Who Revolts? Empirically Revisiting the Social Origins of Democracy

Sirianne Dahlum[1], Tore Wig[2], and Carl Henrik Knutsen[3]

[1]sirianne.dahlum@v-dem.net, Varieties of Democracy Institute (V-Dem), Gothenburg University
[2]tore.wig@stv.uio.no, Department of Political Science, University of Oslo
[2]c.h.knutsen@stv.uio.no, Department of Political Science, University of Oslo

June 27, 2018

Abstract

Prominent accounts of democratization suggest that democratic transitions hinge on collective action by certain social groups. We develop and present the first large-n test of the argument that opposition movements dominated by particular urban groups have the requisite motivation and capacity to bring about democratization. To this end, we collect new data on the social composition of anti-regime opposition movements, globally from 1900-2006. Our study indicates that movements dominated by industrial workers or urban middle classes more often yield democracy, both when compared to other movements and to situations without organized mass opposition. As expected, the relationship is stronger in urban than in rural societies, and in more recent decades. When further differentiating the groups and accounting for plausible alternative explanations, the relationship between industrial worker campaigns and democratization turns out very robust, whereas the evidence is mixed for middle class campaigns.

Mass uprisings have played a key role in many episodes of regime breakdown and change throughout modern history. But, does a democratic outcome of such processes hinge on *who* takes part in the revolts? Many comparative social scientists and historians believe so – certain *social groups* are more conducive to democratization than others. Moore (1966), for example, argues that revolutions where the bourgeoisie partake induce democratization. Meanwhile, Rueschemeyer, Stephens and Stephens (1992) propose that the industrial working class is the main agent of democratization (see also O'Donnell, 1973; Collier, 1999), whereas Stephens (1989) holds that peasants – under certain conditions – are instrumental in democratic revolutions. The preferences and capacities of different social groups are prominent explanatory factors also in the more recent democratization literature (Acemoglu and Robinson, 2006; Boix, 2003; Ansell and Samuels, 2014).

Despite the long-standing debate, large-n evidence directly linking the social composition of anti-regime movements to democratization is missing. Numerous excellent country-case and small-n comparative studies exist (e.g., Moore, 1966; Luebbert, 1991; Rueschemeyer, Stephens and Stephens, 1992; Collier, 1999), but conclusions on the relevant (coalitions of) actors vary with the cases under scrutiny. This inevitably leaves question marks concerning the generality of the extant case-based evidence. To be sure, several large-n studies speak indirectly to who the most relevant agents of democratization are. For instance, Ansell and Samuels (2014) assume that the middle class should be the main agent promoting democratization under certain conditions, while Acemoglu and Robinson (2006) expect that a group defined as the "the poor" is the greatest threat to dictatorships (see also Boix, 2003). Yet, these studies rely on fairly distant proxies of the preferences or capacities of social groups, such as urbanization, GDP per capita, or land or income inequality (e.g., Lipset, 1959; Boix, 2003; Ansell and Samuels, 2014). Finally, recent work on non-violent conflict has measured and linked characteristics of mass opposition movements to democratization, but these studies focus mainly on tactical choices, strategic interactions, and macro-structural conditions enabling successful movements (e.g., Chenoweth and Stephan, 2011; Celestino Rivera and Gleditsch, 2013; Bayer, Bethke and Lambach, 2016). Against this backdrop, we offer the first global large-n study of its kind on

1

the social composition of mass opposition movements and democratization.

We develop a theoretical argument clarifying how and why the social class composition of mass opposition movements affects democratization, implying that collective action organized by *urban groups* is particularly democracy enhancing. More specifically, we detail how industrial workers and the urban middle classes have *both* the requisite motivation *and* the capacity for effectively bringing about democratic regime change – at least in fairly urban and industrialized (or even industrializing) societies. Other social groups often lack either the capacity (e.g. peasants) or motivation (e.g. military officers) to pursue democratization.

We collected new data on the social composition of opposition movements to test these expectations, measuring presence and level of participation of six major social groups in about 200 anti-regime opposition campaigns globally from 1900–2006 (identified in the NAVCO dataset Chenoweth and Stephan, 2011). We first test the cruder proposition that opposition movements dominated by one of the highlighted urban groups are positively related to democratization. Second, we distinguish between opposition movements dominated by industrial workers and by the urban middle classes to investigate whether any one of them are especially conducive to democratization. We employ different relevant specifications and estimators to carefully assess robustness and account for plausible alternative explanations, and our analyses comprise both cross-sectional models at the opposition-campaign level and cross-section time-series models. Hence, we provide fresh evidence bearing on a long-standing debate in the macro-comparative social science literature concerning which particular urban group is the most relevant agent of democratization.

Based on this novel dataset, we present the first large-N evidence that the social composition of social movements is systematically related to democratic outcomes. In particular, we show that anti-regime opposition movements are systematically more likely to yield democratization – particularly in societies that display fairly high degrees of urbanization – if they are dominated by one of the two urban groups in question. Opposition movements dominated by other groups, such as peasants, are far less likely to induce democratization – indeed, they do not even raise the likelihood of democratization relative to situations *without any opposition*

2

mobilization. This finding is robust, and holds for instance when controlling for other features of the opposition movement, such as its size, the regime's tactical responses, or when accounting for urban opposition movements being more common in certain countries and time periods. The finding also holds up in specifications trying to account for the possible endogeneity of opposition campaigns and their social profile (some groups may, e.g., more likely bandwagon on movements perceived to have greater chances in successfully inducing democratization). When further distinguishing between the urban middle classes and industrial workers, there are indications that campaigns dominated by either one of the groups are conducive to democratization. Still, our results are much more robust for industrial worker campaigns, corroborating arguments by Rueschemeyer, Stephens and Stephens (1992) and Collier (1999).

In the next section we review the relevant strands of literature. After that, we present our argument elaborating on why both industrial workers and urban middle classes may have the requisite motivation *and* the capacity to effectively engender democratization. Next, we present our new data on opposition movements, and discuss coding protocol and issues and tests of reliability and validity. In the analysis section, we present results from, first, cross-sectional campaign-level analysis and, second, our core country-year panel specifications. Before concluding, we evaluate alternative explanations and potential sources of bias by introducing several additional tests.

Existing theories and evidence

Democratic transitions often emerge from organized mass movements that force the incumbent regime from power, often referred to as "democratic revolutions". Such organized collective action has been instrumental for democratization episodes in Latin American (Collier, 1999), Eastern European (McFaul, 2002), Sub-Saharan African (Bratton and van de Walle, 1997) and even Middle-Eastern (Tunisia, Stepan, 2012) countries. Democratization can also result from other processes, including coups (Powell, 2012), transitions negotiated between elite groups (O'Donnell and Schmitter, 1986), the incumbent relinquishing office more or less voluntarily (Bunce and Wolchik, 2010), or guided "top-down" liberalization (Acemoglu and

Robinson, 2006). Yet, even such processes are often preceded, and likely even triggered by, mass uprisings. Mass uprisings can put pressures on elites bargaining for a "pacted transition" (Collier, 1999), outright scare autocratic incumbents to provide democratic concessions (Acemoglu and Robinson, 2006), or increase chances of a coup d'état against the old regime (Casper and Tyson, 2014). While relevant throughout modern history, the share of regime changes stemming from actions by broad popular movements have been increasing in recent decades (Kendall-Taylor and Frantz, 2014). As Teorell (2010, 100) notes "collective action undertaken by the mass public appears to have been a widely occurring phenomenon, with alleged democracy-enhancing effects" during the "third wave of democratization".

But, far from all opposition movements have been successful in overthrowing dictators and ushering in democracy. Employing a recent, global dataset on major protest movements (Chenoweth and Stephan, 2011), an emerging literature has investigated why some movements succeed and induce democratization whereas others do not. However, these studies mostly focus on tactical choices by opposition movements, their cooperation with external agents, or strategic interactions with the regime, notably finding that *non-violent* mass mobilization is an effective democratizing force (e.g., Celestino Rivera and Gleditsch, 2013; Bayer, Bethke and Lambach, 2016). Other studies suggest that structural factors, such as industrialization, help explain features and outcomes of opposition movements (Butcher and Svensson, 2014; Chenoweth and Ulfelder, 2015).

Yet, these studies have not investigated whether the *social background* of the mobilizing actors matters for democratization. Promising clues to answering this question come from the theoretical and case-study literature on revolutionary movements. Marx and Engels (2010/1848) famously claimed that liberal bourgeois "democracy" is brought about by the capitalist bourgeoisie revolting against the old feudal order, whereas the next phase of revolution is acted out by the working class, turning against the bourgeoisie. While Marxist theory no longer dominates democratization scholarship, one key notion has survived: Economic development and industrialization is often presumed to empower "new" social groups, notably the urban middle class and industrial workers, with an interest in democratic change. Where scholars part ways,

4

however, is in their favored candidate for which group is most important when it comes to mobilizing for democracy.[1]

Many contributions, including variants of "modernization theory", highlight that industrialization and urbanization strengthen the *urban middle classes*, which will subsequently bring about democratization (e.g., Lipset, 1959; Glaeser, Ponzetto and Shleifer, 2007).[2] Inglehart and Welzel (2006), for instance, suggest that middle class citizens – due to favorable educational and material circumstances – develop liberal values, resulting in a strong desire for democracy. Pursuing a quite different argument, Ansell and Samuels (2014) highlight that urban elites and middle classes prosper from industrialization and wider economic development, and that democracy provides institutional protections from predation by "old" elite groups. They present indirect evidence for this hypothesis, showing that high income inequality – proxying for middle class strength – is related to democratization. Using different proxies and a panel of developing countries from 1985–2013, Chun et al. (2015) also find that strong middle classes promote democratization.

An alternative school of thought contends that industrial workers are the main agents of democratization, and that the urban middle classes, in certain contexts, may actually have incentives to ally with other elites and work *against* democratization (see O'Donnell, 1973).[3]

---

[1]Some contributions provide more complex hypotheses highlighting cross-class *coalitions*. Coalitions between the urban bourgeoisie and workers (Luebbert, 1991), or even peasants and urban middle-classes (Rokkan, 1999), have been proposed as conducive to democratization. Related, Goldstone (2011) argues that broad cross-class-coalitions are successful in virtue of their encompassing nature rather than the *identity* of participants. We address this below, finding clear evidence that broader coalitions, as such, enhance democratization.

[2]While we simplify and mostly treat them as one group ("urban middle classes") below, there may be disparities in income/wealth and political incentives between urban capital owners, professionals, and white collar workers. Rueschemeyer, Stephens and Stephens (1992), for example, use the term *bourgeoisie* for wealthy capital owners, and also distinguish various urban middle income groups. Nonetheless, urban white collar workers and professionals earn incomes far above the median in all real-world contexts (Ansell and Samuels, 2014), and the redistributive and other policy interests of these actors often converge.

[3]Also contrasting with the focus on urban middle classes, "redistributivist theories" suggest

Rueschemeyer, Stephens and Stephens (1992) argue that large numbers and organizational strength will give industrial workers leverage under a democratic system, providing strong incentives to fight for the introduction of mass democracy. In contrast, the upper- and middle classes often lack incentives to expand democratic rights, including suffrage, to poorer segments of the population, whereas peasants lack the capacity to organize a concerted effort to challenge the established regime. Collier (1999) – investigating 27 cases from Western Europe and South America (10 from the 19th and early 20th centuries and 17 from the 1970s onwards) – also attributes a key role to workers in engendering democratization. Related, large-n studies of franchise extensions following revolutionary threats conjecture that industrial worker movements precipitated democratization in the 19th and early 20th century (Aidt and Jensen, 2014; Przeworski, 2009), but without directly measuring the worker-dominance of revolutions.

Despite the many excellent contributions reviewed above, the nature of the current evidence means that we do not yet *know* whether, in general, *which particular social group* push for change matters for democratization. Empirical studies dealing *directly* with the role of social groups are mainly comparative-historical case studies, but findings vary and seem sensitive to the choice of cases. For example, Moore's conclusions about the democratizing role of the bourgeoisie draws on comparing England, France and United States with China and Japan. In contrast, O'Donnell (1973) argues that the urban middle classes may have a detrimental impact on democratization, drawing on Latin American experiences, in particular Argentina and Brazil in the 1960s. Large-n studies only offer *indirect* tests of the role of social groups in democratization, as macro-structural features such as GDP per capita, income inequality, or urbanization are used to proxy for the interests and/or capacities of particular groups (e.g., Ansell and Samuels, 2014; Boix, 2003). There is no neat mapping from socioeconomic structure to the interests or capacities of groups; to exemplify, measures of urbanization do not separate well between the strength of urban middle classes and industrial workers. This approach also misses that hurdles for collective action can vary by group (and across contexts) – peasants may face

that less wealthy actors have the strongest incentives to mobilize for democratization (Acemoglu and Robinson, 2006; Boix, 2003).

higher hurdles for organizing than industrial workers – thereby making it problematic to equate economic structures with the agency of particular groups. In addition to small-n and large-n contributions, excellent qualitative studies on a medium-large number of cases have been conducted (see especially Collier, 1999). Still, they face similar issues as small-N studies, most notably a difficulty of implementing a systematic control strategy and lack of explicit tools for evaluating the uncertainty of the inferences drawn.

We will return to how our new data help alleviate these issues for the empirical study of particular social groups and democratization, but first we outline our theoretical argument.

Theory

In this section, we synthesize insights from the reviewed literatures and develop our theoretical argument on why movements dominated by industrial workers *or* by the urban middles classes should be conducive to democratization. We shall assume that a social group's potential for bringing about democracy through mass mobilization is a function of both *motivation* and *capacities*. For democratization to occur, potential protesters should be motivated to confront the incumbent regime and work towards installing a more democratic system. In addition, they must have access to resources that provide them with strategic leverage over the regime and enable them to organize effectively. We first discuss motivation before turning to capacity.

Members of some social groups may prefer democratization more strongly than members of others. One reason, figuring prominently in the literature, is the differential economic benefits expected from democratization. Relatively poor but numerically strong groups is hypothesized to favor democratization, since democracy allows them to win elections and subsequently legislate progressive redistributive policies (Acemoglu and Robinson, 2006; Boix, 2003). In rural and less industrialized contexts this should make peasants, who are typically both poor and numerous, ardent supporters of democracy. Less numerous and wealthier groups, such as landowners or military officers, should not support democracy according to this logic.

While such redistributive models of democratization have received their share of criticism (Ansell and Samuels, 2014; Houle, 2009), we may relax the assumption of redistribution as the

7

key motivational force for democratization: If we assume, first, that individuals with similar socioeconomic features (type of occupation, education, income, etc.) are more likely to hold converging preferences also over other policies than taxation and redistribution (family policies, work regulations, free trade, etc.) and, second, we maintain that numerical strength yields electoral success, larger social groups should be more likely to prefer systems where policy makers are selected through free and fair elections.[4] Simply put, electoral democracy presents large groups with the institutional framework for enacting legislation that they favor, in different policy areas.[5]

In static economies, social groups with converging preferences and a numerical majority should thus prefer democracy. Industrial workers and urban middle classes, for example, might have preferences for democracy in highly urbanized, industrial societies, but not necessarily in predominantly agrarian, rural societies. Still, societies and economies develop and self-interested actors may be forward-looking. In a dynamic economy – for example characterized by rapid industrialization and urbanization – numerically large but declining groups, such as peasants, should *not always* prefer democracy, whereas groups on the rise, such as industrial workers, might. When a social group anticipates to become a minority in the near-to medium term, majority rule implies that policies will be determined by other groups in the future. Thus, urban middle classes and industrial workers may have particularly strong preferences for democracy in societies that are already fairly urbanized and industrialized, and where de-urbanization and de-industrialization are not expected to occur at rapid rates in the medium term. While processes of industrialization have come at different points in time in different countries – and processes of de-industrialization have set in for some richer countries in recent

---

[4]Even if a single group is not in majority, increased group size should, *ceteris paribus*, increase legislative weight.

