

Latinos in the House? Demographic Change, Ethnic Attrition, and Latino Political Behavior*

Alexander Agadjanian[†], Alejandro Flores[‡], and Marcel F. Roman^{*}

[†]Department of Political Science, UC Berkeley

[‡]Department of Political Science, MIT

[†]Department of Government, Harvard

August 8, 2024

Abstract

Latino population growth rates are declining nationally and locally due to reduced immigration, declining birth rates, and urban gentrification. How do these macro-level demographic trends shape Latino mass behavior? Using a large Latino survey (Study 1, $N = 61,000$), we find local Latino population growth forestalls *ethnic attrition* (opting out of Latino identity) among individuals with at least Latino ancestry. Consistent with demographic trends, Latino growth primarily forestalls shifts toward white identification. Moreover, Latino growth forestalls attrition among “atypical” Latinos disposed to attrition. We replicate our findings in two panel surveys to mitigate residential selection bias (Study 2, $N = 584$; $N = 400$) and with a voter file panel to capture a behavioral means of identity (Study 3, $N \approx 97,500$). We also propose a pre-registered survey experiment priming Latino growth among Latinos (Study 4) for causal leverage and mechanism insights. Given ethnic identity maintenance mitigates political conservatism among putative Latinos, declines in Latino population growth may have reverberating consequences for Latino mass political behavior.

*We thank Feyaad Allie, Steve Ansolobehere, Marco Avina, Jeremiah Cha, Angelo Dagonel, Bernard Fraga, Jennifer Hochschild, Mashail Malik, Benjamin Newman, and Alan Yan for their helpful comments and insights.

Introduction

Latino population *levels* have grown in the past half-century, increasing from 10-64 million between 1970-2020 (Figure 1, Panel A). Concomitantly, decadal Latino *growth* was on a steady, slight upward trajectory between 1970-2000 (51-58%). However, Latino population growth has been declining since 2000 due to lower birth rates (Alvira-Hammond, 2019), reduced immigration from Latin America (Haner and Lopez, 2023), the shift in the source of Latino population growth from immigration to reproduction (Newman et al., 2023), and high rates of inter-ethnoracial marriage among Latinos encouraging ethnic attrition among offspring (Duncan and Trejo, 2010; Alba, 2020). Growth between 2010-2022 now stands at 26%, half of the decadal growth rate between 1970-2000, implying, at some point, that Latino population growth may eventually reach net zero (Figure 1, Panel B). National-level trends extrapolate to the local-level. Average decadal tract-level Latino growth has also declined, which may be exacerbated by white in-migration to Latino neighborhoods as a byproduct of urban gentrification (Mumm and Sternberg, 2023). Given the demographic import of Latinos, the largest non-white ethno-racial group responsible for much of the recent growth in the U.S. population, what are the downstream idenitarian and political consequences of declining Latino population growth on Latino coethnic mass behavior?

We extend *ethnic replenishment theory* (ERT) (Jiménez, 2008; Jiménez, 2010), which posits local coethnic immigrant population growth provides myriad ethnic resources that sustains the salience of ethnic identity for acculturated coethnics (i.e. English-speakers; second, third, and fourth-generation coethnics). We theorize local coethnic growth irrespective of immigration status mitigates the prospect of *ethnic attrition* among immigrant group coethnics (opting out of subjectively identifying as the ethnic group category in favor of a different or dominant category, e.g. Anglo whites). Additionally, consistent with prior insights (Jiménez, 2008; Jiménez, 2010), we theorize local coethnic growth forestalls the prospect of attrition for marginal ethnic group members predisposed toward attrition. We also posit the type of coethnic growth matters. Less acculturated coethnic growth, by virtue of possessing stronger

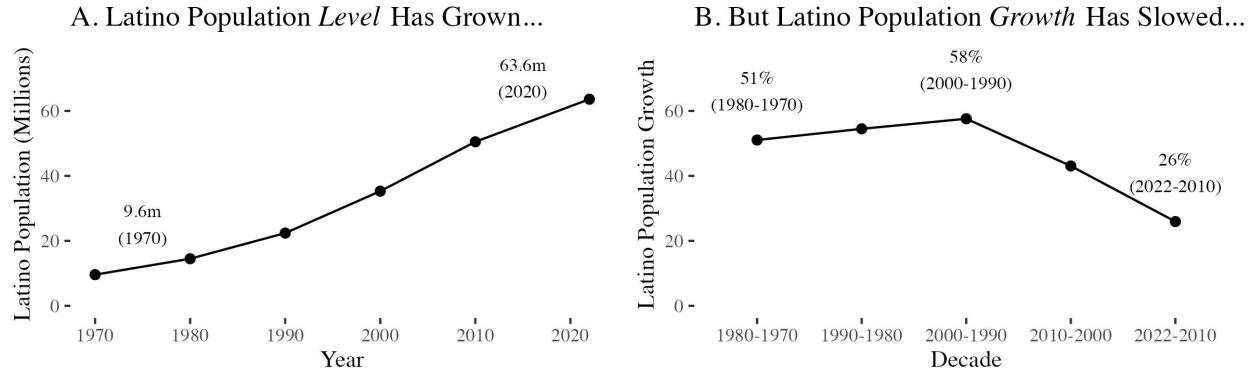


Figure 1: Latino population *level* has grown but Latino population *growth* has slowed (Source: Pew Research Center using augmented Census data)

ethnic resources, may be more likely to forestall ethnic attrition. However, distinct from prior research (Jiménez, 2008; Jiménez, 2010), we find relatively acculturated local coethnic growth still forestalls attrition.

Four studies corroborate our theoretical expectations. We stack cross-sectional Cooperative Election Survey (CES) data between 2010-2022 and subset to Latino respondents in order to garner a uniquely large sample of Latinos with the capability to measure ethnic attrition ($N = 61,000$). We merge these data with zipcode and county Census demographic data and demonstrate local Latino population growth: 1) forestalls ethnic attrition; 2) forestalls ethnic attrition primarily among marginal Latinos predisposed to attrit; and 3) has a heterogeneous influence on attrition conditional on the type of Latino growth, where immigrant and Spanish-speaking Latino growth tends to forestall attrition more strongly than US-born and English-speaking Latino growth (Study 1). To mitigate concerns Study 1 is driven by residential selection bias, we leverage two panel surveys of Latinos (CES '10-'14, $N = 584$; Voter Study Group '11-'16, $N = 400$) to show living in a zipcode with a higher concentration of Latinos undercuts ethnic attrition over the course of 4-5 years but being an ethnic attritor does not determine moving to a more or less Latino zipcode (Study 2). To mitigate concerns that Studies 1-2 may be the result of fickle ethno-racial switching inherent to survey measurement of self-reported attitudes, we construct a two-wave panel from administrative North Carolina voter registration data that measures self-registered Latino/Hispanic ethnicity over four years

($N \approx 97,500$ Latinos). Consistent with Studies 1-2, we show Latinos who initially register they are Latino/Hispanic are more likely to do so four years later conditional on living in a zipcode experiencing relatively higher Latino population growth (Study 3). Finally, given Latino demographic shifts are not random and attrition may be difficult to manipulate via short-term experimental interventions, we propose a survey experimental conceptual replication that heightens the salience of Latino population growth among Latinos and assesses their downstream attachments to the ethnic group short of shifting their ethno-racial identity entirely (Study 4).

This paper clarifies the link between demography, identity, and downstream Latino political behavior, with implications for multiple questions and puzzles related to Latino politics. Most research on ethnic attrition focuses on how it presents statistical artifacts for estimating the relationship between immigrant generational status, health (Antman et al., 2016), socio-economic status (Duncan and Trejo, 2011b; Emeka, 2019), and other dimensions of assimilation (Lopez et al., 2017), especially given ethnic attrition tends to be a product of distance from the immigrant experience (Duncan and Trejo, 2018). Concomitantly, a burgeoning literature has focused on what determines ethnic attrition, with particular attention to ethno-racial intermarriage and discrimination (Alba and Islam, 2009; Hadah, 2023). However, although some inquiries have qualitatively explored the relationship between demography and Latino identity in a few local contexts or imply a link between demography and ethnic attrition without an explicit emphasis on the relationship (Jiménez, 2008; Jiménez, 2010; Fernández et al., 2018; Agadjanian and Lacy, 2021; Manrique and Mouw, 2023), there is relatively less systematic evidence evaluating the relationship between demography and ethnic attrition for the national Latino population with attention to heterogeneity in how distinct types of Latino demographic patterns and individual-level characteristics may condition how demography motivates ethnic identity and attrition. Our evidence explicates the unique role that Latino population *growth*, a demographic dimension that is currently declining due to a number of macro-level factors, plays in undercutting ethnic attrition in addition to relevant

heterogeneity in *how* Latino growth may sustain ethnic identity.

Moreover, recent work has reinforced the notion that the maintenance and adoption of an ethnic identity is *fundamentally political*. It is long-established that in the United States, ethno-racial identity categories are politicized and powerfully shape downstream individual political attitudes and behavior even in spite of intra-group heterogeneity (Kinder and Winter, 2001; Junn et al., 2012). Moreover, ethno-racial self-categorization is not necessarily pre-political, but, in part, a function of political dispositions (Egan, 2020; Agadjanian and Lacy, 2021). Indeed, our own analysis of Cooperative Election Survey data suggests Latino attritors engage in political behaviors and hold political attitudes that are fundamentally distinct, more conservative, and more similar to the dominant ethno-racial group in the U.S. (i.e. Anglo whites) from non-attritors (Section A.1.2). Given the linkages between identity and politics, our theoretical insights and evidence suggest that sustained Latino population growth not only replenishes a strong sense of ethnic identity among coethnics, but replenishes ethnic political behavior. In short, our theory and evidence is critical to understanding the long-term sustainability of *a* Latino politics in light of declining Latino population growth because there cannot be a Latino ethnic politics without Latinos and replenishment of the ethnic group.

