Challenging Demographic Representativeness at State Borders: Implications for Policy Research

Benjamin S. Kay, Albina Khatiwoda

2025-018

Challenging Demographic Representativeness at State Borders: Implications for Policy Research[1]

Benjamin S. Kay

Benjamin.S.Kay@frb.gov

Federal Reserve Board of Governors

Albina Khatiwoda

Harvard Kennedy School

Abstract

This study examines the demographic characteristics of U.S. state border counties, comparing them with those of nonborder counties. The demographic representativeness of border counties is essential for the interpretation of the results in state border-county difference-in-difference analyses, used in state policy evaluations. Our findings reveal that border counties generally have higher proportions of White, older, and disabled populations. We also see occasional instances of wide demographic differences across state boundaries. These differences potentially undermine the external validity and identification of policy evaluations. We illustrate the implications of these finding through a case study, highlighting the need for robustness checks and demographic considerations in border-county policy research.

JEL Codes:

D78 - Positive Analysis of Policy Formulation and Implementation

J15 - Economics of Minorities, Races, Indigenous Peoples, and Immigrants

C21 - Cross-Sectional Models • Spatial Models • Treatment Effect Models

C23 - Panel Data Models • Spatio-temporal Models

Introduction

Effective public policy hinges on the ability to assess which policies work best. Unfortunately for the econometrician, policies are not implemented randomly, and therefore measuring efficacy requires careful econometric work, usually in opportune settings. Unlike in unitary countries like France, the United States' federal system devolves substantial governing authority to state governments (Bognetti and Tate, 2024). This structure fosters significant policy experimentation at the state level, a concept famously described by U.S. Supreme Court Justice Louis Brandeis as the "laboratories of democracy" (*New State Ice Co. v. Liebmann,* 1931). Studying the effects of these state policies has become an important part of the credibility revolution (Angrist and Pischke, 2010), which deployed a variety of techniques to exploit natural experiments for causal identification.

From the econometrician's perspective, properly executed randomized controlled trials (RCTs), which randomly assign subjects to treatment and control groups, present an ideal setting for estimating the causal effects of policy interventions, often using relatively straightforward econometric designs.[2] However, political and ethical constraints often impede the implementation of these experimental designs, complicating the ability to achieve clean causal identification of policy interventions. Despite the absence of RCTs, many studies have used state policy changes to identify and quantify the causal effects of these policies, notwithstanding their nonrandom implementation. To do so, scholars have frequently employed the difference-in-difference (DiD) estimator using the border counties of U.S. states to tease out these effects (see, for example, Card and Krueger, 1994; Holmes, 1998; Hanson and Sullivan, 2009, Kumar, 2018; and Hao and Cowan, 2020).

One precondition for extracting useful causal effect estimates from a nonexperimental setting is that the econometrician must account for the differences in climate, culture, proximity to major markets, welfare generosity, and a myriad of other characteristics (McKinnish, 2005; Giuntella, 2019). And yet, controlling for locational differences in cross-sectional state-level analysis risks improperly specifying the geographic differences and omitting other confounding variables. To better control for these confounders arising from state-level differences, researchers have employed border-county DiD estimates that compares treated border counties to the neighboring border counties in surrounding states to estimate the effects of policy treatments.[3],[4] Because borders are somewhat arbitrary, state lines might be expected to split demographically,

---

[2] In some cases, a basic *t*-test of population means is all that is required. However, see Grossman and Mackenzie (2005) and Cartwright and Munro (2010) for criticisms and limitations of RCTs.

[3] For example, to study a policy change in Florida, its border counties could be matched to neighboring counties in Alabama and Georgia.

[4] There are many other ways to solve this identification problem. Notable examples of these methods include instrumental variables, structural estimation, explicit modeling of the process of selecting into the policy choice, regression discontinuity, and various matching procedures. This paper focuses exclusively on border-county DiD analyses, but the problems identified in this article sometimes also exist with these other methods.

climatically, and economically similar communities. Therefore, there is a strong presumption that comparing counties across state boundaries will control for a lot of unobserved heterogeneity without the need for rich controls that erode statistical power.[5]

This paper challenges this assumption. We examine the reliability of DiD in this setting by investigating how demographically representative border counties are of the national population and the degree to which county demographics differ across state borders. For the DiD estimate to quantify the program's effects, we generally want these differences to be small (more on this later). To be confident these DiD estimates are useful elsewhere, we also want border counties to be similar to nonborder counties.

If border and nonborder counties differ too much, and effects vary across counties, potentially all we have learned is the average effects of the policy in border counties. Since border counties make up less than 20 percent of all counties, this may be a limited finding. And if the border counties used to measure the effects of a policy are unusual even among border counties, the findings may be of even narrower applicability. This point is about external validity—whether the findings generalize to other times and populations—not about identification, which concerns whether the analysis isolates the true effect of the policy change studied. In the next section, we will be more precise about the specific econometric assumptions necessary for the DiD method to identify causal effects. But, as we will show, border-county demographics can affect both external validity and identification.

We contribute to the border-county DiD literature by examining the demographic composition of border counties, assessing how representative these counties are of the national population and the extent of demographic differences across state borders. We also conduct a case study of demographics in the border counties studied in Dube et al. (2010), a seminal border-county DiD paper.

This article is not intended to vilify border-county policy analysis using DiD. Instead, we hope to improve future studies using this approach by strengthening their reliability, external validity, and identification.

---

[5] Lacombe (2004) summarizes this assumption nicely, "The common thread linking all of these studies is their attempt to control for un-observed spatial variation using strategic spatial selection of the sample observations. The rationale for this approach is that geographic differences should be minimized across state borders, while variation in policy impacts are more easily detected, producing more precise estimates of the public policy effects." Students of American history might be skeptical of this assumption. State borders may have a common cause with some of these economic, climactic, and demographic factors. For example, the institution of slavery significantly influenced the borders and demographics of states such as West Virginia and Missouri. Similarly, religious settlement patterns played a crucial role in shaping the borders of Utah and Rhode Island, resulting in enduring demographic impacts.

3

Identification and External Validity

The basic setup of a border-county DiD, as applied to a state policy program, is as follows. Researchers compare the outcomes of a border county that was exposed to the program (treatment state) with those of the border county or counties in the adjacent state(s) that were unexposed (control state) before and after the program implementation. Sometimes, the two states (treated and control) use a dummy variable so that the results of the treatment and control border counties are pooled together. Other times, each border county from the treatment state is matched based on geography or similarity across many dimensions to a single border county from the control state. In yet other instances, each treated border county is matched based on geography or similarity across additional dimensions to one or more untreated border counties. The sample may be limited to only a few states (Cosgrove et. al., 2016), or the entire sample of the contiguous US states (Huang, 2008), depending on if program existed in one or several states. Papers using a multi-state sample can include both dynamic (Thompson and Rohlin, 2013) and static panels (Peng, Xiaohui, and Meyerhoefer, 2020). The causal effect is estimated as the change in the difference between the variable of interest between the pooled/matched border counties of the treatment and control state.

DiD is particularly useful in the border-county setting. The method works in observational studies with modest data requirements and without random assignment. In terms of data, observational studies with DiD design require only a short data panel of treated and control units pre- and post-treatment. Regarding causal identification, the specification is even robust to time-invariant group differences (as in a fixed effect model) and time-varying changes in the variable of interest affecting both the treatment and control groups (assuming the identification assumptions are met).