[5]Social groups might systematically differ in their preference for democracy also for non-instrumental reasons. Modernization theoretic approaches (e.g., Inglehart and Welzel, 2006) suggest that the material security of the middle class – at least in relatively wealthy countries – may generate *inherent* normative preference for liberal institutions and policies. If so, middle class campaigns should display a particularly strong relationship with democratization.

decades – urbanization processes have been widespread especially from the 1950s in most regions (Satterthwaite, 2007), and have yet to show signs of reversal. This suggests that urban groups, at least in many developing countries, may have become more motivated to pursue democratization during the latter half of the 20th century.[6]

Regarding the *capacity* of a social group to bring about democratization, this relates both to members' abilities to *co-ordinate* collective action and to their *leverage*. Leverage refers to the power resources that a group can draw on to inflict costs on the regime, and thereby pressure the regime to accept its demands (for instance political liberalization). Leverage depends, inter alia, to the ability to grind the economy to a halt, thereby reducing tax revenues for the regime. This can be achieved, e.g., through moving capital assets abroad or through carrying out strikes in vital sectors. Groups differ in their ability to inflict such economic damage, and thus in abilities to force regime change. Other sources of leverage – increasing the threat of costly violence – include access to weapons, manpower with relevant training, and a militant ideology that can motivate recruits.

Having the ability to effectively organize large-scale collective action, including riots or mass protests, is vital for bringing about democratization. Large-scale collective action can result in democratization through various path; it may bring the incumbent regime to an involuntary end – either directly (popular revolution) or indirectly (e.g., by triggering a coup d'état Casper and Tyson, 2014) – or by "persuading" the incumbent to liberalize, unless s/he will face even worse consequences (Acemoglu and Robinson, 2006). The ability to create permanent and streamlined organizations with some centralization of decision-making power is one key ingredient to realizing effective collective action. Such organizations group members to alleviate collective action problems by providing infrastructure for effectively transmitting information and dispersing side-payments for participants (Olson, 1965). They also help with recruiting

---

[6]While patterns of relative numerosity and expected growth depend on different factors, one could speculate that industrial workers may fit the description of a large group that continues to grow better in intermediate stages of industrialization, whereas the middle classes fit better in later stages of industrialization and de-industrialization, with labor-saving technological advances and growth in high-income service-sector jobs.

new individuals for the cause, networking with foreign actors, experimenting with and learning effective tactics for pressuring the regime, and maintaining a fairly stable repertoire of crucial skills and manpower. Groups that can build comprehensive and enduring organizations should therefore have a stronger capacity to enact democratic change.

Regarding the specific social groups, industrial workers typically score high on the organizational aspect of movement capacity. In many countries, industrial workers are organized under labor unions and labor parties, which often exist illegally even when autocracies officially ban them. This gives industrial workers a potent infrastructure for orchestrating effective mass resistance. Workers often also have expansive international networks, including international labor organizations and the Socialist International, including more informal and amorphous cross-border networks in the labor movement. Many labor unions are also known for enforcing strict discipline on their members (see, e.g., Olson, 1965). Concerning leverage, industrial workers are often uniquely endowed with a strategic stranglehold over the economy through the ability to organize nationwide or localized strikes targeting crucial sources of revenue for the regime. Further, workers often have fairly high military potential, due to supreme organizational skills, foreign networks, and (historically often) being related to explicitly revolutionary, and sometimes violence-condoning, ideologies (see, e.g., Hobsbawm, 1974).

The urban middle classes also have potent sources of both organizational capacity and strategic leverage over authoritarian regimes. First, they often include a pool of individuals with high human capital and organizational skills, equipping them to alleviate collective action problems. This is reflected in the large number of civil society organizations, professional associations, student organizations that are often associated with the urban middle classes. Similar to industrial workers, urban middle classes often have access to international networks, examples being student networks and human rights advocacy organizations. Further, many urban professionals occupy vital inflection points in the economy, for example in finance and other vital service sectors. They might be particularly skilled at organizing non-violent resistance, utilizing comparatively high levels of human capital to legitimate their movement and strategically targeting the regime's soft spots.

Regarding other groups, military officers obviously score high on leverage, due to their relative supremacy in organized violence, but also on organizational capacity. Yet, it is questionable whether they would *want* to use this capacity for organizing mass movements that engender democratization. We discussed above how peasants, at least in rural societies, could have the requisite motivation to pursue democratization, due to their numbers. Stephens (1989) has then also argued that peasants – under certain conditions – can be instrumental in democratic revolutions. Still, peasants presumably score low on the capacity to effectuate regime change, as they often lack the ability to organize and express a common interest (see, e.g., Rueschemeyer, Stephens and Stephens, 1992; Kimeldorf, 2013). Notably, overcoming collective action problems may be more difficult due to dispersed settlement patterns. Peasants are supposedly also more easily co-opted through clientelistic practices *or* intimidated to follow the landed aristocracy (see, e.g., Ardanaz and Mares, 2013), who are often fierce opponents of democratization (Ansell and Samuels, 2014).

In sum, social group must have both a clear preference for democracy *and* the capacity to remove the incumbent autocratic regime to be an effective agent of democratization. While industrial workers and urban middle class groups may differ regarding what drives their preference for democracy, and particularly regarding the sources of their organizational capacity and leverage, we expect opposition campaigns led by these two groups – especially in fairly urban, industrial societies – to be conducive to democratizing regime changes.

Data

In order to systematically code the social profile of *anti-regime opposition movements* we identify the units using the NAVCO dataset (Chenoweth and Lewis, 2013). NAVCO describes various properties of campaigns globally from 1900 onwards: Each campaign represents

> a series of observable, continuous, purposive mass tactics or events in pursuit of a political objective. Campaigns are observable, meaning that the tactics used are overt and documented. A campaign is continuous and lasts anywhere from days to years, distinguishing it from one-off events or revolts. Campaigns are also

11

> purposive, meaning that they are consciously acting with a specific objective in mind, such as expelling a foreign occupier or overthrowing a domestic regime. (Chenoweth and Lewis, 2013, pp.)

To be considered a campaign, movements must have a discernible leadership, at least 1000 observed participants, and a coherent organization. Into this template, we introduce the concept of a *social group*, defined as *a group of individuals with a common social identity and/or a similar role in the economy giving them converging interests*. This is purposively defined in a rough and inclusive way, allowing for the coding of various – and even some partly overlapping – categories across different contexts. We collect data for all campaigns that NAVCO identifies as aiming for regime change or for "other goals", which include policy concessions or political liberalization. We have *not* coded social group characteristics for secessionist movements, which we consider qualitatively different and not obviously relevant for our research question. There are 203 relevant campaigns in NAVCO, and we collected information on the social composition for 193. 10 cases are completely missing due to no, or too uncertain, information. Our core social group categories are Peasants; Public sector employees; Military employees; Religious or ethnic groups; Industrial workers; and, Urban middle classes (see Appendix A for details).

We focus here on *Industrial workers* and *Urban middle classes*.[7] Industrial workers refer to labor employed in the manufacturing and mining sectors. The urban middle classes category is admittedly more complex to delineate, and we code this as encompassing business elites, smaller merchants, professionals (lawyers, doctors, etc), and students. Importantly, we treat public sector employees as separate from urban middle classes. One reason is that these are often mentioned separately in the sources used. More importantly, their tighter relationship with the state – and thus the government in many instances – means that their motivation to instigate regime change may differ from that of other members of the (broadly defined) middle

---

[7]Campaigns associated with other groups are simply considered as "other campaigns" in our benchmark statistical comparisons. Appendix Table A.20 reports regression models distinguishing further between other campaigns.

class. This choice is surely debatable, and we run analyses with and without public sector employees included in the urban middle classes category.

For each social group, we code the level of involvement in protest campaigns. We draw on information from various sources, notably including the Global Non Violent Action database (Swarthmore, 2015), a global catalogue of non-violent mass movements listing the social-group composition of movements and other features. Other important sources include International Encyclopedia of Revolutions and Protest (Ness, 2015). We also draw on cross-country and country-specific secondary sources. In many sources our social-group categories are clearly described; a source might, e.g., explicitly note that a movement was dominated by industrial workers, and that peasants also participated.

We code *three* dummies representing different levels of involvement. First, we code whether the group participated at some level; i.e., we register whether the movement at some point in time *consisted partly* of members of the given group. This participation dummy is scored 1 if the social group is mentioned in at least two separate sources. This is fairly easy to gauge, especially since the mentioned databases include lists of social-group participation where our categories are included.[8] Second, we code whether a movement *originated* among members of the given group, scored as such if noted explicitly in historical sources. Third, we code whether the opposition movement is *dominated* by the group. These instances include when a social group makes up a majority of campaign members or if it, according to the sources, was highly influential for the movement. While the latter obviously requires subjective evaluation, we specified different procedures to streamline coding decisions. In general, we opted to be conservative when setting the threshold for "highly influential". Our coders (RAs) have explicitly assessed their degree of certainty for each coding, ranging from very uncertain to very certain. Thus, we can test whether results are vulnerable to excluding the more uncertain cases (see Table 5; we remind that the most uncertain cases are always treated as missing).

---

[8]We do not interpret a 0 on this measure to mean that *no* single individual from the social group partook. It simply signifies that the presence of the group was not big enough to be recognized in the source materials.

Table 1 shows descriptives for the set of dummies (*originated; dominated; participated*) for the urban middle classes (MC) and industrial workers (IW). Appendix C provides additional descriptive statistics. While the frequencies have fluctuated over time (see Figure A.1), industrial workers participated in 62 percent of campaigns, overall, whereas the urban middle classes participated in 79 percent. About $1/5$ campaigns were dominated by urban middle classes, and likewise for industrial workers.

Table 1: Descriptives for urban middle classes- and industrial worker campaign variables

	Sum	N	Share
Middle classes dominates	41	187	0.22
Middle classes participate	148	188	0.79
Originated among middle classes	37	180	0.21
Industrial workers dominate	29	192	0.15
Industrial workers participate	116	186	0.62
Originated among industrial workers	19	192	0.10

Note: Highly uncertain codings are treated as missing, meaning that cases coded varies.

Let us illustrate the coding scheme with two examples (see Appendix B for other illustrations). First, the Velvet Revolution against the Czechoslovakian Communist regime in 1989 is described in our sources as a broad-based movement consisting of urban middle classes, industrial workers, and peasants; these three groups are therefore all registered as participating. According to, e.g., the Swarthmore database, the movement was sparked off by student demonstrations on International Students Day in November 1989, which soon developed into larger-scale mobilization including other groups. Moreover, the Swarthmore database describes the Public Against Violence organization – made up of, e.g., artists, scientists and intellectuals – as the "leading force". The movement is therefore coded as having originated among, *and* as being dominated by, urban middle classes. Second, the Senderista Insurgency (Sendero Luminoso) against the Peruvian government from 1980–1999 is described by several sources (e.g., Swarthmore, 2015) as founded by an alliance of peasants and a student group in San Cristobal. Hence, it is coded as having originated among and consisting of peasants and urban middle classes. Global Britannica states that throughout its lifespan, the movement's *main recruiting base* was among indigenous peasants and from poorer urban districts, and that these groups

were in majority and dominated the campaign during its final ten years. Hence, it is coded as dominated by peasants.

In order to assess inter-coder reliability, around 10 percent of the cases (21/203) were coded also by a second coder. 10 cases were double-coded based on brief summaries and references to the main sources used by the original coder, while 11 were coded without prior knowledge of sources used by the first coder.[9] Appendix Section D provides the results, which suggests that, overall, the reliability of our data is good. For instance, 14/16 campaigns where the original coder considered the middle class as participating were coded similarly by the second. 4/5 campaigns originally coded as dominated by the middle class were similarly double-coded. The corresponding numbers for industrial workers were 8/11 (participated) and 4/4 (dominated). The fit between the two coders was equally good regarding "negatives", i.e., observations coded as *not* including/being dominated by particular groups. Below, we leverage this information to explicitly account for measurement errors; this, in fact, strengthens our core regression results.[10]

For our dependent variable, we draw on two measures of (electoral) democracy with extensive coverage. Both capture two central dimensions of most modern notions of democracy, contestation and participation (see, e.g., Munck and Verkuilen, 2002). First, we use the binary measure from Boix, Miller and Rosato (2013) (BMR). BMR requires the presence of contested

---

[9]This distinction does not systematically relate to rates of agreement, and we thus run tests using all 21 double-coded cases jointly.

[10]While the data only include campaigns with at least 1000 participants, that are "observable", "continuous" and have a "discernible leadership" (Chenoweth and Lewis, 2013), and we thus mostly deal with campaigns that are well covered, another issue is that source coverage could correlate with the outcome variable. Nonetheless, different assessments give grounds for optimism: For instance, if data availability and quality is systematically poorer for violent campaigns (often smaller and carried out in remote areas), which extant studies show are less conducive to democratization, this could yield bias if groups differ in the likelihood of participating in violent and non-violent campaigns. Yet, there is virtually no correlation (r=.01) between "uncertainty status" of the coding and violence. Correlations between uncertainty status and participation by urban middle classes (r=-.13) or industrial workers (r=-.06) are also low.

and "free and fair" elections and that $1/4$ of citizens are enfranchised, providing a fairly low threshold of citizen participation, for coding a regime as democratic. Due to its binary nature, democratization will mainly pick up major regime changes, and miss out on non-trivial episodes of "liberalization" in nominally autocratic regimes and "democratic deepening" in regimes above the democracy threshold.

Our main measure, which allows us to also capture more incremental and less rapturous democratizing changes, is the continuous Polyarchy measure from V-Dem (Coppedge et al., 2016a,b). Polyarchy reflects the electoral democracy concept of Dahl (1971), and the theoretical range is from 0–1 (0.01–0.95 in the data). It includes indicators on whether the chief executive is elected (directly or indirectly), the "cleanness" of elections, freedoms of association and speech – due to the importance of free formation of parties and open discussion in ensuring truly competitive elections – and suffrage extension (Teorell et al., 2016). Hence, Polyarchy mainly captures the conditions for elite competition through multi-party elections and the right to participate in these elections.

Empirical analysis

Figure 1 shows the pre- and post-campaign levels of Polyarchy for campaigns dominated by different groups (we include peasants for illustration). Cases on/above/below the diagonal indicate, respectively, no/positive/negative change in Polyarchy. Most middle class- and worker-dominated campaigns are associated either with no change or democratic improvements. Indicatively, the three campaigns associated with the greatest increases in Polyarchy were either dominated by industrial workers (1983-89 Chilean campaign against Pinochet; 1984-85 Uruguayan campaign against the military regime) or by urban middle classes (1989 Czechoslovakian Velvet Revolution). In contrast, peasant-dominated campaigns are often associated with democratic backsliding.

Still, we cannot draw any inferences about social groups and democratization from Figure 1. Different groups may be dominant in very different societies, and campaigns might just happen to co-exist with changes in democracy without playing any active role. Thus, we probe
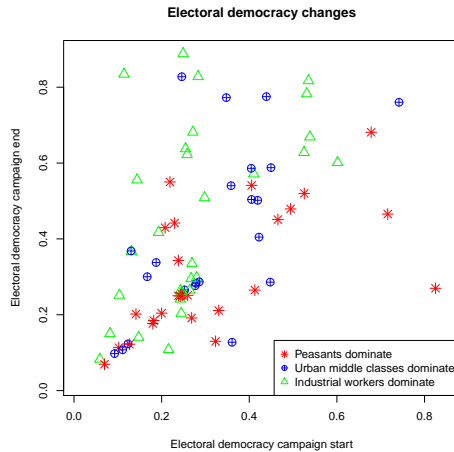
Figure 1: Opposition movements dominated by industrial workers, urban middle classes, and peasants, and Polyarchy scores measured at the campaigns' start-year (X-axis) and end-year (Y-axis).

a series of regression models trying to account for these and other methodological issues. We have two main aims with our analysis: First, we seek a systematic and robust description of the *associations* between social-group composition of opposition movements and democratization. Second, we want to assess hypotheses on the effects of campaigns on democratization. While often implicit, the notion of such effects underpin the reviewed literature linking social groups to democracy. Democratization results *because* specific social groups mobilize.

