When Capacity Encounters Crisis: Subnational Governments' Heterogeneous Response to the "Zero-COVID" Policy in China

Yu Xia *

April 5, 2024

Abstract: Nearly three years after the first outbreak, China relaxed its zero-COVID policy. During the period of strict control, some local officials flexibly achieved both pandemic control and economic recovery with targeted measures, while others insisted on stringent policies with large-scale lockdowns even when the number of infected cases was low. Under the uniform zero-COVID policy, why did subnational governments facing a similar level of infection adopt different lockdown decisions? This study argues that state capacity is key to understanding heterogeneity in policy choice. Specifically, using a novel dateset covering 281 Chinese cities, this study provides robust evidence for these arguments. Firstly, contrary to expectations, fiscal capacity, typically considered a conventional dimension of state capacity, does not significantly influence lockdown decisions. Similarly, capacity less directly related to achieving the policy goal, such as monitoring capacity, also do not explain the variation. Secondly, medical capacity, closely aligned with the policy goal, explains the variation: cities with weaker medical capacity, i.e., fewer doctors for achieving pandemic control, are more likely to prefer strict measures with large-scale lockdowns, while high-capacity cities prefer targeted measures with small-scale lockdowns. Thirdly, different dimensions of capacity are not necessarily correlated: cities with strong fiscal capacity do not necessarily have strong medical capacity. In summary, state capacity is multidimensional and varies across localities, with heterogeneity in policy implementation shaped by capacity contingent on the policy goal, especially in times of crisis.

*Ph.D. student, Department of Political Science, the University of Rochester. Email: yxia24@ur.rochester.edu. I am deeply grateful to Kyle Beardsley, Pablo Beramendi, Christopher Johnston, Melanie Manion, Edmund Malesky, Maria-Giovanna Merli, Casey Petroff, Xiangyu Song, Daniel Stegmueller, and Hongshen Zhu for their valuable feedback and helpful comments. All errors are my own.

1 Introduction

Following a zero-COVID policy, China remained committed to eradicating the virus for nearly three years, implementing extensive lockdowns wherever new cases arose. These lockdowns resulted in food shortages and delayed emergency care, leading to public dissatisfaction with perceived governance failures and challenging the regime's legitimacy. However, lockdowns were neither mandated nor encouraged by the top authorities. To foster innovative solutions and avoid direct responsibility, the Chinese party-state delegated pandemic control to local discretion (Zhu 2023). As a result, subnational governments had considerable policy autonomy between maintaining an open stance and implementing lockdowns. While some cities opted for targeted measures and remained open despite steadily increasing cases, others enforced stringent measures and imposed complete shutdowns even with only a few cases. This paper examines the impact of subnational state capacity on these policy choices, proposing a novel conceptualization of state capacity and evaluating its utility in explaining lockdown decisions.

Past literature on heterogeneity in China's policy processes, particularly in pandemic control, has been focused on incentives. A notable aspect of China's political system is the strong incentive structure for subnational officials, where their performance in achieving policy goals set in Beijing directly impacts their career prospects (Li and Zhou 2005; Xu 2011; Landry, Lü, and Duan 2018). Consequently, variations in policy implementation are often attributed to the unintended consequences of the promotion system. However, it is crucial to recognize that effective policy implementation is heavily reliant on government capacity. North (1990) suggests that, officials without capacity constraints have the discretion to prioritize policies according to their interests. Thus, variations in state capacity at subnational levels should be a primary consideration in explaining heterogeneous policy processes.

State capacity is conceptualized in numerous ways due to the wide spectrum of understanding statehood. The statist approach defines state capacity in terms of the state's most essential functions (Skocpol 1979), while the relational approach considers a broader range of functions and emphasizes the interaction between the state and society (Migdal 1988). Mann (1984) distinguishes despotic power and infrastructural power, providing the most influential understanding of state capacity. Despotic power is entwined with state autonomy and is thus commonly distinguished from state capacity (Hanson and Sigman 2021; Skocpol 1985; Lindvall and Teorell 2016), while conceptualization of state capacity is closely related to infrastructural power, which could be interpreted in three ways: the central state's capabilities, the state's territorial reach, and the state's effects on society (Soifer 2008). These approaches capture the relationship between three actors: central state elites, the radiating institutions of control, and society. Each relationship in this interactive process further provides the basis for identifying measures of state capacity.

Extraction, coercion, and administration are deemed fundamental functions of the state and major activities in the interactive process. Extractive capacity is the state's ability to gather and secure resources. Government revenue is a common measure of extractive capacity, expressed as a proportion of GDP (Besley and Persson 2009), a proportion of total revenue collected (Baunsgaard and Keen 2010), or the amount of tax collected relative to an estimated expected amount of revenue (Kugler 2018). Coercive capacity involves both the state's ability to maintain internal order and protect the state against external threats. It is often operationalized as military size or sophistication - more specifically, military expenditures, military personnel, and security forces (Diehl 1983; Jones, Bremer, and Singer 1996; Wayman, Singer, and Goertz 1983). Administrative capacity pertains to the state's ability to develop and implement policy, provide public services, and regulate economic activities. It could be measured from the condition (presence of state agents) as well as the outcome (public services provision) of effective administration (Acemoglu, García-Jimeno, and Robinson 2015; Acemoglu, Moscona, and Robinson 2016).

When studying multidimensional state capacity, some use one dimension as a proxy for overall state capacity (Dincecco 2011), while others use a composite multi-dimensional index (Hanson and Sigman 2021). Additionally, there are some who use single generic measures (Fearon and Laitin 2003). For these approaches, the basic assumption is that dimensions like extractive, coercive, and administrative capacity are conceptually distinct, but are mutually constitutive and interrelated. Empirical evidence supports the correlation assumption by showing that war generates the state's revenue imperative (Tilly et al. 1992), while effective revenue extraction requires the establishment of sophisticated bureaucratic apparatus (Besley and Persson 2014). However, exploiting one dimension to provide an overall proxy does not always work, because state capacity in different policy sectors and geographic areas could have different starting levels and thus have desynchronized development in later periods. Furthermore, there are tensions between different dimensions of state capacity due to the state's conflicting role as a service provider and a violence monopolizer. Thus state capacity strong in one dimension may even cause weakness in another (Prendergast 2007; Acemoglu et al. 2020).

In addition to the underlying problem of the correlation assumption, broad measures may also suffer from shortcomings that limit their ability to provide policy-relevant guidance (Bersch, Praça, and Taylor 2017). Hanson and Sigman (2021) contend that when dimensions of state capacity do not cohere as a broader construct, measures should be selected to meaningfully represent the narrow concept of interest. As a result, taking the three core dimensions of state capacity as proxies for overall state capacity may be useful for explaining structural questions of state formation and comparing states from historical perspectives (Dincecco 2011; Fearon and Laitin 2003), but it could be misleading for understanding the effect of state capacity on policy implementation.

This paper highlights a new perspective for understanding state capacity. Simply put, the capacity needed for policy implementation is contingent on the specific task. It should not be divorced from what the policy is and what specific conditions should be fulfilled to achieve the policy goal. In addition, different dimensions of state capacity are not necessarily correlated. Given that capacity-building can be influenced by natural endowment, historical legacy, or policy priorities (Straus 2013; Mattingly 2017; Greitens 2017), greater extractive capacity, coercive capacity, or administrative capacity is no guarantee that the state will have stronger capacity in all dimensions. Lastly, subnational variation in state capacity are substantial and can have strong implications for policy outcomes. Capacity at the state level may not fully reflect subnational conditions.

The argument is placed in the context of China during the COVID pandemic. In pandemic control, medical capacity is the crucial shaping factor in policy implementation. Cities with more medical resources choose targeted measures and avoid large-scale lockdowns. To adopt targeted measures, local governments are expected to quickly identify the origins of new cases, efficiently cut off transmission channels, and diagnose patients in a timely way. Equipped with sufficient medical facilities and personnel, the gap between policy goals and local condition in resource-rich cities is relatively narrow. In contrast, for cities with fewer medical resources, pandemic control is challenging, as the shortage of medical resources predicts the unaffordability of an outbreak. Thus, differences in medical capacity are likely to push cities toward two opposite pandemic control strategies, with one group innovatively controlling the outbreak without disrupting economic growth and the other forced to carry out lockdowns at the cost of economic recovery.

Empirically, this paper constructs a novel dataset for 281 Chinese cities in the first month after the COVID outbreak to test the hypothesis that different dimensions of state capacity shape subnational governments' lockdown decisions in response to zero-COVID mandate. There are two main findings. First, medical capacity has a significant effect on local officials' lockdown decisions, while the effects of fiscal capacity and monitoring capacity are insignificant. More specifically, a city with more medical resources is more likely to adopt targeted measures, avoiding large-scale lockdowns and maintaining economic growth. Yet, an overall wealthier city with better surveillance tools does not necessarily show higher tendency to remain open. Additionally, the positive effect of medical capacity on reducing lockdowns is more substantial in cities experiencing more severe waves, and the effect decreases as the pandemic stabilizes. Second, cities with strong fiscal capacity do not necessarily have more medical resources. The mismatch between fiscal capacity and medical capacity demonstrates that state capacity is multidimensional and that different dimensions are not necessarily correlated. It is likely to derive from several conditions, such as the gap between planning and actual demand for public goods due to rapid urbanization and the legacy left by previous policy design and colonial history.

This paper contributes to the literature on state capacity and policy implementation under authoritarianism. First, capacity-based theories focus overwhelmingly on certain aspects of capacity (Cingolani 2013; Berwick and Christia 2018), and the literature on within-country variation (O'Donnell 1993; Fearon and Laitin 2003; Foa and Nemirovskaya 2016) is limited. This paper proposes that state capacity is multidimensional and that different dimensions of capacity are not necessarily correlated. Therefore, taking certain aspects of capacity as proxies for overall capacity is convenient but not necessarily valid. It is more valuable to take policy context into consideration when investigating the effect of state capacity. Furthermore, desynchronized development in state capacity at the subnational level is prevalent and has substantial impacts on governance outcomes. It is thus misleading to ignore within-country variation when examining the effect of state capacity on policy implementation. Second, the existing literature on authoritarian governance is intent-focused, taking state capacity for granted (Harding 1981; Lieberthal and Lampton 2018), even though capacity can be a more constraining factor for many developing or even some developed countries. This paper does not reject the significant role of formal institutions (like promotion incentives) or informal institutions (like patronage networks) in the policy process. Instead, I propose that capacity creates the space for intent, and effective capacity is contingent on policy goals. I demonstrate that, even in a seemingly unified and strong unitary state, capacity constraints can lead to non-compliance by local officials and a fragmented policy process. Unintended policy outcomes can be derived from incapacity rather than lack of incentive.

The rest of this paper is organized as follows. Section 2 provides an overview of existing scholarship on conceptualization and measurement of state capacity. Section 3 follows with a discussion of reconceptualizing state capacity. Section 4 presents an overview of China's pandemic control and subnational governments' response strategy to the central mandate. This sets up the hypotheses of Chinese subnational governments' dichotomous approaches toward the zero-COVID policy and the distribution of different types of capacity. Section 5 moves to empirical strategy, discussing the measurement of key variables and the model. Section 6 presents findings. Section 7 presents robustness checks, parallel trend test, and extensions. Section 8 concludes.

2 Conceptualization and Measurement of State Capacity in Existing Scholarship

2.1 Conceptualization of State Capacity

As interest in understanding the role of the state in development grows (Evans et al. 1985), there has been an expanded understanding of the state, with an increased focus on state capacity. According to Weber's

classic definition, the state is a human community that seeks to monopolize the legitimate use of violence within a given territory (Gerth and Mills 2014). Historical analysis sees the Weberian definition as an idealized endpoint of state development (Dincecco and Wang 2022), conceptualizing the state as a political hierarchy with an extractive apparatus for revenue collection, a military for security, and a judiciary for legal justice administration (Bates 2020). In the framework of Mann (1984), the state is defined as the political relations radiating from a center to cover its territory. This definition not only requires the state to have a professionalized bureaucracy and monopoly of coercive power as emphasized by Weber, but also underscores the extent to which the state's bureaucracy can extend to exert control.

