Political Effect of Economic Data Manipulation: Evidence from Chinese Protests

by

Handi Rita Li

Department of Political Science
Duke University

Date: ______________________

Approved:

____
Timur Kuran, Supervisor

____
Edmund Malesky

____
Kyle Beardsley

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Political Science in the Graduate School of Duke University
2017

Abstract

Political Effect of Economic Data Manipulation: Evidence from Chinese Protests

by

Handi Rita Li

Department of Political Science
Duke University

Date: ____________________
Approved:

____
Timur Kuran, Supervisor

____
Edmund Malesky

____
Kyle Beardsley

An abstract of a thesis submitted in partial fulfillment of the requirements for
the degree of Master of Arts in the Department of Political Science
in the Graduate School of Duke University
2017

Abstract

As the uses of economic statistics broaden, data manipulation occurs more widely across regions, among countries with or without democratic institutions, and at different levels of governments. Government agencies release falsified data in the hope of obtaining higher public evaluations of governmental performance. Their ultimate goal is to maintain public support for the political status quo. Using the case of Chinese protests from 1995 to 2013, this paper explores the effects on the state-society relationship of economic data manipulation. The paper first provides some evidence for China's local GDP falsification. Then, by comparing the influence of reported GDP growth and that of actual economic growth on protests in China's provinces, I find that while actual growth has had a consistently negative correlation with the number of protests, the reported data has little effect. The result holds under a series of robustness checks. It implies that citizens are able to interpret economic performances through their feeling for financial status. Another implication is that falsified economic data does not help alleviate their grievances during economic downturns.

To the beauty of ivory tower.

Contents

List of Tables

List of Figures

Acknowledgements

I wish to express my appreciation to those who generously helped me on this paper. My greatest debt of gratitude goes to my committee members, Prof. Timur Kuran, Prof. Eddy Malesky and Prof. Kyle Beardsley, who have continuously supported me with their time, wisdom and encouragement since the very beginning of this project. I would also like to thank Prof. Chris Johnston, Prof. Erik Wibbels and Prof. Dan Stegmueller, who offered great help when I got stuck. Fengming Lu, Chong Chen, Howard Liu, and many other graduate students also generously contributed to my study by offering very helpful comments and suggestions. Without any of the above supports, I could not have got today's progress.

I am also grateful to my family and friends, who have always been standing by me as I struggle to take my spiritual practice in an intellectual world. I will cherish your love and keep working hard in the future.

1

Introduction

We are living in a time when economic information can be accessed cheaply and easily. As releasing macro-economic data has become a common governmental practice, citizens of both developed and developing countries, in both democratic and non-democratic regimes, are able to learn about economic growth and many other dimensions of economic performances through periodic official reports. Precisely for this reason, governments have incentives to manipulate the data, because economic performance is highly linked with not only meritocratic promotion, but also social stability. During hard times, governments can manipulate data at small and delayed costs. They can thereby maintain the support of citizens, and dampen public opposition, simply by shaping the citizenry's knowledge of economic conditions.

The falsification of economic data is a phenomenon observed across continents at both central and local levels of governments. China's local GDP data has long been doubted as deliberately overestimated, which was recently admitted by the governor of one of its provinces, saying that the province had released fake data for years because its real economy "failed to meet the expected targets."[1] In Argentina,

[1] `http://money.cnn.com/2017/01/18/news/economy/china-province-falsified-economic-data/`

even the central government continuously falsified macroeconomic data, including its backsliding growth as well as rocketing inflation and debts.[2] At the end of 2016, after an unsuccessful coup and under some domestic opposition to the ruling party, the Turkish government revised its GDP statistics upward. This has been demonstrated by the economist Erik Meyersson as largely unexplained and significantly deviate from real GDP predicted by proxies.[3]

These countries' economic data manipulation, blatant or superstitious, raises the question of whether they can achieve their goal of ameliorating people's feeling of bad economic performance. In authoritarian regimes, where data manipulation is most likely, do citizens believe the official data, and insofar as they do, does that stabilize the society? Are people more sensitive to government-published economic statistics or the true economy? This article seeks to answer these questions by comparing correlations between social protest and, on the one hand, with reported Gross Domestic Product (GDP) growth, and, on the other hand, the real economy development of localities in China. After reviewing the received literature, this article will first provide some context, with an emphasis on the local GDP[4] falsification and protests in China during 1995-2013. The next step is to introduce the research design using negative-binomial models on protests with reported and real economic growth, which is proxied by the variation of satellite-generated night-time luminosity. Following an analysis of empirical results, the final section concludes and draws broader implications. The article's main empirical results are a consistently negative

[2] `http://www.economist.com/blogs/americasview/2013/09/argentinas-official-statistics`

[3] Since December 2016, Meyersson has posted a series of reports on his personal website, pointing out that, compared to the old and to other countries' data, the newly released GDP data has strange patterns that would be absent if the data is true. He also find a conspicuous gap between the reported GDP and a series of commonly used economic proxies such as energy consumption, freight, electricity generation, etc. Furthermore, Turkey's stock market performances have shown the period of economic downturn, of which the GDP is upward revised. `https://erikmeyersson.com/blog/`

[4] This article follows China's official denomination for gross regional product, "local GDP", rather than "GRP."

correlation between the count of protests and luminosity growth, and a lack of significant correlation between protests and reported GDP growth, no matter the data falsification is extensive or not. There is evidence, then, that the Chinese people respond to actual growth and not to the reported growth statistics.

This article contributes to the existing literature in three ways: (1) It empirically shows that the Chinese local governments' reported GDP statistics on protests deviate from the indication of the real economy, further supporting the existence of a data manipulation. (2) More importantly, it shows that the manipulated statistics do not reduce protests or contribute to social stability as real economic growth does. (3) With evidence from China, the article also supports the theory that macro-economic growth leads to less, rather than more, political conflicts.

2

Data Manipulation and the Influence on Protests

2.1 Incentives for Economic Data Manipulation

The political science literature on data manipulation focuses on the underlying incentives rather than its effects or outcomes. Political scientists largely argue that poor economic performance can be a great threat to authoritarian regimes, because economic downturn is likely to cause democratization and the demise of autocracy (Geddes, 1999; Gandhi, 2008; Acemoglu and Robinson, 2005). Wallace (2015) suggests that falsified data can conceal an economic downturn, giving the government time to alleviate the problem, and thus help the government survive the crisis. On the local level, the survival of government amounts to the local leaders' relegation or promotion. In autocratic countries like China, despite of the importance of within-elite networks (Shih et al., 2012), local officials get promoted largely on the basis of their regions' economic performance and social stability (Li and Zhou, 2005; Lai, 2010; Lorentzen, 2013). Other than a merit-claim to the local leader's superiors, better economic statistics can also discourage civilians from mobilizing against the government (Hollyer et al., 2011, 2015).

In some countries with democratic electoral institutions, manipulating economic data becomes a way to avoid losing votes for unsatisfying economic performance. Voters' tendency to focus on the macroeconomy close to elections makes the incumbent government highly motivated to falsify economic statistics, especially at the time around elections. Countries where "election-motivated" data manipulation has occurred include Russia, Turkey, Mexico, and even the United States (Healy and Lenz, 2014). During the tenure of the populist president Cristina Fernández, Argentina continuously falsified data to keep citizens' support, even though the fake data cost the country billions of dollars.[1]

Furthermore, compared to the above immediate benefits, the manipulation of economic statistics generates only delayed costs to the government. The main cost of falsifying data is losing governmental credibility, which happens only when the falsification is revealed to the public. The cost is not paid immediately (Wallace, 2015), even though the distrust can persist for quite a long time once data consumers lose confidence in the government. If the untruthful leader only expects a short-term rule, as is the case in China's localities[2], data manipulation costs him/her very little. Therefore, tampering with economic growth data can be an optimal way for authoritarian governments to alleviate the tensions between poor economic results and social stability during their tenures. Empirical evidence also supports this argument. For instance, Hollyer, Rosendorff, and Vreeland (2011, 2015) find in a global public database that non-democratic governments tend to have lower availability of policy-relevant data than their democratic counterparts. Wallace (2015) demonstrates a

[1] Issuing a GDP-indexed bond, Argentina should pay more to its foreign debt holders when it has a higher GDP growth. In 2013, the officially falsified growth data caused Argentina to pay 3.5 billion dollars more foreign debts than it should have paid.
http://www.economist.com/blogs/americasview/2013/09/argentinas-official-statistics

[2] According to the Chinese Constitution, the tenure of local governments is five years. (*The Constitution of People's Republic of China (1982)*, Article 98 and 106 of Section 5, Chapter 3) Even though not written in law, usually the term can be renewable only once.

correlation between China's GDP falsification and leadership turnover on the local level.

2.2 Protests and Official Economic Data

Although the incentives of this seemingly clever scheme constitutes an interesting topic, it is also important to explore the outcomes in order to judge whether it is virtually "clever." Few publications focus on the socio-political effects of economic data manipulation. Hollyer et al. (2015) demonstrate that data transparency leads to higher instability in autocracies, and that this might happen through protests. In other words, intransparency might contribute to autocratic stability by reducing protests. However, their theory only captures the concealment, rather than falsification, of economic data. The latter is even cheaper and more deceptive, because compared to falsification, the withholding of data does not signal a good economic performance to citizens, and risks their immediate mistrust. We may therefore suppose that better-looking, falsified data can help reduce protests against the government and stabilize the society.

Before testing this argument, we should first explore theoretically the influence on protests both of economic data and of real economic development.

The relevant theories are mainly from two fields. In the context of civil conflicts and social protests, many theories suggest that protests are more likely to occur in poor economies. For example, the grievance theory claims that better macroeconomic performance usually provides people the expectation of rising incomes and public welfare, which reduces social grievances (Hendrix et al., 2009). By contrast, when people learn that the social economy becomes worse, dissatisfaction to the government is aroused, and the possibility of protests is heightened (Alesina et al., 1996; Londregan and Poole, 1990), even if the poor economic outcome is sometimes caused by a shock exogenous to the government's performances (Alesina et al., 1993).

From this group of theories, we can assume that economic data falsification is likely to lower the amount of protests, since it suggests that protest is a function of the information of economic performance, which the manipulated statistics can provide.

In another theory, rebellions and social conflicts are determined by relative deprivation, which is defined by Ted Gurr (1968, 1970) as "actors' perception of discrepancy between their value expectations (the goods and conditions of the life to which they believe they are justifiably entitled) and their value capabilities (the amounts of those goods and conditions that they think they are able to get and keep)." Such a discrepancy can be found in comparisons between self and others, as well as between past and future. It is widely used to explain social conflicts, especially urban riots (Walker and Smith, 2002). Relative deprivation theory implies that real economic performance is essential for citizens to assess the government, while upward manipulated economic data might even add to the sense of deprivation, for it heightens people's expectation, and thus further enlarges the gap between their expectation and real financial status.

