



**HAL**  
open science

# Who Revolt? Empirically Revisiting the Social Origins of Democracy

Carl Henrik Knutsen, Sirianne Dahlum, Tore Wig

► **To cite this version:**

Carl Henrik Knutsen, Sirianne Dahlum, Tore Wig. Who Revolt? Empirically Revisiting the Social Origins of Democracy. 2016. hal-03471866

**HAL Id: hal-03471866**

**<https://hal-sciencespo.archives-ouvertes.fr/hal-03471866>**

Preprint submitted on 9 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**SciencesPo**

LABORATOIRE INTERDISCIPLINAIRE  
D'ÉVALUATION DES POLITIQUES PUBLIQUES

## LIEPP Working Paper

November 2016, n°60

# Who Revolt? Empirically Revisiting the Social Origins of Democracy

**Sirianne Dahlum**

University of Oslo

[sirianne.dahlum@stv.uio.no](mailto:sirianne.dahlum@stv.uio.no)

**Carl Henrik Knutsen**

University of Oslo

[c.h.knutsen@stv.uio.no](mailto:c.h.knutsen@stv.uio.no)

**Tore Wig**

University of Oslo

[tore.wig@stv.uio.no](mailto:tore.wig@stv.uio.no)

[www.sciencespo.fr/liepp](http://www.sciencespo.fr/liepp)

© 2016 by the author. All rights reserved.

# Who Revolt? Empirically Revisiting the Social Origins of Democracy

Sirianne Dahlum<sup>†</sup>, Tore Wig<sup>†</sup>, and Carl Henrik Knutsen<sup>†</sup>

<sup>†</sup>Department of Political Science, University of Oslo

October 21, 2016

## Abstract

Different theories of the origins of democracy propose that collective action by particular urban social groups spur democratization. But, recent large-n studies on mass opposition movements and democratization notwithstanding, systematic and direct large-n evidence on this is missing. We further develop the argument that both industrial worker and urban middle class movements are conducive to democratization, albeit for slightly different reasons. We thereafter conduct the first large-n empirical study of its kind on this issue. To this end, we collect new data on social composition of anti-regime opposition movements, globally, for 1900–2012. Results clearly show that movements dominated by one of the two urban groups are associated with increased democratization prospects, both when compared to other movements and to situations without any organized mass opposition movement. We conduct various tests to assess sensitivity to, e.g., measurement strategy, geographic and temporal scope, model specification, and to accounting for the endogeneity of opposition campaigns. The relationship between industrial worker campaigns and democratization is particularly robust, but there is also evidence that middle class campaigns matter. Further tests suggest that these campaigns, as expected, relate more strongly to democratization in urban than in rural societies.

# 1 Introduction

The relationship between the social profile of opposition movements and democratization has received much theoretical attention from comparative social scientists of different stripes, and has been scrutinized in numerous case- and small-n comparative studies (e.g., Moore, 1966; Luebbert, 1991; Rueschemeyer, Stephens and Stephens, 1992; Collier, 1999). A more recent large-n literature also exists on mass opposition movements and democratization, but it focuses mainly on tactical choices within these movement, their strategic interaction with the regime, and the macro-structural conditions enabling them (e.g., Chenoweth and Stephan, 2011; Celestino Rivera and Gleditsch, 2013). Systematic large-n empirical studies that directly link the social composition of anti-regime movements to regime outcomes, however, are missing. This leaves some question marks concerning the generality of extant case-based evidence, as well as the direct relevance of large-n studies on democratization employing distal proxies of the preferences and/or capacities of various social groups, such as urbanization, GDP per capita or (land or income) inequality (e.g., Lipset, 1959; Boix, 2003; Ansell and Samuels, 2014).

We further develop the argument that social class composition of mass opposition movements affects democratization. More specifically, we propose that industrial workers *and* the urban middle classes should – at least in fairly urban and industrialized (or even industrializing) societies – have *both* the requisite motivation and the capacity for effectively bringing about democratic regime change. Other social groups often lack either in capacity (e.g. peasants) or motivation (e.g. military officers). To test this, we collect a novel dataset on the social composition of opposition movements, and provide the first large-n study of its kind on consequences for democratization. Hence, we introduce a new perspective to the emerging empirical literature on opposition movements against autocracies, and provide fresh evidence bearing on some of the most widely celebrated theories in the macro-comparative social science literature.

To be more specific, we collect data on presence and level of participation of 6 major social groups in about 200 anti-regime opposition campaigns, as identified in the extant NAVCO dataset (Chenoweth and Stephan, 2011), from 1900–2012. With these data, we first probe the more general proposition that opposition movements constituted by social groups located in urban centers,

and empowered by industrialization, are conducive to democratization. Second, we conduct tests distinguishing between opposition movements dominated by industrial workers and by the urban middle classes. We employ different relevant specifications and estimators to carefully assess robustness, and our analyses comprise both cross-section models at the opposition-campaign level and cross-section time-series models. The latter not only enable us to control for country-and year-fixed effects, but also to compare “urban” opposition movements both with other movements as well as situations in which there is no organized mass opposition movement.

We find that anti-regime opposition movements are far more likely to induce democratic outcomes – particularly in societies that display fairly high degrees of urbanization – if they are dominated by urban middle classes or by industrial workers. Opposition movements dominated by other groups, such as peasants, are far less likely to induce democratization – indeed, they do not even seem to raise the likelihood of democratization relative to situations without any opposition mobilization. This finding is extremely robust, and holds for instance when controlling for other features of the opposition movement, such as its size, or when accounting for urban opposition movements being more common in certain countries and time periods. It also holds up in instrumental variable and other models accounting for the potential endogeneity of opposition campaigns and their social profile (e.g. because some groups are more likely to bandwagon on movements perceived to have a greater chance in successfully inducing democratization).

The further analysis separating between urban middle classes and industrial workers speaks to a long-standing debate in the democratization literature: While scholars with quite different theoretical leanings propose that urban social groups may engender democracy, there is much more disagreement on which particular group is the most relevant. Moore (1966) proposes that revolutions dominated by the *bourgeoisie* are more conducive to democratic outcomes than other revolutionary movements, and numerous other contributions (e.g., Lipset, 1959; Ansell and Samuels, 2014) consider the urban middle classes as the key agent of democratization. In contrast, scholars such as O’Donnell (1973), Rueschemeyer, Stephens and Stephens (1992) and Collier (1999) argue that the industrial *working class* – due to clearer preferences for mass democracy and strong organizational capacities – is typically the main driver of democratization. While we do find

evidence suggesting that campaigns dominated by both groups are conducive to democratization, our results are much more robust for opposition movements dominated by industrial workers.

In the next section we provide a review of different strands of relevant literatures. After that, we further elaborate on why both industrial workers and urban middle classes (often) have the requisite motivation and capacity to effectively engender democratization. Thereafter, we present the data on opposition movements, and discuss issues of coding protocol and reliability and validity. Next we discuss results from, first, our cross-sectional campaign-level analysis and, second, our core country-year panel specifications. Before concluding, we evaluate different potential sources of bias and introduce relevant additional tests to assess the sensitivity of our core results.

## **2 Argument**

### **2.1 Literature review**

Democratic transitions often emerge from organized mass movements forcing the incumbent regime from power (either by violent or non-violent means). These are often referred to as “democratic revolutions”. In recent decades, organized collective action based in one or more broad social groups has been instrumental for the democratization of countries in Latin America (Collier, 1999), Eastern Europe (McFaul, 2002), Sub-Saharan Africa (Bratton and van de Walle, 1997) and even the Middle-East (Tunisia, Stepan, 2012). Democratization can also be the end-result of other processes, including military coups (Powell, 2012), negotiated transitions 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 by the existing regime (Acemoglu and Robinson, 2006). Still, even in many such cases the outcomes are preceded, and likely even be 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 even increase chances of a coup d’état against the old regime (Casper and Tyson, 2014). The share of regime changes stemming from actions by broad popular movements also seems to have been increasing (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”.

Yet, some opposition movements are successful in overthrowing authoritarian regimes and install democracy – or, at least force incumbent regimes to partly liberalize – whereas others are not. Why is this so? Relying on a recent, extensive dataset on major protest movements (Chenoweth and Stephan, 2011), an emerging literature has offered systematic investigations of opposition movements and their effects, also on democratization. These studies have mostly focused on tactical choices made by opposition movements (particularly use of violence), and their cooperation with external agents or strategic interactions with the regime. A prominent finding is that *non-violent* mass mobilization, in particular, is an effective democratizing force (e.g., Chenoweth and Stephan, 2011; Chenoweth and Cunningham, 2013; Celestino Rivera and Gleditsch, 2013; Bayer, Bethke and Lambach, 2016). Other studies have suggested that structural factors, such as industrialization, help explain features and outcomes of social movements (Butcher and Svensson, 2014; Chenoweth and Ulfelder, 2015).

Yet, we still do not know whether the *social background* of those mobilizing matters. A promising set of clues to answering this question comes from the theoretical and case-study literature on the social composition of revolutionary movements. Marx & Engels (2010/1848) famously claimed that “liberal bourgeois democracy” is brought about by the capitalist bourgeoisie revolting against the old feudal order, while the next phase of revolution is acted out by the working class, turning against the bourgeoisie. While this Marxist theory no longer dominates democratization scholarship, one key notion has survived: Economic modernization and industrialization is widely presumed to empower and articulate the interests of “new” social groups, such as the urban middle class and industrial workers, and these groups have an interest in democratic change. Where scholars part ways, however, is in their favored candidate for “most important group” in mobilizing for democracy.<sup>1</sup>

Many contributions – some subsumed under the heading of “modernization theory” – highlight that industrialization and urbanization strengthen the *urban middle class*, which will be both

---

<sup>1</sup>Some contributions, such as Luebbert (1991), 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 its participants. We test for the importance of cross-class coalitions below, finding clear evidence that broader coalitions, as such, enhance democratization.

capable of, and interested in, bringing about democratic regime change (e.g., Lipset, 1959; Glaeser, Ponzetto and Shleifer, 2007).<sup>2</sup> One stream of research, spearheaded by Inglehart and Welzel (2006), suggests that the middle class – due to favorable educational and material circumstances – should develop particular values, resulting in a strong desire for democratic institutions. Pursuing a quite different argument, Ansell and Samuels (2014) highlight that urban elites and middle classes prosper from industrialization, increased trade, and wider economic development, and that democracy provides an institutional framework for protecting their property against predation by “old” elite groups. They present indirect evidence for the hypothesis that strong urban middle classes induce democratization, showing that income inequality – taken as a proxy for middle class strength – is related to democratization. Using different proxies for middle class strength on a panel of developing countries from 1985–2013, Chun et al. (2015) also find that strong middle classes promote democratization.