Even the identification assumptions are somewhat modest: exchangeability (conditional mean independence of treatment and control units), positivity (0 < probability of a unit being treated < 1), and stable unit treatment value assumptions (SUTVAs, where potential outcomes of each unit are unaffected by the treatment assignment of other units). Commonly, these suppositions are summarized as assuming (a) a common trend and (b) no anticipation of the treatment, that absent the policy, the expected differences in the outcome variables between the treated and control units would have been constant.

If nearby counties are more similar than faraway ones, then this helps satisfy these assumptions. The border-county DiD model reduces the differences between the treated and control populations that are correlated with location. This feature of the model makes it more likely that both units could have been treated and have similar conditional means. It is also intuitively consistent with the SUTVAs, all helping satisfy the common trend assumption critical to DiD-based identification. If border counties are mostly similar, or similar after careful matching, in the absence of the policy treatment, the average difference in the outcome variable, Y between the treated and untreated populations would have stayed relatively constant (Holmes, 1998). Therefore, by comparing the (potentially matched) border counties after the policy treatment, the

4

observed differences between the treated and control groups can be attributed to the treatment and not to other state-level differences and shocks.

Border-county DiD analysis has been used to evaluate policy in an array of disciplines. Notable examples include manufacturing location (Holmes, 1998), labor economics (Dube et al., (2010), unemployment insurance (Dieterle et. al., 2020, Boone et. al., 2021), taxation (Hanson and Sullivan, 2009), political advertising (Huber and Arceneaux, 2007), public health (Lyu and Wehby, 2020), and banking regulation (Huang, 2008).

The method is especially prominent in labor economics. For example, Dube et al al. (2010) is an important paper, using border-county DiD analysis, on unemployment that is frequently cited in policy debate and research summaries supporting the finding that a higher minimum wage does not reduce employment outcomes among lower-skilled workers. However, Jha, Neuman, and Rodriguez-Lopez (2022) have challenged these findings by showing that the conclusion relies critically on defining the local economic areas used to capture spatial economic shocks as pairs of contiguous counties across states. Still, papers that criticize or modify the border DiD technique, like Jha, Neuman, and Rodriguez-Lopez (2022), are few and only recently emerging, while the technique remains an important one for public policy research.

Data and Methodology

Our demographics data are from the 2022 vintage of the American Community Survey (ACS) The ACS is a published survey that measures a broad range of social, economic, and housing characteristics of the U.S. population. We use the five-year estimates (ACS-5) because they cover all counties and equivalents, whereas the annual survey (ACS-1) does not. The five-year estimates also pool the results of the previous five years of annual surveys to provide more precise demographic estimates. We use the most recently available five-year data, from the 2022 survey, which pool the 2018–22 annual survey waves. We do not combine the 2022 data with earlier ACS-5 estimates. Demographic estimates, like race, tend to evolve slowly, and the ACS-5 data are already pooling results from the previous five years to provide a more precise estimate of these variables.[6]

We use five-digit county FIPS (Federal Information Processing Series) codes to match counties in the ACS data to data from Thompson and Rohlin (2013), which classify each county in the contiguous United States as either a border or nonborder county. We use Thompson and Rohlin's (2013) classification approach, which defines as a border county any county that lies on a border with another state, excluding all water borders except those river borders connected by a bridge or commercial ferry service.

---

[6] However, researchers concerned about the potential demographic confounding variables influencing their policy analysis should use demographic statistics contemporaneous to their policy experiment (or close to it).

Our sample data contain all contiguous U.S. counties and equivalents (3,109 counties).[7] However, as a robustness check, we also analyze all counties and equivalents (3,222 counties), since their outcomes are also policy relevant. From the sample of contiguous county and county equivalents, 590 are considered border counties. For our analysis using the ACS-5 demographic data, we use Equation (1) below, where Y is the variable of interest. We join the ACS data to the list of border counties.

$$Y_j = \alpha + \beta \cdot BorderCounty_j + \epsilon_j \qquad (1)$$

Demographic Differences across State Borders

We find statistically precise and economically meaningful differences in the racial composition of border counties (from nonborder counties), as shown in Table 1. Importantly, the White population share is 1.9 percentage points higher for border counties than nonborder counties. Table 1 also shows that border counties have, on net, significantly lower average population shares of Black and biracial Americans than nonborder counties do. Border counties do not differ significantly in their population share that is Asian, Pacific Islander, or Native American, but given the smaller shares of these populations, this lack of significance may reflect small true effect sizes, observed without precision, rather than a null effect.

Notably, Hispanic, or Latino-origin, residents are underrepresented in border counties. The Census Bureau regards anyone as Hispanic based on their "heritage, nationality, lineage, or country of birth of the person or the person's parents or ancestors," so the Hispanic estimate includes Hispanics of any race, including White (U.S. Census, 2022). The Hispanic population share is 1.8 percentage points lower for border counties than nonborder counties. This estimate is large relative to the average county-level Hispanic population share in the United States (10.4 percent in the contiguous U.S.). When we expand the sample to all U.S. counties and county equivalents, including Alaska, Hawaii, and U.S. territories, we find that the racial and Hispanic results are consistent and show, if anything, more significant and larger effects (Appendix Table A.1). Additionally, when basic regression specifications are weighted by county population, the results are broadly similar (Appendix Table A.2).

In Figure 1, we show the entire distribution of counties' the non-white population percentage (a measure of racial makeup) among US border and nonborder counties. In line with our earlier findings, the figure shows more border county representation among the left tail of the distribution meaning that border counties tend to have less non-White population on average compared to all counties. In contrast, towards the middle of the distribution and on the right tail,

---

[7] County equivalents include Alaska boroughs, municipalities, city and boroughs, and census areas; the District of Columbia; Louisiana parishes; Puerto Rico municipios; independent cities in Maryland, Missouri, Nevada, and Virginia (U.S. Census, 2024).

the percentage of non-White populations among all counties appear to be broadly larger than that of border counties.

	White	Hispanic	Black	Asian	Biracial	Islander	Native
Border County	1.9[**]	-1.8[***]	-1.3[**]	-0.2	-0.5[***]	0.01	0.3
	(0.7)	(0.5)	(0.6)	(0.1)	(0.2)	(0.01)	(0.3)
Intercept	79.1[***]	10.4[***]	9.2[***]	1.4[***]	6.0[***]	0.1[***]	1.6[***]
	(0.3)	(0.3)	(0.3)	(0.1)	(0.1)	(0.00)	(0.1)
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109
R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Adjusted R^2	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00

Note: Robust standard errors. `Hispanic` represents those who idenitfy as Hispanic from any race. `Islander` represents Hawaiian or Pacific Islander. `Native` represents Native American.

*$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.

Source: U.S. Census Bureau, 2022 American Community Survey (five year estimates).

