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

Wei-Lin Chen,*Ming-Jen Lin,[†] Tzu-Ting Yang[‡]

November 15, 2022

Abstract

This paper examines 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. In addition, our results suggest that the new curriculum had little impact on people's political preferences related to Taiwan 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.

JEL code: I2, I21

^{*}Department of Economics, University of California San Diego, Email: wec155@ucsd.edu

[†]Department of Economics, National Taiwan University, Email: mjlin@ntu.edu.tw

[‡]Institute of Economics, Academia Sinica and International Master's Program of Applied Economics and Social Development, National Chengchi University, Email: ttyang@econ.sinica.edu.tw. Data Availability Statement: The data used in this article can be obtained from the website of the Survey Research Data Archive (https://srda.sinica.edu.tw/). Additional replication materials are provided in the Online Appendix. Disclosure Statement: We do not have any additional disclosures to report.

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 (Alesina and Reich, 2015; Anderson, 2006; Besley and Persson, 2010; Billig, 1995; Weber, 1976). 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 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,

¹Empirical 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).

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 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 Kaplan, 2007; DellaVigna and Gentzkow, 2010). The estimated persuasion rate is 46%, which is much larger than the estimates (i.e., around 2% to 20%) for news media (Adena et al., 2015; Blouin and Mukand, 2019; Chiang and Knight, 2011; DellaVigna and Kaplan, 2007; DellaVigna et al., 2014; Enikolopov, Petrova and Zhuravskaya, 2011; Gentzkow, Shapiro and Sinkinson, 2009; Gentzkow, 2006; Gerber, Karlan and

²Taiwan 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.

Bergan, 2009; Yanagizawa-Drott, 2014).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 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.

³The estimated persuasion rates in the literature for news media, such as TV programs (DellaVigna and Kaplan, 2007; Enikolopov, Petrova and Zhuravskaya, 2011; Gentzkow, 2006), radio (Adena et al., 2015; Blouin and Mukand, 2019; DellaVigna et al., 2014; Yanagizawa-Drott, 2014), and newspapers (Chiang and Knight, 2011; Gentzkow, Shapiro and Sinkinson, 2009; Gerber, Karlan and Bergan, 2009) are around 2% to 20%. One noticeable exception is Enikolopov, Petrova and Zhuravskaya (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.

 $^{^{4}}$ In section 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.

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 (Akerlof and Kranton, 2000; Alesina and Fuchs-Schündeln, 2007; Bisin and Verdier, 2010; Bisin et al., 2011*a*; Constant and Zimmermann, 2013; Durante, 2020; Georgiadis and Manning, 2013; Jia and Persson, 2019; Manning and Roy, 2010; Masella, 2013; Tajfel and Turner, 1979; Turner et al., 1987). 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 1⁶; however, the ethnic composition 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 (Friedman

⁵For 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.

⁶Note 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 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.

⁷According to government statistics (Copper, 2019; Hsiau, 2003), 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.

⁸One recent study (Chiang, Liu and Wen, 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.

⁹Clots-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,

et al., 2016; Milligan, Moretti and Oreopoulos, 2004), 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 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

choose more 'German' first names for their children, and be less likely to volunteer in World War II.

¹⁰Bai 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.

¹¹For 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.

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 2, we discuss the background of the curriculum reform and analyze the differences between the old and new curricula. Section 3 describes the data and the sample used in this paper, and section 4 discusses our identification strategy—a regression discontinuity design. Section 5 presents our main results, following which we then explore potential mechanisms through a subgroup analysis in section 6 and long-term effects in section 7. Finally, section 8 concludes.

2 Policy Background

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 Appendix A 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 1987. They were the first to study the *Knowing Taiwan* series and were born between September 1984 and August 1985.

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, Huang and Vickers, 2005; Wang, 2001). In particular, the new history textbooks moved away from the "China-oriented" angle seen in earlier textbooks, to a "Taiwan-oriented"

¹²The 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.

¹³This 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.

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.

2.2.1 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

¹⁴In 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.

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.

2.2.2 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

¹⁵Emphasis in this paragraph is added by the authors.

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 **B**, we use several sample paragraphs from such textbooks to show the differences in historical 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.

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.

¹⁶Between 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.

3 Data and Sample

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 longterm 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.

3.2 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

¹⁷Some 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.

¹⁸Note 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.

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 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 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

4 Empirical Specification

4.1 Graphical Evidence

Figure 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).

¹⁹Since 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 6 we find this is not the case.

 $^{^{20}}$ 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.

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

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:

$$Identity_{it} = \alpha_0 + \alpha_1 TextBook_i + f(m;\beta) + X'_i \gamma + \lambda_t + \epsilon_{it}$$
(1)

where *Identity*_{it} indicates the dummy variable defined in section 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 (1) within a bandwidth of 24 months before and 24 months after September

 $^{^{21}}$ 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.

