

NBER WORKING PAPER SERIES

THE LOGIC OF STATE SURVEILLANCE

Gemma Dipoppa
Annalisa Pezone

Working Paper 34492
<http://www.nber.org/papers/w34492>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2025

For suggestions, we are grateful to Oendrila Dube, Anna Grzymala-Busse, Saad Gulzar, John Marshall, Vincent Pons, Massimo Pulejo, Awa Ambra Seck, Jesse Shapiro, Mara Squicciarini, Silvia Vannutelli, Nadia Urbinati, David Yang, and seminar participants at NBER Political Economy, Barcelona Summer Forum in Political Economy, CEPR Applied Micro-Economic History Workshop, Harris PE Seminar, NYU PE Seminar, Columbia CP Seminar, Winsconsin CP Seminar, USC HPE Workshop, and Yale-UB HPE Workshop. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by Gemma Dipoppa and Annalisa Pezone. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Logic of State Surveillance
Gemma Dipoppa and Annalisa Pezone
NBER Working Paper No. 34492
November 2025
JEL No. N43, N44, O33, O38, P00

ABSTRACT

All states adopt systems to surveil political activists. How do they decide whom to watch and why? We study the logic of state surveillance using the first complete individual-level database of those monitored by a state — 152,000 Italians born between 1816 and 1932, encompassing both democratic and authoritarian regimes. We focus on education: exploiting a discontinuous expansion in primary schooling in municipalities above a population and age threshold, we show that cohorts exposed to more years of school experienced an uptick in surveillance. The effect is largest for working classes, who were monitored for longer periods, subjected to harsher measures, and disproportionately targeted when affiliated with communist ideologies. Yet treated cohorts did not become more politically active, indicating that surveillance expanded not in reaction to increased mobilization, but as a preventive strategy rooted in fears of working-class empowerment. These findings reveal how states view educated yet excluded groups as politically threatening and prioritize their surveillance, potentially generating inequalities in groups' ability to influence political change.

Gemma Dipoppa
Columbia University
Department of Political Science
and NBER
gemma.dipoppa@columbia.edu

Annalisa Pezone
New York University
Faculty of Arts and Social Sciences
Department of Politics
ap8979@nyu.edu

An online appendix is available at <http://www.nber.org/data-appendix/w34492>

“For twenty years, we must stop this brain from functioning.” – Michele Isgrò,
Prosecutor in the trial against Antonio Gramsci, 1928.

1 Introduction

States worldwide use systems of surveillance to monitor the political activity of their citizens. In China, a growing apparatus follows activists’ daily activities (Qin, Strömberg and Wu, 2017), and in the United States the government monitors immigrants’ and students’ political behavior (The Guardian, June 18 2025). While surveillance might help maintain security, it may also infringe on individual rights and be used to suppress democratic movements (Zhuravskaya, Petrova and Enikolopov, 2020; Andirin et al., 2022). A recent literature has examined the consequences of surveillance, studying how it impacts dissent and trust (Xu, 2021; Hager and Krakowski, 2022; Beraja et al., 2023; Lichter, Löffler and Siegloch, 2021).

Instead, the determinants of surveillance – which types of people states choose to watch and why – remain largely unexplored, partly due to states’ reluctance to share intelligence data on those they track. Understanding the logic of state surveillance is important to learn about a fundamental strategy of state control and to understand the direction of political change: if certain groups face more scrutiny, the type of change they advocate for might be less likely to take place.

In this paper, we study states’ strategies of surveillance using the first dataset ever assembled on whom a state watched: we scrape newly declassified data with biographical details on individuals surveilled by the Italian state. The data span democratic and autocratic regimes, encompass a range of political movements, from anarchism and communism to fascism, and include detailed information on the surveilled, offering a unique window to observe state choices across contexts, including, unlike previous work, in democratic settings.

Our theoretical point of departure is that states, given limited resources, strategically target those they perceive as threats to their stability: individuals who combine the capacity to mobilize with the potential to seek radical change. A capacity to mobilize can be signaled

by different markers of activism, but education is a particularly important one: a long tradition links education to civic engagement and organizational capacity (Almond, Verba et al., 1963; Verba et al., 1993; Glaeser, Ponzetto and Shleifer, 2007), and police narratives from the surveillance archives point to education as a marker of influence and mobilizing potential. Among those with potential to mobilize, subaltern groups – those who are socially or economically excluded – are described by a long tradition of thought as those most likely to seek radical change (Marx and Engels, 1848; Gramsci, 1975; Fanon, 1963; Gurr, 1970; Scott, 1985).¹ We argue that surveillance operates according to an implicit theory of political threat: states prioritize monitoring those with both the tools to mobilize (education) and the grievances to revolt (subalternity).

Offering evidence for this hypothesis is challenging. In addition to the lack of data on the identities of those monitored, identifying the impact of individual traits such as education is complicated as those are endogenous to other characteristics – family background, group affiliations, or individual fixed attributes – that may cause both social attainment and the likelihood of surveillance.

We address the endogeneity challenge by taking advantage of an exogenous shock to education: a reform that mandated two additional years of primary schooling (third and fourth grade) in municipalities with populations above 4,000. A simple regression discontinuity would be compounded, because other laws also change discretely at this cutoff. Instead, we exploit the interaction of the municipal cutoff with a cohort discontinuity: children who were 7 years old at implementation (i.e., had completed at most second grade) were exposed to the reform, while older cohorts were not. Treatment is thus determined by birth year, not by

¹We use “subaltern” following Gramsci, who defined these as groups beneath the hegemonic power structure and deprived of the means to voice or exit their oppression. While Gramsci referred primarily to class, the concept extends to other marginalized identities such as ethnicity, culture, or religion (Gramsci, 1975, pp. 2286–2288). In other contexts, such as the Chinese Cultural Revolution, elites like landowners could themselves become the subaltern (“black”) classes (Yongyi et al., 2011).

the calendar year of enactment or the timing of subsequent surveillance. Our identification follows a difference-in-discontinuities (DDRD) framework (Grembi, Nannicini and Troiano, 2016), comparing cohorts that completed second grade before versus after the reform in municipalities just below versus just above the cutoff. This design rules out fixed and time-changing confounders unless they vary both at the municipal and cohort threshold. For example, if municipalities above 4,000 received additional funds, and those were used to facilitate surveillance, these improvements would need to enhance surveillance specifically for the treated cohorts, not for all residents.

We start by documenting that the reform was effective: using newly collected data on literacy and a cross-sectional regression discontinuity design, we show that municipalities affected by the additional schooling experienced lower illiteracy rates, consistent with historical accounts that credit the reform with expanding education. We thus use this reform to consider the impact of a shock in primary schooling on surveillance.

Our main result is that more education attracts state scrutiny. Exposure to two additional years of schooling increases surveillance by 3.5 individuals per 100,000 inhabitants, a 64% increase from a 5.5 baseline. This effect is absent in thirty-eight pre-reform cohorts, consistent with the parallel-trends assumption. Importantly, when a later reform removed the 4,000-inhabitant education discontinuity and standardized the educational requirements across municipalities, the effect disappeared, providing additional evidence that the schooling shock – not other factors – drove the increase in surveillance. The effect is present even on cohorts moving to different municipalities, in line with the fact that impacts are not driven by changes in city-characteristics but rather linked to a portable asset – education. In line with a reallocation of monitoring resources, we also observe a drop in surveillance in the control group post-treatment. Results are robust across specifications and absent at placebo population and year of birth cutoffs.

We next ask who bore the brunt of the additional surveillance. Using professions, we classify surveilled individuals into working-, middle-, and upper-class and compute three

outcomes for class-specific surveillance shares. We find that the post-reform increase was driven by surveillance of the working class. Because the reform likely reduced the working-class share of the population, this estimate is a lower bound on the effect of the reform on this group. Consistent with education heightening state fears of working-class mobilization, treated cohorts, and especially working-class individuals, (i) were monitored for longer periods, (ii) faced harsher forms of control, (iii) had more information recorded about them, and (iv) saw larger increases in surveillance among those with socialist or communist affiliations – ideologies tied to the demands of working classes.

This logic of surveillance is similar across regimes and settings: leveraging our extended period, we show that both democratic and autocratic regimes monitored educated working classes at higher rates; that differences in state capacity affect the intensity but not the targets of surveillance; and that increased attention to educated working classes persists after technological innovations in surveillance, underscoring the continued relevance of our findings based on historical data in an era of mass information gathering from online sources.

Having shown that education triggers surveillance of the working classes, our third set of analyses investigates whether the increase in surveillance was driven by increased political mobilization resulting from greater access to education. In our setting, evidence suggests that this force plays a limited role. We combine existing and newly collected data to examine five measures of political activism, ranging from low to high cost actions: voter turnout in national elections, records of notable political figures (scraped from Wikidata), data on protests and strikes (scraped from historical newspapers), the universe of partisans (scraped from an online repository), and episodes of banditry documented by Lecce, Ogliari and Orlando (2022). Across measures, we find no evidence that the education reform increased political participation.

Two explanations are consistent with the null effect of education on political engagement. First, education may not have increased participation in this context, consistent with recent research showing that schooling does not always foster dissent, particularly when designed

to transmit official narratives (Cantoni et al., 2017; Paglayan, 2024). Second, participation might have been suppressed precisely by state preventive surveillance efforts – although, in an heterogeneity analysis, we do not find differences in political engagement between cohorts exposed to surveillance and those who were not. Although we cannot adjudicate between the two, both stories underscore surveillance as a preventive tool: monitoring a large group perceived as possibly threatening for state stability to single out a small number of potential revolutionary entrepreneurs.

The final part of the paper rules out alternative mechanisms operating at the municipality–cohort level that could independently trigger surveillance. First, we test whether the reform produced better-educated policemen who disproportionately surveilled members of their own cohort. We do so by replicating our analysis on surveilled who migrated to other municipalities, where they would not be monitored by former classmates. Second, we replicate our analysis on non-migrants and again find consistent results. This rules out migration as the main driver and addresses the possibility that education-induced moves to larger, more activist cities could explain the findings. Third, we test whether an increase in the number of state officials (teachers) acting as informants could explain the findings. Using a later reform that discontinuously increased the number of state officials in another set of municipalities, without affecting education, we find no associated rise in surveillance. We also show that teachers were not targeted at higher rates or in greater numbers post-reform, making it unlikely that state concern over their potential role as political instigators explains our results. Other alternative explanations are discussed in Section 8.