Figure 1: Histogram of Non-White Population Percentage by County Type
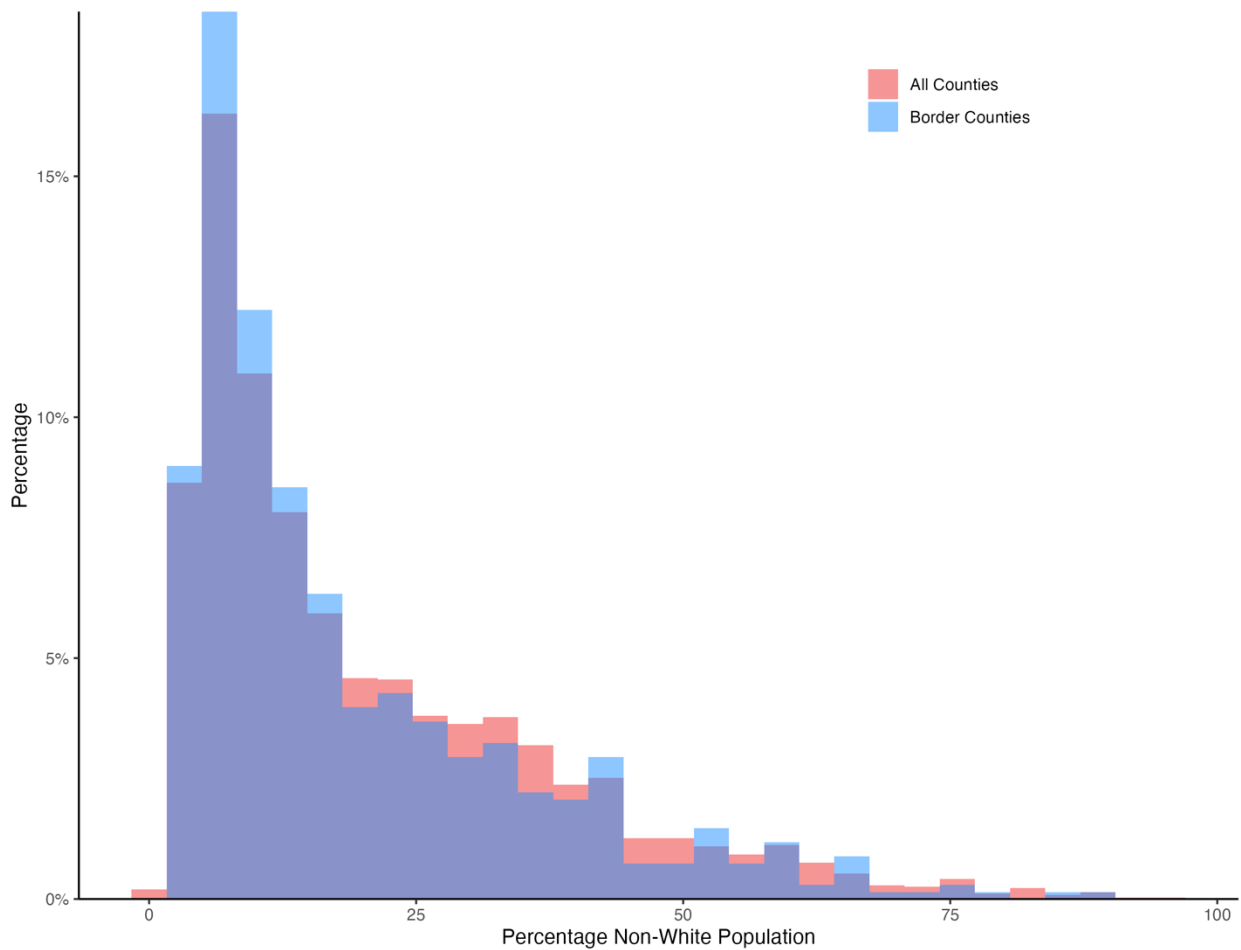
Source: U.S. Census Bureau, 2022 American Community Survey (five-year estimates).

7

Additionally, we find that there occasionally are vast differences in population shares across state boundaries. Table 2 shows the Hispanic population shares in border counties for the 10 states with the sharpest difference in their average concentration of the Hispanic population. The population share averages can differ by almost 40 percentage points. These gaps are evident for most of the other racial groups, including Whites. Appendix B shows that state pair border-county White population shares can differ by as much as 30 percentage points. The Black population shares can differ by similar magnitudes (not shown).

Table 2: Hispanic Population Percentage in Border Counties among the State Pairs with the Greatest Differences

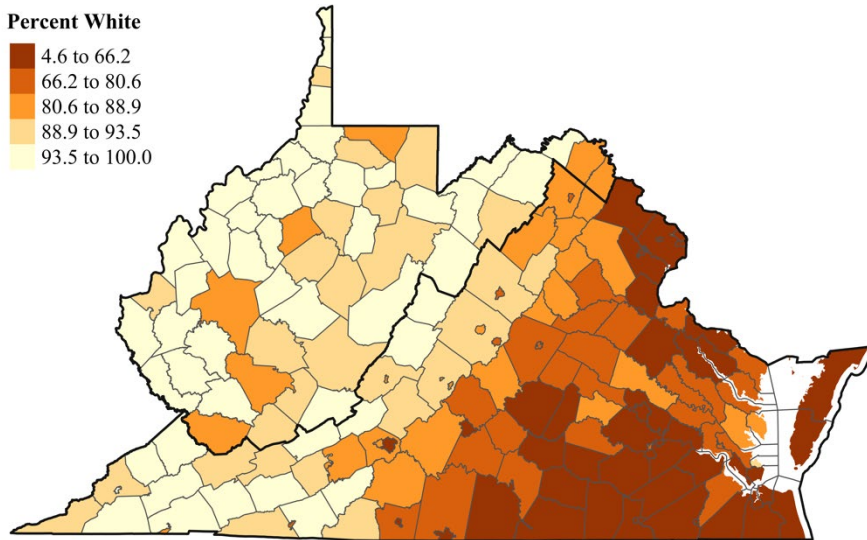| State	Adjacent State	Hispanic (%)	Adjacent State Hispanic (%)	Difference
New Mexico	Utah	46.6	8.0	38.6
New Mexico	Oklahoma	46.6	18.2	28.4
Arizona	Utah	28.6	8.0	20.6
New Jersey	Pennsylvania	20.4	1.7	18.7
New Jersey	New York	20.4	5.9	14.4
Colorado	Utah	21.3	8.0	13.4
Nevada	Utah	20.6	8.0	12.6
Colorado	Kansas	21.3	9.4	11.9
Colorado	Nebraska	21.3	9.6	11.7
Arizona	California	28.6	20.0	8.6
U.S. county-level average:		10.5		
U.S. border counties average:		8.5		
U.S. nonborder counties average:		10.4		

Note: Figures for all U.S. counties follow the same pattern with a starker difference. Source: U.S. Census Bureau, 2022 American Community Survey (five-year estimates); authors' calculations.

If border-county demographics differ substantially from nonborder-county demographics for state pairs, the difference casts doubt on the external validity or the generalization of the results toward the entire state population. For instance, consider an estimation using the border counties of Virginia and West Virginia. Figure 2, Panel A shows that the non-White population share in the border counties of both states is small. Therefore, a DiD analysis conducted with those states would primarily reflect the effects on the White population and we would hesitate to apply these estimates to other contexts where the White population was much smaller and unless we were confident that effects should not differ by race. Panel A also shows that Virgina, away from the West Virgina border, is much less White. Therefore, and unfortunately, border county DiD estimates from this state pair may not even be useful for understanding the effects on Virgina, one of the two states in the exercise.

Conversely, if we see that border-county demographics change abruptly at the state divide, then the DiD estimates would both capture the effects of the policy treatment and reflect the demographic differences in the county pairs, either directly or as an important moderator. Figure 2, Panel B shows an example of a problematic setting where demographics change abruptly at

the state border between Utah and Nevada. The DiD estimates from comparing Utah and Nevada would therefore be contaminated by these direct and indirect demographics effects and may not cleanly identify the effects of the policy under study.

