Farewell Comrades! Power Struggle and Elite Removal during Authoritarian Leadership Succession

Franziska Barbara Keller
Columbia University

Yuhua Wang
Harvard University

1 Franziska Barbara Keller is Post-doctoral Fellow, Harriman Institute, Columbia University (fk2304@columbia.edu). Yuhua Wang is Assistant Professor, Department of Government, Harvard University (yuhuawang@fas.harvard.edu).

The authors are listed alphabetically, and both contributed equally to the research and writing of the article. We would like to thank Jeremy Wallace, Rory Truex, Minxin Pei, Alex Hughes, Adam Przeworski, Junyan Jiang, Eunkyong Shin, and participants in a workshop on elite politics at UCSD and panels at APSA and MPSA for comments. Yuhua Wang wants to thank Yihua Xia for providing excellent research assistance. All errors remain our own.
Abstract

We present the first systematic effort to theorize and empirically test an elite removal strategy in authoritarian regimes using social network analysis. We argue that authoritarian leaders face two challenges when removing powerful elites: counterattack and party split. Unless the leader’s rule is already consolidated, their best strategy is to start the attack from the periphery of the rival’s network—a pattern that we term “encircling.” We combine social network analysis of the recent anti-corruption campaigns in China with two additional historic case studies to demonstrate that the attacks in elite networks there indeed follow an “encircling” pattern. Individuals at the network’s periphery were attacked earlier, until the centers’ power base was sufficiently weakened. Only at this point were the main targets finally removed. Our study contributes to the study of single-party authoritarian regimes by switching the focus from inclusion to exclusion.

Word count: 9,550
Leon Trotsky, the most prominent and popular Bolshevik leader after Lenin, was expelled from the Central Committee in October 1927. But his expulsion, exile and assassination in Mexico in 1940 were just the end of a complicated power struggle between him and Joseph Stalin for Lenin’s succession. In 1925, Stalin had demoted Trotsky’s brother-in-law, Lev Kamenev, from full member to candidate member of the Politburo and instead promoted three new full members loyal to him: Molotov, Voroshilov, and Kalinin. Shortly afterwards, Stalin followed up his victory by sending Sergei Kirov to replace Grigory Zinoviev as party leader of Leningrad—the stronghold of Trotsky and his allies. But Stalin waited three years after Lenin’s death to launch his final attack on Trotsky, and went on to consolidate his power as supreme leader of the Soviet Union. We argue that this is a typical pattern for purges during authoritarian leadership transition.

Recent scholarly work has focused on how authoritarian rulers share power or distribute the spoils of office in order to remain in power (Shirk 1993; Magaloni 2008; Svolik 2009; Bueno de Mesquita et al. 2003 Shih, Adolph, and Liu 2012). But buying them off is only one way in which a leader can deal with powerful elites; the other option is to remove them from the game. While the stories of such purges have fascinated observers, very little systematic research or theorizing has been conducted on how leaders eliminate their rivals to consolidate power around the times of leadership succession. In this paper, we argue that power sharing and elite removal are two sides of the same coin: authoritarian leaders use both to achieve party cohesion, which is the key to the survival of single-party authoritarian regimes (Huntington 1968; Geddes 1999; Smith 2005; Levitsky and Way 2012).
This article focuses on *how* authoritarian leaders remove political elites. We start with the premise that authoritarian leaders’ objective function is to maximize party cohesion—i.e., to eliminate political rivals without splitting the party. An authoritarian leader thus has two goals. The first is minimizing the risk of a counterattack—a *coup d’état*. The leader would prefer to remove the rival and their support network quickly and decisively, but there are at least three reasons for why this is not always possible. First, an incoming leader may not yet have the power to execute such a decapitation (Johnston 2012) during a transition. Second, rivals are less likely to let down their guard in such unstable times. Third, neither the identity of the rival’s supporters nor their strength and support among other elites may be clear. In this situation, the leader’s best strategy may be to start by attacking known peripheral members of the network (who are less able to strike back) in order to weaken the rival.

In addition to the danger of a counterattack, an authoritarian leader also faces the possibility of more generalized resistance, either by citizens who support the rival or by elites concerned they will be targeted later. Therefore the leader must also collect enough information to be able to convince other party members and society at large that removing certain individuals is legitimate, which requires declaring the main target(s) to be guilty of a criminal or moral offense, be it corruption, treason, or holding the wrong ideological view. Peripheral individuals know they are not important enough for anyone to come to their rescue. They are therefore more likely to defect and surrender information on their connections, who can then be cajoled into revealing information against the leader or other more central actors.
In this paper, we formalize these intuitions as the “encircling” hypothesis. Arrests and investigations of the leader’s political rival will start from the periphery of the rival’s network, with less connected individuals. The more powerful central members of the network are only brought down when enough of their connections have been removed.

This strategy is not as self-evident as it may at first seem. Network specialists (Albert, Jeong, and Barabasi 2000) and experts on counterterrorism (Johnston 2012) have instead argued that an attack on the most central individual would be more effective. For example, several attempts were made to assassinate Adolph Hitler in order to end World War II (Gisevius 1947). For a variety of reasons discussed in the next section, the encircling strategy appears to be more common during leadership succession, when an incoming leader needs to consolidate his or her power by eliminating rivals or competitors.

