural E-commerce Innovation Model	Rural social governance	Internet + Industry	Agricultural Industry
Comprehensive Demonstration of E-commerce in Rural Areas	Rural employment		Leading Agricultural Enterprises
E-commerce Poverty Alleviation			Agricultural Technology Promotion
Digital Countryside			Agricultural Cooperatives

Appendix C: Additional Regression Results

Table C1 Baseline results using standardized light as dependent variable

Dep. Var.	Std. Light					
	(1)	(2)	(3)	(4)	(5)	(6)
TBV	-0.008	-0.010	-0.011*	-0.011*	-0.012*	-0.013*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Linear Trend	-	Yes	Yes	Yes	Yes	Yes
Controls × Quadratic Trend	-	-	Yes	Yes	Yes	Yes
Controls × Cubic Trend	-	-	-	Yes	Yes	Yes
Inverse Probability Weighting	-	-	-	-	Yes	-
Coarsened Exact Matching						Yes
# of Clusters	31530	31530	31530	31530	30890	14884
Observations	346,830	346,830	346,830	346,830	339,790	163,724
Adjusted R-squared	0.980	0.981	0.981	0.981	0.982	0.977

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. The dependent variable is standardized value of night light luminosity. All regressions control for township fixed effects, county-by-year fixed effects, and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

Table C2 Baseline results using DiD_{CS} estimator

Dep. Var. log(light+1)	(1)	(2)	(3)	(4)
TBV	-0.027***	-0.038***	-0.038***	-0.038***
	(0.007)	(0.008)	(0.008)	(0.008)
Township FE	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes
Controls× $f(T)$	-	Yes	Yes	Yes
Aggregation Method	OLS	OLS	Doubly Robust (DR)	Improved DR
Dep. Var. Mean	1.552	1.552	1.552	1.552
# of Clusters	31530	31530	31530	31530
Observations	346,830	346,830	346,830	346,830

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

Table C3 Baseline results using synthetic DiD estimator

Dep. Var. log(light+1)	(1)	(2)	(3)	(4)
TBV	-0.043***	-0.044***	-0.041***	-0.042***
	(0.007)	(0.007)	(0.007)	(0.007)
Dep. Var. Mean	1.552	1.552	1.552	1.552
Township FE	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes
Controls × Linear Trend	-	Yes	Yes	Yes
Controls × Quadratic Trend	-	-	Yes	Yes
Controls × Cubic Trend	-	-	-	Yes
Observations	346,830	346,830	346,830	346,830

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. The bootstrap standard errors are reported in the parentheses.

Table C4 Regression results using village sample

Dep. Var.	log(light+1)			IHS(light)		
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Cohort < 2016	Cohort >= 2016	Full Sample	Cohort < 2016	Cohort >= 2016
TBV	-0.015*	-0.018	-0.079**	-0.020*	-0.021	-0.103**
	(0.009)	(0.032)	(0.039)	(0.011)	(0.037)	(0.049)
Observations	45,710	45,710	45,710	45,710	45,710	45,710
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls$\times f(T)$	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the village-year level. All regressions control for village fixed effects, county-by-year fixed effects and treatment-specific year trend. The construction of village sample is outlined in Appendix D. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors clustered at the village level are reported in the parentheses.

Table C5 Robustness to different radius

Dep. Var.	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)	log(light+1)	IHS(light)
Radius	10KM		20KM		30KM		40KM		50KM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
TBV	-0.021**	-0.027**	-0.025***	-0.032***	-0.027***	-0.034***	-0.028***	-0.036***	-0.029***	-0.037***
	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	2.369	2.851	2.197	2.649	2.057	2.485	1.965	2.375	1.893	2.290
# of Clusters	10250	10250	14119	14119	16970	16970	19198	19198	20921	20921
Observations	112,750	112,750	155,309	155,309	186,670	186,670	211,178	211,178	230,131	230,131
Adjusted R-squared	0.973	0.971	0.973	0.970	0.973	0.970	0.972	0.970	0.972	0.970

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the town-year level. All regressions control for township fixed effects, county-by-year fixed effects, treatment-specific year trend and control variables interacted with a 3-order polynomial function of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. In odd columns, the dependent variable is log (1+light), whereas in even columns, the dependent variable is the IHS transformation of night light luminosity. Standard errors reported in the parentheses are clustered at the township level.

Table C6 Robustness to Poisson transformation: night light

Dep. Var. light	(1)	(2)	(3)	(4)	(5)
TBV	-0.026***	-0.022***	-0.021***	-0.021***	-0.017***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Dep. Var. Mean	13.05	13.05	13.05	13.05	13.10
Township FE	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes
Controls × Linear Trend	-	Yes	Yes	Yes	Yes
Controls × Quadratic Trend	-	-	Yes	Yes	Yes
Controls × Cubic Trend	-	-	-	Yes	Yes
Inverse Probability Weighting	-	-	-	-	Yes
Number of Clusters	29761	29761	29761	29761	29140
Observations	326,698	326,698	326,698	326,698	319,865
Pseudo R-squared	0.885	0.886	0.886	0.886	0.869

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the town-year level. All regressions control for township fixed effects, county-by-year fixed effects, treatment-specific year trend and control variables interacted with a 3-order polynomial function of year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. In odd columns, the dependent variable is log (1+light), whereas in even columns, the dependent variable is the IHS transformation of night light luminosity. Standard errors reported in the parentheses are clustered at the township level.