Figure 2: Unrepresentative Border-County State Pairs
Panel A: State Border and Nonborder Counties Differ

Percent White
- 4.6 to 66.2
- 66.2 to 80.6
- 80.6 to 88.9
- 88.9 to 93.5
- 93.5 to 100.0

Panel B: State and Adjacent State Border Counties Differ

Percent Black
- 0.0 to 0.5
- 0.5 to 1.4
- 1.4 to 4.1
- 4.1 to 14.2
- 14.2 to 85.8
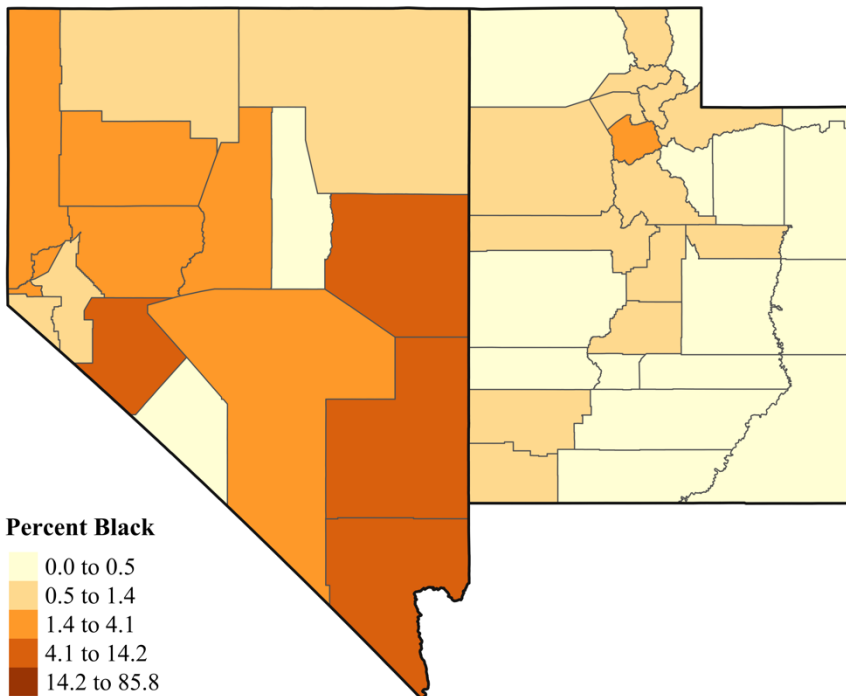
Source: U.S. Census Bureau, 2022 American Community Survey (five-year estimates).

Thought Experiment with a Racial Difference in the Treatment Effect

We next construct a thought experiment to demonstrate the potential importance of racial differences between border counties in state pairs. We consider the estimated effects of a *hypothetical* policy where there are racial differences in *individual* treatment effects of that policy. While the policy is hypothetical, we use *actual* demographic differences.

In the United States, it is well established that for complex reasons, smoking behaviors vary by race. For instance, Black smokers are more likely to be light or non-daily smokers (Trinidad et. al., 2009) consuming fewer cigarettes per day (Trinidad et. al., 2015). Additionally, Black smokers are more likely to start smoking at an older age (Holford et. al., 2016), and to use mentholated cigarettes (Giovino et. al., 2015) while young Black smokers are less likely to use vapes than young White smokers (CDC, 2024). Since smoking behaviors vary by race, anti-smoking policy interventions are likely to have effects that vary with the racial demographics.

In this hypothetical exercise, we aim to estimate the treatment effects of a policy change that increases the tax on a pack of cigarettes by $1. We assume that the true treatment effects, measured in packs of cigarettes purchased per adult per month, are -1.3 for White adults and -0.6 for people of color (POC) adults. The time effect, which represents the average change from the pre-treatment to post-treatment period, absent the policy, is -0.3 packs of cigarettes purchased per White adult per month in both states and -0.1 packs per POC adult per month.[8]

In Table 3, we use actual demographics data to apply this hypothetical exercise and conduct a mock DiD analysis for Mississippi and Tennessee. This state pair has a substantial difference in the share of their White populations. Based on the assumed racially heterogenous treatment effects and actual demographic weights, we find that the DiD estimate for the policy effect derived from the border counties of the treatment and control states is −1.042 packs of cigarettes purchased per adult per month.

However, given the underlying assumptions about individual treatments by race and the demographics, if the treatment were applied to the border counties of Tennessee, the policy effect would be −1.258 packs of cigarettes per adult per month—a larger-in-magnitude effect than with the earlier DiD estimate. Since, by assumption of the exercise, the treatment effects are stronger for White residents and there are material racial differences between the state pairs, the DiD estimates are not externally valid to the residents of Tennessee. Although the sign and magnitude of the effects match, the difference from the true effect is substantial.[9] Importantly,

---

[8] This is a violation of the common trend assumption of DiD, because the White and POC populations have different trends. However, in border county analyses using county level aggregates, as is common, testing the assumption that demographic groups within those counties have common trends may not be possible. Therefore, it may be difficult to know when such violations are a problem.

[9] However, if the sign of the effects differs by an individual's race, then even the *sign* of the true effect might not be externally valid.

the DiD estimate of -1.042 does not clearly reflect the policy effect on any specific population—whether it it be White adults (-1.300), POC adults (-0.600), the border counties of Mississippi (averaging -1.090), or the border counties of Tennessee (averaging -1.258). We therefore argue that identification is compromised, because the DiD estimate does not reflect the policy effect for any relevant population.

Table 3: Mock DiD Analysis for State Pair with Border County Demographic Differences

Mississippi						
Population	**Pre**	**Time Trend**	**Post (w/o Treatment)**	**Effect of Rule**	**Post (w/ Treatment)**	**Border Population Share**
White	4.500	-0.300	4.200	(E) -1.300	2.900	70%
POC	4.000	-0.100	3.900	(F) -0.600	3.300	30%
Border County Avg.	(A) 4.350	-0.240	4.110	(G) -1.090	(C) 3.020	
Tennessee						
White	5.000	-0.300	4.700	-1.300	3.400	94%
POC	4.000	-0.100	3.900	-0.600	3.300	6%
Border County Avg.	(B) 4.94	-0.288	(D) 4.652	(H) -1.258	3.394	
Effect	**White (Actual, All)**	**POC (Actual, All)**	**Pop. Avg. (Border, MS)**	**Pop. Avg. (Border, TN)**	**DiD (Estimated, Border)**	
Notes	(E)	(F)	(G)	(H)	((C)-(D))-((A)-(B))	
Value	-1.300	-0.600	-1.090	-1.258	-1.042	

Note: Outcome variable is packs of cigarettes purchased per adult per month. `POC' means "person of color". Source: Demographics data from U.S. Census Bureau, 2022 American Community Survey (five-year estimates); authors' calculations.

In state pairs with similar demographics, the DiD estimates would be more likely to be well identified.[10] Table 4 shows a second mock DiD analysis, this time for Ohio and Pennsylvania. Only the demographics are changed for this example, the time trend and treatment effects for each demographic group are the same. In this pair, the border counties have almost identical White and POC population shares. The DiD between border counties of the treatment and control states would estimate an effect of the policy of -1.237 packs of cigarettes purchased per adult per month. In this case, the DiD analysis correctly estimates the average causal effect of the policy in both Ohio and Pennsylvania's border counties.