We test the encircling hypothesis using three case studies of elite removal in authoritarian regimes around times of leadership change. The main focus of our empirical analysis is the Chinese Communist Party’s removal of Zhou Yongkang—a former Politburo Standing Committee member and China’s security tsar from 2002 to 2012—during the recent anti-corruption campaign in 2014. Examining Zhou’s network of almost 120 individuals, we find that arrests and investigations follow the expectations of our encircling hypothesis: individuals farther removed from Zhou Yongkang and with fewer connections—“flies” in the parlance of the Chinese anti-corruption campaign—are more likely to be the initial targets of a wave of investigations. Using a contagion hazard model, we show that being connected to an individual under arrest or investigation increases the chances of removal in the following time period. Unlike in epidemiological contagion
models, however, the more central actors close to Zhou Yongkang are only removed after their “power base” has been weakened – operationalized as an adapted version of a known network measure, closeness centrality.

We complement this quantitative analysis with two qualitative historical case studies to show that a similar strategy of encircling and information gathering along network ties also occurs in other regimes: in the late Soviet Union, with Yuri Andropov’s use of anti-corruption campaigns to weaken his opponents and rise to the party secretary position in 1982, and in China during the early 1940s, with Mao Zedong’s purge of the Comintern Gang during the Yan’an Rectification Movement.

This study contributes to the scholarly understanding of authoritarian politics, which has overwhelmingly focused on how the inclusion of social groups and political elites affects authoritarian politics (Gandhi and Przeworski 2007; Magaloni 2008; Boix and Svolik 2013), by instead investigating the exclusion of political elites and its strategies. The term “network” is used in many historical accounts to describe power struggles in authoritarian regimes, yet our study is the first to use it not just as a metaphor (Ward, Stovel, and Sacks 2011), but to employ the methods and concepts of social network analysis. Applying social network analysis in research on authoritarian regimes is, with a few exceptions such as Razo (2008) and Perez-Oviedo (2015), extremely rare.

**Elite Removal and Party Cohesion in Authoritarian Regimes**

Revolution and popular uprisings have long caught the attention of the general and academic audience (Acemoglu and Robinson 2001; Boix 2003). However, as Svolik (2009, 478) shows, more than two-thirds of all authoritarian leaders who lose power by
non-constitutional means are removed by government insiders. Researchers of authoritarian regimes have thus begun to examine the internal politics among political elites, focusing on the extent to which the ruler shares power with other elites (Magaloni 2008): by exploiting his position, the dictator can acquire more power at the expense of the ruling coalition, but the latter may attempt to deter such opportunism by threatening to stage a coup (Svolik 2009). The strategic behavior that emerges as a result of this conflict can explain the dictator’s tenure and the concentration of power in his regime. However, scholars have thus far paid little attention to the survival of other authoritarian elites, and the strategy of shaping the selectorate through systematic elite removal. Our research addresses exactly this question. We use the term elite removal to characterize a situation in which political elites are involuntarily excluded from political life as a result of a power struggle. We distinguish elite removal from elite exit, which involves primarily voluntary behavior such as resignation or retirement (Manion 1993). We argue that elite removal is the flip side of power sharing, and that authoritarian rulers use both to maximize their chance of survival.

In Bueno de Mesquita et al.’s (2003) model of a power struggle between a dictator and a challenger, for instance, the two compete for the support of the selectorate by offering an optimal mix of private and public goods to form a winning coalition. Removing individual selectorate members before the competition has been decided is not an option in this model—unlike in real life, when the eventual winner may only emerge after shaping the selectorate by including supporters and removing opponents. Even after the exile of Trotsky and his supporters, Stalin’s control of the Soviet Union was not predetermined, and required extensive purges of potential challengers’ followers.
Similarly, when Mao died in 1976, Deng Xiaoping slowly sidelined the chosen successor Hua Guofeng and his supporters, albeit in a form of elite removal that was far less deathly than Stalin’s purges. The exclusion of powerful elites, if successful, helps consolidate the power of the incoming leaders by creating “a focal point for reducing uncertainty, achieving consensus, and forestalling a power vacuum” (Brownlee 2007, 597) and seems to contribute to the length of rulers’ tenures: Stalin ruled over the Soviet Union until he died in 1953, and Deng Xiaoping—who became China’s core leader after Hua was ousted—exerted influence until he died in 1997.

However, elite removal, especially of the main challenger, can backfire. For example, there were clashes in Cambodia in 1997 after Hun Sen, the second prime minister, ousted First Prime Minister Norodom Ranariddh when long tensions between the two governing parties broke into factional fighting during a short power-sharing arrangement. Another case is the mutiny of Uganda’s Vice President Mustafa Adrisi’s followers against President Idi Amin Dada after Adrisi was injured in a suspicious auto accident in 1978. Elite struggles also always carry the danger of spilling over into the broader populace, either because the death or removal of a popular leader triggers demonstrations (as in the case of the Tiananmen protests after the death of Hu Yaobang in 1989) or because one of the contenders explicitly turns to the masses for support (arguably Mao Zedong’s strategy during the Cultural Revolution).

We argue that the expected risk associated with elite removal, and the need to legitimize such a removal, incentivize the leader to follow two principles. First, he or she must minimize the risk that their political rivals will stage a *coup d’état*, which is best achieved by completely eliminating the rival’s network. But if the struggle for leadership
or its consolidation is still ongoing, then removing the whole network may not be feasible. Instead, the best strategy may be to whittle away the rival’s followers by attacking the periphery of the network before taking the fight to the center, where resistance is strongest.

Another reason for choosing this strategy is the rival’s information advantage: the leader is less informed about the structure of the rival network, and may not know every member of it. The leader may also lack convincing evidence that would allow him or her to accuse the rival on legal or ideological grounds. This is relevant if we assume that elites in such struggles are not solely driven by the desire to acquire immediate power, but also care about an ideology or about the party’s long-term survival from which they benefit. If this is true, the leader may need to provide evidence of how the opponent endangers this long-term goal in order to avoid facing a united opposition of elites who are afraid of being the next target. Elite removal, therefore, follows the logic of a war of information. The leader’s goal is to collect as much information as possible about the targeted rival. More central actors presumably have more valuable information due to their direct ties to the main target(s) of the campaign. But those directly connected to initial targets of the attack, realizing the danger that they will “spill the beans,” may be willing to intervene. They are likely not willing to do the same for more peripheral network members, who therefore may be more willing to defect.