A third group of theories known collectively as modernization theory suggests that economic growth encourages protest rather than prevent it, because growth is always accompanied with social structural destabilization, wealth redistribution, rapid industrialization and urbanization, all of which require a democratization of the government (Huntington, 1968; Przeworski and Limongi, 1997). According to modernization theory, we should assume that protest is only positively linked with the real economic growth, regardless of any falsification on the data.

Another clue can be found from the context of public support in electoral politics. In economic voting theory, which mainly concerns countries with democratic elections, people vote according to macroeconomic data if they care about the overall national/regional economy (sociotropic voting), but not if they care only about personal economic well-being (pocketbook voting) (Kramer, 1983; Nannestad and

Paldam, 1997; Gomez and Wilson, 2001). In non-democratic countries, where contentious political behavior such as protest, rather than fair election, is the primary channel to release grievances and pressure governments, macro-economic statistics may influence citizen behavior through a similar mechanism in democracies. If people focus only on their "pocketbook," which reflects the real economy rather than falsified statistics, their protests will be driven only by the real economy. By contrast, "sociotropic protests" are more likely to be influenced by manipulated economic data. Empirical studies of countries with more or less democratic elections show different results yet agree that sophisticated voters, or highly educated people, are able to link the government's performance to their personal financial status; the same studies also find that other people are mostly sociotropic (Gomez and Wilson, 2006; Kramer, 1983).[3]

2.3 Theory: Ineffective Data Manipulation on Protests

From observations and previous works on data falsification, we should expect political leaders to have strong incentives to manipulate economic data during economic downturns. In countries where the process of producing and reporting official statistics is not supervised or checked, the only cost of data manipulation is public mistrust once it is disclosed. Moreover, the cost can be avoided during the incumbent's tenure. On the other hand, the government can benefit from the manipulation through higher public support, and through a more stable society that contributes to governance and development. For the central government, the benefits are crucial to regime survival, while their effect on lower-level governments lies mainly in officer promotion.

[3] Gomez and Wilson's study on four cases Canada, Hungary, Mexico and Taiwan confirms their theory that less sophisticated voters are more likely to make decisions on whether to support or oppose the incumbent government according to official economic statistics (Gomez and Wilson, 2006). Kramer shows that in the United States voters are mostly sociotropic (Kramer, 1983), while another study suggests the opposite in Denmark (Nannestad and Paldam, 1997). A study on Turkey shows that people's inclination for sociotropic or pocketbook voting can vary when evaluating different political parties (Başlevent et al., 2005).

Despite the theoretical profitability of falsifying economic statistics from the perspective of government, the strategy's actual influence on citizens may be weak and not comparable to what they feel from the real economy. Based on the theories on the economic determinants of protest, as well as on economic data manipulation, this article assumes that real economy should strongly influence protest in authoritarian countries, while released economic data does not. It argues, in other words, that economic data falsification does not reduce protests. One reason is that citizens are able to learn about the real economy through their personal financial status. In line with the relative deprivation theory, the amplification of economic data will not reduce the grievances stemming from the real economy; on the contrary, it may even amplify dissatisfaction by extending the gap between people's economic expectations and the reality. Some studies also suggest that a lack of data openness and transparency makes people vulnerable to mis-perceptions, increasing the risk of civil conflicts (Kuran and Sunstein, 1999). Empirical findings indicate that the unreported real economy can have a huge impact on social contentions through material issues that are highly related to public welfare, such as income, price, and quality of public services. As a case in point, Hendrix, Haggard, and Magaloni (2009) find that protests in Asian and African countries have a U-shape relationship with food price.

If protest is a way for citizens in authoritarian countries to signal their disappointment with the government, such citizens are mainly "pocketbook protesters," since protests against the officials who give bad overall state or regional economic performance are far more costly and risky than fair elections to select officials with better performance. Participating in protests is highly costly for citizens, not only because of the investment of time, but also the risk of reprisal. The reprisal can be especially substantial in autocratic countries, where civil liberties are poorly protected, and restraint on the power holder is weak (Eisinger, 1973; Casper and Tyson, 2014; Mason, 1994; Chen, 2012). Therefore, the cost-benefit calculation of protesters requires higher and more direct personal benefits from protests. In this sense, "the welfare of the whole society" cannot provide a sufficiently strong incentive for them to participate in protests. Otherwise they would rather hide their preferences, in order to avoid the punishment by the government (Kuran, 1987, 1989). Citizens may assess their government with the information both of macro-economy received through GDP growth data, and of their own financial status received through income, food price, and so on, but by the same token, when it comes to taking actions in the streets, only the improvement of the latter can directly pay actors for the cost of this expensive signaling.

Finally, even if citizens care much about officially released macro-economic data, they might not believe it under governments with low accountability. Some studies hold that social grievances and protests are affected by reported statistics. They claim, in effect, that these statistics provide an "information shortcut" for poorly informed citizens to learn about the economy and evaluate the government's performance, whether it is authoritarian or democratic (Lupia and McCubbins, 1998; Wallace, 2015). However, evidence from surveys has pointed out citizens' mistrust in, or at least skepticism to, non-democratic state authorities and the government-controlled media (Mishler and Rose, 1997; Espinal et al., 2006; Lai, 2010). In the case of China, rumors about local GDP falsification started to appear online at the beginning of 21st century, and was confirmed by Li Keqiang, then the Liaoning CCP Secretary, who told the American Ambassador that Liaoning's GDP was unbelievable. In 2010, this problem began to appear in one of the Chinese central media's report, in which local statistical officials acknowledged the falsification. Citizens with frequent access to media might not believe the data anymore.

Above all, I propose the following hypothesis:

Hypothesis 1: The count of protests has a weaker correlation with officially reported GDP growth, than that with actual economic growth.

Additionally, we should note that not all theories in conflict studies favor a high economic growth. For instance, as I previously mentioned, modernization theory claims that some modernization process that accompany economic growth, such as urbanization and industrialization, can induce frustrations in some groups, generating demands for political democratization. In other words, economic growth can encourage more protests in non-democratic countries. Studies of Tian'anmen Square Movement in 1989 suggests that China's economic growth, which brought higher living standards, and less government control over the economy, contributed to the urban unrest (Mason, 1994). Yet, as I previously discussed, theories such as grievance theory and relative deprivation theory describe a negative relationship between growth and civil conflicts. Some cases of post-Cold-War authoritarian countries show that people's demand for democratization can be lower when they feel their economic status improving (De Mesquita and Downs, 2005). To test whether it is true in my case, here a second hypothesis:

Hypothesis 2: The count of protests is negatively related to actual economic growth.

3

Local GDP Falsification and Protests in China

This study uses China as an empirical case, basically because of three reasons: First, China does not have democratic elections or civil liberties such as the freedom of speech and publication. In democracies, such freedoms release citizens' feelings about the government's performance, whereas in the Chinese case, the feelings of citizens are most effectively represented by protests. Therefore, my results are not at large risk of being biased by public responses to economic performance through means other than protests. Second, China's GDP falsification has happened at the local level over a long time span, which allows the collection of sufficient data for empirical analysis. Third, although China's statistics are not as standardized as those of most developed countries, its local statistics agencies have enough capacity to follow the same rules to collect and calculate comparable data.

3.1 GDP Growth Report as Performance Signaling

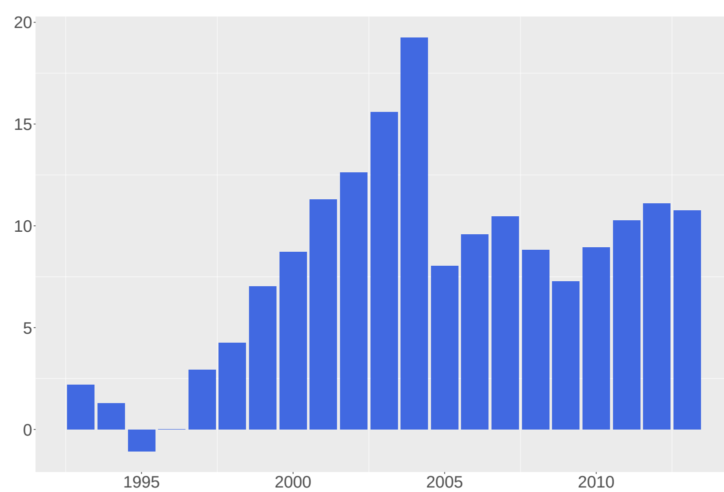
Before exploring GDP falsification in China, we need to discuss how it is released to the public and how citizens process this governmental signal of macro-economic performance. According to China's national statistical regulations, the calculation of annual GDP data contains three steps: *preliminary accounting*, *preliminary verification*, and *final verification*. The accuracy of data is supposedly increasing in that order, because more information is added at each revision. Both national and provincial data calculations are required to follow the three steps. The preliminary accounting data is released on the "Statistical Bulletin on National Economy and Social Development" (SBED) at the Annual Statistical Press Conference, and also on the National Bureau of Statistics of China (NBS) website, about 20 days after the end of each reported year.[1] Usually, by the end of January, the provincial governments follow NBS to release the previous year's GDP data, which is then widely shared by other local media, in a provincial SBED that resembles its national counterpart. In other words, citizens learn about the annual economic growth in their localities every January through a SBED disseminated by the local government as well as the media.

Local governments view this economic data as a strong signal of the government's performance. They are extremely careful with the related information they are about to release to the public. While annual growth of GDP must be reported, if a province has a good performance, such as a high national rank in terms of GDP growth, or an improved rank, its local government and the media will stress the achievement in the news citing SBED. Otherwise, no comparison across provinces will be mentioned. Although people can easily obtain information about relative performance from non-local media, low-growth provincial governments still try to minimize unfavorable impressions from this growth data. However, governments know that simply using euphemism to conceal bad economic performance on news is ineffective. They also know that misrepresenting data can be a less costly way to signal high governmental capability to the public.

