UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Essays on Leader, Local Governance, and National Identity

Permalink
https://escholarship.org/uc/item/9c26b9bg

Author
Chen, Wei-Lin

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Leader, Local Governance, and National Identity

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Wei-Lin Chen

Committee in charge:

        Professor Julianne Berry Cullen, Co-Chair
        Professor Paul Niehaus, Co-Chair
        Professor Samuel Bazzi
        Professor Jeffrey Clemens
        Professor Gordon Boyack Dahl
        Professor Craig McIntosh

2023

The Dissertation of Wei-Lin Chen is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

To my parents Kan-Chu Wang and Chun-Chien Chen

TABLE OF CONTENTS

LIST OF FIGURES

ix

LIST OF TABLES

ACKNOWLEDGEMENTS

VITA

2013   Bachelor of Arts, National Taiwan University

2017   Master of Arts, National Taiwan University

2023   Doctor of Philosophy, University of California San Diego

ABSTRACT OF THE DISSERTATION

Essays on Leader, Local Governance, and National Identity

by

Wei-Lin Chen

Doctor of Philosophy in Economics

University of California San Diego, 2023

Professor Julianne Berry Cullen, Co-Chair
Professor Paul Niehaus, Co-Chair

This dissertation consists of three chapters. In Chapter 1, I study the impact of partisan leaders on traffic stop policing behaviors in North Carolina. Using a difference-in-differences design that exploits sheriff turnovers, I find that offices with a Democrat-to-Republican sheriff turnover rather than a Democrat-to-Democrat sheriff transition have an increase of black drivers' share in traffic stops by 3.2 percentage points, a 13.5% increase compared to baseline. Decomposing the changes in black driver's share along two dimensions: stop purposes and officers, I find that the increase is driven by changes within safety stops instead of investigation stops, and driven by changes in incumbent officers' tendency to stop black drivers. The increase in racial

disparities is not accompanied by an increase in unconditional hit rates.

In Chapter 2, in joint work with Samuel Krumholz, we study the impact of partisan leaders on the political composition of law enforcement agencies in the United States using elected sheriffs in North Carolina as a case study. Using a difference-in-differences design, we find that offices shifting from a Democrat to Republican sheriff experience a 9 percentage point (27%) decrease in the Democratic share of Sheriff deputies relative to counties experiencing a Democrat-to-Democrat turnover. This change is driven both by existing Democratic deputies disproportionately changing their party registration and an increase in the share of Republican deputies among the entrants. Overall, we provide new evidence that leaders can shape the personnel's political composition not just by hiring but by inducing party switches.

In Chapter 3, in joint work with Ming-Jen Lin and Tzu-Ting Yang, we examine the causal effects of textbook content on individuals' national identity, by exploiting a curriculum reform that introduced a new perspective on Taiwan's history for students entering junior high school after September 1997. Using a repeated nationally representative survey and a regression discontinuity design, we show that students exposed to the new textbooks were more likely to hold exclusive Taiwanese identity rather than dual identity (i.e., Taiwanese and Chinese). The effect was greater for academic track students and those living in neighborhoods where fewer people identify as Taiwanese. We find little impact of the new curriculum on people's political preferences related to Taiwan's independence. Finally, we find that the probability of reporting as Taiwanese among old textbook readers converges with that of people reading new textbooks in the long run since the perspectives of old textbooks are in conflict with the recent social trends.

Chapter 1

The Impact of Partisan Politics on Policing Practices: Evidence from North Carolina's Sheriff's Offices

1.1 Introduction

The criminal justice system in the United States is deeply related to and influenced by partisan politics due to the political process of personnel selection. Although leaders of local law-enforcement agencies are often elected, the impact of political preferences of leaders on frontline policing is not well-understood. This paper studies the impact of the political party affiliation of leaders on one of the most frequent interactions Americans have with law-enforcement officers: traffic stops.

I examine the impact of partisan leadership on racial disparities in traffic stops. Racial disparities in traffic stops are well-documented. Black drivers are more likely to be stopped than White drivers, especially before sunset; during the stop process, Black drivers are twice as likely to be searched than White drivers (Pierson et al., 2020). A vast literature studies to what extent the racial disparities come from racial bias and has established evidence of racial discrimination at the officer level (Antonovics and Knight, 2009; Goncalves and Mello, 2021). I start from a different point in the hierarchy of law-enforcement agencies and ask if leaders matter in determining racial disparities of frontline traffic stops.

This paper focuses on Sheriff's Offices in North Carolina. I focus on sheriff's offices instead of police departments since sheriffs are elected through partisan elections. I can thus directly identify sheriff's party affiliations. By exploiting party turnovers of sheriffs induced by elections, I identify the impact of the party affiliation of sheriffs on offices' traffic stop practices. One central challenge in estimating the relationship between party affiliation of local law-enforcement leaders and traffic stop practices is that localities with leaders from different parties may have unobserved differences. Such differences may make officers adopt different traffic stop strategies. In addition, time trends that affect local law-enforcement practices, such as crime rate changes and gentrification development, may evolve differently across such localities.

I adopt a difference-in-differences research design to overcome these challenges. The control group is counties that experience Democrat-to-Democrat (henceforth D-to-D) sheriff transition that does not necessarily involve a leader turnover; the treatment group is counties that experience Democrat-to-Republican (henceforth D-to-R) sheriff turnover. I analyze turnovers from the 2010, 2014, and 2018 elections. For each election, we examine traffic stops in an election cycle defined as from 3 years before the election to 1 year after the election. This definition of election cycle allows us to stack up data from 3 election cycles without having overlapping timing periods.

I find that Republican sheriffs' leadership alters the racial composition of stopped drivers. Republican sheriffs increase the share of Black drivers by 3.2 percentage points, a 13.5% increase compared to the baseline period (two years before the election). To investigate which new policies and instruments the Republican sheriffs adopt that result in an increase in racial disparities, I decompose the changes in the Black driver's share along two dimensions: the initial purpose of stops and the type of patrolling officers.

Law enforcement officers have two goals in conducting traffic stops—maintaining road safety and finding contraband. The two goals motivate the distinction of two types of stops: stops due to moving violations (safety stops) and non-moving violations (investigation stops). How much focus should a law-enforcement agency put on each type of stop is under debate in North

Carolina. The Fayetteville Police Department Chief, in 2013, proposed to minimize the number of stops due to non-moving violations to avoid unnecessary traffic stops. The Mecklenburg County Sheriff proposed a similar policy in 2022 because he was presented with information that Black drivers are disproportionately affected by investigation traffic stops. [1] Fliss et al. (2020) used a synthetical control method and found that the policy in Fayetteville leads to a reduction of traffic crashes and injuries and a decrease in the share of Black drivers in traffic stops.

To see if Republican sheriffs' focus on the two types of stops systematically differs from Democrat sheriffs and thus contributes to the changes in Black drivers' share, I examine the impact of partisan leadership on the share of safety stops. I find that Republican sheriffs decrease the share of safety stops by 5.4 percentage points. Such changes can have racially disparate impacts because, in the counties we analyze, Black drivers account for a lower proportion of safety stops than in investigation stops. However, I find that the change in the share of safety stops can only account for 16.5% of the increase in the Black driver's share. The compositional changes of the types of stops are not the major contributor. Instead, the Black drivers' share *within* each type plays a more critical role. In particular, the change in Black drivers' share *within* safety stops accounts for 68% of the overall change in Black drivers' share. In understanding racial disparities in stops, researchers have focused on investigation stops, in which officers are thought to have more discretion and hence more likely to exhibit racial bias. My results suggest that policies in conducting traffic safety stops may require more attention.

I consider two channels regarding personnel policies that may result in a change in traffic stop practice: (1) reshuffling of officers based on officers' policy preferences regarding traffic stops; (2) change in incumbent officers' stop practices in response to the new leadership. I find evidence supporting the second channel. Regarding personnel reshuffling, I find that D-to-R transitions are associated with more reshuffling of officers. The share of stops conducted by

---

[1] See for a coverage about Fayetteville police department at https://www.usatoday.com/story/news/nation/2021/04/15/police-reform-fayetteville-burlington-nc-traffic-stops-policing/7225318002/ and see https://www.foxnews.com/us/north-carolina-sheriffs-office-stops-pulling-drivers-non-moving-traffic-violations for a coverage about Mecklenburg county sheriff's office.

incumbent officers in D-to-R counties is 19 percentage points (a 34% decrease compared to baseline) lower than in D-to-D counties post-elections. However, the reshuffling does not lead to a change in overall stop practices. The officers who were shuffled *in* are similar to those shuffled *out* regarding the share of Black drivers among their stops.

Do officers alter their traffic practices in response to the new leadership? I find that the incumbent officers, who continued to conduct traffic stops in post-election years in D-to-R counties, increased the Black driver's share in their stops by four percentage points compared to the incumbent officers in D-to-D counties, a 17.5% increase compared to the baseline. Further, I find that the increase in the Black driver' share among incumbent officers is not driven by a few officers but by many officers having small to medium-level changes in the tendency to stop Black drivers. I thus provide a case where the reshuffling of officers does not lead to systematic changes in the observed policy practices, but the leaders reshape policy practices in a way that changes incumbent officers' conduct.

I next analyze an important decision officers make after stopping a driver: whether to search a vehicle. I examine the impact of sheriff's party affiliation on the overall search rates and search rates within racial groups. Note that, with the new sheriff's traffic stop practices, relevant characteristics of the stopped driver composition (regarding suspiciousness of holding contraband, for example) likely change in the post-election year. I thus interpret the impact on search rates (if any) as coming from a combination of changes on whom to stop *and* whom to search. I find no significant impacts of sheriff's party affiliation on the overall and within racial group search rates.

Understanding whether a trade-off between racial disparities in traffic stops and *efficiency* exists is a central focus in the literature (Feigenberg and Miller, 2022). Since finding contraband is at least one of the goals in conducting traffic stops, an reasonable efficiency measure is the unconditional hit rate, defined as the number of searches with found contraband divided by the number of total stops. I find that the D-to-R transition is *not* associated with statistically significant changes in the overall unconditional hit rates.

At last, I examine the long-term impact. I find that the impact of sheriff's party affiliation on traffic stop disparities may be short-lived. I argue that such a short-lived impact may not be surprising given that sheriffs face temporal electoral incentives every four years. In addition, drivers may swiftly change their driving routine in response to the new traffic stop practices.

Overall, this paper contributes to our understanding of sources of racial disparities in the criminal justice system. Previous literature has found partisanship influences sentencing: compared to Democratic-appointed judges, Republican-appointed judges give longer sentences to Black offenders than non-Black offenders with similar crimes (Cohen and Yang, 2019). I provide evidence that the political preferences of leaders matter in determining racial disparities in frontline policing, where literature has identified the importance of the racial composition of voters the leaders face (Facchini et al., 2020), the race of the leaders (Bulman, 2019), and the racial composition of the police force (McCrary, 2007). Very recent literature identified the heterogeneity of racial bias at the officer level (Goncalves and Mello, 2021) and suggested that officers with different levels of bias have varied traffic stop behaviors responding to Trump rallies during his 2015–2016 campaign (Grosjean et al., 2022).

The impact of partisanship on law enforcement is not without ambiguity *ex ante*. Although survey evidence shows that party affiliation of the general public is correlated with attitudes toward policing policies such as body cams and police force size (Hansen and Navarro, 2021), the political preferences of the law-enforcement leaders across parties may not be so dissimilar. Thompson (2020) finds no effect of the party affiliation of sheriffs on compliance with federal requests to detain unauthorized immigrants and suggests that the similar compliance rate may be due to sheriffs sharing similar immigration enforcement views across parties.

I also contribute to the literature that emphasizes the importance of political turnover in personnel in public organizations. Political turnover is often associated with personnel changes on account of patronage. Colonnelli et al. (2020) finds that supporters of the party in power in Brazil are more likely to be hired and are negatively selected on their competence. Akhtari et al. (2022) finds that local mayor election turnovers in Brazil are linked to new personnel turnovers in

schools and are further accompanied by lower student test scores. I provide a case where leaders' political party turnovers are associated with a new assignment of duties (assigned to traffic stop teams or not), but the new assignment seems not to be based on specific policy preferences.

The rest of the paper is as follows. I describe relevant contexts in section 1.2 and introduce the data in section 1.3. I then lay out the empirical methods in section 1.4. Results are discussed in section 1.5. I conclude in section 1.6.

1.2 Background

1.2.1 Law-Enforcement Agencies in North Carolina

Sheriff's offices are the top law enforcement agencies in counties. They perform duties in unincorporated areas within counties. Police departments in municipal governments are in charge of law enforcement in incorporated areas. The main functionality of sheriff's offices includes management of jails and detention centers, crime investigation, immigrants detention, patrol, and document application such as gun permits. In this paper, I focus on the traffic stop and search. Patrol officers account for a fifth of the personnel in sheriff's offices in North Carolina, while jailers and detectives/investigators account for respectively 36% and 10% of the personnel. Police departments do not manage jails, so they assign more personnel to patrol and investigation, 46% for patrol and 14% for investigation.[2] Police officers conduct much more stops than deputy sheriffs. During 2008-2019 (my sample period), on average, deputy sheriffs conducted about a hundred thousand stops a year, while police officers conducted about six hundred seventy thousand stops.

Each of the one hundred counties in North Carolina has one Sheriff's Office. Voters directly elect all sheriffs in North Carolina. The elections are partisan; they occur every four years in November, and there are no term limits. The newly elected sheriffs are sworn in on

---

[2]The personnel numbers are from 2016 Law Enforcement Management and Administrative Statistics (LEMAS) Survey. 22 out of 100 sheriff's offices, 72 out of 189 police departments in North Carolina are in the sample. The included agencies are larger agencies. The median personnel size is 51. The percentage of personnel in each category is the weighted average of the shares, with personnel size in each agency as the weights.

November 30, and the deputies would also take their oath on the same day. All of the elected sheriffs since 1998 are affiliated with either the Democratic Party or the Republican Party. We use sheriffs' turnovers induced by elections as the main variation of change of control. In particular, we focus on sheriff's turnovers that involve party turnovers. Police chiefs, who are the leaders of the police departments, on the other hand, are appointed by the municipal government.

1.2.2 Traffic Stop

Law-enforcement officers stop drivers for two main reasons. First, the driver exhibits reckless driving, such as speeding. Second, officers stop drivers for nonmoving violations. This includes equipment failures such as broken tail lights, vehicle regulation violations such as expired registration, and suspicion in relation to ongoing investigations. Following Baumgartner et al. (2018), I call the first type a traffic safety stop and the second type an investigatory stop. In practice, officers use vehicle regulation violations as a pretext to stop drivers in pursuit of potential criminal investigations or searches for drug possession.

By law, officers can search a vehicle as long as the officers have probable cause to believe that a law has been broken. This is a decision in that officers have much discretionary power. Regardless of whether a search is conducted, a traffic stop leads to four actions: no action, warning, citation, and arrest. During searches, an officer might find contraband, including drugs, alcohol, or weapons.

1.3 Data

I use traffic stop and search data and sheriff elections record to analyze the effect of sheriffs' party affiliation on officers' traffic stop and search behaviors.

1.3.1 Sheriff Election Records.

Sheriff's election results since 2010 are publicly available on the North Carolina State Board of Elections website. We hand-collected the 2006 election data through news reports and

county board of election websites. Party affiliation and the names of the elected sheriffs are used to determine if a county went through sheriff turnovers and party turnovers. Vote shares of the winners are used to assess the competitiveness of the elections.

Table 1.1 reports the sheriff election results from 2010 to 2018. Only four elections involve Republican-to-Democrat turnover. I do not compare counties with elections of Republican-to-Democrat with counties with elections of Republican-to-Republican turnover in this paper due to power concerns. I define the control group as the county-election cycles that experience Democrat-to-Democrat-type elections. The treatment group includes county-election cycles that experience Democrat-to-Republican elections.

Panel D of Table 1.1 shows the winners' vote share distribution. All Democrat-to-Republican elections have winners' vote shares of less than 80%. To match on the winners' vote shares, I confine my sample to the county-cycles where the winners' vote shares are less than 80%. Panel B shows the number of county-cycles in each election type after I apply this restriction.

The last sample restriction is about the number of stops each year within an election cycle. I exclude the county-cycles where a sheriff's office conduct less than 50 stops in at least an election year within the election cycle. Two reasons for this criterion. First, the decomposition analysis in section 1.5.2 and 1.5.3 would not make sense if the number of stops within certain types (safety and investigation stops) and by certain officers (incumbent officers who conduct traffic stops before and after elections and others) is tiny. Second, I aim to have consistent "report" quality across the years. Some counties excluded by this restriction have considerable fluctuations in the number of stops across the years. e.g., New Hanover had four stops in 2009 and 890 stops in 2010. Some counties have zero stops in a year and hundreds of stops in adjacent years. These patterns cast doubt on whether the reported traffic stops reflect a representative sample of all stops in counties where the number of stops fluctuates dramatically. I chose the number 50 based on my judgment of trading off losing too many counties and including bad-quality reports. The resulting number of county-cycles of each election type are presented in

Panel C of table 1.1.

1.3.2 Traffic Stop and Search Records.

The traffic stop and search records are available upon request in North Carolina. The data set contains the driver's race, ethnicity, gender, and age. I construct the share of Black drivers among all traffic stops using the driver's race. Officers have to report the purpose of each stop. Each stop is associated with one of the twelve stop purposes. Following Baumgartner et al. (2018), I exclude the sample associated with the checkpoint because such stops are recorded only when searches are conducted. I classify stops into two types: safety and investigation. Safety stops include ones associated with speed limit violations, stop light/sign violations, driving while impaired, and safe movement violation. Investigation stops include ones associated with vehicle equipment violations, vehicle regulatory violations, seat belt violations, investigation, and other motor vehicle violations. With the categorization, I construct the share of safety stops among all stops and the share of Black drivers within the two types of stops.

Unique officer IDs are included in the data.[3] The IDs are not linked to other information about officers, such as names, races, or ages. I use the IDs to identify two groups of officers: stayers and non-stayers. Stayers are the officers who conduct traffic stops both before and after elections. Non-stayers are the officers who conduct traffic stops only before or after the elections. With this categorization, I construct the variables the share of stops done by stayers and the share of Black drivers in the stops done by stayers or non-stayers.

The data set includes the time and the name of the location of each stop. The name of the location can be a county, a city/town, a census-designated place (CDP), or some location names used by locals. Around 60% of the stops only record the location at the county level. This significantly restricts our analysis of officers' patrolling location decisions.

The dataset also has information about searches and contraband. I use information about whether a search is conducted in a traffic stop and whether any contraband is found during a

---

[3]The officer ID is only unique within the law enforcement agency. We cannot track officers across agencies.

9

search to construct two outcome variables. The search rate is defined as the number of searches divided by the number of stops. The unconditional hit rate is defined as the number of searches with found contraband divided by the number of stops.

Summary Statistics of Traffic Stops and Searches.

Table 1.2 presents the summary statistics of traffic stops and searches in the county-cycles I include in our analysis (Panel C in Table 1.1). I report descriptive shares on race, gender, and traffic stop types. The driver is female in 35% of the stops, black in 25% of the stops, Hispanic in 7% of the stops, and white in 65% of stops. Due to the small share of Hispanic drivers, in the regression analysis, I divide the drivers into Black and non-Black groups. [4] Officers search drivers in 6.7% of stops and find contraband in 2.2% of stops. Black drivers, once stopped, are more likely to be searched than White drivers (7.9% compared to 6.1%). The difference in the search rates between Black and White drivers is much smaller than the one seen in Feigenberg and Miller (2022).

Dividing stops into safety and investigation types, the driver is 28% Black in investigation stops and 24% in safety stops. Officers are more likely to search in investigation stops than in safety stops (8.5% and 5.1%, respectively). The conditional hit rates (number of searches with found contraband divided by the total number of searches) are similar across two types of stops, around 31%.

1.4 Empirical Methods

I aim to identify the causal effect of sheriff's party affiliation on traffic stop practices. To this goal, I adopt a difference-in-differences design, comparing counties that experience elections resulting in Democrat-to-Democrat transitions with counties that experienced Democrat-to-Republican transitions. In the main analysis, I define an election cycle from three years before an election to one year after. This definition allows no overlapping calendar years across election

---

[4]Other races, including Asians, Native Americans, and Other/Unknown, account for around 2% of stops and are included in the non-black group.

cycles but limits the time horizon of the analysis.

To estimate the causal effect of sheriff's party affiliation on traffic stop practices, I estimate an ordinary least square regression with a difference-in-differences type specification:

$$Y_{cle} = \sum_{e=-2}^{1} \beta_e D_{cl}^{D-to-R} \cdot \eta_e + \delta_{le} + \delta_{cl} + \varepsilon_{cle} \qquad (1.1)$$

where Y_{cle} is a variable at county-year level for county c in year e in cycle l. Treatment group status in each election cycle is denoted by D_{cl}^{D-to-R}, δ_{cl} is county-cycle fixed effects. I separate data into three election cycles, denoted as l. I use election results from 2010, 2014, and 2018. Hence l can take three values, 2010, 2014, and 2018. In tables and figures, the time convention is as follows: I denote the year when the election happened as t and other years as $t-2, t-1, t+1$. In regression specifications, the time convention chronologically in an election cycle is denoted as $e = -1, -1, 0, 1$. Since the new sheriff is sworn in on November 30, I define a year starting from December to November. For example, the year t ($e = 0$) in the 2010 election cycle involved observations from December 2009 to November 2010. Hence, δ_{le} uniquely defines the timing of each stop in year e in cycle l. I use the year before the election as the omitted base year. I analyze at the county level instead of the stop level because I am interested in the causal effect of leadership on law enforcement agencies. In the results section, I also report the estimation results weighting the county-year observations with the number of stops of each county at the beginning of the election cycle ($t-3$).

The coefficients of interest are β_e, which captures the differences between control and treatment groups across years within a cycle. All standard errors are clustered at the county level throughout the paper unless stated otherwise.

I examine changes in racial disparities by looking at the share of Black drivers among stops, and the search rates within racial groups. When I analyze the efficiency of traffic stop practices, I look at the unconditional hit rates in the overall stops, and within racial groups. For these outcome variables, specification 1.1 is appropriate. In section 1.5.1, I also compare the

11

change in the number of stops across racial groups. There, I estimate a triple difference-in-differences specification as follows:

$$Y_{cleg} = \sum_{e=-2}^{1} \gamma_e D_{cl}^{D-to-R} \cdot \eta_e \cdot G_g + \sum_{e=-2}^{1} \beta_e D_{cl}^{D-to-R} \cdot \eta_e \tag{1.2}$$
$$+ D_{cl}^{D-to-R} \cdot G_g + G_g + \eta_e \cdot G_g + \delta_{le} + \delta_{cl} + \varepsilon_{cle},$$

where G denotes groups: Black and non-Black. Other notations are defined as in equation 1.1. The coefficients of interest are γ_e and β_e.

1.5 Results

I first present results on racial disparities in the share of Black drivers in traffic stops in section 1.5.1. I then present decomposition analysis along two dimensions of traffic stops: types of traffic stops (safety and investigation) in section 1.5.2 and identity of officers (stayers or non-stayers) in section 1.5.3. I then analyze the second stage during a traffic stop process: the searches in section 1.5.5. After examining racial disparities, I investigate in section 1.5.6 if a change in the efficiency of traffic stops, measured by the unconditional hit rates, is accompanied by changes in racial disparities. Finally, section 1.5.7 discuss the impact of partisan leadership in the longer-term and provide cautions on the interpretation of the changes in racial disparities presented in section 1.5.1.

1.5.1 Black driver's share

Graphical Evidence.

I plot the raw data in Figure **??** to show the data variation captured by the difference-in-differences specification. I compute the Black drivers' share among all stops at the county-year level. I then take the simple averages across counties and election cycles to aggregate the data into D-to-D, D-to-R, and R-to-R groups. D-to-D counties have higher Black driver's shares than D-to-R and R-to-R counties since D-to-D counties are generally more urban areas. Before the

election, the gap in the Black driver's share between the three groups stays roughly constant across the years within an election cycle. One year after the election, however, the Black driver's share in D-to-R counties increased while the shares in D-to-D and R-to-R counties barely changed.

Regression estimation results.

In Figure 1.2, I plot the estimates of interaction terms between the treatment group indicator and election cycle-years from equation 1.1 (β_e) with the Black driver's share as the outcome variable. I report the estimates in Table 1.3. Before the elections ($t-2$ and t), the interaction term estimates are small and non-significant, giving me confidence that the parallel pre-trend assumption, required by the difference-in-differences research design, is satisfied in this setting. Right after the election, the Black driver's share increased by 3.2 percentage points in D-to-R counties one year after the election compared to D-to-D counties. (Table 1.3,Column (1)). Given that the dependent variable mean in D-to-R counties in the year before the elections is 0.24, this amounts to a 13.5% increase in the Black driver's share.

From Columns (2)-(4), I probe the robustness of the impact of sheriff's party affiliation on Black driver's shares by weighting the sample, restricting the sample to close elections, and examining a placebo scenario. In Column (2), I report the regression results with a sample weighted by the number of stops of the county two years before elections ($t-2$). The weight of a county within a cycle is thus fixed. The estimates would be similar to the ones in Column (1) if there is not much causal effect heterogeneity along the number of stops dimension. The standard errors may be smaller when I weight the samples by the number of stops if the number of stops varies tremendously with some county-cycles having very small number of stops and the error term variation mostly coming from within county-cycle. See Solon et al. (2015) for simple examples comparing regression results with and without weights. I find that the magnitude of the estimate from the weighted regression is similar to the unweighted one, suggesting that the effect of sheriff's party affiliation does not vary with the number of traffic stops. The s.e. becomes slightly larger.

In Column (3), I follow the spirit of regression discontinuity designs with close elections and restrict the sample to counties with winners' vote share below 60%. The magnitude of the estimate is similar to Column (1), but the standard errors become much larger, resulting in the statistical insignificance of the estimate.

In Column (4), I look into a placebo scenario, the traffic stops done by the police departments in the D-to-D and D-to-R counties. Although deputy sheriffs and police officers may focus on different neighborhoods in patrolling, the placebo scenario should still capture changes in the driver's population (if any) to some extent. I find that the magnitude of the estimate of the interaction term between $t+1$ and D-to-R dummy variables is much smaller for stops done by police officers. The similar magnitudes of the estimates of the post-election interaction term in Column (1)-(3) and the much smaller magnitudes in Column (4) suggest that the increase of the Black driver's share in D-to-R counties after the elections are driven by the change of traffic stop practices associated with the newly elected Republican sheriffs, instead of changes in the driver's population in specific counties.

Changes in Levels.

Table 1.3 focuses on the change in shares; I now turn to the changes in the levels to know if more Black drivers are stopped. Table 1.4 columns (1) and (2) report the regression estimates from the same estimating specification as in equation 1.1, with the natural log of the number of stops in the separate race groups. Although the magnitudes of the coefficients of the post-election and D-to-R interaction term is large in columns (1), we cannot reject the null of no change in the number of Black stops at the 10 percent significance level.

To compare the changes in the number of stops across racial groups, I report the estimates of γ_e and β_e in equation 1.2 in column (3). The number of stops associated with Black drivers marginally significantly increase more than the number of stops associated with non-Black drivers by 15%. Combining the estimates in Columns (1)-(3), I interpret the changes in the Black driver's shares observed in Table 1.3 driven by an increase in the number of stops of the Black drivers, instead of a decrease of the number of stops of the white drivers.

A notable pattern, the decrease of the number of stops in the election year t is shown in Table 1.4, Column (1)-(3). Since D-to-R elections are more competitive than D-to-D ones (Table 1.1, Panel B), I test whether the competitivenss of the D-to-R elections drive the lower number of stop. In Column (4), I test the hypothesis by comparing the number of stops between counties with close elections (winner's vote share below 60%) and others. The counties included in the estimation are the same as in Columns (1) and (2). I denote the counties with close elections as one with the Close dummy variable, zero otherwise. The magnitude of the estimate of the interaction term between the election year t dummy and the Close dummy variables is much smaller than the magnitude of the estimate of the interaction term between the election year t dummy and the D-to-R dummy variable seen in Columns (1) and (2), suggesting that the competitiveness of the sheriff elections does not drive the decrease in the number of stops.

In this section, I establish evidence that Republican sheriffs increase the number of traffic stops of Black drivers, increasing the Black driver's share. In subsequent sections, I examine whether the changes in the focus of specific types of traffic stops, the changes in personnel, and the changes in patrolling location and time can explain the observed increase in the Black driver's share.

1.5.2 Initial purpose of traffic stops

The first traffic stop policy dimension we examine is the initial purpose of traffic stops. Motivated by the policy proposals seen in the Fayetteville Police Department and the Mecklenburg County Sheriff's Office, and the literature which finds that officers enjoy more discretionary power in investigatory stops (Roach et al., 2022), we examine if the share of safety stops changes as the counties elected new Republican sheriffs. In Table 1.5, column (2), I display the estimation results of equation 1.1 with the outcome variable the share of safety stops among all stops. Compared to the year before the election, the share of safety stops decreases by 3.3 p.p. in the D-to-R counties (relative to D-to-D counties) in the election year, and decreases by 8.8 percentage points after the elections. Given that I do not observe a pre-trend in the share of stops

15

in $t-2$, the decrease in the election year might come from the Democrat incumbents' policy during campaign seasons in response to strong Republican candidates. A reasonable estimate of the effect of Republican leadership on the share of safety stops is 0.0545 (0.0882 - 0.0337). Compared to the dependent variable mean in D-to-R counties in the year before the election, this is a 10% decrease.

Changes in the focus on safety and investigatory stops can have racial disparate impact. Black driver's share is generally higher in safety stops than in investigation stops (see Table 1.2). Assuming that the Black driver's share within the safety and investigation stop stays constant after the election in each county, the mere change in the share of safety stops can generate changes in the overall Black driver's share. On the other hand, sheriffs may adopt policies that induce officers to change their practices of conducting specific types of stops, resulting in a change in Black drivers' share *within* the safety and investigation stops. Following this logic, I decompose the changes in the Black driver's share into four parts: 1) the part contributed by the changes in the share of safety stops (while holding the Black driver's share within two types of stops constant), 2) the part contributed by the changes within the safety stops, 3) the part contributed by the changes within the investigation stops, and 4) the left-over second order changes. The derivation of the decomposition is in Appendix.

I report the decomposition results in Table 1.5, Columns (3)-(6). Note that coefficients in Columns (3)-(6) add up to the coefficient in Column (1). Column (3) shows that the changes in the share of safety stops contribute to the change in the Black driver's share but to a small extent. Only 16% of the changes in the Black driver's share can be explained by the changes in the share of safety stops. The major contributor is the change within the safety stops instead of the investigation stops. Changes within the safety stops account for 68% of the total changes (Column (4)).

1.5.3   Personnel policies

Officers play essential roles in shaping racial disparities in traffic stops (Antonovics and Knight, 2009; Goncalves and Mello, 2021; Grosjean et al., 2022). Literature, however, knows little about how officers respond to leadership and whether leaders assign traffic stop tasks based on officers' traffic stop styles that may be related to the share of Black drivers in traffic stops. I test two mechanisms related to officers that may lead to a change in Black drivers' share. First, officers respond to the new sheriff's leadership by changing traffic stop practices. Second, officers do not change their traffic stop practices, but the reshuffling of the personnel by the new sheriffs makes the agencies have a higher Black drivers' share in traffic stops.

To test the two mechanisms, I decompose the difference in Black drivers' share at the agency level across years into four parts, in the same way as in section 1.5.2. Here, the stops are categorized based on who conducted the stops: stayers or non-stayers. The changes in the Black driver's share across years can be decomposed into first, holding the Black driver's share within stayer and non-stayer stops the same as in the base year, changes in the share of stops done by stayers. The second and third parts are changes in the Black driver's share within stayer and non-stayer stops, holding the share of total stops done by stayers the same as in the baseline year. The fourth part is the second-order changes. For details of the decomposition, see Appendix.

The first mechanism, the officers' response to new leadership, would be captured by the second decomposed part: the changes in the Black driver's share *within* stayer stops. The second mechanism, the personnel reshuffling, would be captured by the first and third decomposed parts. The first decomposed part would explain some of the total changes in the Black driver's share if the new sheriff shuffled specific types of officers out of the patrolling team, making stayers and non-stayers (before the election) stops have different levels of the Black driver's share. The third decomposed part would contribute to the total changes if the new sheriffs shuffled specific types of officers out *or* in, making the Black driver's shares within non-stayers vary over time.

Table 1.6 reports the decomposition results. The total changes in the Black driver's share

in Column (1) are decomposed into four parts in Columns (3)-(6). Column (2) shows that D-to-R transitions are significantly associated with a smaller share of stops done by stayers post elections, a 19 percentage points decrease. This is consistent with a scenario where new sheriffs assign patrolling duties to different officers after elections. Although the share of non-stayers stops increases after the elections, such change cannot explain the changes in the Black driver's share. The interaction term between the post-election and D-to-R dummy in Column (3) is small and insignificant. This suggests that the selection of officers continuing the patrolling duty among all who conducted stops before the elections is not based on the Black driver's share at the officer level. The bulk of the changes in the Black driver's share at the agency level is explained by the second decomposed parts, shown in Column (4), the changes of the Black driver' share *within* stayers. Within-stayer changes (holding the share of stayer stops constant) account for 79% of the total changes in Black drivers' shares. The changes within non-stayers, shown in Column (5), are a non-negligible magnitude but not significant. Overall, the decomposition results in Table 1.6 offers evidence in favor of the mechanism where officers' responses to new leaders contribute to the changes in the Black driver's share.

I now directly examine the changes in Black driver's share within stayers and non-stayers by estimating a regression with specification 1.1 with outcome variables: Black drivers' share within stayer stops and Black drivers' share within non-stayer stops. The estimation results are reported in Columns (1) and (2) in Table 1.7. Column (1) shows that stayers in D-to-R agencies, on average, increase the Black driver' share by four percentage points after elections relative to the changes in stayers in D-to-D agencies. The non-stayers in post-election years in D-to-R agencies do not behave very differently compared to pre-election years relative to the behavior changes in non-stayers in D-to-D agencies. Column (1) suggests that the stayers as a whole group change their traffic stop practices in D-to-R agencies, but it speaks little to whether the changes come from a small set of officers or a wide set of officers. The evidence presented next supports that the changes come from a common practice change.

To examine how widespread it is that stayers change their traffic stop practices against

18

certain racial groups, I measure the "tendency to stop Black drivers" in the following way and examine the changes in the tendencies at the officer level across the years. The tendency to stop Black drivers at the officer level is measured in two steps. First, I regress a dummy variable of whether a driver is Black on stop time fixed effects and stop location fixed effects. Stop times are at the quarter-period level. There are four quarters in a year and four time periods in a day divided by three time points: six am, noon, and six pm. Stop locations are the finest geography level recorded for the stop. They can be county, city, census-designated places (CDP), or intersections. Second, for each officer in an election cycle, I take two averages: one comes from stops before, and one comes from stops after elections. I then take the differences in the tendencies within each officer and plot the cumulative distribution function of the differences in Figure 1.3. Figure 1.3 goes against the hypothesis that the practice change is confined to a small set of officers. If the behavior changes in Column (1) in Table 1.7 are driven by a few officers, I would expect the top ten percent of officers in D-to-R counties to have larger tendency changes than the ones in D-to-D counties. On the contrary, I find that the officers ranked in the top ten percent in terms of their tendency changes in D-to-D and D-to-R counties have similar levels of differences.

I conclude the personnel analysis by examining how much more reshuffling happens in D-to-R counties than in D-to-D counties. I estimate a regression in specification 1.1 with outcome variables being the share of officers who are non-stayers and who are new officers at the agency level. An officer is a new officer in that year if the first traffic stop done by him/her in that agency is recorded in that year. [5] Column (3) in Table 1.7 shows that D-to-R counties have an increase in the share of non-stayers by sixteen percentage points, compared to D-to-D counties after elections. The increase in the share of non-stayers, not just the share of stops done by non-stayers (Column (2) in Table 1.6), suggests that the new Republican sheriffs shuffle in patrol teams *many* officers who did not conduct traffic stops in the two years before the elections.

---

[5]I can only identify unique officer IDs within agencies so I cannot identify the first traffic stop in an officer's career in North Carolina.

In particular, many newly shuffled-in officers have not conducted any traffic stops in the agency before the elections (Column (4)).

I provide two takeaways from the analysis of officers. First, a large set of officers in D-to-R counties seem to change their traffic stop practices in response to the new Republican leadership. Second, new Republican sheriffs reshuffle the patrolling teams by assigning new officers to the teams. But the officers shuffled out and in behave similarly in terms of the share of Black drivers they stopped. The two takeaways contribute to the literature by showing that officers' behavior may be malleable by a leader's management/policy. Policymakers who aim to reduce racial disparities in traffic stops can potentially learn from the differences in the management/policies done by law enforcement leaders from different party affiliations.

1.5.4   Patrol Policies

The last policy dimension I look at is the patrolling time and locations. To see if the Republican sheriffs focus on patrolling at times and locations with more Black drivers on the road, I conduct an exercise to see if predictions on whether a stopped driver is Black in post-election years based on time and locations using pre-election data can explain the changes in Black driver's share seen in Table 1.3 Column (1).

The exercise consists of two steps. First, using only the pre-election data within an election cycle, I regress a dummy variable indicating whether the stopped driver is Black on stop location or stop time fixed effects. As defined in the previous sections, the stop times are at the quarter-period level. There are four quarters in a year and four time periods in a day divided by three time points: six am, noon, and six pm. Stop locations are the finest geography level recorded for the stop. They can be county, city, census-designated places (CDP), or intersections. Unfortunately, only 60% of the stops contain geographical information finer than the county level. I then use the OLS coefficients on the stop time or stop location fixed effects (unique to each county) to predict the probability of a stop with a Black driver for all observed pre and post-election stops. Second, I compute the averages of the predicted probabilities at the

county-year level and estimate a regression in specification 1.1 with such averages as the outcome variable.

Table 1.8 reports the estimation results. Across columns, I find that the predicted probabilities of a stop associated with a Black driver based on time or location do not significantly change in D-to-R counties in post-election years. This holds true for both safety (Column (3)-(4)) and investigation stops (Column (5)-(6)). The regression estimation results suggest that the changes in the Black driver' share under the new Republican sheriffs' leadership are not driven by a shift of focus in patrolling specific neighborhoods or times of the day. I conclude on the patrolling policies by providing a caution: around 40% of the stops do not have stop neighborhood information in the estimation sample. A shift of focus in the neighborhood may not be detectable with such data. Further research on the impact of leaders on traffic stops should try to find a setting with better stop location data.

1.5.5 Search Rate

Thus far, I examine if partisan leadership affects whom to stop. I now turn to the behaviors after stopping a driver: whether to search a vehicle or not. I report the changes in the search rate for all stops and stops in different racial groups. I then further examine the search rate separately for safety and investigation stops. Since the stop decision is shown to be affected by the previous sections, the changes in the search rates should be interpreted as the *combined* impacts of stop and search policy changes associated with the new Republic sheriff. In particular, one should not interpret the changes in the search rates (if any) as the changes in the officer's search behavior, holding the stopped driver's population the same as before elections. [6] Instead, the thought exercise here is to hold the at-risk population of being stopped the same. In particular, the proportion of drivers with contraband and drivers with unsafe driving behaviors in each racial group is thought to be unchanged right before and after elections.

---

[6]The thought exercise of holding the stopped driver's population the same is often evoked in papers that aim to explore the racial bias of officers in search behaviors (Antonovics and Knight, 2009) and to explore the efficiency of searches across racial groups (Feigenberg and Miller, 2022)

Table 1.9 reports the regression estimation results in specification 1.1 with overall search rates, Black driver search rates, and non-Black driver search rates at the county-year level being outcome variables. Results in Panel A show that search rates in all stops (combining safety and investigatory) do not significantly change in D-to-R counties in post-election years.

Next, I examine the search rates separately for safety and investigation stops in Panel B and C. Sheriffs may have specific policies for different types of stops, creating heterogeneity. I find that, if anything, the search rates in safety stops increase in D-to-R counties after elections, and the increase seems to appear in all racial groups. Again, this potential increase in search rates should be considered as the impact of combining (a) the decreased share of safety stops (Column (2) in Table 1.5) and (b) any search behavior changes. Overall, there are no changes in search rate racial disparities associated with sheriff's party affiliation.

1.5.6 Efficiency

Finding contraband has long been considered an important part of a law enforcement agency's objective function. The unconditional hit rate, defined as the number of searches with found contraband divided by the number of stops, can thus be seen as an efficiency measure of the law-enforcement agency's traffic stop performance. Slightly different from the search rate racial disparity versus unconditional hit rate trade-offs more commonly seen in the literature (Feigenberg and Miller, 2022), here, the trade-off is between the stop racial disparity and the unconditional hit rate.

Table 1.10 reports estimation results of specification 1.1 with unconditional hit rates as the outcome variables. Results in Column (1) in Panel A show that the overall unconditional hit rates do not change in D-to-R counties in post-election years. Although the unconditional hit rates in Black stops marginally significantly increase, especially in safety stops (Column (2) in Panel A and Panel B), the magnitude is not large enough to drive an increase in the overall unconditional hit rates.

Taking the results in Table 1.3 and Table 1.10 together, the newly elected Republican

sheriffs enact policies that induce larger racial disparities in traffic stops without a discernible increase in the efficiency measured by the unconditional hit rates.

1.5.7 Long(er)-term impacts

In previous sections, I focus on the short-term impacts of partisan leadership, comparing traffic stop practices right after the elections with those before the elections. A natural request is to examine the long-term impact permitted by the research design restrictions. To this purpose, I extend the analysis period to four years before and after the elections and estimate the partisan leadership impacts with a specification similar to equation 1.1. The only difference is that one election cycle now contains eight years, so $e \in \{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$, where 0 denotes the year the elections happened.

Two caveats should be kept in mind in the longer-term analysis. First, drivers may respond to the new traffic stop practices initiated by the new sheriffs in the longer term. One would then be unable to estimate the causal impact of partisan leadership on the racial composition of traffic stops holding the at-risk driver population constant. Second, the newly elected sheriffs in the D-to-D and D-to-R counties may face different pressure for their next election. Among the counties included in the estimation sample, 60% of D-to-D counties have the winner's vote share smaller than 80% in the next election, while 40% of D-to-R counties fall into such category. The parallel trend assumption thus may fail as the counties progress into the next elections. The number of county-cycles that satisfy the sample selection criterion (the number of stops is larger than 50 every year) decreases from 61 to 47 once I extend the election cycle to eight years.

I present the Black driver's share in the longer cycles in Figure 1.4. The gap between D-to-D and D-to-R groups decreases significantly right after the elections, the same pattern as in the shorter election cycles in Figure **??**. Progress to the end of the election cycle, the gap widens to a similar level to pre-election periods. The increase in gaps is driven by D-to-R counties having a lower Black driver's share three and four years after the elections.

Table 1.11, Column (1) confirms the pattern seen in Figure 1.4. Black driver's share

23

increases by 2.7 percentage points in D-to-R counties one year after the election compared to D-to-D counties. The magnitude of the estimate is similar for the year after, but the standard errors become larger. Three and four years after the election (or one and two year before the next election), The difference in Black driver's share in D-to-D and D-to-R counties become much smaller and are not statistically significantly different from differences in the baseline year $(t-1)$. Weighting the observations by the number of stops at the beginning of the cycle increases the magnitude of the coefficients for all post-election periods (Column (2)), suggesting that some small agencies may drive the decrease in the magnitudes in Column (1). Unfortunately, the standard errors also become larger in Column (2), making the estimates non-significant. The decrease in magnitudes in Column (1) can also not be explained by sheriff's offices responding to any policy changes in the police departments. Column (4) shows that Black drivers' share of stops done by police officers in D-to-D and D-to-R counties exhibit a similar trend along the whole electoral cycle.