Given the broad range of interpretations of statehood, state capacity is conceptualized in various ways. From a state-centered perspective, Tilly (1975) argues that state capacity is rooted in the historical role of war in state formation, emphasizing the ability to establish an apparatus that effectively mobilizes resources from the local population while overcoming resistance to resource extraction. Similarly, Skocpol (1979) examines the state's role in revolutions, focusing on administrative capacity to implement central policies and the dynamics between state apparatus, ruling elites, and bureaucratic agents. In contrast to the state-centric view, other scholars adopt a more sociocentric perspective, emphasizing a relational approach where state capacity emerges from interactions between the state and society. For instance, Migdal (1988) defines state capacity as the ability of state leaders to compel compliance from the populace through state agencies. This perspective prioritizes capacities such as resource collection, social regulation, and territorial penetration. Moreover, it is noted that in developing countries, state autonomy and strength cannot be assumed, as the state-centric approach suggests.

The most influential understanding of state capacity is often attributed to Mann (1984), who delineates state capacity as a form of power and distinguishes between despotic power and infrastructural power possessed by states and state elites. Despotic power refers to the range of actions that state elites can undertake without routine negotiation with civil society, whereas infrastructural power pertains to the state's ability to implement intended policy goals, particularly in the face of societal opposition or challenging socioeconomic circumstances. Scholars tend to separate despotic power from the concept of state capacity due to its association with political regime characteristics and issues of state autonomy (Hanson and Sigman 2021). Similarly, Skocpol (1985) argues that state capacity is the state's ability to implement intended policy goals, while state autonomy refers to the state's control over territories and pursuit of goals that may not align with societal interests. Lindvall and Teorell (2016) further elaborate that state capacity is a reflection of the power the state projects to control outcomes it aims to achieve, whereas state autonomy signifies the extent to which the state is not controlled by external powers. Thus, state capacity should be conceptually distinct from mechanisms for societal involvement in political decisions or the power of society to resist state actions

(Hanson and Sigman 2021).

In contrast to despotic power, infrastructural power is the capacity of the state actually to penetrate civil society and to implement logistically political decisions throughout the realm (Mann 1984). Soifer (2008) breaks it up in three ways: the central state's capabilities, the state's territorial reach, and the state's effects on society. In the first approach of national capabilities, infrastructural power is considered a characteristic of the central state, relating to the extent of material resources and organizational competencies available to the state for exercising its power. The second approach does not assess power in terms of resources, but captures the weight of the state and highlights how states are limited and constructed by members of society. The third approach focuses on subnational variation in the state's ability to exercise control within its territory. In this way, Mann's concept of infrastructural power captures the relationship between three actors: central state elites, the radiating institutions of control, and society. Each relationship in this interactive process further provides the basis for identifying measurements of state capacity.

2.2 Measuring State Capacity

State capacity is regarded as a latent feature as it cannot be directly observed and thus a range of visible indicators are employed to measure state capacity (Hanson and Sigman 2021; Lindvall and Teorell 2016). The selection of these indicators involves defining what functions a capable state is expected to perform. Some focus on the most essential features and functions, determined by to what extent a state holds a monopoly on the legitimate use of force in its territory following the Weberian definition (Linz et al. 1996). While others consider a much broader range of functions, assigning states a large number of roles such as maintenance of internal order, protection of external threats, development of the economy, and provision of public services (Besley and Persson 2011; Tilly et al. 1992). In addition to breaking state capacity down by forms and functions, Berwick and Christia (2018) propose to distinguish between different types of state goals: extraction, coordination, and compliance. Among different approaches to slicing and dicing the concept of state capacity, extraction, coercion, and administration are the three most frequently used dimensions when illustrating state activities.

Extractive capacity, the state's ability to gather and secure resources, is deemed the most fundamental form of state capacity. As Levi (1988) and Tilly et al. (1992) contend, the income of governments is the major limitation for a government to exert control and extend rule, measures of extractive capacity often come in the form of government revenue. Government revenue can be expressed by a proportion of GDP (Besley and Persson 2009), a proportion of total revenue collected (Baunsgaard and Keen 2010), or the amount of tax collected relative to an estimated expected amount of revenue (Kugler 2018). Scholars select different revenue

indicators for their particular purposes. Also, the greater availability of revenue data for most countries than alternative types of data makes extractive capacity the most commonly investigated aspect of state capacity by scholars. For instance, Dincecco (2011) collects annual fiscal data for major countries in Western Europe from 1650 to 1913. Lee and Paine (2022) focus on non-Western states, gathering revenue data from the nineteenth century for more than 40 countries. Nevertheless, the debate over the government's ability versus its willingness to engage in revenue extraction indicates that assessing the level of revenue may reveal the government's policy preference instead of capacity (Alesina, Glaeser, and Sacerdote 2001).

Under the Weberian definition of the state as an entity to monopolize the legitimate use of violence, coercive power is central to state capacity. Coercive capacity involves both the state's ability to maintain internal order and protect the state against external threats. The national military is regarded as key to the state's ability to deter or repel challenges to its authority from both inside and outside (Hendrix 2010). Thus coercive capacity is operationalized as military size or sophistication, more specifically, military expenditures, military personnel, and security forces (Diehl 1983; Jones, Bremer, and Singer 1996; Wayman, Singer, and Goertz 1983). However, Henderson and Singer (2000) dissents that coercive force is not necessarily equal to coercive capacity, as large military force can be associated with a higher likelihood of conflict onset or a sign of insecurity. Furthermore, since military and security forces are not the only way to evoke compliance and maintain order (Levi 1988), Berwick and Christia (2018) expands the militaristic emphasis on coercion, integrating more nuanced aspects of compliance such as ideology and economic incentives with outright coercion.

Beyond political power and the monopoly of legitimate force that are claimed to be essential for the state according to Weberian definition, Mann (1984) argues that the transition from the state to a modern state is the addition of professionalized bureaucracy. Administrative capacity is therefore essential for state development. It pertains to the state's ability to develop and implement policy, provide public services, and regulate economic activities. Given that administrative capacity is an encompassing dimension of state capacity, it is measured from the condition (presence of state agents) as well as the outcome (public services provision) of effective administration. For instance, Acemoglu, García-Jimeno, and Robinson (2015) conceptualize state capacity as the presence of state functionaries and agencies, and measure it through the historical presence of colonial state officials and agencies. Similarly, Lee (2019) examines the presence of village officials in India, taking it as a key institutional input to local state power. Acemoglu, Moscona, and Robinson (2016) proxies the local reach of the federal government in the U.S. by the presence of post offices. Still, there is skepticism about whether capacity or willingness is being measured.

In addition to extractive, coercive, and administrative capacity, information capacity, the state's ability to obtain accurate information about the population and territory, has gained increasingly more attention

(Lee and Zhang 2017). Also, in the review of Cingolani (2013) on state capacity scholarship, four other dimensions in use are identified: transformative/ industrializing, relational/ territorial, legal, and political capacity. Moreover, there are attempts to incorporate different measures associated with extraction, coercion, and administration into a single indicator of state capacity (Hanson and Sigman 2021).

3 Reconceptualization of State Capacity

Studies using one dimension as a strong proxy for overall state capacity, a composite multi-dimensional index, or single generic measures, often assume that dimensions like extractive, coercive, and administrative capacity are conceptually distinct but mutually constitutive and interrelated. For instance, Fearon and Laitin (2003) take GDP per capita as an overall representation of the state's military, administrative, and bureaucratic capacity, while Rogers and Weller (2014) validate income taxation as a good state capacity indicator. Literature pursuing the origins of state capacity describes the relationship between different dimensions. The bellicist hypothesis of Tilly et al. (1992) argues that war generates the state's revenue imperative. In other words, the need for coercive capacity pushes the state to establish and increase extractive capacity. Herbst (1990) explores the relationship with a different approach, contending that the absence of external threats leads to the lack of high-quality institutions in postcolonial African countries. Meanwhile, effective revenue extraction requires sophisticated bureaucratic apparatus. Therefore, states lacking the ability to coordinate, monitor, and enforce tax collection rely more on trade taxes or mineral resource rents, which are less demanding for administration (Besley and Persson 2014; Lieberman 2002). These linkages between different dimensions of state capacity appear to provide the basis for the correlation assumption.

However, exploiting one dimension to provide an overall proxy does not always work. As Gingerich (2013) argues, national-level indicators for capacity are not informative about within-country and cross-agency diversity. Skocpol (1985) notes the potential for the uneven distribution of state capacity across policy realms by using the plural "state capacities". Soifer (2008) also highlights the subnational variation approach, distinguishing it from the approach that considers state capacity as a characteristic of the central state. Therefore, state capacity in different policy sectors and geographic areas could have different starting levels and thus have desynchronized development in later periods. For instance, Skocpol and Finegold (1982) study U.S. government interventions in economic activities in the 1930s and find that the government had greater capacity to intervene autonomously in agricultural affairs than in industry. Historical reasons such as ethnic diversity (Alesina, Glaeser, and Sacerdote 2001), direct or indirect governance arrangement (Mamdani 1996), and settlers' policy preference (Acemoglu, Johnson, and Robinson 2002) could all provide explanations for variations in the development of capacities. In sum, state capacity with a high level in one dimension

need not have an equivalent level in another, and taking one dimension as a general indicator is convenient but not necessarily solid.

Furthermore, there are tensions between the state's role as a service provider and violence monopolizer, or between its priorities on domestic and international affairs (Evans et al. 1985), leading to conflicting relations between different dimensions of state capacity. Prendergast (2007) shows the potential for policy implementors to enhance certain types of state capacity, such as public goods provision, while being unlikely to develop others, such as law enforcement. Acemoglu et al. (2020) investigate the effect of high-powered incentives on Colombian governance, arguing that state-building efforts that place the main emphasis on security objectives may hinder the development of other important governmental institutions. Similarly, the contradiction between legal and administrative capacity, determined by whether higher intervention potential is associated with higher or lower capacity (Cingolani 2013), also adds to the point that state capacity strong in one dimension may cause weakness in another.

When dimensions of state capacity are not coherent into a broader construct, Hanson and Sigman (2021) contend that measures should be selected to meaningfully represent the narrow concept of interest. In other words, taking a commonly-used dimension as an overall proxy to examine the effect of state capacity on policy implementation may be misleading. For instance, in examining the determinant of the strikingly effective implementation of genocidal policies in Rwanda in 1994, Straus (2013) argues that the pervasive control of the state was not realized through extractive capacity but through the permissive geographic context and the tradition of massive mobilization. The analysis of subnational variation in frontier states by Foa and Nemirovskaya (2016) shows that effective public goods provision is driven by strong social institutions instead of formal mechanisms to sustain public order. Thus the capacity to coordinate non-state actors (Berwick and Christia 2018) is key to governance outcomes, while strong coercive capacity may be counterproductive. In short, we should be more careful in using general measures of extractive, coercive, or administrative capacity regardless of the policy context.

The three core dimensions of state capacity, which are minimally necessary to carry out key state functions, are useful for explaining structural questions of state formation and comparing states from historical perspectives (Dincecco 2011; Fearon and Laitin 2003). However, as mentioned above, broad measures may suffer from shortcomings that limit their ability to provide policy-relevant guidance (Bersch, Praça, and Taylor 2017). After all, daily governance and state development are no less important than state formation. Further, broad state goals such as growth and security may consist of smaller policy goals, whose realization depends on a more nuanced investigation of the state's internal structures and capacity. In sum, the state's "ability to implement policy" should not be divorced from the policy itself and the specific conditions that must be met to achieve the policy objective. Also, this policy-based capacity is not always reflected in

traditional dimensions of state capacity and should be evaluated on its own merits.

4 Testing the Concept: COVID-19 in China

Crisis time is an ideal window for examining state capacity's effect on policy implementation. In a crisis, political intents of officials at different levels are expected to be aligned in dealing with the crisis, while monitoring from the top is more effectively carried out (Hart 1993; Boin and Otten 1996; Thornton 2009). Therefore, heterogeneity in policy outcomes is more likely to derive from differences in capacity constraints. This paper chooses China's COVID-19 pandemic control to test the effectiveness of the state's general capacity and policy-based capacity on policy implementation.

4.1 Policy-based Capacity in Pandemic Control

In December 2019, a new virus was discovered by researchers in China. It was later named COVID-19 and has since spread to nearly every country. On January 23, 2020, the first lockdown decision to halt the transmission of the virus in Wuhan, a city of more than eleven million, was made by the top authorities. The Chinese government's intent of eradicating the virus through various measures like large-scale lockdowns and testing remained until December 2022, upending life and derailing the economy for nearly three years. Under the highly uniform zero-COVID policy, however, subnational officials are given the discretion to choose between targeted measures and strict measures, as the center deliberately permits subnational governments the flexibility of implementation to stimulate innovative solutions (Zhu 2023).