Note however that that this estimate may not have external validity to contexts with different demographics. For example, the estimated effects from this second thought experiment do not

---

[10] In addition, if the treatment effects do not vary by race, then differences in racial demographics would not affect identification or external validity.

recover the causal effects for the border counties of Mississippi or Tennessee (-1.090 and -1.258, respectively) from the prior example.

Even when the border counties of treatment and control states are similar and the regression successfully identifies the causal effect on the border counties, the DiD may still fail to capture the causal effect on the treated or control states as a whole. This is another instance where failure in external validity may occur. Appendix C.1 and Appendix C.2 show an example of how this can happen using Virginia and Kentucky.

Table 4: Mock DiD Analysis for State Pair <u>without</u> Border County Demographic Differences

Ohio						
Population	Pre	Time Trend	Post (w/o Treatment)	Effect of Rule	Post (w/ Treatment)	Border Population Share
White	4.500	-0.300	4.200	(E) -1.300	2.900	91%
POC	4.000	-0.100	3.900	(F) -0.600	3.300	9%
Border County Average	(A) 4.455	-0.240	4.173	(G) -1.237	(C) 2.936	
Pennsylvania						
White	5.000	-0.300	4.700	-1.300	3.400	91%
POC	4.000	-0.100	3.900	-0.600	3.300	9%
Border County Average	(B) 4.910	-0.288	(D) 4.628	(H) -1.237	3.391	
Effect	White (Actual, All)	POC (Actual, All)	Pop. Avg. (Border, OH)	Pop. Avg. (Border, PA)	DiD (estimated, Border)	
Notes	(E)	(F)	(G)	(H)	((C)-(D))-((A)-(B))	
Value	-1.300	-0.600	-1.237	-1.237	-1.237	

Note: Outcome variable is packs of cigarettes purchased per adult per month. `POC' means "person of color". Source: Demographics data from U.S. Census Bureau, 2022 American Community Survey (five-year estimates); authors' calculations.

Differences in Health and Social Characteristics across State Borders

Next, we analyze the health and social characteristics of border counties and find statistically and economically significant differences in their population shares of older and disabled individuals. Specifically, we find that the residents of border counties are, on average, 0.6 years older (1.4 percent older) than the residents of nonborder counties (Table 4). Also, the share of the population that is a senior citizen (age 65+) is on average, 0.6 percentage points higher (3.1 percent more senior citizens) higher. There are similar but smaller differences in the share of the population that is 85+. The share of the population with a disability is also 0.3 percentage points

higher (1.9 percent more individuals with disabilities).[11] There is no statistically significant difference between border and nonborder counties in the shares of the population that are veterans or have health insurance.

These results are consistent when the sample is expanded to all available counties (Appendix D.1). Regression results when weighted by county population (weighted regressions) are very similar to the unweighted regressions except that the weighted regressions show that residents of border counties, when weighted by population, are more likely to have health insurance (Appendix D.2).

Table 5: Health and Social Characteristics, by County Type (counties in the contiguous United States)

	Median age	Age 65+	Age 85+	Disabled	Insured	Veteran
Border County	0.6[***]	0.6[***]	0.1[**]	0.3[*]	-0.1	-0.02
	(0.2)	(0.2)	(0.04)	(0.2)	(0.3)	(0.2)
Intercept	41.9[***]	19.6[***]	2.3[***]	16.1[***]	79.4[***]	7.9[***]
	(0.1)	(0.1)	(0.02)	(0.1)	(0.2)	(0.1)
Observations	3,109	3,109	3,109	3,109	3,109	3,109
R^2	0.00	0.00	0.00	0.00	0.00	0.00
Adjusted R^2	0.00	0.00	0.00	0.00	-0.00	-0.00

Note: Robust standard errors. `Age 65+` represents 65 years and older and `Age 85+` represents 85 years and older. Source: U.S. Census Bureau, 2022 American Community Survey (five year estimates).

A Case Study of the Racial Demographic Differences in a Border-County DiD Paper

As this paper has emphasized, failure to account for differences between border counties and nonborder counties can reduce the external validity of program evaluations. Indeed, demographic differences across state borders can compromise identification entirely.[12] Still, in practice, these differences depend on the specific state pairs and may be too small to matter. We next explore the demographic differences in a prominent paper that uses a DiD setup with border counties.

As mentioned, Dube et al al. (2010) uses a DiD design to identify the effects of minimum wages on earnings and employment in restaurants and other low-wage sectors in the United States. The

---

[11] In the ACS, disability is defined as difficulty with hearing, vision, cognition, ambulation (movement), self-care, or living independently (Census, 2021).

[12] If external validity is compromised, the estimated effects may be correct but *only* for the kinds of populations that are more likely to be in border counties. Such effects may be invalid for populations with different demographics. This finding is demonstrated in the exercise shown in Appendix C.1.

When identification is compromised, the effects of the policy are combined with the effects of the differential populations on either side of the border. For example, White workers overwhelmingly and disproportionately hold occupations involving sales (Yau, 2024). Therefore, DiD estimates of the effect of a policy shock affecting sales, like a sales tax change, changes in noncompete rules, or changes in blue laws, could have no meaningful interpretation in the presence of substantial racial differences across the state borders.

paper finds no adverse employment effects of these changes, exploiting local differences in minimum wage policies between 1990 and 2006. The paper identifies 17 states and the District of Columbia in this time frame with minimum wage requirements above the federally mandated minimum wage. The authors then use all states with minimum wage differentials with respect to their neighboring states to estimate the effects of minimum wage policy on earnings and employment by comparing their border counties.

For instance, California had a state law requiring a minimum wage above the federal one, and its bordering states (Arizona, Nevada, and Oregon) did not have any state minimum wage laws (only the federal minimum wage law). The authors compared the border counties in California (treatment group) with the border counties in each of the surrounding state (control group) to conduct a DiD analysis.

As with other papers in the literature, this approach is used to control for unobserved heterogeneity. The paper assumes that "contiguous border counties represent good control groups for estimating minimum wage effects if there are substantial differences in treatment intensity within cross-state county-pairs, and a county is more similar to its cross-state counterpart than to a randomly chosen county" (Dube et al., 2010, p. 949). However, our results show why readers should not casually assume that such results are externally valid to nonborder counties. The border-county demographic differences can be substantial. That is not to say that the authors' results are unidentified, or that they are not externally valid, but the demographic differences give cause for concern.

Figure 3, shows how those differences may confound the analysis. It shows the median age in California counties (which has a minimum wage law) alongside the median ages of the counties of California's surrounding states (without such a law). California border counties are similar in median age to the border counties in Nevada and Oregon (which is good). Unfortunately, the nonborder counties in Arizona, California, and Nevada are noticeably younger, raising external validity concerns. The border counties in Arizona are much older than the border counties in California, of further concern because that is the set of counties to which the authors are comparing them. Because of the hump-shaped profile of income over the lifecycle, this is potentially a problem (Mincer, 1958). If the effects of age on income were linear, that is unlikely to be a meaningful problem.