Yet, as with other macro-level large-n studies, identifying effects with a fair degree of certainty is challenging. Here, endogeneity bias can result if particular social groups self-select into particular movements; if certain groups have a stronger *preference* for democracy, this may lead them to join movements that are more likely to engender democratization (with or without their participation). There are also several plausible confounders. For example, if movement size affects outcomes (e.g., Chenoweth and Stephan, 2011), this may yield biased estimates if unaccounted for. On the flip-side, controlling for movement size could introduce post-treatment bias; as we have argued, differential capacities between social groups in organizing large-scale collective action constitute an inherent property of likely causal relevance. This suggests that there is no single model that can handle all potential issues, and we employ a series of different specifications that deal with different alternative explanations and potential biases.

17

Cross-sectional analysis at the campaign level

We first investigate a simple cross-section of opposition campaigns using OLS. Models 1–4, Table 2, present results pertaining to the broader question of whether campaigns dominated by (*either* of the two) urban groups differ from other campaigns. Model 1 employs a dummy of democratization – scored 1 if the country was coded autocratic by BMR at campaign-start and democratic one year after campaign-end – and only includes observations that were autocratic at campaign-start (143).[11] The coefficient for urban group-dominated campaign (UG) in Model 1 suggests that the probability of observing post-campaign democratization is roughly 19 percentage points higher for campaigns dominated by one of the urban groups than for other campaigns – the overall probability of post-campaign democratization in the sample is 27 percent.

This association could, however, be due to urban group-dominated campaigns occurring more frequently in, for instance, wealthy countries. We thus condition on potentially relevant country- and campaign-level covariates. At the country-level, we control for ln population, ln (real, PPP-adjusted) GDP per capita, and urbanization, all registered at campaign outset and collected from V-Dem (see Coppedge et al., 2016a). At the campaign-level, we control for ln participants. Urban campaigns may enhance democratization simply because they are larger. (But, we remind that campaign size can be considered a relevant post-treatment feature, which is why we drop it in our main models). UG drops – and loses statistical significance, with $t = 1.4$ – when conditioning on these covariates in Model 2, although urban campaigns are predicted to be followed by a 12 percentage point higher probability of democratization.

We also tested the more fine-grained Polyarchy measure. In order to consider correlates of changes in democracy, we condition on the lagged dependent variable (i.e., Polyarchy level at campaign start). Model 3 includes 180 campaigns, and reveals a highly significant UG ($t = 2.7$). The t-value decreases to 1.5 when conditioning on the campaign- and country-level

---

[11]In our main country-year analysis, we employ logit models on this dependent variable, but since the purpose here is description we report OLS models for ease of interpretation. Results, in terms of statistical significance, are very similar for logit specifications (Table A.10).

Table 2: Campaign-level (cross sectional) correlates of democratization.

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	+Δ BMR		Polyarchy		+Δ BMR		Polyarchy	
Ind work OR mid cl dominate (UG)	0.192***	0.108	0.078***	0.047				
	(2.76)	(1.38)	(2.68)	(1.46)				
Industrial workers dominate (IW)					0.189**	0.153	0.126***	0.111***
					(2.17)	(1.65)	(3.38)	(2.77)
Middle classes dominate (MC)					0.144*	0.078	0.023	0.002
					(1.78)	(0.91)	(0.71)	(0.06)
Ln campaign participants		-0.000		-0.000		0.000		0.000
		(-0.15)		(-0.08)		(0.02)		(0.17)
Urbanization		0.463		0.159		0.476		0.159
		(1.60)		(1.36)		(1.66)		(1.39)
Ln population		-0.013		-0.004		-0.009		-0.003
		(-0.54)		(-0.42)		(-0.37)		(-0.31)
Ln GDP per capita		0.047		0.031		0.039		0.028
		(0.70)		(1.10)		(0.58)		(1.03)
Lagged dependent variable			0.647***	0.481***			0.661***	0.509***
			(7.94)	(5.11)			(8.21)	(5.45)
N	143	124	180	158	143	124	180	158
R^2	0.051	0.130	0.286	0.297	0.071	0.149	0.313	0.324

Notes: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. T-values in parentheses. OLS regressions with campaign as cross-section unit.

covariates in Model 4. Yet the results are strengthened in terms of coefficient sizes and t-values when we allow for some effect-lag and measure democratization up to 5 years after campaign end (see Tables A.9 and A.10).

The four rightmost columns in Table 2 display structurally similar models, but where we exchange UG with two dummy variables distinguishing industrial worker- (IW) from urban middle class-dominated (MC) campaigns. Results are quite strong on the association between IW and democratization, especially when employing Polyarchy. MC coefficients are comparatively smaller, and t-values range from 0.1–1.8. This provides an initial hint that industrial worker-dominated campaigns are more clearly associated with democratization middle class-dominated campaigns.

Cross-section time-series analysis

While indicative, campaign-level cross-section models are not fully adequate for answering the question of whether campaign social profile matters for democratization. First, they do not include information on observations without campaigns, disallowing comparisons between campaigns with a given social profile and situations where no campaign exists. Second, including information on covariates from additional years will expectedly allow for more efficient comparisons also between campaigns. Third, controlling for country- and year-fixed effects is

not feasible in the cross-section (campaign) set-up. There may well be country-specific features, related e.g. to geography or political culture, that correlate with the motivations and capacities of different groups to mobilize opposition *and* with democratization. There may also be global trends both in democratization and campaign profiles, for instance with increased chances of democratization and more urban opposition campaigns right after the Cold War. Therefore, we move to a country-year setup, and estimate models of the form:

$$DEM_{i,t+1} = \beta_0 + \beta_1 IW_{i,t} + \beta_2 MC_{i,t} + \beta_3 OC_{i,t} + \beta_4 DEM_{i,t} + \mu \mathbf{X}_{i,t} + \zeta_i + \theta_t + \epsilon_{i,t} \tag{1}$$

$DEM_{i,t+1}$ represents democracy led by one year $(t+1)$ in country i.[12] IW and MC register the presence of ongoing campaigns dominated by, respectively, industrial workers and urban middle classes. Alternatively, we substitute these two dummies with the combined UG dummy. OC registers presence of other campaigns, i.e. those dominated by other groups such as peasants or where no group clearly dominates. The models further include level of democracy at t ($DEM_{i,t}$), and vectors of country-year covariates $\mathbf{X}_{i,t}$, country-fixed effects ζ_i, and year-fixed effects θ_t. Errors are clustered by country to account for panel-level autocorrelation.

Table A.7 presents two versions of our benchmark fixed-effects (FE) OLS model on Polyarchy, including about 9000 observations from more than 140 countries and maximum time series from 1900–2006. Model 1 employs the cruder UG dummy, and Model 2 uses IW and MC. The inclusion of OC implies that the "reference category" is no ongoing opposition campaign. According to Model 1, an ongoing campaign dominated by one of the two urban groups systematically increases Polyarchy in the following year by 0.04 compared to when there is no ongoing opposition campaign, and UG is highly significant ($t = 3.06$). Since OC is negative, opposition campaigns dominated by urban groups are systematically associated with democratic change also relative to situations with other ongoing campaigns. In Model 2, the IW coefficient is significant at 1% and virtually similar (0.54 instead of 0.40) to UG in Model 1, whereas MC is smaller (0.23) and insignificant at conventional levels ($t = 1.27$).

---

[12]We also test and discuss models where change in democracy from t to $t+1$ is dependent variable, including models that only count positive/democratizing changes.

Table 3: Core cross-section times series models

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimator	FE OLS	FE OLS	FE OLS	FE OLS	FE Logit	Logit	Logit
Dependent variable	V-Dem Polyarchy		Δ Polyarchy	$+\Delta$ Polyarchy	BMR democratization		
Ind. wor. OR mid. cl. dominate (UG)	0.040***				0.659	1.305***	
	(3.06)				(1.07)	(2.62)	
Industrial workers dominate (IW)		0.054***	0.054***	0.047***			1.085**
		(2.84)	(2.84)	(2.82)			(2.53)
Middle classes dominate (MC)		0.023	0.023	0.029*			1.289***
		(1.27)	(1.27)	(1.90)			(3.16)
Other campaign (OC)	-0.002	-0.003	-0.003	0.006**	0.893*	0.789**	0.736**
	(-0.57)	(-0.84)	(-0.84)	(2.03)	(1.80)	(2.10)	(2.09)
Ln GDP per capita	0.001	0.001	0.001	-0.002	1.092**	0.523**	0.396*
	(0.38)	(0.30)	(0.30)	(-0.55)	(2.02)	(2.27)	(1.74)
Ln population	-0.001	-0.001	-0.001	0.003	-0.810	-0.027	0.002
	(-0.58)	(-0.46)	(-0.46)	(1.05)	(-0.82)	(-0.31)	(0.03)
Urbanization	0.025**	0.024**	0.024**	0.023**	-5.048	-0.316	0.228
	(2.23)	(2.22)	(2.22)	(2.10)	(-1.58)	(-0.36)	(0.27)
Lagged Polyarchy	0.949***	0.950***	-0.050***	-0.035***			
	(169.46)	(172.86)	(-9.18)	(-7.52)			
Country dummies	Y	Y	Y	Y	N	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y
N	9099	9099	9099	9099	2769	3153	4705
R^2	0.939	0.939	0.087	0.105			
Pseudo R^2					.307	0.091	0.061
Countries	147	147	147	147	68	121	121

Notes: *$p<0.1$; **$p<0.05$; ***$p<0.01$. T-values in parentheses. Except for in Model 5, errors are clustered by country.

Hence, there is evidence for a positive association between industrial worker campaigns and democratization, but no clear evidence for middle class campaigns.

Results are similar if we rather employ changes in Polyarchy score (from t to $t+1$) directly as dependent variable (both when conditioning and not conditioning on level of Polyarchy; see Table A.11). Model 3 is otherwise similar to the benchmark specification, and IW and MC are identical in size and significance to in Model 2. We also run models considering relationships with "democratic upturns" and "downturns" separately (see Teorell, 2010). In the theoretical discussion we mainly focused on how social groups affect democratization, and not so much how they guard against democratic backsliding. Tables A.12 and A.19 show that results are robust and even clearer (for UG, IW, and MC) when considering only democratic upturns. This is illustrated by Model 4, which is a version of the benchmark specification (downturns are here scored as 0 on the dependent variable). IW remains highly significant, whereas UM now turns weakly significant with $t = 1.90$. We find some indications that ongoing industrial worker campaigns, though not middle class campaigns, also guard against subsequent democratic backsliding, though coefficients are smaller and only weakly significant in some specifications.

Models 5–7 are logit equivalents of our benchmark, using the dichotomous BMR measure of democratization as dependent variable and only including autocratic observations in t. This reduces variation in the dependent variable; only 102 democratization episodes are included for Model 5, and we now ignore any impact opposition campaigns may have on de-democratization. We also ignore situations where institutions change in a more "liberal direction" without tilting regimes across the "minimum democracy" threshold (countries with as different regimes as North Korea and Taiwan were BMR autocracies in 1995). Similarly, we ignore any impact campaigns may have in increasing level of democracy in regimes above the minimum threshold (in 1995, Colombia, Mali, and Pakistan had all passed this threshold).[13] Thus, these logit models may fail to pick up effects, even if they exist. Indeed, the fixed effects logit model (Model 5) fails to find a statistically significant UG at conventional levels (t=1.07), although UG is positive and sizeable. However, the issues discussed above are exacerbated by including country-fixed effects. As shown by Beck and Katz (2001), fixed effects logit models often yield Type II errors when the dependent variable changes seldomly or not at all (only 68 countries inform estimates in Model 5). When omitting the country-fixed effects, but keeping the year-fixed effects, in Model 6, the relationship between UG and BMR democratization retains statistical significance at 1%. Model 7 alters Model 6 by substituting UG with IW and MC, and both turn out statistically significant at 5%. While results do not survive the inclusion of country-fixed effects, there is thus some evidence that opposition campaigns dominated by industrial workers or urban middle classes are related to democratic transitions, as measured by BMR.

Still, the associations in Table A.7 could, potentially, be influenced by particular model specification- or measurement choices, or that opposition campaign profile is endogenous to anticipated regime changes. We thus turn to alternative tests. Table 5 presents a selection of

---

[13]Indeed, additional estimates using V-Dem's Polyarchy as dependent variable suggest that industrial worker campaigns are associated with liberalization in BMR autocracies *and* further "democratic deepening" in BMR democracies. In contrast, middle class campaigns are positively related to liberalization in autocracies, but negatively related to democratic deepening in democracies. See Tables A.14–A.17.

these tests, and several others are reported in Appendices E–H. We employ our main measure of (electoral) democracy, Polyarchy, which ensures sufficient variation in the dependent variable. We note at the outset that results for the combined UG dummy are robust (see Table A.18). Hence, we mostly focus on the two separate industrial worker and middle class dummies, allowing us to discuss the more nuanced propositions concerning specific groups and democratization.

First, results are not weakened when measuring the dependent variable with a longer lag (e.g., 5 or 10 years) than the baseline 1-year lag. Model 1 in Table 5 employs a 2-year lag, and the IW point estimate increases considerably, from 0.054 to 0.090, suggesting that a 1-year lag may be too short to capture the full effect of opposition campaigns on regime change. As in the baseline, IW is significant at 1%, whereas MC is positive but statistically insignificant at conventional levels. *Second*, we note that the weakly significant result for MC comes despite a substantially high point estimate. Our benchmark includes country- and year-fixed effects, and employs clustered errors. When combined with relatively few campaign observations, even in our extensive sample, this yields large standard errors. Model 2 employs classical rather than clustered errors, and MC is now highly significant with $t = 4.31$. The t-value of IW also increases dramatically, from 2.84 to 9.02. Yet, in the remaining models, we employ the more conservative benchmark approach, clustering on country to, e.g., account for autocorrelation affecting errors.

Third, our benchmark results could be afflicted by post-treatment bias. We thus tested models omitting potentially "bad controls". Controlling for urbanization, in particular, could wash out parts of the theoretically relevant effect. As discussed, the capacity of industrial workers to bring about democratic change is partly attributed to workers living closely together in cities, easing coordination problems. However, more parsimonious models consistently show results similar to the benchmark. This is exemplified by Model 3, which excludes population, income, and urbanization. *Fourth*, the opposite concern is that we are not including all relevant controls, and that results are afflicted by omitted variable bias. In particular, other aspects of opposition movements than *who* participates may drive democratization. Hence, Model 4 controls also

23

Table 4: Robustness tests

Model	(1) 2-year lag	(2) No clustering	(3) Parsimonious	(4) Extra controls	(5) Public empl.	(6) Nr groups	(7) No uncertain	(8) Camp. origin	(9) Lexical ind.
Ind. workers dominate (IW)	0.090***	0.054***	0.054***	0.053***	0.050***	0.043**	0.054***		
	(3.09)	(9.02)	(2.94)	(2.99)	(2.74)	(2.22)	(2.63)		
Middle classes dominate (MC)	0.033	0.023***	0.025	0.023	0.030**	0.026	0.032		
	(1.15)	(4.31)	(1.43)	(1.23)	(2.20)	(1.39)	(1.60)		
Industrial worker origin								0.025	
								(1.22)	
Middle class origin								0.031***	
								(2.79)	
Industrial worker index									0.017**
									(2.56)
Middle class index									0.016***
									(3.25)
Other campaign	-0.004	-0.003	-0.003	-0.016	-0.003	-0.043***	-0.004	-0.001	-0.021***
	(-0.81)	(-1.27)	(-0.94)	(-1.04)	(-0.82)	(-3.20)	(-0.95)	(-0.25)	(-3.59)
Ln GDP p.c.	0.002	0.001		0.001	0.001	0.001	0.001	0.000	0.001
	(0.36)	(0.42)		(0.30)	(0.34)	(0.34)	(0.21)	(0.20)	(0.23)
Ln Population	-0.004	-0.001		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.79)	(-0.46)		(-0.34)	(-0.46)	(-0.26)	(-0.30)	(-0.59)	(-0.37)
Urbanization	0.054**	0.024**		0.025**	0.025**	0.026**	0.026**	0.025**	0.025**
	(2.36)	(2.53)		(2.26)	(2.23)	(2.32)	(2.28)	(2.28)	(2.21)
Ln campaign participants				-0.000					
				(-0.54)					
Regime violence				0.019					
				(1.28)					
Nr of groups participating						0.013***			
						(3.14)			
Lagged Polyarchy	0.873***	0.950***	0.956***	0.950***	0.949***	0.950***	0.949***	0.948***	0.950***
	(71.79)	(261.81)	(214.00)	(170.51)	(163.54)	(161.27)	(164.93)	(166.55)	(164.14)
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	9240	9244	14407	9244	9234	9124	9142	9240	9124

Notes: * $p<0.1$; ** $p<0.05$; *** $p<0.01$. T-values in parentheses. Polyarchy in $t+1$ is dependent variable.
All models are OLS with country- and year fixed effects. Except for in Model 2, errors are clustered by country.