As discussed by the literature on crisis management, during the pandemic, subnational officials' interests were aligned with the center's zero-COVID policy through well-designed inducements and sanctions. On the one hand, the performance in pandemic control was prioritized and integrated into officials' evaluation system.[1] On the other hand, heavy sanctions against ineffective pandemic control were imposed. By mid-April 2020, more than three thousand Party members and cadres in Hubei Province had been punished for dereliction of duty in pandemic control, including more than ten at provincial and municipal levels and more than one hundred at county level.[2] Furthermore, the problems of asymmetric information and monitoring in the principal-agent relationship were also mitigated. The zero-COVID policy is readily measurable and the center has established effective controls that lead implementors to define the tasks as policymakers wish.[3]

[1] On January 28, 2020, Beijing announced that pandemic control should be linked to Party members' performance evaluation. See http://www.xinhuanet.com/politics/2020-01/28/c_1125508545.htm

[2] See https://theinitium.com/article/20211117-mainland-acountability-officials-COVID/

[3] On January 24, 2020, Chinese Center for Disease Control and Prevention started to operate the System of Dynamic Surveillance on Coronavirus, requiring subnational units to report cases on a daily basis via the system. See https://www.

Also, public scrutiny and expectation over pandemic control were high, effectively monitoring subnational governments' performance through social media (Xu et al. 2020). All these conditions incentivize officials to concentrate resources to hit the policy goal. Therefore, subnational officials are expected to show more compliance with the center during the pandemic, and heterogeneity in policy choice is likely to be determined by the difference in capacity for achieving the policy goal across localities.

In the COVID-19 pandemic, the determining capacity for tackling the crisis is medical capacity. Medical capacity could be defined as the capacity to conduct pandemic control and meet the zero-COVID goal through timely diagnosing and treating infected patients, efficiently identifying the origin of the outbreak, quickly cutting off transmission channels, and other relevant medical approaches. It is reliant on the number of hospitals, medical facilities, medical personnel, and other medical resources. The potential threat of putting an overwhelming burden on medical resources by the sudden outbreak is claimed to be the reason for China's policy consistency in zero-COVID.[4] More specifically, the burden refers to the situation that, in the early stage of the outbreak, every infected patient has to be transferred to nearby hospitals, greatly outnumbering the medical personnel and beds in hospitals in some places (Yin et al. 2021). If the increase of infected cases greatly exceeds the capacity of hospitals, a city's medical system will break down, leading to more infected cases. Furthermore, daily prevention requires an abundant supply of personal preventive equipment such as N95 respirators, facemasks, isolation gowns, and gloves. These are medical resources for emergencies, so the reserves are insufficient compared to the resources used for routine medical needs for many localities. As panic buying by the public led to supply shortages in the market during the early stages of the outbreak, the existing stocks of medical resources that could be allocated by subnational governments became critical for pandemic control. In short, medical capacity plays a key role in pandemic control and policy consistency in zero-COVID.

4.2 Mismatched Dimensions of State Capacity

As capacity is multidimensional and different dimensions may not necessarily correlate, cities that are strong in traditional dimensions of state capacity like fiscal capacity may not have an equivalent level of medical capacity. Research suggests that among cities with the greatest number of doctors per thousand population, most are the third and fourth tier cities in terms of population and size. Similarly, among cities with the greatest number of hospital beds per thousand population, 6 out of 10 locate in the northeast, which is the

chinacdc.cn/zxdt/202001/t20200125_211441.html

[4]See https://covid-19.chinadaily.com.cn/a/202004/21/WS5e9e2c62a3105d50a3d17880_5.html

less developed area in China.[5]

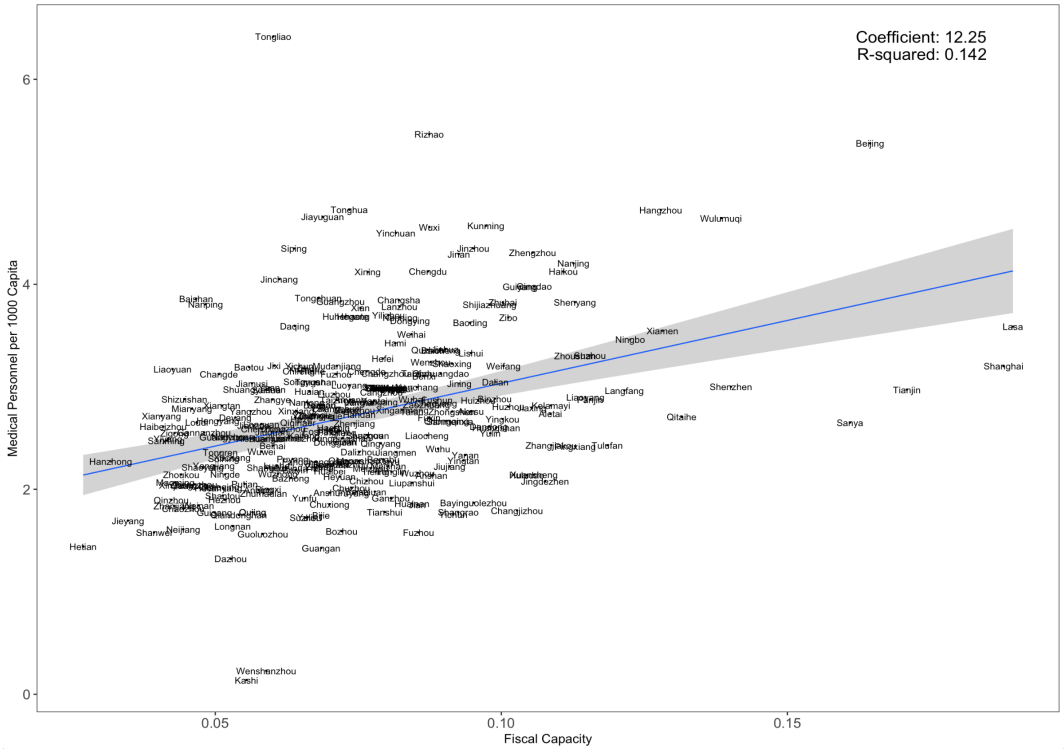
When correlation assumption for different dimensions of state capacity does not hold, the effect of state capacity on policy implementation should consider the policy context, and the influence of the specific policy-based capacity should be investigated separately, independent from traditional dimensions. In this case, specifically, the distribution of fiscal capacity, the traditional dimension, is expected to be mismatched with medical capacity, the policy-based dimension, across the country. Figure 1 offers a systemic look at the relationship. The scatterplot and the boxplot illustrate that there is no strong positive correlation between the two dimensions of state capacity. In the scatterplot, most observations cluster around the bottom left. It shows that cities with the same level of fiscal capacity could vary greatly in medical capacity. Similarly, the boxplot further illustrates that cities in different quantiles of fiscal capacity appear to have the same level of medical capacity. For instance, the plot shows that Shanghai has the strongest fiscal capacity in the country, but its medical capacity drops below the average. Also, Tongliao ranks first in medical capacity, yet its fiscal capacity is around the middle level. These evidences show that state capacity is multi-dimensional. A state with high capacity in one policy realm needs not to have an equivalent level in another.

The mismatch between fiscal capacity and medical capacity could stem from various conditions, which are beyond the scope of this paper. Therefore, only brief discussions are provided. First, the rapid speed of urbanization and the flow of the migrant population are likely to lead to the gap between planning and actual demand. The planning period of building medical capacity is generally five years, consistent with the five-year plan for national economic and social development. However, the planning period for the layout of land is twenty years, following the overall urban plan. Meanwhile, people migrating from poor cities in the west to wealthy cities in the east or crowding in urban areas from rural areas all add up to the disproportional increase in city population and public health burden (Gong et al. 2012). Without a dynamic evaluation mechanism for the implementation of planning, the supply often lags the demand out of the plan.[6] Therefore, the burden on medical resources could be aggravated in major and wealthier cities, while the trend of hollowing out in less developed cities conversely alleviates the problem.

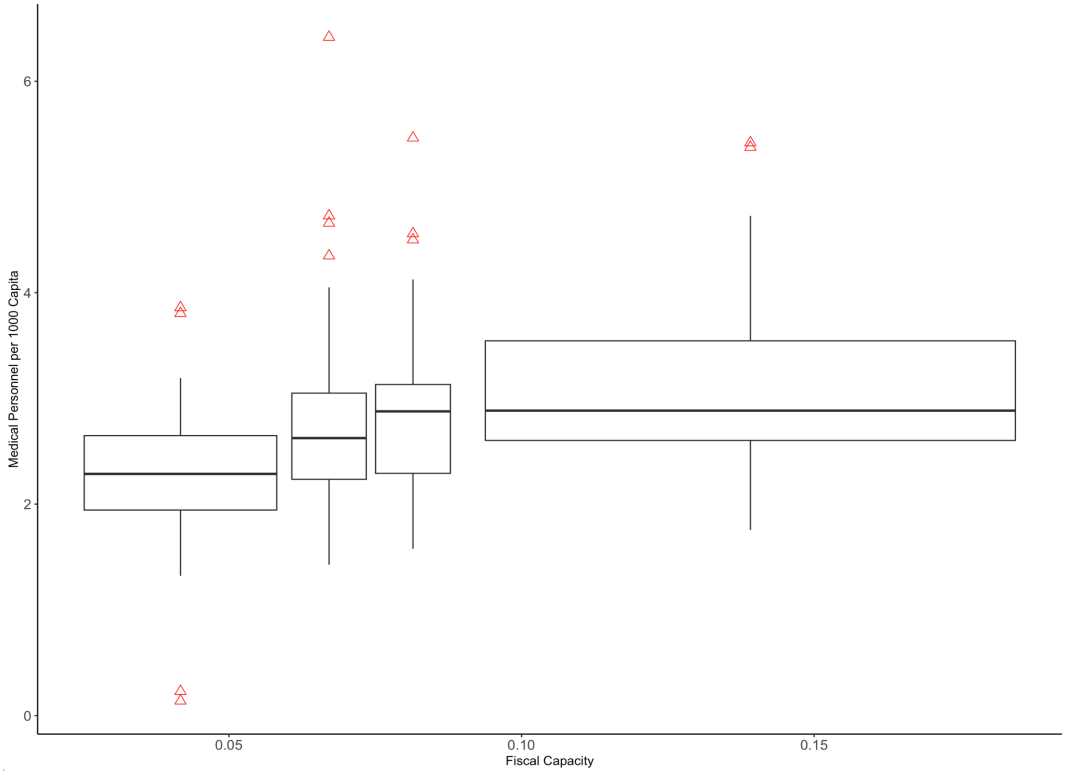
Also, even if the plan catches the dynamic of demand, the implementation is constrained by other factors such as the public expenditure budget (Xu et al. 2021). Subnational governments that have more revenues should be able to provide more public goods provision such as health facilities, but these governments of developed cities could confront with greater difficulty affording the price of building hospitals and medical

[5]Chinese Academy of Social Sciences released a report on Chinese cities' competitiveness, listing cities with the most medical resources and concluding that the geographical distribution of medical resources is uneven. See https://www.guancha.cn/politics/2020_03_09_540664.shtml

[6]See https://www.thepaper.cn/newsDetail_forward_6770718



(a) Scatterplot of Fiscal Capacity and Medical Capacity



(b) Boxplot of Fiscal Capacity and Medical Capacity

Figure 1: The Mismatch between Fiscal and Medical Capacity at the Subnational Level

schools that require large areas of land in populous zones, while governments of less developed regions may face fewer constraints in providing public health services given lower land prices.

Second, policy and historical legacies could also explain the geographical distribution of medical resources. For instance, Mattingly (2017) claims that Japanese colonization of northeast China had a positive long-run effect on state institutions—with persistent increases in the density of health facilities. Also, the West China Development plan initiated in 2000 has provided great policy support for the development of people's welfare, promoting the improvement of medical conditions in western China. Moreover, as mentioned above, cities that experienced SARS pandemic in 2003 could be more likely to invest in medical capacity regardless of the level of fiscal capacity. These examples illustrate how medical capacity might be built without reliance on fiscal capacity.