Alternative arguments contend that the industrial working class is the main agent of democratization, simultaneously highlighting that urban upper- and middle classes, in certain contexts, have direct incentives to ally with other elites and work against democratization (see O’Donnell, 1973).<sup>3</sup> Notably, Rueschemeyer, Stephens and Stephens (1992) argue that large numbers and organizational strength will give industrial workers leverage under a democratic system (something the success of labor parties testaments to), providing strong incentives to fight for the introduction of mass democracy. In contrast, the bourgeoisie often lacks the incentives to expand full democratic rights, including suffrage, to the poorer segments of the population, whereas peasants do not have the capacity to organize a concerted effort challenging 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 proposes a key role for the working class in en-

---

<sup>2</sup>While we simplify and treat them as one group (“urban middle classes”) below, there may be disparities in income/wealth and in political incentives between urban capital owners, urban professionals and urban 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 redistributive and other policy interests of these urban actors often converge.

<sup>3</sup>Also contrasting with the focus on urban middle classes, different “redistributivist theories” suggest that less wealthy actors have the strongest incentives to mobilize for democratization (e.g., Acemoglu and Robinson, 2000, 2006; Boix, 2003). However, the relevance for industrial workers has been challenged on grounds that industrial workers typically earn more than the median income earner (Ansell and Samuels, 2014).

gendering democratization. Extant large-n studies finding that franchise extensions often followed revolutionary threats also refer to threats of revolutions by industrial workers, more specifically, as precipitating democratization in the 19th and early 20th century (Aidt and Jensen, 2014; Przeworski, 2009). Still, there is no direct measurement of the worker-dominance of revolutions in the latter analyses.

More generally, despite the many excellent contributions reviewed above, the nature of the extant evidence means that we do not yet *know* whether it matters for democratization, in general, *which particular social group* is pushing for change. Empirical studies that deal *directly* with the role of social groups are mainly comparative-historical case studies. Findings vary substantially, and seem sensitive to the choice of cases. For example, Moore’s conclusions on democratic outcomes being indebted to the *bourgeoisie* draws on contrasting 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 have also considered the role of social groups, but then only *indirectly*, by relying on macro-structural features of the economy (such as GDP per capita or economic inequality) to proxy for the interests and/or capacities of particular groups (e.g., Lipset, 1959; Ansell and Samuels, 2014; Boix, 2003). While often plausible, this approach also faces pitfalls. Crucially, there is no neat mapping from socioeconomic structure to social group mobilization (nor to group interests or capacities), and proxies used are, at best, very noisy measures. To exemplify, measures of urbanization and the industrial sector’s share of GDP do not separate well between the strength of urban middle classes and industrial workers.<sup>4</sup> This approach also misses that hurdles for collective action may vary by group (and across contexts) – peasants may face higher hurdles for organizing than industrial workers – making it problematic to equate economic structures with the agency of particular groups.<sup>5</sup>

---

<sup>4</sup>To take another example, Ansell and Samuels (2014) argue that while previous literature has assumed that high inequality entails that workers and middle classes should prefer redistribution (and thus democracy as an institutional guarantee) from above-median income groups to below-median income groups, this assumption does not withstand scrutiny. Workers and urban middle classes are often located well above the median of the income distribution.

<sup>5</sup>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 arguably face issues also faced by 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.

## 2.2 Urban groups as agents of democratization: Motivation and capacity

In this section, we synthesize arguments from the reviewed literature and further clarify why movements dominated by industrial workers *or* by the urban middle classes should be conducive to democratization. We shall assume that a social group's potential for bringing about democracy through mobilizing a mass movement is a function of the group's *motivation* and *capacities*. For successful 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 face a favorable opportunity structure which is, in no small part, a function of the group's ability to organize effective collective action. We first discuss motivation before turning to capacity aspects.

Some social groups may simply prefer democratization more strongly than others. One straightforward reason, which figures prominently in the literature, is differential economic benefits from democratization. One influential argument is that relatively poor but numerically strong groups should 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. Groups that are (typically) less numerous and wealthier, such as landowners or military officers, should be less likely to support democracy according to this logic.

While redistributive models of democratization have received their share of criticism, also on empirical grounds (e.g., Houle, 2009; Teorell, 2010), we can relax the assumption of redistribution as the key motivational force for democratization, but maintain the assumption that numerosity is important: If we assume, first, that individuals with similar socioeconomic features (type of occupation, education, and income, etc.) are more likely to hold converging preferences also over other policies than 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 democracy. Simply put, democracy presents large groups with the institutional framework for enacting legislation that they favor (in different policy areas).<sup>6</sup>

---

<sup>6</sup>This argument is simplified and not comprehensive. Social groups might systematically differ in their preference for democracy also for other reasons than merely instrumental. Modernization theoretic approaches (e.g., Inglehart and Welzel, 2006; Welzel, 2013) suggest that the material security of the middle class – at least in relatively wealthy countries – may generate *inherent* normative preference for more liberal institutions and policies. It presumably

In static economies, social groups with converging preferences and numerical majority should prefer democracy. Industrial workers and urban middle classes, for example, should have a clear preference for democracy in highly urbanized, industrial societies, but not necessarily in predominantly agrarian, rural societies. Still, in a dynamic economy – e.g. characterized by rapid industrialization and urbanization – numerically large but declining groups, such as peasants, should *not always* prefer democracy. In contrast, groups on the rise, such as industrial workers, might. When a social group expects to become a minority in the near to medium term, majority rule may imply that policies will be determined by other groups in the future. Thus, urban middle classes and industrial workers should 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 for different countries – and processes of de-industrialization have set in for some richer countries in recent decades – urbanization processes have been widespread especially from the 1950s in most regions (Satterthwaite, 2007), and have yet to show signs of reversal. Thus, urban groups, at least in many developing countries, may expectedly have turned more motivated to pursue democratization during the latter half of the 20th century.<sup>7</sup>

Regarding the *capacity* of a social group to bring about democratization, this relates both to abilities to *co-ordinate* collective action and to more general *leverage*. Leverage refers to the power resources that a group can draw on in order to inflict costs on the regime, and thereby pressure the regime to achieve its demands (for instance political liberalization). This relates, for instance, to the ability to grind the economy to a halt (thereby reducing tax revenues), e.g. through moving capital assets abroad or through carrying out strikes in vital sectors. Groups will differ in their ability to inflict such economic damage, and thus presumably in their abilities to force regime change. Other sources of leverage are easy access to weapons and manpower with relevant training, or having a militant ideology that can motivate recruits.

---

follows that middle class opposition campaigns should, everything else equal, display a particularly strong relationship with subsequent democratization.

<sup>7</sup>While patterns of relative numerosity and expected growth vary according to 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, and the middle classes could fit better in later stages of industrialization and de-industrialization with labor-saving technological advances and high-income service sector jobs on the rise.

Having the ability to organize large-scale collective action, such as riots or mass protests, is also vital for bringing about democratization. Large-scale collective action can result in democratization through various paths, as noted in the literature review. First, 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; Wig and Rød, 2014). Second, it may do so by “persuading” the incumbent to liberalize, or else s/he will face even worse consequences (see, e.g., Acemoglu and Robinson, 2006). The key question then becomes: Which features determine a group’s ability to coordinate large-scale collective action? We surmise that this relates especially to the ability to create permanent and streamlined organizations (with some form of centralization of decision making power). Such organizations help with alleviating collective action problems among group members by, e.g., providing tools for effective information and dispersing side-payments for participants (Olson, 1965). But, creating organizations also help with recruiting new individuals for the cause, networking with foreign actors, experimenting with and (over time) 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 different groups, industrial workers should 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 in autocracies banning them), giving them a potent institutional infrastructure for successfully orchestrating mass resistance. Workers often also have expansive international networks, either formally through international labor organizations or the Socialist International, or informally through more amorphous cross-border networks in the labor movement. Labor unions, for instance, are also known for enforcing strict discipline and organization on their members (see, e.g., Olson, 1965). Regarding the leverage aspect, 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. Second, workers will also have fairly high military potential, due to their supreme organizational skills, foreign networks, and historically often being related to explicitly revolutionary (and sometimes

violence-condoning) ideologies (see, e.g., Hobsbawm, 1974).

Similarly, urban middle classes should have potent sources of both organizational capacity and strategic leverage over authoritarian regimes. First, urban middle classes will often include large a pool of individuals with organizational skills and resources, equipping them to overcome collective action problems. This is reflected in the large number of civil society organizations, liberal parties, professional associations, student organizations and educational institutions that are often associated with the urban middle classes. As for industrial workers, urban middle classes will also often have access to international networks that enable successful collective action, examples being student networks, human rights advocacy organizations, and international media. Regarding leverage, urban professionals often occupy vital inflection points in the economy, example in vital service sectors such as finance. They might also be particularly skilled when it comes to putting pressure on the regime through non-violent conflict; utilizing comparatively high levels of human capital to legitimate their movement and strategically targeting the regime's soft spots (see, e.g., Dahlum, 2015).

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

In sum, we surmise that social group must have both a clear preference for democracy *and* the

capacity to remove the incumbent autocratic regime in order to be an effective agent of democratization. Some groups, such as peasants, may in some contexts have the requisite motivation, but lack in capacity. Other groups, such as landowners or military officers, are strong in terms of capacity, but may lack the motivation to pursue democratization. In contrast, industrial workers and urban middle classes should – especially in fairly urban, industrial societies – display the required combination of capacity and motivation. At a more detailed level, the two groups may differ regarding what drives their preference for democracy, and particularly their sources of organizational capacity and leverage. Nevertheless, we expect both groups to be strong candidates to lead opposition campaigns that induce democratizing regime changes.

### 3 Data

#### 3.1 Social groups and opposition movements

In order to systematically code the social profile of *anti-regime opposition movements* we identify our units using the recent NAVCO dataset (Chenoweth and Lewis, 2013). NAVCO describes various properties of (fairly large-scale) such opposition 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 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 inclusion of various – and even some partly overlapping – categories. We collect data on the presence and participation of different social groups in all campaigns that

NAVCO identifies as aiming for regime change or “other goals”, including policy concessions or political liberalization.<sup>8</sup> There are 203 relevant campaigns in NAVCO, and we have collected information on social composition for 193. The 10 remaining are missing due to no – or too uncertain – information. We discuss the categorizations in Appendix A, but the core categories are Peasants; Public sector employees; Military employees; Religious or ethnic groups; Industrial workers; and, Urban middle classes.

We here restrict our focus to *Industrial workers* and *Urban middle classes*. Campaigns “associated with” (as further specified below) all other groups are therefore considered as “other campaigns” (see Appendix Table C.12 for regression models distinguishing between these “other” campaigns). Industrial workers refer to labor employed in the manufacturing and mining sectors. The Urban middle classes category is complex, and we have coded this to encompass business elites, smaller merchants, professionals (lawyers, doctors, etc), and students. Importantly, we code and treat public sector employees as separate from urban middle classes. First, these are often mentioned separately in the sources used. Second, and 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. However, this choice is surely debatable, and we thus run analyzes with and without public sector employees included in the urban middle classes category (results turn out very similar, as displayed in Table 6).

For each social group, we code involvement in protest campaigns. The detailed coding rules and procedures are discussed in Appendix A. But, in brief, we have drawn on different sources, including some covering multiple movements and countries, notably 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), but we also draw information from (cross-country and country-specific) secondary sources where necessary. In many sources, our social-group categories are often explicitly listed; a source might, e.g., explicitly note that a movement was dominated by industrial workers, and that peasants also participated.