[1] "Annual GDP Accounting Instructions of China", in "Bulletin on the Final Verification Data of GDP in 2014", National Bureau of Statistics of China, January 7, 2016. `http://www.stats.gov.cn/english/PressRelease/201601/t20160108_1301700.html`

3.2 Inferior Local GDP Data Quality

Economists have been debating heatedly over China's national GDP data quality (Rawski, 2001; Wang and Meng, 2001; Holz, 2014; Mehrotra and Pääkkönen, 2011), but a consensus has reached regarding the bias of local GDP in China (Cai, 2000; Holz, 2014; Brandt et al., 2014; Wallace, 2015). The most obvious evidence is that the sum of provincial GDP data always exceeds the national GDP data (Holz, 2014; Brandt et al., 2014). The reason could be that the statistics at different levels are calculated by different organs: the national GDP is calculated by NBS, while the local governments report data calculated by their own statistical organs. But it is hard to believe that the discrepancy between the local sum and the national figure can be positive almost always, and that it can reach such a high amount, given that organs at both levels are required to followed the same rules of data calculation. Figure 3.1 shows that each year since 1999 the provincial sum exceeded national GDP by more than 5 percent. In 2004, the discrepancy reached nearly 20 percent.



Notes: The percentages are calculated by annual provincial sum GDP minus national GDP, divided by national GDP. GDP data for this calculation is annual nominal GDP. Source: *Statistical Yearbook*, NBS.

FIGURE 3.1: Percentage by which the Provincial Sum of GDP Exceeds National GDP (1993-2013)

Further evidence lies in the covert data adjustments of provincial GDP in the Statistical Yearbook, whose updates are not mentioned by governments in SBED or other official reports. It is reasonable to assume that if a province falsifies its GDP data of in a particular year, it is more difficult for it to get a favorable growth data in the next year without continuing falsification of the next year's GDP. And the data discrepancy from the true economy will grow over years. A solution is to lower the previous year's statistics when releasing the next year's. Here I disclose the reason for introducing the three steps of GDP calculation in the preceding subsection. According to "Annual GDP Accounting Instructions of China", while the preliminary accounting data of GDP, through which citizens learn about the previous year's local economic performance, is released in January, the preliminary verification data is published in Statistical Yearbook one year later, and the final verification data of annual GDP is published in Statistical Yearbook after an additional year.[2] In other words, if the province reports the adjusted data for the previous year, the statistics shown in two years' Statistical Yearbooks can differ because of the revision. By comparing the 54 revised GDP data reported in Statistical Yearbooks and their originally reported amount from 1993 to 2013, I find that more than 80 percent of the adjusted statistics go down in the next year. This finding suggests that the provinces are likely to be manipulating their GDP statistics (Table 3.1).

3.3 Rising Protests since the 1990s

During the same time period of when local GDP manipulation was apparently taking place, the count of protests was also rising in China. Although there are various definitions of protests, studies agree on the increasing trend of Chinese protests since the 1990s. In his study on two Chinese provinces, Xi Chen (2012) notes that the

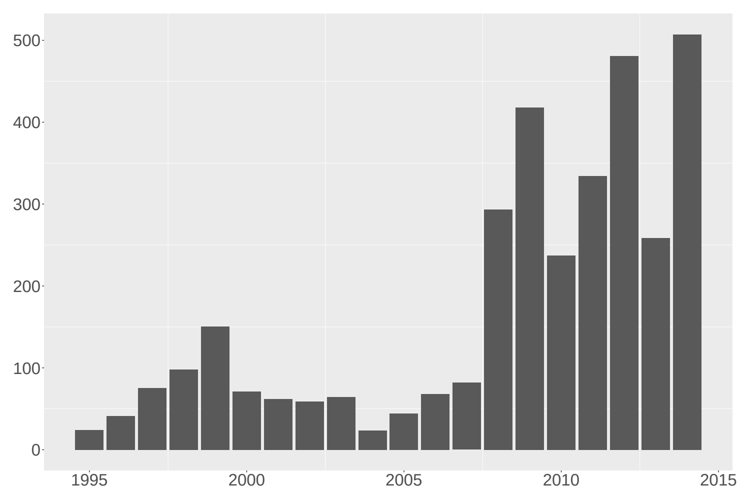
[2] "Annual GDP Accounting Instructions of China", in "Bulletin on the Final Verification Data of GDP in 2014", National Bureau of Statistics of China, January 7, 2016. http://www.stats.gov.cn/english/PressRelease/201601/t20160108_1301700.html

Table 3.1: Upward and Downward Provincial GDP Adjustments

	Downward	Upward	All adjustment
Amount of adjusted cases	44	10	54
Percentage of total cases	11	2.5	13.5
Percentage of adjusted cases	81.5	18.5	100

Notes: For some unknown reason, the Statistical Yearbook stopped update preceding years' data since 2009. And data for the three "economic-census year," 1995, 2006, and 2010 are calculated with different sources and methods from other years. These two kinds of years are seen as incomparable and are not included in the counting for this table. Source: *Statistical Yearbook*, NBS.

amount of social protests during 1991-2000, usually in the form of collective petition, was rising rapidly. A study of "mass incidents" by the Chinese Academy of Social Sciences shows a similarly conspicuous increase in the 21st century (Chinese Academy of Social Sciences, 2014). In my dataset derived from Integrated Crisis Early Warning System (ICEWS) Dataverse, from 1995 to 2014, there is a rise over time in the amount of protests, with a dramatic jumping beginning in 2008 (Figure 3.2).



Source: *ICEWS Dataverse*

FIGURE 3.2: Number of Protests in China (1995-2014)

4

Research Design and Empirical Findings

4.1 Influences of GDP Growth and Real Economic Growth

To explore whether the Chinese economic protesters are more sensitive to officially released economic statistics or the real economy, and whether manipulated GDP data constrain protests, this article estimates two sets of negative-binomial models to measure how the count of protests' correlates with reported GDP growth and with real economic growth. The first set includes GDP growth and a group of control variables to estimate the count of protests. The second set replaces GDP growth with real economic growth.

Dependent Variable and Explanatory Variables

The model's dependent variable is the count of protests. Studies on China's protests are not rare, but a systematic dataset on protests at the local level is unavailable. The available datasets concentrate mostly on the labor movement, including the China Labour Bulletin (CLB) Strike Dataset and the China Strike Dataset by Manfred Elfstrom. The most complete dataset on Chinese protests at provincial level is the Mass Incident Dataset collected by Chinese Academy of Social Sciences. However, the data is not open to the public, and it is collected only from the reports of legal media in P.R.C., which may report protests selectively. The Global Database of Events, Language and Tone (GDELT) Event Database built by GDELT Project[1] uses machine-learning method to mine reports of events from global websites in multiple languages and based in various countries. Although some studies are still using GDELT data, it is now agreed that ICEWS data, which is also a machine-learning-based event dataset, has a more specific and reliable data source.[2] So I instead use ICEWS data for creating my dependent variable, and use the result generated from GDELT as an additional check. Through aggregating the number of protests by the incident location, I calculated the count of protests for each province-year during 1995-2014 from ICEWS dataset, containing about 3000 individual protest events in total. For some protests whose locations are not specified in the dataset, I looked into media sources and recoded them. I also checked and removed some repeated calculation of events.

In model specifications that include GDP growth as an explanatory variable, I collected the GDP growth from each provincial SBED for each year. Interestingly, whether reporting provincial or national GDP data, the annual SBED only contains GDP level by nominal GDP calculation, but GDP growth rate by real GDP calculation (in 1993 price). To avoid the effect of inflation, and because citizens find it easier to interpret GDP growth than GDP levels, I use provincial GDP growth data from each year's SBED. GDP growth data is lagged by one year, because, as discussed in Section 3.1, citizens are informed of any given year's GDP growth data only at the beginning of the following year. Lagged economic growth variables also somewhat help reduce the problem of reverse causality, because theoretically a year's economic

[1] For more information about GDELT, please refer to `http://www.gdeltproject.org/`

[2] For more information about ICEWS, please refer to `https://dataverse.harvard.edu/dataverse/icews`

growth cannot be affected by future protest number.

As a proxy for actual economic growth, I use nighttime luminosity growth. Data for more commonly used proxies, such as steel production, freight volume and electricity consumption, are either very unsystematic or susceptible to manipulation at the local level. Although the usage of luminosity data is relatively new in the social sciences, some works have discussed its usefulness for approximating real economy. Nordhaus and Chen suggest that luminosity data can be used to compensate for the imprecisions of national accounts; it can serve as a good proxy for the real economy, they add, in countries without a mature statistical system, but is of little use for developed countries (Chen and Nordhaus, 2011; Nordhaus and Chen, 2015). However, they did not consider the possibility of data falsification. For his part, Brian Min uses nighttime luminosity as a proxy for the provision of public goods, as it indicates the quality of energy provision. His work on several developing countries demonstrates also that luminosity can be a good predictor of energy consumption rates, outdoor public lightening and household electrification rates (Min, 2010), all of which relate to the real economy.

Provided by the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS)[3], the data is the yearly-average nighttime light intensity captured by satellites in 30 arc second grid cells, with each cell value ranging from 0 to 63. Sunlit, moonlit, glares and observations with clouds are excluded from the data. I use the cleaned up data removing background-noise lights, such as the light from cities, towns, and other sites with persistent lighting, as well as ephemeral events, such as fires. I obtain the provincial luminosity intensity by combining the satellite map with China's administrative map, and then calculating the sum of raster points within each province's polygon. Taking the logarithm of the intensity, I calculate the annual growth rate.

[3] For more information about DMSP-OLS, please refer to: `http://ngdc.noaa.gov/eog/`

There are three reasons for setting the time period as 1995-2013. First, China lacked a reliable GDP calculation and report system until 1993 (Xu, 2009), so GDP data before and after this year are not comparable. Second, ICEWS provides event data since 1995, and nighttime luminosity data are available only until 2013. The third reason is theoretical, which is to try to reduce the immediate influence of the 1989 Tian'anmen Square Movement on protests. This student movement, endorsed by many other citizens, claimed political liberalization and civil liberties, and brought considerable nation-wide attention. Yet it was eventually violently repressed by the Chinese government, causing significant citizen grievances on the one hand, and the government's stronger political coercion on the other hand, in the following several years.