Overall, the long(er)-term results provide a caution to the interpretation of the results in section 1.5.1. The impact of partisan leadership on racial disparities in traffic stops may be short-lived. The short-lived impact is perhaps not surprising: law-enforcement leaders' policy choices may be influenced by temporal incentives over time, e.g., election pressure from the upcoming elections. Drivers may respond to the new traffic stop policies in a short period of time. I conclude the long(er)-term discussion by cautioning that identifying the long-term impact of leaders on traffic stops may be more challenging than other law-enforcement practices.

1.6   Conclusion

I present evidence that partisan leadership affects traffic stop behaviors. A Democratic-to-Republican sheriff turnover, compared to a Democratic-to-Democratic turnover (which may or may not involve sheriff turnover), leads to an increase of 3.2 percentage points in the Black driver's share among all stops. Speaking to the recent policy proposals that aim at reducing

racial disparities by changing the composition of traffic safety and investigation stops, I find evidence that most of the Black driver' increase comes from changes *within* safety stops, rather than changes in the composition of safety and investigation stops. In relation to the importance of officer-level practices in determining racial disparities, I find evidence that the same set of officers can behave differently in their tendencies to stop Black drivers in response to leadership changes. In particular, I find evidence more consistent with the increase in the Black driver's share driven by medium-level changes across a large set of officers, instead of drastic changes concentrated in a small set of officers. With the limited amount of geographical information recorded in the dataset, I find no evidence that the increase in the Black driver's share is driven by Republican sheriffs focusing on patrolling neighborhoods or at times of the day different from the previous sheriffs.

The increase in the racial disparities in traffic stops, however, does not come with an increase in the efficiency measured by the unconditional hit rates, despite that the Republican sheriffs seem to put more focus on crime investigation than traffic safety.

Given the empirical evidence I find that Republican sheriffs are associated with changes in the share of traffic safety stops, one interesting direction of future research is to examine if efficiency measures on traffic safety stops (e.g., number of crashes) respond to the party affiliation of sheriffs.

1.7   Figures and Tables
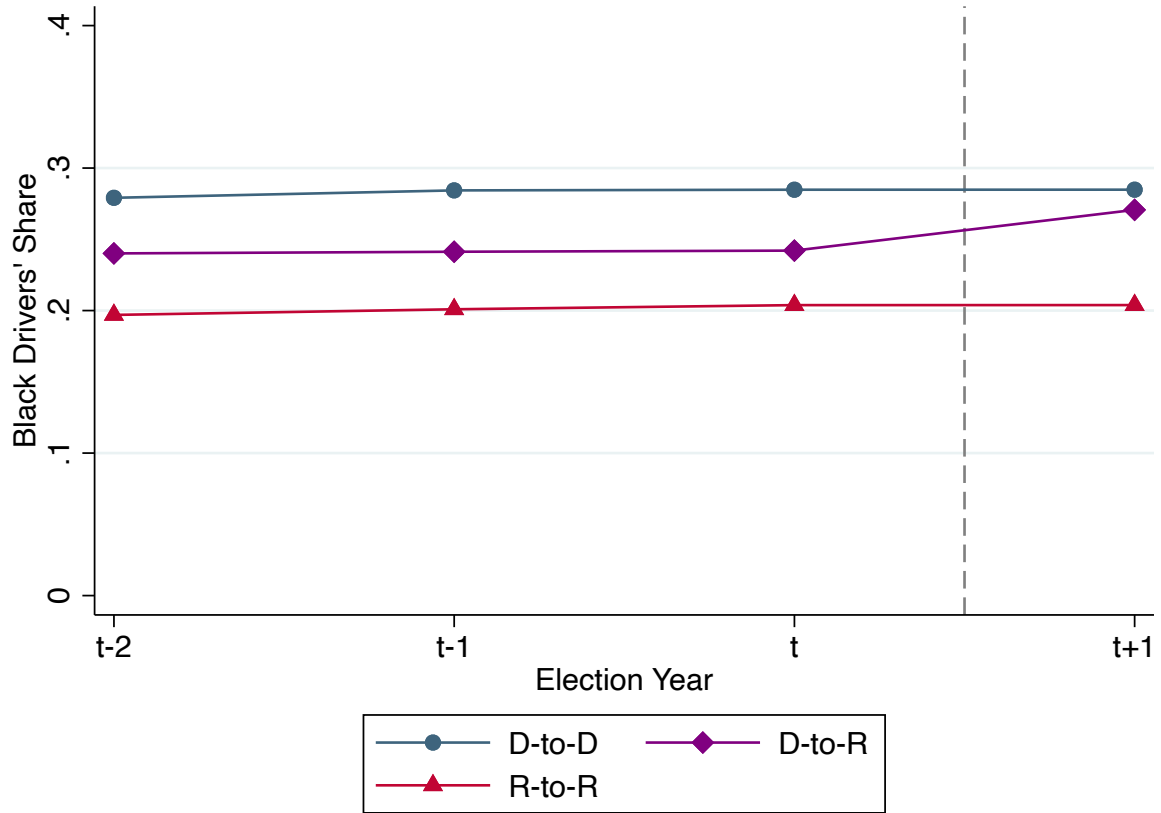
Figure 1.1. Black Drivers' Share Among All Stops

Notes: This figure plots the raw data pattern. I first compute the black driver's share at county-year level. I then compute the simple average of the black driver's share within D-to-D/D-to-R/R-to-R groups, stacking up the three election cycles. Each dot thus contains samples from three years.
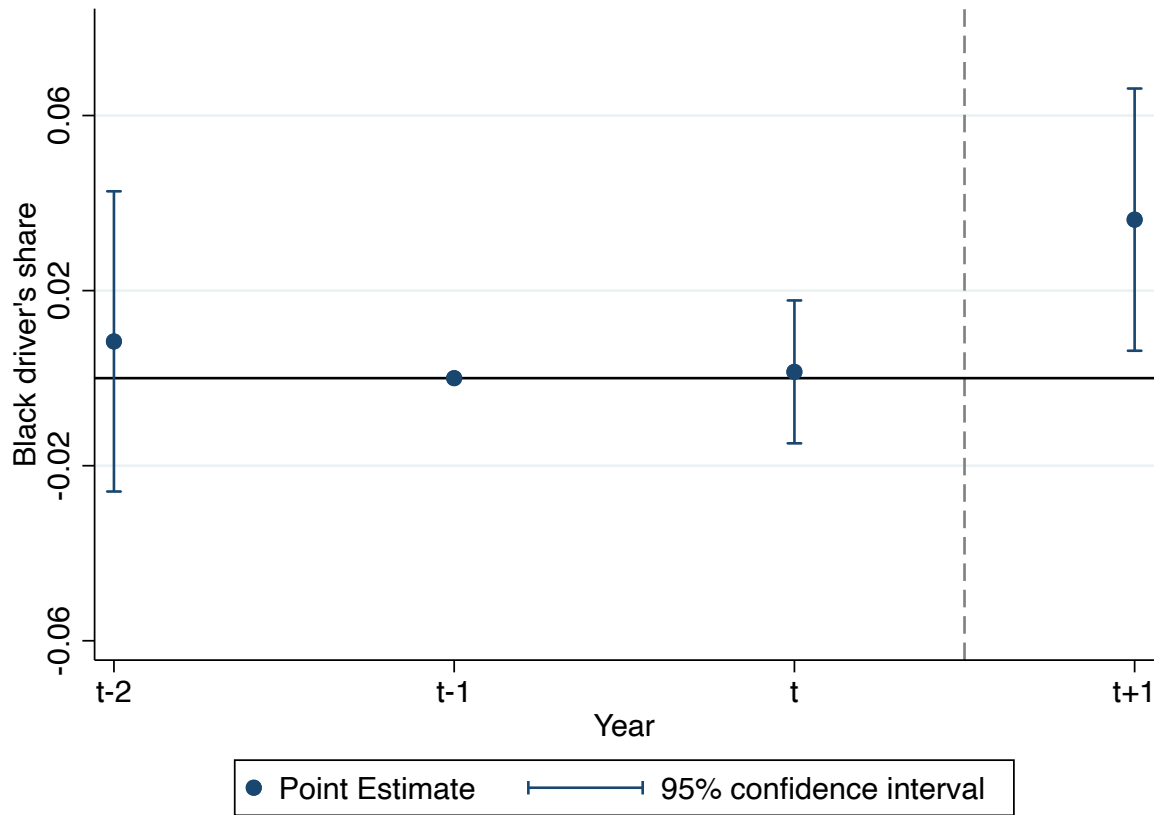
Figure 1.2. The Impact of Partisan Sheriffs on Black Driver's Share

Notes: This figure plots the point estimate and 95% confidence intervals of β_e (coefficients on interaction terms of D-to-R dummy and election cycle-year dummy variables) from a regression estimation of equation 1.1 with the Black driver's share as the outcome variable. t denotes the year when the election happened.
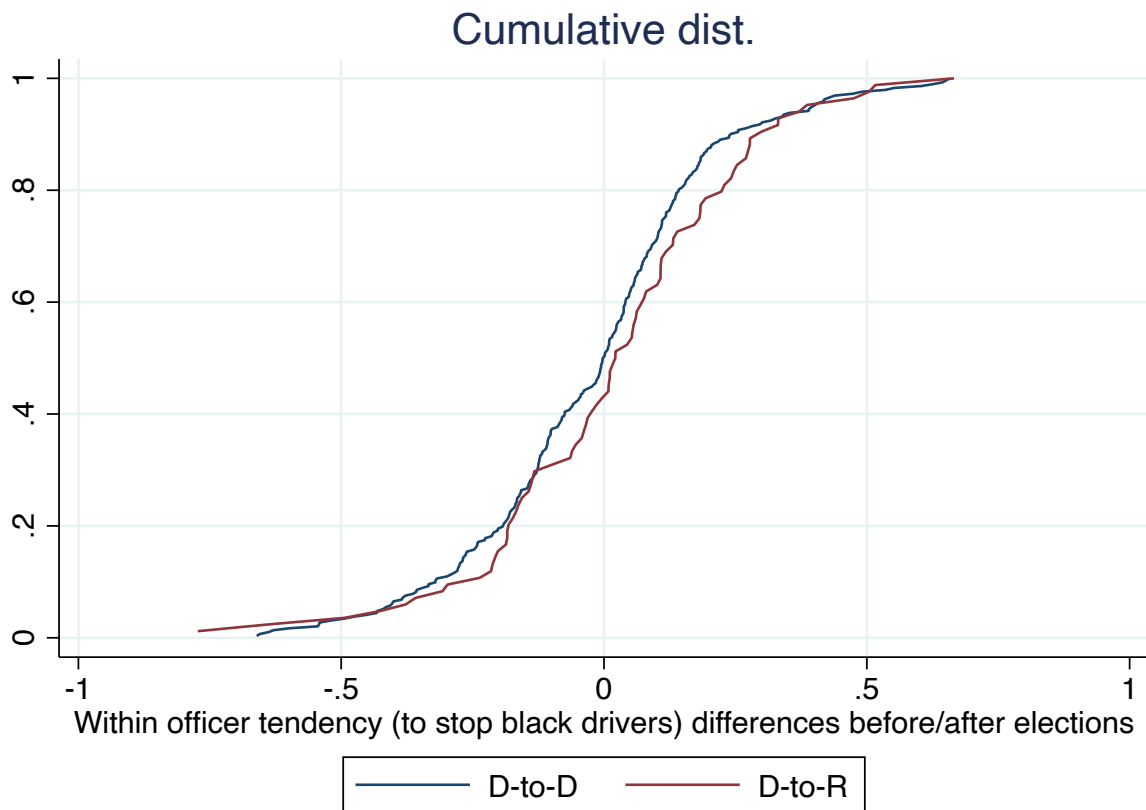
Figure 1.3. Cumulative distributions of the differences in the tendency of stopping black drivers before and after elections among stayers

Notes: This figure plots two cumulative distributions of the difference in the tendencies to stop black drivers before and after elections at the officer level, one for the officers in the D-to-D counties and one for the officers in the D-to-R counties. The tendency to stop black drivers is derived from two steps. First, I regress Black stop (one if the stop driver is black, zero otherwise) on stop location and stop time fixed effects, and get the residuals. Stop locations are counties or cities/towns. I divide a day into four time periods by three time points: 6 am, 12 am, and 6 pm. Stop time is quarter (four quarters in a year) × time period. Second, I compute the average of the residuals for each officer. Only stayers are included in this graph since I need the officers to conduct stops both before and after elections.
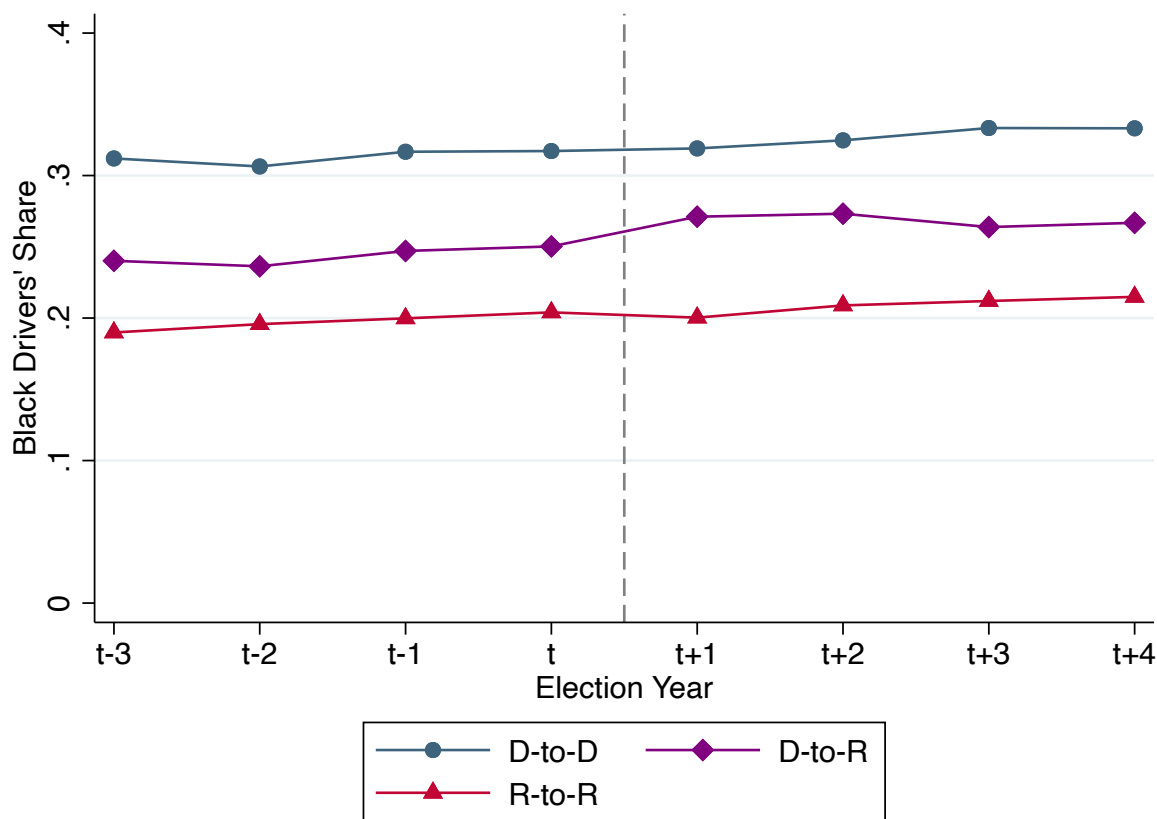
Figure 1.4. Black Drivers' Share Among All Stops in Longer Electoral Cycles

Notes: This figure plots the raw data pattern. I first compute the black driver's share at county-year level. I then compute the simple average of the black driver's share within D-to-D/D-to-R/R-to-R groups, across the three election cycles. Each election cycle is eight year, four year before and after the elections. Each dot contains samples from three years.

Table 1.1. Sheriff Election Results in North Carolina

Panel A: All Sheriffs' Offices						
Election Year	R to R Turnover	R to R No Turnover	R to D	D to D Turnover	D to D No Turnover	D to R
2010	10	24	0	15	46	5
2014	5	33	1	14	37	10
2018	13	32	3	16	27	9
Panel B: Offices with Winners' vote share < 80%						
2010	8	17	0	12	26	5
2014	3	16	1	8	21	9
2018	4	12	3	6	8	8
Panel C: Offices with Winners' vote share < 80% and number of stops > 50 every year						
2010	3	14	0	4	14	4
2014	3	12	0	6	15	6
2018	3	7	3	4	3	5
Panel D: Winners' vote share distribution in all D to D and D to R elections						
	2010		2014		2018	
Winner's vote share	D-to-D	D-to-R	D-to-D	D-to -R	D-to-D	D-to-R
<=0.6	13	3	11	8	5	7
0.6 − 0.7	15	1	8	1	7	0
0.7 − 0.8	11	1	10	1	2	1
>= 0.8& < 1	4	0	4	0	6	0
1	18	0	18	0	23	1

Notes: D refers to the Democratic party, and R refers to the Republican party. North Carolina has 100 sheriff's offices, one for one county. Panel A presents the party turnover distributions in all elections from 2010 to 2018. Panel B drops elections in which the winner's vote share is larger than 80%. This criterion is chosen to match the vote share support of D-to-R elections. Panel C drops elections that are dropped in Panel B and further drops the ones in which the county had at least one year that had fewer than 50 traffic stops in that four-year cycle (from 3 years before the election to 1 year after the election). Panel D presents the winner's vote share distribution in all D-to-D (turnover and no turnover) and D-to-R elections. An election with the winner's vote share being one means there was only one candidate in that election. I use county-cycles in Panel C in the estimation.

Table 1.2. Summary Statistics of Traffic Stops and Searches

| | Stops by Motorists' Group | | | Stops by Types | | All |
	Black	Hispanic	White	Safety	Investigation	
Share Black	1.000	0.000	0.000	0.238	0.278	0.257
Share Hispanic	0.000	1.000	0.000	0.068	0.070	0.069
Share White	0.000	0.000	1.000	0.669	0.634	0.652
Share Female	0.361	0.239	0.359	0.357	0.343	0.350
Share Safety Stops	0.478	0.511	0.530	1.000	0.000	0.517
Share Investigatory Stops	0.522	0.489	0.470	0.000	1.000	0.483
Search Rate	0.079	0.087	0.061	0.051	0.085	0.067
Unconditional Hit Rate	0.024	0.017	0.021	0.016	0.027	0.022
Observations	84,595	22,600	214,132	169,809	158,730	328,539

Notes: This table presents summary statistics including all county-cycles included in Panel C in Table 1. All stops can be categorized into safety or investigatory stops. Safety stops includes stops due to Speed Limit Violation, Stop Light/Sign Violation, Driving While Impaired, Safe Movement Violation. Investigatory stops include stops due to Vehicle Equipment Violation, Vehicle Regulatory Violation, Seat Belt Violation, Investigation, and Other Motor Vehicle Violation.

Table 1.3. The Impact of Partisan Sheriffs on Black Driver's Share: Regression Estimates and a Placebo Test

| | | $\frac{\text{\# of black driver}}{\text{\# of all stops}}$ | | |
	(1)	(2)	(3)	(4)
		Sheriff's offices		Police departments
t-2 x D-to-R	0.0080	0.0094	-0.0196	-0.0123
	(0.0173)	(0.0082)	(0.0277)	(0.0120)
t x D-to-R	0.0007	0.0084*	-0.0116	-0.0018
	(0.0082)	(0.0049)	(0.0146)	(0.0137)
t+1 x D-to-R	0.0326**	0.0312*	0.0319	0.0039
	(0.0151)	(0.0172)	(0.0230)	(0.0145)
County-Cycle	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Weight	Agency	# of stops	Agency	Agency
Sample	All	All	Close election	All
N	244	244	104	164
Dep. mean	0.2413	0.1878	0.2425	0.2293

Notes: Clustered standard errors at the county level in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All outcome variables are at county-year level. t refers to the year of election in that election cycle. I report the coefficients of the interaction terms between the (reletative) election year dummy variables with the D-to-R dummy variable. The D-to-R dummy variable is one if the county experienced a D-to-R election in that cycle and zero if the county experienced a D-to-D election. Column (1)-(3) reports regression results with traffic stop samples from sheriff's offices. Column (4) reports results with samples from police departments in the same set of counties as in Columns (1) and (2). The sample size is smaller in Column (4) because not all counties have police departments. All regression specifications include county-cycle and election-year fixed effects. I weight the county-year observations by the number of stops of that county in $t - 2$ in Column (2). In Column (3), I restrict the samples to counties where the winner's vote share is below 60%. Dep. means are computed from D-to-R counties in year $t - 1$, one year before the sheriff election.

Table 1.4. The Impact of Partisan Sheriffs on the Number of Stops by Race and Stop Purposes

| | ln(number of stops) | | | |
| | (1) | (2) | (3) | (4) |
	Black	Non-Black	Diff b/w races	All stops
t-2 x D-to-R	-0.211	-0.198	-0.194	
	(0.173)	(0.148)	(0.145)	
t x D-to-R	-0.437**	-0.505***	-0.495***	
	(0.178)	(0.170)	(0.169)	
t+1 x D-to-R	0.183	0.0032	0.0137	
	(0.288)	(0.279)	(0.276)	
t-2 x D-to-R x Black			-0.0210	
			(0.0944)	
t x D-to-R x Black			0.0481	
			(0.0584)	
t+1 x D-to-R x Black			0.158*	
			(0.0918)	
t-2 x Close				-0.490***
				(0.153)
t x Close				-0.126
				(0.166)
t+1 x Close				-0.0112
				(0.217)
County-Cycle	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
N	244	244	488	244
Average # of stops	268	1042		1336

Notes: Clustered standard errors at the county level are in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All outcome variables are at the county-year level. t refers to the year of election in that election cycle. I report the coefficients of the interaction terms between the (relative) election year dummy variables with the D-to-R dummy variable in equation 1.1 in Columns (1)-(2). The D-to-R dummy variable is one if the county experienced a D-to-R election in that cycle and zero if the county experienced a D-to-D election. Column (3) reports the regression estimation results from specification 1.2, where Black is a dummy variable being one if the county-year observation is the number of stops on Black drivers, 0 otherwise. Column (4) reports estimation results with specification 1.1 with the log of the number of all stops as the outcome variable. I include the same county-cycles as in Column (1) in the estimation reported in Column (4). Close is a dummy variable being one if the county experienced an election in which the winner's vote share is below 60%, 0 otherwise. The average number of stops is computed from D-to-R (Close election) counties in year $t - 1$, one year before the sheriff election.

Table 1.5. Decomposition of the Changes in Black Driver's Share: Type of Traffic Stops

	(1) All Black Stops / All Stops	(2) All Safety Stops / All Stops	(3) $\Delta S_{i,(-1,t)}(B_{1i,-1}-B_{2i,-1})$ Changes in the share of safety stops	(4) $S_{i,-1}\Delta B_{1i,(-1,t)}$ Changes within safety stops	(5) $(1-S_{i,-1})\Delta B_{2i,(-1,t)}$ Changes within investigation stops	(6) $\Delta S_{i,(-1,t)}(\Delta B_{1i,(-1,t)}-\Delta B_{2i,(-1,t)})$ Second order changes
t-2 x D-to-R	0.0080	-0.0075	-0.0025	0.0062	0.0039	0.0005
	(0.0173)	(0.0288)	(0.0029)	(0.0076)	(0.0091)	(0.0014)
t x D-to-R	0.0007	-0.0337*	0.0026*	0.0026	-0.0044	-0.0000
	(0.0082)	(0.0188)	(0.0014)	(0.0049)	(0.0066)	(0.0013)
t+1 x D-to-R	0.0326**	-0.0882***	0.0054**	0.0224**	0.0072	-0.0023
	(0.0151)	(0.0234)	(0.0022)	(0.0111)	(0.0107)	(0.0034)
County-Cycle	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
N	244	244	183	183	183	183
Dep. mean	0.2413	0.5281	0	0	0	0

Notes: Columns (1) and (2) in the table report estimation coefficients from an OLS regression with specification as in equation 1.1. Column (3)-(6) reports estimation coefficients from an OLS regression with specification as in equation 1.3 in the Appendix. All outcome variables are at the county-year level. t refers to the year of election in that election cycle. Estimation results in Columns (3)-(6) are the decomposition of the results in Column (1). Adding up coefficients from Columns (3)-(6) would equal the coefficient in Column (1). I denote B_{1it} and B_{2it} as the share of black drivers in safety and investigation stops for county i in year t. There are four time periods, $t = -2, -1, 0, 1$. I set $t = -1$ as the baseline period. I denote S_{it} as the share of safety stops of all stops. Then $1 - S_{it}$ is the share of investigation stops of all stops. I denote $\Delta S_{i,(-1,t}$ as the difference of the share of safety stops for county i between period -1 and t. Column (3) represents the contribution to the changes in the black driver's share from changes in the share of safety stops of all stops (while keeping the black driver's share within each type of stop constant). Columns (4) and (5) represent the contribution from changes in the black drivers' share within safety and investigation stops. Column (6) is the leftover second-order changes (contribution from deviation from both the share of safety stops and black driver's share in safety and investigation stops). See the Appendix for the derivation of the decomposition. Clustered standard errors at the county level are in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dep. mean computed from D-to-R counties in year $t - 1$.

Table 1.6. Decomposition of the Changes in Black Driver's Share: Officer

	(1) All Black Stops / All Stops	(2) All Stayer Stops / All Stops	(3) $\Delta S_{i,(-1,t)}(B_{1i,-1}-B_{2i,-1})$ Changes in the share of stayer stops	(4) $S_{i,-1}\Delta B_{1i,(-1,t)}$ Changes within stayer stops	(5) $(1-S_{i,-1})\Delta B_{2i,(-1,t)}$ Changes within non-stayer stops	(6) $\Delta S_{i,(-1,t)}(\Delta B_{1i,(-1,t)}-\Delta B_{2i,(-1,t)})$ Second order changes
t-2 x D-to-R	0.0080	-0.0246	-0.0013	0.0196	-0.0052	-0.0049
	(0.0173)	(0.0352)	(0.0035)	(0.0133)	(0.0081)	(0.0061)
t x D-to-R	0.0007	-0.0050	-0.0003	0.0032	-0.0078	0.0056
	(0.0082)	(0.0762)	(0.0092)	(0.0062)	(0.0084)	(0.0137)
t+1 x D-to-R	0.0326**	-0.191**	0.0088	0.0258*	0.0131	-0.0151
	(0.0151)	(0.0822)	(0.0104)	(0.0147)	(0.0095)	(0.0115)
County-Cycle	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
N	244	244	183	183	183	183
Dep. mean	0.2413	0.5520	0	0	0	0

Notes: Columns (1) and (2) in the table report estimation coefficients from an OLS regression with specification as in equation 1.1. Column (3)-(6) reports estimation coefficients from an OLS regression with specification as in equation 1.3 in the Appendix. Clustered standard errors at the county level are in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All outcome variables are at the county-year level. t refers to the year of election in that election cycle. Dep. mean computed from D-to-R counties in year $t - 1$. Estimation results in Columns (3)-(6) are the decomposition of the results in Column (1). Adding up coefficients from Columns (3)-(6) would equal the coefficient in Column (1). I denote B_{1it} and B_{2it} as the share of black drivers of all stops done by stayers and non-stayers, respectively, for county i in year t. There are four time periods, $t = -2, -1, 0, 1$. I set $t = -1$ as the baseline period. I denote S_{it} as the share of stops done by stayers. Then $1 - S_{it}$ is the share of stops done by non-stayers. I denote $\Delta S_{i,(-1,t}$ as the difference of the shares of stops done by stayers in county i between period -1 and t. Column (3) represents the contribution to the changes in the black driver's share from changes in the share of stops done by stayers. Columns (4) and (5) represent the contribution from changes in the black drivers' share within stops done by stayers and non-stayers. Column (6) is the leftover second-order changes. See the Appendix for the derivation of the decomposition.

Table 1.7. Officer Behavior Change and Personnel Turnover

	(1) Black Stops by Stayers / All Stops by Stayers	(2) Black Stops by Non-Stayers / All Stops by Non-Stayers	(3) # of non-stayers / # of all officers	(4) # of new officers / # of all officers
t-2 x D-to-R	0.0454	-0.0360	-0.00735	-0.00261
	(0.0303)	(0.0399)	(0.0316)	(0.0592)
t x D-to-R	0.00951	-0.0306	0.0265	0.0394
	(0.0130)	(0.0311)	(0.0457)	(0.0589)
t+1 x D-to-R	0.0403**	-0.00510	0.167***	0.217***
	(0.0194)	(0.0323)	(0.0579)	(0.0603)
County-Cycle	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
N	244	244	244	244
Dep. mean	0.2294	0.2661	0.6125	0.3756

Notes: This table reports regression estimation results with specification 1.1 with four outcome variables listed at the head of the table. Stayers are officers who conduct traffic stops both before and after elections. Non-stayers are officers who conduct traffic stops either before or after elections. An officer is a new officer in that year if his/her first traffic stop record in that agency is observed in that year. The D-to-R dummy variable is one if the county experienced a D-to-R election in that cycle and zero if the county experienced a D-to-D election. Clustered standard errors at the county level are in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dep. means are computed from D-to-R counties one year before the election.

Table 1.8. Patrol Location and Time Policy

	(1)	(2)	(3)	(4)	(5)	(6)
			Predicted Black stops			
			Stops			
	All		Safety Stops		Investigatory	
	Location	Time	Location	Time	Location	Time
t-2 x DtoR	0.00307	-0.00125	0.000880	-0.00234	0.00331	0.00317
	(0.00476)	(0.00453)	(0.00395)	(0.00497)	(0.00672)	(0.00417)
t x DtoR	-0.000966	0.000437	0.000549	0.00184	-0.00301	-0.000249
	(0.00369)	(0.00347)	(0.00415)	(0.00370)	(0.00411)	(0.00419)
t+1 x DtoR	0.00505	-0.00212	0.00685	-0.00288	0.000816	-0.00327
	(0.00439)	(0.00410)	(0.00451)	(0.00417)	(0.00512)	(0.00460)
County-Cycle	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
N	244	244	244	244	244	244
dep_mean	0.2417	0.2401	0.2396	0.2352	0.2444	0.2462

Notes: Clustered standard errors at the county level in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dep. mean computed from D-to-R counties before the election. For Columns (1), (3), and (5), we predict whether the stop is associated with a Black driver (Black stop) by the share of Black stops pre-election in each location cell. Locations are places where at least 40 traffic stops were recorded under that place name in the estimation sample. For Columns (2), (4), and (6), we predict whether the stop is a Black stop by the share of Black stops pre-election in each time group x county cell. A day is divided into four time groups by four points: 6 am, noon, 6 pm, midnight.

Table 1.9. The Impact of Partisan Sheriffs on Search Rates by Drivers' Race

	(1) All searches All stops	(2) Black searches Black stops	(3) Non-black searches Non-black stops
Panel A: All stops			
t-2 x DtoR	0.0107	-0.0109	0.0169
	(0.0122)	(0.0176)	(0.0123)
t x DtoR	-0.00139	-0.0216	0.00214
	(0.00930)	(0.0234)	(0.00973)
t+1 x DtoR	0.0177	0.0331	0.0168
	(0.0156)	(0.0246)	(0.0157)
Dep. mean	0.0832	0.1102	0.0768
Panel B: Safety stops			
t-2 x DtoR	0.00871	-0.00472	0.00472
	(0.0175)	(0.0315)	(0.0179)
t x DtoR	-0.00625	-0.0333	-0.00587
	(0.0109)	(0.0232)	(0.0114)
t+1 x DtoR	0.0344*	0.0500	0.0263
	(0.0180)	(0.0354)	(0.0179)
Dep. mean	0.0788	0.0982	0.0723
Panel C: Investigation stops			
t-2 x DtoR	0.0124	-0.0224	0.0260*
	(0.0119)	(0.0178)	(0.0144)
t x DtoR	0.00397	-0.00748	0.0122
	(0.0130)	(0.0302)	(0.0141)
t+1 x DtoR	0.00310	-0.000444	0.00853
	(0.0174)	(0.0313)	(0.0183)
Dep. mean	0.1045	0.1038	0.1078
N	244	244	244
County-Cycle FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes: Clustered standard errors at the county level are in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All outcome variables are at the county-year level. t refers to the year of election in that election cycle. Dep. means are computed from D-to-R counties in year $t-1$, one year before the sheriff election.

Table 1.10. The Impact of Partisan Sheriffs on Unconditional Hit Rates by Drivers' Race

	(1) All contraband All stops	(2) Black contraband Black stops	(3) Non-Black contraband Non-black stops
Panel A: All stops			
t-2 x D-to-R	0.0103	0.0169	0.0105
	(0.00625)	(0.0102)	(0.00637)
t x D-to-R	-0.00285	0.00629	-0.00178
	(0.00563)	(0.0145)	(0.00702)
t+1 x D-to-R	0.00746	0.0181*	0.00578
	(0.00824)	(0.0101)	(0.0102)
Dep. mean	0.0304	0.0337	0.0296
Panel B: Safety stops			
t-2 x D-to-R	0.0154*	0.0109	0.0124
	(0.00847)	(0.0186)	(0.00915)
t x D-to-R	-0.00616	-0.00639	-0.00683
	(0.00629)	(0.0114)	(0.00764)
t+1 x D-to-R	0.0169**	0.0249*	0.0109
	(0.00808)	(0.0137)	(0.0111)
Dep. mean	0.0244	0.0257	0.0247
Panel C: Investigation stops			
t-2 x D-to-R	0.00671	0.0117	0.0105
	(0.00866)	(0.0116)	(0.0102)
t x D-to-R	0.00141	0.0148	0.00635
	(0.00854)	(0.0208)	(0.0108)
t+1 x D-to-R	-0.000221	0.0117	0.00144
	(0.0109)	(0.0137)	(0.0151)
Dep. mean	0.0376	0.0428	0.0354
N	244	244	244
County-Cycle FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes: Clustered standard errors at the county level are in parentheses.Statistical significance is denoted: * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$. All outcome variables are at the county-year level. t refers to the year of election in that election cycle. Dep. means are computed from D-to-R counties in year $t-1$, one year before the sheriff election. Contraband refers to searches that found contraband successfully.

Table 1.11. Longer-term Impact of Partisan Sheriffs on Black Driver's Share

	(1)	(2)	(3)	(4)
		$\frac{\text{\# of black driver}}{\text{\# of all stops}}$		
	Sheriff's offices			Police departments
t-3 x D-to-R	-0.0006	0.0106	-0.0173	0.0101
	(0.0119)	(0.0098)	(0.0159)	(0.0098)
t-2 x D-to-R	0.0039	0.0051	-0.0250	-0.0175
	(0.0186)	(0.0072)	(0.0248)	(0.0156)
t x D-to-R	0.0031	0.0070*	-0.0013	0.0021
	(0.0082)	(0.0038)	(0.0137)	(0.0240)
t+1 x D-to-R	0.0278**	0.0312	0.0091	0.0096
	(0.0122)	(0.0193)	(0.0146)	(0.0258)
t+2 x D-to-R	0.0262	0.0300	-0.0013	-0.0130
	(0.0194)	(0.0200)	(0.0241)	(0.0205)
t+3 x D-to-R	0.0061	0.0269	-0.0228	-0.0032
	(0.0170)	(0.0194)	(0.0171)	(0.0258)
t+4 x D-to-R	0.0106	0.0146	0.0031	-0.0028
	(0.0126)	(0.0198)	(0.0166)	(0.0239)
County-Cycle	Yes	Yes	Yes	Yes
Year-Cycle	Yes	Yes	Yes	Yes
Weight	Agency	# of stops	Agency	Agency
Sample	All	All	Close election	All
N	376	376	144	232
Dep. mean	0.2471	0.1720	0.2446	0.2867

Notes: Clustered standard errors at the county level are in parentheses. Statistical significance is denoted: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All outcome variables are at the county-year level. t refers to the year of election in that election cycle. Dep. means are computed from D-to-R counties in year $t-1$, one year before the sheriff election.

1.8 Appendix

Decomposition of the Total Changes in the Black Driver's Share

Let B_{it} denote the share of black drivers in all stops for county i in year t. Following the timing convention in this paper, $t = -2, -1, 0, 1$, I set $t = -1$ as the baseline period. Let S_{it} be the share of safety stops of all stops. Then $1 - S_{it}$ is the share of investigation stops of all stops. I denote B_{1it} and B_{2it} as the share of black drivers in all safety and investigation stops. I can then write:

$$B_{it} = S_{it} \times B_{1it} + (1 - S_{it}) \times B_{2it}.$$

Re-writing the level of shares as the baseline level plus deviations, we have:

$$B_{it} = B_{i,-1} + \Delta B_{i,(-1,t)},$$

$$S_{it} = S_{i,-1} + \Delta S_{i,(-1,t)},$$

$$B_{1it} = B_{1i,-1} + \Delta B_{1i,(-1,t)},$$

$$B_{2it} = B_{2i,-1} + \Delta B_{2i,(-1,t)}.$$

Taking the difference $B_{it} - B_{i,-1}$, we have:

$$B_{it} - B_{i,-1} = \underbrace{[S_{i,-1} \cdot \Delta B_{1i,(-1,t)}]}_{\text{Changes within Safety Stops}} + \underbrace{[(1 - S_{i,-1}) \cdot \Delta B_{2i,(-1,t)}]}_{\text{Changes within Investigation Stops}}$$

$$+ \underbrace{[\Delta S_{i,(-1,t)} \cdot B_{1i,-1} - \Delta S_{i,(-1,t)} \cdot B_{2i,-1}]}_{\text{Changes from Shares of Safety Stops}}$$

$$+ \underbrace{[\Delta S_{i,(-1,t)} \cdot (\Delta B_{1i,(-1,t)} - \Delta B_{2i,(-1,t)})]}_{\text{Second Order Changes}}.$$

Decomposing the difference, the first bracket is the contribution from the changes in the share of

black drivers in all safety stops; the second bracket is the contribution from the changes in the share of black drivers in all investigation stops. The first and second brackets are the outcome variables in Column (4)-(5) in Table 1.5. The third bracket is the contribution from changes in the share of safety stops of all stops, while the fourth bracket is the leftover second-order term. The third and fourth brackets are the outcome variables in Columns (3) and (6) in Table 1.5.

To see that the estimation results for the coefficients of interest are the same no matter I have the difference between two periods or the level in the year as outcome variables, we duplicate equation 1.1 below:

$$Y_{cle} = \sum_{e=-2}^{1} \beta_e D_{cl}^{D-to-R} \cdot \eta_e + \delta_{le} + \delta_{cl} + \varepsilon_{cle}.$$

Taking the difference $Y_{cle} - Y_{cl,-1}$, we have:

$$Y_{cle} - Y_{cl,-1} = \sum_{e=-2}^{1} \beta_e D_{cl}^{D-to-R} \cdot (\eta_e - \eta_{-1}) + (\delta_{le} - \delta_{l,-1}) + (\varepsilon_{cle} - \varepsilon_{cl,-1}). \qquad (1.3)$$

Hence, I can use the terms in the four brackets above as outcome variables, and estimate four regressions with specifications 1.3 (similar to equation 1.1 but without county-cycle fixed effects), and have four sets of regression coefficient estimates that would add up to the coefficient estimates using the black driver's share as outcome variables.

The decomposition analysis in section 1.5.3 is done in the same procedure by defining B_{1it} and B_{2it} as the share of black drivers within stops done by stayers and non-stayers for county i in year t.

Chapter 2

The Impact of Partisan Politics on Personnel Composition: Evidence from North Carolina's Sheriff's Offices

2.1   Introduction

Over the past century, many local police forces have transformed from politicized, spoils-based organizations to professional organizations governed by civil service laws (Ornaghi, 2019). However, the extent to which civil service protections prevent the politicization of law enforcement offices in the United States today remains unclear. In particular, unlike most police officers, many law enforcement officers working in Sheriff's Offices ("sheriff deputies") are not protected from dismissal due to their political affiliation.[1] Sheriff offices play an important role in United States law enforcement; these offices are the primary law enforcement offices for a significant proportion of the United States population, are responsible for the county jail in almost all counties across the United States and employ 25% of sworn local law enforcement officers in the United States (Brooks, 2019). Even among the police officers and sheriff deputies that are protected from dismissal due to their political affiliation, ideological misalignment with

---

[1]For instance, in *Jenkins v Medford* the Fourth Circuit Court of Appeals ruled that politically-motivated dismissals of sheriff deputies are constitutional because these deputies have "policy making" responsibilities. The court also noted that "The circuits which have examined the interplay between the voters, the sheriff and his policies, and the role of deputies in implementation of policy, have concluded that political affiliation and loyalty to the sheriff are appropriate job requirements. These circuits have held that the position of a deputy sheriff is sufficiently political to allow patronage and politically-motivated dismissals under the exception established by Elrod and Branti."

the office's leader and disagreement with the leader's policies may still lead to differential exit (and entry) from the law-enforcement agency in response to a given leader's policies, creating greater political homogeneity even absent politically-motivated dismissals (Besley and Ghatak, 2005).

In this paper, we study the extent to which law enforcement leaders may affect the political composition of their departments in the context of North Carolina Sheriff's Offices. North Carolina sheriffs are popularly-elected and North Carolina sheriff deputies are not protected from dismissal due to their political affiliation. Understanding the impact of law enforcement leaders' ideology in this context is important to better understand how the removal of civil service protection may impact politicization in other contexts, a question that has increased relevance with recent Republican proposals to remove civil service protections from a large-swath of federal workers.[2]

To study this question, we compare Sheriff's Offices transitioning from a Democratic to Republican sheriff (henceforth D-to-R) [3] to counties who transition from one Democratic sheriff to another (henceforth D-to-D). We find that the replacement of a Democratic incumbent sheriff with a Republican sheriff leads to a large shift in the political composition of the Sheriff's Offices; by two years after the election, the Republican share of sheriff deputies increases by 9 p.p., or 27%, relative to counties in the control group in the same year relative to one year before the election—there is also a decline of similar magnitude in the share of Democratic deputies. We find no evidence for pre-trends two years before the election and a "placebo" test using boards of education in the same counties as treated and control Sheriff's Offices finds no effect of the policy change.

We next decompose the causes of this change in composition and find that it is driven by two main factors. First, the election of a Republican sheriff substantially increases the number of

---

[2]For example, President Trump's Executive Order 13597 would have removed civil service protections from all "policy or rule making officials" had it not been rescinded by President Biden. President Trump has endorsed this policy should he be elected to a second term.

[3]We study only this direction of transition because only a small number of counties transition from a Republican to Democratic sheriff during our sample period.

Democratic deputies that change their voter registration. In the election year, Democrat deputy sheriffs in D-to-R sheriff's offices are 12 p.p. more likely to switch their party registrations than their counterparts in D-to-D sheriff's offices, relative to two years before the election. This is a more than 350% increase given that only 3.4% of deputy sheriffs in D-to-R offices switch their party affiliation two years before the election. Second, the share of Republican deputies among the officers who newly joined the agencies increases by 14 p.p. in the election year (or 33% compared to the baseline share) when the newly elected sheriffs are sworn in with the new deputies. The increase persists for two years after the elections. A third channel that Democratic deputies are more likely to leave the agency after the election contributes to the change in composition to a limited extent.