This study builds on several strands of research. A literature across social sciences has examined how states crack down on political dissent, focusing on overt forms of repression, such as arrests, killings, and censorship (e.g., Besley and Persson (2009); Enikolopov, Petrova and Zhuravskaya (2011); King, Pan and Roberts (2013); Hassan, Mattingly and Nugent (2022); Wen (2025)). A smaller and more recent literature has turned to examine *covert* repression – surveillance. Surveillance differs fundamentally from overt repression: instead of

being a public, performative deterrent, explicitly silencing dissent, it is a covert, information-gathering exercise of state capacity that sidesteps reputational costs, but that demands substantial bureaucratic capacity. Therefore, limited resources are the main constraint for covert action. While previous work on surveillance has focused on its consequences, this paper shifts the focus to its determinants, investigating the strategic factors shaping whom the state chooses to monitor. Before this, Davenport (2005)’s seminal study offered correlational evidence that black neighborhoods were targeted at higher rates in the context of a black nationalist movement in Detroit. We exploit more than a century of archival data spanning multiple regimes and ideologies to show that states target groups most capable of undermining regime stability – newly educated members of subaltern groups.

This paper also relates to a foundational area of inquiry for social science, studying who participates in politics (Almond, Verba et al., 1963; Glaeser, Ponzetto and Shleifer, 2007) and the conditions under which revolutions can take place (Kuran, 1989; Acemoglu and Robinson, 2001; Bueno de Mesquita and Smith, 2017). Scholars have identified subalternity – defined by class, income, race, religion, or other factors (Marx and Engels, 1848; Gramsci, 1975; Fanon, 1963; Gurr, 1970; Scott, 1985) – as a key predictor of revolutions. We build on this tradition of inquiry to argue that states are particularly concerned with individuals having both high mobilization potential (through education) and high propensity for revolt (through grievances induced by subalternity).

By linking education to surveillance, our study also speaks to research on education and state control. Studies show that state-led education can foster divergent citizen responses, from greater participation (Sondheimer and Green, 2010; Marshall, 2016), to greater obedience (Cantoni et al., 2017; Paglayan, 2024). We show that it can also spark a reaction from the state: anticipatory surveillance motivated by fear of unrest. This perspective offers a nuanced picture of states’ trade-offs in expanding schooling, as in (Bourguignon and Verdier, 2000).

Our findings have significant implications for the understanding of state strategic choices and of political change. They reveal an important pattern: as subaltern groups gain

education and political voice, they also face disproportionate surveillance compared to their equally mobilized but higher-status counterparts. This asymmetry can undermine political agency, potentially muting the demands of society’s most vulnerable members, precisely when their political participation increases.

2 Origins, Institutions, and Targets of Surveillance

2.1 The System of Surveillance in Italy

Historical Evolution: The first surveillance files were written even before Italy existed: in the 1820s, the fragmented kingdoms of pre-Unitary Italy confronted waves of uprisings that culminated in the independence movements and revolutions of 1848. The largest states, the Kingdom of Sardinia–Piedmont and the Bourbon monarchy, responded by building Napoleonic-style political police: centralized offices that cultivated spies and informants and closely monitored public spaces (Di Fiore, 2018). After unification in 1861, the Kingdom of Italy inherited and expanded these practices. Rising socialist and anarchist mobilization in the late nineteenth century prompted the Interior Ministry to intensify oversight and, in 1894, to create a national registry of suspects by consolidating local dossiers that predated unification (Tosatti, 1990, pp. 447–49). World War I further expanded state capacity for surveillance through the introduction of wiretapping technologies. Under Mussolini, surveillance became more centralized and pervasive. Drawing inspiration from the Soviet Cheka, the Fascist regime created the OVRA (Organization for the Vigilance and Repression of Anti-Fascism) in 1926, expanding the use of informants and spies (Leto, 1951, p.46) and sharply increasing the number of individuals placed under surveillance each year.

Institutional Architecture: Qualitative archival work, such as Dilemmi (2010)’s study of Verona’s surveillance archives, sheds light on how individuals came under state scrutiny. Surveillance was most often triggered by participation in political parties, unions, or opposition

movements, involvement in protests or strikes or subscribing to critical publications (Dilemmi, 2010, p. 215). In the Verona archives, nearly half of all files originated from the provincial Police Headquarters (43%). Others were initiated by the central Ministry of Interior (18%) and the Carabinieri (18%) (Dilemmi, 2010, p.206). Information on political activists also came, at times, from informants – often prisoners who reported on cellmates in exchange for reduced sentences (Ciconte, 2019, Ch. 9). With the rise of fascism, additional actors, such as regime officials and affiliated organizations, became increasingly active in initiating surveillance cases (Dilemmi, 2010, p.258).

Individuals under state monitoring were generally unaware of being surveilled. A significant share, however, were also subjected to special measures – some covert, others overt – that varied in severity. These included the *rubrica di frontiera* (monitoring movements across national borders, sometimes discreetly, 23% of cases), *diffida* (a formal warning to maintain good conduct, 5.5%), and *confino* (forced relocation, 30%). Surveillance files were closed when the individual was no longer deemed a threat. This could follow a period of inactivity, withdrawal from the activities that had raised suspicion, or broader political and social changes that shifted official priorities. Closure decisions typically reflected the belief that the person no longer endangered public order or regime stability.

2.2 Targets of Surveillance: Police Narratives

To qualitatively investigate which individual characteristics the state paid attention to, we collect a sample of verbatim police reports on the surveilled. In absence of comprehensive digital archives of these narratives, we draw on the transcriptions by Ciancarini (2019), who documented the entirety of the dossiers from Civitavecchia (N=254), and by Carolini and Ecce (2015), who transcribed files on 1,000 antifascists. Neither set is random – their geographic focus and ideological mix differ from the broader population – but the narrative detail allows us to see which attributes surveillance officers flagged as noteworthy.

Qualitative evidence: Many of the narratives contained in the police documents refer to individuals who are described as having both potential for mobilization and belonging to subaltern social classes. For example:

“ Day laborer, (...) He organized armed defense through the Arditi del Popolo (note: armed defense against early fascist violence in 1921), collecting funds to arm the individual squads. He took part in all demonstrations and strikes.”

Capacity to mobilize can be linked to organizational skills, as in the example above, or simply to personal influence. This is often recognized in the police narratives, and at times is explicitly linked to education:

“Mechanic (...) is one of the few to have a fair level of education, having attended up to the first year of technical school, and for this reason he has great influence within the local movement. (...) In every political manifestation, he always stands out as the main proponent of the anarchist idea, assuming in certain circumstances the characteristics of a dangerous subversive.”

In general, reports often made explicit reference to both education and social class, often in conjunction with markers of influence.

“His profession is blacksmith [...] He enjoys a good reputation in the public. He has a bold and enterprising character. His education is above his status. Quick intelligence. Limited culture but above his status. He completed primary education.”

When dealing with individuals openly in the opposition, the state regarded knowledge as so threatening that it sought to suppress not only its dissemination but also its production. This was the case for anti-Fascist politician Riccardo Bauer, who was kept under surveillance and, once imprisoned, was forbidden not only to write but even to read (AANPIA, Riccardo Bauer records). The regime’s fear of knowledge generation is perhaps best captured in the case of one of the most prominent figures in our data, Communist Party founder Antonio Gramsci, at whose trial the prosecutor declared:

*“For twenty years we must stop this brain from functioning.”*²

Alongside education, the state systematically recorded each subject’s profession. As shown by the archival excerpts and detailed in the data section, profession was recorded for 100% of individuals under surveillance, underscoring that this attribute was fundamental to the state, on par with documenting dissidents’ ideology.

LLM Classification of Police Narratives: We analyze these narratives systematically using a machine-learning pipeline designed to extract individual traits of the surveilled. We first segment each biography into sentences, and a large language model (OpenAI GPT-4o; temperature = 0, to maximize reproducibility) flags those describing a personal trait, characteristic, or aspect of personality. For the flagged sentences, we extract trait descriptors and cluster them by semantic similarity. Human-readable labels are then assigned to each cluster by prompting the LLM for concise descriptions (temperature = 0). Out of the 25 categories initially produced, we exclude singletons and merge overlapping descriptors into broader groups, yielding twelve categories. The topic generation process is fully unsupervised – we do not provide the LLM with examples of typical traits of the surveilled. We include more details about this process in Section A.

The output of this unsupervised analysis offers additional descriptive evidence consistent with our theoretical framework. Table 1a reports the twelve individual traits emerging from this pipeline. Beyond ideology – systematically recorded across all files as the stated reason for surveillance – the unsupervised learning identifies “character” and “subversivity” as key dimensions noted in the narratives. These include attitudes against authority or traditional roles and explicit incitement to revolt, aligning closely with the theoretical expectation that states prioritize the surveillance of those likely to seek destabilizing change. The model also highlights “leadership” skills: recognized influence, mobilization capacity, being a good

²Closing speech to the Special Tribunal for the Defence of the State, Rome, 1928. In Antonio Gramsci: Pre-Prison Writings, ed. Richard Bellamy, trans. Virginia Cox, chap. XVIII, 181–183. Cambridge: Cambridge University Press, 1994.

speaker and protest organizer, are all mentioned as features of those targeted. These traits correspond to the second pillar of the theoretical framework, in which the capacity to mobilize predicts state attention. The next topic identified by the LLM is “education”, including individuals’ level of schooling, general culture, and intelligence. “Socioeconomic” characteristics, such as social status, background, upbringing, and profession, follow suit.

Both the qualitative readings of the narratives and the unsupervised text analysis underscore how the state remarked the potential for mobilization and for radical change, and in particular education and social status, as salient traits of those it surveils. While descriptive evidence indicates that those variables shaped how authorities understood those they monitored, it does not establish causality. It remains possible that it is not the potential to mobilize itself that mattered, not education, but rather factors correlated with it – such as family background, area of origin, or other fixed attributes – that actually drove surveillance. We move towards causality using a shock to education affecting otherwise similar individuals.