---

<sup>8</sup>This means that we have not coded social group characteristics for secessionist movements. We consider secessionist movements to be qualitatively different, and not obviously relevant for our research question.

Specifically, we code *three* indicators representing different degrees of involvement. First, we code whether the social group had *some* level of participation; i.e., we register whether the movement at some point in time *consisted partly* of members of the social group. Operationally, for this “low-threshold” dummy to be scored 1, the social group must at least be mentioned in one source. This is fairly easy to gauge, especially since the mentioned databases include lists of social-group participation where our categories are included.<sup>9</sup> Second, we code whether a movement *originated* among members of the given social group, scored as such if mentioned explicitly in historical sources. Third, we code whether the opposition movement is *dominated* by the group. Such instances include when a social group include makes up a majority of campaign members or if it, according to sources, was highly influential for the movement. While the latter obviously requires subjective evaluation, we carefully streamlined the coding through different procedures, as clarified in the Appendix. In general, we opted to be conservative when setting the threshold for “highly influential”. For all variables, our coders (RAs) have explicitly assessed their degree of certainty for each coding, ranging from very uncertain to very certain. Thus, we can run tests to check whether results are vulnerable to excluding the more uncertain cases (see Table 6; we remind that the most uncertain cases are always treated as missing).

One challenge is that source coverage might correlate with the outcome variable. While the data only include campaigns with at least 1000 participants, that are “observable”, “continuous” and have a “discernible leadership” (Chenoweth and Lewis, 2013) – we thus mostly deal with campaigns that are well covered, this can not be excluded. Nonetheless, different assessments provide room for optimism: For instance, if data availability and quality is systematically poorer for violent campaigns (often smaller and carried out in poor, remote areas), which extant studies have shown 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. Meanwhile, the correlation between uncertainty status and participation by urban middle classes ( $r=-.13$ ) or industrial workers ( $r=-.06$ ) are also low. Without downplaying reliability and validity issues by too much, we find it

---

<sup>9</sup>We emphasize that we do not interpret a 0 on this measure to mean that *no* single individual from the social group partook. Rather, the presence of the social group was not big enough to be recognized in the source materials.

Table 1: Descriptives for Urban middle classes (MC) and Industrial worker (IW) campaign profile variables

	Sum	N	Share
MC dominates	41	187	0.22
MC participates	148	188	0.79
Originated among MC	37	180	0.21
IW dominate	29	192	0.15
IW participate	116	186	0.62
Originated among IW	19	192	0.10

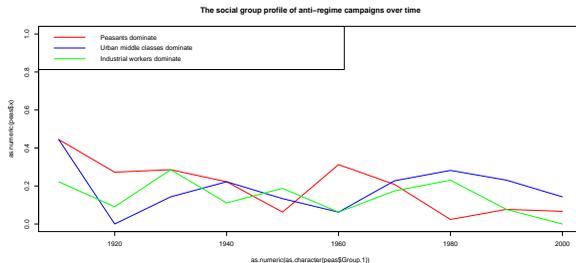
Note: Number of cases coded varies between variables due to varying levels of uncertainty. Highly uncertain cases are treated as missing.

unlikely that measurement error issues should drive the robust results presented below.

Let us briefly illustrate the coding scheme with two examples (see Appendix for further illustrations). First, the Velvet Revolution against the Czechoslovakian Communist regime in 1989 is described in the identified sources as a broad-based movement consisting of urban middle classes, industrial workers and peasants. This campaign is thus coded as consisting of all these three groups. According to, for instance, the Swarthmore database, the movement was sparked off by student demonstrations on International Students Day in November 1989, which soon developed into larger-scale demonstrations 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” of the movement. The movement is therefore coded as having originated among, *and* as being dominated by, the urban middle class. Second, the Senderista Insurgency (Sendero Luminoso) against the Peruvian government from 1980-1999 is described by several sources (e.g., Swarthmore, 2015) as being 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.

Table 1 shows descriptives for the set of binary variables (*originated; dominated; participated*)

Figure 1: Social group profile of opposition movements over time. Campaigns dominated by peasants, urban middle classes, and industrial workers.



for urban middle classes (MC) and industrial workers (IW). Appendix B provides additional descriptive statistics. Overall, industrial workers participated in 62 percent of our registered campaigns, whereas urban middle classes participated in 79 percent. However, only about 1/5 campaigns were dominated by the urban middle classes, and likewise for industrial workers. Figure 1 charts the share of ongoing campaigns dominated by a given social group – we here include peasant campaigns for illustration – as distributed by decade. While the shares fluctuate, the figure reveals that the number of peasant-dominated protests went down towards the end of the 20th century, and that industrial-worker and urban-middle class dominated campaigns increased. This is not surprising as ongoing processes of urbanization and industrialization should empower these urban groups in many countries.

### 3.2 Democracy

For our dependent variable, we draw on two (quite different) measures of electoral democracy with extensive coverage. First, we code democratization using the binary democracy measure from Boix, Miller and Rosato (2012) (BMR). BMR basically registers the presence of contested and “free and fair” elections, while also requiring that 1/4 of citizens are enfranchised. Due to its binary nature, democratization (from 0 on BMR in  $t$  to 1 in  $t + 1$ ) will mainly pick up major regime changes, and – as we will elaborate on below – it may miss out on non-trivial episodes of “liberalization” in nominally autocratic regimes and “democratic deepening” in regimes that

already pass the democracy threshold.

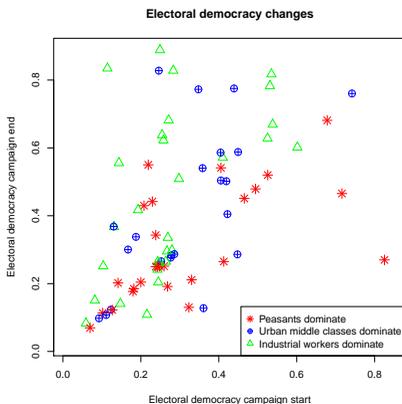
Thus, our main measure – allowing us to capture more incremental and less rapturous democratizing changes, including many franchise extensions – is the continuous Electoral Democracy measure, also called Polyarchy, from Varieties of Democracy (V-Dem) (Coppedge et al., 2015). The theoretical range is from 0–1 (0.01–0.95 in the data). Polyarchy includes indicators on whether or not the chief executive is (directly or indirectly) elected, the extent to which elections are “clean”, the 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. Hence, it mainly caters to the conditions for elite competition through multi-party elections and the right to participate in these elections.

## 4 Empirical analysis

Figure 2 shows the pre- and post-movement levels of Polyarchy – measured at the campaign’s start and end years – for campaigns dominated by different groups (we include peasants for illustration). We note that cases on/above/below the diagonal indicate, respectively no/positive/negative change in Polyarchy. Most urban middle class-dominated and worker-dominated campaigns are associated either with no change or with democratic improvements. Indicatively, the three campaigns associated with the greatest increases in Polyarchy score in our sample 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 notably associated with democratic backsliding.

Still, we cannot draw any inferences about social groups and democratization from this. Different types of groups may dominate in very different societies, and campaigns might just happen to co-exist with changes in democracy without playing any systematic causal role. Thus, we probe 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 provide estimates of causal effects. While often implicit, the notion of such an effect also underpins the reviewed literature linking social groups to democracy. Democratization results

Figure 2: Opposition movements dominated by industrial workers, urban middle classes, and peasants, and changes in Electoral Democracy Index/Polyarchy (V-Dem)



*because* specific social groups mobilize.

As with other macro-level large- $n$  studies, identifying effects with a fair degree of certainty is challenging. Here, likely sources of endogeneity bias include particular social groups self-selecting into particular movements. If certain groups have a stronger *preference* for democracy, this may lead them to self-select into movements that are more likely to engender democratization (with or without their participation). Second, there are several plausible confounding variables. For example, if the majority of a country’s citizens are industrial workers, an observed movement will likely be dominated by this group (the probability approaches 1 when group size approaches the total population). Further, 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 might induce 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 “quick fix” to handle all potential issues in one single model. Therefore, we employ a series of slightly different approaches and designs that, when taken

together, help shore up confidence in the causal interpretation of our results.

#### 4.1 Cross-sectional analysis at the campaign level

In order to first describe associations between opposition movement profile and post-movement regime outcome, we start by investigating a simple cross-section of opposition campaigns. Results pertaining to the broader question of whether campaigns dominated by (*either* of the two) urban groups differ from other campaigns are presented in the four left-most columns of Table 2. Model 1 employs a dummy of democratization as dependent variable – scored 1 if the country was coded autocratic by BMR at campaign start and democratic one year after campaign end – and thus only include observations that were autocratic at campaign start. 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, see Appendix Table C.2). The coefficient for urban group dominated campaign (*UG*) in Model 1 suggests that the probability of observing democratization increases by 18 percentage points when comparing a campaign dominated either by industrial workers or urban middle classes to other campaigns. This is substantial, particularly given that overall probability of democratization in the sample of 140 campaigns (in autocracies) is 27 percent.

This association could, however, be due to industrial worker or middle class 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, PP-adjusted) GDP per capita, and urbanization, all registered at campaign outset and collected from V-Dem (see Coppedge et al., 2015). At the campaign-level, we control for ln participants in campaign – urban campaigns may enhance democratization simply because they happen to be larger. (But, we remind that campaign size can be considered an inherently relevant feature, as our theoretical argument on capacity to organize collective action suggests, which is why we drop it in our main models). *UG* drops – and loses statistical significance, with  $t = 1.5$  – when conditioning on these covariates in Model 2, although urban campaigns are predicted to be followed by a 12 percentage point higher probability of observing democratization when holding campaign size, population size, income and urbanization level constant. Results are strengthened, and weakly

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

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 (UG)	0.183** (2.59)	0.116 (1.46)	0.088*** (3.03)	0.062* (1.97)				
Industrial workers dominate (IW)					0.213** (2.39)	0.151 (1.59)	0.126*** (3.30)	0.101** (2.53)
Middle class dominate (MC)					0.118 (1.41)	0.089 (1.01)	0.033 (0.97)	0.023 (0.66)
Ln campaign participants		0.000 (0.10)		0.000 (0.50)		0.000 (0.18)		0.000 (0.60)
Urbanization		0.514* (1.77)		0.211* (1.85)		0.511* (1.77)		0.204* (1.80)
Ln population		-0.009 (-0.38)		-0.001 (-0.12)		-0.008 (-0.31)		-0.001 (-0.11)
Ln GDP per capita		0.052 (0.77)		0.035 (1.28)		0.046 (0.68)		0.033 (1.24)
Lagged dependent variable			0.666*** (8.18)	0.470*** (5.08)			0.677*** (8.38)	0.494*** (5.32)
Constant	0.135*** (3.42)	-0.253 (-0.45)	0.139*** (4.84)	-0.117 (-0.52)	0.131*** (3.40)	-0.248 (-0.44)	0.136*** (4.80)	-0.116 (-0.52)
N	140	122	176	155	139	121	175	154
R <sup>2</sup>	0.046	0.146	0.307	0.345	0.073	0.165	0.329	0.364

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

significant, for this (and other) model(s) when measuring democratization in the somewhat longer term, including transitions up to 5 years after campaign end (see Appendix Tables C.1 and C.2).