Ethnic Attrition

We define ethnic attrition as the process of putative immigrant origin coethnics opting out of self-identifying as a particular ethno-racial group category, often instead identifying as the dominant ethno-racial group category in the immigrant receiving society. In the context of Latinos/Hispanics and their relationship with the U.S. immigrant-receiving host society, ethnic attrition would mean ceasing to ethno-racially identify as Latino/Hispanic and instead as a different ethno-racial group. Prior evidence shows U.S. Latinos, on average, are likely to assimilate and ethnically attrit to whiteness. Although Latinos are ethno-racially

heterogeneous, many of those living in the United States are perceived as white or may identify as such in light of their European ancestry (Alba, 2020). Latinos also tend to marry whites if they intermarry, leading their offspring to be more inclined to claim a white identity despite Latino origins (Alba and Islam, 2009). Moreover, to the extent inter-ethnic contact may encourage attrition, many Latino immigrants and their offspring tend to move to (and prefer to live in) increasingly white neighborhoods as a function of length of stay in the United States and generational status (Robertson and Roman, 2023; Roman, 2023). Likewise, to the extent Latino neighborhoods become less Latino over time such that the remaining Latino population is exposed to a different populace along ethno-racial lines, Latino neighborhoods primarily become increasingly white, perhaps as a function of macro-level demographic trends such as white growth and gentrification in urban areas (Mumm and Sternberg, 2023).

Ethnic attrition among Latinos is a well-documented phenomenon. Although “Latino” or “Hispanic” are politically meaningful and potent categories the putatively Latino mass public can identify with, the origins of these categories are still a political construction. The Latino ethnic category is a product of political organizing from disparate national origin groups to garner governmental resources. The inherent ambiguity that determines inclusion in the Latino group category was intentional for the purposes of increasing group size and subsequently developing a politically relevant demographic bloc (Mora, 2014). The politically constructed nature of the category may also give way to fluidity in terms of individual ethno-racial self-categorization among segments of the putative Latino mass public. Indeed, prior research leveraging Census and survey data identifies a number of putative Latinos who do not self-identify as Latino or a Latino national origin group (e.g. Mexican-American) despite possessing Latino ancestry or Latin American origins (Alba and Islam, 2009; Duncan and Trejo, 2011a; Emeka and Vallejo, 2011; Lopez et al., 2017; Antman et al., 2016). A recent descriptive assessment finds initial Hispanic identifiers are one of the racial groups most prone to shedding their ethnoracial label across various panel surveys, with an average of 1 in 5 doing so over the course of four years (Agadjanian, 2022). Generational status is also

highly prognostic of attrition, with Antman et al. (2020) identifying 98%, 94%, 81% and 77% of Mexican-origin children ages 17 and younger who are 1st, 2nd, 3rd, and 4th-generation immigrants and identify as ethnically Mexican-American.

How Coethnic Growth Sustains Ethnic Identity

Although prior research has highlighted the role of inter-ethnic marriage, generational status, and discrimination in motivating ethnic attrition (Alba and Islam, 2009; Antman et al., 2020; Hadah, 2023), there is relatively little theoretical formalization and evidence explicating the role that demography, specifically local-level coethnic population growth, plays in undercutting ethnic attrition among Latino individuals nationally. However, prior research does provide some insight on the link between demography and different dimensions of ethnic identity. Some research finds living among coethnics or in contexts with coethnic resources (e.g. ethnic restaurants) is positively associated with stronger affective expressions of in-group identity (Wilcox-Archuleta, 2018). A common thread in political science is that coethnics living in coethnic geographic contexts are more likely to engage in pro-group political behavior (Ichino and Nathan, 2013; Fraga, 2018; White and Laird, 2020). Manrique and Mouw (2023) evaluate a different dimension of ethnic identity (ethnic language maintenance) and show US-born Latinos who live in metropolitan and neighborhood contexts with more US-born coethnics that speak Spanish are more likely to be bilingual. In terms of research on the demography-identity link where *ethnic attrition* is the outcome, Fernández et al. (2018) find that Latinos who live in neighborhoods with more coethnics are more likely to consistently report Latino/Hispanic self-identification across censuses. Although not the center of their analysis, Agadjanian and Lacy (2021) show non-whites are more likely to switch their ethno-racial identity to “white” conditional on living in a predominantly white area.

Perhaps the strongest formalization in the preexisting literature on how coethnic population growth may sustain ethnic identity is *ethnic replenishment theory* (ERT) (Jiménez, 2008;

Jiménez, 2010). Jiménez (2010) examines how growth in unacculturated coethnic populations (i.e. immigrant, Spanish-speaking) affect expressions of ethnic identity, perceptions of ethno-racial discrimination, and Spanish language maintenance among later-generation Mexican-Americans (i.e. third, fourth) living in two cities (Garden City, KS; Santa Maria, CA). Jiménez (2010) posits an influx of less acculturated immigrant, Spanish-speaking coethnics sustains different dimensions of ethnic identity because the immigration of new coethnics to local contexts reinforces both inter- and intra-group boundaries, motivating downstream identification with the ethnic group among individuals who do not fit the stereotypical schema of the in-group (e.g. 4th generation citizens who only speak English). For instance, regarding *intra-group boundaries*, new immigrant Spanish-dominant coethnics may discipline later-generation coethnics for not maintaining the Spanish language, which may encourage the maintenance of the Spanish language or alternative expressions of coethnicity in the face of intra-group exclusion among acculturated coethnics. Likewise, regarding *inter-group boundaries*, new immigrant Spanish-dominant coethnics may generate xenophobic and ethno-racist backlash on part of dominant group members (i.e. Anglo whites) within or near the local geographic context. Xenophobic backlash sharpens intergroup boundaries, potentially subjects later-generation Mexican-Americans to discrimination, and motivates later-generation Mexican-Americans to adopt a stronger sense of ethnic identity to display solidarity to new immigrants.

Although *ERT* is highly informative concerning the demography-identity link, it may not speak broadly as to how coethnic population growth sustains ethnic identity and undercuts attrition for four reasons. First, *ERT* speaks to how *less acculturated* growth maintains ethnic identity. Left unclear is if and how local-level Latino population growth writ large and more acculturated Latino population growth (e.g. non-immigrant, English-dominant, later-generation Latinos) may sustain ethnic identity. Second, *ERT* primarily focuses on how ethnic identity is maintained among more acculturated, later-generation segments of the Mexican-American population. Although ethnic identity primarily weakens among

acculturated Latino coethnics, attrition is still possible even among first-generation Latinos proximate to the immigrant experience. Therefore, it is unclear how Latino population growth may influence the prospect of ethnic attrition among less acculturated Latinos. Third, the evidence buttressing *ERT* from Jiménez (2010) is intrinsic to a few city contexts and the Mexican-American experience. It is unclear if the empirical implications inherent to *ERT* replicate for the Latino population writ large across the nation. Fourth, although *ERT* speaks to the salience of identity among later-generation Mexican-Americans, it does not explicitly theorize about and evaluate the consequences of coethnic growth on ethnic attrition.

To this end, we extend *ERT* and posit local-level coethnic growth writ large can undercut the prospect of ethnic attrition among Latino coethnics. First, coethnic growth can replenish and increase ethnic human capital and concomitant resources. New coethnics in a particular geographic space provide the human capital for ethnic resources, such as restaurants, civic organizations, and linguistic pressure that sustain the salience of individual-level ethnic identity (Wilcox-Archuleta, 2018). Second, psychologically, putative Latinos may choose to not attrit in order to facilitate the maintenance of their individual self-esteem. Social identity theory (SIT) posits that people identify with groups and subsequently support said groups in order to facilitate and maintain self-esteem (Tajfel et al., 1979). An assumption inherent to SIT is that group members are more likely to dissociate from their in-group or identify with another group if their in-group is perceptibly derogated, weaker within a social hierarchy, and/or provides relatively limited benefits in terms of material or psychological resources (Tajfel et al., 1979; Bedolla, 2003; Bedolla, 2005). coethnic population growth may provide a prospective signal of coethnic demographic, and possibly by extension, sociopolitical, dominance in the local geographic vicinity. These signals may incentivize coethnics to maintain their ethnic identity in order to continue to be a part of an increasingly dominant group that provides both psychological and material benefits.

- **H1:** Local Latino population growth will reduce Latino ethnic attrition

Heterogeneity by Marginal *Latinidad*

Coethnic growth may influence attrition in a heterogeneous manner conditional on different types of putative coethnics. Some coethnics may be “marginal coethnics,” that is, individuals who possess characteristics that predispose them to attrit from the ethnic group category. In the context of Latinos in the United States, these may be Latinos who are of a higher generational status, live in predominantly non-Latino areas, and are protestant Christians (as opposed to Catholics), all of which prior evidence identifies as being more likely to attrit (Antman et al., 2016; Lopez et al., 2017; Fernández et al., 2018; Agadjanian, 2022). Conversely, some coethnics may be more prototypical or fit stereotypical schemas of what a Latino coethnic is (i.e. immigrant, non-citizen, undocumented, Spanish-dominant) (Jones et al., 2019). Given marginal coethnics are more likely to attrit, the marginal influence of local coethnic growth may be psychologically and materially stronger for them relative to their more prototypical, stereotypical, less acculturated counterparts. Indeed, Jiménez (2010) primarily focuses on how a local-level influx of immigrants affects later-generation Mexican-Americans precisely because the salience of ethnic identity prior to the influx is not sufficiently primed by the ethnic resources immigrants have to offer.

- **H2:** Local Latino population growth will reduce ethnic attrition primarily among *marginal Latinos*

The expectation that coethnic growth will mitigate ethnic attrition among marginal Latinos is not necessarily guaranteed. If Latino population growth is driven by immigration or births of group members that are less acculturated, prototypical, and/or stereotypical, marginal Latinos may feel less prototypical and therefore less inclined to identify with the ethnic group (Goldman and Hogg, 2016). Moreover, given marginal Latinos may also hold negative attitudes toward new coethnics who possess stereotypical Latino characteristics (Roman, 2023; Fraga et al., 2024), they may perceive the influx of new coethnics as a stigmatizing threat and therefore be inclined to dissociate from the ethnic group (Bedolla,

2003; Bedolla, 2005).

Lastly, it's worth noting this moderating role of marginal Latinidad may not be linear. Among the most acculturated, increasing coethnic presence may do little to undercut disassociation from Latino identity because the forces behind this attrition are already so strong. Instead, it may be those in the middle — having a moderate degree of acculturation, experiencing some ambiguity in racial position, and living in more heterogeneous contexts — who prove most reactive. This nuance demands more from the data, and thus we only explore it closely in Study 3 when the fine-grained nature of marginal Latinidad measurement makes this possible.