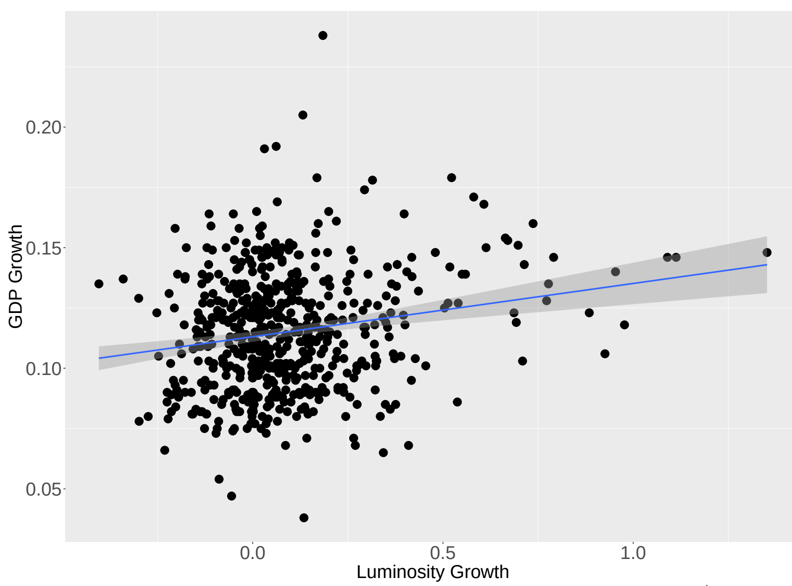


FIGURE 4.1: GDP Growth against Luminosity Growth (Full Sample)

Controlling for Confounding Factors

I first add fixed-effects on provinces for each model to control for provincial level variations that can influence both the number of protests and economic growth, including whether a province is an autonomous region, whether it is a municipal city, and its distance to the capital. Autonomous regions in China accommodate a high proportion of ethnic minority groups. They also have more ethnic conflicts and are relatively less developed economically. Municipal cities are cities governed directly by the central government. They are developed and get more attention from the central power. The central government pay more attention also to the social stability of regions located close to Beijing, the capital, in that they are more likely to affect Beijing's stability.

The model controls for three demographic variables: population, education and unemployment. According to the literature on civil wars and conflicts, highly populated regions are relatively more likely to suffer conflicts than other regions (Fearon and Laitin, 2003; Collier and Hoeffler, 2004), so I include population in the model. Unemployment's possible influence on protests accommodates both theories driven by grievances, as well as ones based on relative deprivation, because unemployment causes both a decline in income, and a divergence of economic welfare between the employed and the unemployed (Gurr, 1970). Hence, the number of unemployed workers is used as a control variable. Education is measured by the percentage of people who hold a higher-education degree. Although there is little doubt that educated labor contributes to economic development, this variable's effect on protest is controvertible. People who lack education are more likely to rebel against the government, because they have lower life prospects and higher opportunity to participate in the rebellion (Humphreys and Weinstein, 2008; Collier and Hoeffler, 2004). As discussed in Section 2.2, however, only highly-educated citizens are able to evaluate the government's performance through the real economy; poorly educated people are relatively more vulnerable to information manipulation, so this variable may well affect protest in the opposite direction to the real economy.

By including the household income-expenditure ratio in both rural and urban areas in the model, I control for a key factor in grievance-based theories of protest. The higher the proportion of a person's income is spent on living expenses, the more likely she is to feel unsatisfied about her economic status. Studies on both democratic and authoritarian countries show that low-income groups tend to use protests as a political tool to obtain policies expected to improve their welfare (Lipsky, 1968; Chen, 2012). Furthermore, since household income and expenditure are correlated also with real economic growth, if Chinese citizens are more likely to protest against a poor real economy because of their own financial status, the ratio between their income and expenditure should also be negatively linked with the count of protests. In China, the income and expenditure of rural and urban areas have divergent patterns, so that they are usually calculated separately with different methods. Therefore, I include the two areas' income-expenditure ratio as two variables.

Two variables that play a key role in the modernization theory, urbanization and industrialization, are also added to my model. Some studies on China suggest that industrialization brings greater employment stratification and worker grievances (Mason, 1994). Scholars have also noticed that the urbanization process causes growing horizontal inequality between cities and villages (Huang, 2008), which may heighten the grievances of rural citizens. Industrialization is measured by the ratio between the number of people who work in the secondary industries and that in primary industries.[4] Because the Hukou, or household registration system in China substantially reduces the chance of rural people who work in cities to become legal urban residents, I use the ratio of urban workers to rural workers, instead of the urban-rural resident ratio, as a measure of urbanization. If the theory that the

[4] General principles of classifying industry sectors are similar across countries, while the classifications may differ. Primary industry is defined as an industry that directly produces materials from natural resources. In China, it roughly equals agriculture, including farming, forestry, fishery and animal husbandry. Secondary industry refers to an industry that produces finished products from the materials. It includes narrowly defined "industry" and construction in China. Source: "Classification of Industrial Sectors," National Bureau of Statistics of China, January 14, 2013. http://www.stats.gov.cn/tjsj/tjbz/201301/t20130114_8675.html

number of protests is positively correlated with real economic growth holds true in this case, the two variables, industrialization and urbanization, should also be related positively to the dependent variable.

In addition, I add protests lagged by one period as a control variable to control for protest experience. The opportunity protest theory suggests that the possibility of protest depends on the actor's perception of how likely a protest can be carried out successfully. Thus, a large number of past protests convey a higher feasibility, and perhaps lower risk, to stage protests in the near future. Furthermore, because potential demonstrators often draw information on risk levels from others, lower risk levels can trigger a cascade. (Kuran, 1989; Bikhchandani et al., 1992).

Data for population, the ratio of income to expenses, the number of urban and rural workers, and the number of secondary- and primary-industry workers are derived from China Yearly Macro-Economics Statistics provided by the University of Michigan China Data Center.[5] I also use higher education amount data released by China's Ministry of Education, and unemployment data collected by National Bureau of Statistics of China. Table 4.1 summarized the data used in this study.

Table 4.1: Summary of Data across 31 Provinces (1995-2013)

Statistic	N	Mean	St. Dev.	Min	Max
protest number (ICEWS)	362	5.014	17.712	0	163
luminosity growth	362	0.116	0.237	−0.404	1.352
GDP growth	362	0.121	0.025	0.038	0.238
higher education rate	362	7.344	5.840	0.430	41.210
rural income/expenditure	362	0.757	0.098	0.484	1.064
urban income/expenditure	362	0.751	0.059	0.605	0.923
population	362	8.090	0.816	5.553	9.335
urbanization	362	0.337	0.172	0.115	0.940
unemployment	362	22.594	14.015	1.000	72.000
industrialization	362	0.973	1.294	0.079	9.570

[5] Please refer to `http://chinadataonline.org/`

Results

The result meets my expectation regarding the relationship between the dependent and independent variables. The growth of luminosity has a significant negative coefficient at the 95% confidence level, while the reported GDP growth is not significantly correlated with the count of protests. The income-expenditure ratios in both rural and urban areas are negatively linked with the dependent variable, although only the urban relationship has a statistically significant effect (G1 in Table 4.2 and L1 in Table 4.3). The coefficients with 90 percent confidence intervals are plotted in Figure 4.2.[6]

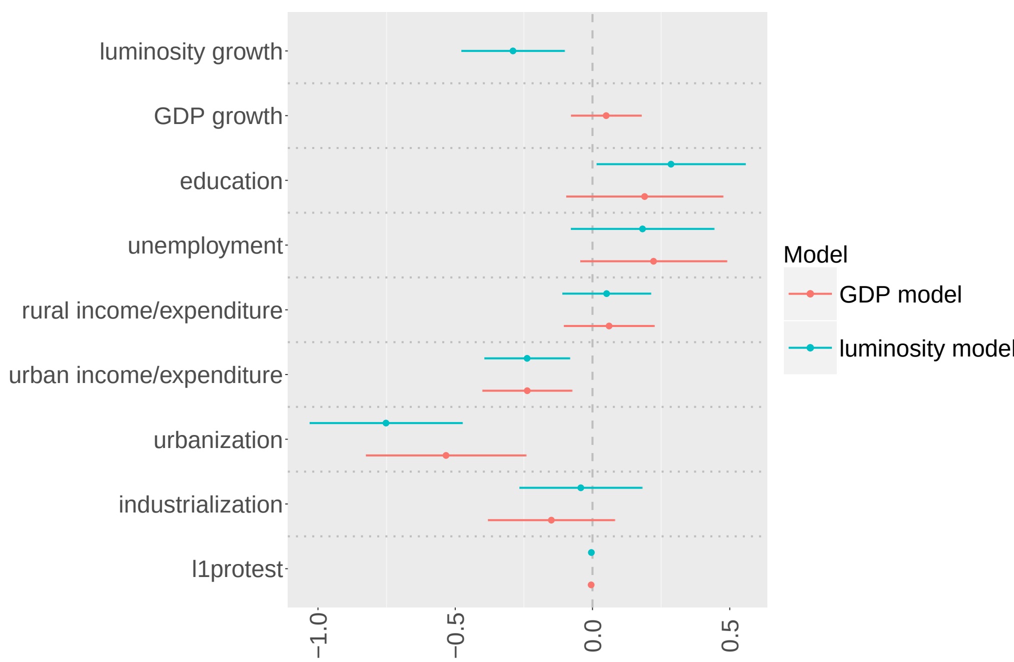


FIGURE 4.2: Coefficients for GDP Growth Model and Luminosity Growth Model

I also calculate the expected count change from the 5th percentile to 95th percentile of each variable, through a simulation that takes 1000 random draws from

[6] The coefficient of population is not shown in the figure because of its too large scale. It is not statistically significant in either model.

Table 4.2: Negative-Binomial Model on Protest Number (GDP Growth Model)

	G1	G2	G3	G4	G5	G6
past protest	0.00	−0.00	0.02	0.01^{*}	0.00	0.04^{***}
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
GDP growth	−4.59	2.58	-6.35^{*}	−3.21	3.99	−6.22
	(3.44)	(4.26)	(3.60)	(3.55)	(4.44)	(3.65)
population	−1.95	−1.71	1.61	−0.33	−0.45	−0.21
	(1.81)	(1.52)	(1.98)	(0.29)	(0.30)	(0.28)
education	0.10^{**}	0.05	0.25^{***}	0.08^{**}	0.03	0.22^{***}
	(0.05)	(0.04)	(0.06)	(0.04)	(0.03)	(0.06)
unemployment	0.01	0.01	0.01	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
rural income/expenditure	1.31	0.85	1.88	0.72	−0.05	1.38
	(1.55)	(1.37)	(1.68)	(1.34)	(1.26)	(1.35)
urban income/expenditure	-11.63^{***}	−1.75	-10.83^{***}	-7.81^{***}	0.25	-7.32^{***}
	(2.45)	(2.20)	(2.73)	(2.10)	(2.17)	(2.26)
urbanization	-2.87^{*}	-2.99^{**}	-5.97^{***}	−0.30	−0.83	-3.32^{**}
	(1.58)	(1.31)	(1.79)	(1.43)	(1.26)	(1.65)
industrialization	0.11	−0.05	−0.32	0.02	0.01	−0.21
	(0.17)	(0.13)	(0.20)	(0.15)	(0.14)	(0.15)
post-2008		2.67^{***}			2.58^{***}	
		(0.82)			(0.85)	
GDP growth:post-2008		−7.73			−8.08	
		(6.38)			(6.58)	
Fixed-effects on province	Yes	Yes	Yes	Yes	Yes	Yes
Zero-inflation	No	No	No	Yes	Yes	Yes
AIC	1328.61	1252.13	1149.08	1367.27	1307.34	1182.11
BIC	1488.17	1419.47	1303.05	1417.86	1365.72	1232.15
Log Likelihood	-623.31	-583.07	-534.54	-670.63	-638.67	-578.05
Deviance	346.14	359.39	324.91			
Num. obs.	362	362	347	362	362	347
Num. groups: Province	31	31	30	31	31	30