3 Data

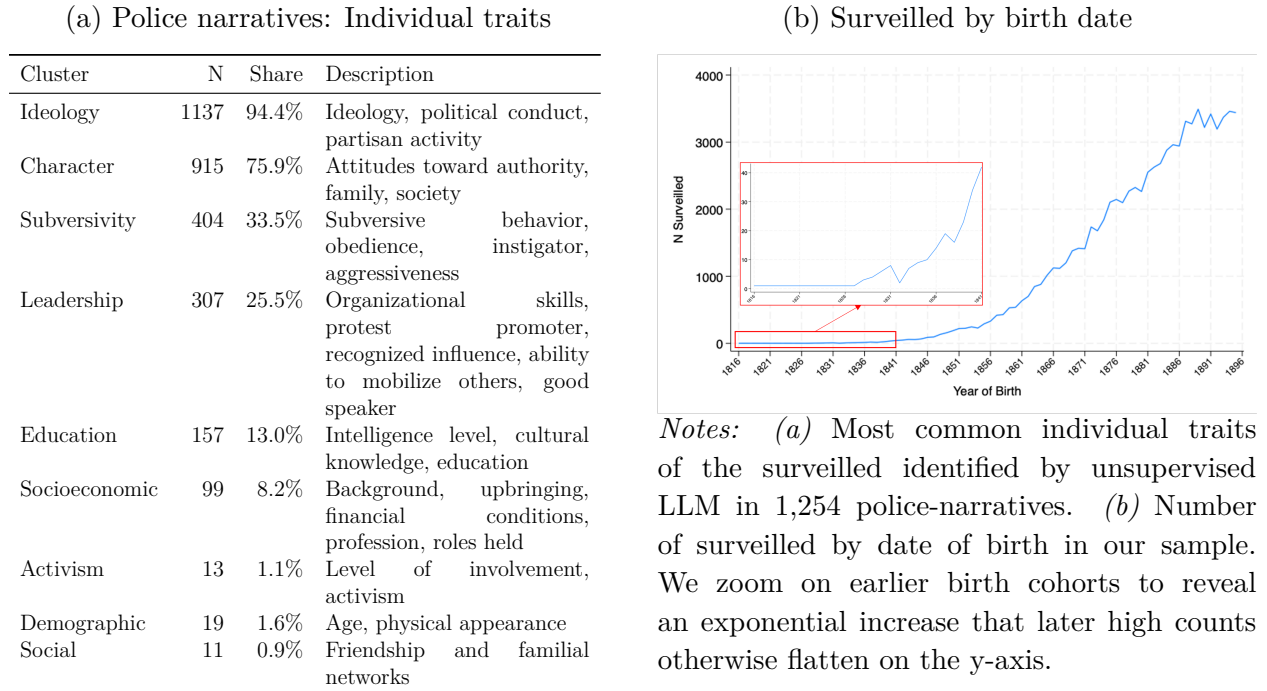
We rely on a range of historical, census, and administrative data partly collected for this project. Below, we outline the key outcomes and regressors. In Section A in the Appendix, we discuss additional historical data we collected and partly digitized.

Targets of Surveillance: Our primary outcome of interest is the identity and frequency of individuals placed under political surveillance by the state. To the best of our knowledge, this is the first study that brings detailed individual-level data on this dimension. We scrape newly released data on 152,434 individuals surveilled by the Italian State from the *Italian Central Political Surveillance Office* (Casellario Politico Centrale), an archive maintained under the authority of the Ministry of Interior. Each entry lists the individual’s name, birth date, birthplace, residence, political ideology, occupation, and years under surveillance (Figure B.1a shows an example of a record). After linking with the first population Census

of 1861 and excluding individuals born outside of Italy, we obtain a sample of 102,535 individuals born from 1816 to 1932, and surveilled from 1869 to 1947. Our empirical strategy leverages a discontinuity ending for those born after 1896, so the sample used in the main analysis includes 64,915 individuals born across 80 cohorts in all 7,713 Italian municipalities in 1861, for a total of 617,040 observations.

In Figure 1b, we plot the number of surveilled in our study period by their date of birth, the time dimension used in our analyses. Surveillance expanded steadily across cohorts, reflecting the state’s growing capacity to monitor its population. Even for the earliest cohorts, where the curve appears flatter, the number of surveilled grows exponentially. Figure B.1b presents spatial variation in the number of individuals surveilled in 100,000 inhabitants: at least one individual was surveilled in 85% of the municipalities, with a larger concentration of surveilled in Central Italy, an area that saw intense socialist activity. We also show the number of surveilled by date of surveillance (Figure B.2a), and the full distribution of surveilled by date of birth, including those outside our study period (Figure B.2b).

Figure 1: Descriptive statistics on the surveilled



Class: The archive contains information on the profession of each individual surveilled, which we use to categorize subjects into upper, middle, and working class. Because the social meaning of professions changed over time (some occupations slid up or down the status ladder), we adopt the historical scheme developed by Labini (1972) for the 1881 census, which maps occupations to social classes as understood at the time. We then adopt an iterative process to classify professions. First, we prompt a LLM classifier (GPT-4o, temperature 0.3) with Labini’s descriptions to tag every recorded occupation. Second, we read a sample of 500 classified jobs, correct classes attributed incorrectly, and feed the corrected list back into the classifier. At the next iteration, we ask the LLM to produce a confidence score for each classified profession, and review cases with low scores. More details about this process are included in Section A, SI. Figures B.3 and B.4a report, respectively, the most common jobs in each class and the prevalence of each class among surveilled, highlighting that working class individuals are the most surveilled on average.

We also use information in the files to code other characteristics of the surveilled: about half of those surveilled were communists and socialists (52%), followed by anti-Fascists (23%), anarchists (18%), and other ideologies (7%, Figure B.4a). About 3% of those surveilled were women most of whom worked as housewives (28%) or tailors (5%). Their numbers grew over time at a faster rate than men’s (Figure B.4b).

Population and other covariates: Our treatment is defined based on population data, which we gather from the Historical Censuses made available by the Italian Institute of Statistic. Our year of reference is 1861, the first Census in Italian history.³ We also digitize new data on the share of illiterates at the municipal level from the 1911 Census, the first including municipal level data on education. We collect and partly digitize a number of other

³Municipalities in the regions of Friuli Venezia Giulia, Trentino Alto Adige, in the provinces of Belluno, Frosinone, Latina, Padova, Rome, Rovigo, Treviso, Trieste, Udine, Venezia, Viterbo, started being part of Italy only after the start of our treatment, and are therefore excluded from the main sample and included in a robustness.

historical data which we discuss in greater details in Section A, SI. We also collect new data on political engagement, which we describe in Section 7.

4 Empirical Strategy

The Education Reform: Our treatment of interest is a shock in the number of years of primary schooling generated by the first education reform of the Italian Kingdom. The Casati Law, enacted across unified Italy in 1861,⁴ introduced free elementary schooling for children aged six and older and made the first two years of schooling compulsory. The law expanded the system of public primary schools, making them the main provider of basic education nationwide. Official statistics released by the Ministry of Education show that in 1884–86 about 69–72% of children aged 6–12 attended school, of whom 92% were in public schools.⁵ These high shares indicate that public schooling was not a phenomenon restricted to any single social stratum: only the nobility and the uppermost elite typically used private tutors for home instruction. Both working and middle classes (and, to some extent, segments of the upper class) sent their children to public schools.

Although the law applied universally, towns with more than 4,000 inhabitants, judged to have sufficient administrative and financial capacity to enforce the requirement, had to extend compulsory schooling by an additional two years, covering ages eight to ten (n. 3725/1859, art. 319–321), creating a discontinuity in education at this population cutoff.

Around this time, however, other laws applied above the same population cutoff. In particular, Law n. 2248/1865, Art. 15, allowed municipalities above the threshold to create fractions of the municipality, and the Royal Decree n. 20/1896 applied different provisions for customs to municipalities above the threshold. To account for this, we leverage also the

⁴The law was initially approved in 1859 in the Kingdom of Sardinia, but its implementation was delayed due to the unification war. Before this time, no laws requiring compulsory education had been implemented (Sotgiu, 1984).

⁵Ministero dell'Istruzione, Condizioni dell'istruzione elementare in Italia e progressi dal 1861 (Rome, 12 April 1890), Table V, p. 32.

temporal component of the education reform: children who were seven or younger at the time of the law enactment (e.g., born in or after 1854) and living in larger municipalities experienced an education shock compared to their peers born just before or in smaller towns.

This discontinuity ended in 1904, when the Orlando Law extended compulsory schooling to age twelve in all municipalities, regardless of population size. To avoid contamination from this reform, we restrict our analyses to cohorts born before 1896, who would have completed schooling before the new law took effect.

Identification Strategy, DDRD Design: We estimate the effect of the education reform on surveillance using a difference-in-discontinuities (DDRD) design (Grembi, Nannicini and Troiano, 2016). This design combines two exogenous sources of variation: (i) a cross-sectional discontinuity at the 4,000-inhabitant cutoff, above which municipalities extended compulsory schooling by two years, and (ii) a temporal discontinuity between cohorts who were seven years old or younger in 1861, and therefore remained in school longer. Because treatment is assigned by birth year, treatment effects are unrelated to the calendar year of either the law’s enactment or later surveillance activities: individuals could be monitored at any point in their adult lives. Formally, for municipality i and cohort t we estimate:

$$\begin{aligned} Surveilled_{it} = & \beta (Post \times Above4000_{it}) + \lambda(f(Pop1861) \times Post_{it}) \\ & + \eta (f(Pop1861) \times Post \times Above4000_{it}) + \alpha_i + \theta_t + \kappa_{r \times t} + \epsilon_{it} \end{aligned} \quad (1)$$

where $Surveilled_{it}$ is the population share of individuals surveilled in each municipality–cohort, $Post$ is an indicator equal to 1 for cohorts born after 1854 (seven or younger in 1861), $Above4000$ is a dummy for municipalities above the threshold, and $Pop1861$ is the running variable. All regressions include municipality fixed effects, cohort fixed effects, and reign–cohort fixed effects. We use reign rather than region fixed effects because the sample begins before unification, but we present robustness using regions.

The coefficient of interest is β , capturing the average causal effect of the reform at the

discontinuity. We control flexibly for population size on either side of the threshold (λ), and of the temporal discontinuity (η) to isolate the effect of the reform itself. Standard errors are clustered at the municipal level.

Concretely, we compare children born in the same year who lived in similar towns differing only in whether they crossed the 4,000-inhabitant threshold, and we difference this contrast against slightly older cohorts who were already too old to be affected by the reform. This design rules out confounders that vary only at the municipal or cohort threshold – for example, if municipalities above 4,000 had higher ex-ante capacity, or received more funds after the reform, this is not an issue for identification, which relies also on the difference across cohorts within the same city.

Identifying Assumptions: We provide several pieces of evidence to support the validity of this approach. First, we find no evidence of sorting: a manipulation test shows that the density of the running variable is constant across the cutoff (Figure C.1, $p = 0.796$). Second, we assess the credibility of the continuity-in-potential-outcomes assumption by showing that pre-treatment covariates are balanced across the discontinuity threshold (Table C.1). Third, there should be no interaction between treatment and the confounding policies (municipal fractions and customs policy, Assumption 3 in Grembi, Nannicini and Troiano (2016)). This assumption would be violated if, for example, cohorts above the population threshold reacted to the creation of municipal fractions or to the customs policy in different ways also that impact surveillance. We show, however, that there are null effects of the 4,000 cutoff on surveillance for cohorts not affected by the education reform but affected by the creation of municipal fractions and the customs policy (Section 5, Placebos). Finally, as for the difference-in-differences design, there should be no pre-existing differences in surveillance between treatment and control municipalities before the onset of the education reform after controlling for fixed effects. We present evidence supporting the absence of pre-trends in the event study analysis (Figure 3a).