A strength of their approach is that it accounts for local economic conditions and is robust to allowing for long-term effects of minimum wage changes (Dube et al., 2010). Unfortunately, because they use county-level data, there are confounders for which they cannot control. If individual-level data were available, then individual controls might help address treatment effects that vary by worker demographics. But with aggregate county data, the limited sample constrains the researcher's ability to include sufficient controls to allow for treatments to vary by demographics. Having just tens of observations leaves the econometrician with insufficient

14

power and variation to control for many distinct aspects of demographics (and additional non-demographic confounders only worsen this problem).[13]

Figure 3: Median Age in California and Surrounding Border Counties
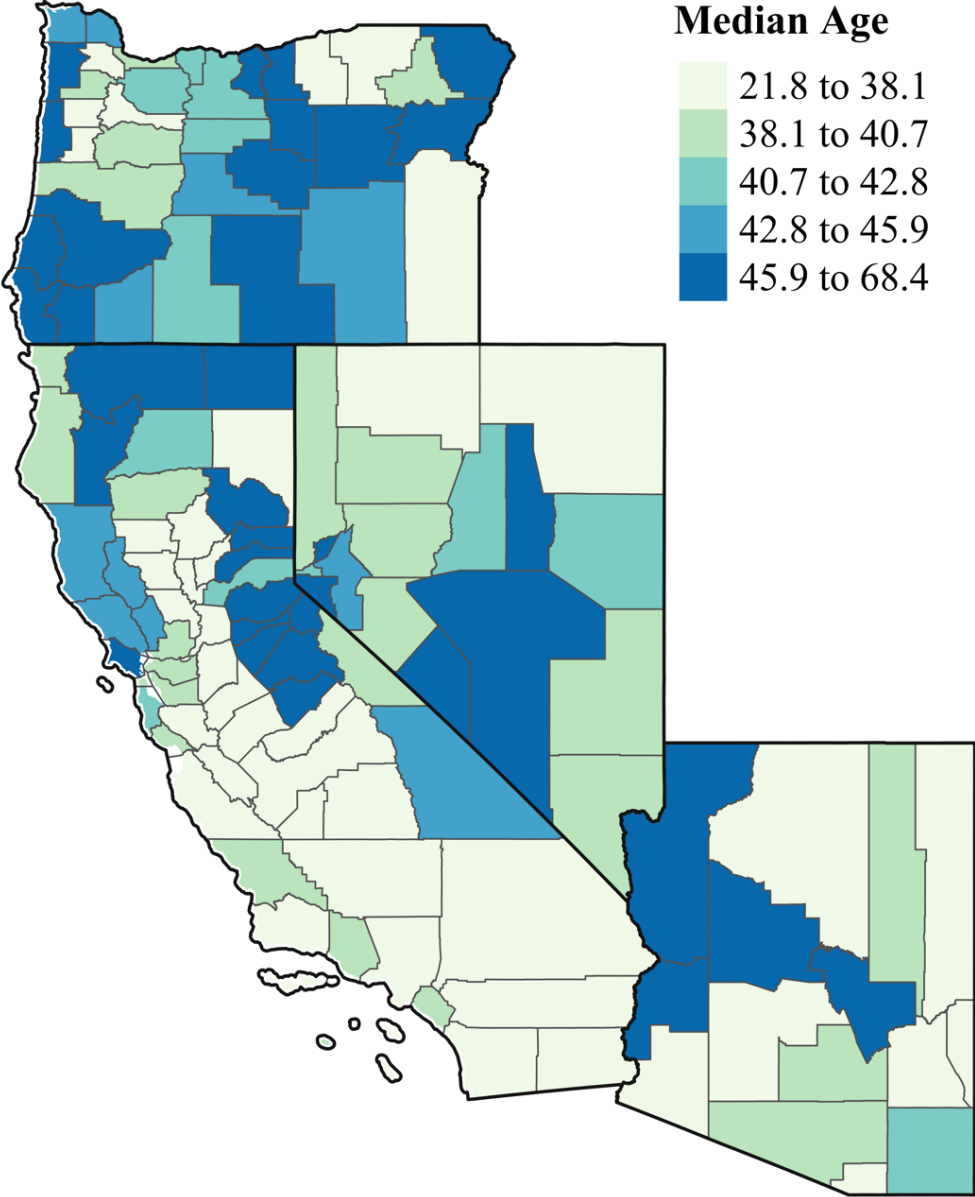
Source: U.S. Census Bureau, 2022 American Community Survey (five-year estimates).

---

[13] The econometrician could, in some cases use county demographics interacted with time dummies right. However, the power issues, potentially severe depending on the setting, would remain.

Conclusion

Policy analysis frequently uses difference-in-difference methods on border-county populations to reduce the confounding geographic variables from state-level analysis. While this method may provide a plausible control population for state-level policies, it can introduce (or at least fail to fix) other problems caused by the differences between border and nonborder counties.

We encourage practitioners conducting border-county-based policy analysis to proceed with caution. Border counties are, on average, systematically whiter, older, and more disabled than nonborder counties. In addition, there are examples where demographics change substantially across state borders that may complicate or compromise causal identification and external validity.

Richer econometric specifications can help address these challenges. Researchers have developed various methodologies to strengthen causal inference and deal with confounders. For instance, synthetic control methods, as introduced by Abadie et al. (2010), estimate the effect of a policy intervention by constructing a synthetic counterfactual group which closely resemble the treated unit before the intervention greatly reducing demographic confounders while helping to control for unobserved confounders. Similarly, Angrist and Pischke (2009) presented propensity score matching, a method where treated and control groups are matched on a score based on observed covariates like demographics to reduce confounding influence. They also discussed instrumental variable (IV) techniques, which use external sources of variation—such as policy changes or natural experiments—as instruments to account for endogeneity and identify causal effects. By leveraging these instruments, IV methods help control for unobserved confounders that might otherwise bias estimates. While these methods present their own challenges (e.g. determining appropriate synthetic control groups, issue of unmeasured confounders, and the difficulty of selecting valid instruments for instrumental variables), they offer creative solutions to reducing the impact of demographic differences in state policy studies.

At a minimum, researchers conducting matched border-county analysis should perform ample robustness checks to ensure that these demographic differences do not introduce important confounders in their questions of interest.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American Statistical Association, 105(490), 493-505. https://doi.org/10.1198/jasa.2009.ap08746

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, 24(2), 3-30.

Bognetti, Giovanni , Tate, C. Neal , Fellman, David and Shugart, Matthew F.. "Constitutional Law". *Encyclopedia Britannica*, 3 Sep. 2024.

Boone, C., Dube, A., Goodman, L., & Kaplan, E. (2021). Unemployment insurance generosity and aggregate employment. *American Economic Journal: Economic Policy*, 13(2), 58-99.

Card, D., & Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review*, *84*(4), 772–793.

Holmes, T. J. (1998). The effect of state policies on the location of manufacturing: Evidence from state borders. *Journal of political economy*, 106(4), 667-705.

Cartwright, N., & Munro, E. (2010). The limitations of randomized controlled trials in predicting effectiveness. *Journal of evaluation in clinical practice*, 16(2), 260-266.

Cosgrove, B. M., LaFave, D. R., Dissanayake, S. T., & Donihue, M. R. (2015). The economic impact of shale gas development: A natural experiment along the New York/Pennsylvania border. *Agricultural and Resource Economics Review*, 44(2), 20-39.

Dieterle, S., Bartalotti, O., & Brummet, Q. (2020). Revisiting the Effects of Unemployment Insurance Extensions on Unemployment: A Measurement-Error-Corrected Regression Discontinuity Approach. *American Economic Journal: Economic Policy*, 12(2), 84-114.