Going after peripheral individuals can serve to temporarily hide the identity of the eventual target. If the common findings from social network analysis, namely that most individuals are only 4-5 degrees removed from any other individual (Travers and Milgram 1969) hold true, then a peripheral individual is indirectly connected to many
possible targets. For instance, an inspection team finding “rampant problems” (South China Morning Post, July 20, 2014) in Tianjin, China led some observers to speculate about an impending attack on Zhang Gaoli, a Politburo Standing Committee member who was party secretary in Tianjin beforehand. But such a move has not materialized, and it seems likely that many other elites could be directly or indirectly connected to the individuals arrested in Tianjin. Yet targeting the few people who are close to the rival makes the intention very obvious, especially as routine investigations of such powerful people are extremely unlikely.

Information collection is also important because it enables the leader to convince the public that removing the rival is necessary and legitimate—thus forestalling possible collective action in support of the rival. Party cohesion has been the key to the survival of single-party authoritarian regimes (Huntington 1968; Geddes 1999; Smith 2005; Levitsky and Way 2012). To remove rivals without splitting the party or triggering a popular uprising, the leader must justify the action to both other elites and the public. Many removed elites discussed later in this article were charged with criminal offenses, such as corruption or treason, or accused on an ideological basis. Their removal is often followed by public trials, such as China’s trials of the Gang of Four and former Politburo members Chen Xitong, Chen Liangyu, or Bo Xilai.

The encircling strategy has a few other advantages. For instance, it allows the leader to test the waters, and to backpedal if he faces stronger resistance than expected. The resources a rival is able to muster to protect his peripheral connections give the leader an idea of the rival’s true strength. An immediate arrest or investigation of the rival
or their closest associates is much more likely to trigger a full-blown counterattack, the
strength of which may be hard to estimate.

This discussion leads us to expect the following more formalized outcomes:

**Information hypothesis**: *Individuals whose network neighbors have been arrested or are
currently being investigated are more likely to be arrested or investigated in the
following time period.*

This hypothesis follows from the theory of the “war of information” and the
necessity of convincing other elites (and perhaps the wider populace) that the removal is
justified. The investigation and arrest thus become “contagious,” as anyone connected to
the targeted individual runs the risk of being implicated by him or her.

**Encircling hypothesis 1**: *Individuals who are removed (arrested or investigated) first—
i.e., without network neighbors who were arrested or under investigation—tend to be
more peripheral.*

**Encircling hypothesis 2**: *Individuals who are at the center of the network are only likely
to be investigated after many of the individuals constituting their power base have
already been removed.*

These two hypotheses capture the encircling theory proposed above. Individuals
far away from the powerful actors at the center, or who have few connections, are more
likely to defect, knowing that no one will come to their aid even if they hold out. Once a
few peripheral actors have been “turned,” the investigation will spread toward the center,
but the powerful central figures will not be removed until after most of their supporters have been removed.

While the encircling strategy may appear intuitive, there are also arguments against it. One obvious downside is that it gives the rival time to organize countermeasures, which could be prevented if the leader arrests the opponent and powerful lieutenants first. This decapitation strategy is commonly employed by the police when targeting criminals, or by the military in counterterrorism operations (Johnston 2012). Social network studies have also focused on strategies that target the most central actors first. In typical social networks (i.e., those with skewed degree distributions and transitivity, see below) such an approach is most likely to split the network into multiple, disconnected parts (Albert, Jeong, and Barabasi 2000), which makes it harder to organize a counterattack.

But the removal of a political rival and the leader of a criminal gang may be quite different situations. Instigators of the latter rarely have to fear for their own lives—in the worst case, the criminal or terrorist leader simply escapes punishment. In addition, the attacker is not usually connected to the targeted network. The attackers can thus focus on cleaning up the opposing network as efficiently as possible. Yet leaders taking out a rival not only face much higher personal risks; they are also embedded in a network of complex relationships involving their rivals and supporters. Removing rivals could change their own network position in unpredictable ways, and splitting the opponent’s network may lead to a wider elite split. Particularly if their own position is not yet secured, they may prefer to employ a more cautious approach.
Empirical Tests

We test our theory using three case studies: Xi Jinping’s removal of Zhou Yongkang and his associates during China’s anti-corruption campaign in 2014, Andropov’s weakening of the Brezhnev network in the Soviet Union in the 1970s, and Mao’s removal of the Comintern Gang during the Yan’an Rectification Movement in the 1940s. Our main focus is on the case of Zhou Yongkang, for which we are able to quantitatively test our hypotheses using data collected during China’s anti-corruption campaign from public sources.

The three cases all fit our scope condition: the removals occurred during leadership transitions, and in all three instances the actions were targeted at a large number of individuals in an elite’s network rather than a single individual. They all occurred in communist or post-communist regimes, but in different time periods and geographic locations, showing the broader applicability of the strategy proposed.

Case Study One: Xi Jinping’s Removal of Zhou Yongkang

Starting in late 2012, the new leadership of the Chinese Communist Party led by Xi Jinping announced an anti-corruption campaign to attack both “flies” (low-ranking officials) and “tigers” (high-ranking officials). The true intention of the campaign is open to debate. While optimists argue that cleaning the party and government is a necessary step towards further economic reforms, skeptics maintain that it is motivated by Xi’s need to consolidate his power (Jiang and Xu 2015).2 Investigating the true motive of the

2 For a debate, see http://www.chinausfocus.com/political-social-development/debunking-misconceptions-about-xi-jinpings-anti-corruption-campaign/ and
anti-corruption campaign is beyond the scope of our study, which focuses on how it was carried out.