Alternative Strategy, RD Design: The difference-in-discontinuities strategy requires a demanding set of data, including individual-level information on date and place of birth. For many of our secondary analyses, this level of detail is often unavailable as we can rely only on municipal-level or municipal-year variation. When only municipality-level data are present, in a minority of cases, we use a simple regression discontinuity design (RDD) that exploits only the population threshold and controls for region fixed effects:

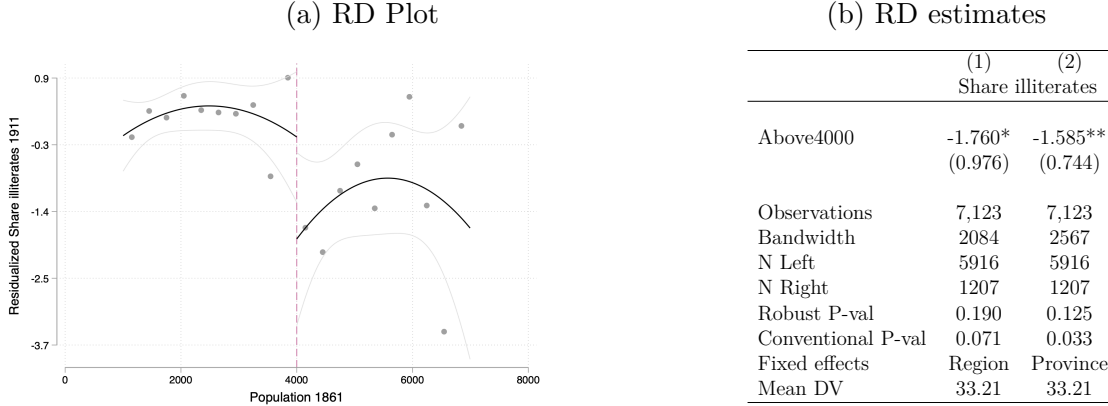
$$Y_{it} = \mu Above4000_i + \nu f(Pop1861)_i + \xi f(Pop1861) \times Above_i + \rho_r + \pi_i \quad (2)$$

When year is present, as in most of the secondary analyses, we also include year fixed effects and show year-by-year RDD estimates. While we provide evidence in line with continuity of potential outcomes (Table C.1) and absence of sorting (Figure C.1), this specification is less precise than the DDRD, since other laws also introduced discontinuities at 4,000 inhabitants. Given that the 4,000-inhabitant cutoff was not unique to the education reform, the DDRD is our preferred identification strategy, and we use the RDD only when data limitations preclude adopting the DDRD. Results from the latter strategy should therefore be interpreted with more caution.

Effect of the Reform on Literacy: We use the alternative identification strategy described above to test whether the Casati reform expanded schooling by examining its long-run effects on literacy. To do so, we digitize the first available municipal-level illiteracy data from the 1911 Census – forty years after enactment and before later schooling reforms – and estimate a cross-sectional regression discontinuity design. Treated municipalities exhibit 1.6–1.8 percentage points lower illiteracy rates, corresponding to a 4.8–5% reduction from baseline levels (Table 2b). These estimates are statistically significant only considering conventional (not robust) p-values.

Two additional years of schooling represented a 100% increase over average school attainment in our setting, and at a time when most Italians were illiterate. The first two years of schooling focused on literacy, mathematics, and religion; the additional two introduced

Figure 2: Education Reform Impact on Illiteracy (RDD)



Notes: (a) RD results (Equation 2) of the impact of the 4,000 population cutoff at which the education reform applied on municipal-level illiteracy rates in 1911. We consider residualized illiteracy rates after accounting for province fixed effects. (b) RD coefficients with region and province fixed effects. Standard errors are clustered at the municipal level. Estimates are computed with triangular weighting kernel, data-driven optimal bandwidth selection and default binning.

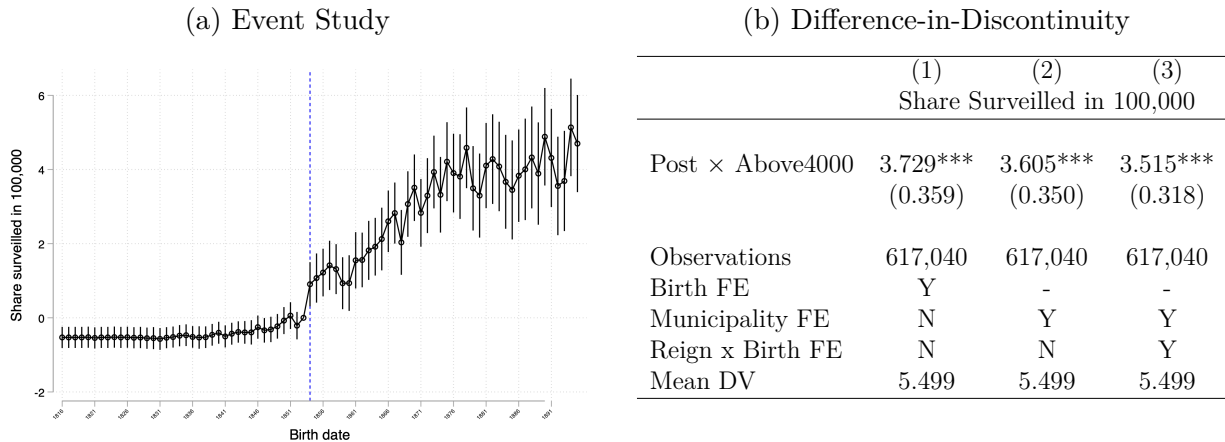
history, geography, the physical and natural sciences, and literature. Given this curricular expansion, the reform's effects are unlikely to be confined to basic reading skills: it plausibly broadened intellectual horizons, encouraged reading beyond the prescribed curriculum, and raised the likelihood of continuing past elementary school. Although literacy is the only outcome systematically measured for all municipalities, this is a proxy for these other broader gains: any observed reductions in illiteracy likely understate the reform's total impact on educational development.

5 Effect of the Education Reform on Surveillance

Our main question is whether increased schooling led to a higher probability of a municipality-cohort being targeted by state surveillance. In Table 3b, we present our main results. Municipality-cohorts exposed to the education reform experienced 3.5-3.7 more individual surveilled in 100,000 inhabitants compared to control group municipalities-cohorts. This effect corresponds to a 64% increase with respect to the mean of the dependent variable in municipalities just below the 4,000 cutoff. Assuming 50 cohorts were alive at the same time

per municipality,⁶ the number of surveilled in the average treated municipality passed from 27.5 to 47.5 in 10,000 inhabitants. Effects are similar in magnitude across specifications that do and do not include fixed effects, in line with the possibility that we are isolating the impact of a quasi-random treatment. We present this analysis as an event study in Figure 3a. In the thirty-eight cohorts preceding the first treated cohort, there was no significant difference in surveillance between municipality-cohorts that will later be exposed to the reform and those that will not. The unusually long pre-treatment period in this setting provides stronger support for the parallel trends assumption than is typically possible in difference-in-differences designs. The pre-trend is slightly negative, driven by a modest increase in the four years preceding treatment among municipalities above 4,000. This pattern is consistent with some individuals becoming treated early as a result of grade repetition.⁷

Figure 3: Effect of the Education Reform on Surveilled in 100,000



Notes: (a) Event-study and (b) difference-in-discontinuity estimates (Equation 1). *Above4000* is=1 for municipalities > 4,000 inhabitants and *Post* is=1 for cohorts affected by the reform (born ≥ 1854). The reference cohort is 1853. Regressions include municipality, and reign × birth-year fixed effects. We report the mean of the dependent variable for municipalities with population just below 4,000 after the reform within the RDD optimal bandwidth. Standard errors are clustered at the municipality level.

⁶Based on life expectancy in 1910 in Italy (Caselli, 1990).

⁷For instance, a child born in 1853 who fails in second grade would repeat the year once the reform is in place, thereby continuing their education beyond grade two and qualifying as treated.

Beginning with the 1854 cohort, we observe a stark increase in surveillance, corresponding to an additional 0.4 individuals surveilled per 100,000 in the first treated cohort. This effect grows over time, reaching an additional 4 individuals surveilled per 100,000 by 1890. The increasing impact of the reform over time is consistent with the state’s gradual expansion of its capacity to enforce mandatory education, and the broader uptake of educational resources, as schooling became less of an elite privilege. While detailed time-varying data on the expansion of students’ enrollment in schools in this period is absent, preventing us from providing direct evidence on when and where education became more inclusive, we include fine-grained location fixed effects and time-varying regional trends to account, for example, for time-varying underdevelopment in different regions of Italy.

End of the discontinuity: In 1904 the population discontinuity ended: compulsory schooling was extended to all municipalities, eliminating differences between treatment and control areas. If the increase in surveillance was produced by the education reform, we should see it disappearing together with the differences between cities. The extended event study in Figure D.1 shows that the difference in surveillance for municipalities above 4,000 decreases precisely starting from 1896 – the birth-date of individuals who turned 8 years old when the new reform was enforced, now exposed to the same education requirements as children in control municipalities. As for the main result, the difference between previously treated and control municipalities does not disappear immediately: places that had built school capital over time maintain an advantage in effectively providing education compared to newly treated locations, resulting in a slowly decreasing impact on surveillance. Over time, the impact converges to zero, with surveillance rates equalizing across municipalities on either side of the original discontinuity.

Redirection of surveillance resources: If states face resource constraints, expanding surveillance in treated municipalities should crowd-out surveillance elsewhere. We test this by comparing post-reform surveillance rates among cohorts born after the reform in control

municipalities. Specifically, we re-estimate our baseline specification without the flexible date-of-birth-by-reign controls to recover separate post-reform effects for treated and control municipalities. We find a 12% decline in surveillance in below-cutoff municipalities after the reform (Table D.2, Column 1), an effect smaller than the contemporaneous increase in treated areas. This pattern is consistent with both resource reallocation and additional investment to expand overall surveillance capacity.