This paper contributes to two main strands of literature. First, we show evidence that in our context, political turnover leads to an increase in the share of politically-aligned law enforcement officers (to the new leaders). This finding builds on several papers examining the impact of political turnover on personnel in political organizations. For example, work on public employees in Brazil shows that supporters of the party in power are more likely to be hired and are negatively selected on their competence (Colonnelli et al., 2020) and that bureaucracies not shielded from political influence experience higher personnel turnover following a change in governing party (Akhtari et al., 2022). Conversely, studies examining federal workers in the United States have found a more muted effect including Spenkuch et al. (2023), who show that among career federal workers there is little evidence of departures due to ideological misalignment, but instead that ideologically misaligned workers may exert less effort and Bolton et al. (2020), who show increased turnover only among senior career executives in agencies whose views likely diverge from the President's. Our results show that in a local US setting without civil service protection, unlike US federal workers but similar to Brazil, political affiliation of local leaders does appear to have large effects on the composition of the agency workforce.

Second, by showing that a political leader's political affiliation is an important driver of the political composition of officers within his or her agency, we also contribute to a growing

literature on determinants and consequences of the composition of law enforcement agencies. Ba et al. (2022, 2021); Miller and Segal (2019) have all shown that a law enforcement agency's composition along racial, gender and political dimensions likely greatly affects its efficacy. However, aside from Miller and Segal (2012), who show that employment discrimination lawsuits have persistent effects on the share of Black officers employed by a department, there is little evidence on what drives the different composition of law enforcement personnel across agencies. We build on this work by showing that one important determinant of the political composition of an agency's force, at least in the elected Sheriff's Offices, is the political affiliation of the agency's leader.

The rest of the paper proceeds as follows. Section 2.2 provides the institutional context for our analysis. Section 2.3 describes our data and empirical strategy. Section 2.4 describes our primary results. Section 2.5 presents suggestive evidence on the mechanism of these results and Section 2.6 concludes.

2.2   Institutional Context

The context for our study is Sheriff's Offices in North Carolina. Each of the 100 counties in North Carolina has an elected sheriff, which acts as the chief law enforcement officer in the county. The sheriff is elected in a partisan election in the November of Midterm election years and serves a four-year term without term limits. Sheriffs are typically sworn in the first week of December following their election. Sheriff's Offices have county-wide jurisdiction over crimes, but typically focus activities in unincorporated areas and municipalities that do not have their own municipal police departments. Sheriffs also have responsibility for running and managing the county jail as well as other administrative tasks such as processing gun permits.

Sheriff Deputies

Law enforcement officers working in sheriff's offices are called sheriff deputies. Our study uses changes in the political affiliation of elected sheriffs to test the impact of political

leadership on the composition of deputies within the sheriff's offices. One important feature of North Carolina Sheriff's offices in the context of this question is that sheriff deputies, unlike many law enforcement officers and other civil servants, lack civil service protection and thus are much easier to fire and may also be easier to hire.

The law on the constitutionality of dismissing sheriff deputies for their political affiliations has differed across courts; at question is whether political loyalty is a necessary condition for sheriff deputies to perform their jobs. However, in the context of North Carolina specifically, both the Fourth Circuit Court of Appeals and the North Carolina Supreme Court have determined that such firing is permissible due to the deputy sheriff's "policymaking" role. This question was settled most recently in 2016, in a series of cases consolidated in *Young v Bradley* where the court found that "after considering these statutory and decisional factors, we conclude that, by standing in the elected sheriff's shoes, a deputy sheriff fills a role in which loyalty to the elected sheriff is necessary to ensure that the sheriff's policies are carried out."

2.3 Data

2.3.1 Empirical Strategy

The primary empirical goal of this paper is to determine the effect of a change in the political affiliation of an elected sheriff on the political composition of deputies in the sheriff's offices. To answer this question, we compare the political affiliation of an office's deputies before and after an election-driven change in the political affiliation of the elected sheriff to the political affiliations of deputies in counties that also elected a new sheriff but for which the political affiliation of the sheriff did not change. This condition is important because it allows us to isolate the effect of the change in political affiliation independent of any effects that may be driven by sheriff turnover.

Due to data limitations, we study turnovers in three election cycles: 2010, 2014, and 2018. We focus on the three years before and two years after an election. Although this prevents

us from observing long-run effects, we adopt this limitation for two reasons. First, our data begins in 2008 and ends in 2020, allowing for three pre-period years in the first election cycle and two post-period years in the final election cycle. Thus, a five-year window allows us to maintain a similar set of pre and post-years across all election cycles. Second, the assumptions necessary for control counties to serve as reasonable counterfactuals for treated counties increase as the time period we are studying lengthens (e.g., although both sets of counties had sheriff turnovers in a given election cycle, these counties may appear less alike the further one gets from the election).

Our regression specification is as follows:

$$Y_{icey} = \sum_{y=-2}^{2} \beta_y Treat_{ce} \cdot \eta_y + \alpha_{iey} + \delta_{ce} + \varepsilon_{icey}. \tag{2.1}$$

The outcome variable Y_{icye} is equal to 1 if individual i who resides in county c is a registered Republican (Democrat) in year y, which is part of election cycle e. Years within an election cycle take five values -2, -1, 0, 1, 2. The year the election was held is the year 0. The coefficients of interest, $\beta_y, y \in \{-2,0,1,2\}$, are the coefficients on interaction terms between whether an individual works in a treated county and election cycle-year dummy variables ($\eta_y, y \in \{-2,0,1,2\}$). We use the year before the election as the baseline year ($y = -1$). Timing notations in the tables are: $t-2, t-1, t, t+1, t+2$ corresponds to $y = -2, -1, 0, 1, 2$. Treated counties are defined as counties that undergo a transition from a Democratic to Republican sheriff—while we are interested in transitions in both directions, during the time period studied in North Carolina, nearly all interparty turnovers involved Democratic-to-Republican transitions (See Table 2.1). Control counties are counties that underwent a Democratic-to-Democratic sheriff turnover.

The vector α_{iey} includes a vector of the calendar year by 5-year age bin fixed-effects. Age is measured in each calendar year. This vector control for any party affiliation trends correlated with age—these trends are a particular concern because in North Carolina, as in many

47

Southern states, there is a large cohort of older conservative Democrats who have increasingly been changing their voter registration. The vector δ_{ce} includes county by election-cycle fixed effects, and ε_{icey} is an error term. The election cycles overlap with each other. For example, the calendar year 2012 serves as (t+2) for 2010 and (t-2) for the 2014 cycle. We thus make any calendar-year-related fixed effects unique at the election cycle level. For instance, 2012-age 5-10 group would have two dummy variables, one for 2010 and one for the 2014 election cycle. Considering the duplicated observations, we two-way clustered our standard errors at the county level and the unique person-agency-calendar-year level.

Importantly, despite treatment occurring at the county-by-year level, the regression is estimated at the individual-by-year level. We do the analysis at this level in order to control for individual age in our primary specification. We also control for sex-by-calendar-year and job-classification-by-calendar-year fixed effects in our robustness checks. We group jobs into three categories: law enforcement (deputy sheriffs), other law-enforcement-related ones (jailers, protective service workers), and administrative ones. Estimating the regression at the individual level implicitly weights larger departments more. We believe such a weighting scheme is appropriate for three reasons. First, we believe this is the policy-relevant parameter. If smaller counties drive the effect, relatively few individuals would be affected by law enforcement office politicization. Second, by giving more weight to large agencies, we can ensure our results are not driven by the replacement of one or two top deputies at the agency. Replacing these top deputies would have a large impact in small agencies but a negligible impact in larger ones. Third, performing our analysis at the individual level increases our statistical power; small agencies experience a large amount of year-to-year variation in political composition, which can make it difficult to detect the effect of the leader's change. However, we also include robustness checks that show our results are largely robust to weighting each agency identically rather than by the number of deputies they employ.

The identifying assumption in this analysis is the parallel trends assumption. In our setting, the parallel trends assumption says that the average *change* in the outcomes for the

D-to-D and D-to-R sheriff's offices would have been the same in the absence of the sheriff party turnover. The primary outcome we consider is the share of Democrat/Republican deputies. We check the pre-trend assumption by showing the pre-election differences in the average outcomes for deputy sheriffs in the control and treated counties. Note that the year t is considered post-treatment in our setting. Newly elected sheriffs and other newly joined deputy sheriffs are typically sworn in in the December of the election year. Since we observe the working status for each officer at the calendar year level instead of the month level, the personnel outcomes in the year t are affected by the new sheriffs. We also show several "placebo" tests, including testing the effect of the party turnover of sheriffs on the political composition of municipal law enforcement departments and county teachers.

In addition to estimating the impact of the D-to-R turnover on political composition, we attempt to determine what mechanisms drive any observed effects. We consider the share of Democrat and Republican deputies among the officers who join the agency during the election cycle. We also consider the share of officers who exit the agency and switch their party registration, given their party affiliation at the beginning of the calendar year. We describe these analyses in greater detail in Section 1.5.3.

2.3.2   Data

To carry out the analysis above, we need three main sets of data: data on sheriff elections, agency officer rosters, and the political affiliation of agency personnel. We describe each of these sources of data below.

North Carolina Sheriff Data

We create a county-by-year panel of the political party in control of the Sheriff's office for each of North Carolina's 100 counties from 2007 to 2022. For the 2010, 2014, and 2018 election cycles, we gather this information from election results available on the North Carolina State Board of Education website. For the 2006 election cycle, we hand collect these data through

news reports and county board of election websites. Table 2.1 shows the number of counties by election cycle experiencing various types of sheriff turnover (e.g., D-to-R, D-to-D). In total, 24 counties undergo a D-to-R sheriff transition over our sample period.

North Carolina Pension Data

We create a yearly roster of Sheriff's Office officers using information from the North Carolina pension system obtained through a public records request from the North Carolina Department of the Treasury. The data set provided in response to our public records request contains the employer and salary history of the universe of workers in the public sector in North Carolina since 2008, along with information about years of service. Importantly, the pension files are yearly snapshots so we only observe if an officer collects any salary in a given year–we do not observe exact quit dates.[4] We use employer names, retirement systems, employee categories, and job classifications to identify sheriff deputies, police officers, and school teachers in the data.

North Carolina Voter Registration

Voter registration snapshot files are publicly available in North Carolina. They can be accessed from the North Carolina State Board of Elections website [5]. We merge voter registration records from 2008-2020 with individuals in the pension data file. We use voter registration snapshot files at the beginning of each year except for 2008 due to data availability, when we use a file from November. We use information on names, race, ethnicity, gender, age, party affiliation, registration status (active or not), and registration location (county) from the voter registration snapshot files.

Matching between voter snapshot and public pension data.

We use first name, middle name, last name, age, and commuting zone to match voter registration records with public pension records. County information in public pension records is

---

[4]While there is a termination date variable, it does not appear to align with the observed salary histories provided in the data.

[5]Voter snapshot files can be downloaded from this link: https://dl.ncsbe.gov/?prefix=data/Snapshots

derived from employer names. In less than 5% of cases when a unique first name, middle name, last name, age, and commuting zone combination have multiple records in the voting file, we calculate the mean of all relevant variables (e.g., political affiliation, race, sex) and categorize the person with a certain characteristic if the mean is higher than 0.67, and treat the characteristic as missing otherwise. Both public pension records and voter registration snapshot files contain unique IDs for each person across the years. From the matching process, we derive the gender, race, ethnicity, voter registration status, and party affiliation of the public sector workers. In the even rarer case ($< 2\%$ of observations) where a unique first name, middle name, last name, age, and commuting zone combination have multiple records in the pension file we merge each individual to the corresponding voter file record. Our matching rates for sworn officers in sheriff's offices and police departments are respectively 70.5% and 62.1%, respectively.

2.4   Results

2.4.1   Officer Composition

We begin by showing the impact of a D-to-R sheriff turnover on the political affiliations of the office's deputies. Figure 2.1 shows the raw changes in the Democratic and Republican shares of sheriff deputies in treated counties relative to control counties before and after the elections. We also plot the Democratic and Republican share of voters in corresponding counties in dash lines to show the overall voter party affiliation trend. D-to-R counties have a lower baseline share of Democrats (both voters and sheriff deputies) and a higher baseline share of Republicans. The two types of counties exhibit similar trends in the two years before the election year, while D-to-R counties experience a drop in the share of Democratic deputies and a corresponding increase in the share of Republican deputies in the years following the election.

Figure 2.2 plot the estimates and 95% confidence intervals of β_y in a regression with specification 2.1. Similar to the raw data plots, we see a small decrease (increase) in the Democratic (Republican) share of deputies in the election year. In the two years after the

elections, we see large decreases (increases) in the Democratic (Republican) share of deputies on the order of 7-8 percentage points.

Table 2.2 shows the estimation results of specification 2.1. Odd columns show results weighted by the number of deputies, and even columns show results with each agency weighted equally. Panel A (B) shows the effects of the D-to-R turnover on the share of Democratic (Republican) deputies. Column (1)-(2) use our base specification, while Columns (3)-(4) shows the effects when including additional job-classification-by-year fixed effects and sex-by-year fixed effects.

In the year of an election, Democratic deputy shares in D-to-R counties are 2.6 percentage points larger than in D-to-D counties, compared to the baseline differences between the two groups of counties. Republican deputy shares in D-to-R counties are three percentage points lower than in D-to-D counties (Column (1)). Two data features are worth noting here. First, the party affiliation of individuals is measured on January first, so any party switch during year t would not be reflected in the party affiliation share in year t. Second, everyone with a positive salary in the agency in year t is included in the computation of the party affiliation share in year t. These individuals may include deputy sheriffs who left the agency in the latter part of the year and those sworn in with the newly elected sheriff at the end of the year.

Two years after the elections, the election of a Republican sheriff leads to a nine p.p. decline in the share of Democratic deputies and a corresponding increase in the share of Republican deputies. Compared to the baseline party affiliation share (evaluated in the D-to-R counties in $t - 1$), this is a sizable 25% increase (decrease) in the Republican (Democratic) share. Note that the change in the Democrat share is not one-to-one to the Republican share since Unaffiliated officers are also included in the sample.

The magnitudes and standard errors barely change after adding gender-by-calendar year and job-classification-by-calendar-year fixed effects (Column (3)). The point estimates stay roughly the same (sometimes shrink a little bit) when we weigh each agency equally, suggesting that the size of the treatment effects do not vary much on the agency size dimension (Columns

(2) and (4)).

We perform two "placebo" tests to check the validity of our identifying assumption. Table 2.3 uses the same specifications as in Table 2.2, but with two samples of local employees for which we expect there should be little effect of a sheriff turnover: municipal law enforcement officers and teachers. Municipal law enforcement officers are employed by municipal governments and therefore not hired by sheriffs. However, it is possible that the leaders of the police departments, the police chiefs, also experience political party turnover. The municipal council may experience an electoral party turnover and appoint a new police chief whose party affiliation differs from the previous one. Teachers are employed by local boards of education and face no change in the political composition of their leadership from the sheriff's election unless the majority party of school board members also changes in the mid-term elections. Table 2.3 also reports the regression estimates of specification 2.1 in a sample of all registered voters in North Carolina.

Indeed, the share of Democrat (Republican) voters decreases (increases) in the election year t and $t + 1$ more in D-to-R counties than in D-to-D counties, compared to the baseline period $t - 1$. But the magnitude is minimal, about 0.9 (0.6) p.p. in $t + 1$, or 1 (2) % of the baseline mean for Democrat (Republican) shares (Table 2.3, Panel A, Column (1) and (2)). The trend in teacher's party affiliation shares is almost exactly the same as the trend in voters (Table 2.3, Panel A, Column (3), (5)). Turning our attention to police officers, we find that the trend in party affiliation composition of police officers is already different between D-to-D and D-to-R counties two years before the elections. Police officers are 1.7 p.p. more likely to be Democrats in D-to-R counties than in D-to-D counties, compared to the baseline. This hints that factors other than the general voter party affiliation trend drove such pre-trend and suggests that police officers might not be a good candidate for placebo tests. For post-election years, police officers are 2-5 p.p. more (less) likely to be Republicans in D-to-R than in D-to-D counties, compared to the baseline. One interpretation is that this reflects the general party affiliation trend of general law-enforcement officers in those counties. If this is true, then the estimate in Table 2.2 overestimates the impact of the party turnover of sheriffs. Another interpretation is that

the turnover of political leadership in police departments is correlated with the turnover of the sheriff's offices. In this case, police officers are not an appropriate group for placebo tests. In section 2.4.2, by examining the decomposition of the changes in the party affiliation, we provide suggestive evidence that the first interpretation is more likely to be true.

Changes in other characteristics

With the changes in the party affiliation share, a natural next step is to examine the changes in other characteristics of the deputy sheriffs, especially the racial composition. In Appendix Table 2.8, we examine the impact of the sheriff party turnover on the share of male workers and black workers in sheriff's offices, police departments, and schools. We find no impact of sheriff party turnover on the gender composition. We find a marginally significant two p.p. decrease in the share of Black officers two years after the elections. The magnitude is large (22% compared to the baseline mean), but estimate is not precise.

2.4.2 Why did the Political Composition of Affected Sheriff Offices Change?

The political composition of deputies could change after the election of a Republican sheriff through three types of officer actions: exiting the agency, switching party registration, and entering the agency. We test the quantitative importance of each of these channels by performing a decomposition analysis in two groups of officers: incumbents and entrants. Incumbents are officers who received positive salaries in the agency at the beginning of the election cycle ($t-2$). Entrants are officers who started to receive positive salaries in the later period of the cycles ($t-1, t, t+1, t+2$). We examine the exits and party registration switches for the incumbents and entry and party registration switches for the entrants. Specifically, we implement the following:

- **Exit**: To test the effect of exit, we assign each incumbent their voter registration as of two years before the election ($t-2$). We then test the impact of a D-to-R transition on the political composition of the incumbents by estimating a regression with specification 2.1. Because we have eliminated new entrants and registration switching, any detectable effect

54

on composition must be driven by exit alone.

- **Exit + Switch**: To see how much more share changes can be explained by registration switches, on top of the exit channel among incumbents, we assign each incumbent their actual voter registration for a given year. This essentially allows registration switches. As above, we estimate specification 2.1 in the sample of incumbents. We interpret as the impact from registration switches the difference in the estimates between **Exit** and **Exit +** **Switch**.

- **Entry**: To test the effect of entry, we assign each entrant their voter registration as of the year they joined the agency. We then test the impact of a D-to-R transition on the political composition of the entrants, estimating specification 2.1. Implicitly we allow the entrants to exit the agency, so the impact here should be interpreted as the impact through the *net entry* among entrants.

- **Entry + Switch**: Entrants can also change their party registration over the years. With the same sample as in the **Entry** analysis, we assign each entrant their actual voter registration for a given year. We then estimate equation 2.1 again. We interpret as the impact from registration switches among the entrants the difference in the estimates between **Entry** and **Entry + Switch**.

Table 2.4 shows the result of these analyses. Two years after the election, differential exit rates alone lead to a marginally significant decrease of 2 p.p. in the Democratic deputy shares. The impact on Republican shares is a significant 2 p.p increase (Column (1)). Incumbents switch their party registration in response to the sheriff's party turnovers. The changes in Democratic share in the election year (t) jump from -0.005 to a significant -0.02 when we allow the switch channel to work in incumbents. Note that we measured the party registration on January 1st, so party switches in year $t - 1$ would result in party share changes in t. The impact of switches on party composition is more pronounced in years $t + 1$ and $t + 2$. Cumulatively, two years

after the elections, the party registration switches account for a 5-7 p.p. decrease (increase) in Democratic (Republican) deputy shares, a 15-19% change compared to the baseline mean measured in D-to-R counties in $t - 1$.

Turning to entrants (Column (3)), we find that entrants in the election year are 6.7 p.p less (more) likely to be Democrats (Republicans). This shows that the newly elected Republican sheriffs bring in proportionally more Republican deputy sheriffs when they were first worn in at the end of the year. Cumulatively, and disallow entrants to switch their party affiliations, two years after the election, the Democratic (Republican) share in entrants' population is 10 (8) p.p lower (higher) in D-to-R than in D-to-D sheriff's offices, compared to $t - 1$, an 18-38% change. Allowing the entrants to switch their party registration (Column (4)), the magnitudes of the estimate become larger, meaning that among entrants, party registration is also more toward Republican in D-to-R than in D-to-D sheriff's offices. The changes in the magnitudes are much smaller in entrants than in incumbents. Overall, two years after the elections, the Democratic (Republican) share of incumbents (80% of the personnel) in D-to-R sheriff's offices decreased (increased) by 9 (7) p.p. And the Democratic (Republican) share in entrants (20% of the personnel) decreased (increased) by 12 (10) p.p.

Both incumbents and entrants change their party affiliation composition around the elections but through different channels. Incumbents exit the agencies to a limited extent but switch their party registration. Entrants are selected on party affiliation and switch their party registration to a limited extent. Importantly, these results suggest that the large shift in the political composition of the sheriff's offices right before and after the election of a Republican sheriff is only *partially* driven by political patronage.

We conduct the same decomposition analysis on police officers and report the estimates in Appendix Table 2.9. We find some evidence that the increase in the Republican share of police officers is driven by incumbents switching their party affiliations (Panel B, Column (2)). The magnitudes for the entrants are large (Panel B, Column (4)), but not significant. It is hard to distinguish whether the party registration switches come from a general trend of law-enforcement

56

officers changing their party affiliation or officers' response to potential new leadership in police departments. Given that we do not observe strong evidence of differential entry based on party affiliation, we are inclined to interpret the party registration switches as a general trend. This implies that we might overestimate the switch effect in deputy sheriffs.

In this section, we do not distinguish behaviors in Democrat and Republican deputy sheriffs. In the next section, we examine the exit, registration switch, and entry behavior separately for Democrat and Republican deputy sheriffs and pay specific attention to the timing of the change of behaviors.

2.5 Mechanisms

2.5.1 Exit Rates

The previous section demonstrated that differential exit rates by deputy political party are an important contributor to changes in political composition after a D-to-R Sheriff turnover. We now examine changes in exit rates among Democratic and Republican deputies relative to Democrat and Republican deputies in control counties. We use the same regression specification as in equation 2.1, but limit our sample to only deputies of the same political party in treatment and control counties and examine how exit rates vary before and after the election. The party affiliation of each deputy-year is measured on January first, so the composition of Democratic and Republican officers in each agency is different across the years. The outcome variable is a dummy variable indicating exits. The variable is one if the year is the last year we observed a deputy earning a positive salary from an agency.[6] Note that by this definition of exits, an exit in year t would impact the party affiliation composition in year $t + 1$. Since we use six years of personnel records in each election cycle in the composition analysis, we include exits in five years, except the last year in the cycle. We choose $t - 2$ as the baseline period because we want

---

[6]The pension database has a variable that indicates termination date, but it does not always align with observed salary information—for example, sometimes an individual appears to no longer be employed by an agency (i.e., does not draw any salary in the data), but does not have a termination date. Because only one-third of exits have a termination date, we doubt the accuracy of this measure and do not use it in our analysis.

to show whether exit rates in $t-1$ contribute to the deputy sheriff composition changes we observed in t in Figure 2.2.

Figure 2.3 shows the regression estimation results in event study form. Exits of Democratic officers increase in the election year and the year after, while the exit rates of Republican deputies are not affected by the elections. Table 2.5 shows the estimation results. Democratic deputies are roughly five percentage points more likely to quit in the election year and the year after in counties experiencing a D-to-R turnover relative to nearby counties experiencing a D-to-D turnover (Table 2.5, Column (1)). Results are robust to the addition of job classification-by-year and gender-by-year fixed effects. The magnitudes of some coefficients become much larger when each agency receives the same weight (the interaction term with the election year for Democrat deputy sheriffs), indicating that the exit behavior is heterogeneous along the agency size. As discussed in section 2.3, we prefer the results from weighting each individual equally (Table 2.5, Columns (1) and (3)), since such results are less likely to be driven by changes in a few officers in small agencies.

2.5.2 Officer Entry

We examine changes in the share of Democrat and Republican entrants around the elections in D-to-D and D-to-R sheriff's offices. We only include the first year the officer joins the agency in the sample. The outcome variables are whether an entrant is Democrat (Republican) or not. The regression specification is the same as in equation 2.1. By the definition of entrants, we observe entrants since $t-1$. We use $t-1$ as the baseline period.

Figure 2.4 plots the point estimates and the 95% confidence intervals of the regression estimates. Since the election year, Democrats (Republicans) account for a smaller (larger) share of entrants in D-to-R than in D-to-D sheriff's offices. Table 2.6 shows that the decrease in Democrat share is 11-15 p.p., or 42-56% compared to the share of Democrats in D-to-R sheriff's offices in $t-1$. The increase in Republican share is 8-14 p.p., or 19-33% compared to the baseline share. Elected sheriffs seem to shape the personnel by hiring deputy sheriffs who are

politically more aligned with them.

Finally, we examine the racial composition of the entrants in response to the party turnover of sheriffs. In general, the estimates are noisy, but we do find a very large reduction of the Black share of entrants in D-to-R than in D-to-D sheriff's offices. Two years after the election, the Republican sheriffs seem to almost stop hiring Black deputy sheriffs. The coefficient of the interaction term between D-to-R and $t+2$ is -0.10, while the baseline share in D-to-R in $t-1$ is 0.11.

2.5.3   Voter Registration Switching

Table 2.4 showed that the most critical factor driving shifts in political composition among incumbents following a D-to-R sheriff turnover is deputies switching their voter registration. Such switching may occur out of a desire to signal their loyalty to the new sheriff and vote for the sheriff in any primaries or may occur because the election serves as a galvanizing event to switch their registration to a party that better represents their underlying ideology. This latter motive may be particularly predominant in a state like North Carolina that due to its location in the South likely has a large number of older legacy Democrats.

We estimate the timing and magnitude of voter registration switching by estimating equation 2.1. The outcome variable is now "registration switches," which we define as a change in worker's party registration from one political party to another (or to unaffiliated). Since the party affiliation is observed on January first, a switch that happened in t would impact the party composition of deputies in $t+1$.

Figure 2.5 shows the results in event study form. There is an enormous increase in vote switching among Democratic deputies in the election year. The magnitude becomes much smaller and only marginally significant in the year after. Republican deputy sheriffs in D-to-R sheriff's offices do not exhibit any different registration switch behaviors from the deputies in D-to-D sheriff's offices during the election cycle. Table 2.7 shows that the increase in registration switching rate among Democrats is 12 p.p., or 350% compared to the registration switching

rate in D-to-R sheriff's offices in $t - 2$. The magnitudes of the estimate in other years are much smaller and non-significant, increasing confidence that the observed result is not due to different trends across the two groups of counties. There is also no impact of the election on switching among Republican deputies. The results are robust to control job classification-by-year and gender-by-year fixed effects (Column (3)). The magnitudes of the estimates become much smaller when we weight the agencies equally, suggesting that the switching behavior is more pronounced in larger agencies.

Two possibilities could explain these results. First, deputies switch their registration after the election results are known in early November. Second, deputies switch their registration before the election either to curry favor with a candidate ahead in the polls or because they agree with the policy positions the candidate is putting forward. The timing of the observed effect provides one clue for the mechanism that may be driving it. If switching registrations was due to a growing realization that the new sheriff's ideology (and party) better aligned with the deputies' own, we might expect switching to occur over an extended period of time. Instead, most of the switching we observe ended by January after the election, supportive of the hypothesis that deputies switch deputies in response to the expected or actual election of a new Republican sheriff, likely to curry favor or protect their jobs.

2.6   Conclusion

This paper shows that the election of a Republic sheriff appears to lead to a sharp increase in the share of Republican deputies employed by the sheriff's offices and a corresponding decrease in the share of Democratic deputies. This change is driven by two factors—Democratic deputies disproportionately changing their voting registration and Republican deputies disproportionately increasing their likelihood of entering the agency.

These results create several important questions for future research. First, future work can focus on the importance of civil service protection by comparing changes in political composition

after a change in the ideological orientation of a sheriff in jurisdictions with differential levels of civil service protections for law enforcement officers.

Second, we show evidence that a substantial number of Democratic deputies change their voter registration after the election of a Republican Sheriff. However, we are unable to test whether changes in voter registration lead to changes in other behavior such as voting or the manner in which these deputies carry out their law enforcement duties. Future work examining this question will provide useful information about whether registration switches in the face of leadership changes are merely cosmetic or can lead to (or are correlated with) larger changes in individual behavior.

61

2.7 Figures and Tables

(a) Trends in Democrat Share of Deputies and Voters in Treated and Control Counties
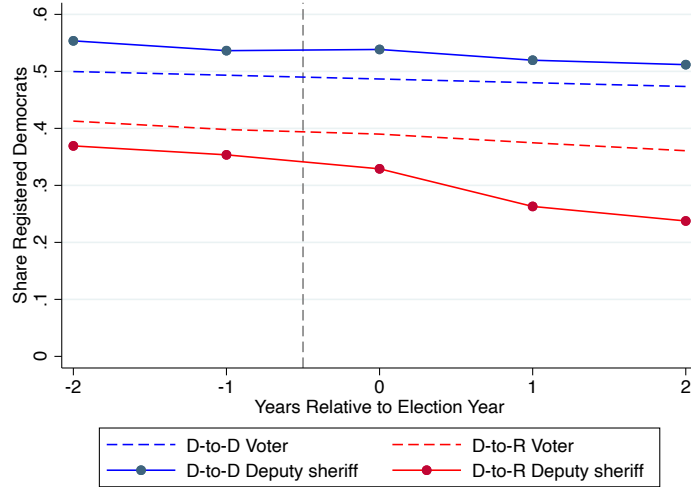
(b) Trends in Republican Share of Deputies and Voters in Treated and Control Counties
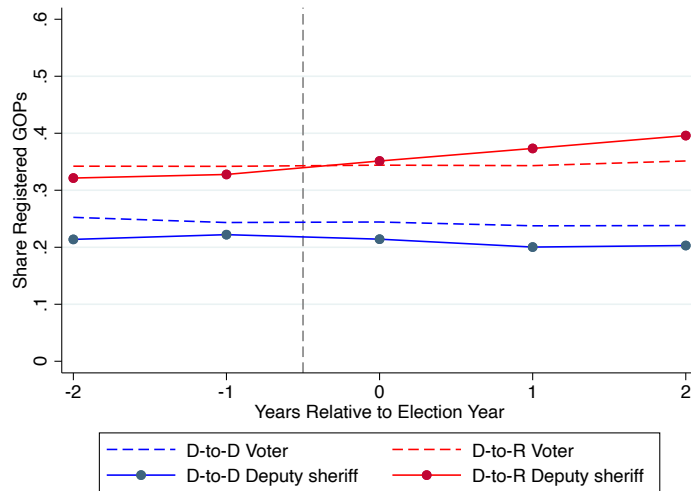
Figure 2.1. Raw Trends in the Share of Democratic and Republican Deputies and Voters in Treated and Control Counties

Notes: These figures show the share of Democratic and Republican deputies and voters in treated and control counties in the three years before and two years after an election. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover.
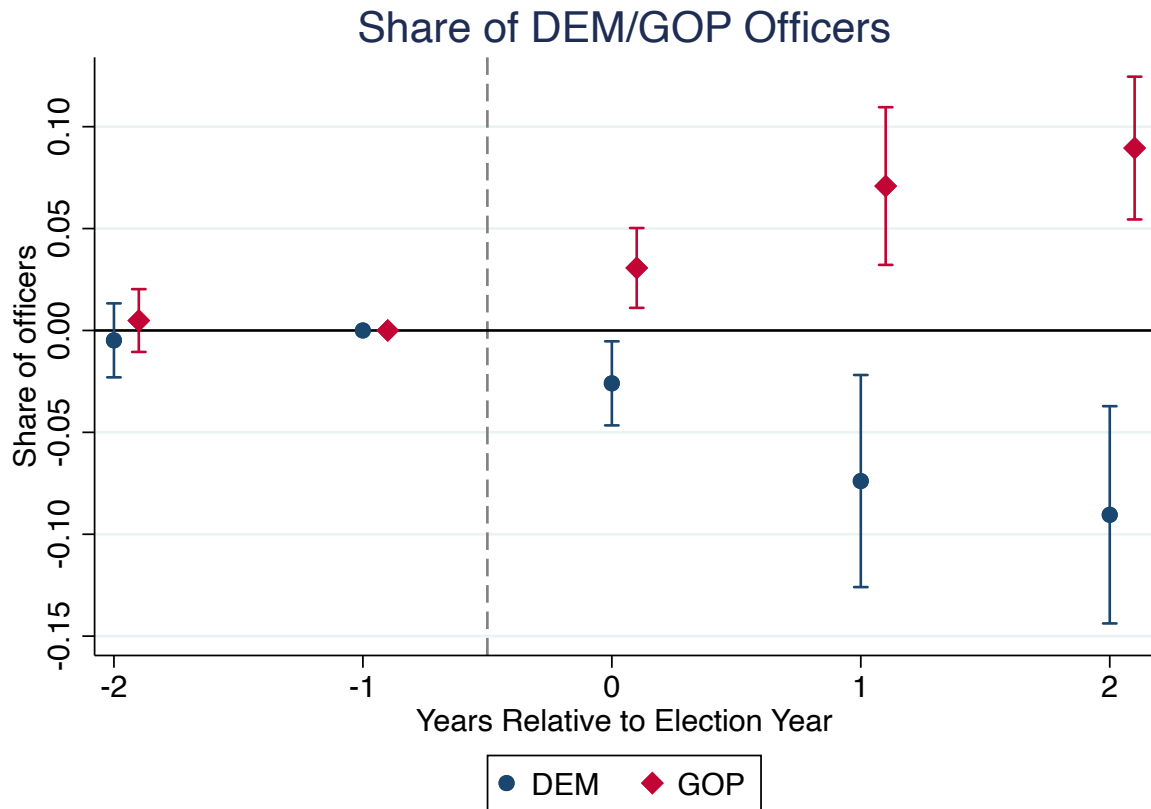
Figure 2.2. Effect of D-to-R Turnover on Democratic and Republican Share of Sheriff Deputies

Notes: This figure shows an event study of the effect of a D-to-R Sheriff turnover on the share of Democratic and Republican Deputies in a given county. Coefficients come from a regression of an indicator variable equal to 1 if a deputy is Democrat (Republican) and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and the cycle-year dummy variables. The specification also includes five year age bins by year fixed-effects and county by election cycle fixed-effects. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover. Standard errors are clustered at the county level. 95% confidence intervals are reported
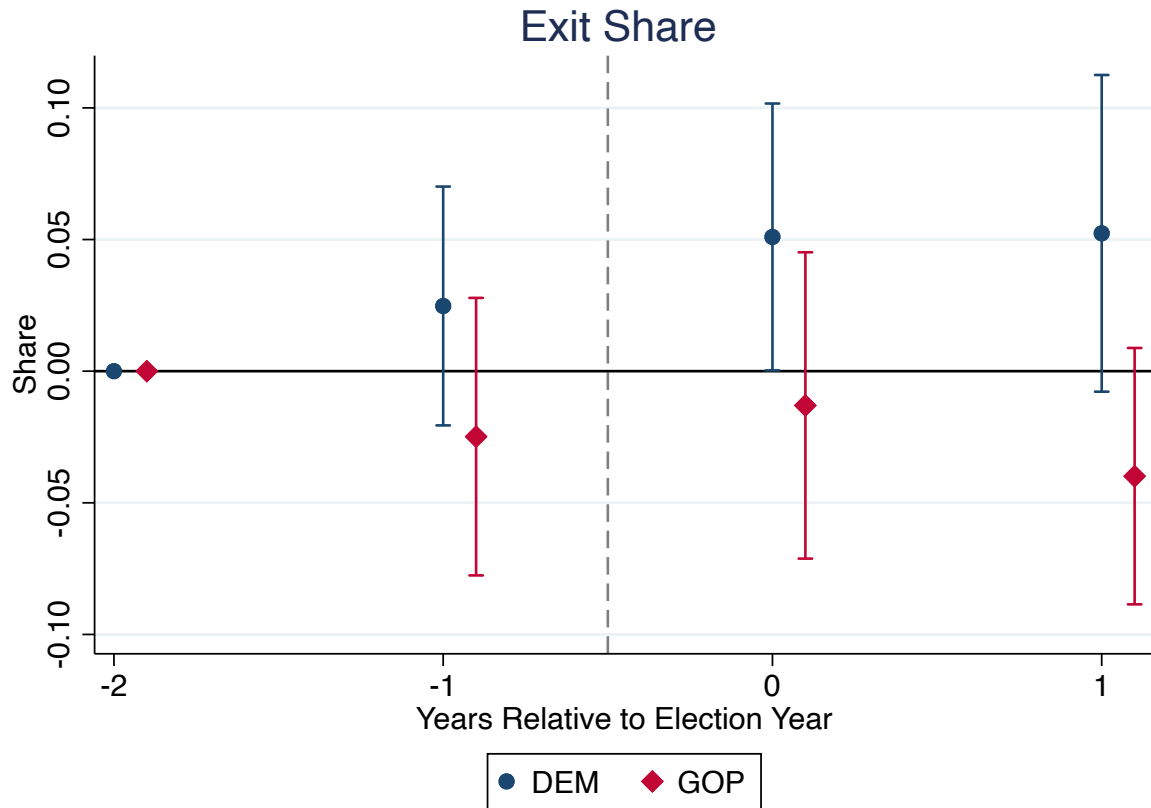
Figure 2.3. Effect of D-to-R Turnover on Sheriff Deputies Exit Share

Notes: This figure show an event study of the effect of a D-to-R Sheriff turnover on the exit rates of Democratic or Republican deputy incumbents in a given county. Incumbents are officers who have already worked for the agency in $t-2$. Coefficients come from a regression of an indicator variable equal to 1 if a deputy left an agency in a given year and 0 otherwise on interactions between an indicator variable for if a deputy had been employed by a treated county two years before the election and election cycle-year dummies. Party registration is measured as of January 1st of a given year. Leaving an agency is defined as having a year of no salary from an agency following a year of positive salary. The specification also includes five-year age bins by year fixed-effects and county by election cycle fixed-effects. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the election cycle. Standard errors are clustered at the county level. 95% confidence intervals are reported
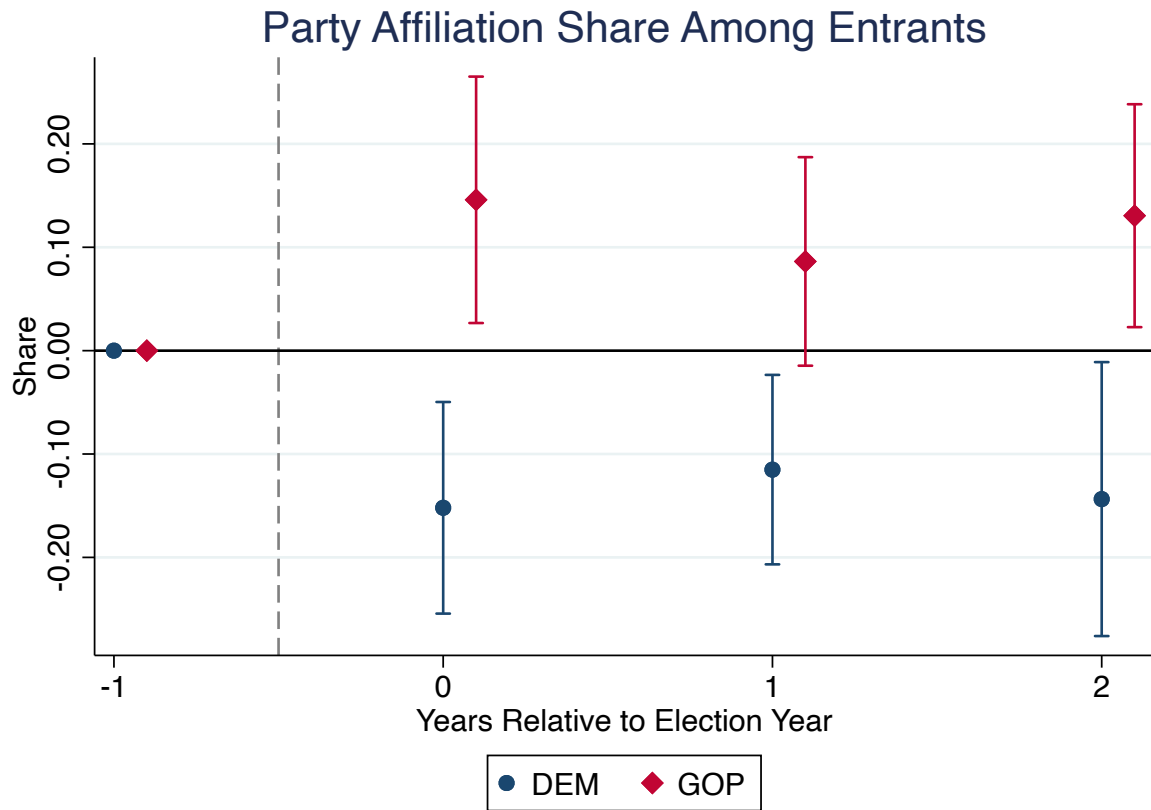
Figure 2.4. Effect of D-to-R Sheriff Turnover on Democratic and Republican Deputy Share in Entrants

Notes: This figure shows the effect of a D-to-R Sheriff turnover on the share of Democratic or Republican deputy entrants. Coefficients come from a regression of an indicator variable equal to 1 if an entrant deputy who enters the agency in that year and is a Democrat (Republican) and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and election cycle-year dummies. Party registration is measured as of January 1st of a given year. The specification also includes five year age bins by year fixed-effects and county by election cycle fixed-effects. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. Standard errors are clustered at the county level. 95% confidence intervals are reported
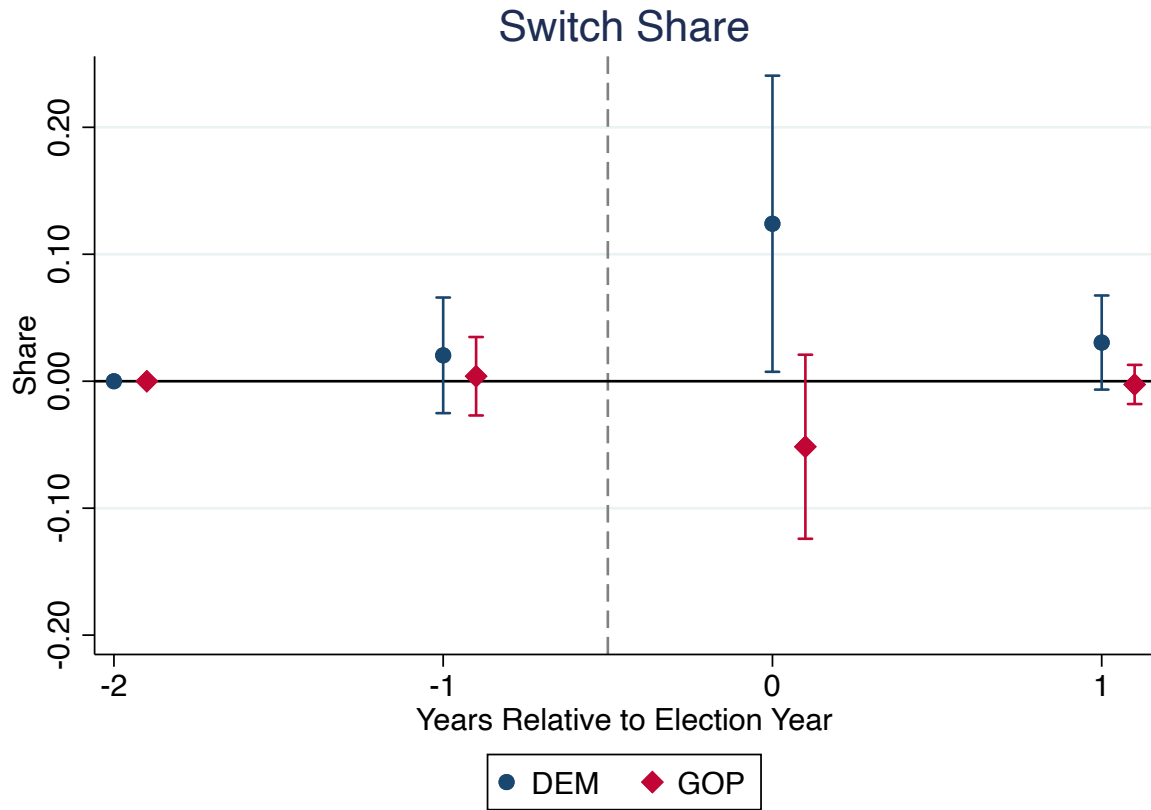
Figure 2.5. Effect of D-to-R Sheriff Turnover on Democratic and Republican Deputy Party Registration Switches

Notes: This figure shows an event study of the effect of a D-to-R Sheriff turnover on the party switch rates of Democratic or Republican deputy incumbents in a given county. Incumbents are officers who have already worked for the agency in $t-2$. Coefficients come from a regression of an indicator variable equal to 1 if a deputy switched their party registration in a given year and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and election cycle-year dummies. Party registration is measured as of January 1st of a given year. The specification also includes five year age bins by year fixed-effects and county by election cycle fixed-effects. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. Standard errors are clustered at the county level. 95% confidence intervals are reported

Table 2.1. Sheriff Election Results in North Carolina

All Sheriffs' Offices						
Election Year	R to R Turnover	R to R No Turnover	R to D	D to D Turnover	D to D No Turnover	D to R
2010	10	24	0	15	46	5
2014	5	33	1	14	37	10
2018	13	32	3	16	27	9

Notes: D refers to the Democratic party, and R refers to the Republican party. North Carolina has 100 sheriff's offices, one for one county. This table presents the party turnover distributions in all elections from 2010 to 2018.