Lastly, Zhou and Xin (2021) contend that subnational leaders that have working experience in public health department may place more importance to building medical capacity and respond more quickly and effectively in pandemic control. Therefore, a city with weaker fiscal capacity but with an experienced leader could have stronger medical capacity. Conversely, local authorities' lack of attention to the provision of public goods may add up to the uneven relationship between medical capacity and fiscal capacity.

In sum, there are many factors shaping the process of capacity building, leading to the uncorrelated relationship. The results illustrate the gap between the ability to provide public goods and the actual level of provision in China, challenging the assumption that different dimensions of state capacity are correlated and that a single dimension could be used as an indicator for the overall state capacity. The mismatched relations also require examining independent effects of different dimensions of state capacity on policy implementation separately.

4.3 Capacity Constraints on Policy Implementation

In general, there are two groups of policy choices in pandemic control. Targeted measures are to control the pandemic with the minimum social cost, achieve rapid tracking of the virus, and ensure the normal operation of production and life. This approach requires subnational governments not to arbitrarily and unnecessarily expand the high-risk zone, extend lockdown periods, close sites providing daily necessities, and restrict public mobility.[7] It is exemplified by precise and small-scale lockdowns. In contrast, strict measures involve indiscriminately and large-scale lockdowns. These measures could be taken when the city is on the verge of breakdown, but could also be used as overly-cautious preventive methods when infection

[7]The central government listed nine prohibited policies that are deemed unreasonable for pandemic control. See http://www.gov.cn/xinwen/2022-07/13/content_5700806.htm

cases are low, resulting from intensifying restrictions on a layer-by-layer basis that are commonly seen in China's bureaucratic system (Zhou 2022). In sum, officials are endowed with the discretion to choose between targeted and strict measures for pandemic control, and the scale of lockdowns implies their policy choices.

In terms of specific regulations, "community closed-off management" represents targeted measures and "family outdoor restrictions" represents strict measures for pandemic control. Figure 2 shows the adoption of targeted and strict measures across cities except those in Hubei Province. Hubei is excluded here because the mandate for large-scale lockdowns was issued directly by the central government, leaving subnational governments with no discretion to choose between different measures. However, cities outside Hubei were granted autonomy in their decision-making. Between January 28 and February 20, over 250 prefecture-level cities in China implemented "community closed-off management" measures (Qiu, Chen, and Shi 2020), typically involving maintaining a single entrance for each community, permitting entry and exit solely for community residents, conducting body temperature checks for entrants, promptly testing and isolating individuals displaying fever symptoms, and tracing and isolating close contacts of suspected cases. In cities adopting closed-off management, communities kept open and people were allowed to move freely.

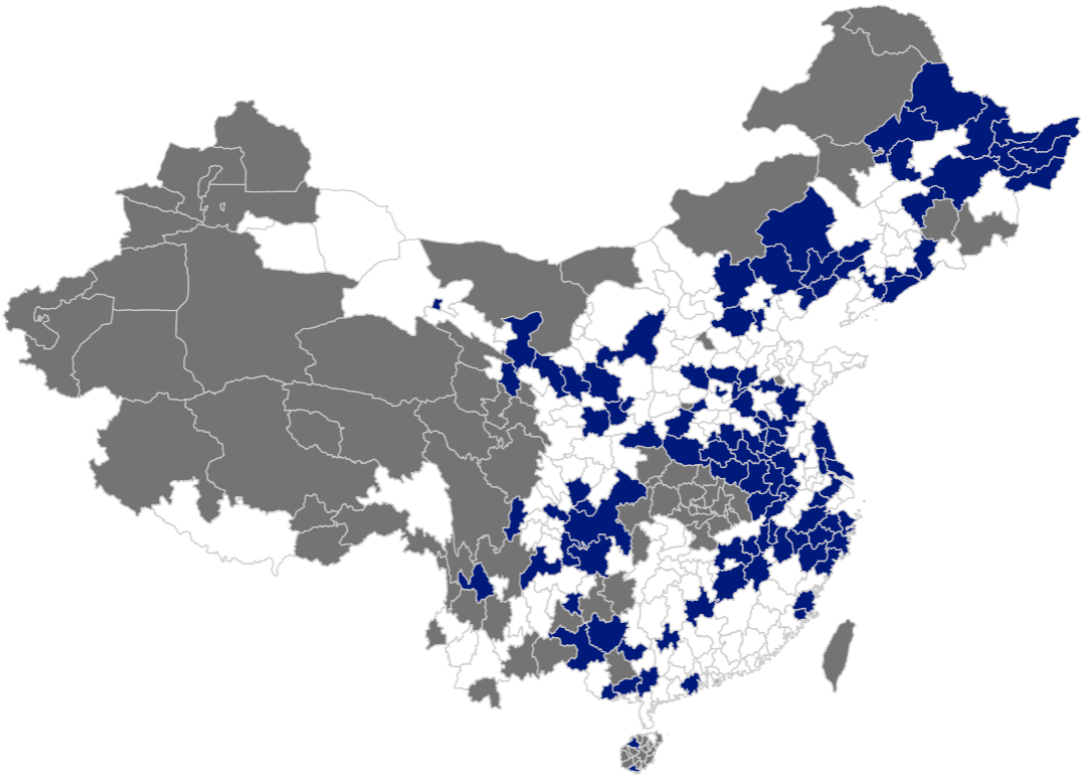


Figure 2: Closed-off Management and Outdoor Restriction for Pandemic Control

Conversely, the local governments of 127 cities implemented stricter "family outdoor restrictions," with residents either confined to their homes or strongly encouraged to stay indoors, except for limited circumstances such as one person per family being allowed to go out for essential shopping once every two to five days (Qiu, Chen, and Shi 2020). Exit permits were typically issued to families in advance and collected upon re-entry to the community, while contacts of affected individuals were also traced and quarantined. Both "closed-off management" and "outdoor restrictions" measures were applied to all communities with a moderate risk level. More stringent measures, such as "total outdoor restriction" for 14 days or longer, could be found in areas with a rapid increase of infected cases.

Similarly, figure 3 uses intracity mobility change as the indicator of the lockdown scale. Each dot represents a city. Dots below zero indicate that the city experienced decreased mobility change from the previous day and vice versa. The red dashed line is January 20, when the warning of human-to-human transmission of COVID was announced. The figure shows the heterogeneity in local officials' policy choice between targeted and strict measures in pandemic control. For example, there was a sharp decrease after January 20, and it dropped to the lowest point after the first lockdown was announced in Wuhan city on January 23. These trends of decreased mobility imply increasing lockdown decisions by local officials.

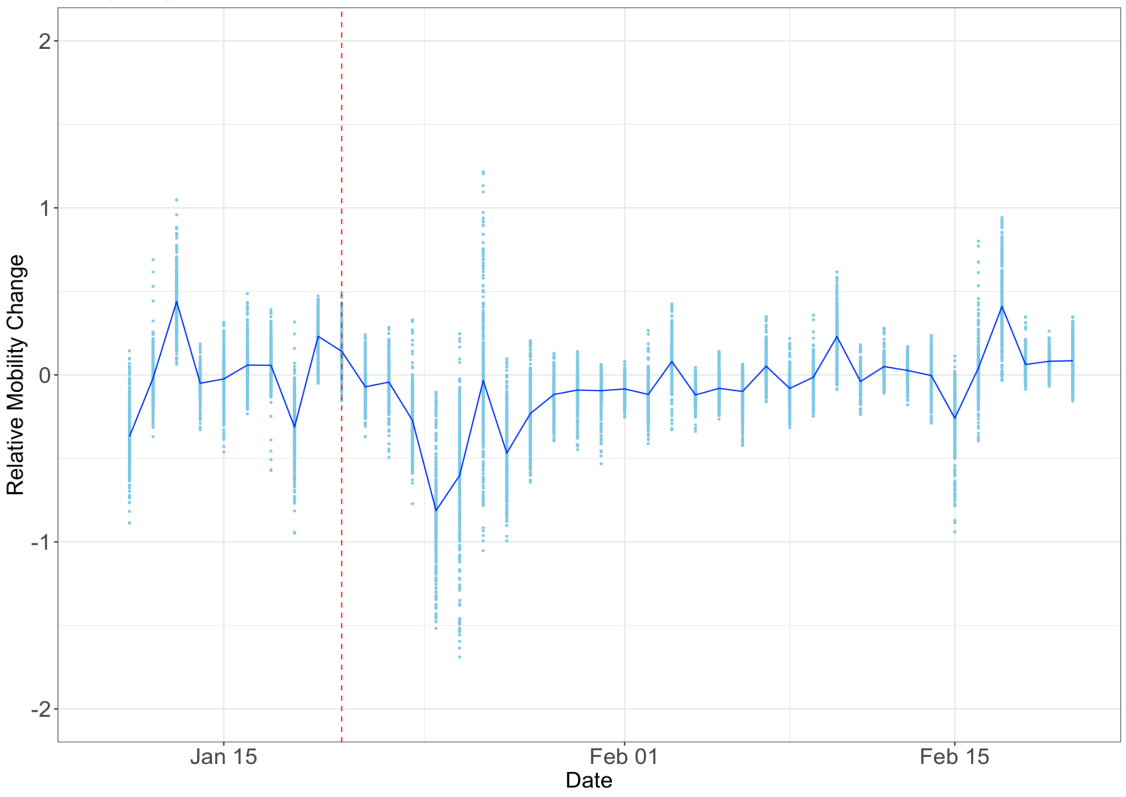


Figure 3: Intracity Mobility Change and Lockdown Decisions

With different levels of medical capacity, resource-rich cities and resource-poor cities seem to be pushed toward different pandemic control strategies. Cities with more medical resources are expected to be more confident in choosing targeted measures. To adopt targeted measures but get the pandemic under control, subnatinonal governments are expected to have the ability to quickly identify the origins of new cases, efficiently cut off transmission channels, and timely diagnose patients. Equipped with sufficient medical facilities and personnel, the gap between policy goals and the local condition in resource-rich cities is relatively narrow. For example, cities with more hospitals could designate more areas to quarantine and treat COVID patients, and more medical personnel would allow for more testing and diagnosing. Contact tracing is critical as well. With a large group of epidemiological researchers and disease control staff in charge of contact tracing and community-level governance, a city could operate the early warning system effectively, controlling the pandemic without a large-scale lockdown (Gu and Li 2020).

Moreover, driven by the promotion system, local cadres with strong capacity could be less risk-averse as they seek promotion rather than merely survival (Solinger and Jiang 2016; Lü and Landry 2014; Liu 2019). Cities with stronger capacity are more confident in achieving daily prevention and could be more willing to take the risk of a new round of outbreak, as the promotion reward is incentivizing. Therefore, with less pressure to deal with the pandemic, officials attach importance to economic recovery, choosing targeted measures as a promotion strategy.

In contrast, for cities with fewer medical resources, pandemic control is challenging as the shortage of medical resources predicts the unaffordability of an outbreak. On the one hand, lockdown effectively cuts down the transmission of the virus, but it apparently collides with the economic growth goal as the whole city has to come to a standstill, let alone the pandemic has already posed a great challenge to economic growth.[8] On the other hand, adopting targeted measures and removing lockdowns are less disruptive but more risky and demanding. It could be devastating for cities without enough medical personnel and hospital beds to confront a large-scale outbreak. Even when a substantial case increase is low, daily prevention work burdens subnational governments and further diverts resources from economic revitalization. With neither enough contact tracing staff to stop transmission nor sufficient medical resources to deal with a new wave of outbreaks, strict lockdown appears to be easier to deliver than the targeted measure when infected cases increase. With the payoff analysis, localities that are incapable of carrying out state policies may choose to devote all effort to tasks to flag the strong effort as a survival strategy (Chen, Li, and Lu 2018).

In addition, weak cities are also likely to choose strict measures out of a performative strategy. The

[8] Chen et al. (2022) estimate that imposing a full-scale lockdown on the four largest cities for one month would reduce the national real GDP by 8.6%, of which 11% is contributed by the spillover effects.

difference between substantial and performative governance in pandemic control is how much the preventive measures match the infection status. Overly-cautious policies are detrimental, hurting both the economy and public trust, and thus not encouraged by Beijing.[9] However, the party-state only administers mild punishment against the overly-cautious measures, while losing control of outbreaks could face a more serious outcome. Meanwhile, improved economic performance would be rewarded, yet a slower growth rate due to the pandemic would not be punished. By choosing strict measures, resource-poor cities do not have to take the risk of a new outbreak, and they would not be blamed for the economic recession. Thus, officials with weak medical capacity to manage the situation but high political intent to display compliance are likely to overreact and choose large-scale lockdowns.