Placebos: We present two sets of placebo analyses to rule out the possibility that other factors cause the jump we observe at the education reform discontinuity. First, we consider effects at other population thresholds below the 4,000 discontinuity. We find null effects for all the round placebo thresholds below the correct cutoff in Figure D.3a. Second, we repeat the same exercise for cohort of birth and ask whether being born in another cohort unexposed to the reform is likely to trigger the same effect by replacing the *Post* indicator with a set of placebo cohorts. Findings presented in Figure D.3b show that there was no significant increase in surveillance in municipalities above 4,000 in non-treated cohorts.

Robustness: We start by considering robustness to alternative specifications in Table D.1. First, we control for the rate of surveillance in different periods of Italian history. We create a vector of controls, one for each decade in the calendar years of surveillance, measuring the number of individuals from each birth cohort who started being surveilled during that decade. Results are similar in magnitude and significance when including calendar-year controls (column 1). Second, we include areas of Italy that were annexed to Italy after 1861, where the education reform was implemented later, and are thus excluded from the main analysis. Results are robust to estimating a staggered DDRD which accounts for the effective year of the reforms' implementation in later-annexed areas (column 2). Column 3 shows that replacing reign (defined pre-treatment) with region (post-treatment) times birth-year fixed effects does not affect our results. In column 4 we show consistent results using a Poisson specification. columns 5 and 6 consider higher-order polynomials.

Table D.2 tests robustness to alternative transformations of the dependent variable. Using a dummy equal to one if any individual was surveilled in a municipality-cohort, we find a 28% higher likelihood of surveillance following the education reform (column 2). By contrast, considering the absolute count yields a null effect (column 3), which is consistent with a highly zero-inflated and right-skewed distribution. In fact, by trimming the upper tail via winsorization (column 4) or modeling proportional changes via a log transformation (column 5), we observe again positive and significant effects.

In Table D.3, we replicate our results using different designs. We start from a less restrictive difference-in-differences (DID) which does not flexibly control for proximity to the population cutoff and leverages the entire distribution (columns 1-3). We also use a more restrictive model flexibly controlling for proximity to both the population and the year of birth discontinuities, such that we focus on observations across both the geographic and the temporal discontinuity, in the spirit of a “double” difference-in-discontinuity design (columns 4-6). Finally, we remove the interaction with Post from the flexible controls to estimate a regression discontinuity version of Equation 1 in columns 7-8. Our findings are robust across these designs.

6 Surveillance of the Working Class

The education reform led to increased surveillance among treated cohorts. To assess whether this effect is driven by state concerns over the education of subaltern groups, we classify the professions of surveilled individuals into distinct social classes, as detailed in the Data section and illustrated in Figure B.3. We then use the same specification as in Equation 1 on three distinct dependent variables: surveilled from the working, middle, and upper class. The results, presented in event study (Figure 4a) and table format (Table 4b), indicate that the increased surveillance following the reform is primarily driven by the working class. The increase for this group is large: there are 2.9 more individuals surveilled in 100,000 – a 60%

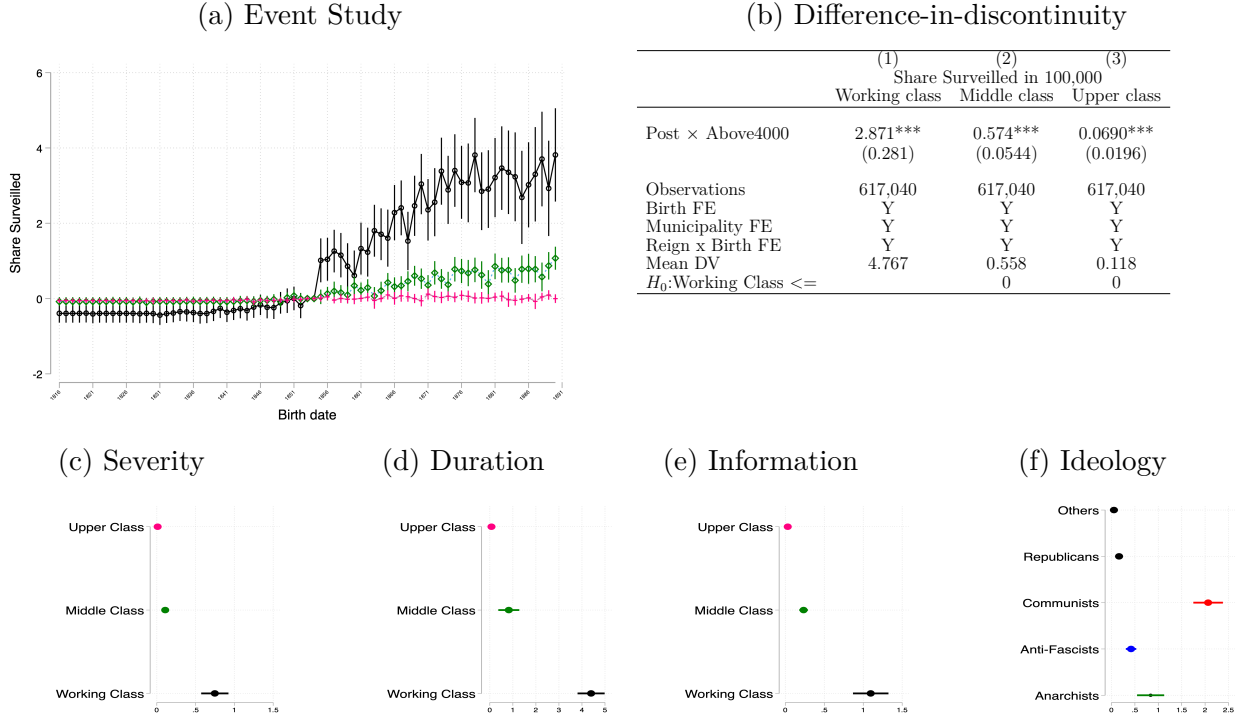
increase with respect to municipalities just below the 4,000 the cutoff after the reform. We observe a positive but smaller effect for the middle ($\beta=0.57$, column 2) and upper classes ($\beta=0.07$, column 3). The last row reports the p-value from a t-test comparing working class to middle and upper class coefficients. Both p-values indicate that effects for the working class are significantly larger than for other classes. The differential increase in surveillance for this social group is also visible in the event study plot. Coefficients remain consistent (if anything, larger) when we use a binary indicator for the presence of any surveilled individuals in the working class: we observe a 0.24 percentage points increase for working classes, as opposed to 0.04 and 0.004 for the middle and upper class respectively (Table E.1).

A potential concern with interpreting this analysis is that the number of individuals in the working class could itself be influenced by the education reform. However, given that the reform expanded access to literacy and likely improved the labor prospects of children born in low income families, it most likely reduced the overall share of working class individuals in the post-reform period. This would imply that our findings represent a lower bound of the effect, as we cannot directly observe the fraction of the population moving to the middle class due to improved opportunities. Therefore, the observed effects of surveillance in this analysis should be interpreted as a conservative estimate of the actual impact.

6.1 State Behavior Consistent with Threat Perception

If surveillance increases because the state fears that education equips subaltern groups with political skills, then we should observe not only higher surveillance rates but also more intensive forms of monitoring directed at educated working classes. We evaluate this prediction using three measures of surveillance intensity – the severity of control, the duration of surveillance, and the volume of information collected – and by examining whether surveillance disproportionately targeted certain ideological orientations.

Figure 4: Effect of the Education Reform on Surveillance by Class



Notes: (a): DDRD eventstudy using as dependent variables the population share of surveilled in the working (black), middle (green) and upper (pink) class. (b): DDRD coefficients (Equation 1). We report the mean of the dependent variable for municipalities just < 4,000 inhabitants after the reform. The bottom row reports the p-value from a one-sided test of the null hypothesis that the effect on the working class is statistically lower than that of the other class. (c-e): DDRD coefficients from separate regressions using as dependent variable the population share of individuals under severe forms of surveillance, the average number of years a municipality-cohort spends under surveillance, and the population share of individuals with a high volume of information collected on them in the working (black), middle (green) and upper (pink) class. (f): DDRD coefficients from separate regressions using as dependent variable the population share of surveilled from each ideology. All regressions include municipality, and reign times birth-year fixed effects. Standard errors are clustered at the municipal level in all regressions.

Severity: First, we consider whether treated cohorts were subjected to more intensive forms of surveillance: monitoring movements across borders, notice of holding a good conduct, and political confinement. Table E.2, column 1, shows that municipality-cohorts affected by the education reform experienced higher rates of severe surveillance. Consistent with the state perceiving a threat and responding with harsher measures, the increase in severe control is concentrated among the working class (Figure 4c and Table E.2, $\beta=0.77$, column 2) and less pronounced for other groups ($\beta= 0.11$ and 0.01 , columns 3-4).

One could wonder if the increase in surveillance for educated working classes is entirely the result of the state switching to overt forms of control as a strategy to advertise state monitoring among these groups. However, the increase in surveillance is also present when we consider surveillance that does not use intensive methods of control, indicating an overall uptick in control of subaltern groups by the state (Table E.3).

Duration: The second measure considers how long individuals remained under surveillance. Duration provides a measure of the intensity and persistence of state attention: longer surveillance implies greater resource investment, consistent with a higher perception of threat. Given that our time dimension is a cohort, we express this duration in terms of how long a cohort was surveilled for in a municipality. On average, cohorts in control municipalities after the reform are surveilled for 2.1 years. The reform significantly extends surveillance duration: treated cohorts are surveilled 5.2 years longer, a 2.5-fold increase (Table E.4, column 1). As with the severity of surveillance, this effect is most pronounced among the working class, where surveillance duration increases by 4.4 years (Figure 4d and Table E.4, column 2), exceeding the increase observed for middle ($\beta=0.82$, column 3) and upper-class individuals ($\beta= 0.07$, column 4).

Volume of Information: The third measure examines whether authorities collected extensive information on the surveilled, as indicated by the presence of a biographical file (“scheda biografica”). These files were maintained for all significant political figures or for individuals

considered particularly dangerous. Descriptive statistics show that such files were more frequently recorded for working class individuals (81.17%) compared to the middle (15%) and upper-classes (2.81%). Was this pattern further enhanced by quasi-exogenous exposure to education? Table E.5, column 1, shows that treated cohorts were twice as likely to have a biographical file, with the share of surveilled individuals with biographical documentation increasing by 1.36 percentage points from a mean of 1.17. This increase is driven by the working class (Figure 4e and Table E.5, $\beta = 1.1$, column 2), which experiences an increase larger than for any other class (0.23 and 0.03 for middle and upper class individuals, respectively, in columns 3 and 4). Across all three measures, the evidence shows that the state intensified surveillance following the education reform, adopting harsher methods and more intensive control practices on newly educated cohorts, and that the sharpest increase was targeted at working-class individuals.