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 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 C.1 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

²²Junior 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.

more likely to have a Taiwanese identity (i.e., call themselves Taiwanese only).²⁴ Figure C.2 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 influencing national identity formation, such as local support for a certain political party.²⁶ In the Online Appendix D, we provide detailed definition of these individual characteristics. Finally, standard errors are clustered at the birth cohort level (i.e. birth year-month).

Table 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 1 suggest that the characteristics of treatment and control groups are quite

balanced.

²⁴Since 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.

²⁵We 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.

²⁶There 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.

5 Results

5.1 The Effect of Curriculum Reform on Taiwanese Identity

Figure 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 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 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 suggests that the discontinuity of *Identity* is roughly 20 percentage points around the cut-off.

Table 2 shows the regression results of the estimating specification (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, 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

²⁷The graph is at the birth year-quarter level, so the first dot in Figure 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 6, Figure 8, and Figure 9.

²⁸This is the mean of *Identity* across all those in the sample who were born between September 1982 and August

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.

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 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% (Adena et al., 2015; Blouin and Mukand, 2019; Chiang and Knight, 2011; DellaVigna and Kaplan, 2007; DellaVigna et al., 2014; Enikolopov, Petrova and Zhuravskaya, 2011; Gentzkow, Shapiro and Sinkinson, 2009; Gentzkow, 2006; Gerber, Karlan and Bergan, 2009; Yanagizawa-Drott, 2014). Our estimate, however, aligns with the persuasion rate found in Cantoni et al. (2017), 1984 (i.e. the control group).

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.

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.

5.3.1 Choices of Polynomial Order and Bandwidth

To examine whether our results are sensitive to different parametric specifications, Table 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 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 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 2 are not sensitive to bandwidth choice.²⁹

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

5.3.2 Exclude Specific Birth Cohorts

Based on Figure 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 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.

5.3.3 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 and Figure C.3 in the Online Appendix suggests that the estimated magnitudes of textbook effect are fairly similar across different waves.³⁰

5.3.4 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 (1) without controlling for the covariates X_i . The regression results are shown in Table 4. Most observable characteristics do not exhibit significant discontinuities at the cut-off.

 $^{^{30}}$ 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).

The only exception is the share of Hoklo people in the hometown. The sixth column of Table 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 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. 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.

5.3.5 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, Jansson and Ma, 2018, 2020; McCrary, 2008). Figure C.4 displays the results for the density test and suggests that there is no discontinuity in the distribution of the running variable at the threshold.

5.3.6 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 (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 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 comparable to our main results.

Panel A to D of Table 5 show the results of the falsification regressions. The estimated "treatment effects" are generally insignificant and the magnitudes are quite small. 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 December 1981 – to obtain the distribution of the placebo estimates. Figure 5 compares the real estimates with these placebo ones. We find that among the estimates based on 384 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 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 E. 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 F.1 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.

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

5.3.7 Difference-in-Differences Design

In this section, we generalize the placebo tests in Table 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 September in each year and estimate the following regression:

$$Identity_{it} = \kappa_0 + \kappa_1 A fter Sep_i + \kappa_2 B_{1984} + \kappa_3 A fter Sep_i \times B_{1984} + s(m;\beta) + X'_i \gamma + \delta_t + \epsilon_{it}$$
(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 (1).

Panel D of Table 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 curricu-

³²There 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 1984.

lum 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 around 18 percentage points.

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:

³³People 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.

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 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 (1) and use the above dummy variables as outcomes. Panel A and B of Table 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 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 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 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.

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 (1) and conduct a similar RD analysis to that seen in section 5.1.

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.³⁴ 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. We categorize the respondents into two groups: academic track and vocational track.³⁵

Figure 6 displays the relationship between Taiwanese identity and birth cohorts by academic track (Figure 6a) and vocational track (Figure 6b) respondents. We observe a distinct jump around

³⁴These 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.

³⁵The 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.

the cut-off in Figure 6a but no such pattern in Figure 6b. Table 7 presents RD estimates based on equation (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*.³⁶ Specifically, we estimate the following regression.

$$Identity_{it} = \delta_0 + \delta_1 A cademic_i + \delta_2 Text Book_i + \delta_3 Text Book_i \times A cademic_i + g(m; \beta) + X'_i \gamma + \mu_t + \epsilon_{it}$$
(3)

Panel C of Table 7 display the estimated coefficient on *TextBook* \times *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 the textbooks. Our results suggest that the effect of a curriculum varies substantially according to the degree of exposure to textbook content.

 $^{^{36}}$ Equation (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.