Table C7 "First stage" effects of key word frequency on TBV establishment

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	TBV	# of TBV	TBV	# of TBV	TBV	# of TBV
	Full Sample		Cohort >= 2016		Cohort < 2016	
KW: Rural E-commerce	0.005***	0.016**	0.004***	0.012**	0.001	0.005
	(0.002)	(0.007)	(0.002)	(0.005)	(0.000)	(0.004)
KW: Rural Revitalization	-0.003	-0.015	-0.003	-0.013**	0.000	-0.003
	(0.002)	(0.011)	(0.002)	(0.006)	(0.001)	(0.007)
Total Frequency	0.001	0.008	0.002	0.007	-0.000	0.001
	(0.002)	(0.010)	(0.002)	(0.005)	(0.001)	(0.006)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.0155	0.0476	0.0155	0.0476	0.0155	0.0476
# of Clusters	340	340	340	340	340	340
Observations	347,017	347,017	343,211	343,211	333,531	333,531
Adjusted R-squared	0.372	0.363	0.253	0.267	0.539	0.409

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is either the dummy variable of TBV establishment (in odd columns) or the number of TBVs established (in even columns). The sample period is 2010-2020. All observations are at the township-year level. All regressions control for township fixed effects and year fixed effects. Standard errors reported in the parentheses are clustered at the prefecture level.

Table C8 "First stage" effects of lagged key word frequency on TBV establishment

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	TBV	# of TBV	TBV	# of TBV	TBV	# of TBV
	Full Sample		Cohort >= 2016		Cohort < 2016	
L1.KW: Rural E-commerce	0.005**	0.019**	0.005**	0.014**	0.001	0.006
	(0.002)	(0.009)	(0.002)	(0.006)	(0.001)	(0.004)
L1.KW: Rural Revitalization	0.002	0.003	0.001	-0.001	0.001	0.005
	(0.003)	(0.013)	(0.003)	(0.008)	(0.001)	(0.007)
L1.Total Frequency	-0.002	-0.007	-0.001	-0.001	-0.001	-0.006
	(0.003)	(0.011)	(0.002)	(0.007)	(0.001)	(0.007)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var. Mean	0.0155	0.0476	0.0155	0.0476	0.0155	0.0476
# of Clusters	340	340	340	340	340	340
Observations	315,470	315,470	312,010	312,010	303,210	303,210
Adjusted R-squared	0.409	0.402	0.278	0.294	0.598	0.454

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is either the dummy variable of TBV establishment (in odd columns) or the number of TBVs established (in even columns). The sample period is 2010-2020. All observations are at the township-year level. All regressions control for township fixed effects and year fixed effects. Standard errors reported in the parentheses are clustered at the prefecture level.

Table C9 Robustness to data reliability on road density

Dep. Var. log (road)	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample		Cohort >= 2016		Cohort < 2016	
PreTBV	0.023**	0.023**	0.017**	0.017**	0.009	0.011
	(0.010)	(0.010)	(0.008)	(0.008)	(0.033)	(0.033)
TBV	-0.021*	-0.021*	-0.036**	-0.035**	0.009	0.008
	(0.013)	(0.013)	(0.016)	(0.016)	(0.029)	(0.029)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls$\times f(T)$	No	Yes	No	Yes	No	Yes
Dep. Var. Mean	3.282	3.282	3.282	3.282	3.282	3.282
# of Clusters	31074	31074	30732	30732	29851	29851
Observations	279,666	279,666	276,588	276,588	268,659	268,659
Adjusted R-squared	0.844	0.844	0.843	0.843	0.841	0.841

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is logarithm value of road length. The sample period is 2012-2020. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

Table C10 Robustness to Poisson transformation: firm entry

Dep. Var. # firm entry	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample		Cohort >= 2016		Cohort < 2016	
TBV	0.031	0.022	-0.366***	-0.358***	0.285***	0.284***
	(0.069)	(0.067)	(0.131)	(0.128)	(0.067)	(0.068)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	No	Yes	No	Yes	No	Yes
Dep. Var. Mean	1.179	1.179	1.179	1.179	1.179	1.179
# of Clusters	23954	23954	23616	23616	22765	22765
Observations	178,711	178,711	175,756	175,756	168,466	168,466
Adjusted R-squared	0.707	0.707	0.702	0.702	0.692	0.692

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is the number of firm entries. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitati on (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