Heterogeneity by Types of Latino Population Growth

The extent to which coethnic growth mitigates ethnic attrition may be conditional on the characteristics of the new coethnic population. Consistent with Jiménez (2010), less acculturated coethnics who possess stronger ethnic resources (e.g. Spanish-language skills, ethnicized tastes) or are more proximate to the immigrant experience may be more capable of sharpening inter- and intra-group boundaries in a manner that sustains the salience of ethnic identity for other coethnics.

- **H3:** Less acculturated local Latino population growth will reduce ethnic attrition among Latinos more strongly than more acculturated local Latino population growth.

Again, the expectation outlined in **H3** is not necessarily obvious. More acculturated coethnic growth may be more likely to mitigate ethnic attrition if, on balance, coethnics who live in particular geographic contexts feel less acculturated growth signals they are atypical members of the ethnic group.

Study 1

Data

To test **H1-H2**, we use stacked data from the nationally representative Cooperative Election Study (CES) between 2010-2022. The CES is critically advantageous for testing our hypotheses for two reasons. First, the CES allows us to measure ethnic attrition through two survey questions. The first question asks respondents to report their race/ethnicity, with an option to choose Latino/Hispanic. The second question asks respondents if they have Latino/Hispanic origins conditional on *not* reporting they are Latino/Hispanic to the first question. Therefore, we can identify two sets of respondents: those who initially identify ethno-racially as Latino/Hispanic (non-attributors); and those who do not initially identify ethno-racially as Latino/Hispanic but report they have Latino/Hispanic origins (attributors). Second, by stacking CES data between 2010-2022, the CES offers an unprecedentedly large sample of Latinos (attritor-inclusive, N = 61,000), who tend to be underrepresented in political surveys (Barreto et al., 2017). A large Latino sample is necessary to test our hypotheses since we are assessing the relationship between aggregate geographies and individual-level ethnic attrition in addition to the heterogeneous influence of geography on attrition.

Although the CES is advantageous, there are caveats. The CES primarily samples English-speaking registered voters, generating a sample that may not reflect the national Latino population (many of whom only speak Spanish, are not registered, and are not citizens) and is likely much more acculturated to the U.S. host society (Roman, 2023). Given our theory implies coethnic population growth will have a stronger influence on undercutting attrition among more acculturated Latino coethnics, our results may overstate the relationship between Latino population growth and ethnic attrition. However, we provide evidence our results hold in less acculturated populations reflected in the national Latino population that the CES underrepresents. Consistent with prior research demonstrating statistical relationships in unrepresentative surveys generalize to more representative surveys (Vitriol et al., 2019),

our results replicate when we subset to non-citizen Latinos accidentally sampled into the CES, a much less acculturated segment of the Latino public (Roman, 2023).¹ Regardless, the CES target sample is the *politically relevant* sample political scientists should prioritize studying. Although Latinos in the CES are unrepresentative of the entire Latino population by virtue of their registration status, they have the capability to vote and produce electoral consequences that may be the byproduct of their decision to maintain their ethnic identity.²

Outcome

The outcome of interest is *attrition*. As previously mentioned, in the CES, attritors are those who do not report Hispanic in response to the following question: “what racial or ethnic group best describes you?”³ But, they report “yes” to the follow-up question: “Are you of Spanish, Latino, or Hispanic origin or descent?”⁴ Thus, *attrition* is a binary indicator equal to 1, 0 otherwise, if a respondent does not initially ethno-racially identify as Hispanic but does report they are of Spanish, Latino, or Hispanic origin/descent.

Importantly, this *attrition* measure characterizes cross-sectional *snapshot attrition*, that is, a self-report of a conscious choice to not initially ethno-racially identify as Hispanic before they indicate that they possess Latino/Hispanic ethnic origins. Thus, the decision to attrit may be the product of longer-term factors (e.g. socialization in a non-ethnic context, intermarriage with a non-ethnic, being of a higher immigrant generational status, mixed ethno-racial parentage). Indeed, relative to panel data measuring attrition or switches in ethno-racial self-identification over the course of a few years (Egan, 2020; Agadjanian and Lacy, 2021; Agadjanian, 2022), the proportion of putative Latinos in our sample that are attritors is relatively high (27%).

¹Although the correlation is certainly weaker, consistent with **H2** (Figure 4).

²We also conceptually replicate Study 1’s findings using the Latino subset of the UCLA Nationscape survey, which targets a nationally representative population sample irrespective of registration status, implying Study 1’s results generalize to the broader Latino population and are not driven by sample composition (Section XXX).

³With options to otherwise report White, Black, Asian, Native American, Mixed, Other, or Middle Eastern

⁴The follow-up is only asked of those who do not report Hispanic to the first question.

We theoretically validate the *attrition* measure. If we are actually measuring something tantamount to ethnic attrition, then we should expect *attrition* to be associated with a number of socio-demographic and political characteristics that are akin to Anglo whites. Indeed, most attritors in the CES primarily identify as white or mixed-race (Figure A1, which may reflect high rates of Latino-white intermarriage, see Alba and Islam (2009)). Moreover, *attrition* is positively associated with higher socio-economic status, identifying as non-Catholic, and political conservatism (broadly construed), characteristics more akin to Anglo whites on average.

Measuring Latino population growth

We use data from the 2000 Census in addition to the 2010-2021 5-year American Community Surveys to measure our main independent variable of interest, Latino population growth ($\Delta \% \text{ Latino}$). We measure $\Delta \% \text{ Latino}$ at the zipcode-level to approximate neighborhood-level dynamics. Generally, $\Delta \% \text{ Latino}$ characterizes relatively long-term shifts. We take the difference in the proportion of the zipcode population that identifies as Latino between 2000 and the year prior to the relevant CES respondent being surveyed.⁵ For instance, $\Delta \% \text{ Latino}$ measures 11 years of demographic change for respondents interviewed in the 2012 CES wave, and 15 years of demographic change for respondents interviewed in the 2016 CES wave.⁶ Our measure of $\Delta \% \text{ Latino}$ is largely consistent with prior work evaluating the political consequences of local-level demographic shifts (Newman, 2013; Newman and Velez, 2014; Hill et al., 2019). Although demographic shifts can be measured in several different ways, we later show our results are invariant to alternative measures of Latino demographic change and measuring $\Delta \% \text{ Latino}$ at higher levels of geographic aggregation. Evaluating shifts in demographic *change* (as opposed to *levels* characterizing the proportion of the

⁵We choose the year prior to the interview year to ensure estimates are less perturbed by post-treatment bias.

⁶Given the absence of 2009 ACS zipcode data, we use 2010 ACS zipcode data to measure $\Delta \% \text{ Latino}$ for the 2010 CES wave. The 2010 CES wave respondents are the only respondents where $\Delta \% \text{ Latino}$ uses data measured concurrent to the survey.

zipcode population that is Latino) is also advantageous for theoretical and empirical reasons. Theoretically, demographic *change* better approximates our theoretical insights suggesting replenishment of coethnic human capital can sustain a strong(er) sense of ethnic identity (Jiménez, 2010). Empirically, demographic *change* is less susceptible to residential selection bias since demographic change (as opposed to levels) are relatively difficult to predict and control among respondents (Hopkins, 2011).

To test **H3**, we use 2000 Census and 2010-2021 5-year American Community Survey data to measure different types of Latino population growth that allow us to evaluate the relationship between *attrition* and more versus less acculturated growth. First, we measure change in the proportion of the zipcode population that is Spanish-speaking Latino ($\Delta \%$ *Spanish-speaking Latino*, less acculturated) and change in the proportion of the zipcode population that is English-speaking Latino ($\Delta \%$ *English-speaking Latino*, more acculturated) between 2000 and the year prior to CES respondent survey interview. According to the Census, Spanish-speaking Latinos are individuals who are 5 years or older that can Speak Spanish or are bilingual. English-speaking Latinos are individuals who are 5 years or older that *only* speak English. Second, we measure change in the proportion of the county population that is immigrant Latino ($\Delta \%$ *Immigrant Latino*, less acculturated) and change in the proportion of the county population that is non-immigrant Latino ($\Delta \%$ *Non-immigrant Latino*, more acculturated).⁷ We measure the count of immigrant Latinos at the county-level by using data on the number of immigrants from Latin America within each county. We then measure the count of non-immigrant Latinos by taking the total number of Latinos within each county and subtracting the number of immigrant Latinos from that value.

⁷Due to small sample size, measurement of foreign-born Latinos is only possible at the county-level using ACS data, therefore we use county-level growth metrics when we assess the association between change in immigrant/non-immigrant Latino populations and *attrition*.

Measuring Marginal Latinos

To test **H2**, we measure marginal Latinos in three theoretically and empirically motivated ways. Prior research suggests Latinos who have a weaker sense of ethnic identity and hold attitudes that are politically unfavorable to their in-group tend to be Latinos who a) live in predominantly non-Latino areas (Telles and Sue, 2019), b) are further from the immigrant experience (i.e. higher generational or citizenship status) (Antman et al., 2020; Roman, 2023), and c) are non-Catholic (Calvillo and Bailey, 2015). Indeed, these are also the types of Latinos who are predisposed to attrit from the ethnic group category. To this end, we evaluate the heterogeneous association between $\Delta \% \text{ Latino}$ and *attrition* by a) baseline zipcode-level $\% \text{ Latino}$ in 2000; b) generational and citizenship status (*immigrant status*); and c) *non-Catholic* (as opposed to Catholic) religious identity. We measure *immigrant status* as binary indicators for Latinos who self-report in the CES they are immigrant citizens and second, third, fourth generation Latinos (non-citizen immigrants = reference). *Non-Catholic* is a binary indicator equal to 1 if a respondent indicates they are not Catholic, 0 otherwise. Our expectation is that the association between $\Delta \% \text{ Latino}$ and *attrition* will be stronger for Latinos that live in neighborhoods that are less Latino at baseline; stronger for Latinos who citizens and of a higher generational status; and stronger for non-Catholic Latinos.

Controls

Our models adjust for a number of theoretically relevant demographic, socio-economic, political, and geographic covariates. Demographic covariates include: age, gender, marital status, parental status, Catholicism, Evangelicalism, national origin,⁸ immigrant status; socio-economic covariates include: income, unemployment, union member, college-education; political covariates include: partisanship and ideology; geographic covariates include: median household income, % college-educated, % unemployed, total population, and population

⁸National origin is measured with binary indicators for Mexican, Puerto Rican, Dominican, Cuban, and Central American. National origin is not measured in CES waves prior to 2015, therefore, a national origin missingness indicator is used for CES waves prior to 2015.

density at the zipcode-level using 2000 Census data. Moreover, all models adjust for binary state-level indicators for California, Florida, Texas, New York, and Arizona.