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 C.2 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., 2011*b*), 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 G, We provide a map demonstrating Hoklo ethnicity distribution and find the 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.³⁷ Figure 7 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 8 displays the relationship between Taiwanese identity and birth cohorts separately for respondents living in the towns with a low Hoklo share (Figure 8a) and a high Hoklo share (Figure 8b). For the former group of respondents, Figure 8a 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 8b). Consistent with the graphical evidence, Panel A of Table 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 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) 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 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 arrange-

³⁷In order to include those adults whom children are more likely to meet, we drop any respondents aged below 25.

ment 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 F.2 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).³⁸ In other words, the curriculum effect is heterogeneous for individuals living in the area with different proportions of Hoklo people.³⁹

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 F.3 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 F.3) but had a small impact on those living in areas populated with larger Hoklo and Hakka ethnic groups (see Panel B of Table F.3). 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 F.3). 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 C.5 in the Online Appendix).⁴⁰

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'

³⁸Specifically, 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, 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.

³⁹We 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).

 $^{^{40}}$ 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 7 using variations in the share of Hoklo people.

parents, dividing respondents into those whose parents both have Hoklo ethnicity, and others. Figure C.6 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%).⁴¹ 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 F.4), but it had insignificant impact on respondents whose parents had Hoklo ethnicity (see Panel B of Table F.4), 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 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.

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 H.2 compares the characteristics of long-term

⁴¹Again, we use 1992, 1995, and 1998 TSCS data and restrict the sample to people who are 25 years old or above.

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.⁴²

Figure 9 displays the relationship between Taiwanese identity and birth cohorts for the longterm sample.⁴³ 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 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 H and I, we examine the validity of the RD design for the long-term sample (see Tables H.2 to H.3 and Figure I.2).⁴⁴ In addition, we conduct a series of robustness checks and find that our estimates are robust to different specifications (see Table H.4 and Figure I.3) and bandwidth choices (see Figure I.1).⁴⁵

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 10 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

⁴²The schooling years of parents are also different. However, the difference is small (around 5% differences) compared to baseline mean.

⁴³As in our main results, we measure birth cohorts at the year-month level and plot average *Identity* after controlling for survey year fixed effects.

⁴⁴Based on the results in Tables H.2 and H.3, we find that the observable characteristics are fairly comparable between the treatment and control groups. Moreover, Figure I.2 suggests that the density of the running variable (birth cohort) is quite smooth at cutoff.

⁴⁵We also implement similar placebo tests shown in section 5.3.6 and find null effects.

period.46

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,⁴⁷ 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.

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

⁴⁶The 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%).

⁴⁷For 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.

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.
References

- Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya. 2015. "Radio and the Rise of the Nazis in Prewar Germany." *The Quarterly Journal* of Economics, 130(4): 1885–1939.
- Aghion, Philippe, Xavier Jaravel, Torsten Persson, and Dorothee Rouzet. 2018. "Education and Military Rivalry." *Journal of the European Economic Association*, 17(2): 376–412.
- Akerlof, George A, and Rachel E Kranton. 2000. "Economics and identity." *The Quarterly Journal of Economics*, 115(3): 715–753.
- Alesina, Alberto, and Bryony Reich. 2015. "Nation Building." National Bureau of Economic Research 18839, Cambridge. NBER Working Paper.
- Alesina, Alberto, and Nicola Fuchs-Schündeln. 2007. "Good bye Lenin (or not?): The effect of communism on people's preferences." *American Economic Review*, 97(4): 1507–1528.
- Anderson, Benedict. 2006. Imagined Communities: Reflections on the Origin and Spread of Nationalism. London:Verso.
- **Bai, Yu, and Yanjun Li.** 2018. "Good Bye Chiang Kai-shek? The Long-Lasting Effects of Education under the Authoritarian Regime in Taiwan." *Working Paper*.
- **Besley, Timothy, and Torsten Persson.** 2010. "State capacity, conflict, and development." *Econometrica*, 78(1): 1–34.
- Billig, Michael. 1995. Banal Nationalism. London: SAGE.
- **Bisin, Alberto, and Thierry Verdier.** 2010. "The economics of cultural transmission and socialization." National Bureau of Economic Research 16512, Cambridge. NBER Working Paper.