24

for campaign size (log thousand participants) from NAVCO. We discussed how including campaign size could yield post-treatment bias, but this could also be a confounder since the social profile of a campaign could be endogenous to how many people participate; if the whole population participates, a majority social group will *by definition* dominate. Model 4 further includes whether the regime uses violence against protestors, since particular groups might "self-select" into campaigns opposing less violent regimes. Adding these controls does not alter the estimated relationships between IW/MC and democracy; both remain positive, but only IW is clearly separable from zero. The control for regime violence speaks to a more general concern about the emergence, composition, and success of campaigns being moderated by interactions with the extant regime. In Appendix F, we discuss this further and conceptualize the conditional relationship between regime strategies and composition of opposition movements as an omitted variable problem. We address this issue by estimating the Controlled Direct Effect (CDE) (following Acharaya, Blackwell and Sen, 2017) of opposition characteristics when holding regime-strategies fixed at given levels. This analysis gives quite similar results to the benchmark, suggesting that our results are not driven by differential regime strategies towards opposition movements.

Fifth, our baseline measure of MC campaigns does not include campaigns dominated by public sector employees. While counted as part of the "middle classes" in many categorization schemes, we excluded public employees since we expect them to often have different interests and leverage over regime change, given their special relationship with the state. The number of public employee-dominated campaigns is about $1/3$ of MC campaigns narrowly defined. When employing the more extensive operationalization – counting public sector employees as urban middle class – the coefficient increases and turns significant at 5%. In other words, the evidence of middle class campaigns engendering democratization is stronger if we include public sector employees. IW remains highly significant also in this model. *Sixth*, some groups may self-select into campaigns that are more likely to induce democratization because they are large-coalition campaigns. Broad coalitions of social groups may have more leverage, and thus better prospects for forcing the regime to liberalize. Model 6 controls for number of social

groups participating in the campaign. This variable is positive, sizeable, and highly significant. Including this count variable does not qualitatively alter the benchmark results, however, as MC remains positive and statistically insignificant ($t = 1.60$) whereas IW remains sizeable (0.43) and significant at 5%.

Seventh, we want to check whether measurement errors in our coding affect results. Some regions and historical periods may be associated with fewer – or conflicting – sources, and Model 5 drops all 92 country-years where our coders reported low certainty. Results are not changed by much when doing so. To assess this issue further, we estimated Errors-in-Variables models (Tables A.7–A.8), informed by measures of agreement from our inter-coder reliability tests. These models show that when we explicitly incorporate measurement error our results are strengthened, as one might expect from econometric theory.[14] Even MC is now statistically significant at conventional levels. Hence, the lack of significance for urban middle class campaigns in the benchmark *might* result from attenuation bias due to noisy measurement. *Eighth*, we tested alternative ways of measuring campaign profile. Model 8 changes the independent variables from social group *dominance* to whether movements *originated* among the given group(s). MC is substantially strengthened in terms of statistical significance ($t = 2.79$), whereas IW now turns insignificant ($t = 1.22$). However, IW is stronger in yet other specifications that alter how the independent variables are constructed: When measuring whether a group at least *participated*, we mostly obtain strong results, especially for industrial workers. In addition, we tested lexical scales (MC Index; IW index), scored 3 if the given group dominated the movement, 2 if it originated from the group and the group participated, 1 if the group only participated, and 0 if the group did not have a recorded presence. These information-rich measures of level of involvement give strong results for both industrial workers and the middle classes. In Model 9, the IW Index is significant at 5% and the MC Index at 1%.

Although controlling for, e.g., income level and time- and country-specific effects should mitigate many sources of bias, other sources could remain. We therefore estimate instrumental

---

[14]Whereas predictions are less clear in multivariate regressions, unsystematic error in the regressor always induces attenuation bias in bivariate regressions.

variable (IV) models where we endogenize the IW and MC profiles of the campaigns. An extra benefit of running IV models is that they mitigate attenuation bias from measurement errors in the endogenous regressors, which, as discussed, could artificially weaken the estimated relationships. We discuss and assess the validity of our instruments, which draw on the social profile of opposition campaigns in other countries in the geographic region, and present results from several 2SLS specifications in Appendix G. In brief, these results suggest that not only IW campaigns, but also MC campaigns, have a substantial positive effect of democratization.

Contextual factors

We finally assess whether results are sensitive to the time period under study, geographic region, and whether societies are urbanized or not. As we will clarify, some of these regressions constitute tests of additional implications from our argument rather than robustness tests. When investigating particular time periods or omitting key regions, the number of campaigns dominated by one group can become quite low. We thus mostly focus on the "combined" UG dummy (but see Table A.22 for tests on IW and MC separately), and use our benchmark model. We first test whether the relationship between UG and democratization is sensitive to time period under study. Given the fairly general nature of our argument, we would be concerned if we only identify the relationship in some parts of recent modern history. But, as argued above, most countries globally (except for some Western countries) have observed processes of industrialization and urbanization mainly after the mid-20th century. Hence, the capacities and motivations of urban actors to pursue democratization have expectedly increased in many countries in more recent decades. Indeed, when splitting the 1900–2010 time-period in two, UG is almost three times larger post-1955 (Model 2) than pre-1955 (Model 1). Still, also the latter result is (weakly) significant, despite including only 2728 country-year observations, which is in line with previous case-study evidence for urban middle class or industrial worker movements in democratization coming from both historical and more contemporary cases (e.g., Collier, 1999).

We would also be concerned if our results basically stem from developments in one particular geographic region. However, we do not expect this: Although many early studies on urban

Table 5: Sensitivity and scope: Split-sample tests and interaction models

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	Time		Omitting particular geographic region						Urban and rural societies			
Sample	≤ 1956	> 1955	$\div WE/NA$	$\div EE$	$\div SSA$	$\div ME/NA$	$\div A/P$	$\div LA$	Rural	Urban	Full	Full
IW or MC dominate	0.017*	0.049***	0.031**	0.029**	0.037**	0.040***	0.044***	0.065***	0.021**	0.060**	-0.013	
	(1.71)	(2.96)	(2.50)	(2.14)	(2.34)	(2.95)	(3.05)	(4.14)	(2.05)	(2.43)	(-0.47)	
IW/MC*Urbanization											0.130*	
											(1.86)	
MC dominate												-0.004
												(-0.07)
IW dominate												-0.021
												(-0.56)
MC*Urbanization												0.093
												(0.60)
IW*Urbanization												0.161*
												(1.72)
Other campaign	-0.012	-0.002	-0.000	-0.002	-0.002	-0.002	-0.003	-0.002	-0.003	-0.006	-0.003	-0.004
	(-1.46)	(-0.49)	(-0.10)	(-0.64)	(-0.46)	(-0.61)	(-0.76)	(-0.66)	(-0.71)	(-0.78)	(-0.52)	(-0.63)
OC*Urbanization											0.004	0.003
											(0.20)	(0.16)
Ln GDP p.c.	-0.014*	0.001	0.003	0.001	0.003	0.003	0.002	0.001	-0.000	0.003	0.002	0.002
	(-1.87)	(0.45)	(1.00)	(0.59)	(1.12)	(0.94)	(0.57)	(0.25)	(-0.14)	(0.82)	(0.91)	(0.87)
Ln Population	-0.004	-0.007	-0.003	0.000	0.000	0.001	-0.001	-0.003	-0.001	-0.009*	-0.000	-0.000
	(-0.36)	(-1.58)	(-0.92)	(0.16)	(0.04)	(0.54)	(-0.34)	(-1.27)	(-0.15)	(-1.86)	(-0.21)	(-0.15)
Urbanization	-0.103***	0.034**	0.016	0.028**	0.027**	0.026**	0.012	0.021*	0.029	0.061**	0.018*	0.018*
	(-2.92)	(1.99)	(1.29)	(2.41)	(2.37)	(2.33)	(1.10)	(1.95)	(1.35)	(2.54)	(1.72)	(1.70)
Lagged Polyarchy	0.940***	0.936***	0.943***	0.943***	0.953***	0.948***	0.948***	0.951***	0.917***	0.949***	0.949***	0.950***
	(74.99)	(130.47)	(139.77)	(154.41)	(153.06)	(157.89)	(149.21)	(158.07)	(76.68)	(118.07)	(159.28)	(169.37)
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	2728	7011	7635	9155	7246	8901	7974	7784	4870	4870	9739	9738

Notes: *$p<0.1$; **$p<0.05$; ***$p<0.01$. T-values in parentheses. Polyarchy in $t + 1$ is dependent variable.
All models are OLS with country- and year fixed effects. Errors are clustered by country.

28

groups and democratic transitions focused on Europe or Latin America, several recent studies on third-wave democratization experiences have focused on Eastern European (McFaul, 2002) and Sub-Saharan African countries (Bratton and van de Walle, 1997). Even more recently, the urban middle-class composition of the movement orchestrating the Tunisian revolution in 2009 is one proposed reason why this revolution induced democratization, while other Arab Spring revolutions culminated in new autocracies (Beissinger, Jamal and Mazur, 2015). Models 3–8 reveal that the relationship is fairly stable and always significant at 5% when omitting major world regions. UG is lowest when omitting Eastern Europe (.029; Model 4) and highest when omitting Asia-Pacific (.065; Model 8).

Finally, we expected that industrial workers and urban middle class actors have clearer motivation and stronger capacity to induce democratization in more urban societies. Models 9 (rural) and 10 (urban) split the sample at median level of urbanization (0.41). Model 10 reveals a highly significant coefficient almost four times the size of the coefficient in Model 9. However, UG is still fairly sizeable (.021) and the t-value is 2.05 even in Model 9. Further, we added multiplicative terms (urbanization multiplied with, respectively, urban and other campaigns) to the benchmark. This model (11) predicts that the interaction is sizeable, although not very precisely estimated ($t = 1.86$). The point estimates actually suggest a negative (but insignificant) relationship between urban groups and democratization when less than about 10 percent of the population lives in cities. Model 12 nuances this result by once again separating industrial worker from middle class campaigns. The interaction term is larger in size (.16 vs .09) and t-value (1.72 vs 0.60) for IW than for MC. While these interaction terms are not clearly distinguishable from each other, this (mildly) suggests that the overall relationship observed in Model 11 is largely driven by industrial workers becoming increasingly effective agents of democratization as urbanization proceeds.

Conclusion

Various scholars have proposed that urbanization and industrialization empowers industrial workers and/or the urban middle classes, and that these groups, in turn, act as agents of democ-

ratization. In this paper, we have elaborated on why these groups may be motivated to pursue democratization, particularly in urban societies, and why they often have the requisite capacity to succeed. We also provide large-n empirical evidence speaking directly to this issue. To this end, we collected data on the social groups partaking in organized opposition campaigns throughout the world from 1900–2006.

Our main findings can be summarized as follows: Social movements dominated by these urban groups are linked to democratization. This goes both when we compare to opposition movements dominated by other groups and to situations where no organized opposition movement exists. While apparent in different contexts, the relationship seems stronger in more recent decades and in urbanized societies. When making the distinction between opposition movements dominated by the urban middle classes and by industrial workers, we find much more robust support for the democratizing role of industrial worker movements. While statistical models accounting explicitly for measurement errors and the endogenous composition of opposition campaigns suggest that also urban middle class movements engender democratization, our results, overall, thus corroborate the accounts presented by scholars, such as Rueschemeyer, Stephens and Stephens (1992) and Collier (1999), who propose that industrial workers are key agents of democratization.

Our findings inform different, though partly overlapping, literatures. First, we provide a novel empirical take on the social origins of democratic revolutions, thus contributing to long-standing debates in comparative-historical political sociology (Moore, 1966; Luebbert, 1991; Rueschemeyer, Stephens and Stephens, 1992). Second, our analysis informs the recent literature on the origins of dictatorship and democracy (Acemoglu and Robinson, 2006; Boix, 2003; Ansell and Samuels, 2014), which is, at least in part, motivated by the aforementioned comparative-historical literature. Finally, we expand on the recent large-n literature on opposition movements, starting with Chenoweth and Stephan (2011), by linking the outcomes of these movements to their social composition.

References

Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Acharaya, Avidit, Matthew Blackwell and Maya Sen. 2016. "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110(3):512–529.

Aidt, Toke S. and Peter S. Jensen. 2014. "Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 1820–1938." *European Economic Review* 72(November):52–75.

Ansell, Ben W and David J Samuels. 2014. *Inequality and Democratization*. Cambridge: Cambridge University Press.

Ardanaz, Martin and Isabela Mares. 2013. "Labor Shortages, Rural Inequality, and Democratization." *Comparative Political Studies* 47(12): 1639-1669.

Bayer, Markus, Felix S Bethke and Daniel Lambach. 2016. "The democratic dividend of nonviolent resistance." *Journal of Peace Research* 53(6):758–771.

Beck, Nathaniel and Jonathan N. Katz. 2001. "Throwing out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon." *International Organization* 55(2):187–195.

Beissinger, Mark R, Amaney A. Jamal and Kevin Mazur. 2015. "Explaining Divergent Revolutionary Coalitions: Regime Strategies and the Structuring of Participation in the Tunisian and Egyptian Revolutions." *Comparative Politics* 48(1):1–24.

Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge: Cambridge University Press.

Boix, Carles, Michael Miller and Sebastian Rosato. 2013. "A complete data set of political regimes, 1800–2007." *Comparative Political Studies* 46(12):1523–1554.

Bratton, Michael and Nicholas van de Walle. 1997. *Democratic Experiments in Africa: Regime Transitions in Comparative Perspective*. Cambridge: Cambridge University Press.

Bunce, Valerie J. and Sharon L. Wolchik. 2010. "Defeating Dictators: Electoral Change and Stability in Competitive Authoritarian Regimes." *World Politics* 62(1):43–86.

Butcher, Charles and Isak Svensson. 2014. "Manufacturing Dissent Modernization and the Onset of Major Nonviolent Resistance Campaigns." *Journal of Conflict Resolution.* 60(2):311–339.

Casper, Brett Allen and Scott A Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'état." *Journal of Politics* 76(2):548–564.

Celestino Rivera, Mauricio and Kristian Skrede Gleditsch. 2013. "Fresh carnations or all thorn, no rose? Nonviolent campaigns and transitions in autocracies." *Journal of Peace Research* 50(3):385–400.

Chenoweth, Erica and Jay Ulfelder. 2015. "Can Structural Conditions Explain the Onset of Nonviolent Uprisings?" *Journal of Conflict Resolution* 61(2):298–324.

Chenoweth, Erica and Maria J. Stephan. 2011. *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict.* New York: Columbia University Press.

Chenoweth, Erica and Orion A. Lewis. 2013. "Unpacking nonviolent campaigns. Introducing the NAVCO 2.0 dataset." *Journal of Peace Research* 50(3):415–423.

Chun, Natalie, Rana Hasan, Muhammad Habibur Rahman and Mehmet A Ulubaşoğlu. 2015. "The role of middle class in democratic diffusion." *International Review of Economics & Finance* 42(March):536–548.

Collier, Ruth Berins. 1999. *Paths Towards Democracy: The Working Class and Elites in Western Europe and South America.* Cambridge: Cambridge University Press.