In short, medical capacity are likely to determine subnational governments' choices in pandemic control strategies, with one group innovatively controlling the outbreak without disrupting citizens' daily life and the other forced to carry out lockdowns at the cost of the economy.

4.4 Hypotheses

Given that medical capacity appears to play a decisive role in pandemic control and the distribution of different dimensions of state capacity could be mismatched, the effect of targeted and policy-based capacity (i.e., medical capacity) and the traditional dimension of state capacity (i.e., fiscal capacity) on lockdown decisions should be investigated independently. In addition, the effect of less targeted capacity (i.e., monitoring capacity) could be examined. This yields a testable hypothesis and an alternative hypothesis for targeted and non-targeted capacity:

H1: With the same level of infection, cities with stronger medical capacity are likely to adopt targeted measures with small-scale lockdowns, while cities with weaker medical capacity are likely to adopt strict measures with large-scale lockdowns.

H2a: With the same level of infection, cities with stronger fiscal capacity are likely to adopt targeted measures with small-scale lockdowns, while cities with weaker fiscal capacity are likely to adopt strict measures with large-scale lockdowns.

H2b: With the same level of infection, cities with stronger monitoring capacity are likely to adopt targeted

[9] Chinese president Xi urged subnational government officials to refrain from more restrictive measures. See https://www.reuters.com/article/us-china-health-xi-economy/xi-warned-officials-that-efforts-to-stop-virus-could-hurt-economy-sources-idUSKBN2050JL

measures with small-scale lockdowns, while cities with weaker monitoring capacity are likely to adopt strict measures with large-scale lockdowns.

5 Data

The unit of analysis is city-day. Hubei Province is left out of the analysis as the lockdown decision was made by Beijing, and thus subnational leaders' consideration of medical capacity is not applicable. The main analysis covers 281 cities over the period from January 20 to February 20, 2020. The time span is chosen for several reasons. First, the human-to-human transmission was confirmed on January 20, and the first full-scale lockdown order was announced in Wuhan on January 23. Second, outside of Hubei, the surge of infected cases was under control by the end of February, with a daily increase of infected cases nationwide dropping to below 100 for three consecutive days on February 20 [10]. Third, this analysis chooses a one-month period at the early stage, as the medical capacity measure may not be constant for a longer period. Subnational governments are adaptive; thus the decisive capacity for dealing with the crisis could be built after its significance has been revealed. More contact tracing staff will be hired, and more mobile cabin hospitals will be built, leading to increasingly imprecise estimation as time passes.

5.1 Dependent Variable

This paper examines how state capacity, especially targeted capacity, influences the policy choices of Chinese subnational governments in pandemic control. The dependent variable investigated is local officials' lockdown decisions, which are measured in two ways. First, it is directly measured by pandemic control policies announced by prefecture-level cities, including targeted measure "closed-off management" and strict measure "outdoor restriction". Data are collected from official government accounts on WeChat, the most popular social media app in China. These accounts serve as reliable sources of government regulations for the public. Between January 28 and February 20, over 250 cities in China implemented "closed-off management", while 127 cities implemented stricter "outdoor restrictions" (Qiu, Chen, and Shi 2020). As discussed previously, the distinction between these two regulations is whether people were allowed to move freely. Cities with "outdoor restrictions" are coded as 1, and cities with "closed-off management" are coded as 0. It should be noted that not every city made an official announcement on lockdown decisions. If a city did not make any official announcement of strict outdoor restriction, it is coded as 0,

Second, using Zhu (2023)'s measure, lockdown decisions are reflected by differences in intracity mobility

[10] See http://www.xinhuanet.com/politics/2020-02/20/c_1125601020.htm

intensity, which is the proportion of city residents who change their locations, from the previous day according to new infection cases. Specifically, it measures how much people's movement is influenced by the newly-increased cases in the previous day. People's mobility is measured by relative change instead of level. According to Zhu (2023), mobility change captures new lockdown decisions in response to newly increased infections, while mobility level captures the result of lockdown. Since subnatioinal governments' decision making is the main concern, mobility change is the more appropriate choice.

The intracity mobility data are collected from the China Data Lab of Harvard Dataverse. Data are originally provided by Baidu Location-Based Service, which is the most extensive location data service platform in China. Baidu developed the widely-used application Baidu Map and provides service for many third-party platforms, which is similar to Google Map. Thus it collects extensive location data from the map application and users' information on related platforms. As the data only cover users of the service, it is incomplete. However, most users aggregate in urban areas, and there was little spread of the virus in rural communities (Wang et al. 2021). Also, given the high user coverage, the subset is expected to be a substantial enough measure of the whole population [11].

5.2 Independent Variable

My independent variable of theoretical interest is medical capacity, measured by the number of medical personnel per thousand population. Figure 4 shows the distribution of medical capacity in China. Medical personnel refers to professional staff employed by public health institutions. There are four categories of professional staff, including medical care, nursing, pharmacy, and technicist. The population refers to the city's permanent resident population. Medical resources could also be measured by hospital beds and health facilities, the results of which will be shown in the extensions part. This analysis deems the number of medical personnel a better indicator of resources than hospitals and beds medical in the time of the pandemic. As self-quarantine was not allowed for infected cases in China until late 2022, every infected patient had to be transferred to a designated hospital nearby or quarantine camp.[12] Many community hospitals and grassroots clinics were unqualified to treat COVID-19 patients but medical personnel could be deployed to support qualified hospitals. Furthermore, medical personnel includes epidemiological researchers who are critical for identifying the origin and cutting off the transmission of the virus, and lockdown decisions relied greatly on their analysis. Thus medical personnel could be relatively more determinant as an indicator of medical

[11] The users of mobile maps reached nearly 1 billion in 2021 and the penetration rate of mobile internet reached 76 percent among the urban population of China in 2020. See http://bj.news.cn/2022-08/30/c_1128959674.htm and http://www.gov.cn/xinwen/2020-09/29/5548176/files/1c6b4a2ae06c4ffc8bccb49da353495e.pdf

[12] On December 7, 2022, China lifted the most severe COVID policies by allowing infection patients to quarantine at home. https://www.bbc.com/news/world-asia-china-63855508

resources that influences lockdown decisions. The results of using hospitals and beds to measure medical capacity are shown in the extension part.

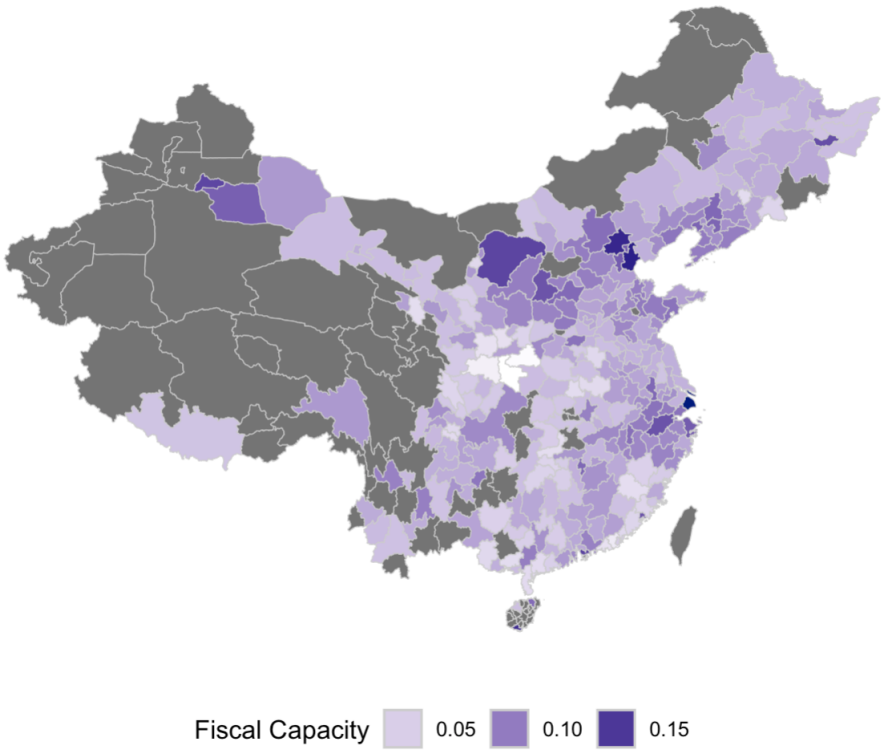
To test the second hypothesis, I am interested in the effect of fiscal capacity and monitoring capacity on lockdown decisions when holding infection cases constant. According to the conventional understanding of state capacity, fiscal capacity is essential to achieve the state's policy goal and it correlates with other dimensions of state capacity. Therefore the effect of fiscal capacity on lockdown decisions is expected to represent the effect of overall state capacity. Fiscal capacity is measured by a subnational government's revenue in the general public budget as a share of GDP. It indicates a subnational government's capacity to extract resources from the local community in the form of tax revenue. A city-level government's revenue in the general public budget refers to the overall revenue subtracted from the tax-sharing by the provincial government and the central government. Transfer payments for redistribution from the upper level are not included. Figure 4 shows the distribution of fiscal capacity across the country.

Data fabrication is a reasonable concern given early coverups of COVID infection cases by subnational governments. However, it soon became less of a problem due to the nature of the pandemic, effective monitoring from the top authorities, and high public scrutiny and expectation. It was extremely difficult to conceal infection cases as virus quickly spread to neighboring areas and suppressing reports outside the jurisdiction would not be easy for subnational officials (Zhu 2023). In addition, the zero-COVID policy is readily measurablethe and the System of Dynamic Surveillance on Coronavirus started to operate on January 24, requiring subnational governments to report cases on a daily basis via the system [13]. Also, the public effectively supervised subnational governments' performance through social media (Xu et al. 2020), further decreasing the possibility of reporting fake statistics.

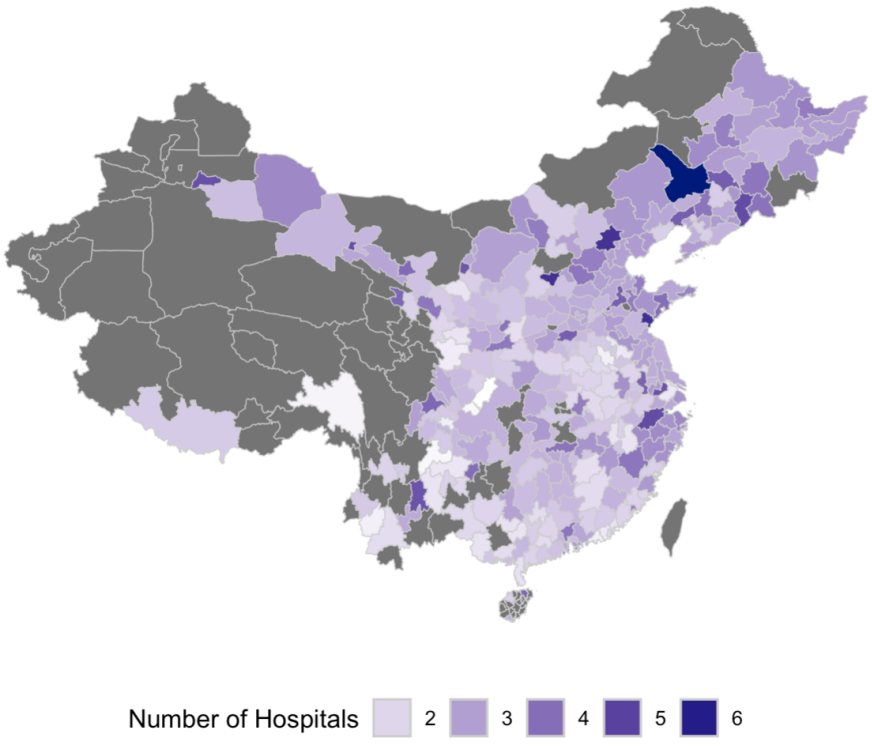
Similar to fiscal capacity, the pre-existing monitoring capacity could be an alternative explanation to variations in lockdown decisions. If a city has strong monitoring capacity, for instance, it has more effective tools for collecting citizens' information or implementing targeted pandemic control measures, it is more capable of avoiding a complete lockdown. Therefore, this analysis control monitoring capacity, measured by the number of high-resolution surveillance cameras per thousand capita purchased by city governments up to the end of 2019 Beraja, Yang, and Yuchtman (2023). This measure captures the stock of newer surveillance cameras at the time, but not the older ones. The focus on newer cameras is appropriate given their higher resolution and thus greater usefulness in collecting information and monitoring citizens.