- **Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou.** 2011*a*. "Ethnic identity and labour market outcomes of immigrants in Europe." *Economic Policy*, 26(65): 57–92.
- **Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou.** 2011*b*. "Formation and persistence of oppositional identities." *European Economic Review*, 55(8): 1046–1071.
- Blouin, Arthur, and Sharun W Mukand. 2019. "Erasing Ethnicity? Propaganda, Nation Building, and Identity in Rwanda." *Journal of Political Economy*, 127(3): 1008–1062.
- Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang. 2017. "Curriculum and Ideology." *Journal of Political Economy*, 125(2): 338–392.
- Cattaneo, Matias D, Michael Jansson, and Xinwei Ma. 2018. "Manipulation testing based on density discontinuity." *The Stata Journal*, 18(1): 234–261.
- **Cattaneo, Matias D, Michael Jansson, and Xinwei Ma.** 2020. "Simple local polynomial density estimators." *Journal of the American Statistical Association*, 115(531): 1449–1455.
- Chiang, Chun-Fang, and Brian Knight. 2011. "Media bias and influence: Evidence from newspaper endorsements." *The Review of Economic Studies*, 78(3): 795–820.
- Chiang, Chun-Fang, Jin-Tan Liu, and Tsai-Wei Wen. 2019. "National identity under economic integration." *Journal of Population Economics*, 32(2): 351–367.
- Clots-Figueras, Irma, and Paolo Masella. 2013. "Education, Language, and Identity." *The Economic Journal*, 123: F332–F357.
- **Constant, Amelie F, and Klaus F Zimmermann.** 2013. "Immigrants, ethnic identities and the nation-state." In *International Handbook on the Economics of Migration*. Edward Elgar Publishing.
- Copper, John Franklin. 2019. Taiwan: Nation-state or province? Routledge.

- **DellaVigna, Stefano, and Ethan Kaplan.** 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics*, 122(3): 1187–1234.
- **DellaVigna, Stefano, and Matthew Gentzkow.** 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics*, 2(1): 643–669.
- DellaVigna, Stefano, Ruben Enikolopov, Vera Mironova, Maria Petrova, and Ekaterina Zhuravskaya. 2014. "Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia." American Economic Journal: Applied Economics, 6(3): 103–132.
- **Durante, Ruben.** 2020. "Building Nations Through Shared Experience: Evidence from African Football." *The American Economic Review*, 110(5): 1572–1602.
- Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. "Media and political persuasion: Evidence from Russia." *American Economic Review*, 101(7): 3253–85.
- Fouka, Vasiliki. 2019. "Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I." *The Review of Economic Studies*, 87(1): 204–239.
- Friedman, Willa, Michael Kremer, Edward Miguel, and Rebecca Thornton. 2016. "Education as Liberation?" *Economica*, 83(329): 1–30.
- Gentzkow, MA, Jesse M Shapiro, and Michael Sinkinson. 2009. "Media market structure and political participation: historical evidence from US newspapers." *University of Chicago mimeo*.
- Gentzkow, Matthew. 2006. "Television and voter turnout." *The Quarterly Journal of Economics*, 121(3): 931–972.
- Georgiadis, Andreas, and Alan Manning. 2013. "One nation under a groove? Understanding national identity." *Journal of economic behavior & organization*, 93: 166–185.

- Gerber, Alan S, Dean Karlan, and Daniel Bergan. 2009. "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.
- Ghitza, Yair, and Andrew Gelman. 2022. "The great society, Reagan's revolution, and generations of presidential voting." *American Journal of Political Science*, forthcoming.
- Hsiau, A-chin. 2003. Contemporary Taiwanese cultural nationalism. Routledge.
- Hsu, Yu-Chin, and Shu Shen. 2019. "Testing treatment effect heterogeneity in regression discontinuity designs." *Journal of Econometrics*, 208(2): 468–486.
- **Ito, Koichiro.** 2015. "Asymmetric incentives in subsidies: Evidence from a large-scale electricity rebate program." *American Economic Journal: Economic Policy*, 7(3): 209–37.
- Jacobs, J Bruce, and Peter Kang. 2017. Changing Taiwanese Identities. Routledge.
- Jia, Ruixue, and Torsten Persson. 2019. "Individual vs. social motives in identity choice: Theory and evidence from China." National Bureau of Economic Research 26008, Cambridge. NBER Working Paper.
- Liu, Mei-Hui, Li-Ching Huang, and Edward Vickers. 2005. "History Education and National Identity in East Asia.". 1 ed., Chapter Identity Issues in Taiwan's History Curriculum. Abingdonon-Thames:Routledge Taylor Francis Group.
- Manning, Alan, and Sanchari Roy. 2010. "Culture Clash or Culture Club? National Identity in Britain." *Economic Journal*, 120: F72–F100.
- Masella, Paolo. 2013. "National identity and ethnic diversity." *Journal of Population Economics*, 26(2): 437–454.
- **McCrary, Justin.** 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of econometrics*, 142(2): 698–714.

- Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does education improve citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics*, 88(9): 1667–1695.
- Mitrunen, Matti. 2018. "Can You Make an American? Compulsory Patriotism and Assimilation of Immigrants." *Working Paper*.
- Tajfel, Henri, and John C Turner. 1979. "An integrative theory of intergroup conflict." In Organizational Identity: A Reader., ed. Mary Jo Hatch and Majken Schultz, 56–65. Oxford:Oxford University Press.
- Turner, John C, Michael A Hogg, Penelope J Oakes, Stephen D Reicher, and Margaret S Wetherell. 1987. Rediscovering the social group: A self-categorization theory. Oxford:Basil Blackwell.
- **Voigtländer, Nico, and Hans-Joachim Voth.** 2015. "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.
- Wang, Fu-Chang. 2001. "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.
- Weber, Eugen. 1976. *Peasants into Frenchmen: the modernization of rural France, 1870-1914.* Stanford:Stanford University Press.
- **Yanagizawa-Drott, David.** 2014. "Propaganda and conflict: Evidence from the Rwandan genocide." *The Quarterly Journal of Economics*, 129(4): 1947–1994.

Figures



Figure 1: Trends of National Identity in Taiwan: 1992-2015

Source: Election Study Center, National Chengchi University



Figure 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: 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 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 4: RD Estimates across Different Bandwidth Choices

Notes: We run regressions as column (4) in Table 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.



Figure 5: 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 December 1981 (384 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).



Figure 6: 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 6a includes respondents whose final education level is senior high school or university. Figure 6b 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 7: 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

Figure 8: 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 8a includes respondents living in the towns with low share of Hoklo people before age 15. Figure 8b 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 9: 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 *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 10: 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.

Tables

	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	(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
1	(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	

Table 1: Descriptive	Statistics for	Treatment	Group and	Control	Group
----------------------	----------------	-----------	-----------	---------	-------

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 D, 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.

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

Table 2: The Effects of the Curriculum Reform on Taiwanese Identity: Main Results

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 (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 firstorder 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.

Taiwanese Identity				
	(1)	(2)	(3)	(4)
Panel A: 2nd Order Poly	nomial			
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
	17	N7	V	N
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

Table 3: Robustness Check: Different Specification, Sample, Identity Questions

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 (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 2. Standard errors are clustered at the birth year-month level in parentheses. *** p < 0.01, ** p < 0.05, and * p < 0.1.

	lat	ole 4: Ko	bustness Ch	eck: Validity of	Regression Disco	ontinuity Desig	u	
VARIABLES	Female	Age	Years of Schooling	Father's Schooling years	Mother's Schooling years	Share of Hoklo people	Hoklo Fathers	Hoklo Mothers
TextBook	$\begin{array}{c} 0.044 \\ (0.065) \end{array}$	$\begin{array}{c} 0.002\\ (0.005) \end{array}$	-0.219 (0.250)	0.015 (0.599)	0.115 (0.431)	-0.075* (0.043)	-0.017 (0.076)	-0.053 (0.059)
Sample Size	427	427	427	427	427	427	427	427
tes: We pool data	from the 2(003, 2004,	and 2005 TSC	S waves and use th	e sample born betwe	en September 196	82 and Aug	ust 1986. The above
le reports the coel	ficient of 7	<i>TextBook</i> ł	based on equat	tion (1), which is or	ie if the birth year-n	nonth of the respo	ondent is af	ter September 1984
o otherwise. Note	that we do	o not inclue	de any covaria	ttes X_i since they at	e outcome variables	now. All columi	ns include t	he survey year fixed
ect and the first-or	der polyno	mials of bi	irth year-mont	h m interacting full	y with <i>TextBook</i> (i.e	e. linear spline). 7	The definiti	ons of the individua

	ы
•	\mathbf{S}
(ร
•	nuty
•	contir
;)ISC
ĥ	
•	gression
	õ
,	~
1	-
ç	H
	0
;	Ξ
÷	≚
Ľ.	đ
÷	>
	4
	20
7	Ă
`	-
	S
	g
	Ë
	us
÷	ā
	Q
۴	Y
_	· .
1	4
÷	9
-	9
F	5

(vocational) high school (12 years of schooling); e) two-year college (14 years of schooling); f) University or vocational university (16 years of a) no education (zero years of schooling); b) elementary school (6 years of schooling); c) junior high school (9 years of schooling); d) senior schooling). 3) Hoklo fathers/mothers: If an individual's father/mother is Hoklo assigned 1, otherwise 0. In the Online Appendix D, we provide characteristics are as follows: 1) Female: If an individual is female assigned 1, otherwise 0. 2) Respondent/Father/Mother's schooling years: 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. Notes table 1 effect I zero (

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

Table 5: Robustness Check: Placebo Test of Fake Textbook Reform

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 (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 (2). In this specification, we combine all available cutoffs used in the main estimation and placebo tests to implement a DID design. Specifications in each5column are the same as in the corresponding columns in Table 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.