Table C11 Additional robustness checks on road construction and firm entry

	(1) Using log(# TBV+1)	(2) Population as weight	(3) Cluster at county level	(4) Radius <=30	(5) Using log(# TBV+1)	(6) Population as weight	(7) Cluster at county level	(8) Radius <=30
Dep. Var.	log(road)				log(# firm entry+1)			
PreTBV	0.014	0.021**	0.021**	0.022**	0.012	0.048***	0.049***	0.048***
	(0.011)	(0.009)	(0.010)	(0.010)	(0.013)	(0.013)	(0.012)	(0.013)
log(# TBV+1)	-0.041**				-0.151***			
	(0.017)				(0.019)			
TBV		-0.018	-0.022*	-0.021*		-0.064***	-0.066***	-0.066***
		(0.012)	(0.012)	(0.013)		(0.018)	(0.018)	(0.017)
Dep. Var. Mean	3.296	3.296	3.296	3.296	0.338	0.338	0.338	0.338
Township FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls× $f(T)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Clusters	31530	31513	2818	16970	31530	31515	2818	16970
Observations	283,770	283,553	283,770	152,730	346,830	346,569	346,830	186,670
Adjusted R-squared	0.846	0.854	0.846	0.859	0.666	0.675	0.666	0.678

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend.

Table C12 Changes in Taobao related firm entry upon the establishment of TBVs

Dep. Var.: log (# firm entry+1)	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample		Cohort >= 2016		Cohort < 2016	
PreTBV	0.043***	0.041***	0.063***	0.061***	-0.103***	-0.102***
	(0.009)	(0.009)	(0.011)	(0.011)	(0.025)	(0.025)
TBV	-0.041***	-0.039**	-0.105***	-0.102***	0.187***	0.183***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.042)	(0.042)
Township FE	Yes	Yes	Yes	Yes	Yes	Yes
County-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Treatment Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls$\times f(T)$	No	Yes	No	Yes	No	Yes
Dep. Var. Mean	0.134	0.134	0.129	0.129	0.123	0.123
# of Clusters	31530	31530	31183	31183	30302	30302
Observations	346,830	346,830	343,013	343,013	333,322	333,322
Adjusted R-squared	0.493	0.495	0.478	0.480	0.462	0.464

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is the logarithm value of the number of private firm entries plus 1. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

Table C13 Effects of top-down directives & political incentives on road and firm entry

	(1)	(2)	(3)	(4)
Dep. Var.	log (road)	log (# firm entry+1)	log(road)	log (# firm entry+1)
PreTBV×High KW	0.001	0.047**		
	(0.020)	(0.021)		
TBV×High KW	-0.014	-0.009		
	(0.020)	(0.021)		
PS Turnover×PreTBV			-0.058	-0.009
			(0.038)	(0.032)
PS Turnover×TBV			-0.095*	-0.078*
			(0.052)	(0.043)
PreTBV	0.028	0.039**	0.086***	0.068***
	(0.018)	(0.017)	(0.028)	(0.023)
TBV	-0.005	-0.036	0.026	0.008
	(0.018)	(0.022)	(0.034)	(0.034)
Dep. Var. Mean	3.296	0.338	3.199	0.309
Township FE	Yes	Yes	Yes	Yes
County by Year FE	Yes	Yes	Yes	Yes
Treatment Linear Trend	Yes	Yes	Yes	Yes
Controls× $f(T)$	Yes	Yes	Yes	Yes
Number of Clusters	31530	31530	15088	15088
Observations	283,770	346,830	135,792	165,968
Adjusted R-squared	0.846	0.666	0.830	0.632

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable is the logarithm value of road length, as well as the number of private firm entries plus 1. All observations are at the township-year level. All regressions control for township fixed effects, county-by-year fixed effects and treatment-specific year trend. The control variables include topography, ruggedness, average slope, agricultural potential yields (2000), precipitation (2000), sunlight duration (2000), distance to the county government, distance to Hangzhou and distance to the coast. All controls are interacted with a flexible polynomial function of year trend. Standard errors reported in the parentheses are clustered at the township level.

Appendix D: Construction of Village Panel

Since there is no currently publicly available village-level geo-referenced information (e.g., village-level shape files), we can only construct the village panel based on the provided longitude and latitude. What's more, as we only have accessed village-level information of TBVs, we can only construct the village panel based on this subset of villages. Collecting the geo-information on other villages is feasible but it will take a substantial amount of effort and may not apply to the approach we proposed in the following. We therefore focus our analysis only on TBVs.

To construct the village panel, we first use the geocoding techniques of Gaode Map to transform the village address (combined with detailed township, county, city and province name to ensure exact match) into longitude and latitude. We then create a 1km×1km grid level data covering the entire mainland China, and map the villages into specific grids based on their geo-information. It worth note that there are instances that multiple villages fall into one grids, but such circumstances are relatively rare as most TBVs are geographically dispersed.[54] To ensure that each grid only maps with a unique village, we exclude grids with multiple TBVs. We then map the night-time light data and other relevant variables (e.g., control variables constructed from other raster data) to the grid level. This yields a quasi-village level panel as the actual location of the village may not fully coincide with the 1km grid. However, since the resolution of night-time light and other related variables are at best at the resolution of 1km, we can only reduce our measurement error to such an extent. Compared to the township level measurement, which on average can cover 10-20 1km-grids, the process above greatly improve the precision of the measurement on economic activities of TBVs, and is less prone to the disturbance of other non-TBVs.

[54] The issue that multiple villages map into one grid would rise if we utilize more