Results

Figure 2 characterizes evidence supporting **H1**.⁹ $\Delta \% Latino$ is negatively associated with *attrition* (Panel A). The min-max coefficient is substantively large (-0.50 , $p < 0.001$), equivalent to 185% of the outcome mean (0.27). We assess zipcode-level growth in other ethno-racial groups to garner a sense of how Latino population growth drives *attrition*. $\Delta \% White$ is strongly positively associated with *attrition* (Panel B), with a min-max coefficient of 0.32 ($p < 0.001$). However, $\Delta \% Black$ and $\Delta \% Asian$ are not associated with *attrition* (Panels C-D). These statistical patterns are sensible, given growth in the Latino population is strongly inversely related to growth in the non-Latino white population (Pearson's $\rho = -0.72$) but is not correlated with growth in the Black ($\rho = -0.15$) or Asian population ($\rho = 0$). However, although $\Delta \% Black$ is not associated with *attrition*, it is associated with Black self-identification (as opposed to Hispanic) (Table A4, Model 3). But, $\Delta \% Asian$ is not associated with Asian self-identification (Table A4, Model 4). These patterns make sense given many Latinos may be African-descendent and Black-Hispanic intermarriage rates are higher than Asian-Hispanic intermarriage rates (Livingstone and Brown, 2017; Morgan et al., 2024).

We test an additional implication of **H1**. To the extent Latinos primarily attrit to identifying as ethno-racially white, does Latino population growth primarily forestall identification as white among those with Latino ancestry? Figure 3 characterizes min-max $\Delta \% Latino$ coefficients from independent fully-specified models where the outcomes are binary indicators for self-identifying as white, Black, Asian, mixed, or other. The min-max $\Delta \% Latino$ coefficients are statistically significant irrespective of the outcome (at least $p < 0.05$). However, the min-max $\Delta \% Latino$ coefficient is larger when the outcome is white self-identification (-0.32)

⁹For Study 1 estimating equations, see Section A.2 for details.

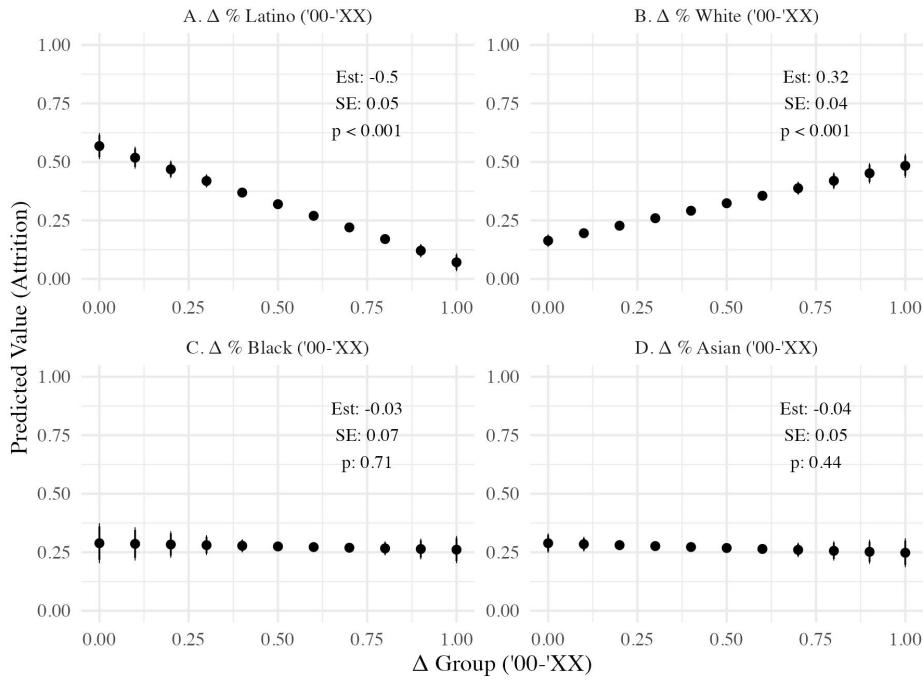


Figure 2: Latino population growth mitigates *ethnic attrition*. X-axis is change in Latino (Panel A), white (Panel B), Black (Panel C), and Asian (Panel D) zipcode population proportion. Y-axis is the *ethnic attrition* predicted value from a fully specified model with covariates held at their means. 95% CIs displayed from zipcode-clustered robust SEs.

relative to Black (-0.07), Asian (-0.02), mixed (-0.05), and other self-identification (-0.04). Formal $\Delta \% \text{ Latino}$ coefficient difference tests for the white self-identification outcome relative to the Black, Asian, mixed, and other self-identification outcome are statistically significant ($p < 0.001$). Thus, although Latino population growth forestalls ethno-racial attrition to all other relevant ethno-racial categories, growth primarily forestalls ethno-racial attrition into whiteness.

Does coethnic growth undercut attrition among marginal, atypical, Latinos?

We find evidence in support of **H2**. First, $\Delta \% \text{ Latino}$ has a stronger association with *attrition* among Latinos who, at baseline, live in less Latino zipcode contexts (Figure 4, Panel A). For Latinos who already live in highly Latino contexts, *attrition* is functionally zero and $\Delta \% \text{ Latino}$ has a a statistically null and substantively minimal relationship with *attrition*.

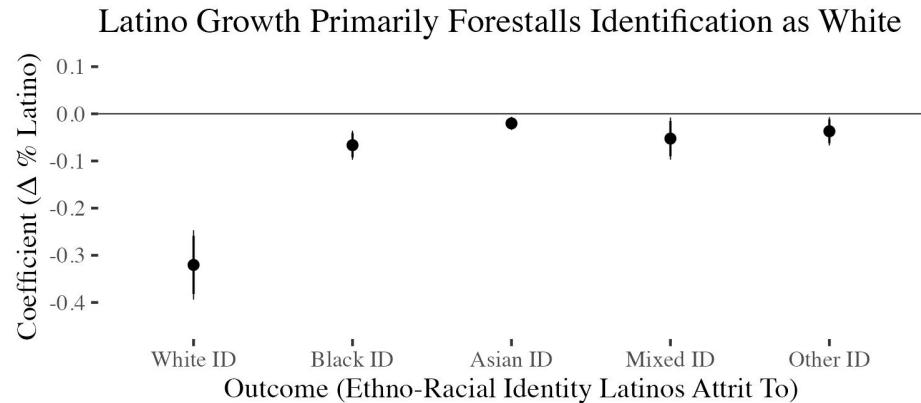


Figure 3: Latino population growth primarily forestalls identification as white. X-axis is the outcome for a separate fully-specified regression model, y-axis is the $\Delta \% \text{ Latino}$ coefficient. 95% CIs displayed from zipcode-clustered robust SEs.

Conversely, for Latinos who live in places with nearly no Latinos at the zipcode-level and the middle level of $\Delta \% \text{ Latino}$, attrition is 41%. But, for Latinos who live in places with nearly no Latinos and at the highest level of $\Delta \% \text{ Latino}$, attrition is functionally zero and converges with Latinos who live in predominantly Latino contexts. Second, the association between $\Delta \% \text{ Latino}$ and *attrition* is heterogeneous by immigrant and generation status (Figure 4, Panel B). The min-max $\Delta \% \text{ Latino}$ coefficient is statistically larger for immigrant (-0.57), second-generation (-0.53), third-generation (-0.44), and fourth-generation+ (-0.48) citizens relative to immigrant non-citizens (-0.26). Finally, $\Delta \% \text{ Latino}$ has a stronger association with attrition among non-Catholic Latinos relative to Catholic Latinos (Figure 4, Panel C). On average, non-Catholic Latinos have higher rates of attrition than Catholic Latinos. But, for non-Catholic Latinos who live in zipcodes with higher rates of Latino growth, rates of attrition converge with Catholic Latinos at zero. In sum, consistent with **H2**, sustained Latino population growth undercuts attrition among Latinos who are predisposed to attrit from the ethno-racial group category.

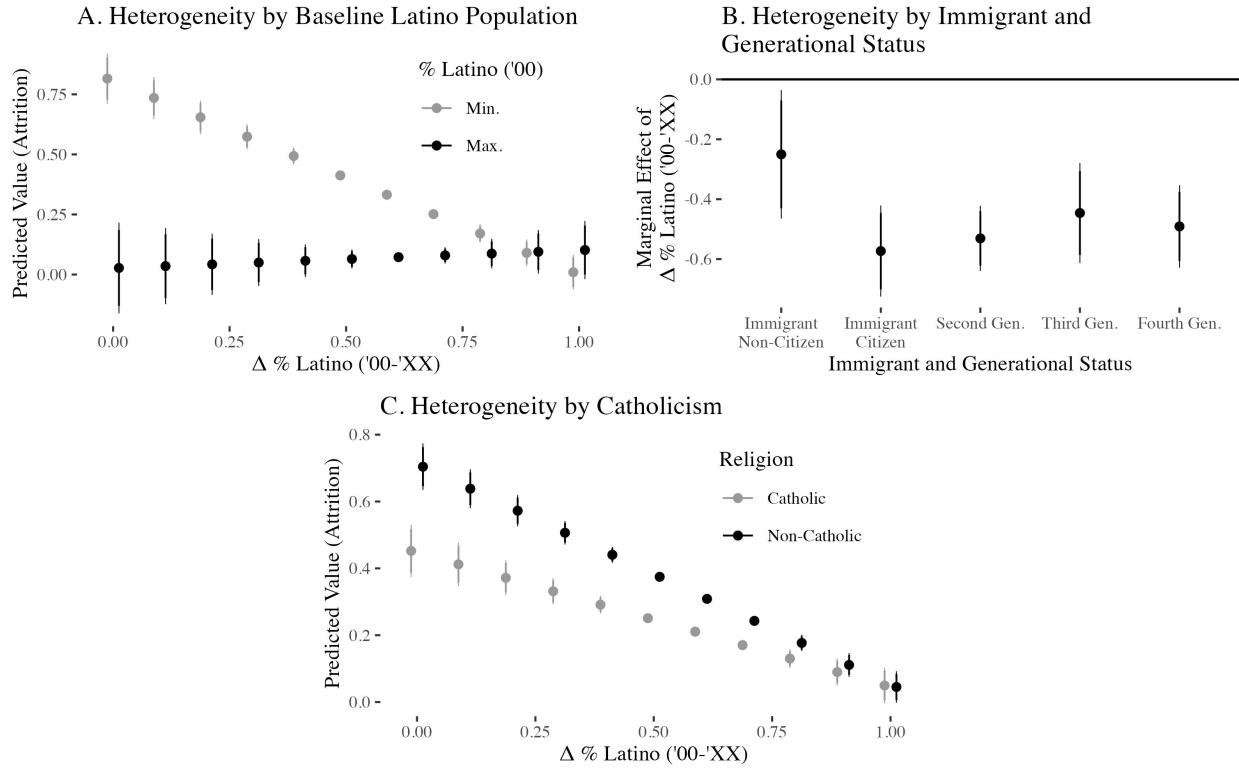


Figure 4: The influence of Latino population growth on mitigating ethnic attrition is stronger for marginal Latinos. 95% CIs displayed from zipcode-clustered robust SEs.