Notes: Past protest is measured by the count of protests lagged for one year. I also tried a two-year lag, and the significance remains unchanged. Province factor coefficients are not shown.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4.3: Negative-Binomial Model on Protest Number (Luminosity Growth Model)

	L1	L2	L3	L4	L5	L6
past protest	0.00	-0.00	0.02^{*}	0.01^{*}	0.00	0.04^{***}
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)
luminosity growth	-0.65^{**}	-1.63^{**}	-0.81^{***}	-0.56^{*}	-1.55^{**}	-0.79^{**}
	(0.31)	(0.65)	(0.31)	(0.30)	(0.68)	(0.31)
population	-1.88	-2.47^{*}	1.75	-0.32	-0.51^{*}	-0.17
	(1.78)	(1.43)	(1.98)	(0.29)	(0.31)	(0.28)
education	0.10^{**}	0.06^{*}	0.26^{***}	0.09^{**}	0.03	0.23^{***}
	(0.05)	(0.03)	(0.06)	(0.04)	(0.03)	(0.06)
unemployment	0.01	0.01	0.00	0.02	0.01	0.02
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
rural income/expenditure	1.05	0.80	1.60	0.63	0.02	1.40
	(1.56)	(1.31)	(1.67)	(1.35)	(1.24)	(1.37)
urban income/expenditure	-11.43^{***}	-1.54	-10.28^{***}	-7.91^{***}	0.05	-7.05^{***}
	(2.44)	(2.05)	(2.71)	(2.09)	(2.06)	(2.24)
urbanization	-2.88^{*}	-3.99^{***}	-5.99^{***}	-0.45	-1.79	-3.40^{**}
	(1.52)	(1.21)	(1.75)	(1.43)	(1.23)	(1.66)
industrialization	0.15	0.04	-0.30	0.04	0.04	-0.19
	(0.17)	(0.12)	(0.20)	(0.15)	(0.12)	(0.15)
post-2008		1.90^{***}			1.76^{***}	
		(0.18)			(0.19)	
luminosity growth:post-2008		0.41			0.41	
		(0.70)			(0.74)	
Fixed-effects on province	Yes	Yes	Yes	Yes	Yes	Yes
Zero-inflation	No	No	No	Yes	Yes	Yes
AIC	1325.24	1225.86	1142.03	1364.64	1287.97	1178.74
BIC	1484.79	1393.20	1296.00	1415.24	1346.34	1228.78
Log Likelihood	-621.62	-569.93	-531.02	-669.32	-628.98	-576.37
Deviance	345.47	363.17	326.39			
Num. obs.	362	362	347	362	362	347
Num. groups: Province	31	31	30	31	31	30

Notes: Past protest is measured by the count of protests lagged for one year. I also tried a two-year lag, and the significance remains unchanged. Province factor coefficients are not shown.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

the model, holds all other variables at their central values, and averages across the means. In the GDP growth model, the reported GDP growth does not have a significant effect on the expected protest count, while one percentage change in education and urban income-expenditure rate lead, respectively, to an increase of 3.5 and a decrease 3.2 in the expected count of protests (Table A.1 in Appendix A). In the luminosity growth model, one percentage increase in luminosity growth leads to a decline of 0.5 in the expected count of protests; the influence of a one-percentage rise in education and the urban income-expenditure rate are about an increase of 3.6 and a decrease of 3.1 (Table A.2 in Appendix A).[7]

The result supports Hypothesis 1. Indeed, the count of protests is significantly correlated with luminosity growth, and it is not correlated with reported GDP growth. Evidently, the manipulation of GDP statistics provides little help in reducing Chinese protests. Hypothesis 2 is also sustained. Luminosity growth has a consistently negative correlation with protest number. The negative effect of urban income-expenditure supports the assumed mechanism that lower actual growth triggers more protests because it is linked with individual citizens' financial status. The positive influence of education suggests that it raises people's likelihood of protesting. The reason is probably that it enables people to link their economic status with the government's performance, as economic-voting theory suggests; in other words, it makes them less likely to be misled by the manipulated data. Furthermore, neither of the two variables proxying for modernization theory, *urbanization* and *industrialization*, is significant in the two models.

[7] In the two tables, the statistically significant changes in expected count are labeled with asterisks.

4.2 Robustness Checks

High Falsification Period

A potential challenge to the theory is that the levels of GDP falsification are probably not randomly assigned by the governments across years. High falsification is especially likely during years when the real economic growth declines, or in years when there are more social conflicts. According to Section 3.3, it is possible that only during several later years of the analyzed time period does GDP manipulation reach levels influential to protest counts. Yet my two models covering a period of twenty years cannot identify that issue. Hence, the results of Models G1 and L1 show only that the data falsification in low-falsification years is insufficient to affect people's incentives to protest. Therefore, we need to determine whether the regression result still holds during high-falsification years.

I first presume that in the years from around 2008, when the number of protests began to spike (Figure 3.2) and real economic growth started to trend downward, local GDP falsification might be significantly higher than in previous years. In order to test this conjecture, I regress each province's reported GDP growth against the luminosity growth; then I conduct structural breaks tests. The test result shows that in most of the provinces the relationship between luminosity growth and GDP growth began a significant discrepancy at around 2008.

In order to test whether this discrepancy influences the results, I add an interaction term between high falsification years and the explanatory variable to each model. High falsification years is coded as a dummy variable, with years 2008 to 2013 coded 1 and the preceding years coded 0. The estimation of the new models (Models G2 in Table 4.2 and L2 in Table 4.3) gives an insignificant coefficients for the interaction terms. Figure 4.3 and Figure 4.4 show that, during the supposed high-falsification years, the change in expected count of protests seems to be dragged downward by a GDP growth change from 5th percentile to 95th percentile, although the higher bound of confidence interval is higher than the no-interaction model. This implies that high-level manipulation of GDP growth does tend to have a negative effect on protest compared to the low-levelled, but it effect is still not statistically significant. From 2008 to 2013, one percentage of luminosity growth leads to about 5 more decrease in expected protest number than the years before 2008, holding other variables at their central values.

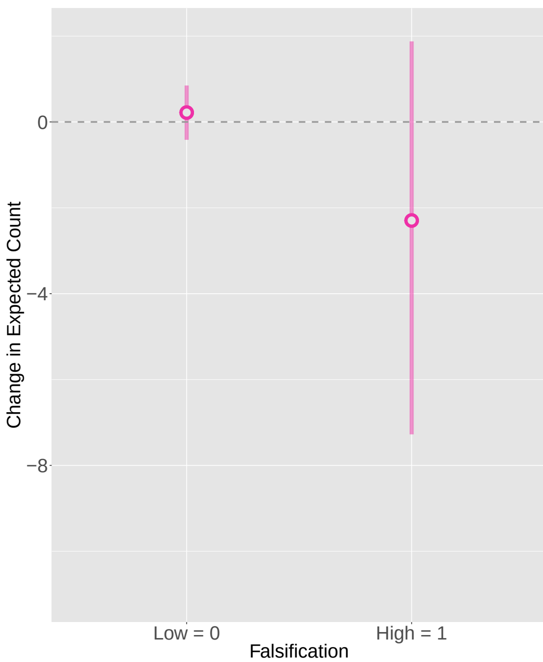


FIGURE 4.3: Change in protest with post-2008 interaction (GDP growth)

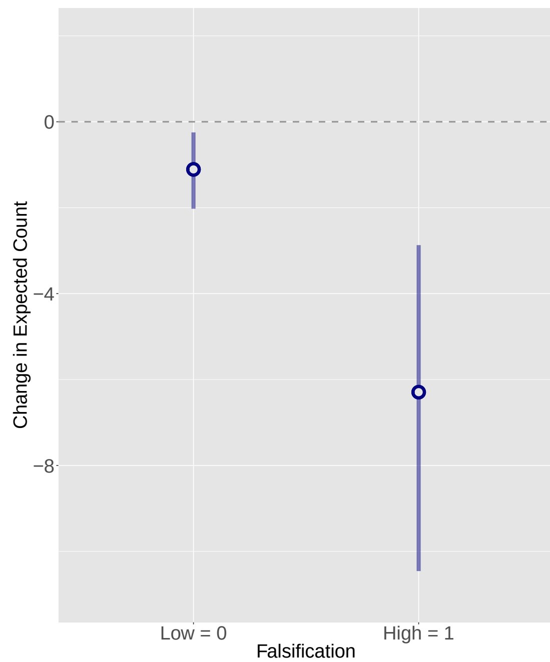


FIGURE 4.4: Change in protest with post-2008 interaction (luminosity growth)

The two figures are drawn from Model G2 and L2 with simulation by taking 1000 random draws from the models of the standardized data, calculating the GDP/luminosity growth at 5th and 95th percentiles, and holding all other variables at their central value. “Low” refers to pre-2008 years whereas “High” refers to post-2008 years.

Removing the Capital and Zero-Inflation

Examining the data, we see that a significantly higher number of protests occurred in Beijing than in other individual provinces. This is possibly because some protesters travel to the capital for demonstrations in order to get more influence, or because Beijing gains more media attention than other regions, so that its protests are more likely to be reported. In order to eliminate the possible bias caused by the observations of Beijing, I introduce Models G3 and L3 to re-estimate G1 and L1, respectively, with the removal of Beijing observations. The results are still robust.

Because a high proportion of the province-year observations of protest count is zero, I adopt zero-inflation in the three pairs of models (Models G1-G3 into G4-G6, Models L1-L3 into L4-L6), and find that the results for the main explanatory variables are still robust, and most of the signs and significance of the control variables do not change.

Endogeneity

There can be a reverse causality between GDP growth data and the number of protests, because it is possible that the local governments tend to heighten the GDP statistics more greatly after seeing a high number of protests. Theoretically, lagged GDP growth should not be resulted by protests, for the government cannot falsify the data according to the number of protests in the future. Statistically, however, future GDP growth is correlated with previous year's GDP growth, so the endogeneity problem cannot be solved simply by using lagged GDP growth.