-	(1)	(2)	(3)	(4)
Panel A: Support Taiwar	n Independ	dence Unc	conditional	ly
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

Table 6: Effects of the Curriculum Reform on Preferences for Taiwan's Independence

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 (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 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 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.

Taiwanese Identity				
	(1)	(2)	(3)	(4)
Panel A: Academic Trac	k			
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 Trac	:k			
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 Heterogen	eity			
TextBook	0.055	0.059	0.028	0.017
	(0.104)	(0.101)	(0.096)	(0.095)
TextBook imes 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

Table 7: Subgroup Analysis: By Education Track

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 (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 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 academic track students Academic. Compared to equation (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.

	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 Hokle	o Proportion	l		
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 Heterogene	eity				
TextBook	0.024	0.013	0.016	0.028	
	(0.103)	(0.109)	(0.103)	(0.104)	
TextBook imes 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	

 Table 8: Subgroup Analysis: By Hometown Ethnicity Distribution

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 (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 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 (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, 58 ** p<0.05, * p<0.1.

Taiwanese Identity				
	(1)	(2)	(3)	(4)
TextBook	0.006 (0.052)	0.023 (0.052)	0.014 (0.050)	0.011 (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

Table 9: The Effects of the Textbook Reform on Taiwanese Identity: Long-Term Results

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 (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 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.

Online Appendix: For Online Publication

Section A	1997 Curriculum Reform: More Information			
Section B	Comparison of Textbooks: Sample Paragraphs from Old and New Textbooks			
Section C	Additional Figures			
Section D	Definitions of Individual Characteristics			
Section E	Questions on Other Social Values			
Section F	Additional Tables			
Section G	Hoklo Ethnicity Distribution in Taiwan			
Section H	Robustness Checks for Long-term Results: Ta- bles			
Section I	Robustness Checks for Long-term Re- sults:Figures			

A 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.

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

Figure B.1 and B.2 display the table of contents of the old history textbook. As mentioned in section 2.2.1, the old History textbook dedicated 25 chapters to Chinese history and only one chapter and a section to Taiwan. Figure B.1 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 B.2 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 B.3 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 2.2.2, 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 B.4 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 B.5 displays a similar paragraph in the new textbook on Chinese history, which explicitly mentions "China" as originating from Peking Man and cavemen.

Figure B.6 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 B.7 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".

國立編譯館



Figure B.1: Table of Contents from the Old History Textbook



Figure B.2: Table of Contents from the Old History Textbook

	認識臺灣(歷史篇)		第六章 第一節	清領時代後期 開港與國際貿易	45 45
			第二節 第三節	日軍侵臺與清廷治臺政策的改變 建省後的積極建設	49 52
第一章	導論	1	做上来	日本庭民统选电机研究达网络演	57
2463 -252			第七早	日本但以就行时期的政府與輕荷	27
第二章	史前時代	5	· · · · · · · · · · · · · · · · · · ·	· 如何以上國央政权化目 和公布(社会物)	57
第一節	文化演進	5	5行—ED 202 一 202	政 石 兴 虹 唐 12 时 7 雨 弓 20 1 声 6 声 同	62
第二節	原住民社會	10	弗二即	用山风海里的其时为亚根	00
Nov	End they doly with units (bit)		第八章	日本殖民統治時期的教育、學術	
第三章	國際競爭時期	14		與社會	71
第一節	漢人與日本人的活動	14	第一節	教育與學術發展	71
第二節	荷蘭人與西班牙人的統治	18	第二節	社會變遷	76
101			第三節	社會運動	81
第四章	鄭氏治臺時期	23			
第一節	政治與文教	23	第九章	中華民國在臺灣的政治變遷	86
第二節	墾殖與貿易	27	第一節	初期的政治	86
267 J.W			第二節	中央政府遷臺後的政治發展	90
第五章	清領時代前期	31	第三節	外交與兩岸關係	96
第一節	政治演變	31			
第二節	經濟活動	35	第十章	中華民國在臺灣的經濟、文教	
第三節	社會與文教發展	39		與社會 101	
	. Manufactoria	1	第一節	經濟發展 101	
			第二節	教育與文化 106	
			第三節	社會變選 111	
	THE WE ADDRESS OF THE TATE OF THE	W.m.	htte 1	5	