Coppedge, Michael, et al. 2016. "V-Dem Country-Year Dataset v6.1." Varieties of Democracy Project.
   URL: *https://www.v-dem.net/en/data/data-version-6-1/*

Coppedge, Michael, et al. 2016. "V-Dem Methodology v6." Varieties of Democracy Project.
   URL: *https://www.v-dem.net/files/42/Methodology%20v6.pdf*

Dahl, Robert A. 1971. *Polyarchy: Political Participation and Opposition.* New Haven: Yale University Press.

Glaeser, Edward L, Giacomo AM Ponzetto and Andrei Shleifer. 2007. "Why does democracy need education?" *Journal of economic growth* 12(2):77–99.

Goldstone, Jack A. 2011. "Cross-class Coalitions and the Making of the Arab Revolts of 2011." *Swiss Political Science Review* 17(4):457–462.

Hobsbawm, Eric J. 1974. "Labor history and ideology." *Journal of Social History* 7(4):371–381.

Houle, Christian. 2009. "Inequality and Democracy: Why Inequality Harms Consolidation but Does Not Affect Democratization." *World Politics* 61(4):589–622.

Inglehart, Ronald and Christian Welzel. 2006. *Modernization, Cultural Change and Democracy – The Human Development Sequence*. Cambridge: Cambridge University Press.

Kendall-Taylor, Andrea and Erica Frantz. 2014. "How Autocracies Fall." *The Washington Quarterly* 37(1):35–47.

Kimeldorf, Howard. 2013. "Worker Replacement Costs and Unionization Origins of the US Labor Movement." *American Sociological Review* 78(6):1033–1062.

Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53(1):69–105.

Luebbert, Gregory M. 1991. *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe*. Oxford: Oxford University Press.

Marx, Karl and Frederick Engels. 2010/1848. *Manifesto of the Communist Party*. Marxists Internet Archive.

URL: *http://www.marxists.org/archive/marx/works/download/pdf/Manifesto.pdf*

McFaul, Michael. 2002. "The fourth wave of democracy and dictatorship: noncooperative transitions in the postcommunist world." *World politics* 54(2):212–244.

Moore, Barrington. 1966. *Social Origins of Democracy and Dictatorship: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Munck, Gerardo L. and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35(1):5–34.

Ness, Immanuel. 2015. "The International Encyclopedia of Revolution and Protest".

URL: *http://www.revolutionprotestencyclopedia.com*

O'Donnell, Guillermo. 1973. *Modernization and Bureaucratic-Authorianism: Studies in South*

American Politics. Berkeley: Institute of International Studies, University of California.

O'Donnell, Guillermo and Phillipe C. Schmitter. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions About Uncertain Democracy*. Baltimore: Johns Hopkins University Press.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Powell, Jonathan. 2012. "Determinants of the Attempting and Outcome of Coup." *Journal of Conflict Resolution* 56(6):1017–1040.

Przeworski, Adam. 2009. "Conquered or granted? A history of suffrage extensions." *British Journal of Political Science* 39(2):291–321.

Rokkan, Stein. 1999. *State formation, nation-building, and mass politics in Europe*. Wotton-under-Edge: Clarendon Press.

Rueschemeyer, Dietrich, Evelyne Huber Stephens and John D. Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.

Satterthwaite, David. 2007. *The transition to a predominantly urban world and its underpinnings*. Human Settlement Discussion Series.
URL: *http://pubs.iied.org/pdfs/10550IIED.pdf*

Stepan, Alfred. 2012. "Tunisia's transition and the twin tolerations." *Journal of Democracy* 23(2):89–103.

Stephens, John D. 1989. "Democratic Transition and Breakdown in Western Europe, 1870-1939: A Test of the Moore Thesis." *American Journal of Sociology* 94(5):1019–1077.

Stock, James H and Motohiro Yogo. 2005. "Testing for weak instruments in linear IV regression." In *Identification and inference for econometric models*, pp. 80–108. New York: Cambridge University Press.

Swarthmore. 2015. "Global Nonviolent Action Database".
URL: *http://nvdatabase.swarthmore.edu*

Teorell, Jan. 2010. *Determinants of Democracy: Explaining Regime Change in the World, 1972–2006*. Cambridge: Cambridge University Press.

Teorell, Jan, Michael Coppedge, Svend-Erik Skaaning and Staffan I. Lindberg. 2016. "Measuring Electoral Democracy with V-Dem Data: Introducing a New Polyarchy Index." Gothenburg University: V-Dem Working Paper 2016:25.

   URL: *https://www.v-dem.net/media/filer_public/b7/1f/b71f18e0-852e-4e52-adc4-9923f7baaac6/v-dem_working_paper_2016_25_edited.pdf*

Weyland, Kurt. 2009. "The diffusion of revolution:'1848'in Europe and Latin America." *International Organization* 63(3):391–423.

Weyland, Kurt. 2014. *Making Waves: Democratic Contention in Europe and Latin America since the Revolutions of 1848*. Cambridge: Cambridge University Press.

Appendices

The Appendices below have the following contents: Appendix A presents and discusses the coding rules for the data collection on the social composition of opposition campaigns. Appendix B presents examples of the brief summary reports that accompany the coding for two campaigns. Appendix C presents a more comprehensive table with descriptive statistics than presented in the paper, including all key variables. Appendix D discusses and presents results on the inter-coder reliability tests, and displays the benchmark regression results when we use the coding conducted by the replication coder rather than the original coder for the double-coded cases. Appendix E presents results from Errors-in-variables regressions that draw on information from the inter-coder reliability tests to estimate the amount of measurement error in our core variables. Appendix F presents results from a procedure (estimating Controlled Direct Effects using sequential-G estimation) that accounts for the success of various opposition campaign being conditioned by various features of the regime in place. Appendix G presents and discusses our instrumental variable results, using the opposition campaign profile in other countries in the region to instrument for our core independent variables. Finally, Appendix F displays several additional tests and extensions briefly noted, but not reported, in the paper. The sequence of, and rationale behind, these tests are further specified at the beginning of Appendix G.

A   Coding scheme for indicators of protest movements social composition

We here present the coding scheme and guidelines for the variables recording the social background of participants in protest campaigns. For each campaign, we always consulted several different sources, including databases and encyclopedias such as "The international encyclopedia of revolution and protest" (Ness, 2015) and the "Global Non-Violent Action Database" (Swarthmore, 2015), various reports by international or non-governmental organizations such as UNHCR or Freedom House, and books and journal articles that discuss the relevant campaigns, examples being Ackerman and DuVall (2001) and Nepstad (2011). We coded each campaign based on the participants that were reported to have been a part of the campaign from the start-year listed in NAVCO 1.0 to the end-year.

As noted in the paper, the sources are usually relatively clear as to whether members of different social groups participated in movements at all, as there is usually some kind of description of the protest participants. Likewise, there is usually some description of where and amongst whom the campaign originated. When it comes to the variable recording whether a particular social group dominated the campaign, we relied on a coding rule requiring that social group members should either make up at least a majority of the campaign members or they should have had a critical impact on the strategies and outcomes of the campaign, following a high threshold for judging impact as critical. We note that the fact that one RA conducted all the original coding ensures some level of cross-country and inter-temporal consistency in terms of applying the latter subjective criterion, but we also note that the inter-coder reliability tests discussed in the paper and later in this appendix showed remarkable consistency on the coding of the "dominance" variables.

The coding scheme distinguished between the following social groups:

- **Peasants**, referring to both farmers and rural workers.

- **Industrial workers**, referring to labor employed in the industrial sector.

- **Public sector employees**, referring to employees in the public sector such as bureaucrats and teachers

- **Urban middle classes**, including business elites, smaller merchants, professionals (including e.g. lawyers, doctors, etc.), students, and university professors, but excluding public sector employees.

- **Military employees**, referring to military officers, soldiers as well as veterans

- **Members of religious or ethnic groups**

Some of these groups constitute mutually exclusive categories, whereas other do not. Notably, religious or ethnic groups may overlap with the socio-economically based groups. As discussed in the paper, public sector employees can plausibly be considered part of the urban middle class, rather than as a separate category. If so, the dummies for the two original categories may be collapsed with an "OR" rule, i.e., the combined middle class dummies take the value 1 if (at least) one of the corresponding dummies for the disaggregated categories take the value 1.

The coding rules for each indicator, that was coded separately for each social group, are described below:

1. **Participated in the campaign**

   Records whether a social group is mentioned in the sources as having participated in the campaign at some point in time. As noted, we always consult several sources to determine the profile of each campaign. To be assigned the value "1", at least 2 sources should mention that the social group participated in the campaign. To be assigned the value "0", at least 4-5 sources should *not* mention these groups as having participated in the campaign, to ensure that the non-mentioning is not simply a missing data or poor data quality problem. Campaigns were also coded as 0 if at least 2 sources described the campaign as consisting exclusively of other groups, such as peasants or workers.

- 0: The social group was *not* reported to have participated in the campaign

- 1: The social group, at a minimum, is reported to have participated in the campaign at some point in time

2. **Dominated the campaigns**

Records whether the campaign was dominated by the given social group. Dominance is scored as 1 if the campaign is in line with at *least one* of the following criterias: 1) The relevant social group made up a majority of campaign members or 2) The relevant social group had a critical impact on the outcome and/or the strategies of the campaign. To be considered to be in majority, the sources should mention this explicitly, for instance through formulations such as "most of the protesters were peasants". Conversely, if sources mention explicitly that other social groups were in majority, the majority criteria is considered as not being fulfilled. The "critical impact" criterion is more reliant on subjective judgment. It should be considered fulfilled if the sources describe the relevant social group as having been the most important force behind the strategies and/or the outcome of the movement.

- 0: The campaign was not dominated by the social group

- 1: The campaign was dominated by the social group

3. **Initiated the campaign**

Records whether the campaign originated among or was initiated by the relevant social group. To code this variable, the year that is listed as "start year" in NAVCO is considered as the campaign's starting point. Hence, events taking place before this starting point are excluded from the consideration. If the movement originated among several groups, each group can be considered as having initiated the campaign although other groups also helped to initiate it. Hence, if the movement was described as originating among, e.g., peasants as well as other groups, the campaign will still be assigned a "1" on the initiation by peasants variable.

- 0: The campaign did not originate among the social group

- 1: The campaign originated among the social group

References

Ackerman, Peter and Jack DuVall. 2001. *A force more powerful: a century of non-violent conflict*. New York: Palgrave Macmillan.

Nepstad, Sharon Erickson. 2011. *Nonviolent revolutions: Civil resistance in the late 20th century*. Oxford: Oxford University Press.

Ness, Immanuel. 2015. "The International Encyclopedia of Revolution and Protest".
   URL: *http://www.revolutionprotestencyclopedia.com*

Swarthmore. 2015. "Global Nonviolent Action Database".
   URL: *http://nvdatabase.swarthmore.edu*

B   Examples of coded campaigns and sources

For each campaign, our coder created a brief summary containing the key information which was used to determine the social profile of the campaign, as well as the most important sources. Two examples of such summaries are given below:

B.1   South Korean Revolution against Rhee regime, 1960

1. Participated in the campaign: About 30.000 students started the demonstration in the morning, but more and more joined in, reaching a peak 100.000 of South Koreans. The sources also report the professors to have joined in after a while, along with what is often reported to be general "citizens". One source mentions workers in particular along with students at the march 26 April. As the revolt involved mostly urban areas, the rural population remained to a large degree unaffected. In addition to students, the urban participants included the unemployed, working class (mechanics, factory workers, taxi drivers etc.), and clerks, officials and staff members of political parties. A few were listed as merchants and farmers.

2. Dominated the campaign: Students initiated, organized and led the campaign. The spokesman for the campaign was also a student. Further, the protestors are often referred to exclusively as "the students".

3. Initiated the campaign: The campaign was a one-day long demonstration against the fraudulent election of President Rhee, spurred by the discovery of the body of a protestor student about one week before. The first protest happened on 18 April, and escalated as the protesters were assaulted on their way home. Students from seven different universities thus came together to protest the election.

Sources:

- http://nvdatabase.swarthmore.edu/content/south-korean-students-force-dictator-resign-new-elections-1960   Students initiated the campaign, the spokesman was a student, lots of South Koreans joined.

- https://www.marxists.org/archive/grant/1960/05/skorea.htm Workers also demonstrated

- Kim, Quee-Young (1996). "From Protest to Change of Regime: The 4-19 Revolt and the Fall of the Rhee Regime in South Korea". In Social Forces 74(4), p. 1179-1208. Students and professors led the protests. Those who joined were unemployed citizens, workers, clerks, officials, staff members of political parties, merchants and farmers. Article also argues for the important role of students.

B.2   Argentinean ERP/Monteneros against Peronist government(s), 1973-1977

1. Participated in the campaign: A recurrent topic in the sources regarding the composition of the two guerilla groups, is that they employ young people. The ERP began as a small Marxist group trying to organize workers, while the Monteneros eventually became a large group originating from students and young professionals. An article argues for the important role of the middle class, but mentions the role of working class armed resistance groups due to blurry definitions of what "working class" is. Peasants were documented to not have played a part since the guerilla groups were urban.

2. Dominated the campaign: Literature stresses that the groups were urban guerilla groups, leftist oriented, and some sources argue for the lack of workers, as revolutionary protest in this style was a middle class phenomenon.

3. Initiated the campaign: The two guerilla groups ERP and Monteneros were respectively Marxist and socialist groups who fought for a communist Argentina. ERP was formed as a military wing of the Worker's Revolutionary Party, while Monteneros was formed by the revolutionary left wing elements of the Peronist youth movement.

Sources:

- http://ucdp.uu.se/statebased/502 ERP originated from worker's party, Monteneros from Peronist youth movement.

- http://upsidedownworld.org/archives/international/argentina-the-creation-of-an-urban-guerrilla/

  ERP began as a small Marxist political group trying to organize workers. Nearly all founding members of the Monteneros were students or young professionals. Peasants played no role.

C Descriptive statistics

The social group profile of anti–regime campaigns over time
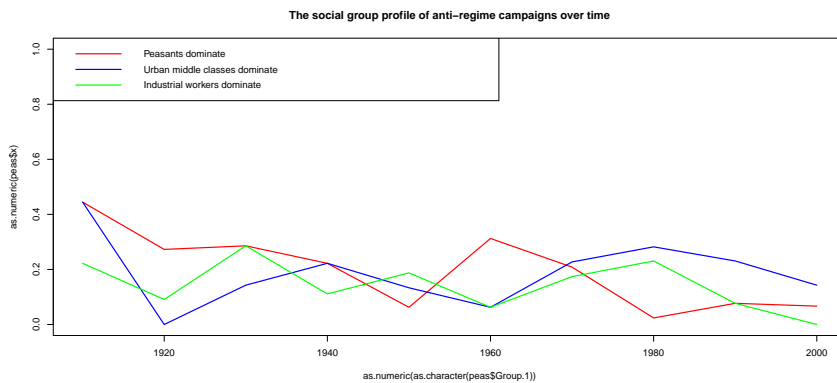
Figure A.1: Social group profile of opposition movements over time. Campaigns dominated by peasants, urban middle classes, and industrial workers.