As expected, results are also much clearer when employing the more fine-grained Polyarchy measure from V-Dem, also allowing us to capture more incremental changes in countries that remain autocratic and those that are already minimum democracies. To consider correlates with changes in democracy, we condition on the lagged dependent variable (i.e., we consider correlates of post-campaign democracy, holding constant the level of democracy at campaign start). Model 3 includes 176 campaigns, and reveals a highly significant *UG* coefficient ( $t = 3.0$ ). The t-value decreases, but only to 2.0, when conditioning on the above-mentioned campaign- and country-level covariates. Holding these factors constant, urban campaigns are predicted to be associated with an additional 0.06 increase in Polyarchy compared to other campaigns. For illustration, this is comparable to the change observed in Ukraine from the start of the Orange Revolution to the year after it finished (0.05).

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 (*MC*) dominated campaigns. Results are clear on the association between *IW* and democrati-

zation, especially when employing Polyarchy. *MC* coefficients are comparatively smaller (about half of *IW* in BMR models and about one quarter in Polyarchy models), and t-values range from 0.7–1.4. This provides an initial suggestion that the hypothesis concerning industrial worker campaigns leading to more democratic outcomes finds stronger support in the data than the corresponding hypothesis for urban middle class campaigns. This echoes the arguments and historical case-based evidence in, for instance, Rueschemeyer, Stephens and Stephens (1992) and Collier (1999).

## 4.2 Cross-section time-series analysis: Benchmark model

While indicative, campaign-level cross-section models are not fully adequate for answering the question of whether campaign social profile matters. First, they do not include information on years without any campaign, disallowing comparisons between campaigns with a given social profile and situations where no campaign exists. Second, since campaigns with specific profiles presumably relate differently to background covariates, including information on covariates from additional years should improve 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:

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

$DEM_{i,t+1}$  represents democracy led by one year ( $t + 1$ ) in country  $i$ . We also test and discuss models where change in democracy level from  $t$  to  $t + 1$  is the dependent variable, including models only counting positive/democratizing changes, and results are robust. *IW* and *MC* register the presence of an ongoing campaign dominated by, respectively, industrial workers and urban middle classes. (Alternatively, we substitute these two dummies with the combined *UG* dummy). *OC* registers the presence of all other campaigns, i.e. those dominated by other groups such as peasants,

Table 3: Core cross-section time 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	$\Delta$ Polyarchy	Upturns Polyar.	BMR democratization		
IW or MC dominate	0.047*** (3.98)				0.850 (1.36)	1.381*** (2.75)	
IW dominate		0.049** (2.55)	0.049** (2.55)	0.044** (2.59)			0.973** (2.28)
MC dominate		0.034* (1.86)	0.034* (1.86)	0.036** (2.22)			1.450*** (3.68)
Other campaign (OC)	-0.002 (-0.58)	-0.003 (-0.73)	-0.003 (-0.73)	0.006** (2.12)	0.916* (1.83)	0.796** (2.13)	0.777** (2.25)
Ln GDP p.c.	0.002 (0.78)	0.002 (0.74)	0.002 (0.74)	-0.001 (-0.33)	1.026* (1.89)	0.514** (2.23)	0.401* (1.78)
Ln Population	-0.001 (-0.24)	-0.001 (-0.23)	-0.001 (-0.23)	0.003 (1.32)	-0.874 (-0.88)	-0.023 (-0.26)	0.004 (0.05)
Urbanization	0.022** (2.14)	0.022** (2.16)	0.022** (2.16)	0.021** (2.00)	-5.011 (-1.56)	-0.190 (-0.21)	0.358 (0.43)
Lagged Polyarchy	0.948*** (166.15)	0.948*** (169.73)	-0.052*** (-9.27)	-0.037*** (-7.37)			
Country dummies	Y	Y	Y	Y	N	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y
N	9739	9738	9738	9738	2763	3149	4698
Countries	147	147	147	147	68	121	121
Max time series	1900–2010	1900–2010	1900–2010	1900–2010	1900–2010	1900–2010	1900–2010

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

or where no group clearly dominates. We also include level of democracy at  $t$  ( $DEMI_{i,t}$ ), and vectors of country-year covariates  $\mathbf{X}$ , country-fixed effects  $\zeta$ , and year-fixed effects  $\theta$ . Errors are clustered by country to account for panel-level autocorrelation and heteroskedasticity.

Table 3 presents two versions of our benchmark fixed effects (FE) OLS model with Polyarchy as dependent variable. These models draw on almost 10 000 observations, taken from 147 countries and maximum time series from 1900–2010. Model 1 employs the cruder  $UG$  dummy, whereas Model 2 rather includes  $IW$  and  $MC$ . The inclusion of  $OC$  implies that the “reference category” is no ongoing opposition campaign.

According to Model 1, an ongoing opposition campaign dominated by one of the two urban groups systematically increases democracy in the following year compared to when there is no ongoing opposition campaign.  $UG$  is highly significant ( $t = 4.0$ ), and predicts that an urban group-dominated opposition campaign is associated with a Polyarchy score that is about 0.05 points higher in the following year. Since  $OC$  is negative, but close to 0 (and  $t = 0.6$ ), opposition campaigns dominated by urban groups are systematically associated with democratic change also when compared to situations with other ongoing opposition campaigns. In Model 2, the  $IW$  coefficient is

virtually similar (0.49 instead of 0.47) to  $UG$  in Model 1, whereas  $MC$  is somewhat smaller (0.34). Both coefficients are at least weakly significant, with t-values of 2.6 and 1.9, respectively. Hence, there is clear evidence for a positive association between industrial worker campaigns and democratization, and somewhat less clear evidence for an association between middle class campaigns and democratization.

Results are virtually 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 Appendix Table C.3). This is illustrated in Model 3, which is otherwise similar to the benchmark specification, where  $IW$  and  $MC$  are identical in size and significance to the respective coefficients in Model 1. 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. As results in Appendix Tables C.4 and C.10 reveal, results are robust and even clearer (for  $UG$ ,  $IW$ , and  $MC$ ) when considering only the relationship with democratic upturns. This is illustrated by Model 4, again a version of the benchmark specification (downturns are here scored as 0 on the dependent variable). Model 4 shows that both industrial worker campaigns and urban middle class campaigns are significantly related to subsequent increases in Polyarchy at the 5% level, and coefficients are fairly close to those found for Model 1. However, there are also indications that ongoing industrial worker campaigns, though not middle class campaigns, guard against subsequent democratic backsliding, though coefficients are much smaller and only weakly significant in some specifications. Hence, it is the association with subsequent democratization experiences that drive results in our benchmark specification. We return to further robustness tests of these relationships in the next section.

But first, Models 5–7 are logit equivalents of our benchmark model, using our dichotomous measure of democratization (transition from BMR dictatorship in  $t$  to BMR democracy in  $t + 1$ ) as dependent variable and only including autocratic observations (in  $t$ ). This reduces variation in the dependent variable. For instance, only 102 instances of democratization are included for Model 5. In practice, we now ignore any impact opposition campaigns may have on de-democratization

(which could come from opposition campaigns successfully implementing a new and more autocratic regime, or the extant regime cracking down on political rights in response to an ongoing campaign). We also ignore situations where institutions change in a more “liberal direction” without tilting regimes across the “minimum democracy” threshold (to illustrate, 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 that have passed the minimum threshold (to illustrate, in 1995, Colombia, Mali, and Pakistan had all passed this threshold). Indeed, additional estimates using V-Dem’s Polyarchy as dependent variable suggest that urban opposition campaigns may be associated with both liberalization in BMR autocracies and further “democratic deepening” in BMR democracies (see Appendix Tables C.6–C.9). Thus, these logit models may very well fail to pick up an effect, even if there is one.

Indeed, the fixed effects logit model (Model 5) fails to find a statistically significant  $UG$  at conventional levels ( $t=1.4$ ), although  $UG$  is positive and substantial in size. However, the issues discussed above are expectedly exacerbated by including country-fixed effects. As shown by Beck and Katz (2001), fixed effects logit models very often yield Type II errors when the dependent variable changes seldom or not at all (only 68 countries can inform estimates in Model 3). When omitting the country-fixed effects (but keeping 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 also to democratic transitions, as measured by BMR.

Still, the associations reported in Table 3 could, potentially, be sensitive to particular model specification- or measurement choices, or that opposition campaign profile is endogenous to anticipated regime changes. We thus turn to alternative tests.

### 4.3 Additional tests

Table 6 presents some core robustness tests, and several additional ones are reported in Appendix C. 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 *extremely* robust (see Appendix Table C.10). Hence, we mostly focus on tests that employ the two separate industrial worker and middle class dummies, allowing us to discuss the more nuanced propositions concerning specific types of opposition campaigns and democratization.

*First*, results are fairly robust to measuring the dependent variable with a longer lag (e.g., 5 or 10 years) than the baseline 1-year lag. This is exemplified by Model 1, Table 6, which employs a 2-year lag. Indeed, point estimates increase considerably from the baseline model – from 0.49 to 0.83 for *IW* and from 0.34 to 0.49 for *MC* – suggesting that a 1-year lag may be too short to capture the full effects of urban opposition campaigns on regime change. As in the baseline, *IW* is significant at 1%, and *MC* at 10%.

*Second*, we note that the weakly significant result for *MC* comes despite a substantially high point estimate. Our baseline controls for country- and year-fixed effects, and employs clustered errors. When combined with the relatively few campaign observations (88 middle class dominated campaign years), even in our extensive sample, this yields large standard errors. This is illustrated by Model 2, which employs classical rather than clustered errors. *MC* is now highly significant with  $t = 6.1$ . More generally, all robustness tests conducted on classical errors yield similarly strong results. The  $t$ -value of *IW* also increases dramatically, from 2.6 to 8.3. Yet, in the reminder, we employ the more conservative benchmark approach, clustering on country to, e.g., account for autocorrelation affecting standard errors.

*Third*, results in our benchmark could be afflicted by post-treatment bias, and we thus tested models omitting potentially “bad controls”. Controlling for urbanization, in particular, could arguably wash out parts of the theoretically relevant effect. As discussed, the capacity of industrial workers, for instance, to bring about democratic change could partly be attributed to workers living closely together in cities, easing coordination problems (this contrasts sharply with settlement patterns for peasants). However, more parsimonious models consistently show results similar to the benchmark. This is exemplified by Model 3, which excludes population, income and urbanization, where results are virtually identical to the benchmark.