Table 2.2. Effect of D-to-R Sheriff Turnover on Democrat and Republican Share of Sheriff Deputies

	(1)	(2)	3)	(4)
Panel A		Democratic Deputy or not		
t-2 x DtoR	-0.00485	-0.00327	-0.0108	-0.00606
	(0.00907)	(0.00760)	(0.00987)	(0.00807)
t x DtoR	-0.0260**	-0.0111	-0.0262**	-0.0117
	(0.0103)	(0.0108)	(0.0104)	(0.0113)
t+1 x DtoR	-0.0739***	-0.0593**	-0.0742***	-0.0600**
	(0.0260)	(0.0243)	(0.0272)	(0.0254)
t+2 x DtoR	-0.0905***	-0.0685***	-0.0897***	-0.0671***
	(0.0266)	(0.0234)	(0.0268)	(0.0248)
Pre Treat. Dep. Var mean	0.354	0.354	0.354	0.354
Adjusted R^2	0.146	0.218	0.165	0.221
N	37596	37596	37596	37596
Panel B		Republican Deputy or not		
t-2 x DtoR	0.00486	-0.00100	0.00837	0.00196
	(0.00769)	(0.0102)	(0.00755)	(0.00936)
t x DtoR	0.0307***	0.0176	0.0311***	0.0207*
	(0.00978)	(0.0124)	(0.0100)	(0.0121)
t+1 x DtoR	0.0709***	0.0737***	0.0731***	0.0766***
	(0.0193)	(0.0223)	(0.0197)	(0.0228)
t+2 x DtoR	0.0895***	0.100***	0.0919***	0.104***
	(0.0175)	(0.0207)	(0.0172)	(0.0217)
Pre Treat. Dep. Var Mean	0.328	0.328	0.328	0.328
Adjusted R^2	0.064	0.107	0.077	0.111
N	37596	37596	37596	37596
Control	Base	Base	Base+	Base+
Weight	Individual	Agency	Individual	Agency

Notes: Coefficients come from a regression of an indicator variable equal to 1 if a deputy is Democrat (Republican) and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and election cyle-year dummy variables. The specification with the baseline control includes five-year age bins by year fixed-effects and county by election cycle fixed-effects. The Base+ additionally includes gender by year and job classification by year dummy variables. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. Statistical significance is denoted: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the county level and unique observation level (person-agency-calendar year)). Pre Treat. Dep. Var Mean is evaluated at the D-to-R counties in the year before the elections $(t-1)$. Observations are given equal weights (Individual) or weights that is inverse to the number of individuals in that agency such that each agency has equal weights (Agency).

Table 2.3. Effect of D-to-R Sheriff Turnover on Democrat and Republican Share of School Employees and Police Officers

	(1)	(2)	3)	(4)	5)	(6)
Panel A	Voters			Teachers		
	DEM	GOP	DEM	DEM	GOP	GOP
t-2 x DtoR	0.00552	-0.00548*	0.00677**	-0.00510	-0.00412**	0.00135
	(0.00443)	(0.00301)	(0.00285)	(0.0107)	(0.00181)	(0.00974)
t x DtoR	-0.00204*	0.00213**	-0.000591	-0.00515	0.00294	0.0128**
	(0.00107)	(0.00102)	(0.00222)	(0.00500)	(0.00227)	(0.00486)
t+1 x DtoR	-0.00941**	0.00684***	-0.00950*	-0.0143*	0.00693**	0.0191**
	(0.00377)	(0.00227)	(0.00527)	(0.00791)	(0.00285)	(0.00840)
t+2 x DtoR	-0.0146**	0.0132***	-0.0151**	-0.0301**	0.00954**	0.0252**
	(0.00590)	(0.00425)	(0.00718)	(0.0139)	(0.00420)	(0.0114)
Pre Treat. Dep. Var mean	0.398	0.342	0.418	0.418	0.336	0.336
Adjusted R^2	0.903	0.873	0.759	0.546	0.700	0.464
N	6585	6585	5589	5589	5589	5589
Panel B			Municipal Police officers			
			DEM	DEM	GOP	GOP
t-2 x DtoR			0.0169***	0.0106	-0.0125	0.0160
			(0.00605)	(0.0140)	(0.0106)	(0.0142)
t x DtoR			0.00561	0.0126	0.0158	0.0202
			(0.00680)	(0.0244)	(0.0106)	(0.0261)
t+1 x DtoR			-0.0217*	-0.00626	0.0359**	0.0121
			(0.0109)	(0.0332)	(0.0143)	(0.0325)
t+2 x DtoR			-0.0258*	-0.0187	0.0515***	0.0614
			(0.0147)	(0.0361)	(0.0154)	(0.0376)
Pre Treat. Dep. Var mean			0.255	0.255	0.434	0.434
Adjusted R^2			0.063	0.212	0.039	0.118
N			44186	44186	44186	44186
Weight	Individual		Individual	Agency	Individual	Agency

Notes: Coefficients come from a regression of an indicator variable equal to 1 if a deputy is Democrat (Republican) and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and election cycle-year dummy variables. The specification also includes five-year age bins by year fixed-effects and county by election cycle fixed-effects. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. The municipal law enforcement sample consists of all municipal law enforcement officers employed by jurisdictions within treated and control counties. Treatment is assigned as above based on the county in which a municipality is located. Similarly, the teacher sample consists of all teachers employed by jurisdictions within treated and control counties. Again, treatment is assigned as above based on the county in which a jurisdiction is located. Standard errors are clustered at the county and unique observation level (person-agency-calendar year)). Statistical significance is denoted: *** p<0.01, ** p<0.05, * p<0.1. Pre Treat. Dep. Var mean is evaluated at the D-to-R counties in the year before the elections ($t-1$). Observations are given equal weights (Individual) or weights that is inverse to the number of individuals in that agency such that each agency has equal weights (Agency).

Table 2.4. Effect of D-to-R Sheriff Turnover on Democrat and Republican Share of Sheriff Deputies: Channel Analysis

	(1)	(2)	3)	(4)	(5)
Panel A			Democrat or not		
		Incumbents		Entrants	All
	Exit	Exit + Switch	Entry	Entry + Switch	
t-2 x DtoR	-0.00529	0.000371			-0.00485
	(0.00389)	(0.00600)			(0.00907)
t x DtoR	-0.00594	-0.0227**	-0.0677**	-0.0680**	-0.0260**
	(0.00514)	(0.00944)	(0.0295)	(0.0294)	(0.0103)
t+1 x DtoR	-0.0126	-0.0762**	-0.0847***	-0.104***	-0.0739***
	(0.00810)	(0.0305)	(0.0297)	(0.0300)	(0.0260)
t+2 x DtoR	-0.0208*	-0.0916***	-0.103***	-0.128***	-0.0905***
	(0.0113)	(0.0315)	(0.0360)	(0.0369)	(0.0266)
Pre Treat. Dep. Var mean	0.377	0.365	0.268	0.268	0.354
Adjusted R^2	0.137	0.150	0.135	0.138	0.146
N	30537	30537	7054	7054	37596
Panel B			Republican or not		
		Incumbents		Entrants	All
	Exit	Exit + Switch	Entry	Entry + Switch	
t-2 x DtoR	0.00950*	0.00817			0.00486
	(0.00518)	(0.00643)			(0.00769)
t x DtoR	0.00234	0.0127	0.0675**	0.0807***	0.0307***
	(0.00415)	(0.00939)	(0.0310)	(0.0301)	(0.00978)
t+1 x DtoR	0.0141*	0.0551**	0.0793**	0.0850**	0.0709***
	(0.00735)	(0.0218)	(0.0346)	(0.0353)	(0.0193)
t+2 x DtoR	0.0205**	0.0706***	0.0844**	0.101***	0.0895***
	(0.00867)	(0.0199)	(0.0356)	(0.0336)	(0.0175)
Pre Treat. Dep. Var mean	0.311	0.313	0.434	0.434	0.328
Adjusted R^2	0.055	0.060	0.102	0.104	0.064
N	30537	30537	7054	7054	37596

Notes: Coefficients come from a regression of an indicator variable equal to 1 if a deputy is Democrat (Republican) and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and election cycle-year dummy variables. Treated counties are defined as counties undergoing a D-to-R sheriff turnover, and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. The Incumbents are the officers who were employed in the agency in year $t-2$. The Entrants are the officers who started being employed in the agency after $t-2$. The "Exit" specification includes only incumbents that are active in a given year and uses their registration from $t-2$. The "Exit + Switch" specification includes only officers present in year $t-2$ that are active in a given year and uses the officer's actual voter registration in a given year. The "Entry" specification only includes the entrants who are active in a given year and uses their registration from the first year they were observed working in that agency in the cycle. The "Entry + Switch" specification includes the entrants who are active in a given year and the officer's actual voter registration in a given year. Since we only consider entrants since $t-1$, there is no interaction term for $t-2$ and the omitted baseline is still $t-1$. Standard errors are clustered at the county and unique observation level (person-agency-calendar year)). Statistical significance is denoted: *** p<0.01, ** p<0.05, * p<0.1. Pre Treat. Dep. Var mean is evaluated at the D-to-R counties in the year before the elections ($t-1$). Observations are given equal weights (Individual) or weights inverse to the number of individuals in that agency such that each agency has equal weights (Agency).

Table 2.5. Effect of D-to-R Sheriff Turnover on Democrat and Republican Deputies Exit Rates

	(1)	(2)	3)	(4)
Panel A	Democrat Deputy Sheriff			
t-1 x DtoR	0.0248	0.0501	0.0260	0.0465
	(0.0226)	(0.0344)	(0.0224)	(0.0333)
t x DtoR	0.0510**	0.106***	0.0480*	0.0993***
	(0.0253)	(0.0391)	(0.0269)	(0.0371)
t+1 x DtoR	0.0524*	0.0721*	0.0539*	0.0671**
	(0.0300)	(0.0362)	(0.0276)	(0.0312)
Pre Treat. Dep. Var mean	0.082	0.082	0.082	0.082
Adjusted R^2	0.045	0.120	0.048	0.135
N	12989	12989	12989	12989
Panel B	GOP Deputy Sheriff			
t-1 x DtoR	-0.0249	0.0233	-0.0260	0.0134
	(0.0263)	(0.0465)	(0.0278)	(0.0445)
t x DtoR	-0.0130	-0.00832	-0.0109	0.00139
	(0.0290)	(0.0637)	(0.0307)	(0.0633)
t+1 x DtoR	-0.0399	-0.0546	-0.0419*	-0.0477
	(0.0243)	(0.0495)	(0.0238)	(0.0486)
Pre Treat. Dep. Var mean	0.097	0.097	0.097	0.097
Adjusted R^2	0.035	0.181	0.038	0.188
N	5905	5905	5905	5905
Control	Base	Base	Base+	Base+
Weight	Individual	Agency	Individual	Agency

Notes: Coefficients come from a regression of an indicator variable equal to 1 if a deputy left an agency in a given year and 0 otherwise on interactions between an indicator variable for if a deputy had been employed by a treated county two years before the election and election cycle-year dummy variables. Leaving an agency is defined as having a year of no salary from an agency following a year of positive salary. The specification with the baseline control includes five-year age bins by year fixed-effects and county by election cycle fixed-effects. The Base+ additionally includes gender by year and job classification by year dummy variables. Party registration is measured as of January 1st of a given year. Unregistered deputies are dropped from the sample. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. The omitted period is $t-2$. This choice is to respect the timing of the exits. An exit in year $t-2$ causes changes in party affiliation deputy share changes in year $t-1$. Since we find significant share changes since year t, we aim to detect exits since year $t-1$. Standard errors are clustered at the county level. Statistical significance is denoted: *** p<0.01, ** p<0.05, * p<0.1. Pre Treat. Dep. Var Mean is evaluated at the D-to-R counties in the year before the elections ($t-1$). Observations are given equal weights (Individual) or weights inverse to the number of individuals in that agency such that each agency has equal weights (Agency).

Table 2.6. Effect of D-to-R Sheriff Turnover on Democrat and Republican Deputies Likelihood of Entering Sheriff's Offices

	(1)	(2)	3)	(4)
Panel A	Democrat Deputy Sheriff Share			
t x DtoR	-0.152***	0.0111	-0.159***	0.00957
	(0.0511)	(0.0759)	(0.0534)	(0.0782)
t+1 x DtoR	-0.115**	-0.0137	-0.114**	-0.0224
	(0.0458)	(0.0541)	(0.0498)	(0.0549)
t+2 x DtoR	-0.144**	0.0319	-0.156**	0.0184
	(0.0662)	(0.0793)	(0.0676)	(0.0759)
Pre Treat. Dep. Var mean	0.268	0.268	0.268	0.268
Adjusted R^2	0.128	0.174	0.139	0.183
N	3235	3235	3232	3232
Panel B	GOP Deputy Sheriff Share			
t x DtoR	0.146**	0.120*	0.161***	0.137**
	(0.0595)	(0.0663)	(0.0571)	(0.0647)
t+1 x DtoR	0.0863*	0.0589	0.0809	0.0499
	(0.0504)	(0.0567)	(0.0489)	(0.0553)
t+2 x DtoR	0.131**	0.0912	0.128**	0.0867
	(0.0538)	(0.0684)	(0.0544)	(0.0694)
Pre Treat. Dep. Var mean	0.434	0.434	0.434	0.434
Adjusted R^2	0.092	0.121	0.112	0.141
N	3235	3235	3232	3232
Panel C	Black Deputy Sheriff Share			
t x DtoR	-0.0638	-0.0152	-0.0702	-0.0139
	(0.0449)	(0.0465)	(0.0430)	(0.0457)
t+1 x DtoR	-0.0651	-0.0387	-0.0725	-0.0427
	(0.0459)	(0.0461)	(0.0478)	(0.0472)
t+2 x DtoR	-0.108**	-0.00513	-0.133**	-0.0219
	(0.0485)	(0.0522)	(0.0528)	(0.0512)
Pre Treat. Dep. Var mean	0.111	0.111	0.111	0.111
Adjusted R^2	0.255	0.218	0.280	0.235
N	3235	3235	3232	3232
Control	Base	Base	Base+	Base+
Weight	Individual	Agency	Individual	Agency

Notes: This table shows the effects of a D-to-R Sheriff turnover on the share of Democratic or Republican deputies among the entrants in a specific year. The entrants are the officers who started being employed in the agency after $t-2$. We only include in the sample the first year the worker joined the agency. Coefficients come from a regression of an indicator variable equal to 1 if a deputy is a Democrat (Republican) in the first year they join the agency and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and election cycle-year dummy variables. The specification with the baseline control includes five-year age bins by year fixed-effects and county by election cycle fixed-effects. The Base+ additionally includes gender by year and job classification by year dummy variables. Party registration is measured as of January 1st of a given year. Treated counties are defined as counties undergoing a D-to-R sheriff turnover, and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle as a treated county. Since we only consider entrants since $t-1$, there is no interaction term for $t-2$, and the omitted baseline is $t-1$. Standard errors are clustered at the county level. Statistical significance is denoted: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Pretreament dependent variable mean is evaluated at the D-to-R counties in $t-1$. Observations are given equal weights (Individual) or weights inverse to the number of individuals in that agency such that each agency has equal weights (Agency).

Table 2.7. Effect of D-to-R Sheriff Turnover on Democrat and Republican Deputies Likelihood of Switching Voter Registration

	(1)	(2)	3)	(4)
Panel A		Democrat Deputy Sheriff		
t-1 x DtoR	0.0204	0.00573	0.0217	0.00561
	(0.0227)	(0.0176)	(0.0227)	(0.0174)
t x DtoR	0.124**	0.0678	0.128**	0.0693
	(0.0582)	(0.0462)	(0.0592)	(0.0470)
t+1 x DtoR	0.0305	0.0198	0.0317	0.0207
	(0.0185)	(0.0172)	(0.0191)	(0.0176)
Pre Treat. Dep. Var Mean	0.034	0.034	0.034	0.034
Adjusted R^2	0.103	0.111	0.104	0.114
N	12989	12989	12989	12989
Panel B		GOP Deputy Sheriff		
t-1 x DtoR	0.00390	0.000165	0.00651	0.00457
	(0.0154)	(0.0199)	(0.0160)	(0.0214)
t x DtoR	-0.0516	-0.109***	-0.0523	-0.104***
	(0.0362)	(0.0359)	(0.0350)	(0.0378)
t+1 x DtoR	-0.00256	-0.00437	-0.00109	-0.00236
	(0.00770)	(0.0102)	(0.00812)	(0.0114)
Pre Treat. Dep. Var mean	0.013	0.013	0.013	0.013
Adjusted R^2	0.075	0.146	0.075	0.145
N	5905	5905	5905	5905
Control	Base	Base	Base+	Base+
Weight	Individual	Agency	Individual	Agency

Notes: This table shows the effects of a D-to-R Sheriff turnover on the party registration switch rates of Democratic or Republican deputies. Coefficients come from a regression of an indicator variable equal to 1 if a deputy switched their party registration in a given year and 0 otherwise on interactions between an indicator variable for if a deputy is employed by a treated county in a given year and election cycle-year dummy variables. The specification with the baseline control includes five-year age bins by year fixed-effects and county by election cycle fixed-effects. The Base+ additionally includes gender by year and job classification by year dummy variables. Party registration is measured as of January 1st of a given year. The omitted period is $t-2$. This choice is to respect the timing of the party switch. A switch in year $t-2$ causes changes in party affiliation deputy share changes in year $t-1$. Since we find significant share changes since year t, we aim to detect party switches since $t-1$. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle as a treated county. Standard errors are clustered at the county level. Statistical significance is denoted: *** p<0.01, ** p<0.05, * p<0.1. Pre-treament dependent variable mean is evaluated at the D-to-R counties in two year before the elections ($t-2$). Observations are given equal weights (Individual) or weights inverse to the number of individuals in that agency such that each agency has equal weights (Agency).

2.8 Appendix Tables

Table 2.8. Effect of D-to-R Sheriff Turnover on Male and Black Share of Sheriff Deputies, Police Officers, and School Employees

	(1)	(2)	3)
	Deputy Sheriffs	Police Officers	Teachers
Panel A		Male person or not	
t-2 x DtoR	-0.00105	-0.00138	-0.00358**
	(0.00390)	(0.00492)	(0.00152)
t x DtoR	0.00688	0.000340	0.000521
	(0.00653)	(0.00430)	(0.00138)
t+1 x DtoR	0.00259	0.00487	-0.00289*
	(0.00893)	(0.00590)	(0.00169)
t+2 x DtoR	0.00229	0.00448	-0.00240
	(0.0112)	(0.00723)	(0.00196)
Pre Treat. Dep. Var mean	0.835	0.887	0.213
Adjusted R^2	0.050	0.011	0.218
N	37596	44186	5589
Panel B		Black person or not	
t-2 x DtoR	-0.00757**	-0.00373	0.000153
	(0.00366)	(0.00317)	(0.00153)
t x DtoR	-0.00488	0.000501	-0.000226
	(0.00620)	(0.00488)	(0.00197)
t+1 x DtoR	-0.0138	-0.000358	-0.00238
	(0.00861)	(0.00758)	(0.00365)
t+2 x DtoR	-0.0225*	0.00658	-0.00227
	(0.0117)	(0.0101)	(0.00464)
Pre Treat. Dep. Var mean	0.106	0.086	0.152
Adjusted R^2	0.256	0.042	0.834
N	37596	44186	5589

Notes: Coefficients come from a regression of an indicator variable equal to 1 if a worker is male (black) and 0 otherwise on interactions between an indicator variable for if a worker is employed by a treated county in a given year and election cyle-year dummy variables. The specification also includes five year age bins by year fixed-effects and county by election cycle fixed-effects. Treated counties are defined as counties undergoing a D-to-R sheriff turnover and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. The police officer sample consists of all municipal law enforcement officers employed by jurisdictions within treated and control counties. Treatment is assigned as above based on the county in which a municipality is located. Similarly, the teacher sample consists of all teachers employed by jurisdictions within treated and control counties. Again, treatment is assigned as above based on the county in which a jurisdiction is located. Standard errors are clustered at the county and unique observation level (person-agency-calendar year). Statistical significance is denoted: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Pre Treat. Dep. Var mean is evaluated at the D-to-R counties in the year before the elections ($t-1$).

Table 2.9. Effect of D-to-R Sheriff Turnover on Democrat and Republican Share of Police Departments: Channel Analysis

	(1)	(2)	3)	(4)	(5)
Panel A			Democrat or not		
	Incumbents		Entrants		All
	Exit	Exit + Switch	Entry	Entry + Switch	
t-2 x DtoR	0.00676	0.00966			0.0169***
	(0.00615)	(0.00751)			(0.00605)
t x DtoR	0.00924	0.00963	0.0356	0.0232	0.00561
	(0.00628)	(0.00610)	(0.0300)	(0.0337)	(0.00680)
t+1 x DtoR	0.00647	-0.0134	0.00920	-0.00522	-0.0217*
	(0.00883)	(0.00886)	(0.0376)	(0.0388)	(0.0109)
t+2 x DtoR	0.00944	-0.0199*	0.0155	0.00526	-0.0258*
	(0.00972)	(0.00997)	(0.0392)	(0.0416)	(0.0147)
Pre Treat. Dep. Var mean	0.274	0.265	0.194	0.194	0.255
Adjusted R^2	0.068	0.069	0.065	0.069	0.063
N	37890	37890	6281	6281	44186
Panel B			Republican or not		
	Incumbents		Entrants		All
	Exit	Exit + Switch	Entry	Entry + Switch	
t-2 x DtoR	-0.00142	-0.00917			-0.0125
	(0.00734)	(0.0115)			(0.0106)
t x DtoR	-0.00111	0.00231	0.0241	0.0337	0.0158
	(0.00754)	(0.00918)	(0.0400)	(0.0393)	(0.0106)
t+1 x DtoR	-0.00357	0.0173	0.0296	0.0450	0.0359**
	(0.0103)	(0.0160)	(0.0476)	(0.0493)	(0.0143)
t+2 x DtoR	0.00422	0.0367**	0.0383	0.0423	0.0515***
	(0.0124)	(0.0174)	(0.0472)	(0.0467)	(0.0154)
Pre Treat. Dep. Var mean	0.427	0.437	0.417	0.417	0.434
Adjusted R^2	0.042	0.042	0.034	0.036	0.039
N	37890	37890	6281	6281	44186

Notes: Coefficients come from a regression of an indicator variable equal to 1 if a police officer is Democrat (Republican) and 0 otherwise on interactions between an indicator variable for if police departments employ a police officer in a treated county in a given year and election cycle-year dummy variables. Treated counties are defined as counties undergoing a D-to-R sheriff turnover, and control counties are defined as counties undergoing a D-to-D sheriff turnover in the same election cycle. The Incumbents are the officers employed in the police departments in year $t-2$. The "Exit" specification includes only incumbents that are active in a given year and uses their registration from $t-2$. The "Exit + Switch" specification includes only officers present in year $t-2$ that are active in a given year and uses the officer's actual voter registration in a given year. The Entrants are the officers who started being employed in the agency after $t-2$. The "Entry" specification only includes the entrants who are active in a given year and uses their registration from the first year they were observed working in that agency in the cycle. The "Entry + Switch" specification includes the entrants who are active in a given year and the officer's actual voter registration in a given year. Since we only consider entrants since $t-1$, there is no interaction term for $t-2$, and the omitted baseline is still $t-1$. Standard errors are clustered at the county and unique observation level (person-agency-calendar year). Statistical significance is denoted: *** p<0.01, ** p<0.05, * p<0.1. Pre Treat. Dep. Var mean is evaluated at the D-to-R counties in the year before the elections ($t-1$).

Chapter 3

Curriculum and National Identity: Evidence from the 1997 Curriculum Reform in Taiwan

3.1 Introduction

The more homogeneous the people, the easier it is to manage a nation. As a result, state leaders are incentivized to use the education system as an instrument for cultivating national identity—an essential step toward nation-building (Aghion et al., 2018).[1] The extensive literature on the theory of nation-building in economics and political science suggests that governments can homogenize their people through education (Weber, 1976; Billig, 1995; Anderson, 2006; Besley and Persson, 2010; Alesina and Reich, 2015). However, the causal effects underlying the intuition and the transmission mechanism behind the effect of education on national identity formation lack detailed scrutiny. National identity trends in society, and cohort effects, pose challenges to the identification of educational content effects. Specifically, these two effects interact with each other in the sense that students in different cohorts experience societal development and political events in different ways.

In this paper, we overcome these difficulties by exploiting a junior high school curriculum

---

[1]Empirical evidence has shown that, in the past 150 years, investments in mass education by governments have appeared in response to military threats, when patriotic people are required to prepare for future wars (Aghion et al., 2018).

reform in Taiwan. In September 1997, the Taiwanese government published its *Knowing Taiwan* series of textbooks for social subjects, namely, History, Geography, and Society. The History curriculum, in particular, adopted a new perspective on the nation's past and provided abundant Taiwan-related content, all of which had been absent from previous textbooks. The education system in Taiwan mandates that children born after 1 September must enter the education system the following year, such that people born in September will enter later than those born in August. This means that those born in September 1984 (i.e. 13 years old in 1997) would have been the first month's cohort to have studied the new textbooks (i.e. *Knowing Taiwan* series), while those born in August 1984 would have studied the old ones.

These institutional features give us a unique opportunity to identify the causal effects of the junior high school curriculum (i.e. textbook content) on people's national identity in later life, since those born either side of the cut-off would have experienced similar social events and political developments. In addition, the birth timing decisions of parents should be predetermined, which is unlikely to be affected by this reform. Therefore, we can isolate curriculum effect from other confounding factors by comparing the national identities of those born just before and just after September 1984, using a regression discontinuity design. Due to their historic, cultural, and political connections to China, the peoples of Taiwan are confused about their national identity (Jacobs and Kang, 2017). This "identity confusion" or "national identity conflict" means that some identify as Taiwanese whilst others identify themselves as Chinese—or a combination thereof (Jacobs and Kang, 2017).[2] We measure national identity by using a self-reported identity question from a repeated nationally representative survey—the Taiwan Social Change Survey—which has consistently asked respondents about their national identity through the question "Do you consider yourself Taiwanese, Chinese, or both?" over a

---

[2]Taiwan has been governed by several political regimes, such as Netherlands and Spain (1622-1661), Kingdom of Tungning (1661-1683), Qing Dynasty (1683-1895), Japan (1895-1945), and Republic of China (ROC, 1945-present). Therefore, they have no clear consensus regarding national identity. In 1949, Kuomintang (the ruling party of the ROC) lost the civil war to the Chinese Communist Party, and as a result it retreated to Taiwan and took around two million people from China to Taiwan. Since the president of the ROC, Chiang Kai-shek, intended to eventually retake control of mainland China, the ROC government attempted to "sinify" the people of Taiwan by implementing a school curriculum that would cultivate people's Chinese national identity.

long period of time.

We obtain three key findings from our research. First, our results suggest that students who studied the new textbooks are more likely to report themselves as Taiwanese than those who read old textbooks when they were around the age of 20 (18 to 23 years old). The magnitude of the effect is 18 percentage points, which accounts for a 30% increase in the control group mean of 61%. Based on our estimates, we can calculate the persuasion rate, using the formula employed to deduce the persuasive effects of media communications (DellaVigna and Gentzkow, 2010; DellaVigna and Kaplan, 2007). The estimated persuasion rate is 46%, which is much larger than the estimates (i.e., around 2% to 20%) for news media (Enikolopov et al., 2011; Gentzkow, 2006; DellaVigna and Kaplan, 2007; DellaVigna et al., 2014; Yanagizawa-Drott, 2014; Adena et al., 2015; Blouin and Mukand, 2019; Chiang and Knight, 2011; Gerber et al., 2009; Gentzkow et al., 2009).[3] This result should be reasonable, since the intensity of exposure is greater for educational content than for media. Especially, students spend substantial time and effort (i.e. three years) on reading these textbooks to prepare for high school entrance exams. In fact, our result is consistent with Cantoni et al. (2017), suggesting the persuasive effects of the school curriculum are quite significant.

Second, we investigate the possible mechanisms through which school curricula can affect an individual's national identity. Our subgroup analysis suggests that these curriculum effects only appear in academic track students, who generally put more effort into studying textbook materials. This result implies that memorizing and synthesizing textbook content is a possible channel of curriculum effect. In addition, we find that the new curriculum has greater impacts on individuals living in neighborhoods or families where fewer people identify

---

[3]The estimated persuasion rates in the literature for news media, such as TV programs (Enikolopov et al., 2011; Gentzkow, 2006; DellaVigna and Kaplan, 2007), radio (DellaVigna et al., 2014; Yanagizawa-Drott, 2014; Adena et al., 2015; Blouin and Mukand, 2019), and newspapers (Chiang and Knight, 2011; Gerber et al., 2009; Gentzkow et al., 2009) are around 2% to 20%. One noticeable exception is Enikolopov et al. (2011), who utilized idiosyncratic variations in the signal availability of an independent television station (NTV) in Russia and found that people who had access to an NTV were less likely to support the pro-government party in the 1999 election. The estimated persuasion rate was 65%, i.e. around 65 percent of the pro-government party supporters who watched NTV changed their mind and voted for other parties.

as Taiwanese.[4] The result aligns with the predictions made by "belief-based models," in that people with weaker prior belief (i.e. weaker Taiwanese identity) are more likely to be persuaded by new information (DellaVigna and Gentzkow, 2010).

Finally, we study the long-term effects of junior high school curricula on people's national identity around the age of 30 (i.e. 24 to 33 years old). Our results suggest that, a decade after the students left junior high school, people who studied the old textbooks hold a similar level of Taiwanese identity to those who studied the new textbooks, and the Taiwanese identity level held by new textbook readers did *not* decline. Since the perspectives of old textbooks are in conflict with the recent social trend, our interpretation is that in the long run the old-textbook readers eventually "catch up" with the general trend and the identity of individuals who studied the new textbook.

Our paper stands apart from the previous literature in the following ways. First, we provide one of the first pieces of evidence on the effect of a school curriculum (i.e. textbook content) on an individual's national identity. The formation of national identity has drawn substantial attention in the social sciences (Turner et al., 1987; Tajfel and Turner, 1979; Akerlof and Kranton, 2000; Alesina and Fuchs-Schündeln, 2007; Bisin and Verdier, 2010; Manning and Roy, 2010; Bisin et al., 2011b; Masella, 2013; Constant and Zimmermann, 2013; Georgiadis and Manning, 2013; Jia and Persson, 2019; Durante, 2020). Previous studies in this stream of the literature have focused on how ethnic diversity affects national identity.[5] The context of the present study is interesting, because the national identity of people in Taiwan has changed dramatically in the last two decades—the proportion of Taiwanese identity increased rapidly from 17% in 1992 to 60% in 2015, as shown in Figure 3.1[6]; however, the ethnic composition

---

[4]In section 3.6.2, we use ethnic composition at the township level as a proxy for the intensity of Taiwanese identity in individuals' home towns. We also use the variation in ethnic composition of individuals' parents as a robustness check and obtain consistent results.

[5]For example, Constant and Zimmermann (2013) offers a thoughtful and thorough discussion on identity formation and its consequences for economic behavior. Masella (2013) suggests that ethnic diversity might not necessarily weaken the intensity of national feeling, whilst Durante (2020) finds that a victory by a country's national team can strengthen national identity and weaken ethnic identity.

[6]Note that there is a substantial increase in the share of respondents reporting themselves as Taiwanese in 1997. We think this could be related to the fact that Taiwan held its first presidential election in 1996. In addition, China

in Taiwan has been quite homogeneous and stable since 1949.[7] Our results suggest that the revision of educational content could play an important role in shaping people's national identity in Taiwan.[8]

There is a small but growing body of literature identifying the causal effects of education policies on people's political behavior and identity formation. Recently, several studies have examined how language use in education affects ethnic identity (Clots-Figueras and Masella, 2013; Fouka, 2019),[9] whether additional schooling affects civic participation or political attitudes (Milligan et al., 2004; Friedman et al., 2016), the impact of authoritarian education on political ideology (Bai and Li, 2018), and the effect of patriotic activities in school on the assimilation of immigrants (Mitrunen, 2018).[10] Our research complements these works by focusing on the impact of school curricula, which should be the key component of the educational process. Compared to other educational policies, changes in textbook and course contents are more common across the world, so understanding their impacts could have more implications. In addition, the reform used in this study only adjusted the textbook contents of social subjects and was not associated with other changes in the educational system, such as the languages of instruction. This feature allows us to clearly estimate the curriculum effect.[11]

One noteworthy exception is Cantoni et al. (2017), who examined the effect of the

---

fired a series of missiles in response to Taiwan's President Teng-hui Li visiting the United States (i.e., the 1996 Taiwan Strait crisis). These events might have strengthened Taiwanese identity.

[7]According to government statistics (Hsiau, 2003; Copper, 2019), over 95% of Taiwan's population consists of the Han people, split into three main groups: Hoklo, Hakka, and Mainlander. Around 2.3% are Austronesian peoples (i.e. Taiwanese aborigines). Due to the Chinese Civil War (i.e. the Kuomintang-Communist Civil War), more than two million Mainlanders retreated from China in 1949.

[8]One recent study (Chiang et al., 2019) empirically examined how economic integration with China affected Taiwanese identity formation. They found that rising investment in China has strengthened Taiwanese identity, especially for unskilled workers.

[9]Clots-Figueras and Masella (2013) found that changing from single-language (Spanish) to bilingual (Spanish and Catalan) education in Catalonia provided students with a stronger sense of Catalan belonging, which led further to changes in political party preferences in elections. Fouka (2019), on the other hand, documented that children of German immigrants who experienced language prohibition in elementary school were more likely to marry Germans, choose more 'German' first names for their children, and be less likely to volunteer in World War II.

[10]Bai and Li (2018) examined the long-term effects of education under the authoritarian regime in Taiwan, finding that one additional year of exposure to authoritarian education during youth could substantially affect an individual's political behaviors, such as their preference for democracy or voting for an authoritative party.

[11]For example, the reform used in Clots-Figueras and Masella (2013) involves adjustments in languages of instruction and textbook contents. Therefore, their results are mixed with both language and curriculum effects.

school curriculum on individuals' political attitudes by exploiting a high school textbook reform program in China. They conducted a survey of students at Peking University (i.e., elite students) and found that those exposed to the new textbooks showed more trust in government and more skepticism toward unconstrained democracy and free markets, which is consistent with the political aims of a new curriculum. This new curriculum also aims at promoting Chinese ethnic unity. However, their results suggest that the new curriculum had insignificant impacts on people's national identity and ethnic identity. In contrast to the findings in Cantoni et al. (2017), our results indicate that people's national identity can be shaped effectively by the content of a textbook.

Second, using nationally representative survey data, our paper examines the impact of educational content on identity for the more general population. The results in previous studies are usually based on a specific subgroup, such as elite students (Cantoni et al., 2017) or immigrants (Fouka, 2019; Mitrunen, 2018). Nevertheless, these results might not be generalized to the whole population or other groups of individuals; in fact, our subgroup analysis shows that the effect of educational content can be heterogeneous across different types of people. The curriculum only affects the identity of specific subgroups, such as individuals who spend more time on reading textbooks or those with less prior belief. These results help us understand the potential mechanisms of curriculum effects.

Finally, we contribute to the existing literature by investigating the long-term effects of the school curriculum. Understanding long-term effects on political preferences has important implications. Recent evidence shows that significant political events in people's 15-24 (i.e., impressionable years) can influence the political attitudes in their entire lives (Ghitza and Gelman, 2022). It is possible that the school curriculum also has persistent impacts. Different from Cantoni et al. (2017), who examine the short-term impacts of textbook contents (i.e., 1 to 2 years after reading textbooks), our repeated survey data allows us to know how the curriculum effects evolves 10 to 20 years after individuals have read the textbooks.

The paper proceeds as follows. In section 3.2, we discuss the background of the curricu-

lum reform and analyze the differences between the old and new curricula. Section 3.3 describes the data and the sample used in this paper, and section 3.4 discusses our identification strategy—a regression discontinuity design. Section 3.5 presents our main results, following which we then explore potential mechanisms through a subgroup analysis in section 3.6 and long-term effects in section 3.7. Finally, section 3.8 concludes.

3.2 Policy Background

3.2.1 The Curriculum Reform of the *Knowing Taiwan* Series

In 1994, the Taiwanese government announced a new curriculum for the junior high school social subjects: History, Geography, and Society. The major change lay in the design of the first-year content. In earlier textbooks, Taiwan-related content accounted for only a small proportion of the text and was scattered through different volumes. However, the new curricula, especially in terms of the History subject, aimed to provide not only much more Taiwan-related knowledge, but also different angles on the history and social development of the nation.[12] After three years of writing and editing, the government published the new textbooks, and students entering junior high school in September 1997 were expected to utilize them accordingly.[13]

The reform was comprehensive, in that students across Taiwan who entered junior high school after September 1997 would study the series. Though the major changes applied mainly to first-year textbooks, the second- and third-year textbooks were also adjusted. An Online

---

[12]The Geography and Society volumes in the *Knowing Taiwan* series introduce extensive knowledge about Taiwan's geographical features, social values, culture, and religions. This knowledge may indeed also affect people's national identity, but the History textbook is likely to play a major role in identity formation. Wang (2001) discussed how the *Knowing Taiwan* series, namely, the History volume, strengthened Taiwanese consciousness.

[13]This reform aroused fierce debate among political parties on whether the books were "appropriate". Political factions at that time were divided into two groups, with the likes of the Kuomintang and the New Party following the "successor to China" ideology, while the Democratic Progressive Party advocated "Taiwan independence" and considered the Kuomintang government, which had ruled Taiwan since 1945, a foreign regime. Discussions at the time, about whether the History textbook in the *Knowing Taiwan* series should be adopted, centered around three perspectives in the textbook: the "relationship between Taiwan and Japan in history," the "relationship between Taiwan and China in history," and the "judgment of contemporary political events and politicians" (Wang, 2001). According to Wang (2001), in just two months, from June to August 1997, 341 articles (five articles every day on average) about *Knowing Taiwan* appeared in the nation's four main newspapers.

Appendix 3.10.1 provides more details on this issue. Senior high school/vocational school entrance examination for students born after September 1984, compared to examinations for earlier cohorts, were therefore based on different textbooks for all three years, thus ensuring that earlier education cohorts were not exposed to the *Knowing Taiwan* series. Herein, we define the education cohort as students entering the compulsory education system in the same year, and we label them with the year they entered junior high school. For example, the 1997 education cohort entered junior high school in September 1997. They were the first to study the *Knowing Taiwan* series and were born between September 1984 and August 1985.

3.2.2 Comparison between the Old and the New Curricula

This curriculum reform aroused politicians' attention, because it brought to awareness the stark differences between two imagined nationalities, namely Chinese consciousness and Taiwanese consciousness (Liu et al., 2005; Wang, 2001). In particular, the new history textbooks moved away from the "China-oriented" angle seen in earlier textbooks, to a "Taiwan-oriented" view. In general, there are two main differences between the old and new textbooks: 1) The amount of content about Taiwan and 2) the context given about the relationship between Taiwan and China. Therefore, the new history textbook may have cultivated Taiwanese identity in two ways: First, there may have been a priming effect, due to students reading the word "Taiwan" more often, and second, the distinction made by describing Taiwanese and Chinese history separately may have provided students with different information to associate with the two imagined groups, and hence helped them differentiate between Taiwanese and Chinese.

A Substantial Increase in Taiwan-Related Content

Under the old curriculum, junior high school students studied the history of China for a year and a half, and then the history of the world for another year and a half, whereas under the new curriculum they studied the history of Taiwan in the first year (i.e. the History textbook in the *Knowing Taiwan* series), the history of China in the second year, and world history in the

third year. In other words, content on the history of China and the rest of the world in the old version was condensed in the new version so that new materials about Taiwan could be added.

In terms of time, teachers utilizing the new textbooks might have spent much more on the history of Taiwan than they did previously. Under the old curriculum, teachers spent three semesters on the history of China (25 chapters), with only one chapter and a section related to Taiwan.[14] Assuming that teachers spent the same amount of time on each chapter and section in a volume, we approximate that they would have spent less than one-fifth of a semester on history related to Taiwan. In contrast, the *Knowing Taiwan* History volume was designed to cover two semesters, with 116 pages of content. For comparison, the old textbooks contained only 16 pages on the subject.

The explicit aim of the *Knowing Taiwan* series History volume was emphasized by its editors as follows:

> This book aims to introduce students to **the history about how ancestors of different ethnic groups made developments in Taiwan.** As a result, students are expected to cultivate a cooperative spirit, patriotic feelings, and worldwide horizons. Also, it is hoped this will augment their understanding of **Taiwanese cultural assets**, and make them appreciate and treasure them accordingly.[15]

The intention of acquainting students with Taiwanese development was not apparent in the old version—as seen from the editors' preface to the old textbook on the history of China:

> The history of China describes **the evolution of Chinese nationality, the change of the territory, and the development of politics, society, economics, and culture**. In particular, it stresses the long history and the blending of the culture of nationality, in order to strengthen patriotic feelings and a cooperative spirit, and to understand the nation's traditions, its position and the responsibility of the population.

---

[14]In the old textbook series, these 25 chapters were spread across three volumes, i.e. one volume per semester. Only a section in the 15th chapter, entitled "The rebellion of Koxinga against the Qing Dynasty and the development of Taiwan," and the 25th chapter, entitled "The achievement and vision of a base for revival," included Taiwan-related content.

[15]Emphasis in this paragraph is added by the authors.

Distinguishing between Taiwan and China

The new textbook not only contained a substantial increase in content about Taiwan, but it also clearly distinguished between the concepts of Taiwan and China, in a contextual change. Basically, the new textbook treated the history of Taiwan as an entity completely detached from the history of China. In contrast, the old textbook did not emphasize this difference. Furthermore, depending on the context, the old textbook sometimes used "our country" to refer to China but sometimes also to refer to Taiwan. Thus, studying the old textbook could have confused students about their national identity.