Given that the pandemic broke out in January 2020, medical capacity, fiscal capacity, and monitoring capacity are measured in 2019 values. Data on the aggregate number of medical personnel, population,

[13] See https://www.chinacdc.cn/zxdt/202001/t20200125_211441.html



(a) Fiscal Capacity



(b) Medical Capacity

Figure 4: The Uneven Distribution of Fiscal Capacity and Medical Capacity

revenue, and GDP are collected from city-level yearbooks. The infection data are collected from the China Data Lab of Harvard Dataverse. The surveillance camera data are from replication data shared by Beraja, Yang, and Yuchtman (2023).

5.3 Control Variables

City leaders' lockdown decisions and citizens' willingness to move were largely captured by the risk of local outbreak as well as the previous experience with severe acute respiratory syndrome (SARS), another viral respiratory disease similar to COVID-19 spread across China between 2002 and 2003. SARS experience may influence both subnational governments' preparedness for the virus and people's awareness of the virus. On one hand, localities severely hit by SARS are more aware of the demand for adequate and advanced medical facilities and thus are expected to invest more in medical capacity and have a quicker response to the COVID outbreak (Zhou and Xin 2021). On the other hand, people who experienced SARS are more familiar with the effective way to prevent infection, and may voluntarily restrict mobility and follow government instructions. Thus this analysis controls cities' SARS experience. Since Beijing, Tianjin, Guangdong province, Shanxi province, and Inner Mongolia autonomous region are categorized as highly infected areas,[14] cities in these areas are coded as 1, while others are coded as 0.

Another factor that could influence lockdown decisions and people's mobility is the distance to Wuhan, which was the area with the greatest number of cases. Cities that are closer to Wuhan were more likely to experience severer situation, as residents from Wuhan city and Hubei Province could move to neighboring cities just before the lockdown. Also, people in the surrounding areas would feel the epidemic more vividly and be less skeptical about the risks. Therefore, both citizens and city leaders could be more sensitive to the increase of cases.

This analysis also control the weather, including temperature and precipitation. Because weather correlated with the spread of the virus and people's mobility, which could also impact governments' lockdown decisions. The two variables are coded as dummies. Given the outbreak was in winter, a temperature below 23 degrees Fahrenheit is coded as 1, which triggers a weather alert, and outdoor activities are not encouraged. Precipitation over 0.98 inches is coded as 1, which triggers a weather alert as well. Data on precipitation and temperature are collected from the China Data Lab of Harvard Dataverse. Table 1 shows the summary of variables.

[14] See http://www.gov.cn/test/2005-06/28/content_10716.htm

Table 1: Descriptive Statistics of Variables

Variables	Count	Mean	SD	Min	Max
Lockdown Status	281	0.36	0.48	0	1
Relative Mobility Change	8807	-0.08	0.32	-2.74	1.94
Increased Infection Cases	8807	1.46	4.31	0	201
Medical Personnel per 1000 Capita	281	2.74	0.81	0.14	6.42
Fiscal Capacity	281	0.08	0.25	0.02	0.19
SARS Experience	281	0.03	0.07	0	1
Distance to Wuhan	281	992.4	626.55	174	3596
Surveillance Camera per 1000 Capita	281	7.59	46.35	0	758.82
Population (million)	281	4.45	3.59	0.21	31.24
Temperature	8807	0.20	0.38	0	1
Precipitation	8807	0.03	0.16	0	1

6 Results

6.1 Medical Capacity Determines Lockdown Regulations

In my first set of empirical results, presented in Table 2, I describe the association between targeted capacity and the policy choice for pandemic control, measured by lockdown regulations officially announced by local governments. In line with my theoretical expectations, these results indicate a positive association between medical capacity and lockdown levels. That is, conditional upon the number of increased infection cases, cities with more doctors per thousand capita were more likely to keep the city open. In addition, capacities that are less relevant for pandemic control did not influence policy choice. That is, fiscal capacity and monitoring capacity have no significant effect on lockdown decisions.

In column (1), I estimate the baseline model, regressing the lockdown status, where 1 represents strict and indiscriminate outdoor restriction and 0 represents targeted closed-off management, on medical capacity. The number of newly-increased infection cases are controlled in all models. Because city leaders primarily make lockdown decisions based on newly-increased infected cases, which indicates the speed of transmission. Medical capacity does not impact lockdown decisions directly. Instead, the speed of transmission does. In this specification, holding infection cases constant, a one-doctor-per-thousand-capita increase yields a 15.39% decline in the probability of a city being completely locked down.

To explore the effect of different dimensions of state capacity on policy choice, in column (2), I include fiscal capacity and monitoring capacity. I also add controls for the distance to Wuhan, SARS experience, population, temperature, precipitation, latitude, and longitude. The effect of medical capacity remains unchanged. Additionally, I find no evidence that stronger fiscal capacity and monitoring capacity would

Table 2: Different Capacity Constraints on Lockdown Status

	Dependent Variable: Lockdown Status			
	(1)	(2)	(3)	(4)
Medical Capacity	-0.756***	-0.736**	-0.776***	-0.133***
	(0.226)	(0.241)	(0.258)	(0.037)
Fiscal Capacity		0.806	2.552	0.275
		(7.214)	(7.421)	(1.290)
Monitoring Capacity		-0.009	-0.006	-0.0004
		(0.014)	(0.010)	(0.002)
Model	Logit	Logit	Logit	OLS
Controls	No	Yes	Yes	Yes
Observations	278	278	226	278
Pseudo R-squared	0.109	0.166	0.121	
R-squared				0.178

Note: Robust standard errors are reported in parentheses. Dates from January 20 to February 20, 2020. Hubei Province is excluded. The longitude and latitude are controlled. Significance levels: *p<0.1; **p<0.05; ***p<0.01

facilitate a city leader to achieve the zero-COVID goal with more targeted measures. Moreover, it is interesting to see that SARS experience also significantly decreases the probability of a city to lockdown, holding infection cases constant. It is likely that cities with SARS experience were equipped with stronger targeted capability to deal with the pandemic.

Given that if a city did not make any official announcement of strict outdoor restriction, it is coded as 0, which implies that the city adopted more targeted closed-off management. Still, it is possible that cities without clear regulations were implementing strict measures. To avoid incorrectly categorizing this group of cities, column (3) removes observations with ambiguous lockdown policies and replicates the specification presented in column (2). The effect of medical capacity remains similar, and the coefficients of fiscal capacity and monitoring capacity remain insignificant.

Column (4) replicates the specification presented in column (2) using an OLS regression. As stated in the hypotheses, cities with stronger medical capacity adopted more targeted measures to achieve the goal of zero-COVID. In addition, cities with stronger fiscal capacity and monitoring capacity did not adopt significantly different lockdown decisions from others. These results indicate that state capacity is multidimensional, and targeted capacity plays a crucial role in determining policy choices.

6.2 Medical Capacity Determines Intracity Mobility

Though relative mobility change proposed by Zhu (2023) is the outcome of lockdown decisions, it could still be a qualified indicator of lockdown decisions, given the substantial implementation of lockdowns at subnational

levels. On the one hand, the public was highly compliant with governmental regulations during this crisis (Perry 2021). On the other hand, lockdown mandates were made feasible by the pre-existing structure of Chinese society, including urban gated housing, grid management, and the ubiquitous monitoring app (Qian and Hanser 2021; He, Shi, and Liu 2020). To further validate the previous findings, Table 3 presents the results using relative mobility change as the measure of lockdown levels.

Column (1) presents the results of the time-variant baseline model without any control, regressing the average change of intracity mobility on the previous day and medical capacity. Column (2) includes medical capacity, fiscal capacity, and monitoring capacity and presents the results of the intracity mobility change, with additional controls on the distance to Wuhan, SARS experience, population, and geographical coordinates. Column (3) presents the results of two-way fixed-effects OLS regression, with standard errors clustered by city, of the intracity mobility change on the increase of infected cases on the previous day and its interaction with the number of medical personnel per thousand capita. Column (4) replicates the specification in column (3), including the interaction terms of medical capacity, fiscal capacity, and monitoring capacity, and presents the results of the intracity mobility change with additional control variables.

Table 3: Different Capacity Constraints on Lockdown Decisions

	Dependent Variable: Relative Change of Mobility			
	(1)	(2)	(3)	(4)
Medical Capacity	0.011***	0.008***		
	(0.002)	(0.002)		
Fiscal Capacity		0.103		
		(0.063)		
Monitoring Capacity		0.00005		
		(0.00003)		
Medical Capacity			0.025***	0.024***
*log(Infection)			(0.005)	(0.006)
Fiscal Capacity				0.048
*log(Infection)				(0.173)
Monitoring Capacity				0.000
*log(Infection)				(0.0003)
City FE	Yes	Yes	Yes	Yes
Date FE	No	No	Yes	Yes
Controls	No	Yes	No	Yes
Observations	278	278	8,807	8,807
R-squared	0.353	0.456	0.003	0.006

Note: Robust standard errors, clustered by cities in column (3) and (4), are reported in parentheses. Dates from January 20 to February 20, 2020. Hubei Province is excluded. Significance levels: *p<0.1; **p<0.05; ***p<0.01

As stated in hypothesis 1, this paper expects cities with stronger medical capacity to have fewer lockdowns,

that is, greater mobility change. The positive and significant coefficients of medical capacity confirms this hypothesis, indicating that cities with more medical capacity adopt significantly fewer lockdowns. Testing hypothesis 2, the effects of fiscal capacity monitoring capacity also positively increase intracity mobility, which is to decrease large-scale lockdown decisions. However, the effects are insignificant. That said, cities with stronger fiscal capacity and monitoring capacity do not adopt significantly different lockdown decisions from others.

Substantively, column (4) shows that holding the previous day's infection constant, the relative mobility change of a city with one doctor increase for every thousand people will increase by 0.024. The substantive effect seems small, but note that the median relative mobility change in a city with a local outbreak is -0.08. The difference between cities with the highest and the lowest number of doctors per thousand population is more than 6, leading to a 0.144 increase in mobility. Also, the effect of medical capacity may not be fully captured and thus could be underestimated in this analysis. Varying levels of lockdowns were imposed synchronically even when there was no outbreak (Zha et al. 2022). These lockdowns are preventive and could be explained by medical capacity but are not captured as infection case remains zero. In sum, I find the effect of medical capacity is substantial and the hypothesis of targeted capacity is supported, while the effects of fiscal and monitoring capacity on lockdown decisions are insignificant.

7 Robustness Checks, Parallel Trend, and Extension

7.1 Robustness Checks

I conduct three robustness checks and results are presented in table 4. First, one might be concerned that the time period from January to February is too limited to reveal the effect, or wonder what happened after the most acute period of the early outbreak. I extend the period from January to April and see how the results change. I choose the time span given that by the end of March, the rate of work resumption exceeded 90% and people's intracity mobility intensity climbed back to the level before the outbreak [15]. And in early May, all county-level areas in China, including those in Hubei province, were adjusted to below low-risk [16]. I expect to see that the effect of medical capacity will decrease while the effect of fiscal capacity will increase. Because subnational units that have greater fiscal capacity were then capable of building more mobile cabin hospitals or employing more staff for contact tracing and testing. The results in column (1) imply that the effect of medical capacity on lockdown decisions remains consistently significant while the effect of fiscal

[15] See http://www.xinhuanet.com/politics/2020-03/18/c_1125730627.htm
[16] See http://www.xinhuanet.com/politics/2020-05/07/c_1125952945.htm

capacity increases.

Second, one might think that border cities could confound the result. The Chinese government has been cautious about both the spread of the coronavirus within the country and the import of infected cases beyond the country. China borders 14 countries on land and border cities have been under greater pressure over epidemic prevention than inland cities. For instance, Ruili is a southwestern city in China bordering Myanmar. The land border is long and impossible to be fully guarded and therefore infected cases keep flowing in. To minimize the risk of spreading cases to other cities, Ruili went through more than 7 months of lockdown in a year, with the highest daily increase of 15 cases.[17] Thus lockdown decisions in border cities could be explained by a different mechanism other than capacity constraints. I assign 22 border cities in my sample a dummy variable and run the test.[18] Results shown in column (2) are consistent with the original model.