Table A.1: Summary statistics for key variables, restricted to the 9099 observations in the benchmark model

	Obs.	Mean	St.dev.	Min.	Max.
Polyarchy	9099	.3862071	.2831038	.0092174	.9434598
Democracy (BMR)	8110	.4274969	.4947459	0	1
Middle classes OR ind.workers dominate	9099	.0161556	.1260808	0	1
Middle classes dominate	9099	.0103308	.1011197	0	1
Industrial workers dominate	9099	.0087922	.0933586	0	1
Middle class origin	9072	.0126764	.1118797	0	1
Industrial worker origin	9081	.0068274	.0823503	0	1
Middle class participate	9008	.0500666	.2180945	0	1
Industrial workers participate	9049	.0364681	.1874622	0	1
Public sector workers dominate	9088	.003191	.056402	0	1
Religious or ethnic group dominate	9094	.0341984	.1817484	0	1
Peasants dominate	9093	.0272737	.1628889	0	1
Public sector workers origin	9084	.0027521	.052391	0	1
Religious or ethnic group origin	9056	.0310292	.1734061	0	1
Peasants origin	9047	.0102797	.1008717	0	1
Peasants participate	9074	.0639189	.2446219	0	1
Public sector workers participate	9037	.0295452	.1693383	0	1
Religious or ethnic group participate	9042	.0508737	.2197519	0	1
Other campaign	9099	.07869	.2692691	0	1
Ln campaign members	9099	3.023079	12.01612	0	68
Regime violence	9099	.0715463	.2577494	0	1
Ln GDP p.c.	9099	7.803445	1.007891	5.315224	10.66701
Ln Population	9099	15.85347	1.553752	10.12663	20.97741
Urbanization	9099	.420854	.2277943	.0134997	.97279
Industrial worker index	9036	.0499115	.2995089	0	3
Middle class index	8981	.0684779	.3401307	0	3
Middle class dominated campaigns in neighborhood	9099	.0674282	.1097395	0	1.636364
Ind. worker dominated campaigns in neighborhood	9099	.045665	.094836	0	1.454545
Campaigns in neighborhood	9099	.1007844	.1074677	0	.8181818
Democracy score in neighborhood	9099	.3755455	.2171304	.0325952	.8931119
Observations	9738				

D   Inter-coder reliability tests and benchmark regression results using scores from the replication coder

This section presents results from the inter-coder reliability exercise, and subsequent tests assessing how measurement errors may affect results. We used two RAs, one as an "original coder" and one as a "replication coder". The replication coder coded a randomly drawn sample (by the authors, using a quasi-randomization routine in STATA) corresponding to about 10% of the cases coded by the original coder (since the sample was randomly drawn, it includes a couple of cases that were not scored by the original coder due to missing or too uncertain information). The replication coding was conducted several months after the original coding had been conducted and without being in contact with the original coder (who was then no longer working as an RA). The first table shows true-positives (where the coders agree that the score is 1) and the second table shows true negatives (where the coders agree that the score is 0) for each variable. This shows quite high levels of agreement, with true positive shares of .80 (4/5) and 1.00 (4/4) for the two main dummies used in the paper capturing campaigns dominated by, respectively, urban middle classes and industrial workers. The true negative rates for these two dummies were also high, respectively, .77 (10/13) and .86 (12/14).

The following tables replicate the main results on samples where the 10% replicated observations coded by the replication-coder substitute for the same observations as coded by the original coder. If these results strongly diverge from our baseline results, this would indicate that our results hinge on idiosyncrasies of the coding. However, the augmented sample yields quite similar coefficients to the original results reported in the paper.

Table A.2: True positives

Variable	Original coding = 1	Original coding =1 and and replication coding =1	Share "true positives"
At least urban middle class	16	14	.88
At least industrial workers	11	8	.73
Urban middle class dominated	5	4	.80
Industrial workers dominated	4	4	1.00

Table A.3: True negatives

Variable	Original coding = 0	Original coding =0 and and replication coding =0	Share "true negatives"
At least urban middle class	3	2	.67
At least industrial workers	7	7	1.00
Urban middle class dominated	13	10	.77
Industrial workers dominated	14	12	.86

Table A.4: Replicating Table 2 in the paper, replacing original codings by the 21 replicated observations

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Democratic transition (0 to 1 on BMR)				Polyarchy from V-Dem (0–1 scale)			
Ind. work. OR mid. cl. dominate	0.161**	0.089	0.077***	0.053*				
	(2.27)	(1.11)	(2.69)	(1.74)				
Middle classes dominate					0.128	0.096	0.030	0.026
					(1.55)	(1.09)	(0.91)	(0.76)
Industrial workers dominate					0.198**	0.135	0.118***	0.087**
					(2.26)	(1.44)	(3.23)	(2.27)
Ln participants		0.000		0.000		0.000		0.000
		(0.03)		(0.53)		(0.17)		(0.62)
Urbanization		0.495*		0.210*		0.502*		0.203*
		(1.68)		(1.84)		(1.71)		(1.78)
Ln population		-0.009		-0.002		-0.008		-0.001
		(-0.36)		(-0.16)		(-0.31)		(-0.14)
Ln GDP per capita		0.064		0.037		0.052		0.034
		(0.92)		(1.35)		(0.75)		(1.26)
Lagged dependent variable			0.653***	0.459***			0.666***	0.482***
			(7.97)	(4.96)			(8.19)	(5.16)
Constant	0.143***	-0.327	0.144***	-0.120	0.132***	-0.286	0.139***	-0.112
	(3.47)	(-0.57)	(5.02)	(-0.53)	(3.33)	(-0.50)	(4.91)	(-0.50)
N	137.000	120.000	176.000	155.000	136.000	119.000	175.000	154.000
R2	0.037	0.142	0.300	0.341	0.068	0.163	0.323	0.356

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses.
OLS regressions with campaign as cross-section unit.

Table A.5: Replicating table 3: Replacing coded observations with inter-coded observations

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimator	FE OLS	FE OLS	FE OLS	FE OLS	FE Logit	Logit	Logit
Dependent variable	V-Dem Polyarchy		Δ Polyarchy	Upturns Polyar.	BMR democratization		
Ind. work. OR mid. cl. dominate	0.038***				0.645	1.283**	
	(3.01)				(1.05)	(2.55)	
Middle classes dominate		0.023	0.023	0.028**			1.331***
		(1.32)	(1.32)	(2.01)			(3.27)
Industrial Workers dominate		0.053***	0.053***	0.047***			1.119***
		(2.90)	(2.90)	(2.89)			(2.61)
Other campaign	-0.002	-0.003	-0.003	0.006**	0.899*	0.793**	0.716**
	(-0.45)	(-0.70)	(-0.70)	(2.03)	(1.80)	(2.09)	(2.01)
Ln GDP p.c.	0.001	0.001	0.001	-0.001	1.092**	0.527**	0.383*
	(0.48)	(0.36)	(0.36)	(-0.49)	(2.02)	(2.28)	(1.68)
Ln Population	-0.001	-0.001	-0.001	0.003	-0.815	-0.026	0.003
	(-0.52)	(-0.41)	(-0.41)	(1.07)	(-0.82)	(-0.30)	(0.04)
Urbanization	0.024**	0.024**	0.024**	0.023**	-5.064	-0.323	0.261
	(2.12)	(2.10)	(2.10)	(2.02)	(-1.59)	(-0.37)	(0.31)
Lagged Polyarchy	0.948***	0.949***	-0.051***	-0.035***			
	(168.61)	(172.27)	(-9.26)	(-7.63)			
Country dummies	Y	Y	Y	Y	N	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y
N	9102	9102	9102	9102	2769	3153	4705
R2	0.937	0.938	0.087	0.104			

Notes: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. T-values in parentheses.
Except for in Model 3, errors are clustered by country.

E   Accounting for measurement error in the core independent variables

This section estimates our baseline models but incorporating measurement error and running Errors-in-variables models (EIV, using the eivreg command in Stata). This allows us to incorporate the uncertainty of the measured treatment variables by specifying their "reliability" parameter. If we are 100% certain of the measurement, this parameter is 1. We first estimate these models with the reliability parameter fixed to the agreement scores from our inter-coder reliability test presented above. As displayed in the previous section of this Appendix, these tests indicate around 80% agreement, so we accordingly set the reliability parameter (for both core treatment variables) to .80. We also run a set of models where we set the reliability parameters to .60, to reflect an even higher level of uncertainty. It might, for example, be that we were "lucky" in drawing cases that were relatively easy to code in the inter-coder reliability tests (although we have no indications that this is the case, and we used a randomization routine to extract the 10% of cases that should be double-coded).

Nonetheless, the tables presented below, show that our results are not substantially weakened when we incorporate these levels of measurement error. As discussed in the paper, accounting for measurement error (be it the low level suggested by our inter-coder reliability tests, or the higher alternative level) strengthens the result for the urban middle class campaigns considerably, as this dummy is statistically significant at least at the 5% level in all the EIV specifications. We also note that the (originally very clear) results for the industrial worker campaign dummy are further strengthened, with t-values ranging from 7.5 to 9.3 in the six different EIV models.

Table A.6: Replicating cross-section times series analysis with Errors-in-variables models

Model	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	EIV	EIV	EIV	EIV	EIV	EIV
Reliability	.80	.80	.80	.80	.80	.80
Outcome	Polyarchy	Polyarchy	Δ Polyarchy	Δ Polyarchy	Upturns	Upturns
Middle class OR ind. workers dominate	0.056***		0.056***		0.059***	
	(8.59)		(8.59)		(11.37)	
Industrial workers dominate		0.071***		0.071***		0.062***
		(8.49)		(8.49)		(9.33)
Middle classes dominate		0.027***		0.027***		0.036***
		(3.57)		(3.57)		(5.84)
Other campaign	-0.006**	-0.006**	-0.006**	-0.006**	0.002	0.003
	(-2.27)	(-2.37)	(-2.27)	(-2.37)	(1.08)	(1.49)
Ln GDP p.c.	0.001	0.001	0.001	0.001	-0.001	-0.001
	(0.64)	(0.50)	(0.64)	(0.50)	(-0.87)	(-0.98)
Ln Population	-0.001	-0.000	-0.001	-0.000	0.003	0.003
	(-0.28)	(-0.18)	(-0.28)	(-0.18)	(1.60)	(1.59)
Urbanization	0.023**	0.023**	0.023**	0.023**	0.022***	0.022***
	(2.35)	(2.35)	(2.35)	(2.35)	(2.82)	(2.88)
Lagged Polyarchy	0.950***	0.950***	-0.050***	-0.050***	-0.034***	-0.034***
	(257.74)	(258.82)	(-13.65)	(-13.66)	(-11.74)	(-11.80)
Country dummies	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
N	9099	9099	9099	9099	9099	9099

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses. Polyarchy in $t+1$ is dependent variable.
All models are OLS with country- and year fixed effects. Errors are clustered by country.

Table A.7: Replicating cross-section times series analysis with Errors-in-variables models

Model	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	EIV	EIV	EIV	EIV	EIV	EIV
Reliability	.60	.60	.60	.60	.60	.60
Outcome	Polyarchy	Polyarchy	Δ Polyarchy	Δ Polyarchy	Upturns	Upturns
Middle class OR ind. workers dominate	0.095***		0.095***		0.100***	
	(8.63)		(8.63)		(11.45)	
Industrial workers dominate		0.108***		0.108***		0.091***
		(7.50)		(7.50)		(7.98)
Middle classes dominate		0.028**		0.028**		0.043***
		(2.11)		(2.11)		(4.14)
Other campaign	-0.016***	-0.011***	-0.016***	-0.011***	-0.008***	-0.002
	(-4.56)	(-3.93)	(-4.56)	(-3.93)	(-2.91)	(-1.01)
Ln GDP p.c.	0.001	0.001	0.001	0.001	-0.002	-0.002
	(0.42)	(0.29)	(0.42)	(0.29)	(-1.17)	(-1.22)
Ln Population	0.001	0.001	0.001	0.001	0.005**	0.004**
	(0.29)	(0.24)	(0.29)	(0.24)	(2.35)	(2.05)
Urbanization	0.021**	0.021**	0.021**	0.021**	0.020**	0.021***
	(2.14)	(2.20)	(2.14)	(2.20)	(2.54)	(2.72)
Lagged Polyarchy	0.952***	0.951***	-0.048***	-0.049***	-0.032***	-0.033***
	(257.09)	(259.25)	(-13.01)	(-13.29)	(-10.95)	(-11.38)
country dummies	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
N	9099	9099	9099	9099	9099	9099

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses. Polyarchy in $t+1$ is dependent variable.
All models are OLS with country- and year fixed effects. Errors are clustered by country.

F   Accounting for regime-strategies

While we focus on the role and characteristics of the opposition movement in the paper, we also emphasize that the outcomes of civil resistance result from a two-player game between the opposition and the regime. This means that the choices made by opposition actors are conditioned by the actions of the regime, and vice versa. Notably, whether the regime is likely to use violence against the opposition or not is likely to influence opposition tactics, and opposition tactics are likely to influence whether the regime uses violence. We conceive of this strategic interaction as an omitted variable problem. On this understanding, regime-strategies are both post-treatment (conditioned by the opposition movement) and pre-treatment (conditioning the opposition movement).

While we make some headway towards handling the pre-treatment problem by, e.g., conducting an instrumental variable analysis in the paper, the inclusion of regime-strategies, such as regime-violence (see Model 4, Table 4), as a control variable is not fully adequate for addressing the post-treatment problem. Particularly, such an estimation strategy does not tell us whether there is a *direct effect of opposition profile when we hold regime strategies constant.* Furthermore, controlling for regime strategies in a conventional way does not enable us to assess the potential intermediate variable bias arising from the inclusion of the regime-strategy controls. To get at this, we estimate the Controlled Direct Effect (CDE) using the procedure developed by (Acharaya, Blackwell and Sen, 2017). This uses sequential-G estimation to estimate the direct effect of some treatment when net of the mediation effect. It also allows us to investigate sensitivity to intermediate variable bias.

Figures A.2 and A.3 plots the CDE coefficients and the coefficients from the "conventional" model (where we do not perform sequential-G estimation), when we fix the mediators "regime violence", "foreign support",and "military defection" (all taken from NAVCO). Since the CDE coefficients are estimated using two equations, we bootstrap the standard errors using 1000 sampling runs. The figures show that the Controlled Direct Effects are quite similar to the conventional coefficients. This indicates that, when we condition on regime-strategies, holding

variables measuring them fixed at their unit-specific levels, we find very similar results as those reported in the simpler models reported in the paper. More importantly, this exercise suggests that our results hold up even when accounting for important (and measurable) features of the response of the regime to the opposition movement.

Figure A.2: Assessing Controlled Direct Effects
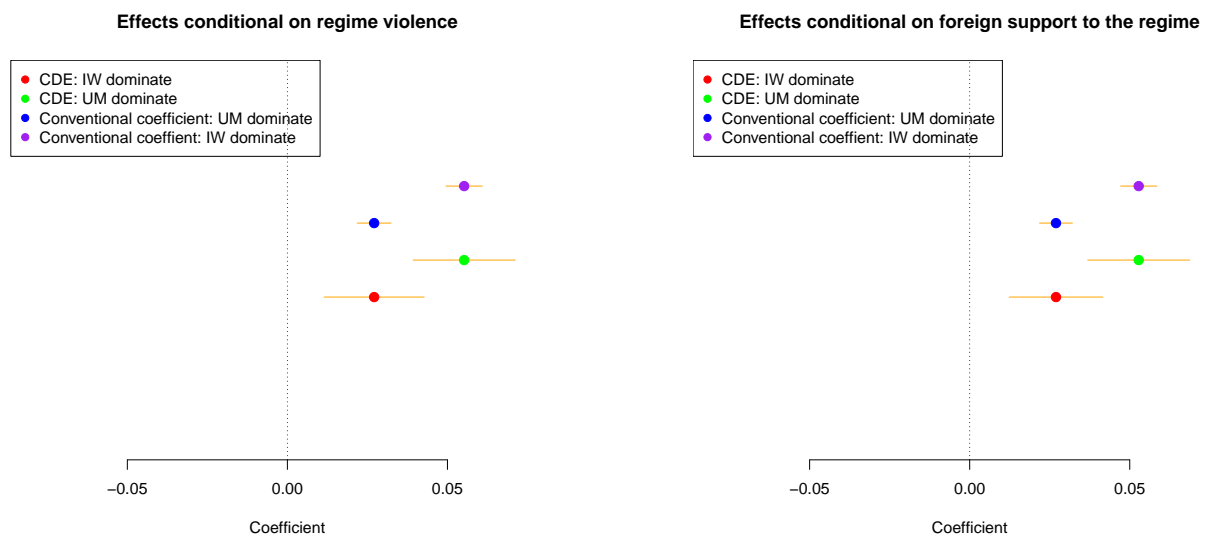
Figure A.3: Assessing Controlled Direct Effects
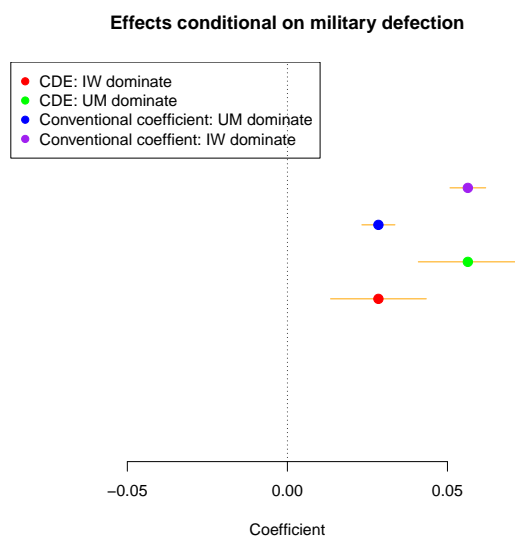
G Instrumental variables regressions

We endogenize the lexical campaign-profile indices to allow for maximum variance. However, results turn out about equally strong for both IW and UMC when we employ the "campaign domination" measures from the benchmark (Table A.21). As instruments, we use the IW and UW profiles of anti-regime opposition campaigns in the neighbourhood of country i, calculated by world regions. We expect, for instance, ongoing industrial worker campaigns in other countries in the region to be a strong instrument for industrial worker campaigns domestically. The assumption is that social movements diffuse across borders and affect the mobilization of similar movements in neighbouring countries, for instance because prospective opposition actors may be inspired or learn from observing and communicating with their counterparts in neighboring countries. This is validated in several studies on how social movements diffuse through mobilizing similar and (often) allied groups in closely related countries (see, e.g., Weyland, 2014, 2009). Indeed, the first-stage F-statistics and t-values of the neighborhood instruments reveal that they are very strong.