More than one hundred officials at or above the ministerial level had been investigated. The biggest tiger that had fallen from grace was Zhou Yongkang, a former Politburo Standing Committee member (2002–12) who, because of his position at the head of the Ministry for Public Security, was dubbed China’s Security Tsar by the media.

There are several reasons why we chose this example for our network analysis. First, the Zhou Yongkang network involves a large number of individuals, which makes a large-N test possible. Indeed, a qualitative analysis of almost 120 individuals and the more than 7,000 possible connections between them would not be feasible. Second, the Zhou Yongkang case has generated hundreds of media reports with often detailed information about connections between individuals. This relative transparency in the ongoing anti-corruption campaign in China makes our empirical study possible.

We construct the Zhou Yongkang corruption network relying on credible media and government reports primarily from Chinese-language sources. We employ a snowball strategy, starting with the individuals who allegedly engaged in corrupt activities with Zhou. Each time a new name was reported, we ran an exhaustive Internet search for the names of further individuals to whom he or she was reportedly connected through corrupt activities, and added these new names to the database. By the end of July 2014, when Zhou Yongkang was officially announced as being under investigation, and the campaign had started to focus on individuals not connected to Zhou, our dataset

contained 119 individuals. This included Zhou Yongkang’s family and colleagues, and businessmen and businesswomen who have had business transactions with people in the network. For individuals with publicly available CVs or short biographies, we also collected and included this information in the dataset.

Snowball samples are representative of the underlying population only in very specific circumstances (Salganik and Heckathorn 2004). However, our goal is not to draw general inferences about the characteristics of the individuals arrested during the current anti-corruption campaign. We are only interested in the timing of the arrests of individuals directly or indirectly connected to Zhou Yongkang through joint corrupt activities. It is possible that we do not observe the whole Zhou Yongkang corruption network: in particular, if an individual enjoys protection from other leaders (note, for instance, that none of the individuals in our dataset has ever worked with Xi Jinping), their corruption ties to Zhou and his associates may not be investigated or revealed to the public. In the current political climate, there is no way to measure the “true” network except through what is revealed in the press and public statements by the government. When measuring an actor’s “power base” we therefore use the absolute numbers of removed supporters as the independent variable. An alternative specification would be to use the percentage instead, which is often done in threshold models (Granovetter 1978), but this would require that we know the numerator, the true size of the power base. Instead we rely on the observed change instead. However, the result is similar if we use the absolute number of remaining supporters or include a covariate for the number of direct network neighbors (degree). We also run our main models on a dataset that leaves
out individuals who have not been investigated or arrested until now (see Table 4A in the online appendix). Results on the number of supporters.

The 119 individuals in the Zhou Yongkang corruption network are connected by a total of 187 corruption ties (defined as joint corrupt activities as reported in the news). Like many social networks, this corruption network is sparse: only 2.7% of all possible ties actually exist. It also displays a higher level of triadic closure or transitivity and degree skewedness (see Figure 4), which is typical of social networks. In other words, one’s partners in crime are likely to commit crimes together as well, and while a few individuals have many ties, the average person has only 3.14.

Figure 1: The Zhou Yongkang Corruption Network

Notes: Size of node proportional to closeness centrality, intensity of node color proportional to betweenness centrality. The layout used is force atlas as implemented by gephi (Bastian, Heymann, and Jacomy 2009, Jacomy et al. 2014).
Figure 1 shows the network of individuals directly or indirectly connected to Zhou Yongkang through an alleged instance of corruption as reported in the newspapers or by government sources. While Zhou Yongkang has the highest degree—that is, the largest number of ties (27)—other individuals have amassed a substantial number of co-conspirators as well, including big names like Jiang Jiemin, former chief executive of China Petroleum (18) and Li Chuncheng, former deputy party secretary of Sichuan (19), as well as tycoon Liu Han (19) and Zhou’s son Zhou Bin (15). Zhou Yongkang is also the most central individual according to most other commonly used centrality measures thought to capture informal power (Brass and Krackhardt 2012). But it is also possible that this result is due to the investigators’ sampling method or strategy (see below). Figure A1 in the Appendix focuses on the ties instead, and shows that Zhou Yongkang, with one exception, is connected only through kinship and coworker ties. This seems to confirm the conclusion that many journalists have drawn from investigating the networks of Politburo Standing Committee members: these leaders are almost never directly involved with businesses, but instead let their family members hold and manage assets.³

In our hypotheses, we propose that the order and timing of arrests or investigations provides insights into the underlying logic of the elite struggle. Figure 2 illustrates how investigations (orange/light grey) and arrests (red/dark grey) occur in the network over

---

³ For example, the New York Times reported that Wen Jiabao, China’s former premier, controls a large amount of wealth through his family:

time by displaying it at different time periods. The top-left corner displays the network at the beginning of our investigation, at the end of November 2011. The only individual already removed is He Bing, who had been arrested almost a decade earlier. Roughly four months later, rumors swirled in Beijing about a military coup involving Zhou Yongkang and the recently dismissed Bo Xilai. But even almost a year into our timeline (top right), when the Communist Party convened to select the new leadership, only a handful of individuals had been arrested. The first tiger, Li Chuncheng, was dismissed from his position as deputy party secretary of Sichuan province at the end of that very year (December 2012). Half a year later, when Zhou Yongkang made one of his last public appearances, many more individuals connected to Li Chuncheng had been arrested, as had an infamous Sichuanese businessman (Liu Han). At the end of August 2013, the Hong Kong South China Morning Post first reported that Zhou was being investigated (an account confirmed almost a year later by official sources). By that time, many individuals connected to another important ally of Zhou, Jiang Jiemin, were under investigation. But before the Chinese Communist Party’s 18th Central Committee met again in 2013, Zhou and most of his network had been arrested, including his son, Zhou Bin. The latter was not announced until July 29, 2014, however.