Does the type of coethnic growth matter for attrition?

We find evidence supporting **H3**. Figure 5 characterizes the association between different types of Latino population growth and *attrition*. Panel A characterizes the association between zipcode-level growth in the Spanish-speaking Latino population and *attrition* in addition to growth in the English-speaking Latino population and *attrition*. Consistent with **H3**, The min-max coefficient for Spanish-speaking Latino population growth (-0.54) is larger than the min-max coefficient for English-speaking Latino population growth (-0.30). This difference is statistically significant ($0.24, p < 0.01$). Likewise, Panel B characterizes the association between county-level immigrant Latino population growth and *attrition* in addition to non-immigrant Latino population growth and *attrition*. Again, consistent with **H3**, the min-max coefficient for immigrant Latino population growth (-0.20) is larger than the min-max coefficient for non-immigrant Latino population growth (-0.09). This difference

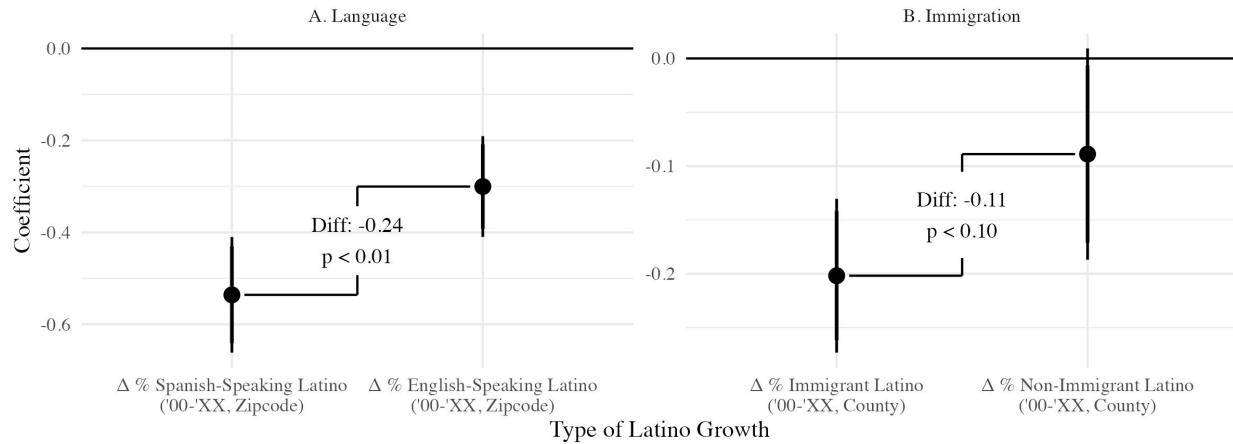


Figure 5: Less acculturated Latino population growth mitigates ethnic attrition. Panel A characterizes the association between Latino population growth and *attrition* decomposed by growth in the Spanish and English-speaking Latino population (zipcode-level). Panel B characterizes the association between Latino population growth and *attrition* decomposed by growth in the immigrant and non-immigrant Latino population (county-level). Coefficients on each panel are from a single model. Annotations characterize formal coefficient difference tests. 95% CIs displayed from zipcode- (Panel A) and county-clustered (Panel B) robust SEs.

is statistically significant ($0.11, p < 0.10$).¹⁰ These results are consistent with the notion that less acculturated local Latino population growth sustains ethnic identity among the Latino population given the ethnic resources that less acculturated Latinos have to offer.

Robustness Checks

We conduct several robustness checks. Our results are not sensitive to independent variable measurement (Table A2). The results replicate if we measure $\Delta\%$ *Latino* based on binary indicators for whether respondents live in zipcodes above median $\Delta\%$ *Latino* (relative to below median $\Delta\%$ *Latino*), in the 2nd-3rd tercile of $\Delta\%$ *Latino* relative to the 1st tercile, and in the 2nd-4th quartile of $\Delta\%$ *Latino* relative to the 1st quartile (Models 1-3). Our re-analyses using “binned” operationalizations of $\Delta\%$ *Latino* are reassuring since they suggest

¹⁰Given each panel on Figure 5 characterizes coefficients from a single model, different types of Latino population growth may be highly correlated with one another such that model estimates are sensitive to multi-collinearity. However, multi-collinearity is not a problem since different types of Latino growth are not strongly related to each other. Growth in the Spanish-speaking and immigrant Latino population is minimally negatively related to growth in the English-speaking and non-immigrant population respectively (Pearson’s $\rho = -0.1, -0.02$).

our results are not driven by our model’s linearity assumption on Figure 2.¹¹ Likewise, the results replicate if we assess: the relationship between $\% Latino$ in 2000 and *attrition* (Model 4); the relationship between *attrition* and percent change in $\% Latino$ between 2000 to the year before being surveyed (Model 5); and the relationship between *attrition* and change in the total Latino population between 2000 to the year before being surveyed (adjusting for change in the total population, Model 6). Moreover, we also demonstrate our results replicate if we measure $\Delta \% Latino$ at the county-level (Table A5), which minimizes the risk our results are driven by choice of geographic aggregation for our main independent variable of interest. Therefore, our theory and evidence is invariant to the type of Latino demography being measured, and in general, a greater concentration of Latinos at the local-level appears to be associated with reduced *attrition*.

We mitigate the risk our results are driven by residential selection. The CES is advantageous since it allows us to measure how long respondents have lived at the current address (up to “5 or more years”). Therefore, we re-analyze our results subsetting our data to Latino respondents who have lived in their present address at the maximum length of self-reported time. The logic here is that respondents with longer residential tenure are less likely to have moved into their place of residence because of recent ethno-racial demographic shifts such that the $\Delta \% Latino$ coefficient with the subsetted data is less likely to be perturbed by residential selection bias (Hopkins, 2011). Our results hold using the subsetted data (Table A6, Model 3).

We re-analyze the relationship between $\Delta \% Latino$ and *attrition* adjusting for state and county fixed effects. These fixed effects allow us to compare ostensibly more like zipcode units within each state and county, mitigating the risk our results are driven by unobserved fixed characteristics at the state and county-level (Pepinsky et al., 2024). The results do not change (Table A7).

¹¹Moreover, descriptive statistics show a clear, mostly monotonic relationship between $\Delta \% Latino$ and *attrition* (Figure A3).

Study 2

Study 1 finds evidence supportive of our theoretical expectations surrounding coethnic growth forestalling attrition, though for a specific type of attrition: “snapshot attrition,” based on cross-sectional survey data comparing reported ethnic origins and self-identification. In light of lessons from other work on Hispanic identity fluidity (Agadjanian, 2022; Fernández et al., 2018), Study 2 offers a theoretically relevant extension on Study 1 by conceptualizing attrition in terms of full disassociation from Hispanic/Latino identity in the short-term and thus centering around changing identities that we know occur within-lifetime.

To accomplish this extension, we turn to high quality panel survey datasets asking racial identity — that includes Hispanic as a race — on waves 4-5 years apart: the Cooperative Election Study panel survey and the Democracy Fund Voter Study Group panel survey. We use the two furthest waves apart for each dataset, 2010/2014 and 2011/2016, respectively. Here we measure attrition as not identifying as Hispanic in wave 2 having done so in wave 1, and then apply a similar OLS specification as in Study 1: modeling attrition as a function of coethnic context and general demographic and political control variables, with standard errors clustered at the zip code level.

Although coethnic growth is negatively related to attrition as expected, the relationship is not statistically significant. This may reflect the small sample sizes of initial Hispanic identifiers that dampens variation on the coethnic growth variable (especially relative to the more expansive CES cross-sectional data). This issue may be less problematic for Hispanic *levels* (e.g., share of Hispanics at the zip code level). As the left panel in Figure 6 shows, this way measuring of coethnic context provides stronger evidence of curtailing ethnic attrition. Importantly, this data also allows us to speak to a potential reverse causal direction at play that could generate this relationship: Hispanics selecting into more Hispanic contexts over the course of the panel periods. Results in the right panel of Figure 6 cast doubt on this dynamic, illustrating null relationships between attritor status and moving into more Hispanic zip codes.

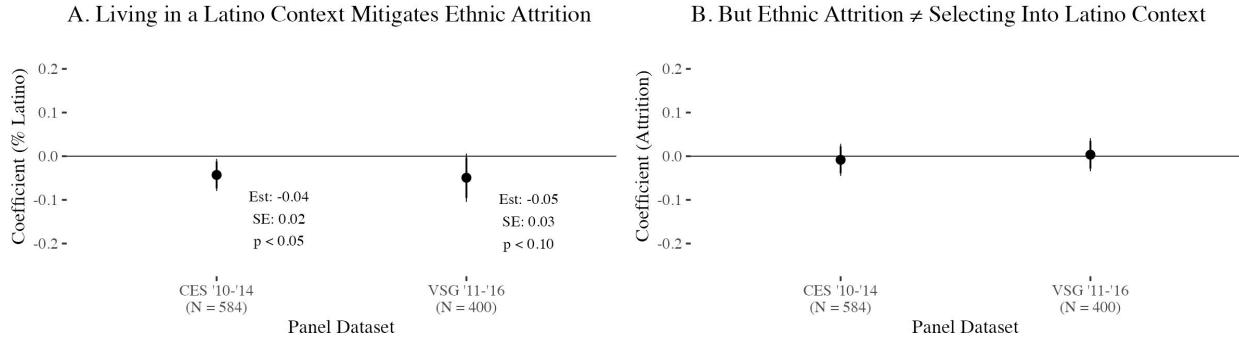


Figure 6: Living in zipcodes with more Latinos mitigates *ethnic attrition*, but *ethnic attrition* does not motivate selection into more Latino zipcodes. X-axis is the panel dataset at use. Y-axis is the % Latino (Panel A) and ethnic attrition (Panel B) coefficient. 95% CIs displayed from zipcode-clustered robust SEs.