Table 4: Robustness Tests

	Dependent Variable		
	Mobility Change	Lockdown Status	Lockdown Status
Medical Capacity	0.014***	-0.719**	-0.715**
	(0.004)	(0.246)	(0.248)
Fiscal Capacity	0.055	1.774	0.941
	(0.113)	(7.381)	(7.384)
Monitoring Capacity	0.0001	-0.008	-0.009
	(0.0002)	(0.014)	(0.014)
Model	OLS	Logit	Logit
City FE	Yes	Yes	Yes
Date FE	Yes	No	No
Controls	Yes	Yes	Yes
Observations	25,847	278	278
R-squared	0.008		
Pseudo R-squared		0.170	0.174

Note: Robust standard errors, clustered by cities in column (1), are reported in parentheses. In column (1), the dependent variable is measured by relative change of mobility, and the dates are from January 20 to April 20, 2020. In column (2) and column (3), the dependent variables are measured by lockdown regulations, and the dates are from January 20 to February 20, 2020. Hubei Province is excluded from all models. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Three, one might ask whether local officials' political intent would confound the result. Zhu (2023) argues that as officials are single-minded, those who have achieved the goal of poverty reduction, another

[17] See https://theinitium.com/article/20211102-mainland-ruili-lockdown/

[18] The list of border cities is retrieved from the government's website. http://www.gov.cn/test/2006-07/14/content_335831.htm

central mandate that was required to be achieved by the end of 2020, would prefer lockdowns as they are unwilling to focus on both economic development and pandemic control. Lockdown would be a good excuse for officials in wealthier cities to ward off economic development goals and only focus on pandemic control. While for localities that were still struggling down the poverty line, officials would prefer fewer lockdowns for promoting economic recovery. I, therefore, assign cities with GDP per capita above the first quantile a dummy variable that captures being a rich city, which is the group I expect to be more likely to choose a lockdown, and run the test. Results shown in column (3) does not challenge the original model.

7.2 Parallel Trend Test

In models with lockdown decisions measured by relative mobility change, the city and date fixed effects control for all time-invariant attributes that could confound the estimate of the effect of medical capacity and infection. Therefore, the causal identification of the treatment effect in the specification relies on a parallel-trends assumption: that the observed relative mobility change among different cities would have the same trend if a large number of infection cases were not present and medical capacity was not in high demand. Note that human-to-human transmission of the virus was confirmed on January 20, 2020, and the first full-scale lockdown order was announced on January 23. In other words, different patterns of mobility induced by lockdown decisions, which are assumed to be determined by different levels of medical capacity and infection, would not emerge earlier than those days. The figure 5 shows that estimates of relative mobility changes within cities before January 26, are consistent with the null hypothesis of no differential trends. The parallel trend thus holds.

7.3 Extension

Medical capacity could be measured not only by medical personnel, given that hospitals and beds are also critical medical resources. It is thus necessary to justify the choice of medical personnel as the measure of medical capacity. To see whether different measures for medical capacity would have a similar effect on lockdown decisions, I collect the data on hospitals and beds and run the test. Table 5 shows the results of three models with the independent variable as the number of medical personnel, beds, and hospitals per thousand population, respectively. The effect of beds is negative but insignificant, while the effect of hospitals is inconsistent with the original model.

Potential causes of the heterogeneous effect of different measures of medical capacity are briefly discussed in the variable selection part. First, there could be a large gap between the recorded number and the number of hospitals and beds actually put into use. As self-quarantined has not been approved as a treatment for

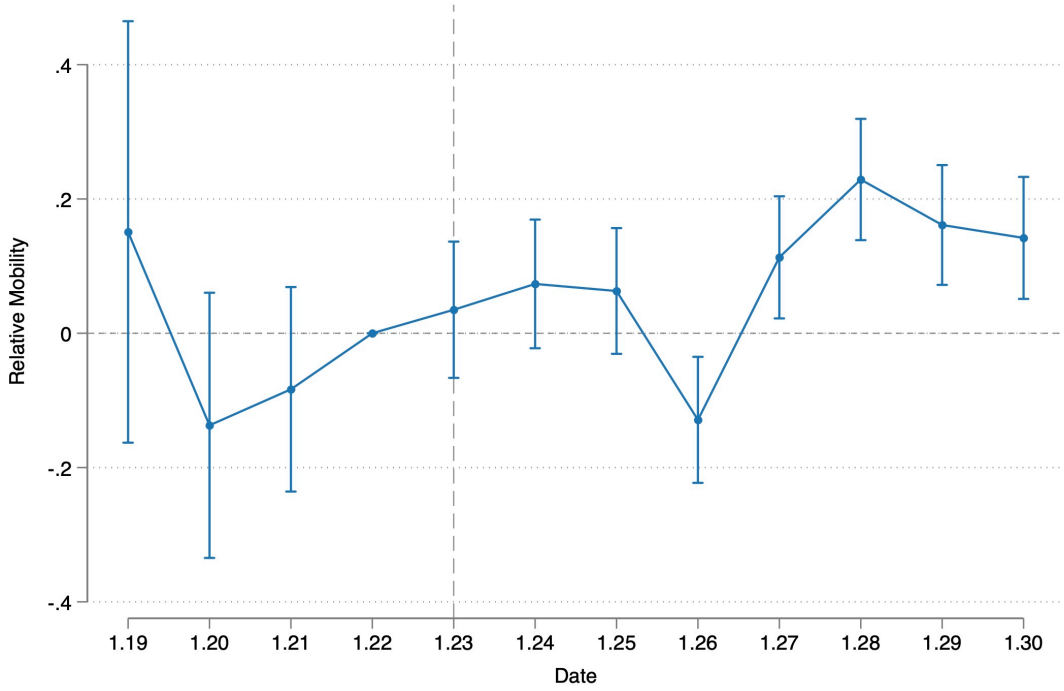


Figure 5: Parallel Trend Test

infected patients in China until the end of 2022, every patient has to be transferred to a nearby hospital or quarantine camp approved to treat COVID-19 patients. Hospital beds are expected to be a valid indicator of medical capacity. However, according to the ninth version of the National Diagnosis and Treatment Guideline released by China's National Health Commission, only designated hospitals are authorized to admit infected cases.[19] As such, many public hospitals are disqualified. Community hospitals and private-run clinics were even forced to close, aiming at reducing the risk of misdiagnose and failure to detect early the spread of the virus. Therefore, the number of hospitals and beds could be less reliable. In addition, by the end of 2019, public hospitals accounted for only 34.73% of the total number of hospitals in China. Yet, public hospitals own 72.46% of beds and employ 78.59% of doctors.[20] This disparity may explain the differing signs observed between hospital and the other two measures.

Second, the number of medical personnel is valid since, unlike facilities, healthcare workers could be flexibly deployed. Doctors and nurses employed by hospitals that are not designated to treat infected patients

[19]The last version of guideline before China lifted the most severe policies by announcing "10 Prevention and Control Measures". See http://www.nhc.gov.cn/jkj/s3577/202206/de224e7784fe4007b7189c1f1c9d5e85/files/504a946af7e744fb9ad7eb1e0f1f9923.pdf

[20]See http://www.nhc.gov.cn/guihuaxxs/s10748/202006/ebfe31f24cc145b198dd730603ec4442.shtml

Table 5: Heterogeneous Effects of Medical Capacity with Different Measures

	Dependent Variable: Lockdown Level		
	(1)	(2)	(3)
Personnel per thousand	-0.736**		
	(0.241)		
Beds per thousand		-0.117	
		(0.120)	
Hospitals per thousand			0.570
			(14.783)
Fiscal Capacity	0.806	-4.326	-3.538
	(7.214)	(6.882)	(6.854)
Monitoring Capacity	-0.009	-0.010	-0.010
	(0.014)	(0.014)	(0.014)
Model	Logit	Logit	Logit
Controls	Yes	Yes	Yes
Observations	278	270	277
R-squared	0.166	0.172	0.150

Note: Robust standard errors are reported in parentheses. Dates from January 20 to February 20, 2020. Hubei Province is excluded. The longitude and latitude are controlled. Significance levels: *p<0.1; **p<0.05; ***p<0.01

could be transferred to support the frontline. At the height of the epidemic in early 2020, healthcare workers across the country were mobilized to support Hubei province.[21] Furthermore, medical personnel includes not only doctors and nurses working in hospitals. There are four categories of professional staff, including medical care, nursing, pharmacy, and technicist, among which epidemiological researchers are as significant as doctors during the pandemic. The capability of contact tracing is the key to cutting the transmission channel of the virus, and lockdown decisions relied greatly on epidemiological researchers' analysis. Thus city leaders may rely more on the number of medical personnel than on hospitals and beds to make lockdown decisions, leading to heterogeneous effects of different measures.

8 Conclusion

China was the last major power to lift its stringent measures against the virus. The large-scale lockdown proved to be an effective short-term tool for containing the spread of a viral epidemic (Yang, 2021). However, overly-cautious countermeasures and excessive lockdowns were detrimental to economic recovery and people's normal life. Similarly, relaxing control without considering the risk of the virus has led to an increase in the death toll and infection. Given the potentially disruptive impact of policy choices at both extremes, the

[21] Sixteen provinces were required by the National Health Commission to support Sixteen cities of Hubei province. See https://news.sina.com.cn/c/2020-02-08/doc-iimxxste9694826.shtml

driving force behind subnational governments' heterogeneous responses to the top authority's zero-COVID mandate deserves attention.

This paper argues that capacity contingent on the policy goal is crucial for understanding the effect of state capacity on policy implementation. Further, state capacity is multidimensional and different dimensions are not necessarily correlated. Taking pandemic control in China as a case, this paper finds that medical capacity shapes Chinese subnational governments' lockdown decisions and it is uncorrelated with fiscal capacity. Specifically, cities with stronger medical capacity are more likely to adopt targeted measures and avoid large-scale lockdowns and vice versa. This effect is more prominent for cities experiencing a more severe wave and could decrease as the pandemic stabilizes. Additionally, fiscal capacity does not significantly influence lockdown decisions. Still, it is expected that cities with stronger fiscal capacity can be more adaptive and resilient in the long run. This paper also demonstrates the multidimensionality of state capacity and that Chinese cities with strong fiscal capacity do not have strong medical capacity.

This paper contributes to the literature on state capacity. It challenges the idea that different dimensions of state capacity are necessarily correlated and underscores the significance of policy context when examining the effect of state capacity. This analysis shows that it is misleading to ignore the relationship between different capacities, the context, and within-country variation when understanding state capacity. Furthermore, this paper contributes to intent-focused literature on policy implementation under authoritarianism. It does not reject the significance of political intent driven by formal and informal institutions on local officials' behavior. However, it argues that capacity creates the space for intent. To understand subnational variation in policy implementation, more attention should be paid to governments' capacity constraints. These findings could be applied to understanding cases beyond China and pandemic control.

Future research may ask: what causes the mismatched distribution across different dimensions of state capacity? How do capacity constraints interact with political intent in decision-making and policy implementation?

This paper has a number of limitations. First, the model cannot rule out spillover effects among cities. A city could adopt lockdowns not because of infection increase but as a measure to prevent the virus from spreading from neighboring areas or people escaping from neighboring cities that plan to implement lockdowns. Still, it should be noted that the scale of outbreaks in early 2020 outside Hubei Province was mostly mild and confined to parts of a city (Zhu 2023; Wang et al. 2021). Therefore, the spillover effects, determined by the scale of outbreaks, may not be substantial. Second, the measurement of medical capacity could be imprecise. The number of contact tracing staff, who were in charge of effectively controlling the spread of the virus, could be a better indicator of medical capacity that may shape lockdown decisions; however, there is no reliable data for such a measure. Contact tracing more directly determines if a city is

capable of conducting a targeted strategy, while the number of doctors and nurses that could treat infected patients is less relevant and more like the safety net when the situation is out of control.

Four years after the first COVID-19 outbreak, Chinese people are still recovering from the life and economic losses caused by a series of unreasonable regulations in lockdowns and reopening. A deeper understanding of the mechanism behind these decisions may help prevent more tragedies in the future.