Ideology of Surveilled: A final piece of evidence in line with threat perception comes from examining whether ideological orientation shaped who was targeted. If surveillance reflected state concern for educated working classes, we should observe disproportionate monitoring of ideologies advocating on their behalf. Table E.6 and Figure 4f present the effect of the education reform on surveillance by ideological affiliation. While all political groups experienced increased monitoring following the education reform, socialists and communists were surveilled at the highest rates. This effect does not simply reflect that all working classes were socialist and communist, as we observe effects also for socialists and communists not in the working class. This pattern aligns with the state targeting ideologies addressing the demands and grievances of subaltern classes.

A test in a similar vein could assess if targeting varied with the ideology of the governing party. Before Fascism, two parties alternated in power: the Historical Left, which supported secularism and centralization, and the Historical Right, which championed monarchical authority, conservative values, and local autonomy. Neither party, however, represented

the interests of the working classes, as both were elite projects dominated by landowners, intellectuals, and the bourgeoisie. The working class was in fact disenfranchised from voting in the entire period until 1921, when Fascism starts. Consistent with this, our results reveal little partisan heterogeneity: surveillance of the working class is statistically indistinguishable between Historical Left and Right (Table E.7), underscoring the exclusion of the working class from meaningful political representation.

6.2 Additional Heterogeneities

Democracy and Dictatorship: A substantively important question is whether the strategic use of state surveillance we document in these analyses varies across regime types. To answer, we leverage our extended time period which spans both democratic governments (1869–1921) and the fascist regime (1922–1947). Table E.8 presents results separately for these two periods. The findings indicate similar patterns across both regimes. Under democracy (columns 1-4), the likelihood of surveillance for newly educated cohorts increases by a factor of 1.54, while under fascism (columns 5-8), it increases by a factor of 1.33. In both cases, the effect is largest among the working class. The consistent patterns across regimes suggest that educated yet subaltern individuals remain a primary focus of state surveillance, regardless of the type of regime in power.

State Capacity: State capacity is another potential mediator of surveillance whose impact is ambiguous. If capacity is high, the state might be able to surveil more people. If limited however, it might have to target even more those representing a threat. We consider the impact of state capacity by using both geographic heterogeneities and a new measure of state presence. To assess if state capacity is a mediator of our findings, we digitize new data on postal offices in the Kingdom of Italy, sourced from an 1861 official postal bulletin. Columns (1) and (2) in Table E.9 show that the targeting of educated groups is more prevalent in areas with postal offices, yet the effect remains significant even where none existed. Columns

(3)-(5) examine geographical variation in treatment effects, revealing that while the impact of the education reform on surveillance is stronger in Central Italy, significant effects are observed across all regions. Taken together, these patterns indicate that greater capacity facilitated intensified surveillance of targeted individuals, while limited capacity did not prevent such targeting, at least within the range of capacity observed in Italy at the time.

Technology: The technology employed for surveillance could also influence states’ strategies, for example by reducing the need for targeting once data collection is easier. We evaluate this possibility by leveraging the introduction of telephonic wiretapping as a tool for intelligence gathering in 1916. Column (6) of Table E.9 shows no statistically significant difference in the rate of surveillance in response to the education reform after the introduction of wiretapping. An innovation enhancing state capacity to collect intelligence did not change the strategic calculus of the state to target educated subjects.

Pre-Reform Surveillance Levels: We consider whether the increase in surveillance is driven by newly monitored areas or by heightened scrutiny in already highly surveilled regions. Table E.10 shows that both channels contribute to the observed effects. Treatment effects are larger in areas with higher pre-reform surveillance, possibly due to spillover from surveilling more individuals, or network-driven mobilization. However, the effect remains significant even when accounting for pre-reform surveillance levels.

7 Education and Political Engagement

The increase in surveillance could, in part or entirely, be a response to increased political mobilization caused by education. We test whether and to what extent this explanation fits the data by measuring the impact of the education reform on political engagement. To explore this possibility extensively, we use and partly collect five different indicators of political engagement that range from low-cost high-frequency actions, like turning out to

vote, to actions requiring substantial commitment and risk, such as taking up weapons against the state. Since each data source varies at a different level of aggregation, we adopt distinct empirical strategies tailored to each measure. In every case, our preferred specification is the most restrictive available, but we also report all results using the minimal common denominator – a cross-sectional RDD.

Voter turnout: We begin by examining the impact of the education reform on voting in national elections. We gather information on the number of votes cast in each municipality and in each election from 1861 to 1913 from the *Electoral Historical Atlas* archive.⁸ Because the number of eligible voters is not reported, we normalize votes by the 1861 population – the last population census before elections began.

The first treated cohorts became eligible to vote in 1880, when the first treated cohort (born 1865) reached the voting age of 25. In this initial treatment year, only a small share of the electorate had received the additional schooling. Because election data is at the municipality-election year level rather than at the municipality-cohort level, and because treatment intensity increases gradually as more cohorts entered the electorate, the time discontinuity is less informative in this context. Thus, we present results using yearly RD coefficients for the effect of the 4,000-inhabitant cutoff (Equation 2), allowing us to trace the evolution of the reform’s impact over time independently of pre-1880 differences. We also report DDRD estimates in Appendix, which yield a consistent picture.

We start by showing that, before the first treated cohort started voting, municipalities above the 4,000-population threshold voted at about the same rate as those below (Figure F.3),

⁸We focus on elections held up to 1913, after which the electorate included cohorts no longer exposed to the discontinuous variation in education (i.e., those born after 1896, for whom the 4,000 cutoff no longer applied). Data on votes by party are unavailable; only the votes for winning candidates are reported. However, since candidates are defined at the constituency level (above the municipality), this prevents us from examining either the effect of the education reform on support for specific parties or on the electoral success of particular candidates.

in line with the parallel trends assumption. We then consider the impact of the education reform after the first cohort gained voting rights.

Our results show a null effect: there is no differences in voting in municipalities-election years affected by the education reform (Figure 5, panel (a)). Findings are robust both considering the year-by-year RD coefficients, and the overall effect of the population discontinuity in the post-period (Figure F.1, panel (a), and Table F.1, columns 1-2), as well as in the DDRD using different transformations of the dependent variable (Table F.1, columns 3-4).

Notable people involved in politics: Our second measure of political engagement is the number of individuals who held political roles, which we scrape from WikiData. We scrape information about all historical figures born in our sample period (born 1816-1895) and collect their birth date and place, occupation, and other significant roles held. Using occupation and significant roles, we code notable individuals who were involved in politics, ranging from members of parliament to activists.

Because in this case we observe individuals' date of birth and have sufficient observations before and after the education reform, we can replicate the same municipality-cohort specification as in our main analysis. The DDRD event study in Figure 5, panel (b), shows that, before the reform, individuals in municipalities above and below the 4,000-inhabitant threshold held political positions at similar rates. After the reform, this pattern remains unchanged – if anything, the estimates suggest a slight negative effect on political engagement. Table F.2 confirms the finding both for individuals in politics and for notable figures in other sectors.

Protests and strikes: Third, we consider a costlier but more widespread form of activism than holding political roles: participation in protests and strikes. We scrape information on the location and timing of protests from the archives of one of the main Italian newspapers, *La Stampa*, the only including articles dating as far back as 1867. We identify the location of the event by writing an algorithm that searches the municipality name in the text of the

article, and excluding municipalities with names having double meaning.⁹

In this case, the data is at the municipality-year level and we do not have a pre-period before the reform, as data is only available from 1867. Because we cannot exploit the temporal component, we present regression discontinuity coefficients for the impact of the 4,000 inhabitants discontinuity in each year. Consistently with the previous results, our findings indicate no significant increase in protests and strikes in municipalities above the 4,000 discontinuity in each post-period year (Figure 5, panel (d)), and overall (Figure F.1), independently on transformations of the dependent variable (Table F.3).

Partisans and Bandits: Finally, we examine two of the highest-cost forms of political activism: participation in the resistance movement and banditism. Both were armed movements against the state, and political action entailed a substantial risk of death.

Partisans were guerrilla fighters who opposed Fascist and Nazi forces through acts of sabotage and armed combat. We scrape data on the universe of partisans born between 1816 and 1895 (N=19,341) from the official database by the Ministry of Culture, which reunites regional sources that, after World War II, collected the lists of those part of the resistance movement.¹⁰ The data on partisans includes their date and location of birth, but too few of them were born pre-education reform to meaningfully exploit the time variation.¹¹ Therefore, also in this case, we present yearly RDD coefficients as the main specification,

⁹We excluded from the sample 0.5% of municipalities called with common nouns, such as Porte (“doors”) or Ora (“hour”). For a random subset of cases, we manually reviewed the articles mentioning them and confirmed that the mentions referred to the common noun rather than the municipality.

¹⁰It is not possible to match partisans and surveilled: partisans started their activity in the summer of 1943, after Mussolini was arrested and a temporary government was put in place in Rome, while Northern Italy was under German occupation. Partisans surveillance files might be present in the German archives, but the Italian secret police largely halted its surveillance activity in this period, as is visible in Figure B.2a.

¹¹While the number of observations in the pre-period is not zero, partisans started operating in 1943, so those born pre-reform were 90 years old when the resistance started.

and show consistent DDRD results in appendix.

We find a null effect of the education reform on the population share of members of the resistance movement (Figure 5, panel (d), and Table F.4, column). Results are similar in a DDRD design (columns 3-4) and considering a RDD with year fixed effects (Figure F.1).