Figure B.3: Table of Contents from the New History Textbook

一、時間悠久 我國的歷史,從黃帝建國算	納起來,其特色有下列四點:	文明的主體,在世界歷史上占有重要的地位。歸	我國歷史的特色 我國歷史悠久,是東亞	形態了。	現代人尚有一段距離,而真人則已具有現代人的	在法國發現的「克魯麥囊人」。猿人的形態比起	著名:一是在我國發現的「山頂洞人」;另一是	的「尼安德人」。在各種真人當中,有兩種比較	三是在我國發現的「北京人」;四是在德國發現	的「東非人」;二是在印尼發現的「爪哇人」;	,比較有名的至少有四種:一是在非洲東岸發現	了很多猿人和真人的情形。目前已經發現的猿人	現。我們從地下發現的史前人類化石之中,知道	和現代人之間的「 猿人 」,後有「真人」的出	史前時代的人類,先是體質特徵介於人形猿	大概的活動情形了。	國中歷史(第一册)
GR		法國長期	同席の後	いのない	No. Com		Ser And	3	5 # A & 2		単 「 」				and the second		

Figure B.4: Old Textbook: Our Country (China)

用火取暖、照明和燒烤食物。

至於<u>中國</u>境內舊石器時代晚期的人類,則以「<u>山頂洞人</u>」為代 表。「<u>山頂洞人</u>」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「<u>北京</u> 人」進步得多。

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



Figure B.5: New Textbook: China

分之八十左右;至七十九學年度業已近於百分之	度,臺灣地區六至十一歲學齡兒童的就學率為百	文教建設 普及教育方面:民國三十九學年	已為舉世所公認。	,躋身「亞洲四小龍」之列,經濟方面的成就,	目前,我國由於雄厚的工業基礎與外貿潛力	十二項建設的延續,具有前瞻性的大工程。	,又推出十四項重要建設計畫,多為前述十大、	推動交通、工業、農業等十二項建設。七十三年	合稱十大建設,皆陸續完成。六十八年,政府又	船廠、鐵路電氣化及桃園國際機場等重要建設;	鐵路、蘇澳港、石油化學工業、大煉鋼廠、大造	並限期五年内完成南北高速公路、臺中港、北迴	經國鄭重宣布:政府除積極興建核能發電廠外,	重大建設;六十二年十一月,行政院院長蔣	破。

Figure B.6: Old Textbook: Our Country (Taiwan)



圖10-7 花蓮國際港開放 中華民國在臺灣的經濟、

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 B.7: New Textbook: Taiwan

C Additional Figures



Figure C.1: 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.


Figure C.2: 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.



Figure C.3: RD Graph for Different Identity Questions

Notes: Figure C.3a is based on 2003 and 2004 TSCS. Figure C.3b 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 C.4: Density of Birth Cohort: Density Discontinuity Test

Notes: This figure displays the results for a density discontinuity test proposed by Cattaneo, Jansson and Ma (2018, 2020). 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.



Figure C.5: 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 C.6: 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.

D 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 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.

E Questions on Other Social Values

We construct binary outcome variables used in Table F.1, 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."

F Additional Tables

	(1)	(2)	(3)	(4)
Panel A: One's success n	elies on co	ming from	a rich fami	ly
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 inc	ome distrib	oution make	es people w	ork 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" h	nave a colle	ge 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 Ŷear FE	Yes	Yes	Yes	Yes
Ethnic variables	No	Yes	Yes	Yes
Demographic variables	No	No	Yes	Yes
Regional Dummies	No	No	No	Yes

Table F.1: Effects of the Curriculum Reform on Attitude Towards Other Social Values

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 (1), which is one if the birth yearmonth 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 E. 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.

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

Table F.2: Heterogeneous RD analysis: By Share of Hoklo People in the Hometown

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.

Taiwanese Identity						
	(1)	(2)	(3)	(4)		
Panel A: Hometown with	n Low Proj	portion 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 Heterogene	eity					
TextBook	0.079	0.097	0.099	0.108		
	(0.079)	(0.076)	(0.078)	(0.081)		
TextBook imes 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		
County Dummies	No	No	No	Yes		

Table F.3: Subgroup Analysis: By Hometown Ethnicity Distribution (Hoklo and Hakka)

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 (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 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 (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.

	Taiwanese	Identity				
	(1)	(2)	(3)	(4)		
Panel A: One of Parents	are not Hok	lo 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: 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 C: Test Heterogene	eity					
TextBook	0.085	0.088	0.106	0.102		
	(0.102)	(0.101)	(0.101)	(0.105)		
TextBook imes 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 Ŷear FE	Yes	Yes	Yes	Yes		
Ethnic variables	No	Yes	Yes	Yes		
Demographic variables	No	No	Yes	Yes		
County Dummies	No	No	No	Yes		

Table F.4: Subgroup Analysis: By Ethnicity of Respondents' Parents

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 (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, while Panel B includes respondents whose parents are both Hoklo. Specifications in each column are the same as in the corresponding columns in Table 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 (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 birth year-month level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

G Hoklo Ethnicity Distribution in Taiwan



Figure G.1: 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.