References

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

Acemoglu, Daron, Camilo García-Jimeno, and James A Robinson. 2015. "State capacity and economic development: A network approach". *American economic review* 105 (8): 2364–2409.

Acemoglu, Daron, Simon Johnson, and James A Robinson. 2002. "Reversal of fortune: Geography and institutions in the making of the modern world income distribution". *The Quarterly journal of economics* 117 (4): 1231–1294.

Acemoglu, Daron, Jacob Moscona, and James A Robinson. 2016. "State capacity and american technology: evidence from the nineteenth century". *American Economic Review* 106 (5): 61–67.

Alesina, Alberto F, Edward L Glaeser, and Bruce Sacerdote. 2001. "Why doesn't the us have a european-style welfare system?".

Bates, Robert H. 2020. *The Political Economy of Development*. Cambridge University Press.

Baunsgaard, Thomas and Michael Keen. 2010. "Tax revenue and (or?) trade liberalization". *Journal of Public Economics* 94 (9-10): 563–577.

Beraja, Martin, David Y Yang, and Noam Yuchtman. 2023. "Data-intensive innovation and the state: Evidence from ai firms in china". *The Review of Economic Studies* 90 (4): 1701–1723.

Bersch, Katherine, Sérgio Praça, and Matthew M Taylor. 2017. "State capacity, bureaucratic politicization, and corruption in the brazilian state". *Governance* 30 (1): 105–124.

Berwick, Elissa and Fotini Christia. 2018. "State capacity redux: Integrating classical and experimental contributions to an enduring debate". *Annual Review of Political Science* 21 : 71–91.

Besley, Timothy and Torsten Persson. 2009. "The origins of state capacity: Property rights, taxation, and politics". *American economic review* 99 (4): 1218–1244.

Besley, Timothy and Torsten Persson. 2011. "Pillars of prosperity". In *Pillars of Prosperity*. Princeton University Press.

Besley, Timothy and Torsten Persson. 2014. "Why do developing countries tax so little?". *Journal of economic perspectives* 28 (4): 99–120.

Boin, R Arjen and Marc HP Otten. 1996. "Beyond the crisis window for reform: Some ramifications for implementation". *Journal of Contingencies and Crisis Management* 4 (3): 149–161.

Chen, Jingjing, Wei Chen, Ernest Liu, Jie Luo, and Zheng Michael Song. 2022. "The economic cost of locking down like china: Evidence from city-to-city truck flows". *OpenScholar@ Princeton*: 1–43.

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 : 84–101.

Cingolani, Luciana. 2013. "The state of state capacity: a review of concepts, evidence and measures".

Diehl, Paul F. 1983. "Arms races and escalation: A closer look". *Journal of Peace Research* 20 (3): 205–212.

Dincecco, Mark. 2011. *Political transformations and public finances: Europe, 1650–1913*. Cambridge University Press.

Dincecco, Mark and Yuhua Wang. 2022. "State capacity in historical political economy".

Evans, Peter P, Peter B Evans, Dietrich Rueschemeyer, and Theda Skocpol. 1985. *Bringing the state back in*. Cambridge University Press.

Fearon, James D and David D Laitin. 2003. "Ethnicity, insurgency, and civil war". *American political science review* 97 (1): 75–90.

Foa, Roberto Stefan and Anna Nemirovskaya. 2016. "How state capacity varies within frontier states: A multicountry subnational analysis". *Governance* 29 (3): 411–432.

Gerth, Hans Heinrich and C Wright Mills. 2014. *From Max Weber: essays in sociology*. Routledge.

Gingerich, Daniel W. 2013. "Governance indicators and the level of analysis problem: empirical findings from south america". *British Journal of Political Science* 43 (3): 505–540.

Gong, Peng, Song Liang, Elizabeth J Carlton, Qingwu Jiang, Jianyong Wu, Lei Wang, and Justin V Remais. 2012. "Urbanisation and health in china". *The Lancet* 379 (9818): 843–852.

Greitens, Sheena Chestnut. 2017. "Rethinking china's coercive capacity: an examination of prc domestic security spending, 1992–2012". *The China Quarterly* 232 : 1002–1025.

Gu, Edward and Lantian Li. 2020. "Crippled community governance and suppressed scientific/professional communities: a critical assessment of failed early warning for the covid-19 outbreak in china". *Journal of Chinese governance* 5 (2): 160–177.

Hanson, Jonathan K and Rachel Sigman. 2021. "Leviathan's latent dimensions: Measuring state capacity for comparative political research". *The Journal of Politics* 83 (4): 1495–1510.

Harding, Harry. 1981. *Organizing China: the problem of bureaucracy, 1949-1976*. Stanford University Press.

Hart, Paul't. 1993. "Symbols, rituals and power: The lost dimensions of crisis management". *Journal of contingencies and crisis management* 1 (1): 36–50.

He, Alex Jingwei, Yuda Shi, and Hongdou Liu. 2020. "Crisis governance, chinese style: Distinctive features of china's response to the covid-19 pandemic". *Policy Design and Practice* 3 (3): 242–258.

Henderson, Errol A and J David Singer. 2000. "Civil war in the post-colonial world, 1946-92". *Journal of Peace research* 37 (3): 275–299.

Hendrix, Cullen S. 2010. "Measuring state capacity: Theoretical and empirical implications for the study of civil conflict". *Journal of peace research* 47 (3): 273–285.

Herbst, Jeffrey. 1990. "War and the state in africa". *International Security* 14 (4): 117–139.

Jones, Daniel M, Stuart A Bremer, and J David Singer. 1996. "Militarized interstate disputes, 1816–1992: Rationale, coding rules, and empirical patterns". *Conflict Management and Peace Science* 15 (2): 163–213.

Kugler, Jacek. 2018. *Political capacity and economic behavior*. Routledge.

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

Lee, Alexander. 2019. "Land, state capacity, and colonialism: Evidence from india". *Comparative Political Studies* 52 (3): 412–444.

Lee, Alexander and Jack Paine. 2022. "The great revenue divergence". *International Organization*: 1–42.

Lee, Melissa M and Nan Zhang. 2017. "Legibility and the informational foundations of state capacity". *The Journal of Politics* 79 (1): 118–132.

Levi, Margaret. 1988. *Of rule and revenue*. Univ of California Press.

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): 1743–1762.

Lieberman, Evan S. 2002. "Taxation data as indicators of state-society relations: possibilities and pitfalls in cross-national research.". *Studies in Comparative International Development* 36 (4).

Lieberthal, Kenneth G and David M Lampton. 2018. *Bureaucracy, politics, and decision making in post-Mao China*, Volume 14. University of California Press.

Lindvall, Johannes and Jan Teorell. 2016. *State capacity as power: a conceptual framework*. Department of Political Science, Lund University.

Linz, Juan J, Juan J Linz, Alfred Stepan, et al. 1996. *Problems of democratic transition and consolidation: Southern Europe, South America, and post-communist Europe*. jhu Press.

Liu, Dongshu. 2019. "Punish the dissidents: The selective implementation of stability preservation in china". *Journal of Contemporary China* 28 (119): 795–812.

Lü, Xiaobo and Pierre F Landry. 2014. "Show me the money: Interjurisdiction political competition and fiscal extraction in china". *American political science Review* 108 (3): 706–722.

Mamdani, Mahmood. 1996. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton University Press.

Mann, Michael. 1984. "The autonomous power of the state: its origins, mechanisms and results". *European Journal of Sociology/Archives européennes de sociologie* 25 (2): 185–213.

Mattingly, Daniel C. 2017. "Colonial legacies and state institutions in china: Evidence from a natural experiment". *Comparative Political Studies* 50 (4): 434–463.

Migdal, Joel S. 1988. *Strong societies and weak states: state-society relations and state capabilities in the Third World*. Princeton University Press.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Political Economy of Institutions and Decisions. Cambridge University Press.

O'Donnell, Guillermo. 1993. "On the state, democratization and some conceptual problems: A latin american view with glances at some postcommunist countries". *World Development* 21 (8): 1355–1369.

Perry, Elizabeth J. 2021. "China's (r) evolutionary governance and the covid-19 crisis". In *Evolutionary Governance in China*, pp. 387–396. Harvard University Asia Center.

Prendergast, Canice. 2007. "The motivation and bias of bureaucrats". *American Economic Review* 97 (1): 180–196.

Qian, Yue and Amy Hanser. 2021. "How did wuhan residents cope with a 76-day lockdown?". *Chinese Sociological Review* 53 (1): 55–86.

Qiu, Yun, Xi Chen, and Wei Shi. 2020. "Impacts of social and economic factors on the transmission of coronavirus disease 2019 (covid-19) in china". *Journal of population economics* 33 : 1127–1172.

Rogers, Melissa Ziegler and Nicholas Weller. 2014. "Income taxation and the validity of state capacity indicators". *Journal of Public Policy* 34 (2): 183–206.

Skocpol, Theda. 1979. *States and social revolutions: A comparative analysis of France, Russia and China*. Cambridge University Press.

Skocpol, Theda. 1985. "Bringing the state back in: strategies of analysis in current research". *Bringing the state back in* 9 .

Skocpol, Theda and Kenneth Finegold. 1982. "State capacity and economic intervention in the early new deal". *Political science quarterly* 97 (2): 255–278.

Soifer, Hillel. 2008. "State infrastructural power: Approaches to conceptualization and measurement". *Studies in comparative international development* 43 : 231–251.

Solinger, Dorothy J and Ting Jiang. 2016. "When chinese central orders and promotion criteria conflict: implementation decisions on the destitute in poor versus prosperous cities". *Modern China* 42 (6): 571–606.

Straus, Scott. 2013. "The order of genocide". In *The Order of Genocide*. Cornell University Press.

Thornton, Patricia M. 2009. "Crisis and governance: Sars and the resilience of the chinese body politic". *The China Journal* (61): 23–48.

Tilly, Charles. 1975. "Reflections on the history of european state-making". *The formation of national states in Western Europe* 38 : 3–83.

Tilly, Charles et al. 1992. *Coercion, capital, and European states, AD 990-1992*. Blackwell Oxford.

Wang, Huan, Markus Zhang, Robin Li, Oliver Zhong, Hannah Johnstone, Huan Zhou, Hao Xue, Sean Sylvia, Matthew Boswell, Prashant Loyalka, et al. 2021. "Tracking the effects of covid-19 in rural china over time". *International Journal for Equity in Health* 20 (1): 1–13.

Wayman, Frank W, J David Singer, and Gary Goertz. 1983. "Capabilities, allocations, and success in militarized disputes and wars, 1816–1976". *International Studies Quarterly* 27 (4): 497–515.

Xu, Chenggang. 2011. "The fundamental institutions of china's reforms and development". *Journal of economic literature* 49 (4): 1076–1151.

Xu, Junfang, Yu Shi, Feng Cheng, and Wannian Liang. 2021. "China's public health system: time for improvement". *The Lancet Public Health* 6 (12): e869–e870.

Xu, Qing, Ziyi Shen, Neal Shah, Raphael Cuomo, Mingxiang Cai, Matthew Brown, Jiawei Li, Tim Mackey, et al. 2020. "Characterizing weibo social media posts from wuhan, china during the early stages of the covid-19 pandemic: qualitative content analysis". *JMIR Public Health and Surveillance* 6 (4): e24125.

Yin, Gaofeng, Hanning Song, Jian Wang, Stephen Nicholas, and Elizabeth Maitland. 2021. "The covid-19 run on medical resources in wuhan china: causes, consequences and lessons". In *Healthcare*, Volume 9, pp. 1362. Multidisciplinary Digital Publishing Institute.

Zha, Hao, Yuxi Zhang, Hui Zhou, Shanghai Jiao Tong, Lijun Wang, Zihan Zhang, Zijia Tan, Longmei Deng, and Thomas Hale. 2022. "Chinese provincial government responses to covid-19". Technical report, Version 2. Working Paper, Blavatnik School of Government, Oxford University.

Zhou, Kai and Ge Xin. 2021. "Who are the front-runners? unravelling local government responses to containing the covid-19 pandemic in china". *China Review* 21 (1): 37–54.

Zhou, Xueguang. 2022. *The Logic of Governance in China: An Organizational Approach.* Cambridge University Press.

Zhu, Hongshen. 2023. "Policy under conflicting mandates: Evidence from 1 billion cellphones during china's covid lockdowns". *Under review*.