4 We did not include Bo Xilai in the network because no reliable source has confirmed this rumor.
Figure 2: Investigations (orange) and Arrests (red) in the Zhou Yongkang Corruption Network at Different Time Periods

While highly suggestive, Figure 2 does not constitute a test of our hypothesis. For a statistical test, we implement a contagion hazard model, in which the hazard or risk of infection—in our case being arrested or investigated—depends on whether one is connected so someone already arrested or under investigation. We implement the hazard
model through a regular logistic regression with the individual in each week as the unit of observation. As Beck, Katz, and Tucker (1998) have shown, such a setup is equivalent to a hazard model if three conditions are met: (1) dummies for the time “at risk” are added to the equation, (2) observations are dropped once they have succumbed to the hazard, and (3) the standard errors are clustered on the individual. Following Carter and Signorino (2010), we replace the time dummies with polynomials. The binary dependent variable is the removal of the individual in question: either their arrest or the official start of their investigation.

Table 1: Predicting the Probability of an Investigation/Arrest in a Given Week

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>has removed neighbor</td>
<td>1.57**</td>
<td>1.57**</td>
<td>1.58**</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.37)</td>
<td>(0.38)</td>
</tr>
<tr>
<td>bureaucratic rank</td>
<td>-0.08*</td>
<td>-0.08*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>distance to Zhou</td>
<td>-0.15</td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.15)</td>
<td></td>
</tr>
<tr>
<td>initial degree</td>
<td>0.05*</td>
<td>0.05*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>People’s Congress member</td>
<td>0.15</td>
<td></td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td></td>
<td>(0.34)</td>
</tr>
<tr>
<td>work experience at center</td>
<td>-0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>-0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>week</td>
<td>-0.06</td>
<td>-0.07</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>week²</td>
<td>0.00*</td>
<td>0.00*</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>week³</td>
<td>-0.00*</td>
<td>-0.00**</td>
<td>-0.00*</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>constant</td>
<td>-6.37**</td>
<td>-6.14**</td>
<td>-6.02**</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(0.95)</td>
<td>(0.98)</td>
</tr>
<tr>
<td>number of observations</td>
<td>12,830</td>
<td>12,830</td>
<td>12,680</td>
</tr>
</tbody>
</table>
Notes: Hazard models (logistic regression with time (week))
polynomials, standard errors clustered on the individual). Standard
errors in parentheses. Levels of statistical significance: ** p<0.01, *
p<0.05, + p<0.1

The sole independent variable in Model 1, apart from the time polynomials, is a
dummy indicating whether any network neighbors have been arrested or are currently
being investigated. This dummy turns out to be a positive and significant predictor:
having one or several such neighbors more than quadruples one’s weekly probability of
getting into trouble with the investigators from 0.27% to 1.26%.

Model 2 tests the finding’s robustness by adding two covariates that measure the
power of the individual and his or her importance to the rival: official rank and closeness
to Zhou Yongkang within the network. An individual’s rank ranges from the highest-
level national positions (10) to the lowest local positions (2). Businessmen and other
individuals without official position are assigned a 0. A higher rank is associated with a
lower weekly probability of arrest or investigation, but the result is not very robust.
Distance to Zhou Yongkang (ZYK) measures the number of “steps” actors are removed
from him. 1 indicates that they are directly connected, 2 that they are connected to
someone who is directly connected to Zhou, and so forth. The coefficient is not
significant. On their own, these two variables are moderately correlated with the timing

5 The Chinese bureaucratic rank has ten levels: state (guoji), deputy state (fu guoji),
minister (buji), deputy minister (fu buji), bureau (juji), deputy bureau (fu juji), office
(chuji), deputy office (fu chuji), section (keji), and deputy section (fu keji).
of the arrest. But their effect works through the necessary number of removed network
neighbors, as we will show later. The model also includes degree, the number of ties (in
other words, the number of partners in corruption) an individual has and a possible proxy
for their level of corruption. It is indeed positively associated with being arrested.

Model 3 includes two additional covariates that might facilitate or hinder an arrest
or investigation: membership in the People’s Congress and having worked at the center
(i.e., Beijing). Neither appears to provide some form of protection. Gender also does not
affect the outcome.

In order to test Encircling Hypothesis 1, we compare two groups of individuals: (1)
those with whom a chain of investigation starts (the “patient zeros” or the “seeds” that
start the contagion) and (2) those who are removed only after one or more of their
neighbors has been investigated or arrested. If the investigators only want to quickly
removing all corrupt figures, their best strategy is to target the most central actors—those
with the largest number of connections or those with positions closest to the center.
Highly corrupt or connected victims can implicate many other network members, and
targeting individuals closest to the center ensures that the investigators will need only a
few waves of roundups to acquire evidence on everyone in the network.¹

It is of course possible that the investigators do not know the shape of the network.
In that case they could either investigate individuals at random and arrest them as soon as
they have gathered enough evidence—in which case the individuals arrested or
investigated initially should not differ much from the overall population—or investigate
transactions at random. In the latter case, the high-degree actors should again be the ones
arrested early: the more crimes, the more partners in crime, i.e., the more connections that investigators could discover during their random inspections.

However, in our theoretical discussions we have proposed that elite removal follows quite a different—political—logic and that investigators should therefore target the least connected actors first. Figure 3 is a histogram of the actors’ number of corrupt connections that shows each group’s degree of distribution. The majority of the seeds have just one connection, and only one has more than three. In other words, the possibly least corrupt get arrested first.