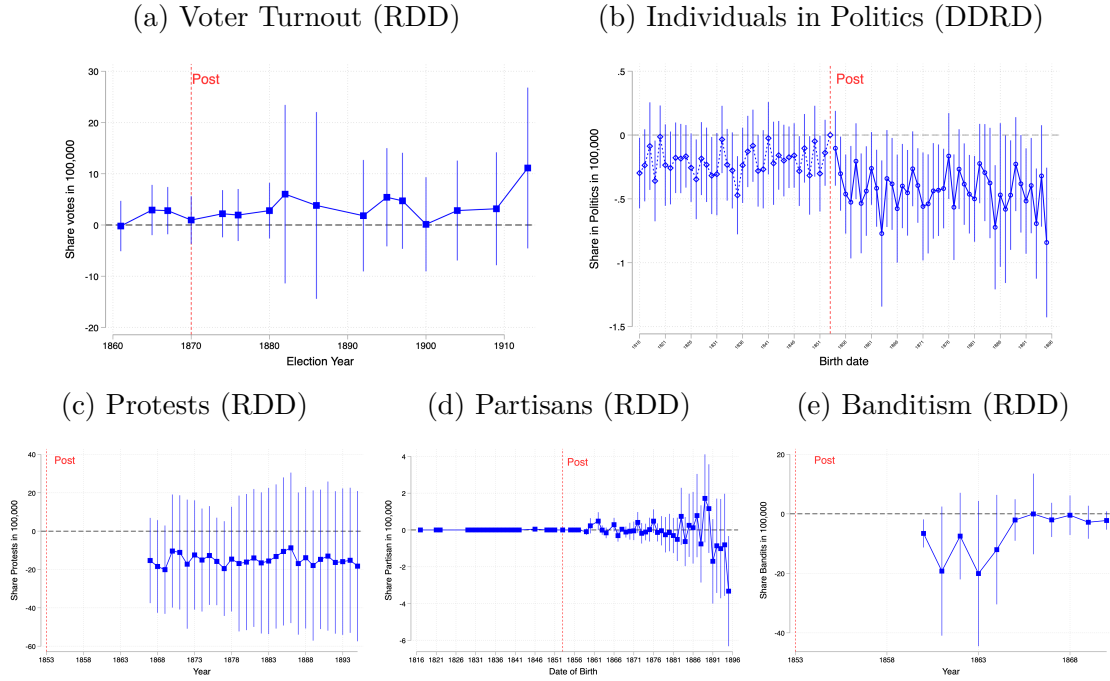
Bandits, or briganti, were individuals who operated outside the law, engaging in raids, kidnappings, and thefts that primarily targeted wealthy landowners. Their actions, while criminal in form, often carried a political dimension, functioning as a form of insurgency against the newly unified Italian state. Particularly widespread in southern and central Italy, regions marked by deep social and economic inequalities, banditry has been described by historians not merely as lawlessness but as a manifestation of popular rebellion against an unjust social order (Ciconte, 2011).

We use data collected by Lecce, Ogliari and Orlando (2022) on the municipal-level number of bandits in Southern Italy (excluding Sicily) from 1860 to 1870. Also in this case, we employ yearly RDD coefficients given the presence of few years of data in the post-period only. Figure 5, panel (e) shows a null effect of the education reform on the number of bandits, results confirmed in Figure F.1e, examining the full period with year fixed effects.

Interpreting the absence of effects on political engagement We have examined a wide range of political engagement outcomes, and the evidence consistently shows that exposure to additional primary education did not translate into higher levels of political participation. This finding may appear to stand in tension with the increase in surveillance we document. If education did not foster greater political activism, why would the state invest additional resources in monitoring these individuals?

We propose two interpretations. The first is that political participation did not increase precisely because the state’s preventive surveillance strategy succeeded in deterring it. Under this account, increased surveillance produced a chilling effect on political mobilization. The second possibility is that participation did not increase in the first place, independently

Figure 5: Effect of the education reform on political engagement



Notes: The outcome is the share of votes cast in 100,000 inhabitants by election year (a), the winsorized share of protests in 100,000 by each calendar year (c), the population share of partisans in 100,000 by date of birth (d) and the share of bandits in 100,000 by calendar year (e). The specification includes RD coefficients (Equation 2) in each available year for being above the 4,000 population discontinuity in panels (a, c-e). Estimates are computed with triangular weighting kernel and data-driven optimal bandwidth selection. In panel (b), we consider a DDRD event study (Equation 1) of the effect of the education reform on the share of individuals in politics in 100,000 inhabitants in a municipality-cohort. Standard errors are clustered at the municipal level in all analyses.

of surveillance, meaning that state surveillance was not a response to an actual increase in activism. This would be consistent with recent studies showing that the content of education matters to determine whether this leads to more activism or more obedience with the state.

Our evidence cannot adjudicate decisively between these two interpretations. Descriptive exercises, however, provide little support for the chilling-effect interpretation: we find no differences in political engagement between cohorts that were and were not subject to surveillance pre-reform (Table F.6) and that were subject to high or low levels of surveillance (Table F.5). While these analyses suggest that surveillance does not appear to explain the absence of greater political engagement, they are endogenous and we do not interpret them as conclusive of one or the other interpretation.

Given that education did not trigger higher participation, does this imply that the state’s heightened surveillance was irrational? We argue that this is not the case. Identifying future political leaders ex-ante is notoriously difficult, and the potential payoffs to early detection are extremely high. As in the case of innovators, a few whose work can have large and lasting spillovers (Akcigit, Baslandze and Stantcheva, 2016), also revolutionary entrepreneurs can have wide-ranging consequences – in this case, for states’ stability (De Mesquita, 2010). Therefore, from the state’s perspective, even a small number of empowered members of subaltern groups could pose enough risk to justify increased surveillance of many of them. For this reason we argue that, even if education did not trigger widespread mobilization, the state’s calculus of targeting this group at higher rates might still be correct.

8 Alternative Explanations

Other municipal-level changes induced by the education reform: The education reform may have influenced other aspects of treated municipalities: they might receive more funds, or gain better capacity. For these changes to represent alternative explanations for our findings, they should vary discontinuously not only by municipality, but also by cohort.

For example, the reform could have increased education levels among local institutions, such as the police, or in the general population. A better-educated police force might be more effective at monitoring individuals, increasing surveillance. For such a mechanism to account for the jump in surveillance we observe, (i) the increase in surveillance would need to be driven specifically by the newly educated cohort of policemen or citizens; and (ii) these actors would need to selectively target individuals from their own age cohort, since our results are tied to the age of those surveilled rather than the timing of surveillance.¹²

We empirically rule out this class of alternative explanations by running our analyses on individuals who relocated away from their municipality of origin. Because these individuals are no longer exposed to changes in municipal-level characteristics, any local institutional or compositional shifts – such as changes in the education level of local institutions or residents – should not affect their likelihood of being surveilled. As shown in columns (4)–(6) of Table G.1, the results for this subsample closely mirror those reported in Table 3b, suggesting that migration does not attenuate the treatment effect. The finding also replicates if we consider individuals moving exclusively to control municipalities, where any local changes produced by the education reform would be absent (columns (7)–(9)).

This finding does more than rule out a confounding channel: it reinforces our interpretation of the mechanism. The persistence of the effect among movers indicates that the treatment operates through a portable individual characteristic –namely, human capital– rather than through fixed local attributes.

Migration of educated individuals: Conversely, we also show that migrants do not drive our findings. This could be a concern if we consider that educated individuals might be more likely to migrate to larger cities for better opportunities, and there they might encounter heightened political activism. In this scenario, increased surveillance could stem

¹²In the case of law enforcement specifically, officers were typically assigned based on institutional needs rather than place of origin, making it unlikely that they would serve in their city of origin. See, for example, Arma dei Carabinieri, Chapter XII – Transfers.

from education-driven migration, as individuals with higher education gravitate toward politically active urban centers. To rule out this possibility, we limit the analysis to those who remained in their birthplace. Results in Table G.1 confirm our primary findings even among non-migrants: treated municipality-cohorts experience an increase in the number of surveilled individuals to about 0.78 times the baseline –a result consistent with those shown in Table 3b. Therefore, while education may indeed have influenced migration and increased exposure to activism, this does not appear to be the main driver of our results.

Teachers as state agents: One possibility is that teachers, as part of the bureaucracy, may have acted as informants for the state, monitoring former students and reporting on their political activity. While this possibility is unlikely in light of the first test on surveilled who migrated away from their municipality, we investigate it further. If state bureaucrats were used as informants, we would expect that also other reforms increasing the size of the bureaucracy would lead to a similar surge in surveillance. To rule out this possibility, we examine a non-education placebo reform: the Crispi Law of 1891, which required municipalities with more than 5,000 residents to expand their welfare boards, thereby substantially increasing the local administrative workforce. The law transferred responsibility for public welfare institutions, such as orphanages, shelters, and charitable organizations, to municipal governments and mandated the creation of local welfare boards to oversee the distribution of aid to vulnerable populations. Because the reform applied discontinuously across a population and time cutoff, we can employ a DDRD to estimate its impact on surveillance.

We find a null effect of the Crispi Law on surveillance in Table G.2. The analysis shows that a considerable expansion in state agents had no measurable effect on surveillance, an indirect test that reinforces the historical accounts, suggesting that an increase in the number of bureaucrats alone did not necessarily lead to heightened state surveillance.¹³

¹³In addition, historical records indicate that informants were generally individuals in dire personal situations and passing informations to attenuate personal difficulties, such as

School records: A related concern is that the reform might have increased the volume of school records, making exposed cohorts easier to identify. However, first, the 4,000 discontinuity did not increase enrollment, which was compulsory on both sides of the threshold, it merely lengthened primary schooling by two years. There is, therefore, no discrete jump in the creation of pupil records at the cutoff. Second, our outcome measures surveillance at any age, often decades after pupils left elementary school. For this force to determine our findings, what a child did in third rather than second grade would need to better inform political policing when the same person is 40, and even if this person migrated away from the hometown where this initial record was kept. Information collected when children were 8 vs 7 years old is unlikely to cause the jump in surveillance we observe.

Teachers as instigators: Another possibility is that the state perceived teachers as instigators of radical ideologies. If this were the case, we would observe that the increase in surveillance is driven by a surge in surveillance of teachers, who become a larger occupational group in treated areas. First, teachers constituted a negligible fraction of those surveilled pre-reform (0.00005 per 100,000 inhabitants), making it improbable that their surveillance alone drives the treatment effect. Second, although the number of teachers in the population likely increases post-reform, the absolute number of teachers who were surveilled did not increase (Table G.3, column 1), implying that the surveillance rate of teachers must have declined, or at most remained constant. Finally, the share of teachers among those surveilled increases (column 2), but by an amount that is similar to or smaller than that observed for other professions with comparable baseline surveillance rates (columns 3–5, Table G.3), despite the post-reform growth in the share of teachers in the population. These facts point to the same conclusion: teachers were not a group the police singled out after the reform.

prisoners seeking reduced sentences by reporting on their jail mates (Ciconte, 2011). Instead, teachers held financially stable and high-status positions, far from the situations of captivity that motivated the typical informants.

Job prospects: An alternative story for the effects on the working class can be found in the downstream consequences of surveillance itself. Surveillance may have reduced individuals’ job prospects, thereby pushing them to select into working class occupations. Table G.4 examines the main results by age at the start of surveillance and shows that the effect is present even among individuals who began being surveilled as adults or late in life, after their careers had been defined. This suggests that the impact of the reform on surveillance of working classes cannot be explained by early-career sorting into this social class.