*Fourth*, the opposite concern is that we are not including all relevant controls, and that results

Table 4: Robustness tests

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	2-year lag	No clustering	Parsimonious	Extra controls	Public employees	Nr groups	No uncertain	Campaign origin	Lexical indices
Ind. workers dominate	0.083*** (2.84)	0.049*** (8.26)	0.049*** (2.65)	0.049*** (2.73)	0.050*** (2.73)	0.043** (2.19)	0.053*** (2.62)		
Middle class dominate	0.040* (1.71)	0.034*** (6.11)	0.035* (1.92)	0.034* (1.82)	0.030** (2.18)	0.022 (1.17)	0.032 (1.58)		
Ind. worker origin								0.025 (1.22)	
Middle class origin								0.031*** (2.78)	
Ind. worker index									
Middle class index									
Other campaign	-0.004 (-0.66)	-0.003 (-1.10)	-0.002 (-0.76)	-0.019 (-1.29)	-0.003 (-0.84)	-0.030*** (-3.11)	-0.004 (-0.96)	-0.001 (-0.28)	0.017** (2.57)
Ln GDP p.c.	0.004 (0.74)	0.002 (1.03)		0.002 (0.74)	0.002 (0.70)	0.002 (0.76)	0.002 (0.68)	0.002 (0.68)	0.015*** (3.22)
Ln Population	-0.003 (-0.60)	-0.001 (-0.23)		-0.000 (-0.14)	-0.000 (-0.22)	-0.000 (-0.13)	-0.000 (-0.07)	-0.001 (-0.34)	-0.021*** (-3.57)
Urbanization	0.050** (2.33)	0.022** (2.47)		0.023** (2.20)	0.021** (2.08)	0.022** (2.08)	0.022** (2.12)	0.022** (2.13)	0.002 (0.70)
Lagged dep. var.	0.874*** (75.31)	0.948*** (272.97)	0.953*** (215.25)	0.949*** (168.97)	0.948*** (168.28)	0.949*** (165.85)	0.949*** (169.43)	0.947*** (171.28)	-0.000 (-0.14)
Ln camp. particip.				-0.000 (-0.31)					0.022** (2.07)
Regime violence				0.020 (1.45)					0.949*** (168.50)
Nr groups in campaign						0.010*** (2.92)			
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	9734	9738	15519	9738	9738	9628	9646	9744	9628

Notes: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01. T-values in parentheses. Polychaety 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.

are afflicted by omitted variable bias. In particular, other aspects of opposition movements (rather than *who* participates) may drive democratization. Hence, Model 4 controls also for the estimated number of people participating in the campaign, from NAVCO (log thousand participants). This could be an important confounder since the social profile of a campaign could be endogenous to how many people participate; if the whole population participates, for example, then a majority social group will *by definition* dominate. (But, it is also a likely post-treatment variable. The particular group organizing the campaign will expectedly affect how many people that take part, as suggested by our argument on social group capacity and, e.g., discussions on strong mobilizing ability of, for instance, industrial worker unions and labor parties.) Model 4 further includes whether the regime uses violence against protestors, since particular campaigns might “self-select” into mostly opposing less violent regimes. Nonetheless, adding these controls does not alter the estimated relationships between *IW/MC* and democracy.

*Fifth*, our baseline measure of urban middle class dominated campaigns does not include campaigns dominated by public sector employees. While counted as “middle class” in many categorizations schemes, we excluded public employees due to the expectation that they often have different interests and leverage over regime change, because of their relationship with the state. In our baseline sample, there 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 decreases slightly (from 0.34 to 0.30), but also becomes more precisely estimated. Indeed, it is now significant at 5%. *IW* remains highly significant also in this model.

*Sixth*, some social 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. As expected, this variable is positive, sizeable, and highly significant. While including this count variable weakens *MC* (0.22;  $t = 1.2$ ) – suggesting that urban middle class dominated campaigns, at least in part, are more successful because they are associated with broader coalitions of social groups – *IW* remains sizeable (0.43) and

significant at 5%. However, controlling for class coalitions may very well introduce post-treatment bias, which is why it is not included in our favored specification: If the capacity to build broad coalitions is an inherent feature of urban middle class actors, Model 6 underestimates the effect of *MC*.

*Seventh*, we want to ensure that measurement errors in our coding are not driving 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 have reported low certainty. Results are not changed by much when doing so, although *MC* just drops below the  $p \leq .10$  threshold. In contrast, the industrial worker results is actually strengthened, but only slightly, suggesting that the benchmark results do not hinge on cases with uncertain coding.

*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 ( $t = 2.8$ ), whereas *IW* now turns insignificant ( $t = 1.2$ ). However, the *IW* result is stronger in yet other specifications altering how the independent variables are constructed: We also run tests on the measures registering whether a given social group at least *participated*. These models mostly show strong results, particularly for industrial workers. In addition, we tested lexical scales (*MC* Index; *IW* index). These are scored as 3 if the given group dominated the movement, 2 if it originated among the group and the group participated, 1 if the group only participated, and 0 if the group did not have a recorded presence. These indices provides more nuanced, and arguably more information rich, measures of the role of the respective social groups in opposition movements, and the related results are strong for both industrial workers and middle classes. In Model 9, where the *IW* Index is significant at 5% and *MC* Index at 1%. Hence, prospects for a more democratic regime increase with how strongly industrial workers and urban middle classes are involved in an opposition movement.

Although controlling for, e.g., income level and time- and country-specific effects should mitigate many sources of bias, other sources could remain. To further investigate whether there *might* be causal effects of the involvement of urban groups in opposition campaigns on democratization, we estimate instrumental variable models where we endogenize the *IW* and *MC* profiles of the

campaigns. More specifically, we discuss IV models where we endogenize the lexical campaign-profile indices described above, in order to allow for maximum variance. However, results turn out to be equally strong for industrial workers and even stronger for the urban middle class when we employ the “campaign domination” measures from the benchmark as endogenous independent variables (see Appendix C.13).

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 in the country in question. 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 motivated by and learn from observing and communicating with their counterparts in neighboring countries. This is validated in string of studies on how social movements diffuse through mobilizing similar and (often) allied groups in closely related countries (see, e.g., Weyland, 2014, 2010, 2009). As expected, the first-stage F-statistics and t-values of these neighborhood instruments reveal that they are very strong.

To achieve consistent estimates we also need to ensure that neighbourhood campaign profile is not otherwise linked to democratic change in the country in question. While we do include all covariates from the benchmark (including country- and year-fixed effects), it may, e.g., be that neighborhood campaigns affect neighborhood democracy, which in turn affects democracy domestically. Therefore, in order to further mitigate concerns about the exclusion restrictions we also condition on *a)* the neighborhood-profile of other campaign variables in the model, and *b)* the neighborhood-profile on democracy. In these specifications, we thus assume that the social profile of a campaign *in the neighborhood – conditional on all the included covariates* – only affects democratic change *via* the campaign profile in country  $i$ .

The instrumental variable (FE 2SLS) results are presented in Table 5. Model 1 shows the second-stage regression treating the *IW* index as endogenous. It shows a positive and precisely estimated ( $t = 3.6$ ) effect on democratic change. 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

Table 5: Instrumental variable models (Fixed Effects 2SLS)

Stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage	1st stage
Dependent variable	Polyarchy	<i>IW</i> index	Polyarchy	<i>MC</i> index	Polyarchy	<i>UG</i> index
Model	(1)	(2)	(3)	(4)	(5)	(6)
Industrial worker index	0.066*** (3.64)		0.001 (0.10)	0.247*** (2.96)		
Middle class index	-0.002 (-0.24)	0.270*** (2.94)	0.054** (2.42)			
Urban groups (combined) index					0.123*** (4.14)	
Other campaign	-0.043*** (-4.14)	0.416*** (3.25)	-0.056*** (-2.63)	0.891*** (7.66)	-0.158*** (-3.86)	1.302*** (13.28)
Ln GDP p.c.	0.003 (1.12)	0.010 (0.57)	0.003 (1.06)	0.015 (0.78)	0.002 (0.64)	0.015 (0.80)
Ln Population	0.001 (0.38)	-0.024 (-0.80)	0.000 (0.13)	-0.015 (-0.65)	0.004 (1.23)	-0.047* (-1.95)
Urbanization	0.017 (1.35)	-0.015 (-0.14)	0.014 (1.34)	0.041 (0.44)	0.004 (0.22)	0.106 (0.96)
Other campaigns in neighb.	-0.005 (-0.38)	-0.255** (-2.00)	-0.002 (-0.21)	-0.412*** (-3.81)	0.039*** (2.65)	-0.611*** (-4.31)
Democracy in neighborhood	0.050*** (3.23)	0.143 (1.00)	0.057*** (3.85)	0.034 (0.30)	0.049** (2.52)	0.109 (0.83)
Ind. work. camp. in neighb.		1.116*** (4.04)	0.076*** (4.78)	-0.337** (-2.18)		0.543** (2.37)
Mid. class camp. in neighb.	0.052** (2.53)	-0.349** (-2.20)		0.785*** (4.53)		0.449** (2.61)
Lagged Dependent variable	0.945*** (140.51)	-0.075 (-1.43)	0.942*** (137.69)	-0.035 (-0.97)	0.952*** (107.98)	-0.106** (-2.16)
Country dummies	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
N	9628		9628		9673	
Hansen J-test p-value	Ex. ident.		Ex. ident.		.64	
Cragg-Donald Wald F-stat.	696.7		377.4		236.6	
Instrument(s)	<i>IW</i> neighb.		<i>UM</i> neighb.		<i>IW</i> neighb, <i>UM</i> 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.

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 causal effect of industrial worker campaigns on democratization.

The estimated effect of urban middle class involvement in campaigns is also fairly clear. The *MC* index is significant at 5% ( $t = 2.4$ ) in Model 3, and the point estimate is only slightly lower than for the *MC* index in Model 1. Again, the first-stage regression reveals neighborhood (middle class) campaign profile to be a strong instrument.

Finally, in Models 5–6, we test for an effect of (any type 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 at least 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 campaigns in the neighborhood – we can run standard overidentification tests. According to the Hansen J-test p-value (.64) the exclusion restriction does not seem to be violated. This provides fairly strong evidence for the (cruder) hypothesis that the involvement of urban groups in opposition movements enhances prospects for democratic change.

It is important to note 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 regional neighborhood. While we have no concrete 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.

#### **4.4 Is the relationship between urban opposition movements and democratization sensitive to contextual factors?**

Before concluding, we briefly discuss whether results are sensitive to the time period under study, geographic region, and whether societies are urbanized or not. These tests may be viewed as “robustness tests”, but, as we will clarify, we consider at least some of them rather as tests of additional implications from our argument. Since we are limiting degrees of freedom by considering

particular contexts, we mostly focus on the “combined”  $UG$  dummy: when investigating particular time periods or omitting key regions, the number of campaigns dominated by one group can become quite low (but see Appendix Table C.14 for tests on  $IW$  and  $MC$  separately). The models below are similar to our benchmark, with, e.g., country- and year-fixed effects and errors clustered by country.

We start out by investigating whether the relationship between urban campaigns and democratization is sensitive to time period under study. Admittedly, given the fairly general nature of our argument, we should be concerned if we only identify the relationship in some parts of recent modern history. But, as argued above, most countries globally (with the exception of some Western countries) have mainly observed processes of industrialization and, particularly, urbanization after mid-20th century. Hence, the capacities and motivations of urban actors to pursue democratization may have increased. When splitting the 1900–2010 time-period in two, we find that  $UG$  is 0.060 when estimated on post-1955 data (Model 2) and 0.017 on pre-1956 data (Model 1). Still, also the latter result is (weakly) significant, despite including only 2728 country-year observations and 34 observations of campaigns dominated by urban groups. This dovetails nicely also with previous case-study evidence for urban middle class or industrial worker movements in democratization coming from both historical and more contemporary cases (see, e.g., Collier, 1999).