Dube, A., Lester, T. W., & Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *The review of economics and statistics*, 92(4), 945-964.

European Committee of the Regions (n.d.). "Cor - France Introduction." Accessed 19 Aug. 2024.

Forum of Federations (n.d.). Federal Countries - Forum of Federations. Accessed 19 Aug. 2024.

Giuntella, O., & Mazzonna, F. (2019). Sunset time and the economic effects of social jetlag: evidence from US time zone borders. *Journal of health economics*, 65, 210-226.

Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: gold standard, or merely standard?. *Perspectives in biology and medicine*, 48(4), 516-534.

Hanson, A., & Sullivan, R. (2009). The incidence of tobacco taxation: evidence from geographic micro-level data. *National Tax Journal*, 62(4), 677-698.

Hao, Z., & Cowan, B. W. (2020). The cross-border spillover effects of recreational marijuana legalization. *Economic inquiry*, 58(2), 642-666.

Holford, T.R., Levy, D.T. and Meza, R., 2016. Comparison of smoking history patterns among African American and white cohorts in the United States born 1890 to 1990. *Nicotine & Tobacco Research*, *18*(suppl_1), pp.S16-S29.

Huang, R. R. (2008). Evaluating the real effect of bank branching deregulation: Comparing contiguous counties across US state borders. *Journal of Financial Economics*, 87(3), 678-705.

Huber, G. A., & Arceneaux, K. (2007). Identifying the persuasive effects of presidential advertising. *American Journal of Political Science*, 51(4), 957-977.

Jha, P., Neumark, D., & Rodriguez-Lopez, A. (2022). What's across the border? Re-evaluating the cross-border evidence on minimum wage effects. *SSRN*.

Kumar, A. (2018). Do restrictions on home equity extraction contribute to lower mortgage defaults? Evidence from a policy discontinuity at the Texas border. *American Economic Journal: Economic Policy*, 10(1), 268-297.

Lacombe, D. J. (2004). Does econometric methodology matter? An analysis of public policy using spatial econometric techniques. *Geographical analysis*, 36(2), 105-118.

Lyu, W., & Wehby, G. L. (2020). Comparison of estimated rates of coronavirus disease 2019 (COVID-19) in border counties in Iowa without a stay-at-home order and border counties in Illinois with a stay-at-home order. *JAMA network open*, 3(5), e2011102-e2011102.

McKinnish, T. (2005). Importing the poor: welfare magnetism and cross-border welfare migration. Journal of human Resources, 40(1), 57-76.

Mincer, J., 1958. Investment in human capital and personal income distribution. *Journal of political economy*, *66*(4), pp.281-302.

New State Ice Co. v. Liebmann, 285 U.S. 262 (1931). *Legal Information Institute*, Cornell Law School.

Peng, L., Xiaohui G., & Meyerhoefer, C. D. "The effects of Medicaid expansion on labor market outcomes: evidence from border counties." *Health economics* 29.3 (2020): 245-260.

Thompson, J. P., & Rohlin, S. M. (2013). The Effect of State and Local Sales Taxes on Employment at State Borders, *FEDS working paper*, 2013-49.

Trinidad, D.R., Pérez-Stable, E.J., Emery, S.L., White, M.M., Grana, R.A. and Messer, K.S., 2009. Intermittent and light daily smoking across racial/ethnic groups in the United States. *Nicotine & Tobacco Research*, *11*(2), pp.203-210.

18

Trinidad, D.R., Xie, B., Fagan, P., Pulvers, K., Romero, D.R., Blanco, L. and Sakuma, K.L.K., 2015. Disparities in the population distribution of African American and non-Hispanic white smokers along the quitting continuum. *Health Education & Behavior*, *42*(6), pp.742-751.

U.S. Census Bureau. (2023) (2009-2022) American Community Survey, ACS 5-Year Estimates (2018-2022) Public Use Microdata Samples. Retrieved using tidycensus.

U.S. Census Bureau. Areas Published. American Community Survey (ACS), 3 Sept. 2024.

U.S. Census Bureau. How Disability Data are Collected from The American Community Survey, 11 Nov. 2021.

U.S. Census Bureau. Why We Ask Questions About...Hispanic or Latino Origin, 1 Mar. 2022.

U.S. Centers for Disease Control and Prevention (CDC). E-Cigarette Use Among Youth, 15 May 2024.

Appendices

APPENDIX A

The main text shows, using an unweighted sample of contiguous U.S. states, that the racial and ethnic makeup of border counties were different from those of nonborder counties. The same analysis on the full sample of U.S. state border counties shows stronger results (Appendix Table A.1). When weighted by county population, the results are broadly similar, although less precise (Appendix Table A.2).

APPENDIX A.1: RACE AND HISPANIC ETHNICITY, BY COUNTY TYPE (ALL U.S. COUNTIES INCLUDING PUERTO RICO)

	White	Hispanic	Black	Asian	Biracial	Islander	Native
Border County	3.4***	-4.4***	-1.2**	-0.2**	-1.1***	-0.02	-0.02
	(0.8)	(0.6)	(0.6)	(0.1)	(0.2)	(0.01)	(0.3)
Intercept	77.6***	12.9***	9.1***	1.5***	6.5***	0.1***	1.9***
	(0.4)	(0.4)	(0.3)	(0.1)	(0.1)	(0.01)	(0.2)
Observations	3,222	3,222	3,222	3,222	3,222	3,222	3,222
R²	0.01	0.01	0.00	0.00	0.01	0.00	0.00
Adjusted R²	0.00	0.01	0.00	0.00	0.01	-0.00	-0.00

Note: Robust standard errors. `Hispanic` represents those who idenitfy as Hispanic from any race. `Islander` represents Hawaiian or Pacific Islander. `Native` represents Native American.
*P < 0.1; **P < 0.05; ***P < 0.01.
Source: U.S. Census Bureau, 2022 American Community Survey (five year estimates).

19

APPENDIX A.2 POPULATION-WEIGHTED RACE AND HISPANIC ETHNICITY, BY COUNTY TYPE (CONTIGIOUS US COUNTIES)

	White	Hispanic	Black	Asian	Biracial	Islander	Native
Border County	2.6	-4.7*	0.3	-0.8	-1.8***	-0.1**	0.3
	(2.8)	(2.5)	(1.6)	(0.8)	(0.6)	(0.05)	(0.2)
Intercept	65.2***	20.2***	12.4***	5.8***	9.2***	0.2***	0.8***
	(1.3)	(1.4)	(0.6)	(0.6)	(0.4)	(0.04)	(0.05)
Observations	3,222	3,222	3,222	3,222	3,222	3,222	3,222
R^2	0.00	0.01	0.00	0.00	0.01	0.00	0.00
Adjusted R^2	0.00	0.01	-0.00	0.00	0.01	0.00	0.00

Note: Results are weighted by population and contain robust standard errors. `Hispanic` represents those who idenitfy as Hispanic from any race. `Islander` represents Hawaiian or Pacific Islander. `Native` represents Native American.
*$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.
Source: U.S. Census Bureau, 2022 American Community Survey (five year estimates).

APPENDIX B

As shown in Table 2, there are some cases of vast differences in the Hispanic population across state boundaries when comparing the average percentages of the Hispanic population among border counties with those of the Hispanic population among nonborder counties. In Appendix Table B1, we show that these vast differences are consistent with some state pairs for the White demographic as well.