Substitution with the Church: One effect of the education reform is that schools began to take on the church’s role in education, as before the reform, churches were largely responsible for educating children. However, since our analysis centers on rebellion against the state, and the reform positioned the state as the provider of education, it is unlikely that state-led schooling would inspire greater rebellion against itself. In fact, research by Paglayan (2024) shows that states typically design education reforms to cultivate loyalty and compliance among citizens. Therefore, if a substitution from church to state oversight occurred, as seems likely, we would expect the opposite effect: with increased exposure to state-run education, surveillance should have decreased, not increased.

9 Conclusions

This study focuses on the strategic considerations states make when deciding whom to subject to political surveillance. Our findings reveal that states are more likely to target individuals who display markers of potential political activism – educated individuals –, and particularly those with the potential to press for radical change – in this setting, the working class. This logic is observable across democratic and authoritarian regimes, and independently on levels of state capacity. We find limited evidence that an increase in political mobilization can explain these effects, consistent with surveillance spurring from state heightened concern rather than from an effectively higher mobilization. Our findings support a view in which

states perceive educated but subaltern individuals as particularly threatening to regime stability, regardless of whether these individuals engage in overt political activism.

To what extent do historical findings on surveillance remain relevant in the age of digital monitoring? Specifically, do states still face resource constraints that force them to choose whom to surveil, or have digital tools eliminated the need for selectivity, enabling widespread and indiscriminate targeting? On one hand, social media and biometric recognition technologies have made it easier for states to gather information on individuals. On the other hand, technological advancements also provide activists with new ways to evade scrutiny, such as encrypted communication and an abundance of alternative channels for correspondence. Moreover, although digital tools have expanded surveillance capabilities, analyzing vast amounts of data remains costly, as it requires transforming unstructured information into actionable intelligence, for which both data analysts and intelligence personnel are needed. In authoritarian contexts, where dissidents are acutely aware of surveillance risks, large datasets may yield less useful intelligence, as activists take steps to avoid leaving digital traces of their actions. As a result, states may still need to complement digital surveillance with traditional intelligence-gathering methods. Considering these factors, we argue that it remains useful to think of states as resource-constrained actors who must make strategic choices about whom to invest on, indicating that our findings based on historical data – the only type of intelligence data on surveilled likely to be shared by states – might still yield insights to understand our time of digital surveillance.

While this study sheds light on the logic of state surveillance, most facets of this important tool of state control remain understudied. For example, resources allocation might vary by target: large numbers of newly educated workers may be monitored lightly to gauge their intentions, whereas a handful of elites might be followed intensively to detect concrete plans. Future research could also chart how covert and overt surveillance evolve over a movement's life cycle, how different agencies share or weaponize intelligence, and how these strategies generalize across contexts where subalternity is defined along dimensions other than class.

This research offers a new perspective on the enduring challenges faced by states in balancing security with democratic inclusivity. As states around the world continue to adopt increasingly sophisticated surveillance technologies, our findings emphasize the need to ensure that these tools do not disproportionately target and disenfranchise already marginalized populations.

References

- Acemoglu, Daron and James A Robinson. 2001. “A theory of political transitions.” *American Economic Review* 91(4):938–963.
- Akcigit, Ufuk, Salomé Baslandze and Stefanie Stantcheva. 2016. “Taxation and the international mobility of inventors.” *American economic review* 106(10):2930–2981.
- Almond, Gabriel Abraham, Sidney Verba et al. 1963. “The civic culture: Political attitudes and democracy in five nations.”
- Andirin, Veli, Yusuf Neggers, Mehdi Shadmehr and Jesse M Shapiro. 2022. Real-time Surveillance of Repression: Theory and Implementation. Technical report National Bureau of Economic Research.
- Beraja, Martin, Andrew Kao, David Y Yang and Noam Yuchtman. 2023. “AI-tocracy.” *The Quarterly Journal of Economics* 138(3):1349–1402.
- Besley, Timothy and Torsten Persson. 2009. “Repression or civil war?” *American Economic Review* 99(2):292–297.
- Bourguignon, François and Thierry Verdier. 2000. “Oligarchy, democracy, inequality and growth.” *Journal of Development Economics* 62(2):285–313.
- Bueno de Mesquita, Bruce and Alastair Smith. 2017. “Political succession: A model of coups, revolution, purges, and everyday politics.” *Journal of Conflict Resolution* 61(4):707–743.

- Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2015. “Optimal data-driven regression discontinuity plots.” *Journal of the American Statistical Association* 110(512):1753–1769.
- Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman and Y Jane Zhang. 2017. “Curriculum and ideology.” *Journal of political economy* 125(2):338–392.
- Carolini, Simonetta and Fabio Ecce. 2015. *Antifascisti alla sbarra*. Palombi.
URL: <https://anppia.it/gli-antifascisti/>
- Caselli, G. 1990. “Mortalità e sopravvivenza in Italia dall’Unità agli anni’30.” *Popolazione, Società e ambiente. Temi di Demografia storica italiana (secc XVII-XIX)* pp. 275–310.
- Cattaneo, Matias D., Michael Jansson and Xinwei Ma. 2018. “Manipulation Testing Based on Density Discontinuity.” *The Stata Journal* 18(1):234–261.
- Ciancarini, Enrico. 2019. “Appunti di vite sovversive.” *Società Storica Civitavecchiese* .
- Ciconte, Enzo. 2011. *Banditi e briganti: una storia senza fine, dal Cinquecento all’Ottocento*. Rubettino.
- Ciconte, Enzo. 2019. *Storia criminale: la resistibile ascesa di mafia, ’ndrangheta e camorra dall’Ottocento ai giorni nostri*. Rubbettino Editore.
- Davenport, Christian. 2005. “Understanding covert repressive action: The case of the US government against the Republic of New Africa.” *Journal of Conflict Resolution* 49(1):120–140.
- De Mesquita, Ethan Bueno. 2010. “Regime change and revolutionary entrepreneurs.” *American Political Science Review* 104(3):446–466.
- Di Fiore, Laura. 2018. *Gli Invisibili: Polizia politica e agenti segreti nell’Ottocento borbonico*. Vol. 21 FedOA-Federico II University Press.

- Dilemmi, Andrea. 2010. *Polizia e sorveglianza del dissenso politico (Verona, 1894-1963)* PhD thesis Università di Verona. Doctoral thesis in Scienze storiche e antropologiche.
- Enikolopov, Ruben, Maria Petrova and Ekaterina Zhuravskaya. 2011. “Media and political persuasion: Evidence from Russia.” *American economic review* 101(7):3253–3285.
- Fanon, Frantz. 1963. “The Wretched of the Earth.” *Grove Weidenfeld* .
- Glaeser, Edward L, Giacomo AM Ponzetto and Andrei Shleifer. 2007. “Why does democracy need education?” *Journal of economic growth* 12(2):77–99.
- Gramsci, Antonio. 1975. *Quaderni del carcere*. Torino: Einaudi. Critical edition edited by Valentino Gerratana.
- Grembi, Veronica, Tommaso Nannicini and Ugo Troiano. 2016. “Do fiscal rules matter?” *American Economic Journal: Applied Economics* pp. 1–30.
- Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.
- Hager, Anselm and Krzysztof Krakowski. 2022. “Does state repression spark protests? evidence from secret police surveillance in communist poland.” *American Political Science Review* 116(2):564–579.
- Hassan, Mai, Daniel Mattingly and Elizabeth R Nugent. 2022. “Political control.” *Annual Review of Political Science* 25(1):155–174.
- King, Gary, Jennifer Pan and Margaret E Roberts. 2013. “How censorship in China allows government criticism but silences collective expression.” *American political science Review* 107(2):326–343.
- Kuran, Timur. 1989. “Sparks and prairie fires: A theory of unanticipated political revolution.” *Public choice* 61(1):41–74.

- Labini, Paolo Sylos. 1972. “Sviluppo economico e classi sociali in Italia.” *Astrolabio*, n.3 pp. 18–31.
- Lecce, Giampaolo, Laura Ogliari and Tommaso Orlando. 2022. “State formation, social unrest and cultural distance: Brigandage in post-unification Italy.” *Journal of Economic Growth* 27(3):453–483.
- Leto, Guido. 1951. *OVRA: Fascismo-Antifascismo*. Bologna: Cappelli.
- Lichter, Andreas, Max Löffler and Sebastian Sieglöcher. 2021. “The long-term costs of government surveillance: Insights from Stasi spying in East Germany.” *Journal of the European Economic Association* 19(2):741–789.
- Marshall, John. 2016. “Education and voting Conservative: Evidence from a major schooling reform in Great Britain.” *The Journal of Politics* 78(2):382–395.
- Marx, Karl and Friedrich Engels. 1848. *Manifesto of the Communist Party*. Workers’ Educational Association. Reprinted in various editions.
URL: <https://www.marxists.org/archive/marx/works/1848/communist-manifesto/>
- Paglayan, Agustina. 2024. *Raised to Obey: The Rise and Spread of Mass Education*. Vol. 134 Princeton University Press.
- Qin, Bei, David Strömberg and Yanhui Wu. 2017. “Why does China allow freer social media? Protests versus surveillance and propaganda.” *Journal of Economic Perspectives* 31(1):117–140.
- Scott, James C. 1985. *Weapons of the Weak: Everyday Forms of Peasant Resistance*. Yale University Press.
- Sondheimer, Rachel Milstein and Donald P Green. 2010. “Using experiments to estimate the effects of education on voter turnout.” *American Journal of Political Science* 54(1):174–189.

- Sotgiu, Girolamo. 1984. *Storia della Sardegna sabauda 1720-1847*. Roma-Bari: Laterza.
- Tosatti, Giovanna. 1990. “Le origini del Casellario politico centrale.” *Istituto per la scienza dell’amministrazione pubblica, Archivio Nuova serie 6 “Le riforme crispine”, vol. I “Amministrazione statale”*.
- Verba, Sidney, Kay Lehman Schlozman, Henry Brady and Norman H Nie. 1993. “Citizen activity: Who participates? What do they say?” *American Political Science Review* 87(2):303–318.
- Wen, Jaya Y. 2025. “State employment as a strategy of autocratic control in China.” *Review of Economics and Statistics* pp. 1–45.
- Xu, Xu. 2021. “To repress or to co-opt? Authoritarian control in the age of digital surveillance.” *American Journal of Political Science* 65(2):309–325.
- Yongyi, Song et al. 2011. “Chronology of mass killings during the Chinese Cultural Revolution (1966-1976).” *Online encyclopedia of mass violence* pp. 315–327.
- Zhuravskaya, Ekaterina, Maria Petrova and Ruben Enikolopov. 2020. “Political effects of the internet and social media.” *Annual review of economics* 12(1):415–438.