To achieve consistent estimates we also need to ensure that neighbourhood campaign profile is not otherwise systematically linked to democratization domestically. While we do include country- and year-fixed effects and other covariates from the benchmark, it may, e.g., be that neighborhood campaigns affect neighborhood democracy, which in turn affects democracy domestically. We therefore condition on *a)* the neighborhood-profile of "other campaigns", and *b)* the neighborhood-profile on democracy. The assumption underlying these specifications is thus that the social profile of a campaign *in the neighborhood – conditional on all included covariates* – only affects democratic change *via* the campaign profile in country i. Granted, this is a quite strong assumption, and we do not consider this estimation strategy a "silver bullet" for identifying causal effects.[15]

---

[15]Yet, we have not come up with a plausible theoretical mechanism that would lead the exclusion restriction to be grossly violated once we condition on the discussed covariates. We note, however, that the IV estimate describes a Local Average Treatment Effect (LATE), interpreted as the effect of campaign profile for those campaigns that "responded" to the instrument, in the sense that they were predicted by similar campaigns in the region. While we have no clear

Table A.8: Instrumental variable models (Fixed Effects 2SLS)

Stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage	1st stage
Dependent variable	Polyarchy	IW index	Polyarchy	MC index	Polyarchy	UG index
Model	(1)	(2)	(3)	(4)	(5)	(6)
Industrial worker index	0.066***		0.001	0.247***		
	(3.64)		(0.10)	(2.96)		
Middle class index	-0.002	0.270***	0.054**			
	(-0.24)	(2.94)	(2.42)			
Urban groups (combined) index					0.123***	
					(4.14)	
Other campaign	-0.043***	0.416***	-0.056***	0.891***	-0.158***	1.302***
	(-4.14)	(3.25)	(-2.63)	(7.66)	(-3.86)	(13.28)
Ln GDP p.c.	0.003	0.010	0.003	0.015	0.002	0.015
	(1.12)	(0.57)	(1.06)	(0.78)	(0.64)	(0.80)
Ln Population	0.001	-0.024	0.000	-0.015	0.004	-0.047*
	(0.38)	(-0.80)	(0.13)	(-0.65)	(1.23)	(-1.95)
Urbanization	0.017	-0.015	0.014	0.041	0.004	0.106
	(1.35)	(-0.14)	(1.34)	(0.44)	(0.22)	(0.96)
Other campaigns in neighb.	-0.005	-0.255**	-0.002	-0.412***	0.039***	-0.611***
	(-0.38)	(-2.00)	(-0.21)	(-3.81)	(2.65)	(-4.31)
Democracy in neighborhood	0.050***	0.143	0.057***	0.034	0.049**	0.109
	(3.23)	(1.00)	(3.85)	(0.30)	(2.52)	(0.83)
Ind. work. camp. in neighb.		1.116***	0.076***	-0.337**		0.543**
		(4.04)	(4.78)	(-2.18)		(2.37)
Mid. class camp. in neighb.	0.052**	-0.349**		0.785***		0.449**
	(2.53)	(-2.20)		(4.53)		(2.61)
Lagged Polyarchy	0.945***	-0.075	0.942***	-0.035	0.952***	-0.106**
	(140.51)	(-1.43)	(137.69)	(-0.97)	(107.98)	(-2.16)
Country dummies	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
N	8980		8980		8980	
Hansen J-test p-value	Ex. ident.		Ex. ident.		.64	
Cragg-Donald Wald F-stat.	696.7		377.4		236.6	
Instrument(s)	IW neighb.		UM neighb.		IW neighb,UM neighb	

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses.
Polyarchy in $t+1$ is dependent variable. Errors are clustered by country.

The IV (FE2SLS) results are presented in Table A.8. Model 1 shows the second-stage regression treating IW as endogenous. It shows a positive and precisely estimated ($t = 3.6$) effect on democratization. Model 2, showing the first-stage regression, displays the strength of the industrial-worker-campaigns-in-neighborhood instrument, with a t-value of 4.0 and F-values far above conventional heuristics for identifying strong instruments (e.g., Stock and Yogo, 2005). While one can never guarantee that the exclusion restriction holds, Model 1 thus suggests a clear effect of industrial worker campaigns on democratization. The estimated effect of urban middle class campaigns is also fairly clear. MC is significant at 5% ($t = 2.4$) in Model 3. Again, the first-stage regression shows that the urban-middle-class-campaigns-in-neighborhood instrument is strong. Finally, in Models 5–6, we test for an effect of (any type

grounds for suspecting this, the LATE(s) estimated in the IV models might not be generalizable to all campaigns, and is in this sense less externally valid than the non-IV estimates.

of) urban campaign profile, employing an index that takes the value 0 if there is no involvement by either industrial workers or the middle class, 1 if at least one of these (only) participated, 2 if the campaign originated from one of the movements, and 3 if one of them dominated the campaign. We find a sizeable and highly significant ($t = 4.1$) coefficient. Since we here include two instruments – measures of both industrial worker and middle class profiles of campaigns in the neighborhood – we can run standard overidenfication tests. According to the Hansen J-test p-value (.64) the exclusion restriction does not seem to be violated. This provides additional evidence for the hypothesis that urban-group opposition movements spur democratic change.

H Additional robustness tests and extensions

The following section contains various other robustness tests and extensions noted in the main text, but not reported there in Tables. The sequence of the tests is as follows:

Table A.9 shows campaign level regressions on democratization reflecting the situation five years – rather than the baseline one year – after the relevant campaign ended. Table A.10 displays logit versions of the (campaign-level) baseline OLS regressions when the dichotomous BMR measure is used as dependent variable.

Moving to the robustness tests on the country-year regressions, Tables A.11–A.13 displays regressions using changes (either all changes, only positive, or only negative) in the Polyarchy measure as the dependent variable, and these tables contains specifications controlling for and not controlling for the initial level of Polyarchy. Tables A.14–A.17 displays various regressions investigating Polyarchy changes with countries defined as minimum democracies and autocracies, according to BMR, and including and excluding instances of democratization (as coded using BMR). This allows for showing how campaigns dominated by urban groups can affect gradual changes in level of democracy, also when not pushing a regime across the threshold for being coded as a minimalist democracy. Table A.18 provides a large set of additional robustness tests when considering the measure on whether campaigns were dominated by either one of the two urban groups (industrial workers or urban middle classes) under focus. Table A.19 replicates the robustness test table (Table 4) in the paper, but now using positive changes

in Polyarchy ("democratic upturns") as the dependent variable.

Further, Table A.20 nuances the core findings in one particular respect: These are the core country-year regressions using Polyarchy as dependent variable from Table 3, but now providing a more detailed categorization of campaigns dominated by other groups than the two urban groups under focus. The results from this table suggests that campaigns dominated by peasants are particularly "bad" for democracy.

Table A.21 replicates the second stages of the 2SLS regressions from Table A.8, but now using the "campaign domination" dummies rather than the more fine grained lexical indices as endogenous regressors. Finally, Table A.22 reproduces the split sample tests reported in Table 5 of the paper, but this time differentiating between industrial workers and urban middle classes rather than using the combined variable for whether one of these "urban groups" dominate the campaign.

Table A.9: Campaign-level (cross sectional) correlates of democratization, measuring democratic transition with longer time horizon. That is, going from 0 on BMR at campaign start to 1 five years after campaign ended

Model	(1)	(2)	(3)	(4)
Middle classes OR ind. workers dominate	0.293***	0.156*		
	(3.80)	(1.86)		
Industrial workers dominate			0.247**	0.140
			(2.61)	(1.42)
Middle classes dominate			0.266***	0.216**
			(2.96)	(2.32)
Ln participants		0.000		0.001
		(0.12)		(0.42)
Urbanization		0.222		0.220
		(0.72)		(0.72)
Ln population		-0.023		-0.019
		(-0.85)		(-0.70)
Ln GDP per capita		0.185**		0.177**
		(2.56)		(2.48)
Constant	0.186***	-0.840	0.169***	-0.882
	(4.24)	(-1.41)	(4.05)	(-1.49)
N	143	124	140	122

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses.
OLS regressions with campaign as cross-section unit.
Only campaigns initiated in BMR autocracies included in sample.

Table A.10: Campaign-level correlates of democratization as measured by change from 0 to 1 in BMR. Logit models

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Dem. trans. measured 1 year after campaign end				Dem. trans. measured 5 years after campaign end			
MC OR IW dominate	1.140***	0.628			1.392***	0.765		
	(2.63)	(1.24)			(3.53)	(1.62)		
IW dominate			1.012**	0.854			1.088**	0.699
			(2.02)	(1.55)			(2.25)	(1.27)
MC dominate			0.818*	0.442			1.440***	1.185**
			(1.70)	(0.81)			(3.24)	(2.21)
Ln participants		-0.003		-0.001		-0.000		0.003
		(-0.24)		(-0.06)		(-0.04)		(0.34)
urbanization		2.879		3.064		1.107		1.279
		(1.52)		(1.60)		(0.60)		(0.67)
Ln population		-0.114		-0.083		-0.155		-0.134
		(-0.62)		(-0.44)		(-0.92)		(-0.77)
Ln GDP per capita		0.305		0.261		1.020**		0.985**
		(0.69)		(0.58)		(2.28)		(2.18)
Constant	-1.866***	-3.192	-1.879***	-3.581	-1.479***	-6.729*	-1.573***	-7.166**
	(-6.26)	(-0.85)	(-6.57)	(-0.95)	(-5.66)	(-1.90)	(-6.07)	(-1.98)
N	143	124	143	124	143	124	143	124

Notes: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. T-values in parentheses.
Logit regressions with campaign as cross-section unit.
Only campaigns initiated in BMR autocracies included in sample.

Table A.11: Core cross-section time series models, with change (from t to $t+1$) in Polyarchy as dependent variable

Model	(1)	(2)	(3)	(4)
Estimator	FE OLS	FE OLS	FE OLS	FE OLS
Middle class OR ind. workers dominate	0.044***		0.040***	
	(3.34)		(3.03)	
Industrial workers dominate		0.057***		0.053***
		(2.89)		(2.80)
Middle classes dominate		0.026		0.023
		(1.37)		(1.26)
Other campaign	-0.001	-0.002	-0.002	-0.003
	(-0.35)	(-0.57)	(-0.56)	(-0.83)
Ln GDP p.c.	-0.002	-0.002	0.001	0.001
	(-0.86)	(-0.98)	(0.51)	(0.43)
Ln Population	-0.001	-0.001	-0.001	-0.001
	(-0.52)	(-0.37)	(-0.51)	(-0.40)
Urbanization	0.019**	0.019**	0.024**	0.024**
	(2.14)	(2.07)	(2.12)	(2.11)
Lagged Polyarchy (level in t)			-0.051***	-0.051***
			(-8.98)	(-9.10)
Constant	0.030	0.027	0.018	0.015
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9099	9099	9099	9099
Countries	147	147	147	147
Max time series	1900–2006	1900–2006	1900–2006	1900–2006

Notes: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. T-values in parentheses.
Except for in Model 3, errors are clustered by country.

Table A.12: Core cross-section time series models, with "upturns"/positive change (from t to $t+1$) in Polyarchy as dependent variable (democratic downturns coded as 0 change)

Model	(1)	(2)	(3)	(4)
Estimator	FE OLS	FE OLS	FE OLS	FE OLS
Middle class OR ind. workers dominate	0.045***		0.042***	
	(4.01)		(3.73)	
Industrial workers dominate		0.050***		0.047***
		(2.86)		(2.79)
Middle classes dominate		0.031**		0.029*
		(1.98)		(1.89)
Other campaign	0.007**	0.007**	0.007**	0.006**
	(2.43)	(2.22)	(2.28)	(2.03)
Ln GDP p.c.	-0.003	-0.004	-0.001	-0.001
	(-1.30)	(-1.38)	(-0.38)	(-0.45)
Ln Population	0.003	0.003	0.003	0.003
	(1.36)	(1.38)	(1.01)	(1.05)
Urbanization	0.020*	0.020*	0.023**	0.023**
	(1.90)	(1.88)	(2.03)	(2.04)
Lagged Polyarchy			-0.035***	-0.035***
			(-7.38)	(-7.49)
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9099	9099	9099	9099
Countries	147	147	147	147
Max time series	1900–2006	1900–2006	1900–2006	1900–2006

Notes: *$p<0.1$; **$p<0.05$; ***$p<0.01$. T-values in parentheses.
Except for in Model 3, errors are clustered by country.

Table A.13: Core cross-section time series models, with "downturns/negative change (from t to $t + 1$) in Polyarchy as dependent variable (democratic upturns coded as 0 change

Model	(1)	(2)	(3)	(4)
Estimator	FE OLS	FE OLS	FE OLS	FE OLS
Middle class OR ind. workers dominate	-0.001		-0.002	
	(-0.15)		(-0.47)	
Industrial workers dominate		0.007**		0.006*
		(2.19)		(1.89)
Middle classes dominate		-0.005		-0.006
		(-0.82)		(-0.95)
Other campaign	-0.008***	-0.009***	-0.009***	-0.009***
	(-3.50)	(-3.78)	(-3.70)	(-4.00)
Ln GDP p.c.	0.001	0.001	0.002*	0.002*
	(1.10)	(1.04)	(1.94)	(1.87)
Ln Population	-0.004**	-0.004**	-0.004***	-0.004***
	(-2.60)	(-2.52)	(-2.68)	(-2.62)
Urbanization	-0.000	-0.001	0.001	0.001
	(-0.06)	(-0.11)	(0.17)	(0.11)
Lagged Polyarchy			-0.016***	-0.016***
			(-5.33)	(-5.26)
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9099	9099	9099	9099
Countries	147	147	147	147
Max time series	1900–2006	1900–2006	1900–2006	1900–2006

Notes: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. T-values in parentheses.
Except for in Model 3, errors are clustered by country.

Table A.14: Excluding all democracies (and excluding democratization years)

	(1) FE-OLS	(2) FE-OLS	(3) FE-OLS	(4) FE-OLS	(5) FE-OLS
Middle class OR ind. workers dominate	0.0637**	0.0639**	0.0629***	0.0470***	
	(3.04)	(2.96)	(7.78)	(5.92)	
Industrial workers dominate					0.0492***
					(5.14)
Middle classes dominate					0.0460***
					(4.78)
Other campaign	0.00151	-0.000964	-0.00391	-0.00558	-0.00831
	(0.35)	(-0.20)	(-0.87)	(-1.26)	(-1.92)
Ln GDP p.c.		-0.00384	-0.00316	0.0224***	0.0222***
		(-1.18)	(-0.89)	(5.63)	(5.59)
Ln Population		-0.000763	0.0330***	0.0128	0.0130
		(-0.71)	(6.62)	(1.73)	(1.76)
Urbanization		0.0222	-0.00348	-0.0224	-0.0218
		(1.80)	(-0.17)	(-0.97)	(-0.95)
Lagged Polyarchy	0.870***	0.859***	0.688***	0.634***	0.634***
	(42.84)	(35.52)	(56.98)	(49.41)	(49.59)
Country dummies	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y
N	5920	4617	4617	4617	4617

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: These models exclude all democracies and instances of democratization (as captured using the BMR democracy measure). This means that these models only look at liberalization short of full-blown democratiztion, in dictatorships.