Unfortunately, I did not find a good instrument for the model, and even instrumental methods do not completely solve the problem. Therefore, I try to check its influence on my result from other ways. My original model cannot avoid endogenous problem because of temporally correlated GDP growth, so controlling for the time trend might to some extent alleviate the problem. I add a time trend variable that interacts with province indicator to control for the province-specific time trend. As Table 4.4 reports, this makes the coefficient of GDP growth significant within 90 percent confidence interval, holding other variables constant. Conducting similar strategy does not change much the luminosity growth model's result. In both models, education and urban income-expense rate no longer stay significant.

Under time-trend control GDP growth only gets a weak negative correlation at 0.1 level with protests – weaker than the effect of luminosity growth, but we are still uncertain about whether its average effect is smaller in scale. Therefore, I standardize all continuous variables and reestimate the two models. Figure 4.5 shows that the difference between the two variables' mean coefficients is quite small. This means that we should not totally ignore the influence from reported GDP growth, although it is far less clear than that of actual growth.

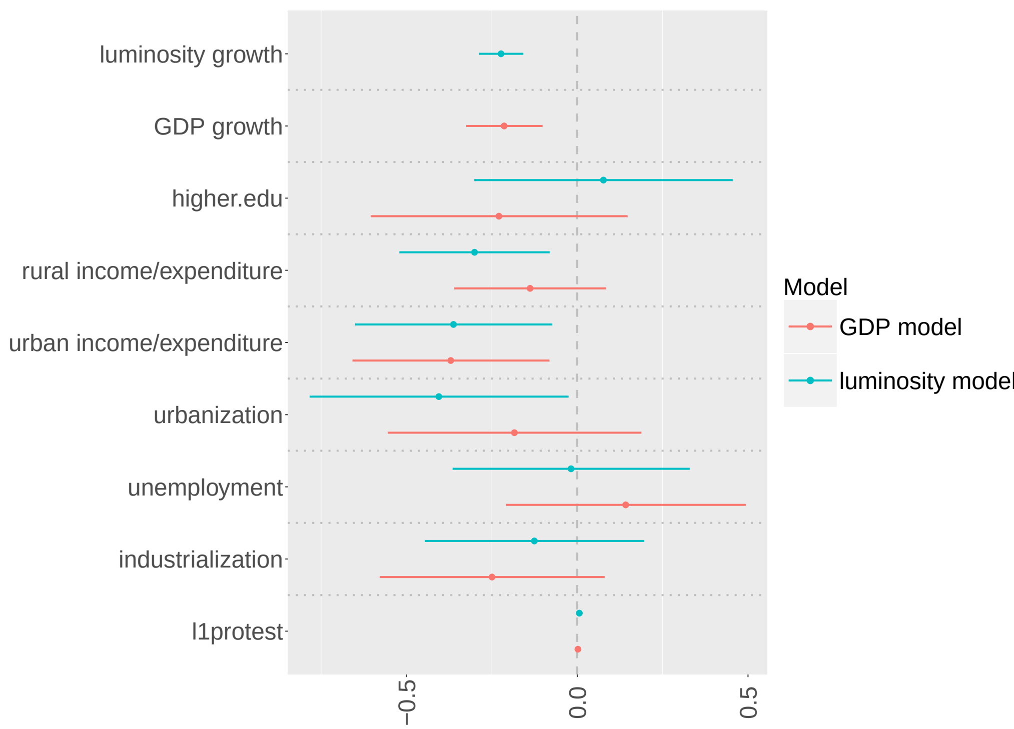


FIGURE 4.5: Coefficients for Two Models under Time-Trend Control

Table 4.4 also shows that, removing control variables and thus restricting the model to only lagged protest number and fixed-effects on provinces, neither the result for GDP growth or luminosity growth changes greatly.

Table 4.4: Additional Checks

	GDP growth model			luminosity growth model		
	FE Only	FE+CV	Time	FE Only	FE+CV	Time
past protest	0.03^{***}	0.00	0.00	0.03^{***}	0.00	0.01
	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
GDP growth	-0.46	-4.59	-7.22^{*}			
	(2.73)	(3.44)	(3.90)			
luminosity growth				-0.66^{**}	-0.65^{**}	-1.12^{***}
				(0.30)	(0.31)	(0.26)
population		-1.95	3.31		-1.88	3.48
		(1.81)	(3.71)		(1.78)	(3.63)
education		0.10^{**}	-0.05		0.10^{**}	0.01
		(0.05)	(0.06)		(0.05)	(0.05)
rural income/expenditure		1.31	-0.53		1.05	-2.19
		(1.55)	(1.80)		(1.56)	(1.73)
urban income/expenditure		-11.63^{***}	-3.62		-11.43^{***}	-3.55
		(2.45)	(3.10)		(2.44)	(2.97)
urbanization		-2.87^{*}	-1.80		-2.88^{*}	-3.47^{*}
		(1.58)	(1.87)		(1.52)	(1.84)
unemployment		0.01	0.01		0.01	-0.00
		(0.02)	(0.02)		(0.02)	(0.02)
industrialization		0.11	-0.20		0.15	-0.08
		(0.17)	(0.23)		(0.17)	(0.21)
time trend			0.21			0.24^{*}
			(0.12)			(0.13)
Fixed-Effects on Province	Yes	Yes	Yes	Yes	Yes	Yes
Time Trend Control	No	No	Yes	No	No	Yes
AIC	2332.54	1328.61	1270.00	2327.48	1325.24	1255.83
BIC	2480.71	1488.17	1550.20	2475.64	1484.79	1536.03
Log Likelihood	-1132.27	-623.31	-563.00	-1129.74	-621.62	-555.92
Deviance	571.38	346.14	326.68	571.21	345.47	332.71
Num. obs.	577	362	362	577	362	362

Notes: Past protest is measured by the count of protests lagged for one year. Coefficients for province factors and their interactions with time trend are not shown.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Another possible source of endogenous confounding is citizens' access to media. As is discussed in Section 2.3, frequent and broad media access provides people with more information about their government, making them more skeptical about officially released statistics and news. This implies that citizens with this kind of access are more likely to protest during economic downturns. On the other hand, as we can learn from development economics literature, the access to media, especially media uncontrolled by government, is positively correlated with economic growth. Unfortunately, I did not find a reliable data source to measure citizens' media access in each province-year. But even if I control for this confounding factor, which has positive correlations both with protests and economic growth, that only strengthens my main finding that actual economic growth is negatively linked with protests.

Additional Data Source

As I mentioned in Section 4.1, another widely used event dataset for civil conflict is GDELT Dataset, though it is controversial because of some unspecified sources. Using the same method to create dependent variable protest number, I obtain a new dataset including observations from 1994 to 2013. Using this new dataset as an additional check, I find that the main results do not change much: the substantive significance for GDP growth shrinks, yet there is still no statistical significance. As for luminosity growth, both its statistical significance and substantive significance remain at a similar level (The summary of new dataset and results are displayed in Appendix B).

5

Conclusion and Further Discussion

Theoretically, manipulating economic data is an ideal scheme for a government to signal its strength and to ameliorate citizens' grievances, thereby stabilizing society without immediate high costs. However, even though the government has an incentive to falsify economic data, it may not succeed in diluting the impact of the real economy that people feel in their life. Exploring this issue may help us understand how people interpret economy, and the extent to which the government can control society through information manipulation.

The empirical result of this study gives a negative answer to this question. Using the China case, it shows that citizens tend to protest against a weak real economy. It also shows that reported economic statistics are poor predictors of protests. Falsified statistics cannot help alleviate people's grievances and reduce protests as actual growth does. This is probably for three reasons. First, the potential protesters are able to feel harmed by their poor personal economic status, which is associated with the real economy rather than manipulated data. Second, as the country's educational level improves, more and more citizens can link their personal status with the performance of the government. Finally, as citizens in authoritarian countries use protests as a substitute for votes, they are probably ego-tropic in deciding to take risky and costly action. Although each of these mechanisms needs further exploration, this article's empirical results may shed light on the studies regarding the social influence of data falsification. The results also contribute to the political economy of protests, especially of protests under authoritarianism.

Besides the further theoretical exploration, looking into representative cases of protests might also be necessary for future studies. Although this article rejects modernization theory, some event case studies have pointed to specific reasons for protests related to high economic growth. Some Chinese residents' protests against urban construction offer us an example. Thus, we should expect more direct evidence to emerge of economic data falsification in authoritarian countries. Futhermore, efforts are needed to reexamine the findings of this study in other contexts, using data from other countries. In addition, with the advancement of informational technology, expression and protest online becomes increasingly popular, which means we should reconsider the spatial distribution of protest, as discontentment with certain characteristics of one region can be mixed up with other regions on the Internet.

If the conclusions of this article are consistently supported by further research, the governments should rethink their citizens' sensitivity to the real economy. They should abandon the notion that information manipulating strategies such as propaganda, censorship, and specifically secret and cheaply administered data falsification, can dampen social antagonisms provoked by bad economic performances. Public satisfaction will not improve in the absence of projects that truly enhance public welfare.

Appendix A

Substantive Results

Table A.1: Expected Count Change of Protests by Simulation (GDP Growth Model)

Variables	mean change	5th percentile	95th percentile
past protest	0.030	-0.153	0.244
GDP growth	-0.380	-0.911	0.109
population	-14, 499.940	-6, 191.059	5.065
education*	3.472	0.084	8.894
unemployment	0.784	-0.820	2.658
rural income/expenditure	0.401	-0.609	1.515
urban income/expenditure*	-3.208	-4.967	-1.806
urbanization	-1.087	-2.640	0.478
industrialization	0.418	-0.748	2.299

Note: The statistically significant changes in expected count are labeled with asterisks.

Table A.2: Expected Count Change of Protests by Simulation (Luminosity Growth Model)

Variable	mean change	5th percentile	95th percentile
past protest	0.045	-0.136	0.260
luminosity growth*	-0.481	-0.872	-0.068
population	-3,030.096	-3,715.144	6.107
education*	3.580	0.292	9.310
unemployment	0.581	-0.932	2.368
rural income/expenditure	0.311	-0.671	1.364
urban income/expenditure*	-3.085	-4.734	-1.690
urbanization	-1.132	-2.558	0.339
industrialization	0.565	-0.646	2.413

Note: The statistically significant changes in expected count are labeled with asterisks.