In sum, we find more mixed evidence in Study 2, but it remains theoretically relevant and generally supportive of the idea that coethnic context constrains ethnic attrition. Moreover, an alternative explanation for this empirical relationship due to selective moving does not appear in the data.

Study 3

Using both cross-sectional and panel survey datasets, Studies 1 and 2 have consistently shown a role for local coethnic growth in constraining Hispanic attrition and that it appears most strongly for individuals disposed to attriting — the “marginal Latino.” In this third study, we replicate and expand on these patterns with a novel data source: voter registration files. In using survey data earlier, we presume it maps on cleanly to real world behavior. Yet prior work offers mixed conclusions about the link between subjective survey self-reports and behavior (e.g., Ansolabehere and Hersh, 2012; Kaiser and Oswald, 2022). For this reason and for the sake of diversifying data sources, we integrate data on real world, naturally occurring expressions of racial identity outside of a survey context: states that ask people to record their racial/ethnic identity as part of voter registration forms.

This extension is significant for other reasons as well. Because it occurs in the real world,

it carries greater potential for real life consequences for individuals. How people appear in administrative records can shape services they receive and life outcomes they experience, and in this context specifically, identity regulates how campaigns perceive and target individuals and constituencies. Moreover, this makes for a hard test of our expectations. Official voter registration represents an environment with real world stakes, which could constrain how individuals report information about themselves. Compared to a survey context, where identity expression and fluctuation across time and context presents no costs, we might expect more identity stability overall and less responsiveness to contextual factors.

Data

Seven states in the American South ask questions about race/ethnicity as part of voter registration forms. Out of convenience and as an initial test, we turn first to North Carolina. Not only are raw voter file datasets easily accessible online,¹² but this state also maintains historical voter file “snapshot” files: point-in-time records for all registrants going back several years. Because many of the same registrants appear multiple times across these snapshots, this data allows us to construct an organic individual-level panel dataset. We specifically choose snapshots taken on Election Days, as this makes it most likely for registrants to have had a chance to update their records.¹³

We use Election Day snapshots from 2016 and 2020, which have sample sizes of 5,933,652 and 6,602,052, respectively, after subsetting to active registrants in both years. Records are merged using a unique voter identification number that the North Carolina State Board of Elections preserves for voters across elections.¹⁴ The merges results in 4,797,842 matches, meaning 81% of active registrants from 2016 appear as active registrants in 2020. North Carolina asks race and ethnicity separately in the following way:

¹²<https://www.ncsbe.gov/results-data/voter-registration-data>

¹³This still does not fully sidestep the issue of “false negatives” — Hispanics attriting in the real world but not showing up this way in the data because they are not updating their identity responses on voter files in real time — though this is more of an immediate issue for establishing correct overall rates of attrition, and less so for how attrition relates to other variables, which is our focus.

¹⁴The uniqueness of this identifier was confirmed in email communication with a NCSBE analyst.

- Provide your demographic information (optional).
- Ethnicity question: Not Hispanic/Latino, Hispanic/Latino
- Race question: African American/Black, American Indian/Alaska Native, Asian, Multiracial, Native Hawaiian/Pacific Islander, White, Other

We focus attention on the ethnicity question for our purposes. The data itself contains possible responses of “HISPANIC or LATINO,” “NOT HISPANIC or NOT LATINO,” and “UNDESIGNATED” across the two years. We start by subsetting to registrants who mark “HISPANIC or LATINO” in 2016 and who also have records in 2020 ($N = 101,973$).¹⁵ We operationalize attrition as marking any response other than this one in 2020. This amounts to a mean attrition rate of 9.00%.¹⁶

For the independent variable, in line with analyses from Studies 1 and 2, we compute home zip code-level data on racial and other demographic traits from the Census and merge this onto registrants’ 2016 reported zip code. The primary independent variable is a 0-1 rescaling of the difference between 2015 and 2000 Hispanic zip code percentages. The same transformation is done for growth in the other three major racial groups. In later formal tests, we introduce control variables that draw on 2000 Census data for contextual variables and 2016 voter file data for individual variables (aiming for “pre” measures in both cases).

Main results

We begin this voter file-based exploration of the relationship between coethnic growth and Hispanic attrition with plotting the data in binned form. In order to avoid assumptions about a linear relationship but also retain as much information as possible, we first turn the continuous racial group growth measure into ventile form: a categorical variable with

¹⁵The final sample is around $N = 97,500$ after adding in right-hand side variables.

¹⁶This comprises 2.91% marking “NOT HISPANIC or NOT LATINO” and 6.09% marking “UNDESIGNATED.” We pool these attrition responses for now but plan to unpack these different types of attrition later.

20 groups of roughly equal size. Each group has several thousand observations. In Figure 7, we plot attrition rates (i.e., not identifying as Hispanic in 2020 among this 2016 Hispanic identifier subsample) against this ventile growth measure for four racial group growth types, overlaying a loess curve in each panel.

This provides initial support for our expectations. The key panel is in the bottom left: moving across the lowest to highest ventiles of Hispanic growth from 2000 to 2015 corresponds to lower attrition moving from 2016 to 2020. From a qualitative look, the relationship for Hispanic growth is steepest across all four types of growth. White growth once again positively relates to attrition, reflecting the complementary mechanical nature of these different comparisons (though growth types are inversely related). Unlike earlier results, Asian growth appears a bit more positively related to attrition. At the same time, this link is not terribly strong, possibly driven by an outlier data point, and, again, fits with what we might expect mechanically speaking.

We now move to formally testing the relationship between growth and attrition with OLS regression models that vary the stringency of included controls. Growth is measured continuously on a 0-1 scale, meaning that coefficients capture change in attrition when moving from minimum to maximum growth. Table 1 contains these results. Each model regresses attrition on four types of racial group growth independently (in terms of rows going down) and shows results across a mix of county fixed effects, inclusion of a baseline (2000) Hispanic share control, and inclusion of a wider set of controls. The latter set approximates earlier study tests as close as possible, entailing total population, median income, college share, unemployment rate, and population density (contextual level) and age, sex, and party registration (individual level). All models cluster standard errors at the level of the independent variable (2016 reported zip code) and, again, only include 2016 Hispanic identifiers.

The first row for Hispanic growth shows that these formal tests confirm the patterns observed from the raw data: across all models, Hispanic growth corresponds to less Hispanic attrition. Though the magnitude of the relationship varies — decreasing from 0.12 points to

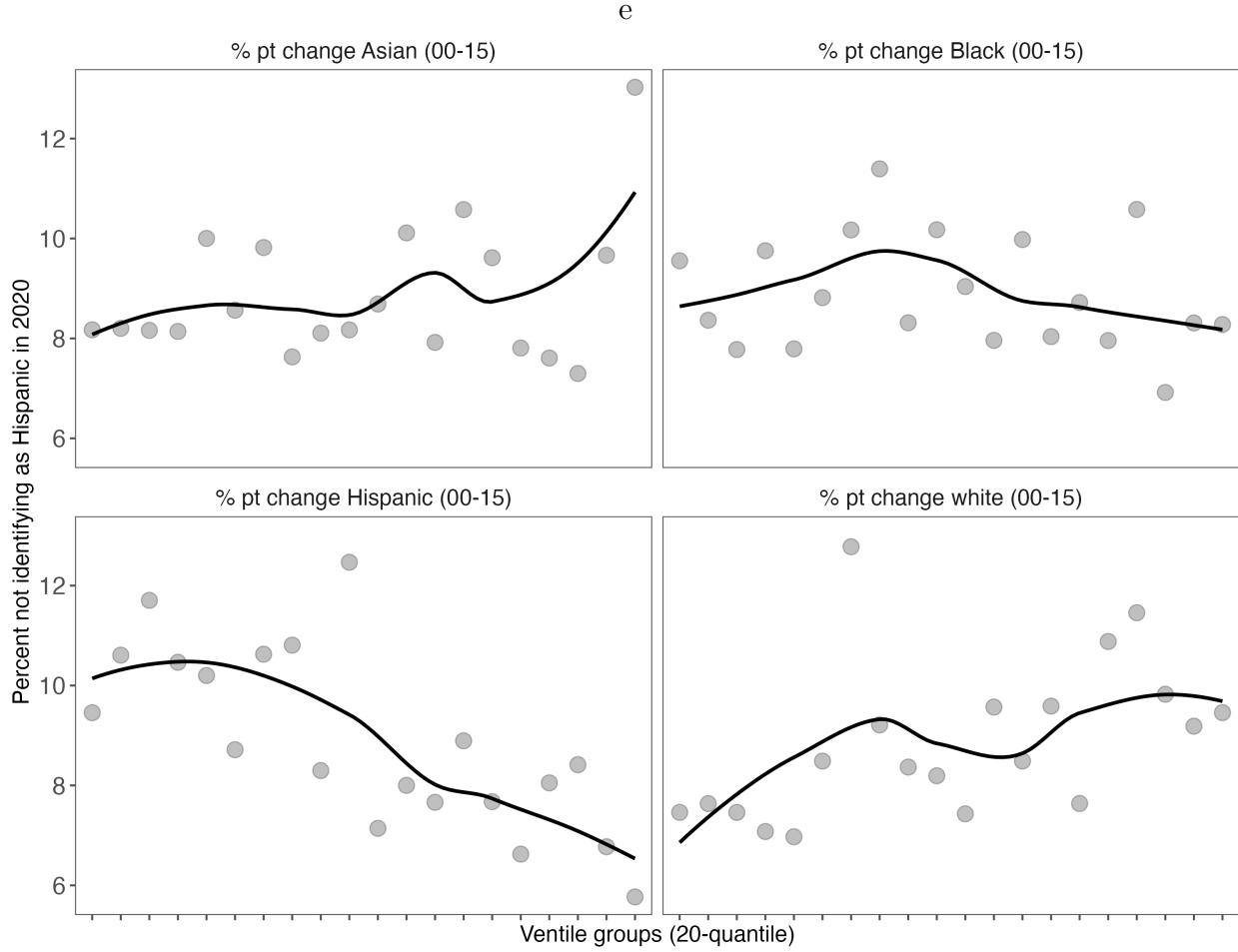


Figure 7: Hispanic attrition (2016-20) as a function of different racial group growth, grouped by ventiles (Study 3)

0.04 points going from the basic to most stringent model — it consistently keeps the right sign (negative) and is always significant at least at the $p < .05$ level (and always at the $p < .01$ level except in Model 6, where $p = .01$). Perhaps the cleanest test (Model 4, without full controls, some of which might be post-treatment of growth) returns very strong evidence in terms of magnitude and significance ($p < .001$). Even the smallest magnitude is substantively large in light of the 0.09 mean attrition rate. Models 2, 4, and 6 are especially illuminating. Even when including county FEs and thus looking only at within-county growth variation across zip codes, the significant constraining force of coethnic growth on attrition holds.