Table A.15: Excluding all democracies (and including democratization years)

	(1) FE-OLS	(2) FE-OLS	(3) FE-OLS	(4) FE-OLS	(5) FE-OLS
Middle class OR ind. workers dominate	0.0637**	0.0639**	0.0629***		
	(3.04)	(2.96)	(7.78)		
Industrial workers dominate				0.0492***	0.0492***
				(5.14)	(5.14)
Middle classes dominate				0.0460***	0.0460***
				(4.78)	(4.78)
Other campaign	0.00151	-0.000964	-0.00391	-0.00831	-0.00831
	(0.35)	(-0.20)	(-0.87)	(-1.92)	(-1.92)
Ln GDP p.c.		-0.00384	-0.00316	0.0222***	0.0222***
		(-1.18)	(-0.89)	(5.59)	(5.59)
Ln Population		-0.000763	0.0330***	0.0130	0.0130
		(-0.71)	(6.62)	(1.76)	(1.76)
Urbanization		0.0222	-0.00348	-0.0218	-0.0218
		(1.80)	(-0.17)	(-0.95)	(-0.95)
Lagged Polyarchy	0.870***	0.859***	0.688***	0.634***	0.634***
	(42.84)	(35.52)	(56.98)	(49.59)	(49.59)
Country dummies	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y
N	5920	4617	4617	4617	4617

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: These models exclude all democracies but include instances of democratization (as captured using the BMR democracy measure). This means that these models only look at instances of partial and full-blown democratization in states that are autocratic.

Table A.16: Excluding all dictatorships (and including democratization years)

	(1) FE-OLS	(2) FE-OLS	(3) FE-OLS	(4) FE-OLS	(5) FE-OLS
Middle class OR ind. workers dominate	0.0675	0.0626	0.0168	0.0177	
	(1.70)	(1.55)	(1.10)	(1.21)	
Industrial workers dominate					0.149***
					(6.06)
Middle classes dominate					-0.0494**
					(-3.11)
Other campaign	-0.0209	-0.0223	-0.0156*	-0.0187**	-0.0148*
	(-1.97)	(-1.69)	(-2.12)	(-2.62)	(-2.11)
Ln GDP p.c.		0.0137**	0.0228***	0.00465	0.00445
		(3.29)	(5.29)	(0.72)	(0.69)
Ln Population		0.00120	0.00263	-0.00603	-0.00589
		(0.58)	(0.36)	(-0.74)	(-0.72)
Urbanization		0.0346	0.226***	0.138***	0.140***
		(1.72)	(8.42)	(4.95)	(5.02)
Lagged Polyarchy	0.869***	0.814***	0.614***	0.617***	0.623***
	(61.48)	(36.50)	(52.95)	(53.04)	(53.63)
Country dummies	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y
N	3647	3450	3450	3450	3450

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: These models exclude all dictatorships but includes instances of democratization. This means that these models look at incremental democratization in states that are already democratic according to the BMR threshold *and* full on democratization from dictatorship to democracy.

Table A.17: Excluding all dictatorships (and excluding democratization years)

	(1) FE-OLS	(2) FE-OLS	(3) FE-OLS	(4) FE-OLS	(5) FE-OLS
Middle class OR ind. workers dominate	-0.0423	-0.0417	-0.0547***	-0.0509**	
	(-1.83)	(-1.57)	(-3.36)	(-3.28)	
Industrial workers dominate					0.0749*
					(2.14)
Middle classes dominate					-0.0619***
					(-3.77)
Other campaign	-0.0181*	-0.0181*	-0.0127	-0.0142*	-0.0152*
	(-2.40)	(-2.25)	(-1.88)	(-2.20)	(-2.37)
Ln GDP p.c.		0.00671**	0.0104**	-0.000931	-0.000790
		(2.86)	(2.72)	(-0.16)	(-0.13)
Ln Population		0.000735	-0.00491	-0.0101	-0.00987
		(0.57)	(-0.75)	(-1.41)	(-1.37)
Urbanization		0.0224*	0.154***	0.0965***	0.0968***
		(2.09)	(6.51)	(3.93)	(3.95)
Lagged Polyarchy	0.931***	0.902***	0.776***	0.781***	0.781***
	(82.32)	(60.53)	(68.00)	(68.52)	(68.59)
Country dummies	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y
N	3512	3328	3328	3328	3328

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: These models exclude all dictatorships and instances of democratization. This means that these models only look at incremental democratization in states that are already democratic according to the BMR threshold.

Table A.18: Robustness tests when using combined dummy for urban middle class OR industrial workers dominate campaign as key measure

Model	(1) 2-year lag	(2) No clustering	(3) Parsimonious	(4) Extra controls	(5) Nr groups	(6) No uncertain cases
Middle class OR ind. workers dominate	0.060***	0.040***	0.041***	0.038***	0.033***	0.045***
	(2.76)	(8.58)	(3.27)	(2.75)	(2.66)	(3.49)
Other campaign	-0.002	-0.002	-0.002	-0.011	-0.047***	-0.003
	(-0.29)	(-0.81)	(-0.70)	(-0.65)	(-3.30)	(-0.71)
Ln GDP p.c.	0.005	0.001		0.001	0.001	0.001
	(0.88)	(0.73)		(0.51)	(0.47)	(0.38)
Ln Population	-0.004	-0.001		-0.001	-0.001	-0.001
	(-0.73)	(-0.52)		(-0.34)	(-0.27)	(-0.30)
Urbanization	0.049**	0.024**		0.024**	0.026**	0.026**
	(2.11)	(2.44)		(2.17)	(2.26)	(2.21)
Ln number of campaign participants				-0.000		
				(-0.86)		
Regime violence				0.018		
				(1.17)		
Number of social groups in campaign					0.015***	
					(3.29)	
Lagged Polyarchy	0.871***	0.949***	0.955***	0.949***	0.950***	0.949***
	(68.94)	(257.67)	(207.99)	(162.43)	(157.04)	(159.22)
Country dummies	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
N	8951	9099	14242	9099	8980	8998

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses. Polyarchy in $t+1$ is dependent variable.
All models are OLS with country- and year fixed effects. Except for in Model 2, errors are clustered by country.

xxx

Table A.19: Robustness tests, with positive change in Polyarchy from t to $t+1$ (democratic upturns) as dependent variable.

Model	(1) 2-year lag	(2) No clustering	(3) Parsimonious	(4) Extra controls	(5) Public employees	(6) Nr groups	(7) No uncertain	(8) Campaign origin	(9) Lexical indices
Industrial workers dominate	0.034***	0.047***	0.045***	0.044***	0.045***	0.041**	0.047**		
	(3.07)	(9.93)	(2.73)	(2.76)	(2.74)	(2.26)	(2.58)		
Middle class dominate	0.016**	0.029***	0.030**	0.028*	0.033***	0.028*	0.033*		
	(1.98)	(6.64)	(2.06)	(1.81)	(2.56)	(1.70)	(1.89)		
Industrial worker origin								0.032*	
								(1.86)	
Middle class origin								0.023**	
								(2.09)	
Industrial worker index									0.016***
									(2.75)
Middle class index									0.013***
									(2.92)
Other campaign	0.004	0.006***	0.007**	0.010	0.006*	-0.019**	0.006*	0.008**	-0.009*
	(1.53)	(3.19)	(2.46)	(1.01)	(1.88)	(-2.11)	(1.85)	(2.35)	(-1.87)
Ln GDP p.c.	-0.000	-0.001		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.14)	(-0.86)		(-0.44)	(-0.39)	(-0.40)	(-0.44)	(-0.53)	(-0.49)
Ln Population	0.002	0.003		0.003	0.003	0.003	0.003	0.002	0.003
	(0.75)	(1.34)		(1.24)	(1.03)	(1.28)	(1.29)	(0.95)	(1.24)
Urbanization	0.024**	0.023***		0.023**	0.023**	0.024**	0.023**	0.023**	0.023*
	(2.13)	(2.95)		(2.06)	(2.01)	(2.05)	(2.03)	(2.06)	(1.95)
Ln nr particip. camp.				-0.000*					
				(-1.74)					
Regime violence				0.006					
				(0.64)					
Nr. groups campaign						0.009**			
						(2.52)			
Lagged Polyarchy	-0.053***	-0.035***	-0.030***	-0.035***	-0.037***	-0.035***	-0.036***	-0.037***	-0.036***
	(-8.95)	(-12.02)	(-7.41)	(-7.32)	(-6.95)	(-6.55)	(-6.80)	(-7.28)	(-6.82)
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	8949	9099	14242	9099	9089	8980	8998	9095	8980

Notes: $*$ p$<$0.1; $**$ p$<$0.05; $***$ p$<$0.01. T-values in parentheses. Polyarchy in $t+1$ is dependent variable.
All models are OLS with country- and year fixed effects. Except for in Model 2, errors are clustered by country.

Table A.20: Core cross-section time series models, further nuancing "Other campaigns"

Model	(1)	(2)	(3)	(4)
Estimator	FE OLS	FE OLS	FE OLS	FE OLS
Dependent variable	V-Dem Polyarchy		Δ Polyarchy	Upturns Polyarchy
Middle class OR ind. workers dominate	0.046***			
	(3.78)			
Industrial workers dominate		0.052***	0.052***	0.045***
		(3.02)	(3.02)	(3.00)
Middle class dominate		0.034*	0.034*	0.033**
		(1.83)	(1.83)	(2.03)
Public employees dominate	0.023	0.020	0.020	0.019
	(0.70)	(0.63)	(0.63)	(0.69)
Regious or ethnic groups dominate	-0.003	-0.002	-0.002	-0.011*
	(-0.35)	(-0.26)	(-0.26)	(-1.71)
Peasants dominate	-0.025**	-0.024**	-0.024**	-0.029***
	(-2.36)	(-2.27)	(-2.27)	(-3.52)
Existence of any campaign	0.007	0.006	0.006	0.020***
	(0.87)	(0.69)	(0.69)	(3.26)
Ln GDP p.c.	0.002	0.002	0.002	-0.000
	(0.96)	(0.92)	(0.92)	(-0.19)
Ln Population	-0.000	-0.000	-0.000	0.004
	(-0.04)	(-0.00)	(-0.00)	(1.54)
Urbanization	0.021*	0.020*	0.020*	0.019*
	(1.94)	(1.94)	(1.94)	(1.84)
Lagged Polyarchy	0.948***	0.949***	-0.051***	-0.036***
	(165.21)	(168.59)	(-9.13)	(-7.32)
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9077	9077	9077	9077

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses.
Except for in Model 3, errors are clustered by country.

Table A.21: Instrumental variable second-stage regression (Fixed Effects 2SLS), with industrial worker- and/or urban middle class campaign domination variables (rather than lexical indices) as endogenous independent variables

Endogenous independent variable	Ind. work. dominate	Mid. cl. dominate	IW or MC dominate
Model	(1)	(2)	(3)
Industrial workers dominate	0.210***	0.012	
	(3.12)	(0.51)	
Middle classes dominate	-0.012	0.150**	
	(-0.58)	(2.44)	
Middle class OR ind. workers dominate			0.261***
			(5.08)
Other campaign	-0.022***	-0.023**	-0.060***
	(-2.97)	(-2.30)	(-3.42)
Ln GDP p.c.	0.002	0.003	0.002
	(0.80)	(0.99)	(0.39)
Ln Population	0.002	0.000	0.006
	(0.76)	(0.13)	(1.56)
Urbanization	0.016	0.017	0.013
	(1.09)	(1.42)	(0.74)
Lagged Polyarchy	0.945***	0.942***	0.952***
	(128.65)	(131.48)	(114.88)
Middle class dominate campaigns in neighborhood	0.055***		
	(2.69)		
Ind. worker dominate campaigns in neighborhood		0.076***	
		(4.79)	
Campaigns in neighborhood	0.001	0.003	0.039***
	(0.04)	(0.30)	(2.93)
Democracy score in neighborhood	0.061***	0.065***	0.057**
	(3.53)	(3.54)	(2.48)
Country dummies	Y	Y	Y
Year dummies	Y	Y	Y
N	9099	9099	9099
Hansen J-test p-value	Ex. ident.	Ex. ident.	.92
Cragg-Donald Wald F-stat.	527.8	294.7	270.7
Instrument(s)	IW neighb.	UM neighb.	IW neighb,UM neighb

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses.
Polyarchy in $t + 1$ is dependent variable. Errors are clustered by country.

Table A.22: Sensitivity and scope: Split-sample tests and interaction models, separating between urban middle class and industrial worker dominated campaigns.

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Sample	Time		Omitting particular geographic region						Urban and rural societies		
Sample	≤ 1956	> 1955	$\div WE/NA$	$\div EE$	$\div SSA$	$\div ME/NA$	$\div A/P$	$\div LA$	Rural	Urban	Full
Industrial workers dominate	0.022	0.075***	0.056***	0.060**	0.045**	0.053***	0.053***	0.048***	0.052***	0.064**	-0.022
	(1.65)	(3.08)	(2.90)	(2.55)	(2.21)	(2.77)	(2.63)	(2.91)	(2.70)	(2.07)	(-0.57)
Middle classes dominate	0.014	0.019	0.011	0.008	0.024	0.023	0.028	0.059***	0.006	0.039	-0.004
	(0.79)	(0.85)	(0.65)	(0.48)	(0.96)	(1.15)	(1.37)	(2.75)	(0.46)	(1.07)	(-0.07)
MC dom. X urbaniz.											0.096
											(0.59)
IW dom. X urbaniz.											0.162*
											(1.71)
Other campaign	-0.014	-0.002	-0.002	-0.004	-0.003	-0.003	-0.004	-0.003	-0.006	-0.005	-0.005
	(-1.59)	(-0.34)	(-0.45)	(-0.98)	(-0.60)	(-0.85)	(-0.97)	(-0.69)	(-1.47)	(-0.55)	(-0.67)
Oth. camp. X urbaniz.											0.004
											(0.21)
Ln GDP p.c.	-0.014*	-0.000	0.002	0.001	0.003	0.002	0.001	-0.000	-0.003	0.002	0.002
	(-1.93)	(-0.01)	(0.68)	(0.22)	(0.84)	(0.57)	(0.37)	(-0.13)	(-0.84)	(0.54)	(0.55)
Ln Population	-0.004	-0.009*	-0.002	0.000	-0.000	0.001	-0.001	-0.005*	-0.000	-0.012**	-0.001
	(-0.35)	(-1.78)	(-0.72)	(0.13)	(-0.17)	(0.32)	(-0.38)	(-1.68)	(-0.02)	(-2.12)	(-0.30)
Urbanization	-0.102***	0.040**	0.017	0.028**	0.031**	0.028**	0.011	0.025**	0.030	0.068***	0.020*
	(-2.90)	(2.11)	(1.24)	(2.17)	(2.51)	(2.40)	(0.90)	(2.10)	(1.38)	(2.61)	(1.75)
Lagged Polyarchy	0.940***	0.937***	0.945***	0.944***	0.954***	0.949***	0.949***	0.951***	0.917***	0.950***	0.950***
	(75.00)	(134.59)	(141.06)	(154.30)	(163.07)	(163.68)	(150.75)	(153.63)	(76.02)	(128.87)	(162.34)
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	2728	6371	7125	8565	6771	8341	7464	7229	4697	4403	9089

Notes: *p<0.1; **p<0.05; ***p<0.01. T-values in parentheses. Polyarchy in $t+1$ is dependent variable. Errors are clustered by country.
All models are OLS with country- and year fixed effects.