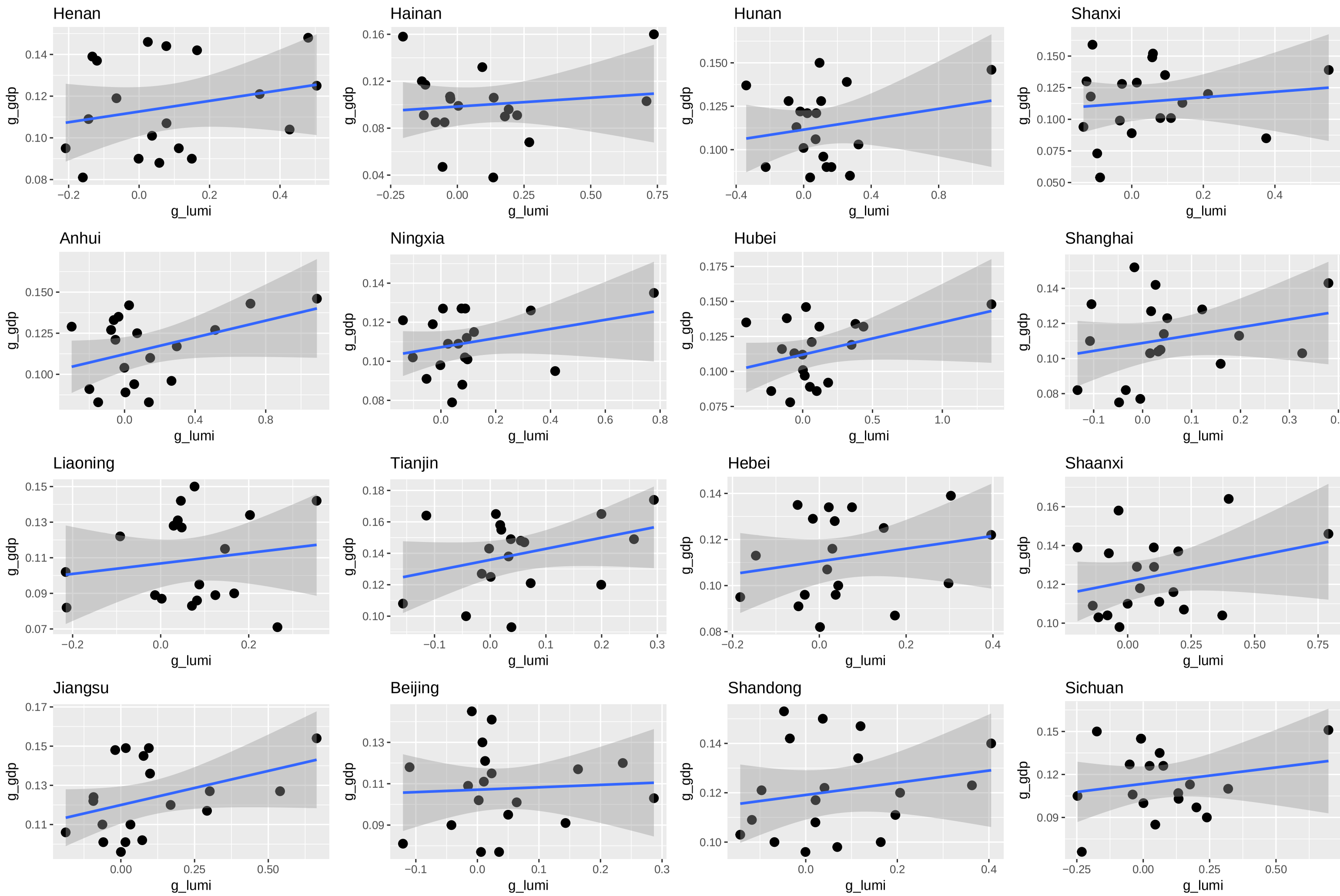


FIGURE A.1: GDP Growth against Luminosity Growth for Each Province (1995-2013)

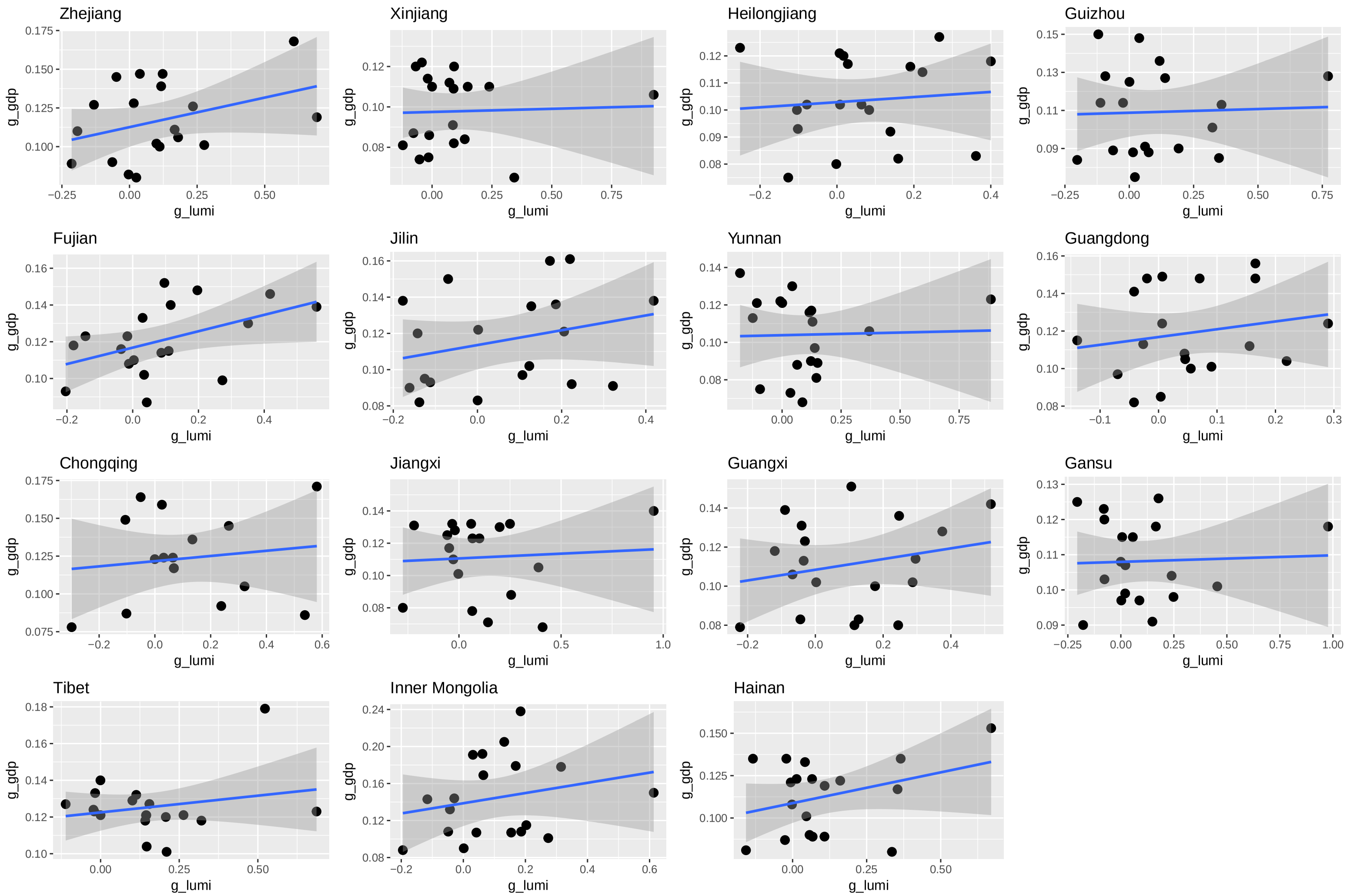


FIGURE A.2: GDP Growth against Luminosity Growth for Each Province (1995-2013, Continued)

Appendix B

Results Using GDELT Data

Table B.1: Summary of Statistics (1994-2013)

Statistic	N	Mean	St. Dev.	Min	Max
protest number (GDELT)	391	4.653	17.144	0	163
GDP growth	391	0.143	0.099	0.054	0.961
luminosity growth	391	0.062	0.186	−0.404	1.352
population (logarithm)	391	8.091	0.815	5.553	9.335
education	391	6.990	5.810	0.430	41.210
unemployment	391	22.185	13.821	1.000	72.000
rural income/expenditure	391	0.763	0.100	0.484	1.064
urban income/expenditure	391	0.756	0.059	0.605	0.923
industrialization	391	0.965	1.292	0.079	9.570
urbanization	391	0.337	0.172	0.115	0.940

Table B.2: Negative-Binomial Model on Protest Number (GDP Growth Model)

	G1	G2	G3	G4	G5	G6
past protest	0.00	-0.01	0.01	0.01	-0.00	0.04^{**}
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)
GDP growth	-1.09	-1.99^{*}	0.35	-0.99	-1.64	-0.35
	(1.11)	(1.08)	(1.16)	(1.05)	(1.02)	(1.11)
population	-1.15	-1.14	1.61	-0.44	-0.58^{*}	-0.27
	(1.63)	(1.36)	(1.75)	(0.30)	(0.32)	(0.29)
education	0.10^{**}	0.04	0.24^{***}	0.10^{**}	0.04	0.22^{***}
	(0.05)	(0.04)	(0.06)	(0.04)	(0.03)	(0.05)
unemployment	0.02	0.02	0.01	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
rural income/expenditure	-0.44	-0.09	-0.02	-0.23	-0.74	0.73
	(1.59)	(1.35)	(1.73)	(1.37)	(1.27)	(1.40)
urban income/expenditure	-11.84^{***}	-2.25	-10.88^{***}	-8.57^{***}	-0.58	-7.71^{***}
	(2.53)	(2.22)	(2.82)	(2.17)	(2.19)	(2.32)
urbanization	-2.24	-2.92^{**}	-4.63^{***}	-0.38	-1.11	-2.77^{*}
	(1.55)	(1.29)	(1.75)	(1.45)	(1.26)	(1.61)
industrialization	-0.02	-0.19	-0.36^{*}	-0.03	-0.09	-0.24
	(0.17)	(0.13)	(0.20)	(0.15)	(0.14)	(0.16)
post-2008		2.73^{***}			2.40^{***}	
		(0.61)			(0.66)	
GDP growth:post-2008		-7.08			-5.68	
		(4.43)			(4.89)	
Fixed-effects on province	Yes	Yes	Yes	Yes	Yes	Yes
Zero-inflation	No	No	No	Yes	Yes	Yes
AIC	1339.26	1255.96	1150.88	1377.97	1314.04	1184.75
BIC	1501.97	1426.61	1307.95	1429.56	1373.57	1235.80
Log Likelihood	-628.63	-584.98	-535.44	-675.99	-642.02	-579.38
Deviance	352.21	370.97	332.60			
Num. obs.	391	391	375	391	391	375
Num. groups: Province	31	31	30	31	31	30

Notes: Past protest is measured by the count of protests lagged by one year. I also tried a two-year lag, and the significance remains unchanged. GDP growth is lagged by one year. Coefficients for fixed-effect intercepts are not shown.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table B.3: Negative-Binomial Model on Protest Number (Luminosity Growth Model)