Figure 3: Degree Distribution of the Actors in the Network
Notes: Histogram indicates number of ties of each individual in the network by whether they were the starting point of an investigation—i.e., whether they were investigated or arrested without having a neighbor under investigation (blue), or not (red).

The results of simple two-tailed t-tests comparing the 16 individuals who were arrested or investigated without having an immediate network neighbor under investigation with the remaining 103 individuals seem to support the first part of the encircling hypothesis: they have a significantly lower degree (1.75 instead of 3.36) and are less (standardized) closeness central. They are also ranked lower in the party hierarchy—if they have a rank at all—and are further removed (2.25 instead of 2.02 steps) from Zhou Yongkang in the network. The latter difference is only significant at the 10% level, however. The initial targets of the campaign are indeed the more peripheral network members.

As far as the later targets is concerned, a similarly striking finding emerges from simple correlations: the correlation between official rank and the number of neighbors under investigation or arrest at the time the person is arrested is 0.38 (or 0.67, if individuals without official rank are coded as missing instead of 0). The correlation between distance to Zhou Yongkang and the number of affected neighbors is -0.55, and the correlation with degree and closeness (both 0.73) is even stronger. Figure 4 shows the scatter plots of the first two relationships mentioned. The results thus support Encircling Hypothesis 2: those who rank higher in the party bureaucracy, are closer to Zhou Yongkang, or are in other ways more central to the network are only officially investigated or arrested when many of their connections have already suffered the same fate.
Figure 4: Correlation between Distance to Zhou Yongkang in Network and Official Rank With the Number of Network Neighbors Under Investigation or Arrest at the Time of Investigation/Arrest

*Note:* Scatter plots (with jitter of up to 0.2).

Figure 4 and the correlation coefficients thus all seems to hint that powerful individuals are treated differently, and that their power first has to be reduced before they can be removed. But how to measure this reduction in an individual’s power? The measures used so far, such as official rank and distance to Zhou Yongkang, remain largely unchanged during the period of investigation.

In this particular network, however, there is a relatively simple way to conceptualize the power base. It is the number of individuals willing to defend the target against the investigators. An investigation of one individual puts his or her neighbors at a direct risk of being investigated as well (as we have seen models 1-3 in Table 1), either because the individual’s interrogation or his or her investigation provides incriminating
information. With the process of committing corrupt activities thus forming co-
dependency (as observers have pointed out), direct network neighbors are likely to
oppose the investigation.

But these are not the only actors who might be opposed to the investigation. After
all, the neighbors of the immediate neighbors could also be worried: once the initial
target has been taken out, the investigation might spread to them via the immediate
neighbors. This is not to say that the whole corruption network will defend every member
under attack, and therefore everyone’s power base is equal. The farther removed an
individual, the lower her awareness of the indirect connection between her and the target,
as she likely does not know all the corrupt activities that others were involved in. There is
also a higher chance that the investigators will fail to discover a specific connection, or
that the campaign will stop before it reaches the individual.

Thus, the farther removed an actor, the less willing he or she will be to stand up
for target. One intuitive way of measuring an individual’s power base thus counts his
direct and indirect neighbors, with the latter discounted by their network distance, e.g. the
number of steps removed from the individuals.\(^7\) Figure 5 shows the size of the power
base of member of the corruption network over time. Some of the “tigers” are marked in
black and different line types. The power base measure has been recalculated every time
an individual was put under investigation or under arrest and thus removed from the
network.

\(^7\) A form of network centrality, closeness centrality, is already calculated in a very similar way, but usually
divides the sum by the size of the network to measure the average distance between the individual and
everyone else. For our purposes, we measure the “size of the power base” as this weighted headcount,
without further dividing the sum.
Figure 5: The power base of each individual in the ZYK network over time.

The individuals identified as powerful, such as Zhou Yongkang and his personal secretaries or associates, as well as his son, do indeed have some of the largest power base at the beginning. But as individuals are gradually removed, even their power melts precipitously. By the time they are put under arrest or investigation, it is often smaller than that of the average member at the beginning.

However, measuring absolute power in this way is still fraught with problems: as mentioned earlier, it seems likely that we do not observe all corruption ties. And actors may have power resources that are not related to having committed corrupt activities with others. For the statistical analysis, we thus take the difference in the power base since the beginning, as opposed to the absolute number.
Table 2 shows that loss of power base is indeed associated with an increasing likelihood of being removed in Model 1. This finding is robust to the inclusion of the same covariates as in Table 1.

Table 2: Predicting the Probability of an Investigation/Arrest in a Given Week

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>loss of power base</td>
<td>0.03*</td>
<td>0.03*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>loss of direct supporters</td>
<td>1.22**</td>
<td>1.29**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>loss supporters (2 steps removed)</td>
<td>-0.04</td>
<td>-0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>loss supporters (3 steps removed)</td>
<td>-0.00</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>loss supporters (4 steps removed)</td>
<td>-0.02</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bureaucratic rank</td>
<td>-0.05</td>
<td>-0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>distance to Zhou</td>
<td>-0.04</td>
<td>0.64+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>initial degree</td>
<td>0.10**</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>People’s Congress member</td>
<td>0.27</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>work experience at center</td>
<td>-0.50+</td>
<td>-0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>-0.37</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>week</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.08)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>week^2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>week^3</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>constant</td>
<td>-7.27**</td>
<td>-7.43**</td>
<td>-6.91**</td>
<td>-8.56**</td>
</tr>
<tr>
<td></td>
<td>(1.14)</td>
<td>(1.24)</td>
<td>(1.31)</td>
<td>(1.81)</td>
</tr>
<tr>
<td>number of observations</td>
<td>12,829</td>
<td>12,679</td>
<td>10,098</td>
<td>10,043</td>
</tr>
</tbody>
</table>