In their first grade of junior high school, students studying the old textbooks started to learn the history of "our country (i.e. China)" through the statement that the earliest human beings lived in "our country (i.e. China)," namely *Homo erectus pekinensis*, in the Palaeolithic age. The "common ancestor" of *Chinese nationality* was Huang Di, and the first dynasty of "our country (i.e. China)" was the Xia Dynasty. The history of "our country (i.e. China)" therefore proceeded through sequential dynasties, from Xia to Qin, to Tang, and all the way to Qing.[16] Interestingly, the old textbook also used "our country" to refer to Taiwan when it mentioned the development of the Kuomintang government in Taiwan after the 1949 Chinese Civil War (i.e. the Kuomintang-Communist Civil War).

In contrast, the term "our country" is used less in the History textbook in the *Knowing Taiwan* series or for the textbook on the history of China in the new curriculum; "Taiwan" and "China" are used instead. More precisely, "our country" only appears in descriptions of Taiwan. Following the divided usage of terms, Taiwanese history stands out not as part of the history of China but as an individual entity in the new History textbooks. In the Online Appendix 3.10.2, we use several sample paragraphs from such textbooks to show the differences in historical

---

[16]Between the Ming and Qing dynasties in this straightforward development line, students saw the first appearance of "Taiwan," identified by the editors as a basis for Koxinga's fight against the Qing regime. It is worth noting that Koxinga is written as "recovering" Taiwan from the Dutch. The usage of the verb demonstrates explicitly the ideology behind the old textbook, showing that the editors viewed the ruling Dutch in the 17th century as a "foreign regime." Simultaneously, this implicitly claimed Taiwan as the territory of "our country (i.e. China)" before Dutch rule.

perspectives between the old and the new curriculum. Basically, in the old curriculum, "Taiwan" was virtually ignored, and "our country" usually referred to "China". In the new curriculum, "Taiwan" and "China" were explicitly separated so that readers had the chance to distinguish between the two.

3.2.3   The Role of Teachers

So far, we have not discussed the role of teachers in this curriculum reform. For example, teachers might change how they conduct a lecture according to the new curriculum. To the best of our knowledge, the Ministry of Education did not request teachers to utilize different ways to teach new textbooks. Basically, they followed the content of the textbooks. In addition, junior high school education in Taiwan is exam-oriented, and the senior high school entrance examination is fully based on textbook content. Therefore, we believe the role of teachers is relatively minor.

3.3   Data and Sample

3.3.1   Data

The data used in this paper is taken from the Taiwan Social Change Survey (TSCS), which is a nationally representative repeated cross-sectional survey for respondents aged above 18 in Taiwan. The sample size of each TSCS wave is around 1,800 to 2,200 respondents.

Three features of the TSCS make it suitable for our analysis. First, it asks respondents consistently about their national identity through the following question:

- In our society, some people call themselves Taiwanese, some Chinese, and some both. Do you consider yourself Taiwanese, Chinese, or both?

This feature allows us to combine different survey waves, in order to compare the short- and long-term impacts of curriculum reform on national identity. Second, the TSCS records the birth year *and* birth month of respondents. Since the school year in Taiwan starts in September, by

exploiting this feature, we can identify the correct educational cohort, which is crucial for our regression discontinuity design. Third, the TSCS holds rich demographic information about respondents, which helps us investigate the mechanism further through subgroup analysis.

3.3.2 Sample

The first educational cohort exposed to the 1997 curriculum reform, born in September 1984 or later, was first surveyed in 2003.[17] To balance out regression analysis respondents before and after the reform, we hence include surveys held from 2003 onward, which contain the national identity question and enough demographic information: These are the 2003, 2004, 2005, 2010, 2012, 2013, 2014, and 2015 waves.[18]

We drop any respondents who reported being born outside Taiwan and those who reported that the place they had lived the longest before they were 15 was outside of the country, since we could not be sure that they had entered junior high school and hence been exposed to the curriculum reform. In addition, we drop respondents whose answer to the national identity question was "Other." These selection rules remove 2% of the main regression sample (i.e. the short-term sample). The main results in this paper are not influenced by the sample selection.

3.3.3 Construction of Outcome Variable

Based on the TSCS's national identity question, we create the outcome variable as a dummy variable *Identity* by assigning one to respondents answering "Taiwanese" and zero to those answering "Chinese" or "Both."[19] In our main regression sample, only 3.8% of respondents

---

[17]Some of birth cohorts (e.g., those born in 1985 or 1986) were not surveyed since they were below 18 years old in 2003 and 2004.

[18]Note that 2009 TSCS had an identity question but did not include the demographic information we need in our regressions. Hence, we do not include this wave in the RD design. In addition, the TSCS held two waves in 2014.

[19]Since the measurement of national identity is based on a self-reported response, the natural question is: Does this measurement truly reflect respondents' national identity? One possible explanation for a change in *Identity* (if observed) is that previous students were afraid to respond that they felt Taiwanese. The new textbooks provided not only a Taiwanese identity, but also the message that viewing oneself as Taiwanese was no longer taboo. We provide two counterarguments to this explanation. First, the simple mean of *Identity* for the control group in our main analysis sample is 0.6. When over half of one's peer group identify themselves as Taiwanese, it is hard to believe that the Taiwanese identity was indeed taboo. Second, the change in *Identity* should be visible in different subgroups if this explanation were indeed true, but in section 3.6 we find this is not the case.

answer "Chinese," indicating that, in this generation, very few people identify as exclusively Chinese. Most of the respondents have an exclusively Taiwanese identity (64.8%) or a dual identity (31.4%), considering themselves to be both Chinese and Taiwanese.[20]

3.4   Empirical Specification

3.4.1   Graphical Evidence

Figure 3.2 plots the simple mean of *Identity* in each educational cohort, using all available data. We observe a roughly 10% increase in Taiwanese identity between the 1996 and 1997 education cohorts (i.e. between the last to study the old textbooks and the first to study the *Knowing Taiwan* series). Two important caveats should be noted in the above analysis. First, compared to people who enter school earlier, those who enter school later are less likely to have been surveyed in the early years, since they are too young to become respondents. In addition, people's national identity might be affected by social events happening in the survey year, so the above change in Taiwanese identity could be confounded by survey year effects. Second, the result in Figure 3.2 might be mixing up the short- and long-term effects of the school curriculum on Taiwanese identity. Since we use all available survey waves from 2003 to 2015 to plot Figure 3.2, this implies that some in the sample would have been surveyed in the early stages of their life, and some in the later stages. To alleviate the above concerns, we control for the survey year fixed effect and restrict our sample to fewer education cohorts in the rest of our analysis, namely, those born between September 1982 and August 1986 (four education cohorts, two of which would have studied the *Knowing Taiwan* series). In addition, we first analyze these cohorts when they were relatively young, aged from 18 to 23 and surveyed from 2003 to 2005 (henceforth **short-term sample**, 5 to 10 years after reading textbooks). To examine if the curriculum effect is persistent, we examine the same education cohorts surveyed from 20010 to 2015, when aged

---

[20]Note that in the 2005, 2010, 2014, and 2015 waves, TSCS further categorizes "both" into two alternatives: 1) Both Taiwanese and Chinese; 2) Both Chinese and Taiwanese. When we construct our outcome variable, these two alternatives refer to "both". In the later section, we show that the estimation is not affected by a particular framing of the questions.

between 24 and 33 (henceforth **long-term sample**, 11 to 20 years after reading textbooks).

3.4.2 Regression Discontinuity Design

Different cohorts of students would have been exposed to different societal trends, which in turn may have affected their national identity formation. Thus, we use a regression discontinuity (RD) design to eliminate this problem by comparing the identities of people born close together (i.e. around September 1984). The reason this work is that close birth cohorts should experience almost the same societal developments while growing up. The major difference is that those born just after September 1984 would have studied the *Knowing Taiwan* series, while those born just before this date would have studied the old textbooks. At first glance, we should conduct an RD design on an education cohort (i.e. academic year) basis, since the treatment status varies at that level. However, people in the same education cohort may have experienced different events that could have altered their national identity.

An example of this relates to voting. Elections in Taiwan are generally held in December, January, and March, and the age at which one becomes eligible to vote is 20. In some elections, people born in the first half of the education cohort would have been eligible, while those born later would not have been. Students in the first cohort exposed to the curriculum reform offer one example in this regard. The sixth legislative election was held on December 11, 2004, splitting the education cohort into two groups: People who had the voting right (born before December 11, 1984) and people who did not have it (born after December 11, 1984). Students in the last cohort studying the old textbook provide another example. The event in this case was the presidential election that took place on March 20, 2004. The reason this is important is that politicians in Taiwan debate fiercely on the subject of national identity in elections. Thus, different "first vote" experiences may affect people's national identity formation. Bearing in mind such differences embedded in respondents within an education cohort, we measure birth cohort at the year-month

level and estimate the following regression:

$$Identity_{it} = \alpha_0 + \alpha_1 TextBook_i + f(m; \beta) + X_i' \gamma + \lambda_t + \varepsilon_{it} \quad (3.1)$$

where $Identity_{it}$ indicates the dummy variable defined in section 3.3.3, for individual i interviewed at year t. The variable $TextBook$ indicates whether the respondent was exposed to the curriculum reform and takes the value one if the respondent reported himself born after September 1984, and zero otherwise.[21] We use birth cohort measured by year-month as our running variable, and we center it on September 1984, the first year-month affected by the reform. In our main specification, we estimate equation (3.1) within a bandwidth of 24 months before and 24 months after September 1984 (i.e. we use the sample born between September 1982 and August 1986).[22] In addition, we specify $f(m; \beta)$ as a linear function but allow the slope to be different on either side of the cut-off. That is, $f(m; \beta)$ is the first-order polynomial of birth cohort m interacting fully with $TextBook$.[23] In a later section, we examine whether our main results are sensitive to the bandwidth choices and different specifications.

Our primary interest is in α_1, which measures any deviation away from the relationship between the birth cohort and Taiwanese identity $Identity_{it}$ at the cut-off (i.e. when the treatment

[21]Although the enrollment cutoff is nationally mandated, it is possible that some parents do not follow the rule. However, we are unable to examine this concern directly, since TCSC data does not provide information about an individual's school enrollment status. Instead, we use 2006-2018 PISA (Programme for International Student Assessment) data, which contains a student's birth year-month and enrollment status, to investigate this issue. We find that most students (around 95%) follow the nationally mandated enrollment cutoff. Thus, we believe the variable $TextBook$ can represent whether the respondent was exposed to the new curriculum or not.

[22]Junior high education curriculum reforms in Taiwan have happened every five to ten years since 1968, when compulsory education was extended from six to nine years. The exact years new curricula were introduced were 1968, 1972, 1983, 1986, 1995, 2001. Note that the new curriculum we looked at was published in 1995, but the textbooks were not adopted until 1997. Curriculum reforms for senior high education happened on average every decade. The exact years were 1962, 1971, 1983, 1995, 2005. The only curriculum reform experienced by the four education cohorts we focused on is the one we looked at. The next closest reform to them was from 1995 for senior high education. The senior high textbooks, edited according to the 1995 curriculum, were adopted in 1998. Hence, each of the four education cohorts we focused on studied the same senior high school textbooks if they entered the academic track. As far as we know, the reform we are looking at is the first since 1968 in junior high education to center on social objects.

[23]We also include a second-order polynomial of the birth cohort m interacting fully with $TextBook$, for a robustness check.

variable *TextBook* switches from 0 to 1). If all factors except textbook content did not change around the cut-off, α_1 can be interpreted as the causal effect of the junior high school curriculum on students' Taiwanese identity.

In order to single out the overall effect of societal trends in each survey year, we include the survey year fixed effect (λ_t) in all specifications. We also include a set of covariates (X_i) which might influence national identity formation, including gender, age, parents' education, parents' ethnicity, share of Hoklo people in the respondents' hometown, and a set of dummy variables indicating the region where an respondent lived in before his/her 15 years-old. The parents' ethnicity and education level capture the family's influence on the respondents' national identity.

Four major ethnicities live in Taiwan: Hoklo, Mainlanders, Hakka, and Aborigines. Using 1992, 1995, and 1998 TSCS data, we display a breakdown of these four ethnic groups in Figure 3.19 of the Online Appendix. About 70% of the Taiwanese people descend from Hoklo immigrants, who originated from Xiamen, Quanzhou and Zhangzhou, China, and arrived on the island around 400 years ago. As the largest ethnic group, compared to other ethnicities, the Hoklo people are more likely to have a Taiwanese identity (i.e., call themselves Taiwanese only).[24] Figure 3.20 in the Online Appendix indicates that about 39% of Hoklo people identify as Taiwanese, which is much higher than the other main ethnic groups, namely Aborigines (27%), Hakka (25%), and Mainlanders (8%).[25] Therefore, in some specifications, we include the share of Hoklo people in the respondents' hometown to control the intensity of Taiwanese identity in individuals' hometowns.

The inclusion of dummy variables for regions help us control for regional factors possibly

---

[24]Since 1945 (the end of Japanese colonization in Taiwan), construction of the concept of Taiwanese has centered on ethnicity groups living in Taiwan before 1945. This includes Aborigines (in Taiwan for thousands of years), Hoklo and Hakka (migrated from southern China since 400 years ago) but excludes Mainlanders (who have migrated from all over China since 1945). Politically, the Hoklo people account for the majority of the population, and they play a more important role in political movements, which often mobilize people via identity politics, than Hakka and Aborigines.

[25]We also use 1992, 1995, and 1998 TSCS data and restrict the sample to people who are 25 years old or above, in order to make sure that the respondents are not affected by curriculum reform.

influencing national identity formation, such as local support for a certain political party.[26] In the Online Appendix 3.10.4, we provide detailed definition of these individual characteristics. Finally, standard errors are clustered at the birth cohort level (i.e. birth year-month).

Table 3.1 reports the summary statistics of related individual characteristics in the empirical analysis, such as the respondent's gender, age, and years of schooling, their fathers'/mothers' education level, their fathers'/mothers' ethnicity (i.e. whether they are Hoklo people), and the share of Hoklo people in the respondents' hometown. We find both treatment and control groups are similar in terms of these variables except for the respondent's age. The treatment group is 1.4 years-old younger than the control group. This result is not surprising since our research design essentially compares the young and old educational cohorts. In the empirical analysis, we will control for the effects of birth cohorts on outcomes using a linear function of birth year-month. To sum up, our findings from Table 3.1 suggest that the characteristics of treatment and control groups are quite balanced.

3.5 Results

3.5.1 The Effect of Curriculum Reform on Taiwanese Identity

Figure 3.3 displays the relationship between Taiwanese identity and the birth cohort. We group up the sample by every three birth year-months to increase the sample size of each dot. Thus, each dot in Figure 3.3 represents the average of variable *Identity* (i.e. Taiwanese identity) by three birth year-month cohorts (i.e. the birth year-quarter cohort), after it has been regressed on the survey year dummies (i.e. controlling for the survey year fixed effect).[27] The lines in Figure

---

[26]There were 23 county/city in Taiwan during the sample period. We categorize them into four regions: northern, middle, southern, and eastern regions. Northern region includes Taipei City, New Taipei City, Yilan County, Taoyuan City, Keelung City, Hsinchu County, Hsinchu City. Middle region includes Miaoli County, Taichung City, Taichung County, Changhua County, Nantou County. Southern region includes Yunlin County, Chiayi County, Chiayi City, Tainan City, Kaohsiung City, Tainan County, Kaohsiung County, Penghu County and Pingtung County. Eastern region includes Hualien County and Taitung County. We use the eastern region as a reference group.

[27]The graph is at the birth year-quarter level, so the first dot in Figure 3.3 represents average *Identity* (after controlling for the survey year fixed effect) for those born in September, October, and November 1982, and the last dot represents average *Identity* (after controlling for the survey year fixed effect) for those born in June, July, and August 1986. In the later sections, we use a similar way to display Figure 3.7, Figure 3.9, and Figure 3.10.

3.3 represent fitted regressions of the cell's mean dots, using first-order polynomials interacting with the dummy variable *TextBook*. In so doing, we eliminate the potential confounding effect of the survey years. The fitted line in Figure 3.3 suggests that the discontinuity of *Identity* is roughly 20 percentage points around the cut-off.

Table 3.2 shows the regression results of the estimating specification (3.1). The first-order polynomials of birth cohort m fully interact with *TextBook*, and the survey year fixed effects are included in all regressions. Column (1) reports our baseline results. Consistent with the graphical evidence in Figure 3.3, the estimate of the coefficient on *TextBook* is 0.16 and statistically significant. In other words, studying the new textbook (i.e. the *Knowing Taiwan* series) can increase one's probability of reporting oneself as Taiwanese by around 16 percentage points.

In columns (2) to (4), we gradually include ethnic/demographic variables to increase the precision of the estimates and lessen any potential bias due to discontinuities in observables at the cut-off. In general, we find qualitatively similar estimated coefficients on *TextBook* across the different specifications. Our results suggest that new curricula significantly increase the likelihood of identifying as Taiwanese by around 18 percentage points. Compared to the baseline mean of *Identity* (i.e. around 61%)[28], the magnitude of the estimated effect is sizeable—accounting for a 30% increase.

Note that the changes in the 1997 curriculum reform include: 1) More materials covering Taiwan and fewer covering China; 2) The wording used in the textbook distinguishes between Taiwan and China. One important caveat is that the estimated effect bundles up all changes in the reform. We need to stress that our research design and data cannot identify which key element within the curriculum reform leads to changes in Taiwanese identity. Thus, the results provide a global evaluation of the 1997 curriculum reform.

---

[28]This is the mean of *Identity* across all those in the sample who were born between September 1982 and August 1984 (i.e. the control group).

3.5.2 Discussion: Persuasion Rate

In this section, we provide the persuasion rate, calculated by the formula used in the literature on persuasive communications (DellaVigna and Gentzkow, 2010):

$$100 \times \frac{y_t - y_c}{e_t - e_c} \times \frac{1}{1 - y_c},$$

where e_i denotes the share of group i receiving the message (the textbook content in our case), and y_i the share of group i adopting the behavior (i.e. considering themselves Taiwanese in our case). The subscripts t and c represent the treatment and control groups. The persuasion rate measures the degree to which the treatment persuades people to adopt the behavior, scaled by the share of people receiving the messages and the share of the control group "to be persuaded" $(1 - y_c)$. In our case, since all students born after September 1984 were exposed to the new textbook, $e_t - e_c = 1 - 0 = 1$. The persuasion rate reported in the first column in Table 3.2 is calculated as $100 \times \frac{0.182}{1} \times \frac{1}{1-0.608} = 46.4$. This 46% persuasion rate is quite high compared to persuasion rates found in the literature studying the persuasive effects of media communications, which are barely higher than 20% (Enikolopov et al., 2011; Gentzkow, 2006; DellaVigna and Kaplan, 2007; DellaVigna et al., 2014; Yanagizawa-Drott, 2014; Adena et al., 2015; Blouin and Mukand, 2019; Chiang and Knight, 2011; Gerber et al., 2009; Gentzkow et al., 2009). Our estimate, however, aligns with the persuasion rate found in Cantoni et al. (2017), in which more than a quarter of the persuasion rates were higher than 20%, and the highest was 50%. The high persuasion rate is not that surprising after taking into account the degree of exposure: Students had to study the *Knowing Taiwan* series for at least a year, and they also spent three years memorizing the materials for the high school admission examinations. This exposure is much greater than typically occurs with specific newspaper, TV, or radio programs.

3.5.3 Robustness Check

We validate the robustness of the main results in two ways. First, we discuss their sensitivity to different empirical settings, such as the inclusion of higher polynomial orders, the choice of bandwidth, and sample selection. Second, we investigate the validity of the identification assumption for RD design, by examining the smoothness of observable covariates and conducting a series of falsification tests.

Choices of Polynomial Order and Bandwidth

To examine whether our results are sensitive to different parametric specifications, Table 3.3 displays estimates based on a specification with a second-order polynomial (i.e. quadratic spline). The estimated results suggest that studying new textbooks, on average, can increase Taiwanese identity by 19-21 percentage points, which is a range quite similar to our main estimates. Next, we examine the robustness of our estimates over a wide range of bandwidths. Figure 3.4 shows the point estimates of the coefficient on *TextBook* and their corresponding 95% confidence intervals, using the same specification as in column (4) of Table 3.2, with bandwidths ranging from two education cohorts (i.e. 24 months) to one (i.e. 12 months) on each side of the cut-off. The magnitudes of the point estimates remain similar as we narrow down the birth year-month window, showing that the results in Table 3.2 are not sensitive to bandwidth choice.[29]

Exclude Specific Birth Cohorts

Based on Figure 3.3, it seems that our RD results is driven by the birth cohorts between -6 to -4 (i.e. individuals born between March to May 1984). In order to investigate this concern, Panel B of Table 3.3 reports the estimates based on the sample excluding these cohorts. We find that the RD estimate decrease slightly to 0.13. But the estimate is still statistically significant and suggests that new curricula raises the likelihood of identifying as Taiwanese by around 13 percentage points.

---

[29]The confidence intervals of point estimates increase slightly. The estimated standard errors increase from 0.083 (bandwidth: 24 months) to 0.097 (bandwidth: 12 months).

Wording of the Identity Question

The framing of the identity question in TSCS varies slightly across years. In addition, the theme of the survey is sometimes "national identity," while it is "civil rights" or "religion" in other years. Specifically, the wording of the identity questions used in the 2003 and 2004 waves is slightly different from the 2005 wave. The identity question for the 2005 TSCS categorizes "both" in two ways: 1) Both Taiwanese and Chinese; 2) Both Chinese and Taiwanese. Therefore, we conduct our RD estimations based on the questions in the 2003-2004 waves and 2005 wave, respectively. Panels C and D of Table 3.3 and Figure 3.21 in the Online Appendix suggests that the estimated magnitudes of textbook effect are fairly similar across different waves.[30]

Smoothness of Observable Covariates at Cutoff

A key identification assumption of RD design is that the individuals' characteristics should be similar on both sides of the cut-off (i.e. born in September 1984). In other words, no other confounding factors should change in September 1984. To investigate this issue, we examine whether the selected observable characteristics are balanced on both sides of the cut-off. We use these characteristics as outcome variables and estimate equation (3.1) without controlling for the covariates X_i. The regression results are shown in Table 3.4. Most observable characteristics do not exhibit significant discontinuities at the cut-off.

The only exception is the share of Hoklo people in the hometown. The sixth column of Table 3.4 suggests this variable exhibits a drop at cutoff, with a size of 7.5 percentage points (i.e. less than 10% decline from baseline mean). In other words, it is more likely that we will observe a respondent who lived in a town with fewer Hoklo people on the right-hand side of the cut-off. However, we find that the statistical significance of this estimate is only marginal at 10% level. Furthermore, Table 3.1 suggests that the change in share of Hoklo people at cutoff is not significant when comparing the observations of two-sides around cutoff directly.

---

[30]The RD estimates based on 2003-2004 waves are around 0.18 to 0.24. Due to smaller sample size, the estimates using 2005 wave is not statistically significant but within the same range (i.e., 0.19 to 0.23).

Therefore, we think the finding of discontinuity in share of Hoklo people in the hometown is not very conclusive. Finally, in order to lessen any potential bias, we include this variable in the specifications and find that our estimates are robust to its inclusion.

Density of the Running Variable around Cutoff

Although the running variable of our RD design – birth cohort – is predetermined and unlikely to be affected by the reform in 1997, it is still possible that the survey might have sampling biases and the number of individuals different around cutoff (i.e., September 1984). We implement a density discontinuity test to examine this issue (Cattaneo et al., 2020, 2018; McCrary, 2008). Figure 3.22 displays the results for the density test and suggests that there is no discontinuity in the distribution of the running variable at the threshold.

Placebo Tests

In this section, we further examine our identification assumption, namely, that no other confounding factors change at the cut-off, by conducting a series of placebo tests. One potential confounding factor could be the mental age effect: People who were born on the left-hand side (i.e. August) of the birth year-month cut-off would have been more mentally mature than those on the right-hand side (i.e. September), since they had entered the school system earlier and thus, at any given time, may have had more work or social experience, which might have affected their Taiwanese identity. That being the case, we should observe similar jumps in September for every birth cohort. To examine this hypothesis, we estimate equation (3.1) for three fake reforms.

We take 1996, 1995, 1994, and 1993 as academic years for the fake curriculum reforms and thus treat September 1983, 1982, 1981, and 1980 as birth year-month cut-offs for placebo tests.[31] We then replicate the results in Table 3.2 for each fake curriculum reform, using the same TSCS waves in 2003, 2004, and 2005. Note that we only include two education cohorts (i.e. 24 months) on each side of the fake birth year-month cut-off, to make the falsification results

---

[31]They are 13 years old in 1996, 1995, 1994, and 1993, respectively.

comparable to our main results.

Panel A to D of Table 3.5 show the results of the falsification regressions. The estimated "treatment effects" are generally insignificant and the magnitudes are quite small. Figure 3.5 show the consistent graphical evidence. Thus, the results of the above placebo tests suggest our main estimates might not be driven simply by the mental age effects or other confounding factors.

Since the choice of these years is rather arbitrary, we also generalize the above analysis to a permutation test, as in Cantoni et al. (2017), by assigning the fake reform to all possible months and years – from January 1950 to September 1983 – to obtain the distribution of the placebo estimates. Figure 3.6 compares the real estimates with these placebo ones. We find that among the estimates based on 405 fake reforms, only four of them are larger than the estimated curriculum effect (0.18). The real estimates are way above the placebo ones, and the p-value is only 0.01. In sum, these placebo tests indicate that the significant estimates in Table 3.2 should be treated as causal and are not just findings made by chance.

Finally, we conduct another type of placebo test by repeating the same RD analysis, but on this occasion we look at the different survey questions that could capture attitude towards other social values (e.g., opinions on social welfare or family issues). We list these questions and alternatives in the Online Appendix 3.10.5. The idea behind this placebo test is that the curriculum reform should not affect these cohorts differently on other social values – only on national identity. Table 3.10 in the Online Appendix suggests that the new curriculum had a negligible impact on other social values, thereby further verifying that our main result in terms of identity is not a chance finding.

Difference-in-Differences Design

In this section, we generalize the placebo tests in Table 3.5 by using a difference-in-differences design. Specifically, we combine all available cutoffs used in the main estimation and placebo tests, following which we narrow down the bandwidth to 6 months before and after

98

September in each year and estimate the following regression:

$$Identity_{it} = \kappa_0 + \kappa_1 AfterSep_i + \kappa_2 B_{1984} + \kappa_3 AfterSep_i \times B_{1984} + s(m;\beta) + X_i' \gamma + \delta_t + \varepsilon_{it}$$

(3.2)

where $AfterSep_i$ is a dummy indicating that individuals were born within the 6 months after September. That is, $AfterSep_i = 1$ if an individual's birth month is between September and the following year's February. $AfterSep_i = 0$ if an individual's birth month is between March and August. Similar to the standard DID design, we include a dummy variable B_{1984} indicating the 1984 cohort – individuals born between March 1984 and February 1985 (i.e. $B_{1984} = 1$) – since they were exposed to different curricula, depending on whether they were born before or after September. For other cohorts (i.e. $B_{1984} = 0$), the textbooks they read are independent of their birth month.[32] We also allow the linear spline of the running variable to be cohort-specific $s(m;\beta)$.

The key variable is an interaction term between $AfterSep_i$ and B_{1984}, which compares the cutoff of students born in 1984, net of the same cutoff differences for neighboring birth cohorts. If our RD estimate is mainly driven by the curriculum effect, we should expect that the jump from August to September would systematically only exist for the 1984 cohort. The remaining notations are defined in the same way as those in Equation (3.1).

Panel D of Table 3.5 shows that the estimated coefficients on $AfterSep_i \times B_{1984}$ range from 0.15 to 0.18. Since our treated cohort only includes those born within six months before and after September 1984, the estimates are less statistically significant, but the magnitudes of the curriculum effect are close to the RD estimates. Our preferred estimate (Column (4)) suggests that new curricula significantly increase the likelihood of identifying as Taiwanese by

---

[32]There are four other cohorts in the DID design. 1980 cohorts: individuals born between March 1980 and February 1981. 1981 cohorts: individuals born between March 1981 and February 1982. 1982 cohorts: individuals born between March 1982 and February 1983. 1983 cohorts: individuals born between March 1983 and February 1984.

around 18 percentage points.

3.5.4 The Effect of Curriculum Reform on other Political Outcomes

So far, we have found that people who read new textbooks are more likely than old-textbook readers to consider themselves as Taiwanese. In this section, we investigate the impact of curriculum reform on other political preferences and attitudes which might be related to the change in national identity. People holding stronger Taiwanese identity could be more likely to support independence or the parties who are against unification with China.[33] Indeed, Clots-Figueras and Masella (2013) found that individuals who had experienced greater exposure to teaching in Catalan not only had stronger Catalan identity, but also were more likely to vote for a Catalan regionalist party and had stronger separatist attitudes. The TSCS has a question in which respondents are asked about their opinion on whether they support Taiwan independence, the status quo, or the unification of mainland China and Taiwan:

- Concerning the future Taiwan mainland-China relationship, some think that Taiwan should be independent, while others think we should unify with mainland China. Which comes closer to your view? 1) Declare independence as soon as possible; 2) Maintain the present condition, but go towards independence in the future; 3) Maintain the present condition forever; 4) Maintain the present condition, but go towards unification in the future; 5) Unify with mainland China as soon as possible.

We create a dummy variable which is equal to one if the respondent selects the first two alternatives (i.e. support for Taiwan independence), zero otherwise. In addition, the TSCS includes a question in which individuals are asked which political party they support:

- Political parties in Taiwan have their own supporters. Among these political parties, which one do you support? 1) Kuomintang; 2) Democratic Progressive Party; 3) People First

---

[33]People in Taiwan aged above 20 are eligible for voting. In the first cohort who studied new textbooks, only half of the cohort who were born before 1985 March became eligible for voting in the president and national legislative election in March 2004; also, the observed turnout for the cohort is extremely low, which limits the sample size that reports voting choice in the short-run data (2003-2005). We thus do not analyze direct voting choices in this section.

Party; 4) Taiwan Solidarity Union; 5) New Party; 6) Taiwan Independence Party; 7) Other political party; 8) Pan-blue; 9) Pan-green.

Democratic Progressive Party, Taiwan Solidarity Union, and Taiwan Independence Party, which are so-called "Pan-green" parties, support Taiwan independence. Thus, we construct a dummy variable equal to one if the respondent chose these parties, zero otherwise. To examine the effect of curriculum reform on individuals' preferences over Taiwan independence, we estimate the equation (3.1) and use the above dummy variables as outcomes. Panel A and B of Table 3.6 displays the estimated effect of the new curriculum on people's preferences over Taiwan independence. RD estimates in this regard suggest that the curriculum reform does not induce people to support Taiwan independence or vote for a political party which is against unification. In contrast to the results for national identity, all estimated coefficients on *TextBook* in Panel A and B of Table 3.6 are quite small and statistically insignificant. There are two possible reasons why our results are distinct from the findings in Clots-Figueras and Masella (2013). First, Catalan reform is more comprehensive than the curriculum reform used in this paper. According to Clots-Figueras and Masella (2013), Catalan reform not only changed language use in class, but also modified course contents, which might affect more political outcomes. Second, the political situations of Taiwan and Catalonia are quite different, in that the former has her own army and sovereignty, while Catalonia belongs to Spain and has limited autonomy. Declaration of independence in Taiwan's context, is thus likely not as important as in Catalonia.

Finally, we argue that the military threat from China might explain the lack of increase of *unconditional* support of independence. People's subjective probabilities on high-stake events is an important input in determining observed political preferences in societies that face huge uncertainty at macro level. It's a non-trivial probability that declaration of Taiwan's independence would trigger a war between Taiwan and China. We utilize the data from the following survey question in TSCS to try to partial out the influence of military threat from China on people's support of independence:

- Some think that if the independence of Taiwan would not lead to war, we should declare independence. To what extent do you agree or disagree with this point of view? 1) Agree strongly; 2) Agree; 3) Disagree; 4) Disagree strongly.

We create a dummy variable which is equal to one if the respondent selects the first two alternatives, and zero otherwise. Panel C of Table 3.6 reports RD estimates using this dummy variable as an outcome. The estimation results suggest that people who studied new textbooks were on average 5 percentage points more likely to support Taiwan's independence in a hypothetical situation where the declaration of Taiwan's independence will not result in war. However, the estimate is not statistically significant.

Since the proportion of respondents whose answers are the first two alternatives accounts for more than 70%, we create a new outcome indicating only the first alternative, in order to capture the shift from modest support for independence to a strong one. In this case, the baseline mean for the share of individuals strongly agreeing with Taiwan's independence if it would not lead to war is only 13%. Our preferred estimate in Panel D of Table 3.6 suggests that new curricula significantly increase the likelihood of strongly supporting Taiwan's independence by around 13 percentage points (see Column (4)). Our research design rules out the possibility that the results are driven by differential malleability of perception about the possibility and the costs of a war. Taken together, exposure to a new textbook did not directly translate into higher *unconditional* support for independence, but it did translate into higher support in a hypothetical state of the world where the major cost of declaring independence is removed. This is consistent with our interpretation of the influence of textbooks: people have a clearer distinction between the two nations Taiwan and China.[34]

---

[34]In the Online Appendix 3.10.5, we also examine the impact of curriculum reform on the responses to questions regarding political participation. We look at the questions, such as "Do you agree that you have the power to affect governmental decisions?" or "How often do you discuss politics with your friends?" or "How often do you read political news in newspaper/TV/internet?" Our results suggest that new curriculum have small and insignificant effect on individuals' willingness for political participation.

3.6   Mechanisms

In this section, we explore the possible mechanisms through which school curricula (i.e. textbook contents) might affect an individual's national identity, by conducting subgroup analysis along two dimensions: Education track and the ethnic distribution of one's hometown. For each subgroup, we estimate equation (3.1) and conduct a similar RD analysis to that seen in section 3.5.1.

3.6.1   Memorization: Subgroup Analysis by Education Track

One possible channel through which school curricula might affect one's national identity is memorization. Students who paid more attention to studying their textbooks should be associated with higher treatment intensity, in the sense that they may have memorized more Taiwan-related texts. Specifically, we examine this mechanism by utilizing a subgroup analysis based on intensity of exposure to the new textbooks. The ideal proxy for this intensity is the grade of social subjects in the high-school entrance exam.[35] Unfortunately, the TSCS data does not include such information, so instead we use students' choice of education track to distinguish roughly between high and low levels of effort devoted to academic subjects in general.

After completing compulsory education, students in Taiwan are divided into two educational streams: The academic track and the vocational track. The choice of track is highly correlated with the effort students put in to studying when in junior high school. Students who were motivated to pursue more academic knowledge would have studied the textbooks far more, to give them a better chance of being selected by their preferred senior high school. On the other hand, common wisdom suggests that parents in Taiwan encourage students who lack motivation but are adept at obtaining excellent grades (for example, they memorize the material more quickly than the average person) to opt for the academic instead of the vocational track. Consequently, the education track implies something about the students' exposure to the content of textbooks.

---

[35]These junior high school graduates in our sample, no matter which education track they proceeded with, took the same national examination and used the grade they achieved to apply for senior or vocational high school.

We categorize the respondents into two groups: academic track and vocational track.[36]

Figure 3.7 displays the relationship between Taiwanese identity and birth cohorts by academic track (Figure 3.7a) and vocational track (Figure 3.7b) respondents. We observe a distinct jump around the cut-off in Figure 3.7a but no such pattern in Figure 3.7b. Table 3.7 presents RD estimates based on equation (3.1) for academic track respondents (Panel A) and vocational track respondents (Panel B), respectively. The estimates for academic track respondents suggest that new curricula significantly increase the probability of such students having a Taiwanese identity, by around 31 percentage points. In contrast, the results for vocational track students are small and statistically insignificant.

Following Ito (2015), we formally test the statistical significance of differences in the curriculum effect between the two subgroups by adding the interaction term *TextBook* and a dummy for the academic track students *Academic*.[37] Specifically, we estimate the following regression.

$$Identity_{it} = \delta_0 + \delta_1 Academic_i + \delta_2 TextBook_i + \delta_3 TextBook_i \times Academic_i$$
$$+ g(m;\beta) + X_i' \gamma + \mu_t + \varepsilon_{it} \tag{3.3}$$

Panel C of Table 3.7 display the estimated coefficient on *TextBook × Academic*. The result indicates that there is substantial and statistically significant heterogeneity in curriculum effect between academic and vocational tracks students. This subgroup analysis complements existing evidence provided by Cantoni et al. (2017) of a curriculum effect. Since Cantoni et al. (2017) conducted their survey at Peking University (i.e. an academic track school), their sample consisted of students who excelled at memorizing textbook materials. Thus, they could not tell whether the school curriculum would influence those who do not put too much effort on studying

---

[36]The academic track includes respondents whose final education level is senior high school or university. The vocational track includes respondents whose educational level is junior high school, vocational high school (including military school), and vocational university.

[37]Equation (3.3) also includes 1) interactions between a dummy for the academic track students *Academic* and a running variable and 2) interactions between *Academic* and survey year fixed effects.

the textbooks. Our results suggest that the effect of a curriculum varies substantially according to the degree of exposure to textbook content.

3.6.2 Prior Belief: Subgroup Analysis by Hometown Ethnicity Composition

According to "belief-based" models, people who possess less prior belief can be affected more by new information (DellaVigna and Gentzkow, 2010). In our case, this leads to a prediction that treatment effect is *decreasing* in the dimension of students' familiarity of Taiwanese identity prior to the exposure of the textbook. We proxy this familiarity by the ethnicity distribution of students' hometown and try different ways to present the treatment effect heterogeneity on this dimension.

As discussed earlier, ethnicity is correlated with Taiwanese identity. Due to historical reasons, Hoklo people hold the strongest Taiwanese identity, followed by Aborigines and Hakka (see Figure 3.20 in the Online Appendix). Since children may randomly pick up cultural ideas from parents or role models in the neighborhoods in which they live (Bisin et al., 2011a), a child growing up in a town with higher share of ethnicity who hold stronger Taiwanese identity is more likely to be exposed to Taiwanese identity and Taiwan-related knowledge before junior high school. Another source of Taiwanese identity exposure in such towns may come from daily political discussions (e.g., election campaigns). People living in towns with fewer Hoklo people would have been exposed to fewer Taiwan-oriented speeches, since politicians running for local elections have to cater to local people's political preferences, including those related to identity.

Our first approach to test the "belief-based" models focus on the distribution of Hoklo people, which hold the strongest Taiwanese identity historically, in a discrete way. We categorize students into two groups: People who lived in towns with high and low proportions of Hoklo people. The definition of towns with high (low) proportions of Hoklo people is that the share of Hoklo people in one's hometown is more (less) than the population median (77.1%). In Online Appendix 3.10.7, We provide a map demonstrating Hoklo ethnicity distribution and find the

105

towns with a high proportion of Hoklo people are located in the southern and western parts of Taiwan (see the white-colored area).

To show that this subgroup criterion distinguishes between local environments with different levels of Taiwanese identity, we utilize the 1992, 1995, 1998, and 2000 TSCS waves and calculate the mean of *Identity* in the towns with high and low shares of Hoklo people during different survey years.[38] Figure 3.8 suggests that people living in towns with low proportions of Hoklo people, on average, would be less likely to report themselves as Taiwanese than those living in towns with a low Hoklo count (i.e. around 10-15% less). This assures us that students living in these two types of area would have faced significantly different social environments in terms of issues regarding national identity when in junior high school and elementary school—the time when they would have absorbed this information from the environment in which they were living. We argue that the curriculum effect would have been greater for students living in the towns with a low proportion of Hoklo people, according to "belief-based" models, since they would have been less familiar with Taiwan-related knowledge beforehand.

Figure 3.9 displays the relationship between Taiwanese identity and birth cohorts separately for respondents living in the towns with a low Hoklo share (Figure 3.9a) and a high Hoklo share (Figure 3.9b). For the former group of respondents, Figure 3.9a suggests there is a substantial increase in Taiwanese identity at the cut-off. However, for the latter group, we find little evidence of any change in Taiwanese identity around the cut-off (see Figure 3.9b). Consistent with the graphical evidence, Panel A of Table 3.8 suggests that the new curriculum significantly increased the Taiwanese identity of respondents living in towns with a low proportion of Hoklo people but had little impact on those living in towns with a high Hoklo share (see Panel B of Table 3.8).

Similar to subgroup analysis by education track, we test the statistical significance of differences in the curriculum effect between two subgroups by estimating equation (3.3)

---

[38]In order to include those adults whom children are more likely to meet, we drop any respondents aged below 25.

but replacing the interaction term *TextBook* × *Academic* with *TextBook* × *LowHoklo*, where

LowHoklo is a dummy variable for individuals from the low-Hoklo area. Panel C of Table 3.8

reports the estimated coefficient on *TextBook* × *LowHoklo* and suggests that the difference in

the curriculum effect between individuals from the high-/low-Hoklo area is large and statistically

significant.

Since we categorize proportions of Hoklo people (i.e., continuous variables) into discrete

groups (i.e., hometowns with high and low Hoklo shares), people might be concerned that this

arrangement may be arbitrary. We examine treatment effect heterogeneity in RD designs using

the method proposed by Hsu and Shen (2019). Consistent with the above findings, Table 3.12 in

the Online Appendix suggests that we can reject the null hypothesis that the effect of exposing to

new curriculum on Taiwanese identity does not vary in line with the share of the Hoklo ethnic

group in the respondents' hometown (p-values are between 0.02 to 0.10).[39] In other words, the

curriculum effect is heterogeneous for individuals living in the area with different proportions of

Hoklo people.[40]

To explore other parts of the ethnicity distribution, we also do an analysis based on the

distribution of both Hoklo and Hakka, the two groups with stronger Taiwanese identity. We now

divide towns based on whether their added share of Hoklo and Hakka people is higher than the

population median (88.2%). The RD estimates shown in Table 3.13 of the Online Appendix

suggest that our results are robust for this grouping. Reading the new curriculum significantly

increased the Taiwanese identity of respondents living in neighborhoods with lower proportions

of Hoklo and Hakka people (see Panel A of Table 3.13) but had a small impact on those living in

---

[39]Specifically, we test whether conditional treatment effects estimated from different subgroups are all the same
as the treatment effect estimated from the whole sample. The construction of the subgroups is as follows. We first
set the largest number of subgroups (Q). Second, we form the subgroups by: 1) form Q subgroups which evenly
divide the hometown's Hoklo share, 2) form $Q - 1$ (q) subgroups which evenly divide the hometown's Hoklo share,
3) so on until q equals to one. For example, when Q equals to 4, we have 10 overlapping subgroups. We then collect
all these overlapping subgroups, estimate conditional treatment effects within each group, and test if all conditional
treatment effects equal to the average treatment effect. We tried three possible Q, which equals to two, three, and
four. We also try different bandwidth 24 months or 12 months. The p-value becomes bigger as Q increases. This is
reasonable since larger Q divides the sample into smaller subgroups.