State	Adjacent State	White (%)	Adjacent State White (%)	Difference
Utah	Arizona	89.5	59.6	29.9
Pennsylvania	New jersey	90.8	63.4	27.4
Utah	New Mexico	89.5	63.9	25.7
Pennsylvania	Delaware	90.8	66.0	24.8
Tennessee	Mississippi	94.2	70.0	24.2
New York	New Jersey	87.3	63.4	23.9
Colorado	Arizona	83.1	59.6	23.4
Pennsylvania	Maryland	90.8	67.5	23.2
Tennessee	Alabama	94.2	71.3	23.0
Virginia	Maryland	89.6	67.5	22.1
U.S. county-level average:		78.2		
U.S. border counties average:		81.0		
U.S. nonborder counties average:		79.1		

Note: Figures for all U.S. counties follow the same pattern with a starker difference. Source: U.S. Census Bureau, 2022 American Community Survey (five-year estimates); authors' calculations.

APPENDIX C

The main text presents a thought experiment with real demographic statistics and made-up treatment effects where the DiD estimator differs considerably from that in the counterfactual (Table 3). But this is sensitive to the demographic differences in border counties. When there are minimal demographic differences, the DiD is properly identified even when the two groups have differential treatment effects and time trends (Table 4). This appendix discusses a related issue, the concequences for identification and external validity when border counties and non-border counties, within the states studied, have substantial differences in demographics.

Appendix Table C.1 shows that for state pairs with narrower differences in demographics, the matched border-county DiD results do not differ considerably from those in the counterfactual. Notice how the White effect (-1.3 packs per adult per month) is close to the DiD effect (-1.228) since the border counties of both states have overwhelmingly White populations. But, in this case, the border counties of Virginia have quite different demographics from the rest of the state. Appendix Table C.2 shows that the effect of the policy on Virginia as a whole (not just the border counties) would be -1.104 packs per adult per month. Therefore, the effects estimate from the border county DiD do not have external validity to the state as a whole.

APPENDIX C.1 MOCK DID ANALYSIS FOR STATE PAIR WITHOUT BORDER COUNTY DEMOGRAPHIC DIFFERENCES

Virginia						
Population	Pre	Time Trend	Post (No Treatment)	Effect of Rule	Post (With Treatment)	Border Population Share
White	4.500	-0.300	4.200	(E) -1.300	2.900	90%
POC	4.000	-0.100	3.900	(F) -0.600	3.300	10%
Border County Avg.	(A) 4.45	-0.240	4.210	(G) -1.230	(C) 2.940	
Kentucky						
White	5.000	-0.300	4.700	-1.300	3.400	91%
POC	4.000	-0.100	3.900	-0.600	3.300	9%
Border County Avg.	(B) 4.910	-0.282	(D) 4.628	(H) -1.237	3.391	
Effect	White (Actual, All)	POC (Actual, All)	Pop. Avg. (Border, VA)	Pop. Avg. (Border, KY)	DiD (estimated, Border)	
Notes	(E)	(F)	(G)	(H)	((C)-(D))-((A)-(B))	
Value	-1.300	-0.600	-1.230	-1.237	-1.228	

Note: Outcome variable is packs of cigarettes purchased per adult per month. `POC' means "person of color". Source: Demographics data from U.S. Census Bureau, 2022 American Community Survey (five-year estimates); authors' calculations.

APPENDIX C.2 MOCK DID ANALYSIS FOR STATE PAIR, SAME STATES, WITH ALL COUNTY DEMOGRAPHIC DIFFERENCES

Virginia						
Population	Pre	Time Trend	Post (No Treatment)	Effect of Rule	Post (With Treatment)	All Population Share
White	4.500	-0.300	4.200	(E) -1.300	2.900	72%
POC	4.000	-0.100	3.900	(F) -0.600	3.300	28%
Border County Avg.	(A) 4.360	-0.244	4.116	(G) -1.104	(C) 3.012	
Kentucky						
White	5.000	-0.300	4.700	-1.300	3.400	92%
POC	4.000	-0.100	3.900	-0.600	3.300	8%
Border County Avg.	(B) 4.920	-0.284	(D) 4.636	(H) -1.244	3.392	
Effect	White (Actual, All)	POC (Actual, All)	Pop. Avg. (All, VA)	Pop. Avg. (All, KY)	DiD (estimated, Border)	
Notes	(E)	(F)	(G)	(H)	((C)-(D))-((A)-(B))	
Value	-1.300	-0.600	-1.104	-1.244	-1.064	

Note: Outcome variable is packs of cigarettes purchased per adult per month. `POC' means "person of color". Source: Demographics data from U.S. Census Bureau, 2022 American Community Survey (five-year estimates); authors' calculations.

APPENDIX D

The main text shows, using an unweighted sample of contiguous U.S. counties, that border counties have disproportionately more older and disabled individuals (Table 5). This appendix conducts two variations on the analysis for robustness purposes. First, when regression include the full sample of U.S. counties (not just contiguous United States counties) we generally reaffirm the results and additionally shows higher health insurance coverage in border counties (Appendix Table D.1). Second, when the regression using the contiguous United States counties are weighted by county population, the results are also broadly similar, although less precise (Appendix Table D.2).

APPENDIX D.1: HEALTH AND SOCIAL CHARACTERISTICS, BY COUNTY TYPE (ALL AVAILABLE U.S. COUNTIES)

	Median age	Age 65+	Age 85+	Disabled	Insured	Veteran
Border County	0.7***	0.6***	0.1**	0.4*	0.8**	0.1
	(0.2)	(0.2)	(0.04)	(0.2)	(0.3)	(0.2)
Intercept	41.9***	19.6***	2.3***	16.1***	78.5***	7.8***
	(0.1)	(0.1)	(0.02)	(0.1)	(0.2)	(0.1)
Observations	3,144	3,222	3,222	3,144	3,222	3,222
R²	0.00	0.00	0.00	0.00	0.00	0.00
Adjusted R²	0.00	0.00	0.00	0.00	0.00	0.00

*p<0.1; **p<0.05; ***p<0.01
Note: Robust standard errors. `Age 65+` represents 65 years and older and `Age 85+` represents 85 years and older.
*P < 0.1; **P < 0.05; ***P < 0.01.
Source: U.S. Census Bureau, 2022 American Community Survey (five year estimates).

APPENDIX D.2 POPULATION-WEIGHTED HEALTH AND SOCIAL CHARACTERISTICS, BY COUNTY TYPE (CONTIGIOUS US COUNTIES)

	Median age	Age 65+	Age 85+	Disabled	Insured	Veteran
Border County	0.7*	0.3	0.03	0.1	1.1	-0.2
	(0.4)	(0.4)	(0.1)	(0.4)	(0.7)	(0.3)
Intercept	39.1***	16.5***	2.0***	12.9***	80.6***	6.7***
	(0.2)	(0.2)	(0.03)	(0.2)	(0.4)	(0.2)
Observations	3,109	3,109	3,109	3,109	3,109	3,109
R²	0.00	0.00	0.00	0.00	0.00	0.00
Adjusted R²	0.00	0.00	0.00	-0.00	0.00	0.00

Note: Results are weighted by population and contain robust standard errors. `Age 65+` represents 65 years and older and `Age 85+` represents 85 years and older.
*P < 0.1; **P < 0.05; ***P < 0.01.
Source: U.S. Census Bureau, 2022 American Community Survey (five year estimates).