H Robustness Checks for Long-term Results: Tables

	Long-run	Short-run	Difference
	Sample	Sample	(long-short)
Female	0.477	0.445	0.032
Age	(0.500)	(0.498)	(0.030)
	28.923	20.396	8.527***
Years of schooling (self)	(2.127)	(1.104)	(0.091)
	13.803	13.761	0.042
Years of schooling (father)	(2.083)	(2.021)	(0.122)
	11.034	10.600	0.435**
Vears of schooling (mother)	(3.453)	(3.450)	(0.206)
Properties of Helde in the howetown	(3.582)	(3.310)	(0.203)
Proportion of Hokio in the nometown	(0.137)	(0.215)	(0.013)
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	

Table H.1: Descriptive Statistics for Long-run and Short-run Sample

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 D, 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.

	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***
6	(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	

Table H 2 [.] Descrip	ptive Statistics for	Treatment Grour	o and Control Group

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 D, 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.

	1401	N .C.H 2		HECK. VAIIUILY U.	I NEBIESSIOII DISU	community mes	Igu	
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* (0.052)	0.044 (0.052)	0.009 (0.273)	0.417 (0.505)	0.258 (0.500)	0.048 (0.030)	$\begin{array}{c} 0.025\\ (0.053) \end{array}$	0.064 (0.042)
Sample Size	822	822	822	822	822	822	822	822
tes: We pool data	from the 2	2010, 2012	, 2013, 2014,	and 2015 TSCS wi	aves and use the san	nple born betwee	en Septemb	er 1982 and August
36. The above tab	le reports t	the coeffici	ent of TextBo	ok based on equatio	in (1), which is one i	if the birth year-r	nonth of th	e respondent is after
stember 1984, zeru	otherwise	e. Note tha	tt we do not ir	sclude any covariate	is X_i since they are c	outcome variable	s now. All	columns include the
vey year fixed effe	set and the	first-order	polynomials c	of birth year-month	n interacting fully w	ith <i>TextBook</i> (i.e	. linear spl	ine). The definitions
		11 J		1I UI .I.	ן. י <u>ן</u> יו	· · · ·		

Table H.3: Robustness Check: Validity of Regression Discontinuity Design

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 university (16 years of schooling). 3) Hoklo fathers/mothers: If an individual's father/mother is Hoklo assigned 1, otherwise 0. In the Online schooling); d) senior (vocational) high school (12 years of schooling); e) two-year college (14 years of schooling); f) University or vocational Appendix D, 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. Notes 1986. Septe surve

Tai	wanese Id	lentity				
	(1)	(2)	(3)	(4)		
Panel A: 2nd Order Polyn	nomial					
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 Questio	on – 2012-	2014 Wav	'es			
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		
Linaar/Quadratia Splina	Vac	Vac	Vac	Vac		
Survey Vear EE	Vec	Ves	Vec	Vec		
Ethnic variables	No	Ves	Ves	Ves		
Demographic variables	No	No	Ves	Ves		
Regional Dummies	No	No	No	Yes		
TextBookSample SizePanel C: Identity QuestionTextBookObservationsPanel D: Identity QuesticTextBookObservationsLinear/Quadratic SplineSurvey Year FEEthnic variablesDemographic variablesRegional Dummies	$\begin{array}{r} 0.009\\ (0.056)\\ 774\\ \hline m-2012-\\ -0.010\\ (0.075)\\ 426\\ \hline m-2010,\\ 0.017\\ (0.067)\\ 396\\ \hline \\ Yes\\ Yes\\ No\\ No\\ No\\ No\\ No\\ No\\ \end{array}$	0.034 (0.056) 774 2014 Wav 0.017 (0.075) 426 2014-201 0.041 (0.070) 396 Yes Yes Yes Yes No No	0.032 (0.055) 774 ves 0.027 (0.076) 426 5 Waves 0.031 (0.071) 396 Yes Yes Yes Yes Yes No	0.030 (0.054) 774 0.029 (0.075) 426 0.022 (0.071) 396 Yes Yes Yes Yes Yes Yes Yes		

Table H.4: Robustness Check: Different Specifications, Sample, and Identity Questions

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 (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 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.

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 C	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		

Table H.5: Robustness Check: Placebo Test of Fake Textbook Reform

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 (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 B₁₉₈₄ in the equation (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 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.

I Robustness Checks for Long-term Results: Figures



Figure I.1: RD Estimates across Different Bandwidth Choices: Long-term Results

Notes: We run regressions as column (4) in Table 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 I.2: 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, Jansson and Ma (2018, 2020). 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.



Figure I.3: RD Graph for Different Identity Questions

Notes: Figure I.3a is based on 2012-2014 TSCS. Figure I.3b 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? A first order polynomial of birth year-quarter interacted with *TextBook* dummy variable.



Figure I.4: Permutation Test

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 December 1981 (384 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).