We would also be concerned if our results are basically stemming from developments in one particular geographic region. However, we did not expect this: Although many early studies on urban groups and democratic transitions focused on Europe or Latin America, several recent studies on third-wave democratization experiences have considered countries from different regions, including Eastern Europe (McFaul, 2002) and Sub-Saharan Africa (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 1% when omitting any major region.  $UG$  is lowest when omitting Eastern Europe (.038; Model 4; benchmark estimate is .047) and highest when omitting Asia-Pacific (.063; Model 8).

Table 6: 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		
	≤ 1956	> 1956	$\div W/E/NA$	$\div EE$	$\div SSA$	$\div ME/NA$	$\div A/P$	$\div LA$	Rural	Urban	Full	Full
IW OR MC dominate	0.017* (1.71)	0.060*** (4.29)	0.039*** (3.50)	0.038*** (3.15)	0.046*** (3.19)	0.048*** (3.93)	0.052*** (4.00)	0.063*** (4.09)	0.021* (1.98)	0.076*** (3.56)	-0.012 (-0.46)	-0.012 (-0.46)
IW/MC*Urbanization											0.130* (1.86)	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 (-1.46)	-0.002 (-0.49)	-0.000 (-0.10)	-0.002 (-0.64)	-0.002 (-0.46)	-0.002 (-0.61)	-0.003 (-0.76)	-0.002 (-0.66)	-0.003 (-0.71)	-0.006 (-0.78)	-0.003 (-0.52)	-0.004 (-0.63)
OC*Urbanization											0.004 (0.20)	0.003 (0.16)
Elect. democracy	0.940*** (74.99)	0.936*** (130.47)	0.943*** (139.77)	0.943*** (154.41)	0.953*** (153.06)	0.948*** (157.89)	0.948*** (149.21)	0.951*** (158.07)	0.917*** (76.68)	0.949*** (118.07)	0.949*** (159.28)	0.950*** (169.37)
Ln GDP p.c.	-0.014* (-1.87)	0.001 (0.45)	0.003 (1.00)	0.001 (0.59)	0.003 (1.12)	0.003 (0.94)	0.002 (0.57)	0.001 (0.25)	-0.000 (-0.14)	0.003 (0.82)	0.002 (0.91)	0.002 (0.87)
Ln Population	-0.004 (-0.36)	-0.007 (-1.58)	-0.003 (-0.92)	0.000 (0.16)	0.000 (0.04)	0.001 (0.54)	-0.001 (-0.34)	-0.003 (-1.27)	-0.001 (-0.15)	-0.009* (-1.86)	-0.000 (-0.21)	-0.000 (-0.15)
Urbanization	-0.103*** (-2.92)	0.034** (1.99)	0.016 (1.29)	0.028** (2.41)	0.027** (2.37)	0.026** (2.33)	0.012 (1.10)	0.021* (1.95)	0.029 (1.35)	0.061** (2.54)	0.018* (1.72)	0.018* (1.70)
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. Polarity in t+1 is dependent variable. All models are OLS with country- and year fixed effects. Errors are clustered by country.

Finally, we turn to regressions that should rather be viewed as additional tests of our argument: We expected that industrial workers and urban middle class actors have clearer motivation and stronger capacity to induce democratization in more urban societies. This is also what we find: Models 9 (relatively rural) and 10 (urban) split the sample according to 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 p-value is .051 even in Model 9.

To further probe the theorized interaction, we added multiplicative terms (urbanization multiplied with, respectively, urban and other campaigns) to the benchmark. This model (11) predicts that the expected interaction is sizeable, although not very precisely estimated ( $t = 1.9$ ). 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 estimated interaction term is both larger in size (.16 vs .09) and t-value (1.7 vs 0.6) for *IW* than for *MC*. While we note that 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.

## 5 Conclusion

A prominent view – linking socio-economic development to democratization – is that urbanization and industrialization empowers industrial workers and/or the urban middle classes, and that these groups, in turn, often act as agents of democratization. In this paper, we have elaborated on why these groups may be motivated to pursue democratization, particularly in urban societies, and explaining why they should have the required capacity to (sometimes) succeed. Importantly, we also shed new empirical light on this fascinating issue – which has predominantly either been scrutinized in historical accounts of selected cases *or* by regression models employing distal proxies of social group motivation and/or capacity, such as GDP per capita. To this end, we collected data on the social groups partaking in organized opposition campaigns throughout the world from 1900–2012.

Our main findings can be quickly summarized as follows: Social movements dominated by these

urban groups are clearly linked to democratization,. This goes both when we compare to opposition movements dominated by other groups, such as peasants or religious or ethnic groups, and to situations where no organized opposition movement exists. The result finds support in different contexts, but the relationship seems stronger in more recent decades and in urbanized societies. Further, when making the distinction between opposition movements dominated by the urban middle classes and by industrial workers, we find evidence that both are associated with democratization. However, results are much more robust for the democratizing role of industrial worker movements. This result corroborates the accounts presented by scholars such as Rueschemeyer, Stephens and Stephens (1992) and Collier (1999).

Our findings inform different, though partly overlapping, literatures. First, we present a novel empirical take on the social origins of democratic revolutions, thus contributing to long-standing debates in comparative-historical political sociology (Moore, 1966; Skocpol, 1979; Rueschemeyer, Stephens and Stephens, 1992; Luebbert, 1991; Tilly, 1996). Second, our analysis significantly informs the recent literature on the origins of dictatorship and democracy in general (Acemoglu and Robinson, 2006; Boix, 2003; Teorell, 2010; 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. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115(4):1167–1199.
- Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. New York, NY: Cambridge University Press.
- Ackerman, Peter and Jack DuVall. 2001. *A force more powerful: a century of non-violent conflict*. New York: Palgrave Macmillan.
- 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(0):52 – 75.
- Ansell, Ben W and David J Samuels. 2014. *Inequality and Democratization*. Cambridge University Press.
- Ardanaz, Martin and Isabela Mares. 2013. "Labor Shortages, Rural Inequality, and Democratization." *Comparative Political Studies* p. 0010414013512609.
- Bayer, Markus, Felix S Bethke and Daniel Lambach. 2016. "The democratic dividend of nonviolent resistance." *Journal of Peace Research* .  
**URL:** <http://jpr.sagepub.com/content/early/2016/09/30/0022343316658090.abstract>
- Beck, Nathaniel and Jonathan N. Katz. 2001. "Throwing out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon." *International Organization* 55:187–195.
- 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. 2012. "A complete data set of political regimes, 1800–2007." *Comparative Political Studies* p. 0010414012463905.
- 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*. 0022002714541843.
- Casper, Brett Allen and Scott A Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'état." *The Journal of Politics* 76(02):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 Non-violent Uprisings?" *Journal of Conflict Resolution* 0022002715576574.
- Chenoweth, Erica and Kathleen Gallagher Cunningham. 2013. "Understanding nonviolent resistance An introduction." *Journal of Peace Research* 50(3):271–276.
- Chenoweth, Erica and Maria J. Stephan. 2011. *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. Columbia Studies in Terrorism and Irregular Warfare.
- 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* .
- Collier, Ruth Berins. 1999. *Paths Towards Democracy: The Working Class and Elites in Western Europe and South America*. Cambridge: Cambridge University Press.
- Coppedge, Michael, John Gerring, Staffan Lindberg, Jan Teorell, David Altman, Michael Bernhard, Steven Fish, Adam Glynn, Allen Hicken and Matthew Kroenig. 2015. "Varieties of Democracy (V-Dem) Codebook."
- Dahlum, Sirianne. 2015. "Students in the Streets: The Role of Education in Non Violent Conflict." Working Paper, University of Oslo.
- 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: Social Classes and the Political Origins of Regimes in Interwar Europe*. 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(02):212–244.
- Moore, Barrington. 1966. *Social Origins of Democracy and Dictatorship: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.
- Nepstad, Sharon Erickson. 2011. *Nonviolent revolutions: Civil resistance in the late 20th century*. Oxford University Press.
- Ness, Immanuel. 2015. "The International Encyclopedia of Revolution and Protest." .  
**URL:** <http://www.revolutionprotestencyclopedia.com>
- O'Donnell, Guillermo. 1973. *Modernization and Bureaucratic-Authoritarianism: Studies in South American Politics*. Politics of Modernization Theories Berkeley, CA: Institute of International Studies, University of California.
- O'Donnell, Guillermo and Phillippe C. Schmitter. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions About Uncertain Democracy*. Baltimore, MD: Johns Hopkins University Press.
- Olson, Mancur. 1965. *The Logic of Collective Action*. MA: Harvard University Press.
- Powell, Jonathan. 2012. "Determinants of the Attempting and Outcome of Coup." *Journal of Conflict Resolution* 56(6):1017–104.
- 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: the theory of Stein Rokkan: based on his collected works*. Clarendon Press.
- Rueschemeyer, Dietrich, Evelyne Huber Stephens and John D. Stephens. 1992. *Capitalist Development and Democracy*. Chicago, IL: University of Chicago Press.
- Satterthwaite, David. 2007. *The transition to a predominantly urban world and its underpinnings*. Human Settlement Discussion Series.
- Skocpol, Theda. 1979. *States and social revolutions: A comparative analysis of France, Russia and China*. Cambridge, UK: Cambridge University Press.
- 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.  
**URL:** <http://www.jstor.org/stable/2780466>
- Stock, James H and Motohiro Yogo. 2005. "Testing for weak instruments in linear IV regression." *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* .

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.

Tilly, Charles. 1996. "European Revolutions: 1492-1992 (Making of Europe)."

Welzel, Christian. 2013. *Freedom rising*. Cambridge University Press.

Weyland, Kurt. 2009. "The diffusion of revolution: 1848 in Europe and Latin America." *International Organization* 63(03):391–423.

Weyland, Kurt. 2010. "The diffusion of regime contention in European democratization, 1830–1940." *Comparative Political Studies* 43(8-9):1148–1176.

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

Wig, Tore and Espen Geelmuyden Rød. 2014. "Cues to Coup plotters: Elections as coup triggers in dictatorships." *Journal of Conflict Resolution* Online First:1–26.

## A Codebook for indicators of protest movements social composition

Below is the codebook for the variables recording the social background of participants in protest campaigns. For each campaign, we always consulted several different sources, including both databases and encyclopedias such as “The international encyclopedia of revolution and protest” (Ness, 2015) and the “Global Non-Violent Action Database” (Swarthmore, 2015), partly on various reports by international or non-governmental organizations such as UNHCR or Freedom House, where the relevant events were described, and partly on books and journal articles that cover the campaigns, such as e.g. 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 who the campaign originated. When it comes to the variable recording whether each social group dominated the campaign, we relied on a coding rule requiring that the movement should at least either have that social group as majority of its members, or, if not, they should have had a critical impact on the strategies and outcomes of the campaign.

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
- **Other urban middle classes**, referring to all urban groups other than public sector employees, including business elites, smaller merchants, professionals (including e.g. lawyers, doctors etc) and university professors.
- **Military employees**, referring to military officers, soldiers as well as veterans
- **Members of religious or ethnic groups**

Most of these groups are mutually exclusive, except from religious or ethnic groups, that can overlap with the other social groups.