Notes: Hazard models (logistic regression with time (week) polynomials, standard
errors clustered on the individual). Standard errors in parentheses. Levels of statistical significance: ** p<0.01, * p<0.05, + p<0.1

The measure for power and loss thereof may strike one as too crude: would really half those two steps removed from the target come to his or her aid? Model 3 and 4 therefore allow the data to speak about the decay function, and enter a variable for each circle of supporters. As it turns out, it is mainly the removal of those directly connected that weaken a target – either because they are those most likely to support the target, or because their removal makes the target lose supporters farther away. The other power loss variables are individually not significant, but together with the direct supporters jointly highly significant. The results are similar (but with reversed sign, obviously) if we measure power instead of loss of power as Table A2 in the online appendix (Model 1). In that table, we also distinguish between losing direct supporters to arrests and only to investigations. Both have a significant effect (Model 2 and 3), but it’s mainly the investigations that have an impact when entered jointly (Model 4). Table A3, finally, does not include in the analysis those individuals that were mentioned in news report, but apparently were investigated or arrested. The results do not change.

In the following sections, we provide suggestive evidence that the three patterns—starting at the periphery, attacking network neighbors of earlier targets, but holding off from powerful figures until they have been weakened—applies to purges during leadership struggles in other times and countries as well.

**Case Study Two: Andropov’s Removal of the Brezhnev Network**
Some observers have argued that Yuri Andropov’s rise to general secretary of the Soviet Communist Party (November 1982 – February 1984) is an example of how an anti-corruption campaign can be used to weaken political opponents (Clark 1993; Solov’ev and Klepikova 1995). In many ways the system itself facilitated the (ab)use of corruption charges for political purposes. It was nearly impossible for any official to fulfill the requirements of the plan imposed by the center without breaking rules in order to, for instance, acquire necessary materials in short supply (Clark 1993). Urban (1986 216) and others have thus argued that this was intentional: the center kept the agents in a double bind, forcing them to violate an expectation or rule no matter their decision. The agents responded by establishing a close network of informal relationships with coworkers and subordinates to reduce the risk that each would report the wrongdoings of the other to superiors or inspectors.

This made almost any official vulnerable to the accusation of corruption. As chief of the Committee for State Security (KGB), Andropov was thus in the perfect position to wield corruption charges as a political weapon (Duhamel 2010) in order to weaken other contenders in line to succeed Brezhnev, and even to attack the party secretary himself.

One of the first cases of interest occurred in the periphery, in Georgia, where first party secretary and candidate member of the Politburo, Vasilii Mzhavanadze, was removed from office in 1972 after a stolen diamond ring had been spotted on the finger of his wife Victoria. Victoria had allegedly been a close friend of Brezhnev’s wife (Clark 1993: 155). The man who had collected the evidence against her, the Georgian Minister of Internal Affairs Shevardnadze, replaced her husband.
Another early move in the periphery that may have strengthened Andropov was the replacement of Azerbaijani Party Secretary Akhundov with the republic’s KGB chairman Gaidar Aliev in 1969, two years after Andropov had become the head of the KGB (Willerton 1992). Aliev also relied heavily on corruption charges to replace Akhundov’s network with his own people (Clark 1993: 157).

But the anti-corruption campaign really gained momentum in 1979 when Brezhnev’s health started to decline (Clark 1993: 162). One line of attack began with the conviction of four individuals involved in housing construction in the city of Sochi in Krasnodar. A year later, the investigations had spread to the mayor of Sochi, Voronkov, and finally reached Brezhnev’s ally, the party boss of Krasnodar district, Sergey Medunov. In 1983, Medunov was officially removed from the Central Committee for “mistakes in his work.” In the meantime, his successor—who supported Andropov’s campaign—continued the effort by indicting several officials in the district. The net tightened around Medunov from multiple sides: an arrested trade union leader, Shibaev, allegedly implicated him, and some threads from the case against Deputy Fishing Industry Minister Rytov may have led to him as well (Clark 1993: 166–72). Also removed in the same session of the Central Committee was another Brezhnev ally, the chief policemen of the Soviet Union, Nikolai Shchelokov. With the investigations closing in, his wife had committed suicide a few months earlier, as did he a year later.

But the anti-corruption campaign got even closer to Brezhnev: in a twist worthy of a soap opera, Boris Buriatia—a clown at the world-famous Soviet Circus—was arrested for stealing tsarist diamonds from a fellow circus performer. Questioned by the KGB, he allegedly testified against Galina Churbanova, Brezhnev’s socialite daughter.
That Brezhnev’s daughter could be interrogated—although admittedly never officially charged—in Moscow, the heart of the Soviet Union (Burns 1982), must have sent a powerful signal to the ailing general secretary’s protégés. The fact that Brezhnev’s brother-in-law and first deputy chairman of the KGB, Semyon Tsvigun, had passed away only a few days earlier did nothing to squelch the rumors. The absence of Brezhnev’s signature on Tsvigun’s obituary—unlike that of many KGB and police officers—raised further speculations about a possible suicide.

Brezhnev’s death in November 1982 and the selection of Andropov over Brezhnev’s preferred candidate, Chernenko, for general secretary did not end the anti-corruption campaign, however. The KGB descended on Moscow’s wholesale and retail grocery network, presumably with the intent of incriminating Moscow’s first party secretary, Viktor Grishin, a supporter of Chernenko (Duhamel 2010). Unsurprisingly, some of the individuals arrested were close friends of Brezhnev’s daughter Galina Churbanova and her husband Yuri (Clark 1993: 185).