[40]We cannot apply Hsu and Shen (2019)'s method to the curriculum effect by education track, which is a
categorical variable (i.e., vocational or academic track).

areas populated with larger Hoklo and Hakka ethnic groups (see Panel B of Table 3.13). Figure 3.23a and 3.23b show the corresponding RD graphs. Although the difference in the curriculum effect between the two subgroups is not statistically significant, the estimated magnitude is still substantial (i.e., 17 percentage points, see Panel C of Table 3.13). We think this result is reasonable, since the Hakka people do have a weaker Taiwanese identity. Therefore, the gap in prior belief between the two subgroups is smaller (see Figure 3.24 in the Online Appendix).[41]

The above analysis might be confounded with other social, political, and economic factors at the regional level. To deal with this concern, we exploit variations in the ethnicity of the respondents' parents, dividing respondents into those whose parents both have Hoklo ethnicity, and others. Figure 3.25 in the Online Appendix shows that the ethnic composition of parents is related to people's Taiwanese identity – individuals with at least one non-Hoklo parent are 20% less likely to report themselves as Taiwanese than people whose parents are both Hoklo. (18% v.s. 39%).[42] Consistent with the results based on regional ethnic distribution, we find that the new curriculum significantly increased the Taiwanese identity of respondents with at least one parent who was non-Hoklo (see Panel A of Table 3.14). The corresponding RD graph is displayed in Figure 3.26a. Note that there is an outlier (see the rightmost dot in Figure), which consists of only two individuals. Both did not have Taiwanese identity so that it is particularly negative compared to other dots. However, our result is robust to exclusion of these two respondents (see Panel B of Table 3.14). In contrast, the curriculum reform had little impact on respondents whose parents had Hoklo ethnicity (see Panel B of Table 3.14 and Figure 3.26b), thereby suggesting that the curriculum effect is greater for individuals with less prior information.

Two points should be noted about the exercise in this section. First, in our sample, the parents' ethnicity leans very heavily towards Hoklo, and the sample size for individuals with non-Hoklo parents is quite small. The results based on this subgroup analysis should be

---

[41]The differences in Taiwanese identity between areas with a low/high share of Hoklo and Hakka people are 5% to 10%, which is smaller than the result shown in Figure 3.8 using variations in the share of Hoklo people.

[42]Again, we use 1992, 1995, and 1998 TSCS data and restrict the sample to people who are 25 years old or above.

interpreted with caution. Therefore, we are more confident in the results looking at the hometown ethnicity dimension. Second, parents' ethnicity is highly correlated with hometown ethnicity distribution. Although we cannot clearly separate out the two dimensions, we do find that our evidence is strongly consistent with the belief-based models: proxied familiarity to Taiwanese identity is correlated with textbook treatment effect sizes.

3.7   Long-Term Results

Up to this point, we have found that the introduction of new textbooks can significantly increase students' Taiwanese identity when they are 18 to 23 years old (short-term sample). The natural question to ask, therefore, is whether or not the impact of the school curriculum was transitory or persistent. We explore this issue by examining the long-term sample, i.e. respondents who were surveyed during 2010 to 2015, when they were 24 to 33 years old (i.e. 11 to 20 years after reading textbooks). In the Online Appendix, Table 3.16 compares the characteristics of long-term sample with the ones of main sample. Since the each wave of survey is nationally representative, we find that characteristics are broadly comparable across survey years. One notable exception is age. The average age and of long-term sample are larger than those of short-term sample (i.e. main sample). But the difference in age is reasonable since long-term sample includes individuals who were older.[43]

Figure 3.10 displays the relationship between Taiwanese identity and birth cohorts for the long-term sample.[44] We find the mean level of Taiwanese identity to be quite similar on either side of the cut-off. Consistent with the graphical evidence, the regression results in Table 3.9 suggest the coefficients of *TextBook* are small and insignificant across all specifications, which are quite different from our main estimates. We find that in the long run, both old and new textbook readers hold similar levels of Taiwanese identity. In the Online Appendix 3.10.8 and

---

[43]The schooling years of parents are also different. However, the difference is small (around 5% differences) compared to baseline mean.

[44]As in our main results, we measure birth cohorts at the year-month level and plot average *Identity* after controlling for survey year fixed effects.

3.10.9, we examine the validity of the RD design for the long-term sample (see Tables 3.16 to 3.17 and Figure 3.29).[45] In addition, we conduct a series of robustness checks and find that our estimates are robust to different specifications (see Table 3.18 and Figure 3.30) and bandwidth choices (see Figure 3.28).[46]

Based on our research design, there are two possible interpretations for this finding. First, the likelihood of a Taiwanese identity among people who read the new textbooks (i.e. *Knowing Taiwan* series) "retreats" to the original level (i.e. that in the control group) in the long run. Second, the likelihood of a Taiwanese identity among people who read the old textbooks "catches up" with that for those who read the new textbooks in the long run. Figure 3.11 compares Taiwanese identity during 2003-2005 and 2010-2015 by treatment status. The result supports the second interpretation. We find that the probability of reporting as Taiwanese among old textbook readers (i.e. control group) catches up with that of people reading new textbooks (i.e. treatment group) during the sample period.[47]

One possible explanation is that although students who studied the old textbook would have weaker Taiwanese identity than students exposed to the new textbook in the short run, the old textbook readers might change their identity after receiving new information,[48] since the content of the old textbook substantially deviates from current situations and recent social trends. However, given our research design and data limitation, we are not able to verify this explanation directly.

---

[45]Based on the results in Tables 3.16 and 3.17, we find that the observable characteristics are fairly comparable between the treatment and control groups. Moreover, Figure 3.29 suggests that the density of the running variable (birth cohort) is quite smooth at cutoff.

[46]We also implement similar placebo tests shown in section 3.5.3 and find null effects.

[47]The difference in Taiwanese identity between two groups is 10 percentage points (61% v.s. 71%) in 2003-2005 but shrinks to 4 percentage points during 2010-2015 (79% v.s. 83%).

[48]For example, Taiwan has already gone through three presidential elections since 2008, with both Kuomintang and Democratic Progressive Party (i.e. two major political parties in Taiwan) won at least once. The successful experience of party alternation may also help build Taiwanese identity.

3.8 Conclusion

In this study, we have shown that school curricula (i.e. the content of textbooks) can shape an individual's national identity. By utilizing a textbook reform which introduced a new perspective on Taiwan's history for students entering junior high school after September 1997, we use a regression discontinuity design to isolate curriculum effects from other confounding factors. Our results suggest that people who studied new textbooks are on average 18 percentage points more likely to report themselves as Taiwanese than those studying old textbooks. The estimated effect is sizable and accounts for a 30% increase in the baseline mean. Moreover, our subgroup analysis indicates that the curriculum effects only appear in academic track students and those living in neighborhoods where fewer people identify as Taiwanese. Finally, we find that in the long run, both old and new textbook readers hold similar levels of Taiwanese identity since "old-textbook" effect is declined.

Our findings point towards some fruitful directions for future research. For example, we provide evidence aligned with "belief-based" models in the persuasion literature, but empirical evidence on whether people holding stronger or weaker prior beliefs are more affected by education policies is mixed. Voigtländer and Voth (2015), for instance, found that people who held a stronger prior anti-Semitic attitude were affected more by anti-Semitic indoctrination between 1933 and 1945 (i.e. they exhibited the largest increases in anti-Jewish attitudes). Why persuasion is effective in different subgroups under different contexts is a potential research question for the future. In addition, one limitation of our analysis is that we cannot pin down which key element of the 1997 curriculum reform leads to the estimated effect. Identifying the major component of the reform that raised Taiwanese identity is an important issue for future research.

3.9 Figures and Tables

3.9.1 Figures
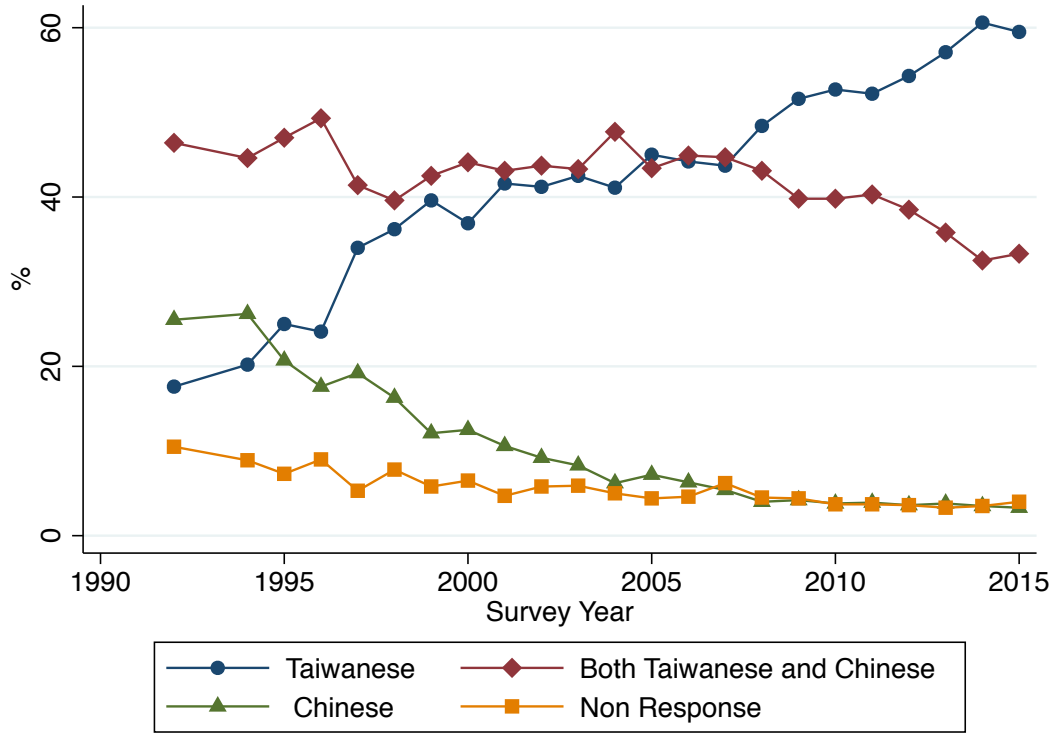
Figure 3.1. Trends of National Identity in Taiwan: 1992-2015

Source: Election Study Center, National Chengchi University
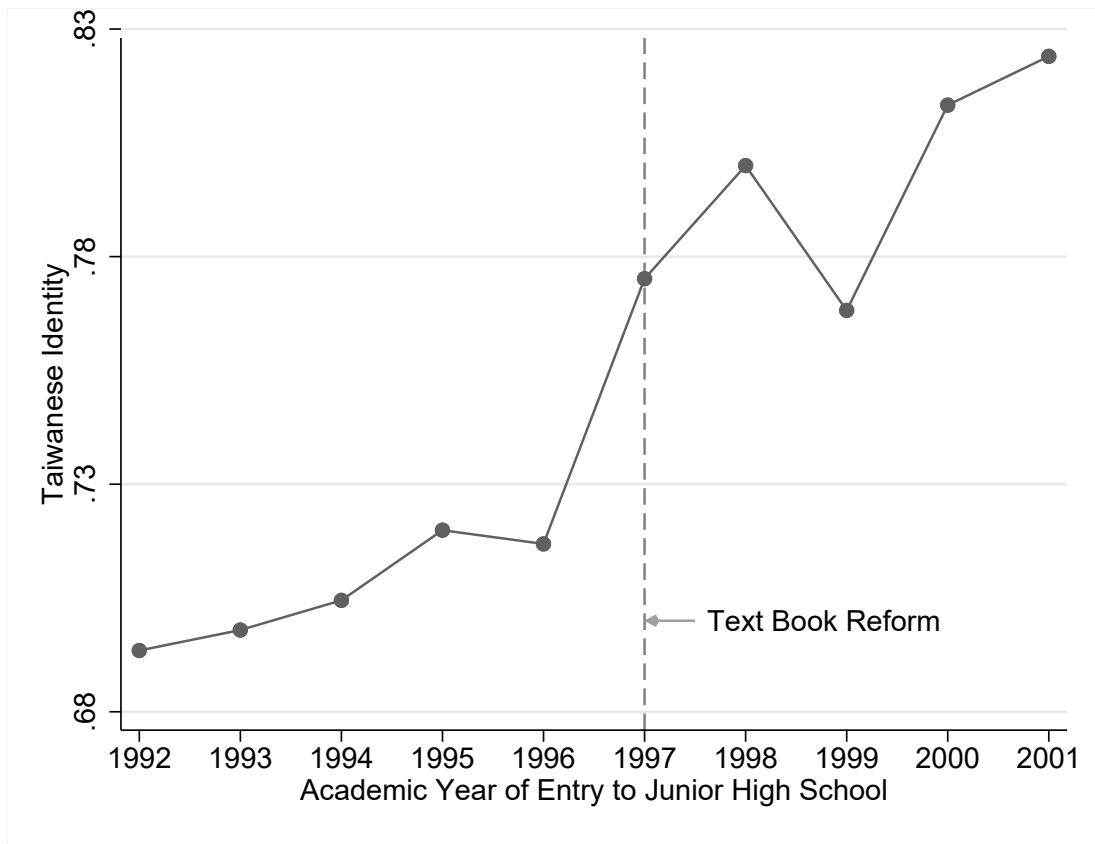
Figure 3.2. Taiwanese Identity and Education Cohorts

Notes: We pool all available TSCS data (i.e. 2003, 2004, 2005, 2009, 2010, 2012, 2013, 2014, and 2015 waves) and include education cohorts from 1992 to 2001. We include the 2009 wave, which is not included in our regression analysis, since we do not require demographic information to draw the graph. Taiwanese identity is measured by a dummy variable *Identity*. It assigns one to respondents answering "Taiwanese" and zero to those answering "Chinese" and "Both". Each dot represents average Taiwanese identity (*Identity*) for specific education cohorts.
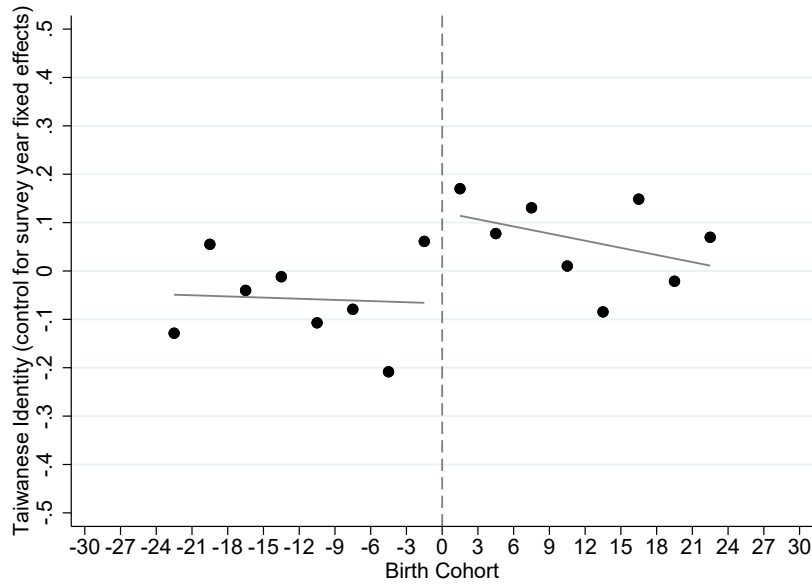
Figure 3.3. Taiwanese Identity and Birth Cohorts: Main Results

Notes: We pool data from 2003, 2004, 2005 TSCS and use the sample born between September 1982 and August 1986. We first regress *Identity* on survey year dummies and then collapse the residuals at birth year-quarter level (i.e. three birth year-month cohorts) to derive the dots. Thus, the first dot in this figure represents average *Identity* (after controlling for the survey year fixed effect) for those born in September, October, and November 1982 and the last dot represents average *Identity* (after controlling for the survey year fixed effect) for those born in June, July, and August 1986. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.

Figure 3.4. RD Estimates across Different Bandwidth Choices

Notes: We run regressions as column (4) in Table 3.2 with different bandwidths: 12 to 24 months on each side of the cut-off, i.e., two education cohorts. The solid line represent the point estimates of coefficients on the *TextBook* dummy variable and the dotted line represents the corresponding 95% confidence interval derived from standard errors clustered at birth year-month level.

(a) 1983 September Cutoff

(b) 1982 September Cutoff

(c) 1981 September Cutoff

(d) 1980 September Cutoff

Figure 3.5. RD Graph for Placebo Tests

Notes: We pool the 2003, 2004, and 2005 TSCS data. We first regress *Identity* on survey year dummies and then collapse the residuals at birth year-quarter level (i.e. three birth year-month cohorts) to derive the dots. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable. Figure 3.5a uses the sample born between August 1981 and September 1985; Figure 3.5b uses the sample born between August 1980 and September 1984; Figure 3.5c uses the sample born between August 1979 and September 1983. Figure 3.5d uses the sample born between August 1978 and September 1982.

Figure 3.6. Permutation Test

Notes: We pool data from the 2003, 2004, and 2005 TSCS waves and assign the fake reform to all possible months and years – from January 1950 to September 1983 (405 fake reforms). This figure display the distribution of placebo estimates (see the histogram) and compare them with our main RD estimate (see the dash line).

(a) Academic Track

(b) Vocational Track
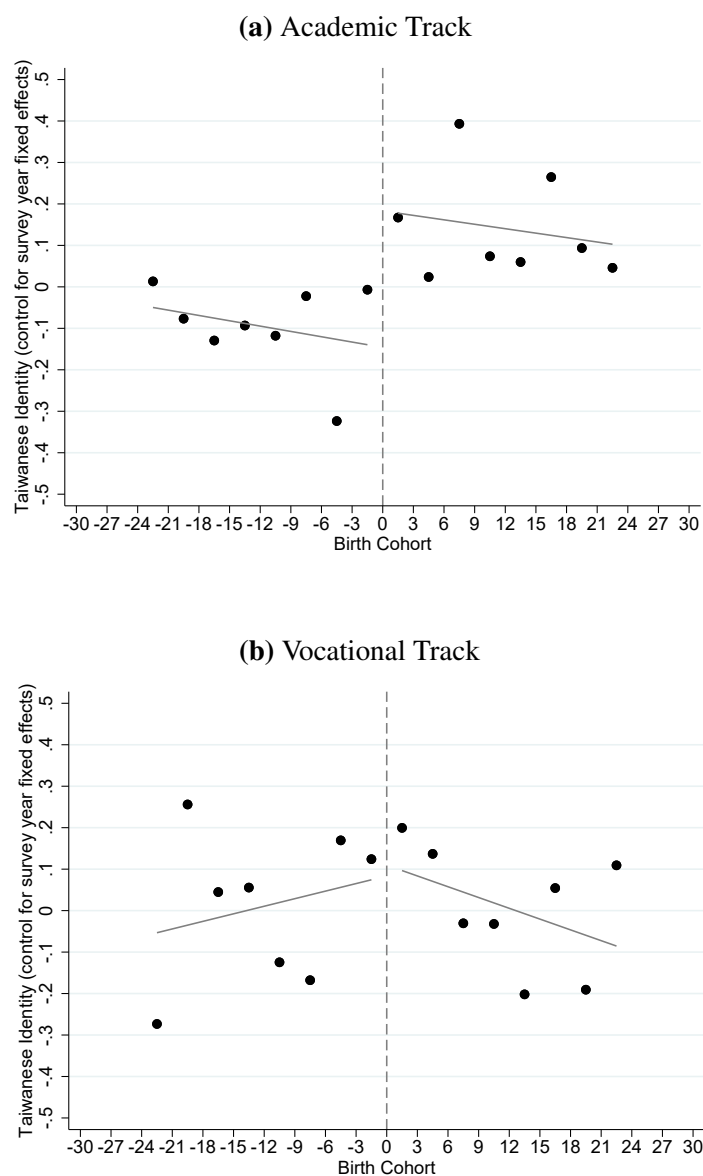
Figure 3.7. Taiwanese Identity and Birth Cohorts: By Education Track
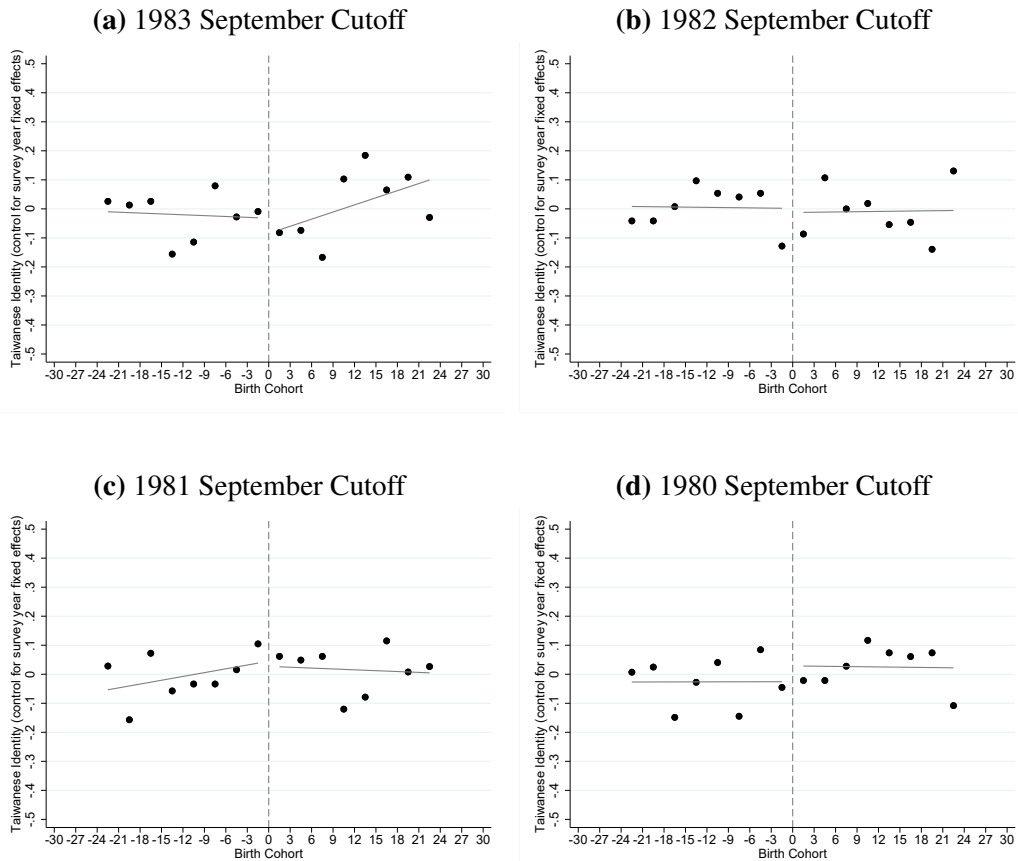
Notes: We pool data from 2003, 2004, 2005 TSCS and use the sample born between September 1982 and August 1986. Figure 3.7a includes respondents whose final education level is senior high school or university. Figure 3.7b includes respondents whose educational level is junior high school, vocational high school, and vocational university. We first regress *Identity* on survey year dummies and then collapse the residuals at birth year-quarter level (i.e. three birth year-month cohorts) to derive the dots. Thus, the first dot in this figure represents average *Identity* (after controlling for the survey year fixed effect) for those born in September, October, and November 1982 and the last dot represents average *Identity* (after controlling for the survey year fixed effect) for those born in June, July, and August 1986. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.
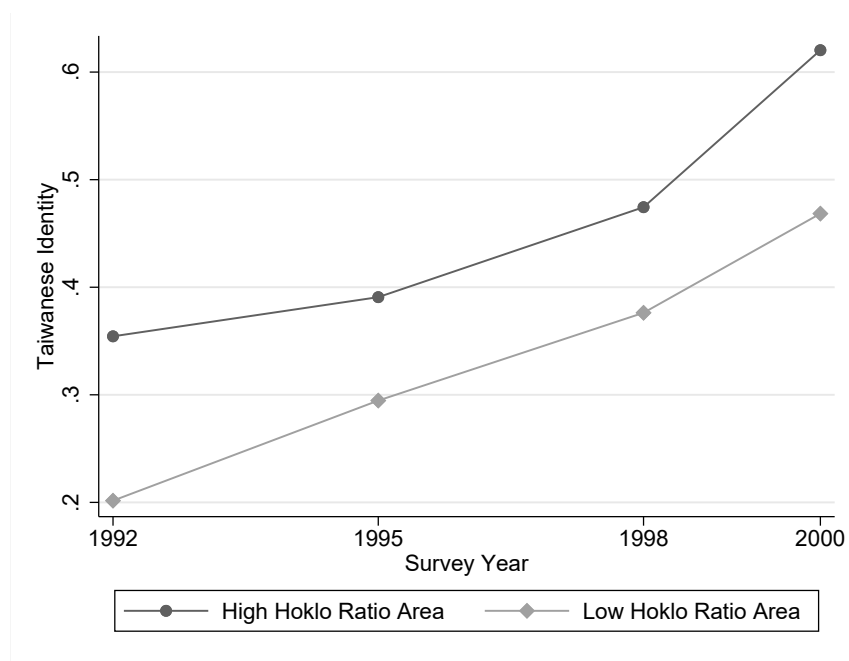
Figure 3.8. Taiwanese Identity Trend in Townships with High/Low Share of Hoklo People in 1990s

Notes: We pool data from 1992, 1995, 1998 and 2000 TSCS waves. In order to include those adults whom children are more likely to meet, we restrict the respondents aged 25 or above. Each dot represents share of people reporting Taiwanese identity in given survey year and area. The circle symbol represents the area with high share of Hoklo people. The diamond symbol represents the area with low share of Hoklo people.

(a) Area with Low Share of Hoklo People

(b) Area with High Share of Hoklo People
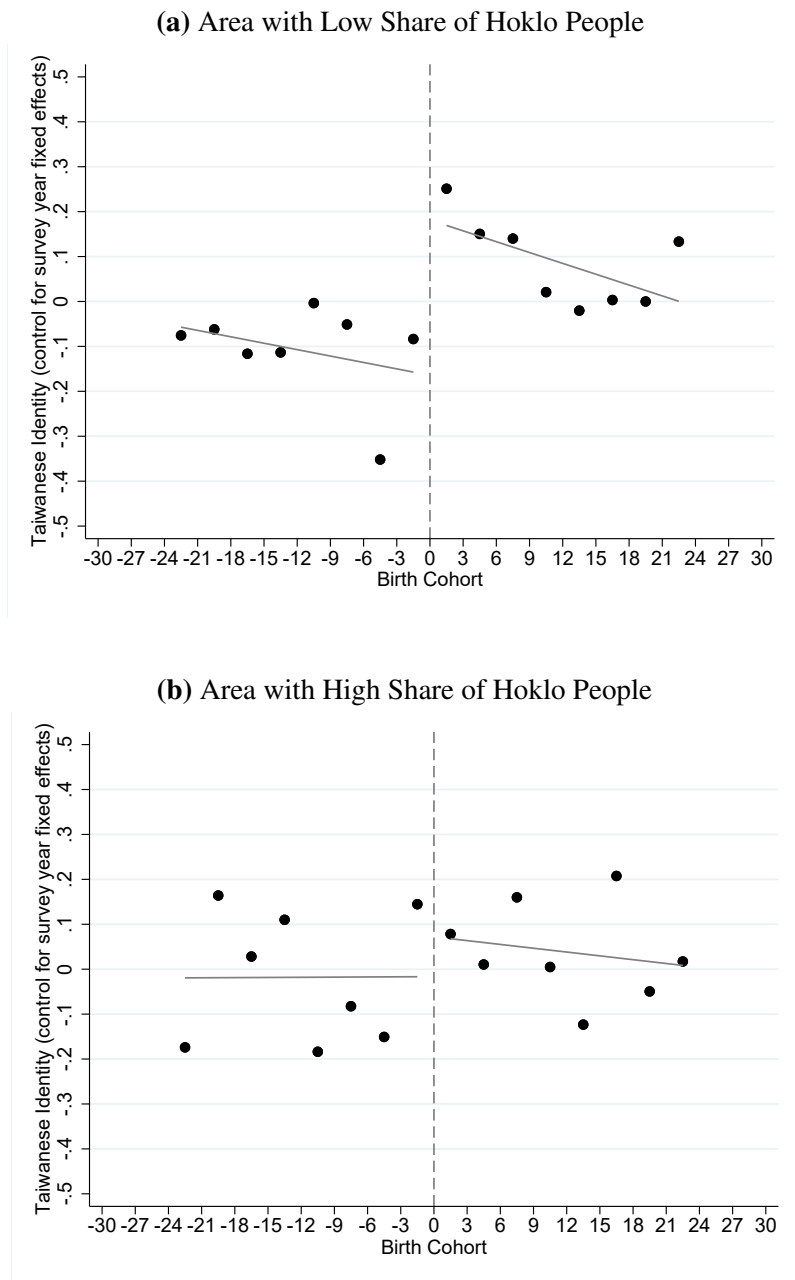
Figure 3.9. Taiwanese Identity and Birth Cohorts: By High/Low *Hoklo Proportion* Areas

Notes: We pool data from 2003, 2004, 2005 TSCS and use the sample born between September 1982 and August 1986. Figure 3.9a includes respondents living in the towns with low share of Hoklo people before age 15. Figure 3.9b includes respondents living in towns with high share of Hoklo people before age 15. We first regress *Identity* on survey year dummies and then collapse the residuals at birth quarter level to derive the dots. Thus, zero in the figure represents September, October, and November 1984. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.
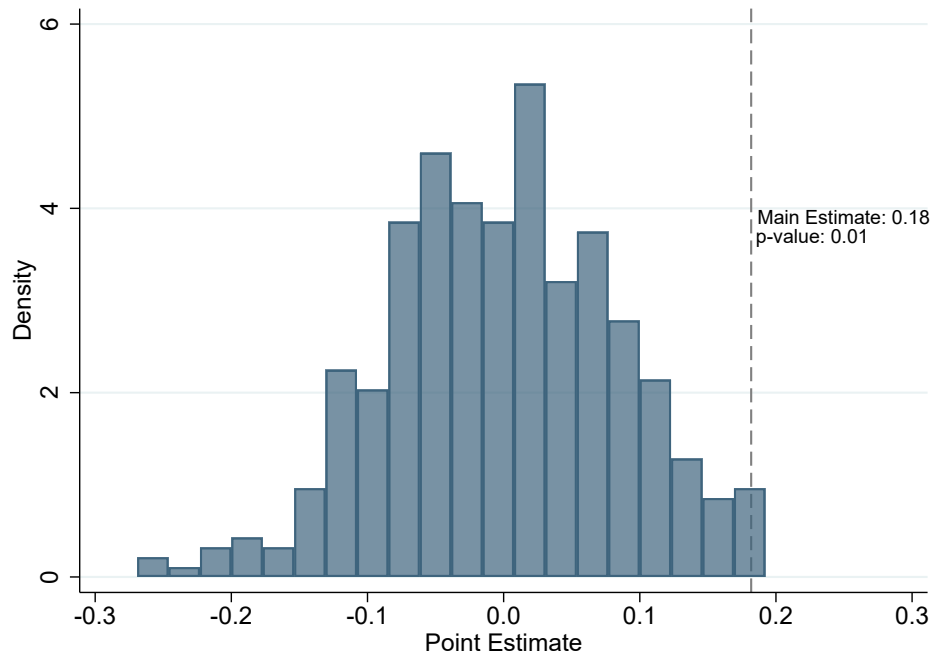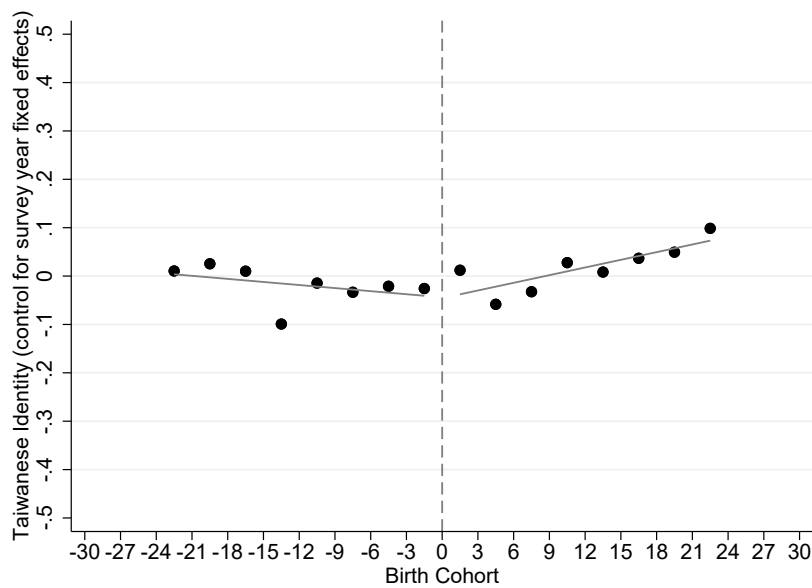
Figure 3.10. Taiwanese Identity and Birth Cohorts: Long-Term Results

Notes:We pool data from 2010, 2012, 2013, 2014 and 2015 TSCS waves and use the sample born between September 1982 and August 1986. We first regress *Identity* on survey year dummies and then collapse the residuals at birth year-quarter level (i.e. three birth year-month cohorts) to derive the dots. Thus, the first dot in this figure represents average *Identity* (after controlling for the survey year fixed effect) for those born in September, October, and November 1982 and the last dot represents average *Identity* (after controlling for the survey year fixed effect) for those born in June, July, and August 1986. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.

Figure 3.11. Trend in Taiwanese Identity: By Treatment Status

Notes: We pool data from 2003, 2004, 2005, 2010, 2012, 2013, 2014 and 2015 TSCS and use the sample born between September 1982 and August 1986. Each bar represents simple mean of *Identity* during 2003-2005 or 2010-2015 by treatment status. Control group includes 1995 and 1996 education cohorts and treatment group includes 1997 and 1998 education cohorts.

Table 3.1. Descriptive Statistics for Treatment Group and Control Group

	Born after September 1984	Born before September 1984	Difference (after - before)
Female	0.445	0.445	0.000
	(0.498)	(0.498)	(0.049)
Age	19.578	20.954	-1.376***
	(0.659)	(0.995)	(0.080)
Years of schooling (self)	13.566	13.894	-0.327
	(2.130)	(1.936)	(0.202)
Years of schooling (father)	10.827	10.445	0.382
	(3.246)	(3.580)	(0.334)
Years of schooling (mother)	10.075	9.760	0.315
	(3.424)	(3.232)	(0.330)
Proportion of Hoklo in the hometown	0.711	0.734	-0.023
	(0.227)	(0.206)	(0.022)
Hoklo father	0.786	0.768	0.018
	(0.411)	(0.423)	(0.041)
Hoklo mother	0.827	0.823	0.004
	(0.38)	(0.383)	(0.038)
# of individuals	173	254	

Notes: We pool data from the 2003, 2004, and 2005 TSCS waves and use the sample born between September 1982 and August 1986. The definitions of the individual characteristics are as follows: 1) Female: If an individual is female assigned 1, otherwise 0. 2) Respondent/Father/Mother's schooling years: a) no education (zero years of schooling); b) elementary school (6 years of schooling); c) junior high school (9 years of schooling); d) senior (vocational) high school (12 years of schooling); e) two-year college (14 years of schooling); f) University or vocational university (16 years of schooling). 3) Hoklo fathers/mothers: If an individual's father/mother is Hoklo assigned 1, otherwise 0. In the Online Appendix 3.10.4, we provide detailed definition of proportion of Hoklo people in the hometown. Standard deviations in parentheses, and standard errors in brackets. *** significant at the 1 percent level, ** significant at the 5 percent level, and * significant at the 10 percent level.

Table 3.2. The Effects of the Curriculum Reform on Taiwanese Identity: Main Results

	Taiwanese Identity			
	(1)	(2)	(3)	(4)
TextBook	0.162**	0.173**	0.183**	0.182**
	(0.080)	(0.081)	(0.082)	(0.084)
Baseline Mean	0.608	0.608	0.608	0.608
Persuasion Rate	41.3	44.1	46.6	46.4
Sample Size	427	427	427	427
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool data from the 2003, 2004, and 2005 TSCS waves and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. All columns include the survey year fixed effect and the first-order polynomials of birth year-month m interacting fully with *TextBook* (i.e. linear spline). Column (2) adds the ethnic variables, such as parents' ethnicity and share of Hoklo in the hometown. For parents' ethnicity, we include a set of dummy variables indicating a respondent's father/mother is Mainlanders, Hakka, Aboriginal and Other. We use Hoklo as a reference group. Column (3) further includes demographic variables, such as gender, fathers'/mothers' education level. For fathers'/mothers' education level, we include a set of dummy variables indicating a respondent's father's/mother's highest degree is junior high school, senior high school, vocational high school, college, university, military school. We use elementary school as a reference group. Column (4) adds a set of dummy variables indicating the region where an respondent lived in before age 15. There were 23 county/city in Taiwan during the sample period. We categorize them into four regions: northern, middle, southern, and eastern regions. We use the eastern region as a reference group. The baseline mean is the simple average of *Identity* of respondents born between September 1982 and August 1984. Standard errors are clustered at birth year-month level in parentheses. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

Table 3.3. Robustness Check: Different Specification, Sample, Identity Questions

	Taiwanese Identity			
	(1)	(2)	(3)	(4)
Panel A: 2nd Order Polynomial				
TextBook	0.186	0.199*	0.188	0.209*
	(0.112)	(0.116)	(0.120)	(0.120)
Sample Size	427	427	427	427
Panel B: Exclude Specific Birth Cohorts				
TextBook	0.123	0.131	0.136*	0.133*
	(0.075)	(0.078)	(0.076)	(0.076)
Sample Size	408	408	408	408
Panel C: Identity Question – 2003-2004 Waves				
TextBook	0.179*	0.196*	0.211*	0.240**
	(0.098)	(0.106)	(0.106)	(0.113)
Observations	243	243	243	243
Panel D: Identity Question – 2005 Wave				
TextBook	0.207	0.193	0.225	0.190
	(0.140)	(0.145)	(0.154)	(0.156)
Observations	184	184	184	184
Linear/Quadratic Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool data from the 2003, 2004, and 2005 TSCS waves and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. All panels include the survey year fixed effect and the first-order polynomials of birth year-month m interacting fully with *TextBook* (i.e. linear spline). Panel A additionally includes quadratic spline. Other covariates are the same as in the corresponding columns in Table 3.2. Standard errors are clustered at the birth year-month level in parentheses. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

Table 3.4. Robustness Check: Validity of Regression Discontinuity Design

VARIABLES	Female	Age	Years of Schooling	Father's Schooling years	Mother's Schooling years	Share of Hoklo people	Hoklo Fathers	Hoklo Mothers
TextBook	0.044	0.002	-0.219	0.015	0.115	-0.075*	-0.017	-0.053
	(0.065)	(0.005)	(0.250)	(0.599)	(0.431)	(0.043)	(0.076)	(0.059)
Sample Size	427	427	427	427	427	427	427	427

Notes: We pool data from the 2003, 2004, and 2005 TSCS waves and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Note that we do not include any covariates X_i since they are outcome variables now. All columns include the survey year fixed effect and the first-order polynomials of birth year-month m interacting fully with *TextBook* (i.e. linear spline). The definitions of the individual characteristics are as follows: 1) Female: If an individual is female assigned 1, otherwise 0. 2) Respondent/Father/Mother's schooling years: a) no education (zero years of schooling); b) elementary school (6 years of schooling); c) junior high school (9 years of schooling); d) senior (vocational) high school (12 years of schooling); e) two-year college (14 years of schooling); f) University or vocational university (16 years of schooling). 3) Hoklo fathers/mothers: If an individual's father/mother is Hoklo assigned 1, otherwise 0. In the Online Appendix 3.10.4, we provide detailed definition of proportion of Hoklo people in the hometown. Standard errors are clustered at the birth year-month level in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

Table 3.5. Robustness Check: Placebo Test of Fake Textbook Reform

	Taiwanese Identity			
	(1)	(2)	(3)	(4)
Panel A: Use September 1983 as Cut-Off				
TextBook	-0.055	-0.070	-0.049	-0.043
	(0.073)	(0.075)	(0.079)	(0.076)
Sample Size	487	487	487	487
Panel B: Use September 1982 as Cut-Off				
TextBook	-0.009	-0.006	-0.027	-0.026
	(0.092)	(0.100)	(0.094)	(0.094)
Sample Size	509	509	509	509
Panel C: Use September 1981 as Cut-Off				
TextBook	-0.018	-0.023	0.012	0.006
	(0.069)	(0.071)	(0.073)	(0.074)
Sample Size	519	519	519	519
Panel D: Use September 1980 as Cut-Off				
TextBook	0.052	0.062	0.006	0.003
	(0.080)	(0.076)	(0.088)	(0.088)
Observations	506	506	506	506
Panel E: Difference-in-Differences Design				
$AfterSep \times B_{1984}$	0.150	0.164	0.171	0.176*
	(0.103)	(0.104)	(0.106)	(0.104)
Observations	656	656	656	656
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2003, 2004, and 2005 TSCS data. Panel A uses the sample born between August 1981 and September 1985; Panel B uses the sample born between August 1980 and September 1984; Panel C uses the sample born between August 1979 and September 1983. Panel D uses the sample born between August 1978 and September 1982. The above table reports the coefficient of *TextBook* based on equation (3.1). In each placebo test, we define dummy variable *TextBook* as respondents born after following cutoffs: September 1983 (Panel A), September 1982 (Panel B), September 1981 (Panel C), or September 1980 (Panel D). Panel E reports the coefficients of $AfterSep \times B_{1984}$ in the equation (3.2). In this specification, we combine all available cutoffs used in the main estimation and placebo tests to implement a DID design. Specifications in each column are the same as in the corresponding columns in Table 3.2. Note that in DID design, we allow the linear spline of running variable to be cohort-specific. Standard errors are clustered at the birth year-month level in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

Table 3.6. Effects of the Curriculum Reform on Preferences for Taiwan's Independence

	(1)	(2)	(3)	(4)
Panel A: Support Taiwan Independence Unconditionally				
TextBook	0.037	0.048	0.056	0.056
	(0.036)	(0.036)	(0.037)	(0.040)
Baseline Mean	0.107	0.107	0.107	0.107
Sample size	424	424	424	424
Panel B: Support Parties that Prefer Taiwan Independence				
TextBook	-0.103	-0.084	-0.078	-0.074
	(0.080)	(0.073)	(0.081)	(0.082)
Baseline Mean	0.353	0.353	0.353	0.353
Sample size	354	354	354	354
Panel C: Support Taiwan Independence if There is no War				
TextBook	0.027	0.035	0.044	0.045
	(0.050)	(0.051)	(0.060)	(0.063)
Baseline Mean	0.714	0.714	0.714	0.714
Sample size	422	422	422	422
Panel D: Strongly Support Taiwan Independence if There is no War				
TextBook	0.109	0.111	0.127*	0.131*
	(0.066)	(0.069)	(0.068)	(0.068)
Baseline Mean	0.131	0.131	0.131	0.131
Sample size	422	422	422	422
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2003, 2004, and 2005 TSCS data and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Panel A displays the results for the preferences on Taiwan independence. Panel B displays results for the preference on the parties supporting Taiwanese independence. Panel C displays the results for the preference on Taiwanese independence under the condition that the independence of Taiwan would not lead to war. Panel D displays the results for the strong preference on Taiwanese independence under the condition that the independence of Taiwan would not lead to war. Specifications in each column are the same as in the corresponding columns in Table 3.2. The baseline mean is the simple average of outcomes of respondents born between September 1982 and August 1984 in the corresponding subgroup. Standard errors clustered at the birth year-month level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table 3.7. Subgroup Analysis: By Education Track

	Taiwanese Identity			
	(1)	(2)	(3)	(4)
Panel A: Academic Track				
TextBook	0.279**	0.293**	0.309**	0.308**
	(0.113)	(0.115)	(0.137)	(0.138)
Baseline Mean	0.583	0.583	0.583	0.583
Sample Size	219	219	219	219
Panel B: Vocational Track				
TextBook	0.055	0.069	0.052	0.040
	(0.104)	(0.104)	(0.109)	(0.114)
Baseline Mean	0.639	0.639	0.639	0.639
Sample Size	208	208	208	208
Panel C: Test Heterogeneity				
TextBook	0.055	0.059	0.028	0.017
	(0.104)	(0.101)	(0.096)	(0.095)
TextBook × Academic	0.225	0.239	0.307*	0.323**
	(0.163)	(0.158)	(0.164)	(0.160)
Sample Size	427	427	427	427
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2003, 2004, and 2005 TSCS data and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Panel A includes respondents whose final education level is senior high school or university. Panel B includes respondents whose educational level is junior high school, vocational high school (including military school), and vocational university. Specifications in each column are the same as in the corresponding columns in Table 3.2. The baseline mean is the simple average of *Identity* of respondents born between September 1982 and August 1984 in the corresponding subgroup. Panel C tests the statistical significance of the difference in curriculum effect between two subgroups by showing the coefficient on the interaction term of *TextBook* and a dummy for the academic track students *Academic*. Compared to equation (3.1), this specification also includes 1) a dummy for the academic track students *Academic*; 2) the interaction term of *TextBook* and *Academic*; 3) the interactions between *Academic* and running variable; 4) the interactions between *Academic* and survey year fixed effects. Standard errors clustered at the birth year-month level are in parentheses. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

Table 3.8. Subgroup Analysis: By Hometown Ethnicity Distribution

	Taiwanese Identity			
	(1)	(2)	(3)	(4)
Panel A: Hometown with Low Hoklo Proportion				
TextBook	0.328***	0.343***	0.381***	0.351***
	(0.113)	(0.120)	(0.122)	(0.123)
Baseline Mean	0.559	0.559	0.559	0.559
Sample Size	193	193	193	193
Panel B: Hometown with High Hoklo Proportion				
TextBook	0.024	0.008	0.030	0.057
	(0.103)	(0.107)	(0.101)	(0.104)
Baseline Mean	0.65	0.65	0.65	0.65
Sample Size	234	234	234	234
Panel C: Test Heterogeneity				
TextBook	0.024	0.013	0.016	0.028
	(0.103)	(0.109)	(0.103)	(0.104)
TextBook × *LowHoklo*	0.304*	0.338**	0.349**	0.309**
	(0.152)	(0.164)	(0.159)	(0.152)
Sample Size	427	427	427	427
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2003, 2004, 2005 TSCS data and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Panel A includes respondents whose hometown has a lower proportion of Hoklo people compared to the median of the population in the National Hakka Population Basic Information Survey Research, while Panel B includes respondents whose hometown has higher proportion of Hoklo people. Specifications in each column are the same as in the corresponding columns in Table 3.2. The baseline mean is the simple average of *Identity* of respondents born between September 1982 and August 1984 in the corresponding subgroup. Panel C tests the statistical significance of difference in curriculum effect between two subgroups by showing coefficient on the interaction term of *TextBook* and a dummy for the individuals from area with low share of Hoklo people *LowHoklo*. Compared to equation (3.1), this specification also includes 1) a dummy for the individuals from low-Hoklo area *LowHoklo*; 2) the interaction term of *TextBook* and *LowHoklo*; 3) the interactions between *LowHoklo* and running variable; 4) the interactions between *LowHoklo* and survey year fixed effects. Standard errors clustered at birth year-month level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.9. The Effects of the Textbook Reform on Taiwanese Identity: Long-Term Results

	Taiwanese Identity			
	(1)	(2)	(3)	(4)
TextBook	0.006	0.023	0.014	0.011
	(0.052)	(0.052)	(0.050)	(0.049)
Baseline Mean	0.794	0.794	0.794	0.794
Sample Size	822	822	822	822
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2010, 2012, 2013, 2014, and 2015 TSCS waves and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Specifications are the same as in Table 3.2. The baseline mean is the simple average of *Identity* of respondents born between September 1982 and August 1984. Clustered standard errors at birth year-month level are in parentheses.*** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

3.10 Appendix

3.10.1 1997 Curriculum Reform: More Information

In Taiwan, junior high school students need to learn three subjects for social studies: History, Geography, and Society. They also have to take a six-semester course for each subject. Under the old curriculum of social studies (i.e. before the 1997 curriculum reform), the subject History focused on the history of China (including Taiwan) during the first three semesters, and world history during the fourth to sixth semesters. Geography focused on the geography of China (including Taiwan) during the first two years, and world geography during the third year, while the Society course covered basic sociology, political science, and economics—all of which students learned from their first year to the third year.