	L1	L2	L3	L4	L5	L6
past protest	0.00	-0.00	0.02	0.01	-0.00	0.04^{**}
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)
luminosity growth	-0.66^{**}	-1.42^{**}	-0.83^{***}	-0.57^{*}	-1.36^{*}	-0.81^{**}
	(0.32)	(0.69)	(0.32)	(0.31)	(0.70)	(0.33)
population	-1.18	-1.33	1.06	-0.47	-0.66^{**}	-0.28
	(1.57)	(1.29)	(1.74)	(0.31)	(0.32)	(0.29)
education	0.12^{**}	0.06	0.26^{***}	0.11^{***}	0.05^{*}	0.24^{***}
	(0.05)	(0.04)	(0.06)	(0.04)	(0.03)	(0.05)
unemployment	0.02	0.02	0.01	0.02	0.02^{*}	0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
rural income/expenditure	-0.96	-0.77	-0.05	-0.50	-1.10	0.61
	(1.53)	(1.30)	(1.65)	(1.37)	(1.24)	(1.39)
urban income/expenditure	-12.23^{***}	-2.41	-10.92^{***}	-8.90^{***}	-0.83	-7.78^{***}
	(2.55)	(2.18)	(2.81)	(2.18)	(2.13)	(2.31)
urbanization	-2.79^{*}	-3.78^{***}	-5.37^{***}	-0.88	-2.07^{*}	-3.40^{**}
	(1.56)	(1.27)	(1.77)	(1.46)	(1.26)	(1.60)
industrialization	0.00	-0.11	-0.32^{*}	-0.03	-0.05	-0.25
	(0.17)	(0.12)	(0.19)	(0.15)	(0.13)	(0.15)
post-2008		1.97^{***}			1.84^{***}	
		(0.19)			(0.19)	
luminosity growth:post-2008		0.15			0.16	
		(0.75)			(0.77)	
Fixed-effects on province	Yes	Yes	Yes	Yes	Yes	Yes
Zero-inflation	No	No	No	Yes	Yes	Yes
AIC	1335.54	1238.68	1143.84	1375.59	1299.35	1178.85
BIC	1498.25	1409.33	1300.91	1427.18	1358.88	1229.90
Log Likelihood	-626.77	-576.34	-531.92	-674.79	-634.68	-576.43
Deviance	350.97	366.94	333.26			
Num. obs.	391	391	375	391	391	375
Num. groups: Province	31	31	30	31	31	30

Notes: Past protest is measured by the count of protests lagged by one year. I also tried a two-year lag, and the significance remains unchanged. Coefficients for fixed-effect intercepts are not shown.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Bibliography

Acemoglu, D. and J. A. Robinson (2005). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

Alesina, A., J. Londregan, and H. Rosenthal (1993). "A Model of the Political Economy of the United States". *American Political Science Review*, 12–33.

Alesina, A., S. Özler, N. Roubini, and P. Swagel (1996). "Political Instability and Economic Growth". *Journal of Economic Growth 1*(2), 189–211.

Başlevent, C., H. Kirmanoğlu, and B. Şenatalar (2005). "Empirical Investigation of Party Preferences and Economic Voting in Turkey". *European Journal of Political Research 44*(4), 547–562.

Bernstein, T. P. and X. Lü (2003). *Taxation without Representation in Contemporary Rural China*. Cambridge University Press.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades". *Journal of Political Economy*, 992–1026.

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2014). "Challenges of Working with the Chinese NBS Firm-Level Data". *China Economic Review 30*, 339–352.

Cai, Y. (2000). "Between State and Peasant: Local Cadres and Statistical Reporting in Rural China". *The China Quarterly 163*, 783–805.

Casper, B. A. and S. A. Tyson (2014). "Popular Protest and Elite Coordination in a Coup D'état". *The Journal of Politics 76*(02), 548–564.

Chen, X. (2012). *Social Protest and Contentious Authoritarianism in China*. Cambridge University Press.

Chen, X. and W. D. Nordhaus (2011). "Using Luminosity Data as a Proxy for Economic Statistics". *Proceedings of the National Academy of Sciences 108*(21), 8589–8594.

Chinese Academy of Social Sciences (2014). *Annual Report on China's Rule of Law (2014)*, Chapter "The Features, Causes, and Solutions of Mass Incidents" (in Chinese). Beijing, China: Social Sciences Academic Press.

Collier, P. and A. Hoeffler (2004). "Greed and Grievance in Civil War". *Oxford Economic Papers 56*(4), 563–595.

De Mesquita, B. B. and G. W. Downs (2005). "Development and Democracy". *Foreign Affairs*, 77–86.

Eisinger, P. K. (1973). "The Conditions of Protest Behavior in American Cities". *American Political Science Review 67*(01), 11–28.

Espinal, R., J. Hartlyn, and J. M. Kelly (2006). "Performance Still Matters: Explaining Trust in Government in the Dominican Republic". *Comparative Political Studies 39*(2), 200–223.

Fearon, J. D. and D. D. Laitin (2003). "Ethnicity, Insurgency, and Civil War". *American Political Science Review 97*(01), 75–90.

Gandhi, J. (2008). *Political Institutions under Dictatorship.* Cambridge University Press Cambridge.

Geddes, B. (1999). "What Do We Know about Democratization after Twenty Years?". *Annual Review of Political Science 2*(1), 115–144.

Gomez, B. T. and J. M. Wilson (2001). "Political Sophistication and Economic Voting in the American Electorate: A Theory of Heterogeneous Attribution". *American Journal of Political Science*, 899–914.

Gomez, B. T. and J. M. Wilson (2006). "Cognitive Heterogeneity and Economic Voting: A Comparative Analysis of Four Democratic Electorates". *American Journal of Political Science 50*(1), 127–145.

Gurr, T. (1968). "A Causal Model of Civil Strife: A Comparative Analysis Using New Indices". *American Political Science Review 62*(04), 1104–1124.

Gurr, T. R. (1970). *Why Men Rebel?* Routledge.

Healy, A. and G. S. Lenz (2014). "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy". *American Journal of Political Science 58*(1), 31–47.

Hendrix, C., S. Haggard, and B. Magaloni (2009). "Grievance and Opportunity: Food Prices, Political Regime, and Protest". In *presentation at the International Studies Association Convention, New York (August, 2009).*

Hollyer, J. R., B. P. Rosendorff, and J. R. Vreeland (2011). "Democracy and Transparency". *Journal of Politics 73*(4), 1191–1205.

Hollyer, J. R., B. P. Rosendorff, and J. R. Vreeland (2015). "Transparency, Protest, and Autocratic Instability". *American Political Science Review 109*(04), 764–784.

Holz, C. A. (2014). "The Quality of China's GDP Statistics". *China Economic Review 30*, 309–338.

Huang, Y. (2008). *Capitalism with Chinese Characteristics: Entrepreneurship and the State*, Volume 1. Cambridge Univ Press.

Humphreys, M. and J. M. Weinstein (2008). "Who Fights? The Determinants of Participation in Civil War". *American Journal of Political Science 52*(2), 436–455.

Huntington, S. P. (1968). *Political Order in Changing Societies*. Yale University Press.

Kramer, G. H. (1983). "The Ecological Fallacy Revisited: Aggregate- versus Individual-Level Findings on Economics and Elections, and Sociotropic Voting". *American Political Science Review 77*(01), 92–111.

Kuran, T. (1987). "Preference Falsification, Policy Continuity and Collective Conservatism". *The Economic Journal 97*(387), 642–665.

Kuran, T. (1989). "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution". *Public Choice 61*(1), 41–74.

Kuran, T. and C. R. Sunstein (1999). "Availability Cascades and Risk Regulation". *Stanford Law Review*, 683–768.

Lai, H. (2010). "Uneven Opening of China's Society, Economy, and Politics: Pro-Growth Authoritarian Governance and Protests in China". *Journal of Contemporary China 19*(67), 819–835.

Li, H. and L.-A. Zhou (2005). "Political Turnover and Economic Performance: the Incentive Role of Personnel Control in China". *Journal of Public Economics 89*(9), 1743–1762.

Lipsky, M. (1968). "Protest as a Political Resource". *American Political Science Review 62*(04), 1144–1158.

Londregan, J. B. and K. T. Poole (1990). "Poverty, the Coup Trap, and the Seizure of Executive Power". *World Politics 42*(02), 151–183.

Lorentzen, P. L. (2013). "Regularizing Rioting: Permitting Public Protest in An Authoritarian Regime". *Quarterly Journal of Political Science 8*(2), 127–158.

Lupia, A. and M. D. McCubbins (1998). *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge University Press.

Mason, T. D. (1994). "Modernization and Its Discontents Revisited: The Political Economy of Urban Unrest in the People's Republic of China". *The Journal of Politics 56*(02), 400–424.

Mehrotra, A. and J. Pääkkönen (2011). "Comparing China's GDP Statistics with Coincident Indicators". *Journal of Comparative Economics 39*(3), 406–411.

Min, B. K.-H. (2010). *Democracy and Light: Public Service Provision in the Developing World*. University of California, Los Angeles.

Mishler, W. and R. Rose (1997). "Trust, Distrust and Skepticism: Popular Evaluations of Civil and Political Institutions in Post-Communist Societies". *The Journal of Politics 59*(2), 418–451.

Nannestad, P. and M. Paldam (1997). "From the Pocketbook of the Welfare Man: a Pooled Cross-Section Study of Economic Voting in Denmark, 1986–92". *British Journal of Political Science 27*(01), 111–155.

Nordhaus, W. and X. Chen (2015). "A Sharper image? Estimates of the Precision of Nighttime Lights as a Proxy for Economic Statistics". *Journal of Economic Geography 15*(1), 217–246.

Przeworski, A. and F. Limongi (1997). "Modernization: Theories and Facts". *World Politics 49*(02), 155–183.

Rawski, T. G. (2001). "What is Happening to China's GDP Statistics?". *China Economic Review 12*(4), 347–354.

Shih, V., C. Adolph, and M. Liu (2012). "Getting Ahead in the Communist Party: Explaining the Advancement of Central Committee Members in China". *American Political Science Review 106*(01), 166–187.

Walker, I. and H. J. Smith (2002). *Relative Deprivation: Specification, Development, and Integration*. Cambridge University Press.

Wallace, J. L. (2015). "Juking the Stats? Authoritarian Information Problems in China". *British Journal of Political Science 46*(01), 11–29.

Wang, X. and L. Meng (2001). "A Reevaluation of China's Economic Growth". *China Economic Review 12*(4), 338–346.

Xu, X. (2009). "The Establishment, Reform and Development of China's National Economic Accounting System". *Social Sciences in China (in Chinese) 06*.