The campaign against corruption continued after Andropov’s death in February 1984 under Chernenko, but took on a quite different format with the initiation of Gorbachev’s reforms a year later.

**Case Study Three: Mao’s Removal of the “Comintern Gang”**

In 1942, Mao Zedong started the Yan’an Rectification Movement to transform the Chinese Communist Party from a Soviet-style Leninist Party under the leadership of the Moscow-based Communist International (Comintern) to a more peasant-based party led by leaders trained in China’s revolution and familiar with Chinese realities (Selden 1971).
Many scholars believe Mao used the campaign to replace the Comintern Gang—a few top party leaders who were trained in the Soviet Union and served as agents of the Comintern, such as Wang Ming, Bo Gu, Zhang Wentian, Wang Jiaxiang, and Kai Feng—with his protégés, such as Liu Shaoqi and Kang Sheng, and consolidate his supreme position in the party (Gao 2000; Seybolt 1986; Teiwes 1976).

Mao started the movement in the periphery without touching the gang members, who were all on the powerful Politburo. On February 1, 1942, Mao made his well-known speech denouncing subjectivism, sectarianism, and party formalism at the opening ceremonies of the reorganized central party school in Yan’an (Selden 1971; Seybolt 1986). Two months later, the rectification became systematized, and party members in party and government organs and schools were required to start self-criticism and to read and discuss Mao’s speeches. The first person who was publicly purged was a journalist named Wang Shiwei, an intellectual who shared the flaws of the Comintern Gang—idealism, out of touch with the peasantry, and bookishness. His essay “Wild Lily” had criticized the concentrated leadership in the party (Goldman 1967; Benton 1975). Wang was later arrested, and remained in detention until his execution in 1947.

The encircling continued with the counterespionage campaign in June 1942. Building on the foundations of the early rectification campaign, the leadership launched a screening in which thousands of cadres were elevated, demoted, or relieved of their duties (Seybolt 1986: 52). Major targets of the counterespionage campaign included people who studied or worked in the Soviet Union (Gao 2000: 443). During the campaign, Mao clearly used a network approach to identify campaign targets. For instance, in an
internal meeting on June 19, 1942, he suggested that “people who closely interacted with the ‘problem people’” should be the “key targets” of the screening (Gao 2000: 416).

Consistent with our information story, the early stages of the campaign primarily involved information collection. Key targets of the screening were asked to write reports on every detail of their whereabouts and meetings, and especially what they knew about the whereabouts, words, and meetings of the gang members (Gao 2000: 479–85).

The final strike occurred in November 1943 when Mao convened the Rectification Meeting of the Central Politburo. By then, Mao had collected all the evidence needed against the Comintern Gang. The meeting turned into a denunciation of Wang Ming and Bo Gu, during which Mao ordered Li Guohua, a party member who used to work in the Comintern, to criticize Wang Ming’s mistakes while working in the Soviet Union. The meeting ended with most of the gang members being labeled as sectarianist or formalist (Gao 2000: 615–41).

The Seventh Party Congress in 1945 marked Mao’s total victory: most of the Comintern Gang were expelled from the Politburo, Liu Shaoqi rose to No. 2, and Mao Zedong Thought was established as a guiding ideology of the Communist Party (Gao 2000: 641–44).

**Discussion and Conclusion**

We present the first systematic effort to theorize and examine strategies used to remove rivals during the power struggles of leadership turnover. We argue that authoritarian leaders face two challenges in removing powerful elites: exposing themselves to the risks of a counterattack and a party split. Faced with these two challenges, the leader’s best
strategy is to start the attack from the periphery of the rival’s network—a pattern that we term encircling. Attacking individuals on the periphery will avoid strong resistance early on and gradually eliminate a rival’s support. Arresting and investigating individuals on the periphery also helps the leader collect useful information that can be used against the target by convincing other party members and the public that the removal is necessary and legitimate. If used carefully, elite removal is the flip side of power sharing: party leaders can use both to achieve elite cohesion.

While we use two additional qualitative case studies to demonstrate the wider applicability of our findings, our focus is on Xi Jinping’s attack on the Zhou Yongkang network in China in 2014. Taking advantage of the rich information revealed in the Zhou Yongkang case and the tools provided by social network analysis, we are able to test these intuitive strategies more systematically. We have shown that the starting points of the purges tend to be the less powerful and well connected, and that the subsequent removal of elites appears to be a contagious process, albeit one in which individuals closer to the powerful center enjoy a higher level of protection.

Our study sheds light on the scholarly understanding of authoritarian politics in three important ways. First, previous research has overwhelmingly focused on how the inclusion of social groups and political elites affects authoritarian politics (Gandhi and Prezworski 2007; Magaloni 2008; Boix and Svolik 2013). We flip the question by asking how authoritarian rulers exclude elites, and show what strategies rulers use to maximize their tenure in office.

Second, we move away from the existing literature’s emphasis on authoritarian institutions (Schedler 2006; Lust-Okar 2006; Gandhi 2008; Malesky and Schuler 2010;
Gehlbach and Keefer 2011) and focus on authoritarian rulers’ strategies to, as Magaloni and Kricheli (2010) argue, “simultaneously minimize threats from the elites and from the masses.” We show that the encircling strategy enables the ruler to simultaneously prevent elite splits and popular uprisings that could be triggered by an abrupt elite removal.

Third, with a few exceptions (Arriola 2009; Francois, Rainer, and Trebbi 2014), most quantitative research on authoritarian regimes focuses on the survival of the leader, rather than other members of an elite inner circle. This approach to regime stability usually fails to capture the complex politics behind the scene described in historical accounts. Our article proposes social network analysis as one alternative approach with which to study this highly relevant aspect of authoritarian stability.
References