The 1997 curriculum reform substantially increased Taiwan-related content. Students in their first year had to read the *Knowing Taiwan* series, which included three volumes: History, Geography, and Society. Therefore, they learned the history of Taiwan (i.e. History volume of the *Knowing Taiwan* series) during their first year and studied the history of China and of the world during their second and third years. Similarly, students studied the geography of Taiwan (i.e. Geography volume of the *Knowing Taiwan* series) during first year, then the geography of China and East Asia during the second year, and world geography during the third year. Instead of studying Society, students read the Society volume of the *Knowing Taiwan* series to obtain knowledge about Taiwan's social values, culture, and religions in their first year, and then learned the Civics subject in their second and third years.

3.10.2   Comparison of Textbooks: Sample Paragraphs from Old and New Textbooks

Figure 3.12 and 3.13 display the table of contents of the old history textbook. As mentioned in section 3.2.2, the old History textbook dedicated 25 chapters to Chinese history and only one chapter and a section to Taiwan. Figure 3.12 shows a section in the 15th chapter, entitled "The rebellion of Koxinga against the Qing Dynasty and the development of Taiwan," which described how a former courtier of the Ming Dynasty, Zheng Cheng-Gong, rebelled against the Qing Dynasty. Figure 3.13 shows the 25th chapter, entitled "The achievement and vision of a base for revival," which is the last chapter of Chinese history and described how Kuomintang developed Taiwan as a base for recovering China. In contrast, Figure 3.14 displays the table of contents of the *Knowing Taiwan* series—History volume. It has eleven chapters, and each chapter describes how the ancestors of different ethnic groups made developments in Taiwan: From the prehistoric era until the 16th century (Chapter 2), Dutch and Spanish rule (Chapter 3), Kingdom of Tungning (Chapter 4), Qing rule (Chapters 5-6), Japanese rule (Chapters 7-8), and Republic of China rule (Chapters 9-11).

As mentioned in section 3.2.2, the new textbook not only contained a substantial increase in content about Taiwan, but it also clearly distinguished between the concepts of Taiwan and China, in a contextual change. In general, the old textbook treated Taiwan as a part of the history of China, but the new textbook treated Taiwan's past as an independent entity. Figure 3.15 shows a sample paragraph from the old textbook which described that "our country" originated from Peking Man and cavemen. Here, "our country" clearly refers to China in this context. Figure 3.16 displays a similar paragraph in the new textbook on Chinese history, which explicitly mentions "China" as originating from Peking Man and cavemen.

Figure 3.17 shows another sample paragraph from the old textbook which described the economic development of "our country" after the 1949 Chinese Civil War (i.e. the Kuomintang-Communist Civil War). In this context, "our country" actually refers to Taiwan. Figure 3.18

displays a similar paragraph mentioning the economic development of Taiwan after 1949 in the History textbook in the *Knowing Taiwan* series. The authors explicitly used the term "Taiwan" rather than "our country".

Figure 3.12. Table of Contents from the Old History Textbook

Figure 3.13. Table of Contents from the Old History Textbook

目　次

認識臺灣（歷史篇）

Figure 3.14. Table of Contents from the New History Textbook

大概的活動情形了。

史前時代的人類，先是體質特徵介於人形猿和現代人之間的「猿人」，後有「真人」的出現。我們從地下發現的史前人類化石之中，知道了很多猿人和真人的情形。目前已經發現的猿人，比較有名的至少有四種：一是在非洲東岸發現的「東非人」；二是在印尼發現的「爪哇人」；三是在我國發現的「北京人」；四是在德國發現的「尼安德人」。在各種真人當中，有兩種比較著名：一是在我國發現的「山頂洞人」；另一是在法國發現的「克魯麥囊人」。猿人的形態比起現代人尚有一段距離，而真人則已具有現代人的形態了。

我國歷史的特色

我國歷史悠久，是東亞文明的主體，在世界歷史上占有重要的地位。歸納起來，其特色有下列四點：

一、時間悠久 我國的歷史，從黃帝建國算

Figure 3.15. Old Textbook: Our Country (China)

用火取暖、照明和燒烤食物。

　　至於中國境內舊石器時代晚期的人類，則以「山頂洞人」為代表。「山頂洞人」距今約兩萬年，體質已和現代人差不多。他們已知埋葬死者，還會用獸骨作成骨針，用獸齒製成裝飾品，生活比「北京人」進步得多。

　　由這些舊石器時代人類化石的發現，可知中國是人類的主要起源地之一；但他們和現代中國人有無直接關係，目前仍無法確定。要追究中國文化的源頭，比較可靠的線索是新石器時代的考古發現。

「北京人」頭蓋骨
從復原的「北京人」頭蓋骨化石可知，「北京人」兩眉相連，前額低平，腦容量不如現代人。

Figure 3.16. New Textbook: China

140

破。

重大建設：六十二年十一月，行政院院長蔣經國鄭重宣布：政府除積極興建核能發電廠外，並限期五年內完成南北高速公路、臺中港、北迴鐵路、蘇澳港、石油化學工業、大煉鋼廠、大造船廠、鐵路電氣化及桃園國際機場等重要建設；合稱十大建設，皆陸續完成。六十八年，政府又推動交通、工業、農業等十二項建設。七十三年，又推出十四項重要建設計畫，多為前述十大、十二項建設的延續，具有前瞻性的大工程。

目前，我國由於雄厚的工業基礎與外貿潛力，躋身「亞洲四小龍」之列，經濟方面的成就，已為舉世所公認。

文教建設　普及教育方面：民國三十九學年度，臺灣地區六至十一歲學齡兒童的就學率為百分之八十左右；至七十九學年度業已近於百分之

Figure 3.17. Old Textbook: Our Country (Taiwan)

圖10-7
花蓮國際港開放

化方面，積極致力於對外貿易、金融、產業經營的自由化。具
體作法分別為解除進口管制、大幅降低關稅稅率、取消銀行利
率的管制、大幅放寬外匯管制①、開放民間設立銀行，以及推
動公營事業民營化等。

國際化方面，具體作法包括放寬外國公司在臺投資、設立
臺灣境外金融中心、致力使新臺幣國際化等。近年又籌設亞太
營運中心，期使臺灣成為亞太地區的運輸、金融、資訊的重鎮。

Figure 3.18. New Textbook: Taiwan

3.10.3   Additional Figures
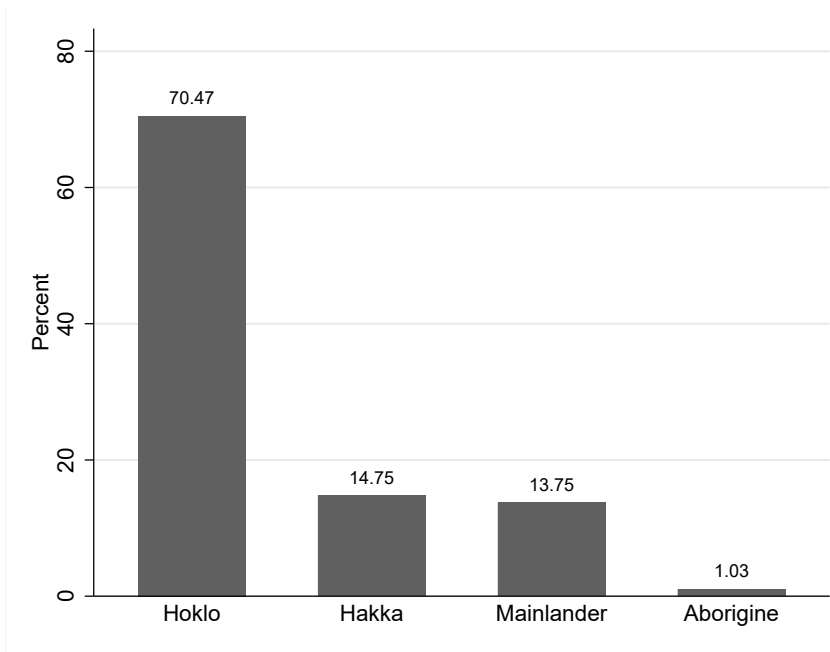
Figure 3.19. Distribution of Four Ethnic Groups in Taiwan

Notes: This figure displays the distribution of four major ethnic groups in Taiwan. Data is from 1992, 1995, and 1998 TSCS waves. We restrict sample to people who are 25 years old or above. The ethnic group that a respondent belongs to is based on his/her father's origins.

3.10.4 Definitions of Individual Characteristics

The definitions of the individual characteristics are as follows: 1) Female: If an individual is female assigned 1, otherwise 0. 2) For parents' education level, we include a set of dummy variables indicating a respondent's father's/mother's highest degree is junior high school, senior high school, vocational high school, college, university, military school. We use elementary school as a reference group. 3) For parents' ethnicity, we include a set of dummy variables indicating a respondent's father/mother is Mainlanders, Hakka, Aboriginal and Other. We use Hoklo as a reference group.

For the proportion of Hoklo in the respondents' hometown, we use a question from the TSCS that reads as follows: "Where did you live longest before you were 15 years old?" The responses are on the township (zip code) level, and so we regard them as reflecting where the respondents lived when in junior high school (i.e., their hometown). This hometown information is combined with township-level ethnicity data to approximate how many Hoklo people the respondents were surrounded by in their daily lives before senior high school. As mentioned in the main text, compared to other ethnic groups, Hoklo people are more likely to consider themselves Taiwanese; therefore, a township with a higher proportion of Hoklo people is considered as a neighborhood with stronger Taiwanese identity. The ethnicity data comes from the National Hakka Population Basic Information Survey Research, conducted in 2004 with a sample size of 37,693, equivalent to about 100 people in each town. We use the responses to the question: "You consider yourself as..?" The six options included 1) Taiwan Hakka, 2) Mainland Hakka, 3) Hoklo, 4) Mainlander, 5) Aborigine, and 6) Foreigner. The respondents could only pick one answer to this question. The proportion of people answering Hoklo would be regarded as the share of Hoklo people in the town.

3.10.5 Questions on Other Social Values and Political Participation

Questions on Other Social Values

We construct binary outcome variables used in Table 3.10, using the 2004 and 2005 Taiwan Social Change Survey waves. The original questions and the ways in which we construct the binary variables are as follows. Some questions are not the same across both years. In such cases, we construct the binary variable with the goal of having similar means of the outcome variable, using samples with all ages across years.

One's success relies on coming from a rich family.

- Original Question

    - 2004: How important is the following factor in determining one's success: one's family background?

      1) Extremely important; 2) Very important; 3) Important ; 4) Not important.

    - 2005: One's success relies on coming from a rich family. To what extent do you agree or disagree with this point of view?

      1) Agree strongly; 2) Agree; 3) Disagree; 4) Disagree strongly.

- Construction of the binary variable for the 2004 and 2005 waves: responses of 1, 2 are coded as 1, 0 otherwise.

More equal income distribution makes people work less

- Original Question in 2005: If we have more equal income distribution, a normal person will be less likely to not work hard. To what extent do you agree or disagree with this point of view?

  1) Agree strongly; 2) Agree; 3) Disagree; 4) Disagree strongly.

- Construction of the binary variable for the 2005 wave: Responses of 1, 2 are coded as 1, 0 otherwise.

A better social welfare system makes people work less.

- Original question in 2005: If we have a better social welfare system, a normal person will be less likely to work hard. To what extent do you agree or disagree with this point of view?

  1) Agree strongly; 2) Agree; 3) Disagree; 4) Disagree strongly.

- Construction of the binary variable for the 2005 wave: Responses of 1, 2 are coded as 1, 0 otherwise.

A male should "at least" have a college degree.

- Original question in 2005: What level of education should a boy "at least" attain?

  1) Elementary school; 2) Junior high; 3) Senior high; 4) Vocational college; 5) Undergraduate; 6) Postgraduate.

- Construction of the binary variable for the 2005 wave: Responses of 5, 6 are coded as 1, 0 otherwise.

A female should "at least" have a college degree.

- Original question in 2005: What level of education should a girl "at least" attain?

  1) Elementary school; 2) Junior high; 3) Senior high; 4) Vocational college; 5) Undergraduate; 6) Postgraduate.

- Construction of the binary variable for the 2005 wave: Responses of 5, 6 are coded as 1, 0 otherwise."

Table 3.10. Effects of the Curriculum Reform on Attitude Towards Other Social Values

	(1)	(2)	(3)	(4)
Panel A: One's success relies on coming from a rich family				
TextBook	-0.049	-0.008	-0.047	-0.029
	(0.0692)	(0.0727)	(0.0781)	(0.0882)
Baseline Mean	0.708	0.708	0.708	0.708
Sample size	320	320	320	320
Panel B: More equal income distribution makes people work less				
TextBook	-0.133	-0.102	-0.074	-0.107
	(0.146)	(0.158)	(0.177)	(0.213)
Baseline Mean	0.534	0.534	0.534	0.534
Sample size	183	183	183	183
Panel C: A Better social welfare system makes people work less				
TextBook	0.127	0.083	0.099	0.029
	(0.157)	(0.150)	(0.158)	(0.174)
Baseline Mean	0.409	0.409	0.409	0.409
Sample size	183	183	183	183
Panel D: A Male should "at least" have a college degree				
TextBook	0.121	0.086	0.058	0.114
	(0.134)	(0.143)	(0.160)	(0.187)
Baseline Mean	0.67	0.67	0.67	0.67
Sample size	183	183	183	183
Panel E: A Female should "at least" have a college degree				
TextBook	0.070	0.039	0.019	0.114
	(0.122)	(0.123)	(0.126)	(0.142)
Baseline Mean	0.659	0.659	0.659	0.659
Sample size	183	183	183	183
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2003, 2004, and 2005 TSCS data and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Questions and outcome variables used in Panel A to Panel E can be found in the Online Appendix 3.10.5. The baseline mean is the simple average of outcomes of respondents born between September 1982 and August 1984 in the corresponding subgroup. Standard errors clustered at the birth year-month level are in parentheses. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

Questions on Political Participation

We construct binary outcome variables used in Table 3.11, using the 2004 Taiwan Social Change Survey wave.

You have the power to affect governmental decisions.

- Original Question: Do you agree that you have the power to affect governmental decisions?

  1) Totally agree; 2) Agree; 3) No opinion; 4) Do not agree; 5) Do not agree anymore

- Construction of the binary variable: responses of 1, 2 are coded as 1, 0 otherwise.

You often discuss politics with your friends.

- Original question in 2005: How often do you discuss politics with your friends?

  1) Very often; 2) Sometimes; 3) Seldom; 4) Never.

- Construction of the binary variable: Responses of 1, 2 are coded as 1, 0 otherwise.

You pay attention to political news in the newspaper.

- Original Question: How often do you read political news in newspapers?

  1) Every day; 2) 3-4 days in a week; 3) 1-2 days in a week; 4) Less than 1-2 days in a week; 5) Never

- Construction of the binary variable: Responses of 1, 2, 3 are coded as 1, 0 otherwise.

You pay attention to political news on TV.

- Original Question: How often do you read political news on TV?

  1) Every day; 2) 3-4 days in a week; 3) 1-2 days in a week; 4) Less than 1-2 days in a week; 5) Never

- Construction of the binary variable: Responses of 1, 2, 3 are coded as 1, 0 otherwise.

148

You pay attention to political news on the internet.

- Original Question: How often do you read political news on the internet?

  1) Every day; 2) 3-4 days in a week; 3) 1-2 days in a week; 4) Less than 1-2 days in a week; 5) Never

- Construction of the binary variable: Responses of 1, 2, 3 are coded as 1, 0 otherwise.

Table 3.11. Effects of the Curriculum Reform on Political Participation

	(1)	(2)	(3)	(4)
Panel A: You have the power to affect governmental decisions				
TextBook	-0.080	-0.052	-0.091	-0.083
	(0.144)	(0.149)	(0.148)	(0.153)
Baseline Mean	0.354	0.354	0.354	0.354
Sample size	144	144	144	144
Panel B: You often discuss politics with your friends				
TextBook	-0.035	-0.038	-0.030	-0.025
	(0.031)	(0.034)	(0.024)	(0.020)
Baseline Mean	0.305	0.305	0.305	0.305
Sample size	144	144	144	144
Panel C: You pay attention to political news in newspaper				
TextBook	0.022	0.009	0.013	-0.008
	(0.122)	(0.115)	(0.119)	(0.123)
Baseline Mean	0.256	0.256	0.256	0.256
Sample size	144	144	144	144
Panel D: You pay attention to political news in TV				
TextBook	0.022	0.003	-0.017	0.008
	(0.170)	(0.165)	(0.177)	(0.176)
Baseline Mean	0.561	0.561	0.561	0.561
Sample size	144	144	144	144
Panel E: You pay attention to political news in internet				
TextBook	-0.109	-0.096	-0.044	-0.026
	(0.100)	(0.102)	(0.100)	(0.101)
Baseline Mean	0.293	0.293	0.293	0.293
Sample size	144	144	144	144
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2004 TSCS data and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Questions and outcome variables used in Panel A to Panel E can be found in the Online Appendix 3.10.5. The baseline mean is the simple average of outcomes of respondents born between September 1982 and August 1984 in the corresponding subgroup. Standard errors clustered at the birth year-month level are in parentheses. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

3.10.6 Additional Tables

Table 3.12. Heterogeneous RD analysis: By Share of Hoklo People in the Hometown

	(1)	(2)	(3)	(4)
H_0: CATE(.)=ATE				
P-value	0.02	0.05	0.10	0.09
Sample size	427	427	427	427
Number of overlapping subgroups	3	6	10	10
Bandwidth	24	24	24	12

Notes: This table examines whether the curriculum effect is heterogeneous across individuals from their hometown with different share of Hoklo ethnic people. We use a test for treatment effect heterogeneity in RD designs proposed by Hsu and Shen (2019). The null hypothesis is that the curriculum effect in each subgroup (i.e., CATE(.)) is equal to average treatment effect (i.e., ATE). In other words, the curriculum effect is homogeneous across all subgroups. The construction of subgroups is as follows. First, we set the largest number of subgroups (Q). Second, we form the subgroups by: 1) form Q subgroups which evenly divide the hometown's Hoklo share, 2) form $Q-1$ (q) subgroups which evenly divide the hometown's Hoklo share, 3) so on until q equals to one. For example, when Q equals to 4, we have 10 overlapping subgroups. We then collect all these overlapping subgroups, estimate conditional treatment effects within each group, and test if all conditional treatment effects from each subgroup equal to the average treatment effect. P-value for such a test is reported.

Table 3.13. Subgroup Analysis: By Hometown Ethnicity Distribution (Hoklo and Hakka)
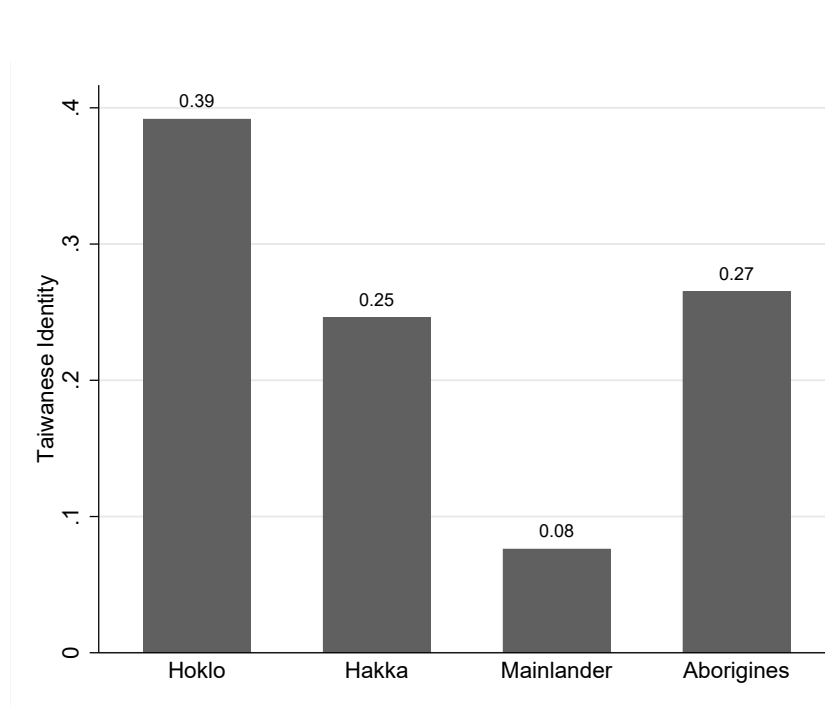
	Taiwanese Identity			
	(1)	(2)	(3)	(4)
Panel A: Hometown with Low Proportion of Hoklo and Hakka Ethnic				
TextBook	0.276**	0.285**	0.338**	0.311**
	(0.133)	(0.133)	(0.132)	(0.132)
Baseline Mean	0.581	0.581	0.581	0.581
Sample Size	211	211	211	211
Panel B: Hometown with High Proportion of Hoklo and Hakka Ethnic				
TextBook	0.080	0.097	0.067	0.090
	(0.079)	(0.080)	(0.099)	(0.097)
Baseline Mean	0.64	0.64	0.64	0.64
Sample Size	216	216	216	216
Panel C: Test Heterogeneity				
TextBook	0.079	0.097	0.099	0.108
	(0.079)	(0.076)	(0.078)	(0.081)
$TextBook \times LowHoka$	0.196	0.182	0.199	0.175
	(0.153)	(0.142)	(0.138)	(0.132)
Sample Size	427	427	427	427
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2003, 2004, 2005 TSCS data and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Panel A includes respondents whose hometown has a lower proportion of Hoklo and Hakka people compared to the median of the population in the National Hakka Population Basic Information Survey Research, while Panel B includes respondents whose hometown has a higher proportion of Hoklo and Hakka people. Specifications in each column are the same as in the corresponding columns in Table 3.2. The baseline mean is the simple average of *Identity* of respondents born between September 1982 and August 1984 in the corresponding subgroup. Panel C tests the statistical significance of difference in curriculum effect between two subgroups by showing coefficient on the interaction term of *TextBook* and a dummy for the individuals from the area with low share of Hoklo and Hakka people *LowHoka*. Compared to equation (3.1), this specification also includes 1) a dummy for the individuals from the area with low share of Hoklo and Hakka people *LowHoka*; 2) the interaction term of *TextBook* and *LowHoka*; 3) the interactions between *LowHoka* and running variable; 4) the interactions between *LowHoka* and survey year fixed effects. Standard errors clustered at birth year-month level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

**Table 3.14.** Subgroup Analysis: By Ethnicity of Respondents' Parents

	Taiwanese Identity			
	(1)	(2)	(3)	(4)
Panel A: One of Parents are not Hoklo Ethnic				
TextBook	0.385***	0.434***	0.538***	0.569***
	(0.118)	(0.142)	(0.152)	(0.162)
Baseline Mean	0.592	0.592	0.592	0.592
Sample Size	118	118	118	118
Panel B: One of Parents are not Hoklo Ethnic (Exclude the Outlier)				
TextBook	0.350***	0.398***	0.515***	0.552***
	(0.116)	(0.145)	(0.157)	(0.164)
Baseline Mean	0.592	0.592	0.592	0.592
Sample Size	116	116	116	116
Panel C: Both Parents are Hoklo Ethnic				
TextBook	0.085	0.086	0.112	0.096
	(0.101)	(0.101)	(0.096)	(0.103)
Baseline Mean	0.617	0.617	0.617	0.617
Sample Size	309	309	309	309
Panel D: Test Heterogeneity				
TextBook	0.085	0.088	0.106	0.102
	(0.102)	(0.101)	(0.101)	(0.105)
TextBook × NotHoklo	0.301*	0.334*	0.317*	0.322*
	(0.154)	(0.168)	(0.167)	(0.177)
Sample Size	427	427	427	427
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2003, 2004, 2005 TSCS data and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Panel A includes respondents with at least one parent who are non-Hoklo. Note that there is an outlier (see the rightmost dot in the Figure 3.26a), which is consist of only two individuals. Both did not have Taiwanese identity so that it is particularly negative compared to other dots. Panel B reports the estimate by excluding these sample. Panel C includes respondents whose parents are both Hoklo. Specifications in each column are the same as in the corresponding columns in Table 3.2. The baseline mean is the simple average of *Identity* of respondents born between September 1982 and August 1984 in the corresponding subgroup. Panel C tests the statistical significance of difference in curriculum effect between two subgroups by showing coefficient on the interaction term of *TextBook* and a dummy for the individuals with at least one parent who are non-Hoklo *NotHoklo*. Compared to equation (3.1), this specification also includes 1) a dummy for the individuals with at least one parent who are non-Hoklo *NotHoklo*; 2) the interaction term of *TextBook* and *NotHoklo*; 3) the interactions between *NotHoklo* and running variable; 4) the interactions between *NotHoklo* and survey year fixed effects. Standard errors clustered at the birth year-month level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 3.20. Four Major Ethnic Groups and Taiwanese Identity

Notes: This figure displays the share of people identifying them as Taiwanese by four major ethnic groups in Taiwan. Data is from 1992, 1995, and 1998 TSCS waves. We restrict sample to people who are 25 years old or above. The ethnic group that a respondent belongs to is based on his/her father's origins.
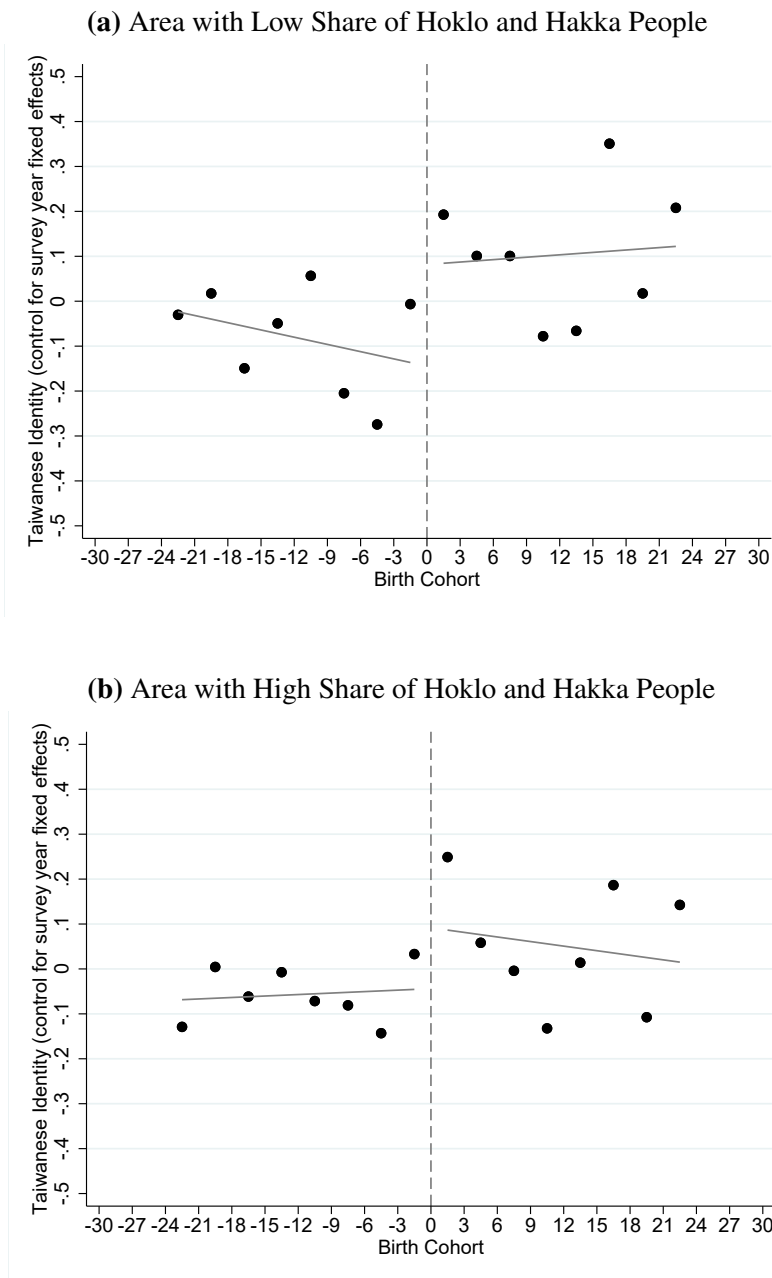
(a) Identity Question Used in 2003-2004 Waves

(b) Identity Question Used in 2005 Waves
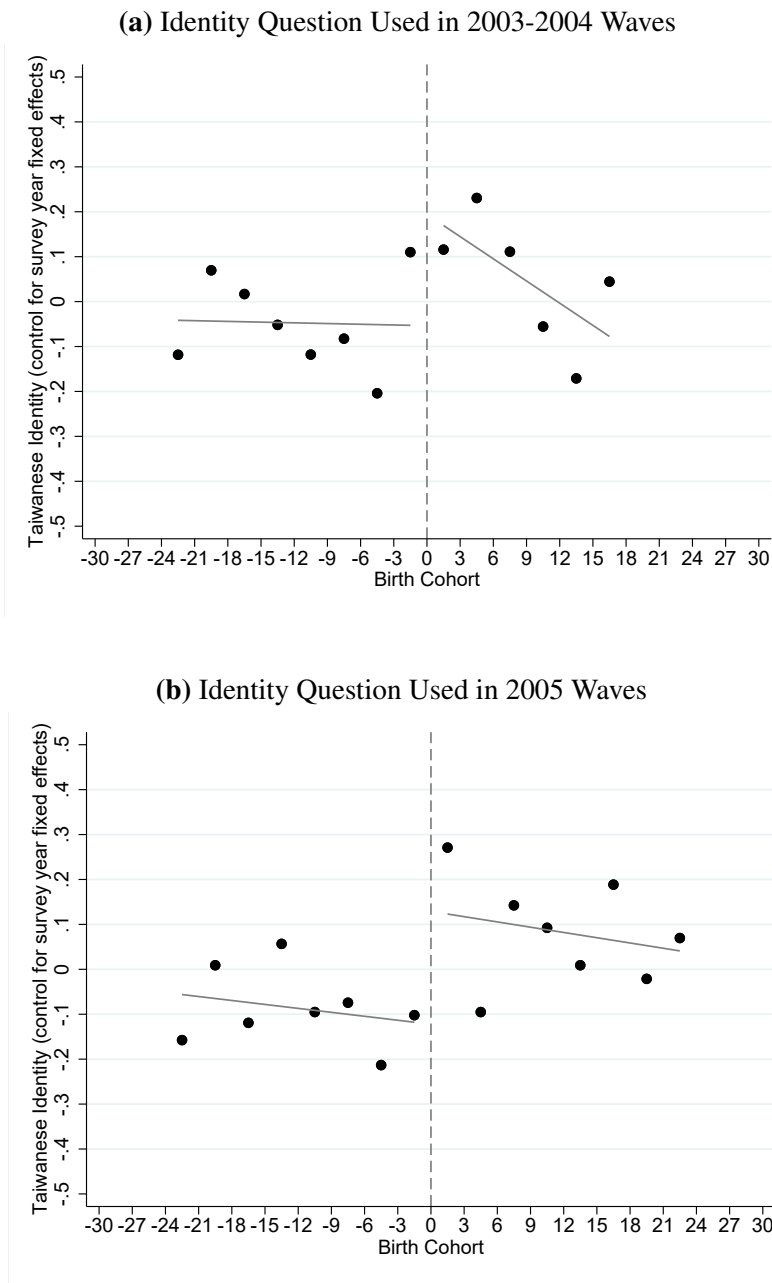
Figure 3.21. RD Graph for Different Identity Questions

Notes: Figure 3.21a is based on 2003 and 2004 TSCS. Figure 3.21b is based on 2005 TSCS. We use the sample born between September 1982 and August 1986. Note that the last two birth cohorts (e.g., those born during March 1986 to August 1986) were not surveyed since they were below 18 years old in 2003 and 2004. We first regress *Identity* on survey year dummies and then collapse the residuals at birth year-quarter level (i.e. three birth year-month cohorts) to derive the dots. Thus, the first dot in this figure represents average *Identity* (after controlling for the survey year fixed effect) for those born in September, October, and November 1982 and the last dot represents average *Identity* (after controlling for the survey year fixed effect) for those born in June, July, and August 1986. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.
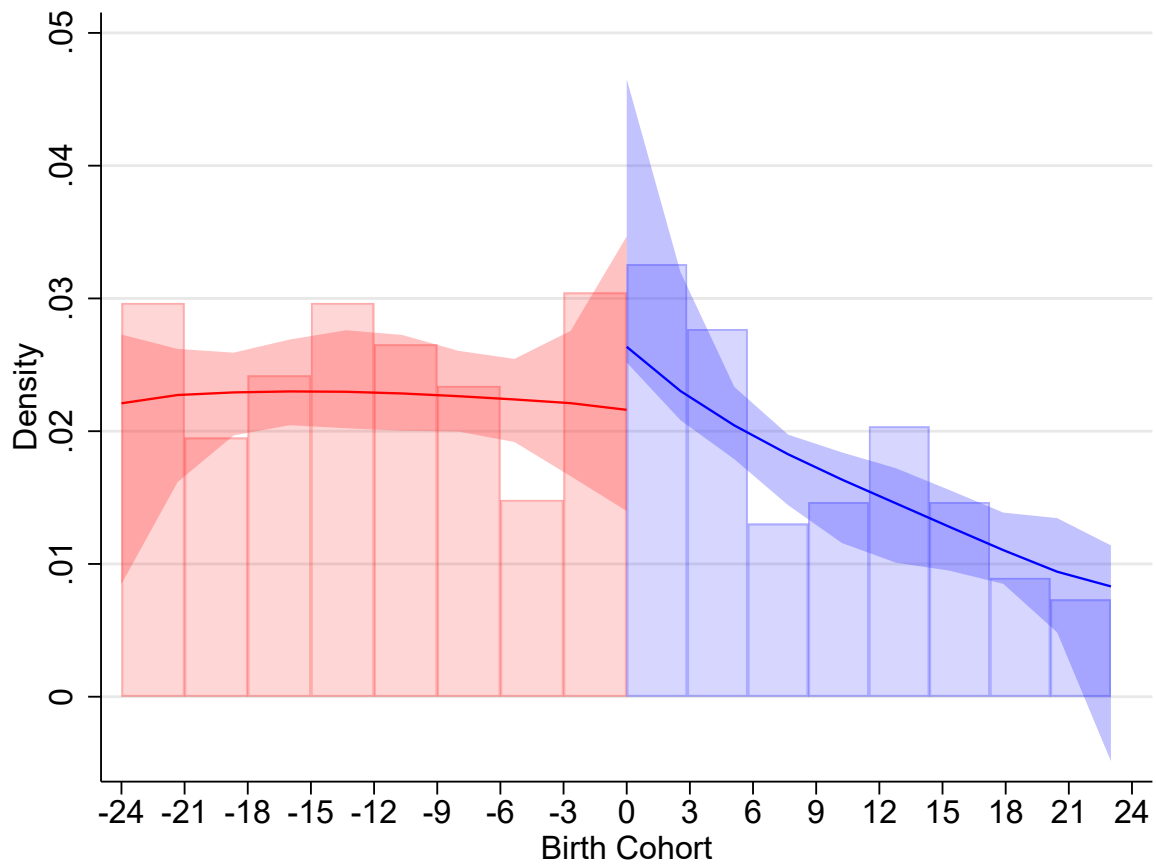
Figure 3.22. Density of Birth Cohort: Density Discontinuity Test

Notes: This figure displays the results for a density discontinuity test proposed by Cattaneo et al. (2020, 2018).
We pool the 2003, 2004, and 2005 TSCS data. Each bar represents the density of birth cohort. The birth cohort
is measured at birth year-quarter level (i.e. three birth year-month cohorts). The shaded area represents the 95%
confidence interval.

(a) Area with Low Share of Hoklo and Hakka People

(b) Area with High Share of Hoklo and Hakka People

Figure 3.23. Taiwanese Identity and Birth Cohorts: By High/Low *Hoklo and Hakka Proportion* Areas

Notes: We pool data from 2003, 2004, 2005 TSCS and use the sample born between September 1982 and August 1986. Figure 3.23a includes respondents living in the towns with low share of Hoklo and Hakka people before age 15. Figure 3.23b includes respondents living in towns with high share of Hoklo and Hakka people before age 15. We first regress *Identity* on survey year dummies and then collapse the residuals at birth quarter level to derive the dots. Thus, zero in the figure represents September, October, and November 1984. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.

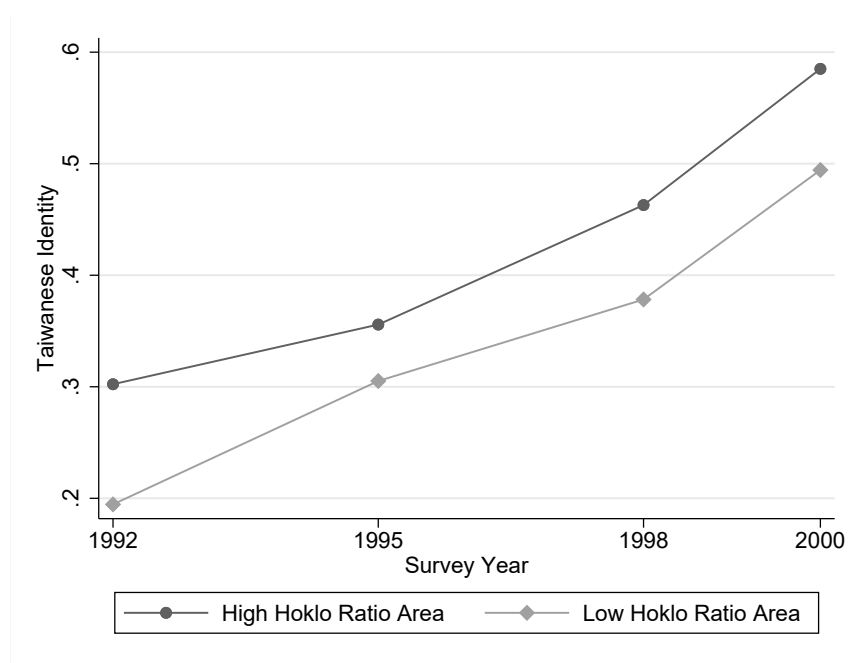**Figure 3.24.** Taiwanese Identity Trend in Townships with High/Low Share of Hoklo and Hakka People in 1990s

Notes: We pool data from 1992, 1995, 1998 and 2000 TSCS waves. In order to include those adults whom children are more likely to meet, we restrict the respondents aged 25 or above. Each dot represents share of people reporting Taiwanese identity in given survey year and area. The circle symbol represents the area with high share of Hoklo and Hakka people. The diamond symbol represents the area with low share of Hoklo and Hakka people.
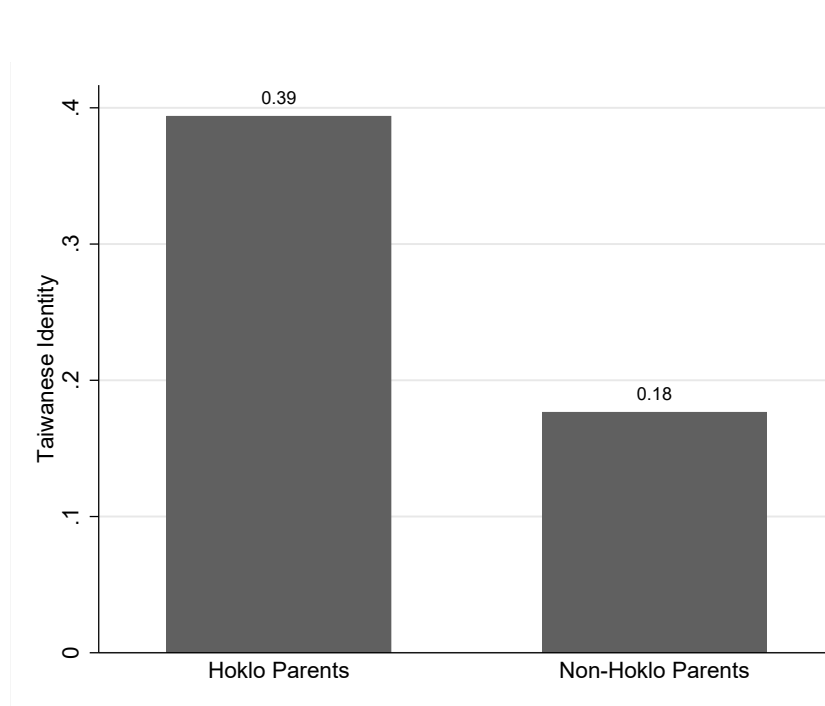
Figure 3.25. Ethnic Composition of Parents and Taiwanese Identity

Notes: This figure display the share of people identifying them as Taiwanese by ethnic composition of parents. Data is from 1992, 1995, and 1998 TSCS waves. We restrict sample to people who are 25 years old or above. Hoklo parents: both father and mother are Hoklo. Non-Hoklo parents: at least one of parents are non-Hoklo.