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

### 1. Participated in the campaign

These variables record 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 get assigned the value “1”, at least 2 sources should mention that the social group participated in the campaign. To get assigned the value “0”, at 4-5 sources should *not* mention these groups as having participated in the campaign, to ensure that the no-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

These variables record whether the campaign was dominated by each social group. The dominance criteria should be considered fulfilled 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“ criteria 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

These variables record 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

## B Descriptive statistics

Table B.1: Summary statistics for key variables (Restricted to the 9738 observations in the baseline model)

	(1)				
	count	mean	sd	min	max
Electoral democracy index	9738	.3988276	.2856613	.0092174	.9559249
Democracy (BMR)	8240	.4302184	.4951366	0	1
Middle class OR ind. workers dominated	9738	.0141713	.1182028	0	1
Middle class dominated	9738	.0088314	.0935644	0	1
Industrial workers dominated	9738	.0081125	.0897082	0	1
Middle class origin	9721	.0120358	.1090512	0	1
Industrial worker origin	9729	.0062699	.0789382	0	1
Middle class participated	9646	.0462368	.2100084	0	1
Industrial workers participated	9692	.0338423	.1808326	0	1
Public sector workers dominated	9737	.0029783	.0544955	0	1
Religious or ethnic group dominated	9733	.0322614	.1767025	0	1
Peasants dominated	9732	.0256885	.1582123	0	1
Public sector workers origin	9733	.0025686	.0506186	0	1
Religious or ethnic group origin	9695	.0292935	.1686365	0	1
Peasants origin	9695	.0096957	.0979934	0	1
Peasants participated	9718	.0595802	.2367195	0	1
Public sector workers participated	9685	.0279814	.1649281	0	1
Religious or ethnic group participated	9689	.0477861	.2133245	0	1
Other campaign	9738	.0734237	.2608441	0	1
lmembers	9738	2.815054	11.6175	0	68
regviol	9738	.0664407	.2490637	0	1
Ln GDP p.c.	9738	7.845352	1.031364	5.315224	10.66701
Ln Population	9738	15.89295	1.556347	10.12663	21.00888
Urbanization	9738	.4302695	.2296751	.0134997	.97432
Industrial worker index	9683	.0465765	.2885249	0	3
Middle class index	9629	.0647004	.3312883	0	3
Middle class dominated campaigns in neighborhood	9738	.063606	.1071837	0	1.636364
Ind. worker dominated campaigns in neighborhood	9738	.0427623	.0921789	0	1.454545
Campaigns in neighborhood	9738	.0949367	.1064943	0	.8181818
Democracy score in neighborhood	9738	.3883268	.2198609	.0325952	.8979004
Observations	9738				

## C Additional robustness tests and extensions

This Appendix contains various other robustness tests and extensions noted in the main text, but not reported there in Tables.

The sequence of tests is as follows: Table C.1 shows campaign level regressions on democratization reflecting the situation five years – rather than the baseline one year – after the relevant campaign ended. Table C.2 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 C.3–C.5 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 C.6–C.7 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 C.10 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 C.11 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 C.12 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 C.13 replicates the second stages of the 2SLS regressions from Table 5, but now using the “campaign domination” dummies rather than the more fine grained lexical indices as endogenous regressors. Finally, Table C.14 reproduces the split sample tests reported in Table 6 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 C.1: 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)
Dependent variable	Democratic transition.			
Ind. workers or mid. class dominate	0.267*** (3.41)	0.145* (1.70)		
Industrial workers dominate			0.267*** (2.76)	0.160 (1.59)
Middle class dominate			0.257*** (2.84)	0.205** (2.18)
Ln campaign participants		0.000 (0.07)		0.001 (0.33)
Urbanization		0.208 (0.66)		0.202 (0.66)
Ln population		-0.022 (-0.80)		-0.021 (-0.79)
Ln GDP per capita		0.189** (2.60)		0.179** (2.51)
Constant	0.187*** (4.27)	-0.874 (-1.44)	0.170*** (4.09)	-0.845 (-1.43)
N	140	122	139	121

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 C.2: 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			
IW or MC dominate	1.092** (2.48)	0.687 (1.34)			1.284*** (3.21)	0.704 (1.48)		
Ind. workers dom.			1.128** (2.20)	0.827 (1.47)			1.248** (2.54)	0.782 (1.40)
Mid. class domi.			0.679 (1.35)	0.529 (0.94)			1.225*** (2.65)	1.065* (1.94)
Ln camp. particip.		0.001 (0.05)		0.002 (0.14)		-0.001 (-0.09)		0.002 (0.22)
Urbanization		3.233* (1.70)		3.327* (1.72)		1.013 (0.54)		1.065 (0.55)
Ln population		-0.081 (-0.44)		-0.064 (-0.34)		-0.147 (-0.87)		-0.142 (-0.82)
Ln GDP per capita		0.323 (0.73)		0.288 (0.65)		1.037** (2.31)		1.014** (2.23)
Constant	-1.854*** (-6.22)	-4.085 (-1.07)	-1.871*** (-6.51)	-4.238 (-1.11)	-1.466*** (-5.61)	-6.899* (-1.93)	-1.570*** (-6.03)	-7.082* (-1.94)
N	140	122	139	121	140	122	139	121

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 C.3: 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	0.051*** (4.13)		0.047*** (3.98)	
Industrial workers dominate		0.054*** (2.68)		0.049** (2.55)
Middle class dominate		0.036* (1.85)		0.034* (1.86)
Other campaign	-0.001 (-0.34)	-0.002 (-0.45)	-0.002 (-0.58)	-0.003 (-0.73)
Ln GDP p.c.	-0.001 (-0.32)	-0.001 (-0.40)	0.002 (0.78)	0.002 (0.74)
Ln Population	0.000 (0.03)	0.000 (0.04)	-0.001 (-0.24)	-0.001 (-0.23)
Urbanization	0.014* (1.79)	0.014* (1.76)	0.022** (2.14)	0.022** (2.16)
Polyarchy (level in $t$ )			-0.052*** (-9.12)	-0.052*** (-9.27)
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9739	9738	9739	9739
Countries	147	147	147	147
Max time series	1900–2010	1900–2010	1900–2010	1900–2010

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 C.4: 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 dominated	0.049*** (4.40)		0.046*** (4.23)	
Industrial workers dominate		0.048*** (2.68)		0.044** (2.59)
Middle class dominate		0.037** (2.22)		0.036** (2.22)
Other campaign	0.007** (2.44)	0.007** (2.33)	0.006** (2.25)	0.006** (2.12)
Ln GDP p.c.	-0.003 (-1.10)	-0.003 (-1.16)	-0.001 (-0.28)	-0.001 (-0.33)
Ln Population	0.004* (1.93)	0.004* (1.86)	0.003 (1.35)	0.003 (1.32)
Urbanization	0.015 (1.60)	0.015 (1.60)	0.021* (1.98)	0.021** (2.00)
Polyarchy (level in $t$ )			-0.037*** (-7.25)	-0.037*** (-7.37)
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9739	9738	9739	9738
Countries	147	147	147	147
Max time series	1900–2010	1900–2010	1900–2010	1900–2010

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 C.5: 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 dominated	0.002 (0.48)		0.001 (0.18)	
Industrial workers dominated		0.006* (1.82)		0.004 (1.45)
Middle class dominated		-0.001 (-0.18)		-0.001 (-0.27)
Other campaign	-0.008*** (-3.51)	-0.008*** (-3.72)	-0.008*** (-3.70)	-0.009*** (-3.93)
Ln GDP p.c.	0.002 (1.59)	0.002 (1.55)	0.003** (2.33)	0.003** (2.29)
Ln Population	-0.004** (-2.60)	-0.003** (-2.56)	-0.004*** (-2.79)	-0.004*** (-2.75)
Urbanization	-0.001 (-0.20)	-0.001 (-0.23)	0.001 (0.25)	0.001 (0.22)
Polyarchy (level in $t$ )			-0.015*** (-5.55)	-0.015*** (-5.48)
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9739	9738	9739	9738
Countries	147	147	147	147
Max time series	1900–2010	1900–2010	1900–2010	1900–2010

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 C.6: Excluding all democracies (and excluding democratization years)

	(1)	(2)	(3)	(4)
	m1	m2	m3	m4
Middle class or industrial workers dominated	0.0685 (1.51)	0.0687 (1.50)	0.0307* (2.03)	0.0307* (2.03)
Other campaign	-0.0205* (-1.98)	-0.0223 (-1.73)	-0.0188** (-2.71)	-0.0188** (-2.71)
Polyarchy (level in $t$ )	0.871*** (64.32)	0.816*** (37.84)	0.624*** (54.55)	0.624*** (54.55)
Ln GDP p.c.		0.0132** (3.30)	0.00400 (0.63)	0.00400 (0.63)
Ln population		0.000983 (0.48)	-0.00717 (-0.90)	-0.00717 (-0.90)
Urbanization		0.0368 (1.85)	0.141*** (5.17)	0.141*** (5.17)
country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	3732	3528	3528	3528

$t$  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 democratization, in dictatorships.

Table C.7: Excluding all democracies (and including democratization years)

	(1)	(2)	(3)	(4)
	FE OLS	FE OLS	FE OLS	FE OLS
Middle class OR ind. workers dominated	0.0735*** (3.62)	0.0740*** (3.62)	0.0768*** (9.50)	0.0603*** (7.57)
Other campaign	0.000941 (0.22)	-0.00171 (-0.36)	-0.00494 (-1.12)	-0.00672 (-1.55)
Polyarchy (level in $t$ )	0.872*** (44.79)	0.864*** (37.74)	0.691*** (58.34)	0.640*** (50.66)
Ln GDP p.c.		-0.00506 (-1.57)	-0.00510 (-1.48)	0.0176*** (4.54)
Ln Population		-0.000654 (-0.61)	0.0337*** (6.91)	0.0133 (1.83)
Urbanization		0.0284* (2.23)	0.00688 (0.34)	-0.0117 (-0.52)
country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	5975	4670	4670	4670

$t$  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 C.8: Excluding all dictatorships (and including democratization years)

	(1)	(2)	(3)	(4)
	m1	m2	m3	m4
Middle class or industrial workers dominated	0.0685 (1.51)	0.0687 (1.50)	0.0307* (2.03)	0.0307* (2.03)
Other campaign	-0.0205* (-1.98)	-0.0223 (-1.73)	-0.0188** (-2.71)	-0.0188** (-2.71)
Polyarchy (level in $t$ )	0.871*** (64.32)	0.816*** (37.84)	0.624*** (54.55)	0.624*** (54.55)
Ln GDP p.c.		0.0132** (3.30)	0.00400 (0.63)	0.00400 (0.63)
Ln population		0.000983 (0.48)	-0.00717 (-0.90)	-0.00717 (-0.90)
Urbanization		0.0368 (1.85)	0.141*** (5.17)	0.141*** (5.17)
country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	3732	3528	3528	3528

$t$  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 C.9: Excluding all dictatorships (and excluding democratization years)