The second through fourth rows largely provide complementary evidence. White growth initially positively relates to Hispanic attrition, as we might expect from earlier results,

Table 1: OLS regressions of Hispanic attrition on different racial group growth (Study 3)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Hispanic growth	-0.12*** (0.02)	-0.08*** (0.01)	-0.10*** (0.02)	-0.07*** (0.01)	-0.05*** (0.01)	-0.04** (0.01)
White growth	0.09*** (0.04)	0.03 (0.03)	0.10*** (0.04)	0.03 (0.03)	0.04 (0.03)	-0.01 (0.03)
Black growth	-0.05 (0.04)	0.03 (0.04)	-0.10** (0.05)	0.00 (0.04)	-0.01 (0.04)	0.03 (0.03)
Asian growth	0.09* (0.05)	0.08*** (0.03)	0.08 (0.05)	0.08** (0.03)	0.03 (0.03)	0.05** (0.02)
Observations	97524	97524	97524	97524	97495	97495
County FEs		✓		✓		✓
Hisp '00 control			✓	✓	✓	✓
Full controls					✓	✓

Notes: Standard errors clustered at 2016 zip code level. Subset to 2016 Hispanic identifiers.
^{*} $p < 0.1$, ^{**} $p < 0.05$, ^{***} $p < 0.01$; two-tailed.

though coefficient significance and magnitude weakens and sign becomes inconsistent when layering on more controls. Black growth similarly shows no strong relationship with attrition. Curiously, the Asian growth measures show stronger signs of a positive relationship with attrition. However, the aforementioned complementary nature makes this less concerning, and most importantly, the magnitude, sign, and significance of Asian growth coefficients are less robust to different modeling choices than the same for Hispanic growth coefficients.¹⁷

Two other things are worth reflecting on for these main tests. First, concerns over reverse causality in this panel data setup entail attritors being more likely to move to contexts with greater Hispanic growth. While causal direction cannot be perfectly pinpointed, casting some doubt on this alternative dynamic is important. In analyses not shown here, we discover some evidence to suggest this alternative explanation is not at play: regressing a variable that captures moving to greater Hispanic growth zip codes (from 2016 to 2020) on an attrition

¹⁷Outliers from Figure 7 for Asian growth play some role in these results too: for example, when removing the highest growth ventile for Hispanic and Asian analyses for Model 4, the latter but not the former becomes non-significant.

indicator returns small magnitudes and insignificant relationships across modeling choices. Second, while notable, results for Hispanic growth in Table 1 are smaller than in Study 1 using cross-sectional data. Although not directly comparable, the divergence might reflect larger relationships accruing for attrition beyond the four-year span used for the linked voter files as well as weaker relationships for more “extreme” forms of attrition (full disassociation from Hispanic/Latino identity, as the voter file and more comparable Study 2 approach capture — unlike the cross-sectional data in Study 1).

Marginal Latino Moderator Tests

Following the spirit of earlier studies, we conclude this section by checking whether relationships for the main test are more pronounced for different manifestations of “marginal Latino” status, i.e., those who might be most disposed to attrition. Results here do not fully support earlier ones, but add some additional nuance to marginal Latino dynamics.

First, we run the same types of models from row 1 in Table 1 but interact Hispanic growth with baseline (2000) Hispanic share in a home zip code. The basic approach without any controls (Model 1) provides supportive evidence — the relationship between growth and attrition is most negative for less Hispanic zip codes at baseline — but this pattern does not withstand county fixed effects (Model 2 approach). Second, using the Census-style race/ethnicity distinction found in this voter file data, one might view “white” racial identifiers as marginal Latinos; once again, these individuals see their attrition a bit more constrained by coethnic growth but this moderating role falls apart with more controls.

Third, ethnic-sounding names may proxy for level of acculturation, tapping into marginal Latino status at another angle. We use the ‘wru’ package in R to predict likelihood of Hispanic/Latino identification for every observation based only on their reported surname from the 2016 voter file snapshot (Imai and Khanna, 2016). We then transform this continuous 0-1 scale to another fine-grained ventile grouping (with around $N = 5,000$ per group) and run the most stringent test (Table 1 Model 6) for each ventile group. Figure 8 shows results

along with an overlaid loess line to illustrate the overall pattern. Strong signs of a nonlinear relationship emerge: while weak magnitudes and possibly positive relationships appear at the highest and lowest ventiles of predicted Hispanic, middling ventile groups produce the strongest evidence of growth constraining attrition. For example, the largest absolute magnitude (-0.11) occurs between the 60th and 65th percentile of predicted Hispanic names. This differs from earlier evidence but may offer a nuanced perspective only available from fine-grained data. Perhaps the lowest ventiles capture hyper-acculturated individuals who react little to local coethnic growth, while it is highly-reactive individuals in the ambiguous name middle who represent the real marginal Latinos.

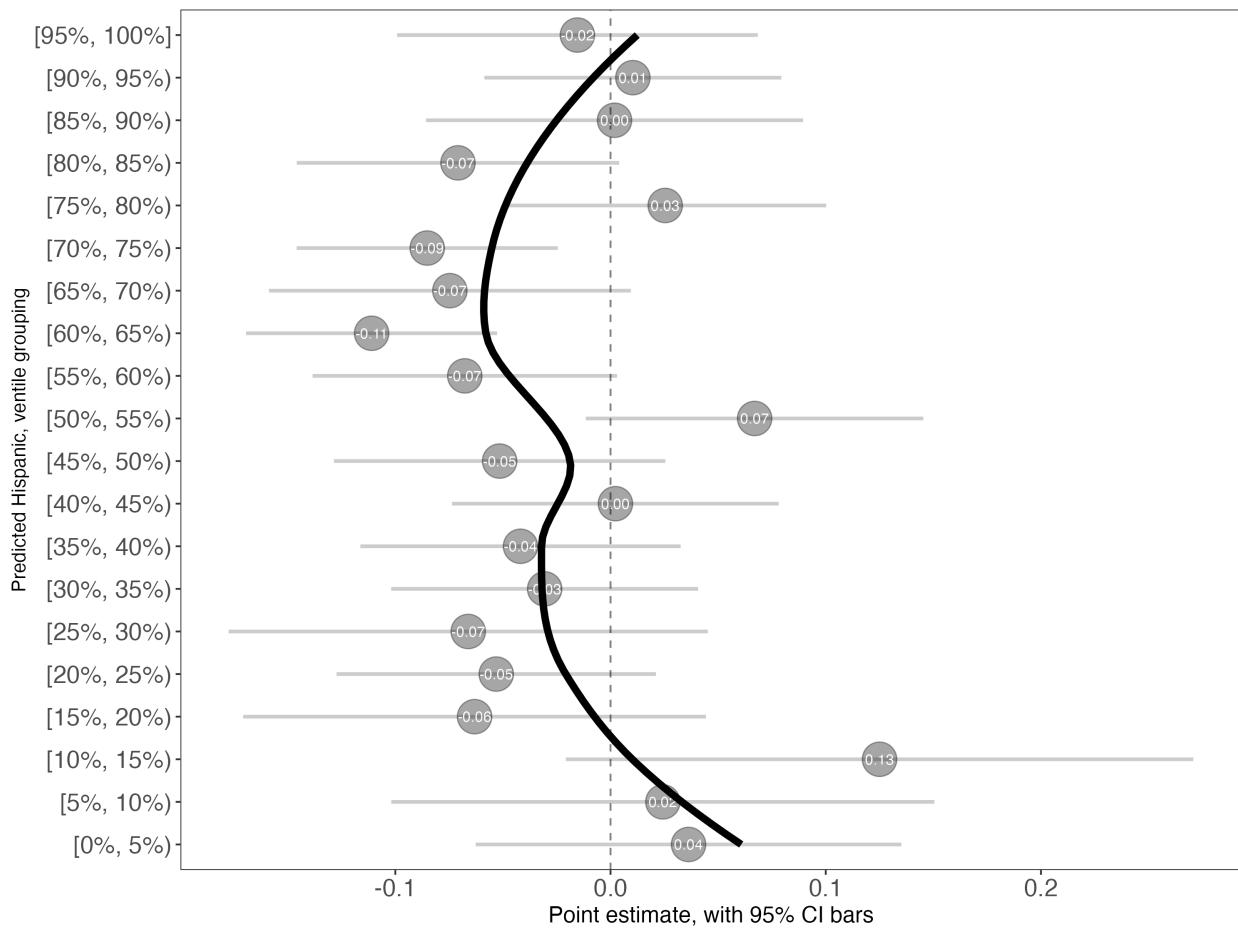


Figure 8: Full model results regressing attrition on growth, by fine-grained categories of predicted Hispanic surnames (Study 3)

The fourth and last check on the marginal Latino hypothesis uses the presence of Hispanic

coethnics in the household as the moderator of interest (lower levels offers another proxy for marginal Latino status). We count both number of Hispanic (2016) identifiers and total people living at each household — i.e., unique home address — in 2016 and then merge this onto the main dataset. We perform a few modifications to make the data more meaningful for this test, subsetting to multi-person households (addresses) but ones with no more than 10 members (excluding colleges, fraternities, etc.). Given weaker data variation than Hispanic name predictions, for example, we transform this moderator into only a three-category group: individuals whose other household member/s are 0% Hispanic, 100% Hispanic, or are a mix of Hispanics and non-Hispanics. We again apply the most stringent OLS test, with results shown in Figure 9. This moderator test produces another nonlinear relationship like before: coethnic growth constrains attrition most strongly for individuals more in the ambiguous middle, namely those living with an ethnic mix of people at home.