(a) One of Parents are not Hoklo Ethnic

(b) Both Parents are Hoklo Ethnic
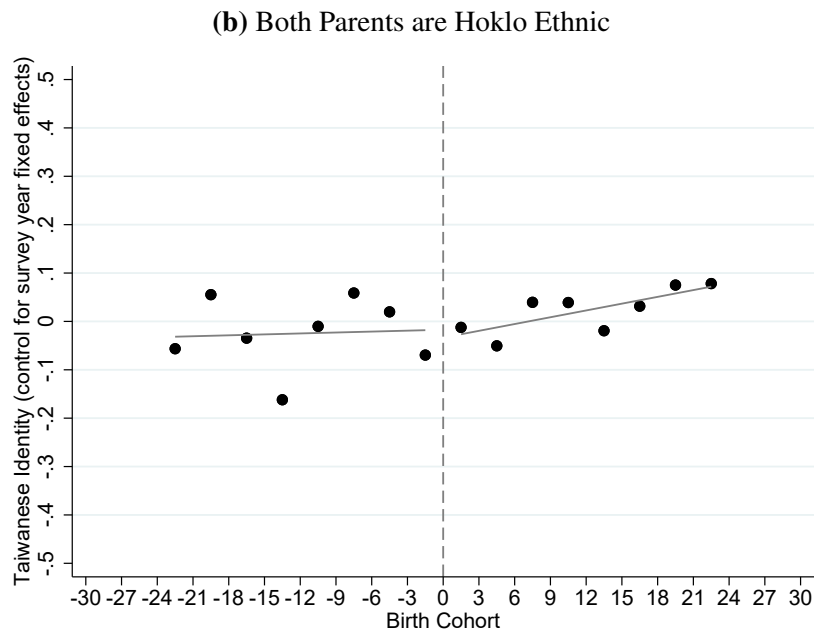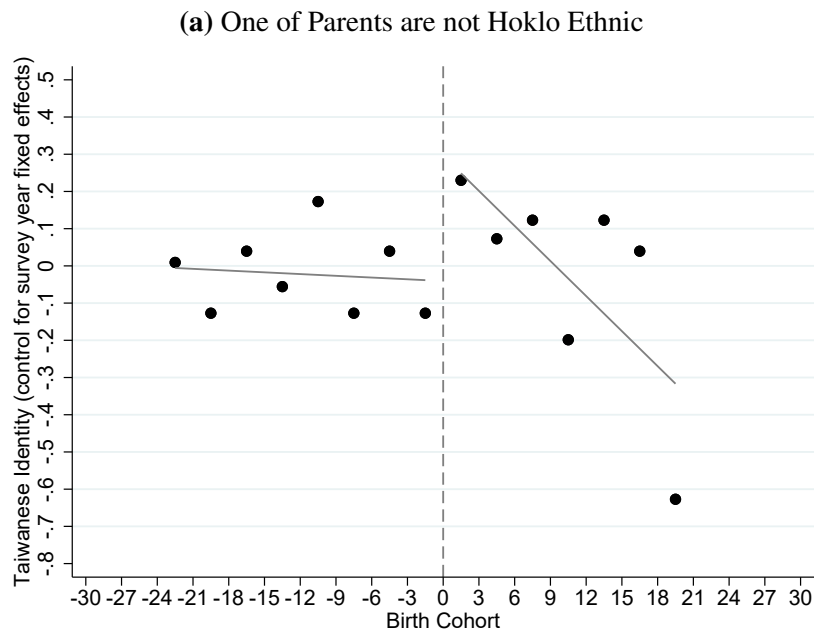
Figure 3.26. Taiwanese Identity and Birth Cohorts: By Ethnicity of Respondents' Parents

Notes: We pool data from 2003, 2004, 2005 TSCS and use the sample born between September 1982 and August 1986. Figure 3.26a includes respondents with at least one parent who was non-Hoklo. Figure 3.26b includes respondents whose parents had Hoklo ethnicity. We first regress *Identity* on survey year dummies and then collapse the residuals at birth quarter level to derive the dots. Thus, zero in the figure represents September, October, and November 1984. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.

3.10.7 Hoklo Ethnicity Distribution in Taiwan
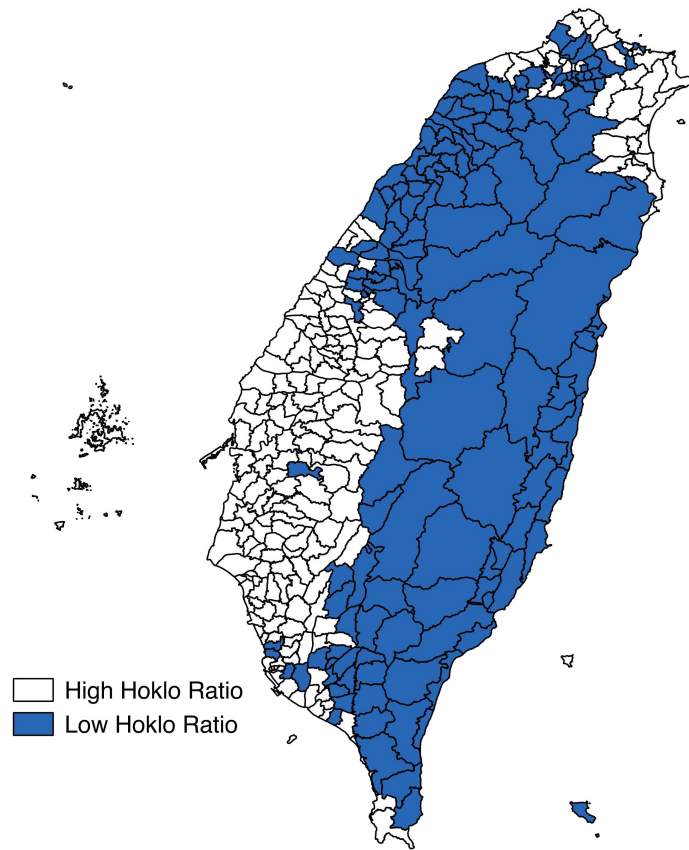
Figure 3.27. Geographical distribution of Towns with High and Low Proportions of Hoklo People

Notes: We compute the proportion of Hoklo people in each town and the population median of the proportion of Hoklo people (with 2004 population as weight), using the data from 2004's National Hakka Population Basic Information Survey Research. The median is 77.1%. High Hoklo area includes the towns where have a proportion of Hoklo people higher than 77.1%; low Hoklo area otherwise.

3.10.8 Robustness Checks for Long-term Results: Tables

Table 3.15. Descriptive Statistics for Long-run and Short-run Sample

	Long-run Sample	Short-run Sample	Difference (long-short)
Female	0.477	0.445	0.032
	(0.500)	(0.498)	(0.030)
Age	28.923	20.396	8.527***
	(2.127)	(1.104)	(0.091)
Years of schooling (self)	13.803	13.761	0.042
	(2.083)	(2.021)	(0.122)
Years of schooling (father)	11.034	10.600	0.435**
	(3.453)	(3.450)	(0.206)
Years of schooling (mother)	10.398	9.888	0.510**
	(3.582)	(3.310)	(0.203)
Proportion of Hoklo in the hometown	0.737	0.724	0.013
	(0.187)	(0.215)	(0.012)
Hoklo father	0.794	0.775	0.019
	(0.404)	(0.418)	(0.025)
Hoklo mother	0.802	0.824	-0.023
	(0.399)	(0.381)	(0.023)
# of individuals	822	417	

Note: We pool data from the 2003, 2004, 2005, 2010, 2012, 2013, 2014, and 2015 TSCS waves and use the sample born between September 1982 and August 1986. The definitions of the individual characteristics are as follows: 1) Female: If an individual is female assigned 1, otherwise 0. 2) Respondent/Father/Mother's schooling years: a) no education (zero years of schooling); b) elementary school (6 years of schooling); c) junior high school (9 years of schooling); d) senior (vocational) high school (12 years of schooling); e) two-year college (14 years of schooling); f) University or vocational university (16 years of schooling). 3) Hoklo fathers/mothers: If an individual's father/mother is Hoklo assigned 1, otherwise 0. In the Online Appendix 3.10.4, we provide detailed definition of proportion of Hoklo people in the hometown. Standard deviations in parentheses, and standard errors in brackets. *** significant at the 1 percent level, ** significant at the 5 percent level, and * significant at the 10 percent level.

Table 3.16. Descriptive Statistics for Treatment Group and Control Group

	Born after September 1984	Born before September 1984	Difference (after - before)
Female	0.483	0.471	0.0117
	(0.500)	(0.500)	(0.035)
Age	27.737	29.969	-2.232***
	(1.873)	(1.759)	(0.127)
Years of schooling (self)	13.784	13.819	-0.0348
	(2.097)	(2.072)	(0.146)
Years of schooling (father)	11.418	10.696	0.723***
	(3.224)	(3.612)	(0.238)
Years of schooling (mother)	10.771	10.069	0.703***
	(3.533)	(3.596)	(0.249)
Proportion of Hoklo in the hometown	0.738	0.736	0.00147
	(0.187)	(0.187)	(0.0131)
Hoklo father	0.779	0.808	-0.0286
	(0.415)	(0.394)	(0.0284)
Hoklo mother	0.800	0.803	-0.0032
	(0.401)	(0.398)	(0.0279)
# of individuals	437	385	

Notes: We pool data from the 2010, 2012, 2013, 2014, and 2015 TSCS waves and use the sample born between September 1982 and August 1986. The definitions of the individual characteristics are as follows: 1) Female: If an individual is female assigned 1, otherwise 0. 2) Respondent/Father/Mother's schooling years: a) no education (zero years of schooling); b) elementary school (6 years of schooling); c) junior high school (9 years of schooling); d) senior (vocational) high school (12 years of schooling); e) two-year college (14 years of schooling); f) University or vocational university (16 years of schooling). 3) Hoklo fathers/mothers: If an individual's father/mother is Hoklo assigned 1, otherwise 0. In the Online Appendix 3.10.4, we provide detailed definition of proportion of Hoklo people in the hometown. Standard deviations in parentheses, and standard errors in brackets. *** significant at the 1 percent level, ** significant at the 5 percent level, and * significant at the 10 percent level.

Table 3.17. Robustness Check: Validity of Regression Discontinuity Design

VARIABLES	Female	Age	Years of Schooling	Father's Schooling years	Mother's Schooling years	Share of Hoklo people	Hoklo Fathers	Hoklo Mothers
TextBook	0.097*	0.044	0.009	0.417	0.258	0.048	0.025	0.064
	(0.052)	(0.052)	(0.273)	(0.505)	(0.500)	(0.030)	(0.053)	(0.042)
Sample Size	822	822	822	822	822	822	822	822

Notes: We pool data from the 2010, 2012, 2013, 2014, and 2015 TSCS waves and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. Note that we do not include any covariates X_i since they are outcome variables now. All columns include the survey year fixed effect and the first-order polynomials of birth year-month m interacting fully with *TextBook* (i.e. linear spline). The definitions of the individual characteristics are as follows: 1) Female: If an individual is female assigned 1, otherwise 0. 2) Respondent/Father/Mother's schooling years: a) no education (zero years of schooling); b) elementary school (6 years of schooling); c) junior high school (9 years of schooling); d) senior (vocational) high school (12 years of schooling); e) two-year college (14 years of schooling); f) University or vocational university (16 years of schooling). 3) Hoklo fathers/mothers: If an individual's father/mother is Hoklo assigned 1, otherwise 0. In the Online Appendix 3.10.4, we provide detailed definition of proportion of Hoklo people in the hometown. Standard errors are clustered at the birth year-month level in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

Table 3.18. Robustness Check: Different Specifications, Sample, and Identity Questions

|	Taiwanese Identity			
	(1)	(2)	(3)	(4)
Panel A: 2nd Order Polynomial				
TextBook	-0.021	-0.018	-0.004	-0.013
	(0.082)	(0.075)	(0.074)	(0.073)
Sample Size	822	822	822	822
Panel B: Exclude Specific Birth Cohorts				
TextBook	0.009	0.034	0.032	0.030
	(0.056)	(0.056)	(0.055)	(0.054)
Sample Size	774	774	774	774
Panel C: Identity Question – 2012-2014 Waves				
TextBook	-0.010	0.017	0.027	0.029
	(0.075)	(0.075)	(0.076)	(0.075)
Observations	426	426	426	426
Panel D: Identity Question – 2010, 2014-2015 Waves				
TextBook	0.017	0.041	0.031	0.022
	(0.067)	(0.070)	(0.071)	(0.071)
Observations	396	396	396	396
Linear/Quadratic Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool data from the 2010, 2012, 2013, 2014, and 2015 TSCS waves and use the sample born between September 1982 and August 1986. The above table reports the coefficient of *TextBook* based on equation (3.1), which is one if the birth year-month of the respondent is after September 1984, zero otherwise. All panels include the survey year fixed effect and the first-order polynomials of birth year-month m interacting fully with *TextBook* (i.e. linear spline). Panel A additionally includes quadratic spline. Other covariates are the same as in the corresponding columns in Table 3.2. Note that 2014 TSCS had two waves and ask slightly different identity questions. Standard errors are clustered at the birth year-month level in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table 3.19. Robustness Check: Placebo Test of Fake Textbook Reform

| | Taiwanese Identity | | | |
	(1)	(2)	(3)	(4)
Panel A: Use September 1983 as Cut-Off				
TextBook	-0.055	-0.070	-0.049	-0.043
	(0.073)	(0.075)	(0.079)	(0.076)
Sample Size	487	487	487	487
Panel B: Use September 1982 as Cut-Off				
TextBook	-0.009	-0.006	-0.027	-0.026
	(0.092)	(0.100)	(0.094)	(0.094)
Sample Size	509	509	509	509
Panel C: Use September 1981 as Cut-Off				
TextBook	-0.018	-0.023	0.012	0.006
	(0.069)	(0.071)	(0.073)	(0.074)
Sample Size	519	519	519	519
Panel D: Use September 1980 as Cut-Off				
TextBook	0.052	0.062	0.006	0.003
	(0.080)	(0.076)	(0.088)	(0.088)
Sample Size	506	506	506	506
Panel E: Difference-in-Differences Design				
$AfterSep \times B_{1984}$	-0.048	-0.049	-0.047	-0.048
	(0.059)	(0.056)	(0.056)	(0.055)
Sample Size	1,097	1,097	1,097	1,097
Linear Spline	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Notes: We pool the 2010, 2012, 2013, 2014, and 2015 TSCS data. Panel A uses the sample born between August 1981 and September 1985; Panel B uses the sample born between August 1980 and September 1984; Panel C uses the sample born between August 1979 and September 1983. Panel D uses the sample born between August 1978 and September 1982. The above table reports the coefficient of *TextBook* based on equation (3.1). In each placebo test, we define dummy variable *TextBook* as respondents born after following cut-offs: September 1983 (Panel A), September 1982 (Panel B), September 1981 (Panel C), or September 1980 (Panel D). Panel E reports the coefficients of $AfterSep \times B_{1984}$ in the equation (3.2). In this specification, we combine all available cutoffs used in the main estimation and placebo tests to implement a DID design. Specifications in each column are the same as in the corresponding columns in Table 3.2. Note that in DID design, we allow the linear spline of running variable to be cohort-specific. Standard errors are clustered at the birth year-month level in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

3.10.9 Robustness Checks for Long-term Results: Figures
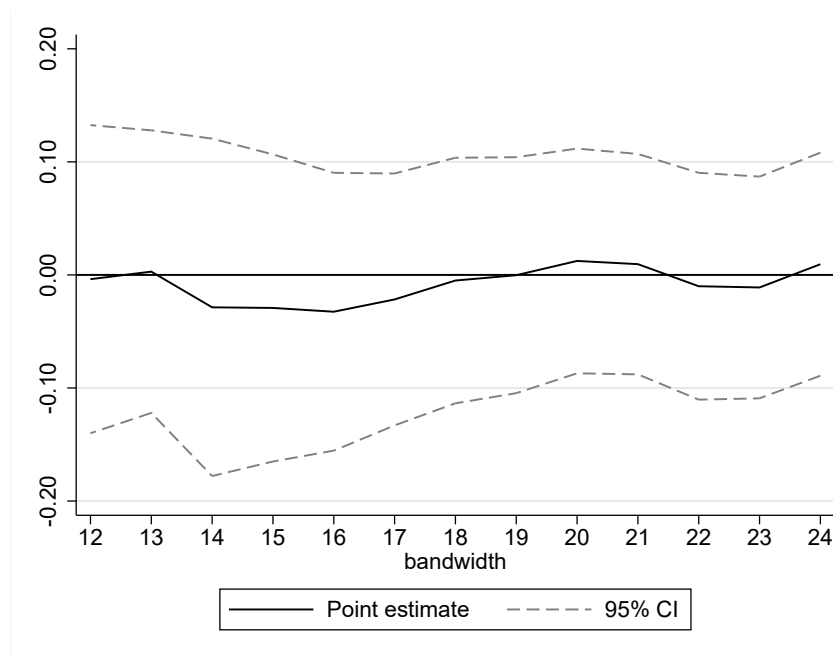
Figure 3.28. RD Estimates across Different Bandwidth Choices: Long-term Results
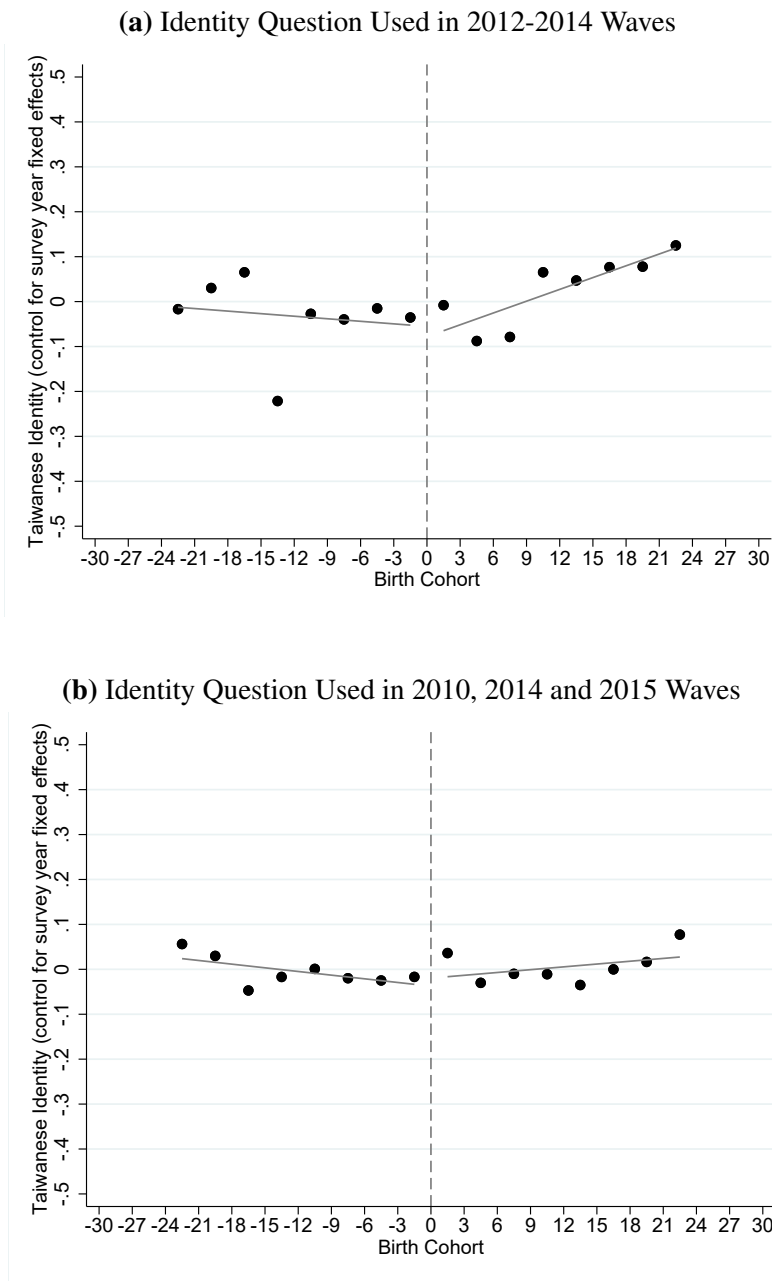
Notes: We run regressions as column (4) in Table 3.9 with different bandwidths: 12 to 24 months on each side of the cut-off, i.e., two education cohorts. The solid line represent the point estimates of coefficients on the *TextBook* dummy variable and the dotted line represents the corresponding 95% confidence interval derived from standard errors clustered at birth year-month level.
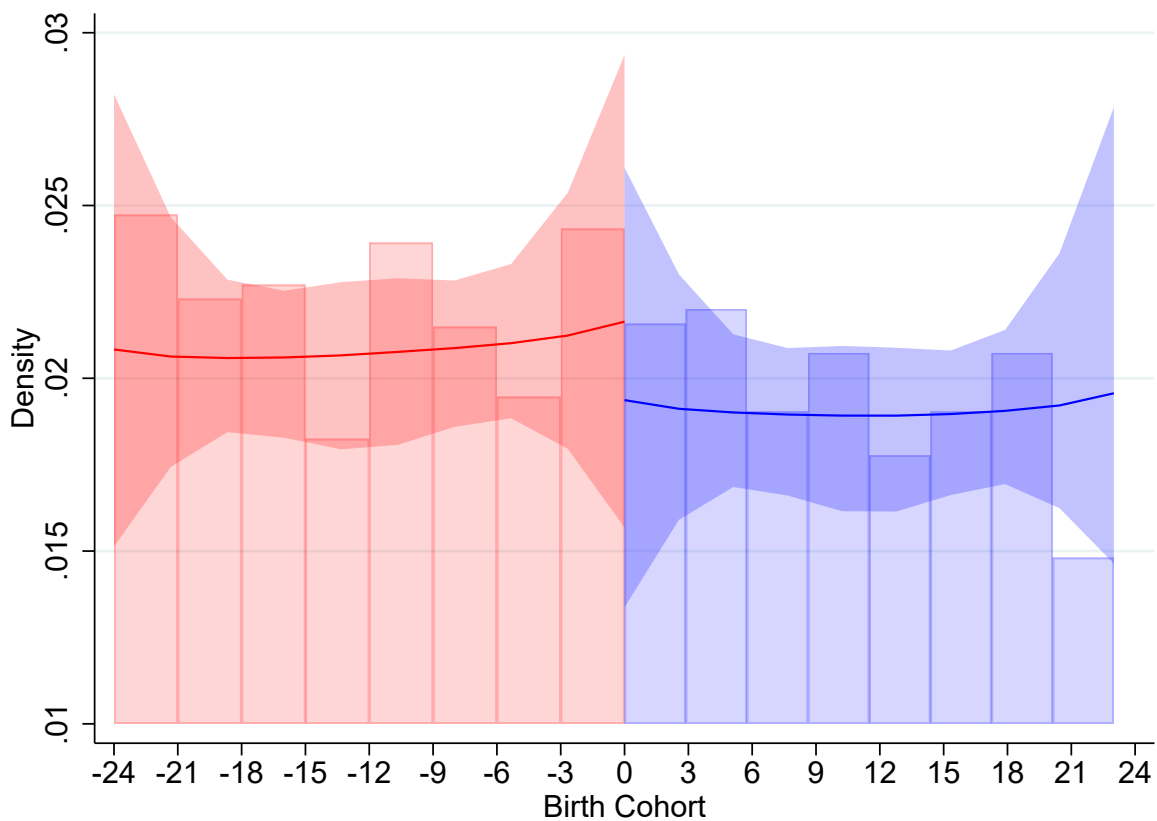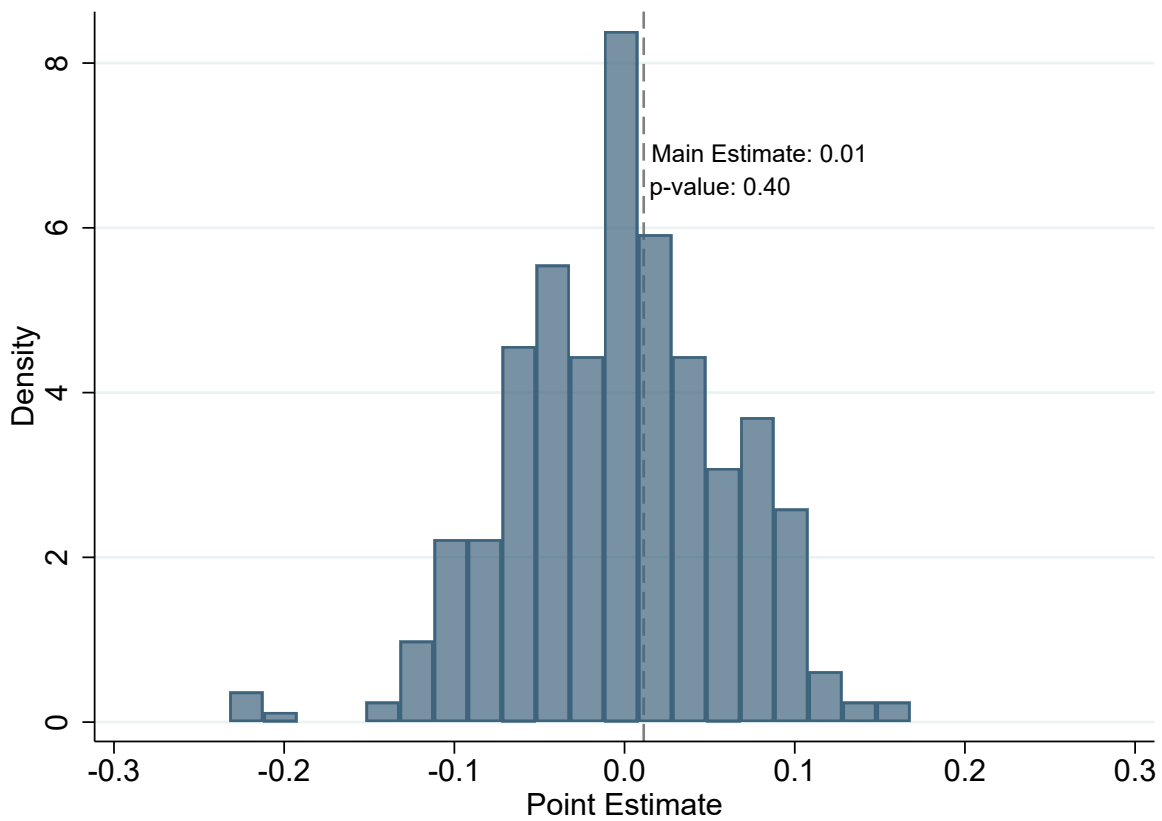
Figure 3.29. Density of Birth Cohort: Density Discontinuity Test for Long-run Sample

Notes: This figure displays the results for a density discontinuity test proposed by Cattaneo et al. (2020, 2018). We pool the 2010, 2012, 2013, 2014, and 2015 TSCS data. Each bar represents the density of birth cohort. The birth cohort is measured at birth year-quarter level (i.e. three birth year-month cohorts). The shaded area represents the 95% confidence interval.

(a) Identity Question Used in 2012-2014 Waves

(b) Identity Question Used in 2010, 2014 and 2015 Waves

Figure 3.30. RD Graph for Different Identity Questions for Long-run Sample

Notes: Figure 3.30a is based on 2012-2014 TSCS. Figure 3.30b is based on 2010, 2014 and 2015 TSCS. Note that 2014 TSCS had two waves and ask slightly different identity questions. We use the sample born between September 1982 and August 1986. We first regress *Identity* on survey year dummies and then collapse the residuals at birth year-quarter level (i.e. three birth year-month cohorts) to derive the dots. Thus, the first dot in this figure represents average *Identity* (after controlling for the survey year fixed effect) for those born in September, October, and November 1982 and the last dot represents average *Identity* (after controlling for the survey year fixed effect) for those born in June, July, and August 1986. Fitted lines are from regression of the dots on a first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.

Figure 3.31. Permutation Test for Long-run Sample

Notes: We pool data from the 2010, 2012, 2013, 2014, and 2015 TSCS waves and assign the fake reform to all possible months and years – from January 1950 to September 1983 (405 fake reforms). This figure display the distribution of placebo estimates (see the histogram) and compare them with our main RD estimate (see the dash line).

Bibliography

M. Adena, R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya. Radio and the rise of the Nazis in prewar Germany. *The Quarterly Journal of Economics*, 130(4):1885–1939, 2015.

P. Aghion, X. Jaravel, T. Persson, and D. Rouzet. Education and military rivalry. *Journal of the European Economic Association*, 17(2):376–412, 2018.

G. A. Akerlof and R. E. Kranton. Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753, 2000.

M. Akhtari, D. Moreira, and L. Trucco. Political turnover, bureaucratic turnover, and the quality of public services. *American Economic Review*, 112(2):442–93, February 2022. doi: 10.1257/aer.20171867. URL https://www.aeaweb.org/articles?id=10.1257/aer.20171867.

A. Alesina and N. Fuchs-Schündeln. Good bye lenin (or not?): The effect of communism on people's preferences. *American Economic Review*, 97(4):1507–1528, 2007.

A. Alesina and B. Reich. Nation building. Technical Report 18839, National Bureau of Economic Research, Cambridge, 2015. NBER Working Paper.

B. Anderson. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso, London, 2006.

K. Antonovics and B. G. Knight. A New Look at Racial Profiling: Evidence from the Boston Police Department. *The Review of Economics and Statistics*, 91(1):163–177, 02 2009. ISSN 0034-6535. doi: 10.1162/rest.91.1.163. URL https://doi.org/10.1162/rest.91.1.163.

B. Ba, J. Kaplan, D. Knox, M. Komisarchik, R. Mariman, J. Mummolo, R. Rivera, and M. Torres. Who are the Police? Descriptive Representation in the Coercive Arm of Government. 2022.

B. A. Ba, D. Knox, J. Mummolo, and R. Rivera. The role of officer race and gender in police-civilian interactions in chicago. *Science*, 371(6530):696–702, 2021. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.abd8694. URL https://www.sciencemag.org/lookup/doi/10.1126/science.abd8694.

Y. Bai and Y. Li. Good bye chiang kai-shek? the long-lasting effects of education under the authoritarian regime in Taiwan. *Working Paper*, 2018.

F. R. Baumgartner, D. A. Epp, and K. S. Shoub. *Suspect Citizens: What 20 Million Traffic Stops Tell Us About Policing and Race*. Cambridge University Press, 2018.

T. Besley and M. Ghatak. Competition and incentives with motivated agents. *American Economic Review*, 95(3):616–636, 2005. ISSN 0002-8282. doi: 10.1257/0002828054201413. URL http://pubs.aeaweb.org/doi/10.1257/0002828054201413.

T. Besley and T. Persson. State capacity, conflict, and development. *Econometrica*, 78(1):1–34, 2010.

M. Billig. *Banal Nationalism*. SAGE, London, 1995.

A. Bisin and T. Verdier. The economics of cultural transmission and socialization. Technical Report 16512, National Bureau of Economic Research, Cambridge, 2010. NBER Working Paper.

A. Bisin, E. Patacchini, T. Verdier, and Y. Zenou. Formation and persistence of oppositional identities. *European Economic Review*, 55(8):1046–1071, 2011a.

A. Bisin, E. Patacchini, T. Verdier, and Y. Zenou. Ethnic identity and labour market outcomes of immigrants in Europe. *Economic Policy*, 26(65):57–92, 2011b.

A. Blouin and S. W. Mukand. Erasing ethnicity? propaganda, nation building, and identity in Rwanda. *Journal of Political Economy*, 127(3):1008–1062, 2019.

A. Bolton, J. M. de Figueiredo, and D. E. Lewis. Elections, Ideology, and Turnover in the US Federal Government. *Journal of Public Administration Research and Theory*, 31(2):451–466, 11 2020. ISSN 1053-1858. doi: 10.1093/jopart/muaa051. URL https://doi.org/10.1093/jopart/muaa051.

C. Brooks. Sheriffs' offices, 2016: Personnel. *Bulletin NCJ*, 252834, 2019.

G. Bulman. Law enforcement leaders and the racial composition of arrests. *Economic Inquiry*, 57(4):1842–1858, 2019. ISSN 0095-2583, 1465-7295. doi: 10.1111/ecin.12800. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecin.12800.

D. Cantoni, Y. Chen, D. Y. Yang, N. Yuchtman, and Y. J. Zhang. Curriculum and ideology. *Journal of Political Economy*, 125(2):338–392, April 2017.

M. D. Cattaneo, M. Jansson, and X. Ma. Manipulation testing based on density discontinuity. *The Stata Journal*, 18(1):234–261, 2018.

M. D. Cattaneo, M. Jansson, and X. Ma. Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455, 2020.

C.-F. Chiang and B. Knight. Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies*, 78(3):795–820, 2011.

C.-F. Chiang, J.-T. Liu, and T.-W. Wen. National identity under economic integration. *Journal of Population Economics*, 32(2):351–367, 2019.

I. Clots-Figueras and P. Masella. Education, language, and identity. *The Economic Journal*, 123: F332–F357, August 2013.

A. Cohen and C. S. Yang. Judicial politics and sentencing decisions. *American Economic Journal: Economic Policy*, 11(1):160–191, 2019. ISSN 1945-7731, 1945-774X. doi: 10.1257/ pol.20170329. URL https://pubs.aeaweb.org/doi/10.1257/pol.20170329.

E. Colonnelli, M. Prem, and E. Teso. Patronage and selection in public sector organizations. *American Economic Review*, 110(10):3071–99, October 2020. doi: 10.1257/aer.20181491. URL https://www.aeaweb.org/articles?id=10.1257/aer.20181491.

A. F. Constant and K. F. Zimmermann. Immigrants, ethnic identities and the nation-state. In *International Handbook on the Economics of Migration*. Edward Elgar Publishing, 2013.

J. F. Copper. *Taiwan: Nation-state or province?* Routledge, 2019.

S. DellaVigna and M. Gentzkow. Persuasion: Empirical evidence. *Annual Review of Economics*, 2(1):643–669, 2010.

S. DellaVigna and E. Kaplan. The fox news effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3):1187–1234, 2007.

S. DellaVigna, R. Enikolopov, V. Mironova, M. Petrova, and E. Zhuravskaya. Cross-border media and nationalism: Evidence from Serbian radio in Croatia. *American Economic Journal: Applied Economics*, 6(3):103–132, 2014.

R. Durante. Building nations through shared experience: Evidence from african football. *The American Economic Review*, 110(5):1572–1602, 2020.

R. Enikolopov, M. Petrova, and E. Zhuravskaya. Media and political persuasion: Evidence from russia. *American Economic Review*, 101(7):3253–85, 2011.

G. Facchini, B. Knight, and C. Testa. The franchise, policing, and race: Evidence from arrests data and the voting rights act. 2020. doi: 10.3386/w27463. URL http://www.nber.org/papers/ w27463.pdf.

B. Feigenberg and C. Miller. Would Eliminating Racial Disparities in Motor Vehicle Searches have Efficiency Costs?*. *The Quarterly Journal of Economics*, 137(1):49–113, 05 2022. ISSN 0033-5533. doi: 10.1093/qje/qjab018. URL https://doi.org/10.1093/qje/qjab018.

M. D. Fliss, F. Baumgartner, P. Delamater, S. Marshall, C. Poole, and W. Robinson. Re-prioritizing traffic stops to reduce motor vehicle crash outcomes and racial disparities. *Injury Epidemiology*, 7(1):3, Dec. 2020. ISSN 2197-1714. doi: 10.1186/s40621-019-0227-6. URL https://injepijournal.biomedcentral.com/articles/10.1186/s40621-019-0227-6.

V. Fouka. Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I. *The Review of Economic Studies*, 87(1):204–239, 2019. ISSN 0034-6527. doi: 10.1093/restud/rdz024. URL https://doi.org/10.1093/restud/rdz024.

W. Friedman, M. Kremer, E. Miguel, and R. Thornton. Education as liberation? *Economica*, 83 (329):1–30, 2016.

M. Gentzkow. Television and voter turnout. *The Quarterly Journal of Economics*, 121(3): 931–972, 2006.

M. Gentzkow, J. M. Shapiro, and M. Sinkinson. Media market structure and political participation: historical evidence from us newspapers. *University of Chicago mimeo*, 2009.

A. Georgiadis and A. Manning. One nation under a groove? understanding national identity. *Journal of economic behavior & organization*, 93:166–185, 2013.

A. S. Gerber, D. Karlan, and D. Bergan. Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics*, 1(2):35–52, 2009.

Y. Ghitza and A. Gelman. The great society, reagan's revolution, and generations of presidential voting. *American Journal of Political Science*, forthcoming, 2022.

F. Goncalves and S. Mello. A few bad apples? racial bias in policing. *American Economic Review*, 111(5):1406–41, May 2021. doi: 10.1257/aer.20181607. URL https://www.aeaweb.org/articles?id=10.1257/aer.20181607.

P. Grosjean, F. Masera, and H. Yousaf. Inflammatory political campaigns and racial bias in policing. *The Quarterly Journal of Economics*, page qjac037, 2022. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjac037. URL https://academic.oup.com/qje/advance-article/doi/10.1093/qje/qjac037/6710386.

M. A. Hansen and J. C. Navarro. Gender and Racial Gaps in Support for Policing and Cor-rectional Reforms: Are the Gaps a Consequence of Political Partisanship? *Crime & Delin-quency*, page 001112872110647, Dec. 2021. ISSN 0011-1287, 1552-387X. doi: 10.1177/

00111287211064788. URL http://journals.sagepub.com/doi/10.1177/00111287211064788.

A.-c. Hsiau. *Contemporary Taiwanese cultural nationalism*. Routledge, 2003.

Y.-C. Hsu and S. Shen. Testing treatment effect heterogeneity in regression discontinuity designs. *Journal of Econometrics*, 208(2):468–486, 2019.

K. Ito. Asymmetric incentives in subsidies: Evidence from a large-scale electricity rebate program. *American Economic Journal: Economic Policy*, 7(3):209–37, 2015.

J. B. Jacobs and P. Kang. *Changing Taiwanese Identities*. Routledge, 2017.

R. Jia and T. Persson. Individual vs. social motives in identity choice: Theory and evidence from China. Technical Report 26008, National Bureau of Economic Research, Cambridge, 2019. NBER Working Paper.

M.-H. Liu, L.-C. Huang, and E. Vickers. *History Education and National Identity in East Asia*, chapter Identity Issues in Taiwan's History Curriculum. Routledge Taylor Francis Group, Abingdon-on-Thames, 1 edition, 2005.

A. Manning and S. Roy. Culture clash or culture club? national identity in britain. *Economic Journal*, 120:F72–F100, 2010.

P. Masella. National identity and ethnic diversity. *Journal of Population Economics*, 26(2): 437–454, 2013.

J. McCrary. The effect of court-ordered hiring quotas on the composition and quality of police. *American Economic Review*, 97(1):318–353, March 2007. doi: 10.1257/aer.97.1.318. URL https://www.aeaweb.org/articles?id=10.1257/aer.97.1.318.

J. McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714, 2008.

A. R. Miller and C. Segal. Does temporary affirmative action produce persistent effects? a study of black and female employment in law enforcement. *Review of Economics and Statistics*, 94 (4):1107–1125, 2012.

A. R. Miller and C. Segal. Do female officers improve law enforcement quality? effects on crime reporting and domestic violence. *The review of economic studies*, 86(5):2220–2247, 2019.

K. Milligan, E. Moretti, and P. Oreopoulos. Does education improve citizenship? evidence from the United States and the United Kingdom. *Journal of Public Economics*, 88(9):1667–1695, 2004.

M. Mitrunen. Can you make an american? compulsory patriotism and assimilation of immigrants. *Working Paper*, 2018.

A. Ornaghi. Civil service reforms: Evidence from u.s. police departments. 2019.

E. Pierson, C. Simoiu, J. Overgoor, S. Corbett-Davies, D. Jenson, A. Shoemaker, V. Ramachandran, P. Barghouty, C. Phillips, R. Shroff, and S. Goel. A large-scale analysis of racial disparities in police stops across the united states. *Nature Human Behaviour*, 4(7):736–745, 2020. ISSN 2397-3374. doi: 10.1038/s41562-020-0858-1. URL http://www.nature.com/articles/s41562-020-0858-1.

K. Roach, F. R. Baumgartner, L. Christiani, D. A. Epp, and K. Shoub. At the intersection: Race, gender, and discretion in police traffic stop outcomes. *Journal of Race, Ethnicity, and Politics*, 7(2):239–261, 2022. doi: 10.1017/rep.2020.35.

G. Solon, S. J. Haider, and J. M. Wooldridge. What are we weighting for? *Journal of Human Resources*, 50(2):301–316, 2015. ISSN 0022-166X. doi: 10.3368/jhr.50.2.301. URL https://jhr.uwpress.org/content/50/2/301.

J. L. Spenkuch, E. Teso, and G. Xu. Ideology and performance in public organizations. *Econometrica*, 91(4):1171–1203, 2023.

H. Tajfel and J. C. Turner. An integrative theory of intergroup conflict. In M. J. Hatch and M. Schultz, editors, *Organizational Identity: A Reader*, pages 56–65. Oxford University Press, Oxford, 1979.

D. M. Thompson. How partisan is local law enforcement? evidence from sheriff cooperation with immigration authorities. *American Political Science Review*, 114(1):222–236, 2020. doi: 10.1017/S0003055419000613.

J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, and M. S. Wetherell. *Rediscovering the social group: A self-categorization theory*. Basil Blackwell, Oxford, 1987.

N. Voigtländer and H.-J. Voth. Nazi indoctrination and anti-Semitic beliefs in Germany. *Proceedings of the National Academy of Sciences of the United States of America*, 112(26):7931–7936, 2015.

F.-C. Wang. National imagination, ethnic consciousness, and history: Content and context analysis of the "Getting to Know Taiwan" textbook disputes,. *Taiwan Historical Research*, 8 (2):145–208, December 2001.

E. Weber. *Peasants into Frenchmen: the modernization of rural France, 1870-1914*. Stanford University Press, Stanford, 1976.

D. Yanagizawa-Drott. Propaganda and conflict: Evidence from the Rwandan genocide. *The Quarterly Journal of Economics*, 129(4):1947–1994, 2014.