	(1)	(2)	(3)	(4)
	FE OLS	FE OLS	FE OLS	FE OLS
Middle class OR ind. workers dominated	-0.0289 (-1.16)	-0.0263 (-0.91)	-0.0287 (-1.76)	-0.0287 (-1.76)
Other campaign	-0.0177* (-2.40)	-0.0179* (-2.27)	-0.0148* (-2.33)	-0.0148* (-2.33)
Polyarchy (level in $t$ )	0.932*** (84.03)	0.900***	0.782*** (69.72)	0.782*** (69.72)
Ln GDP p.c.		0.00732** (3.18)	-0.000834 (-0.15)	-0.000834 (-0.15)
Ln Population		0.000628 (0.48)	-0.0107 (-1.53)	-0.0107 (-1.53)
Urbanization		0.0225* (2.13)	0.0994*** (4.13)	0.0994*** (4.13)
country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	3594	3405	3405	3405

$t$  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 C.10: Robustness tests when using combined dummy for urban middle class OR industrial workers dominate campaign as key measure

Model	(1)	(2)	(3)	(4)	(5)	(6)
	2-year lag	No clustering	Parsimonious	Extra controls	Nr groups	No uncertain cases
Middle class OR ind. workers dominate	0.071*** (3.64)	0.047*** (10.19)	0.047*** (4.09)	0.047*** (3.86)	0.031** (2.44)	0.044*** (3.48)
Other campaign	-0.002 (-0.39)	-0.002 (-0.85)	-0.002 (-0.61)	-0.016 (-1.03)	-0.032*** (-3.19)	-0.003 (-0.72)
Ln GDP p.c.	0.004 (0.77)	0.002 (1.11)		0.002 (0.79)	0.002 (0.79)	0.002 (0.72)
Ln Population	-0.003 (-0.63)	-0.001 (-0.24)		-0.000 (-0.12)	-0.000 (-0.21)	-0.000 (-0.14)
Urbanization	0.050** (2.32)	0.022** (2.46)		0.023** (2.17)	0.022** (2.11)	0.023** (2.14)
Ln number of campaign participants				-0.000 (-0.51)		
Regime violence				0.020 (1.38)		
Number of social groups in campaign					0.011*** (3.14)	
Lagged dependent variable (Polyarchy)	0.873*** (74.33)	0.948*** (272.49)	0.953*** (213.31)	0.949*** (164.49)	0.949*** (163.59)	0.948*** (166.66)
Country dummies	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
N	9735	9739	15520	9739	9628	9646

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 C.11: Robustness tests, with positive change in Polity from  $t$  to  $t + 1$  (democratic upturns) as dependent variable.

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	2-year lag	No clustering	Parsimonious	Extra controls	Public employees	Nr groups	No uncertain	Campaign origin	Lexical indices
Ind. workers dominate	0.034*** (3.07)	0.044*** (9.41)	0.042*** (2.52)	0.042*** (2.61)	0.045*** (2.77)	0.038** (2.16)	0.047*** (2.58)		
Middle class dominate	0.016* (1.82)	0.036*** (7.95)	0.037** (2.32)	0.034** (2.12)	0.033** (2.50)	0.023 (1.39)	0.033* (1.89)		
Other campaign	0.004* (1.73)	0.066*** (3.30)	0.007** (2.38)	0.007 (0.75)	0.005* (1.85)	-0.016** (-2.45)	0.006* (1.86)	0.008** (2.33)	-0.008* (-1.88)
Lagged Polyarchy level	-0.052*** (-9.49)	-0.037*** (-13.16)	-0.031*** (-7.78)	-0.036*** (-7.28)	-0.037*** (-7.35)	-0.036*** (-7.08)	-0.036*** (-7.22)	-0.037*** (-7.70)	-0.036*** (-7.24)
Ln GDP p.c.	0.000 (0.00)	-0.001 (-0.63)	-0.001 (-0.32)	-0.001 (-0.32)	-0.001 (-0.28)	-0.001 (-0.30)	-0.001 (-0.31)	-0.001 (-0.41)	-0.001 (-0.35)
Ln Population	0.003 (1.05)	0.003* (1.71)	0.003 (1.47)	0.003 (1.47)	0.003 (1.34)	0.003 (1.50)	0.004 (1.54)	0.003 (1.26)	0.004 (1.50)
Urbanization	0.021*** (2.01)	0.021*** (2.92)	0.021*** (2.02)	0.021*** (2.02)	0.020* (1.92)	0.020* (1.89)	0.020* (1.92)	0.021** (1.98)	0.020* (1.87)
Ln nr particip. camp.				-0.000 (-1.51)					
Regime violence				0.008 (0.84)					
Nr. groups campaign						0.008*** (3.03)			
Industrial worker origin								0.032* (1.88)	
Middle class origin								0.022** (2.11)	
Industrial worker index									0.016*** (2.77)
Middle class index									0.013*** (2.90)
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	9731	9738	15519	9738	9738	9628	9646	9744	9628

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 C.12: 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	$\Delta$ Polyarchy	Upturns Polyarchy
Middle class OR ind. workers dominate	0.046*** (3.78)			
Industrial workers dominate		0.052*** (3.02)	0.052*** (3.02)	0.045*** (3.00)
Middle class dominate		0.034* (1.83)	0.034* (1.83)	0.033** (2.03)
Public employees dominate	0.023 (0.70)	0.020 (0.63)	0.020 (0.63)	0.019 (0.69)
Regious or ethnic groups dominate	-0.003 (-0.35)	-0.002 (-0.26)	-0.002 (-0.26)	-0.011* (-1.71)
Peasants dominate	-0.025** (-2.36)	-0.024** (-2.27)	-0.024** (-2.27)	-0.029*** (-3.52)
Existence of any campaign	0.007 (0.87)	0.006 (0.69)	0.006 (0.69)	0.020*** (3.26)
Lagged Polyarchy	0.948*** (165.21)	0.949*** (168.59)	-0.051*** (-9.13)	-0.036*** (-7.32)
Ln GDP p.c.	0.002 (0.96)	0.002 (0.92)	0.002 (0.92)	-0.000 (-0.19)
Ln Population	-0.000 (-0.04)	-0.000 (-0.00)	-0.000 (-0.00)	0.004 (1.54)
Urbanization	0.021* (1.94)	0.020* (1.94)	0.020* (1.94)	0.019* (1.84)
Country dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	9726	9726	9726	9726

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 C.13: 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	<i>IW</i> or <i>MC</i> dominate
Model	(1)	(2)	(3)
Industrial workers dominate	0.212*** (3.12)	0.007 (0.30)	
Middle class dominate	-0.008 (-0.36)	0.149** (2.47)	
Middle class OR ind. workers dominated			0.257*** (5.30)
Other campaign	-0.021*** (-3.00)	-0.018** (-2.25)	-0.054*** (-3.61)
Lagged Polyarchy	0.946*** (136.37)	0.942*** (133.14)	0.951*** (115.23)
Ln GDP p.c.	0.003 (1.10)	0.003 (1.15)	0.002 (0.65)
Ln Population	0.002 (0.93)	0.000 (0.14)	0.005 (1.58)
Urbanization	0.014 (1.06)	0.017 (1.61)	0.012 (0.83)
Middle class dominated campaigns in neighborhood	0.054*** (2.67)		
Ind. worker dominated campaigns in neighborhood		0.077*** (4.95)	
Campaigns in neighborhood	-0.003 (-0.28)	-0.003 (-0.26)	0.035*** (2.58)
Democracy score in neighborhood	0.056*** (3.36)	0.056*** (3.47)	0.046** (2.29)
Country dummies	Y	Y	Y
Year dummies	Y	Y	Y
N	9739	9738	9739
Hansen J-test p-value	Ex. ident.	Ex. ident.	.92
Cragg-Donald Wald F-stat.	527.8	294.7	270.7
Instrument(s)	<i>IW</i> neighb.	<i>UM</i> neighb.	<i>IW</i> neighb, <i>UM</i> 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 C.14: 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	≤ 1956	> 1955	$\div WE/N.A$	$\div EE$	$\div SSA$	$\div ME/N.A$	$\div A/P$	$\div LA$	Rural	Urban	Full
Sample	Time	Time		Omitting particular geographic region					Rural	Urban	Full
Ind. workers dominate	0.020 (1.51)	0.068*** (2.73)	0.052*** (2.67)	0.055*** (2.31)	0.040* (1.90)	0.048** (2.49)	0.048*** (2.35)	0.048*** (2.95)	0.051** (2.57)	0.055* (1.66)	-0.021 (-0.56)
Mtd. class dominate	0.015 (0.84)	0.032 (1.51)	0.021 (1.28)	0.019 (1.11)	0.039 (1.53)	0.035* (1.76)	0.040*** (2.00)	0.057*** (2.73)	0.005 (0.43)	0.063 (1.58)	-0.004 (-0.07)
MC dom. X urbaniz.											0.003 (0.60)
IW dom. X urbaniz.											0.161* (1.72)
Other campaign	-0.014 (-1.61)	-0.001 (-0.31)	-0.001 (-0.32)	-0.003 (-0.90)	-0.002 (-0.48)	-0.003 (-0.74)	-0.003 (-0.83)	-0.003 (-0.71)	-0.005 (-1.34)	-0.004 (-0.48)	-0.004 (-0.63)
Oth. camp. X urbaniz.											0.003 (0.16)
Ln GDP p.c.	-0.014* (-1.92)	0.001 (0.49)	0.002 (0.95)	0.001 (0.53)	0.003 (1.10)	0.002 (0.87)	0.002 (0.54)	0.001 (0.25)	-0.001 (-0.37)	0.003 (0.70)	0.002 (0.87)
Ln Population	-0.004 (-0.35)	-0.006 (-1.46)	-0.003 (-0.87)	0.001 (0.29)	-0.000 (-0.04)	0.001 (0.51)	-0.001 (-0.38)	-0.003 (-1.25)	-0.000 (-0.04)	-0.009* (-1.90)	-0.000 (-0.15)
Urbanization	-0.103*** (-2.90)	0.033* (1.96)	0.016 (1.26)	0.027** (2.28)	0.028** (2.45)	0.026** (2.39)	0.013 (1.16)	0.021** (1.99)	0.030 (1.42)	0.063*** (2.63)	0.018* (1.70)
Lagged Polyarchy	0.940*** (74.99)	0.937*** (134.62)	0.944*** (143.69)	0.943*** (154.32)	0.953*** (158.66)	0.948*** (161.83)	0.948*** (152.29)	0.951*** (160.62)	0.916*** (76.49)	0.949*** (124.15)	0.950*** (169.37)
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	2727	7011	7634	9154	7245	8900	7973	7784	4869	4870	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.



***Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex).***

***Ce projet est distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR).***

***Il est financé dans le cadre des investissements d'avenir.***

***(ANR-11-LABX-0091, ANR-11-IDEX-0005-02)***

***[www.sciencespo.fr/liepp](http://www.sciencespo.fr/liepp)***

**Directeurs de publication :**

**Bruno Palier & Etienne Wasmer**

Sciences Po - LIEPP  
27 rue Saint Guillaume  
75007 Paris - France  
+33(0)1.45.49.83.61  
liepp@sciencespo.fr