Study 4

Data and Methods

To effectively advance our understanding of how group size intersects with politics, we plan to experimentally test whether Latinos primed with the realities of the U.S. Hispanic population exhibit greater group consciousness, an elevated sense of political efficacy, or alternatively, a readiness to adopt views typically held by non-Hispanic whites (e.g., racialized policy positions). The full text of the survey instrument is provided in the Appendix. We leverage the controlled setting to more directly get at the political mechanisms that may underlie ethnic attrition. A survey experiment affords us the heightened level of precision required to analyze the fluid nature of self-concept, which is formed across a spectrum of domains that encapsulate a wide array of personal experiences. That is, the ability to explain a loss of ethnic identification is improved by the specificity in measurements of that identity. The link connecting fluctuations in Hispanic affiliation to variations in group composition

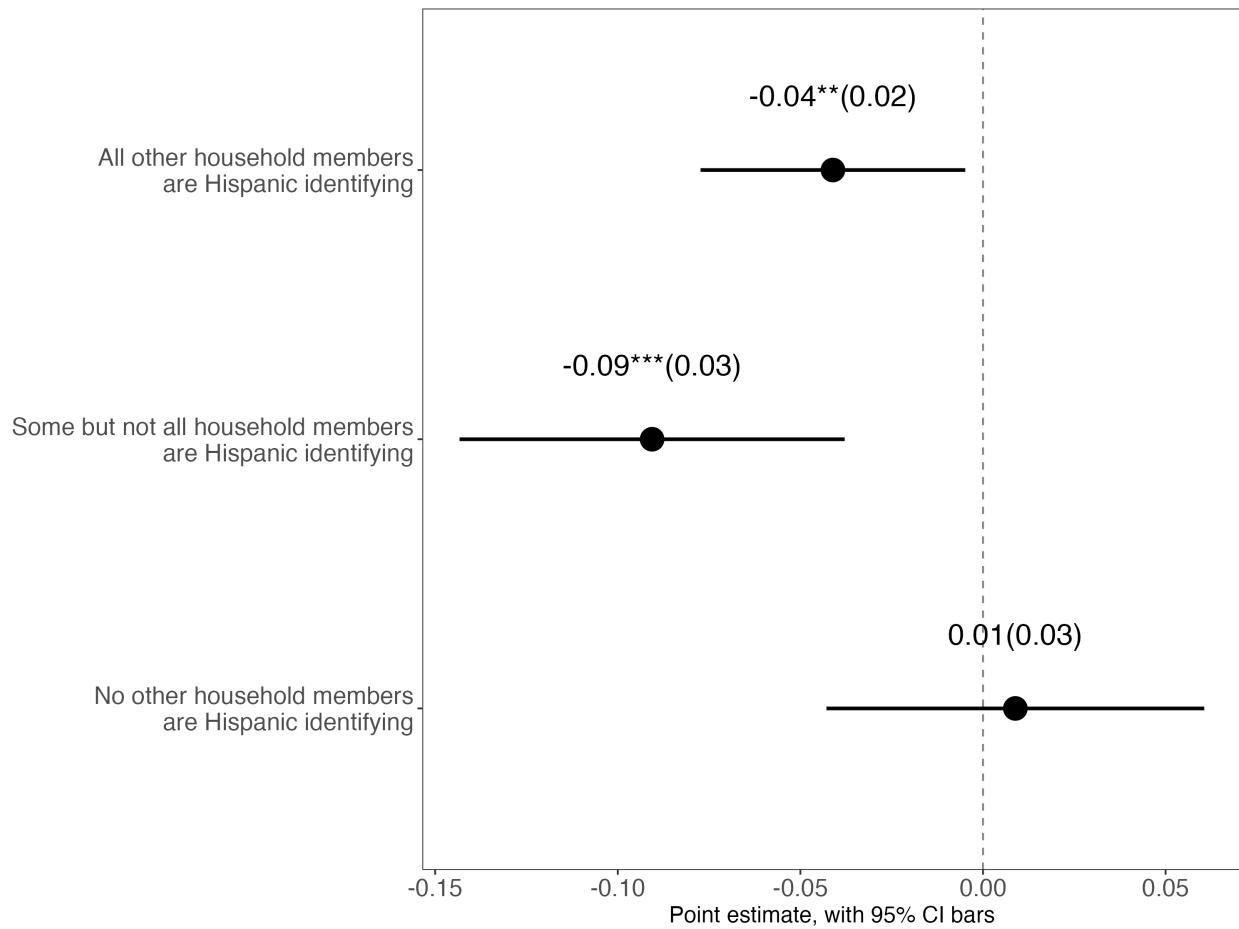


Figure 9: Full model results regressing attrition on growth, by household ethnic composition (Study 3)

is made clearer by a more detailed account of the extent to which demographic shifts are perceived. Distinguishing patterns in this observed sensitivity to the growth of the group is facilitated by identifying which types of Latino members find this a motivating logic. Crucially, this additional component allows us to delve into the intentionality behind distancing from ancestral ties by grounding these decisions in their political context.

Sample

Our experiment breaks new ground in the study of intergroup relations by not only tracing their downstream political effects but also in our choice of focal population. Prior work has primarily explored the impact of changes in the share of one group (e.g., Latinos) from the perspective of other groups (e.g., whites, Asians, Blacks). In contrast, our analytical approach prioritizes the reactions of Latinos themselves. We rely on a sample of Latinos recruited through an online panel provider. We considered the potential for selection bias arising from examining the maintenance and relevance of an ethnic label, while simultaneously using that same label to recruit participants. However, consistent with the notion that group classification does not automatically imply identification—meaning that not all individuals identified as Latino necessarily share a cultural attachment—this concern is significantly mitigated. Thus, while recruitment screeners help assemble the sample, our instrument specifically gets at the affective dimensions of that identity. We also overcome a similar methodological challenge to the representativeness of the Latino sample when the survey being administered solely in English, given that our population of interest (i.e., later generation, “marginal” Latinos) is likely to be English-dominant.

We require a minimum sample of 1050 respondents based on an a-priori power analysis conducted using the Declare Design framework. We estimate the treatment effect in this experimental design using a difference-in-means approach. Based on limited but nevertheless germane initial work on the influence of Hispanic population growth among Hispanic respondents (Craig and Richeson 2017), the average standard error on the difference between

information conditions is 0.075 standard units. Multiplied by 2.8, this indicates that a two-sided test with $\alpha = 0.05$ has 80% power to detect effects of at least 0.21 standard units (Gelman and Hill 2006). Drawing analogies from the effect sizes found among out-groups (e.g., Asian, Black respondents), which range from 0.04 to 0.11, we approximate a generally positive average treatment effect of a larger magnitude for Latino subjects and are well equipped to make these comparisons. We also broaden our scope of inquiry to include the possible conditional basis of the relationship between group attachment and significant group transformations, probing which types of members are predisposed to abandon their Hispanic heritage (e.g., nativity and generational status, geography). To this end, we adjust our power calculations to accommodate such specifications—including covariate models controlling for formative characteristics such as partisanship and education—revealing minimum detectable effects of 0.30 and 0.42 standard units for two- and three-way interactions, respectively. Our sample will consist of all consenting participants, with a primary analytical focus on those deemed “attentive.” Attentiveness is defined by metrics designed to assess engagement throughout the survey. Although we expect most participants to meet these criteria with few drop-offs, analyses will be presented for both the entire sample and the attentive subset.

Design

We will conduct a randomized, two-arm controlled trial with a sample of U.S. Latinos, comparing opinion on key outcomes between those presented with information about the size of the Hispanic population in the country (treatment) and those without it (baseline). The rationale for this choice of estimand is two-fold. First, any meaningful reaction to information about a group evinces the salience, relevance, or accessibility of that group membership itself. We would expect commentary of Latinos to be most consequential and identity affirming to those with the strongest cultural attachment. Second, and more importantly, the relative effect of that count gives clarity the political meaning that we argue is specifically ascribed to demographic changes: mere data about the population does not inherently carry

political implications, thus our intervention should make no difference. However, if treated respondents' attitudes indeed markedly differ from baseline, it would suggest that a group's size has political stakes.

To demarcate the politics of this response, we also take stock of how participants' nativity might be driving or explain in part the patterning of differences. The distinction between being U.S.- versus foreign-born can be a constitutive factor in the social, economic, political, and cultural choices available to an individual. How they come to appreciate and relate to their ethnic identity may also vary. For instance, information highlighting the increasing number of Latinos in the U.S. might feel alienating to those Latinos with a weaker ethnic attachment, who may perceive the newcomers as unrepresentative of the Latinos they identify with, thus increasing their distancing from the group. Conversely, for someone in a predominantly Anglo region, the growth of the Latino population might appear as a noteworthy and welcomed phenomenon, while it could seem routine and unremarkable to someone in an area with a dense concentration of Latinos residents.

In defining our benchmarks, we explored several perspectives to establish both who and what those points of comparison should be. Initially, we considered a baseline scenario in which no information is provided, a typical approach for novel research questions. Instead of a pure control, we also contemplated an alternative design contrasting two informational environments that vary in content but share a common theme of population change. One that describes population composition and the other a population habit (recycling). Both interventions are politically neutral and emphasize the concept of growth. Once finalized, we will secure IRB approval and pre-register our analysis plan.

Procedure

In our study, we also accounted for potential biases stemming from the experimental protocol itself—peripheral yet significant logistical aspects of our experiments that could inadvertently influence outcomes. For instance, the sequencing and timing of questions early in the survey

about identity might prime respondents to think about that identity prior to the experimental manipulation. Similarly, asking participants about their perceptions of the size of the Latino population at the outset could prime them to give undue cognitive weight to this factor for the remainder of the session.

Despite these concerns, recent meta-analyses have suggested that the impact of such biases is generally limited and minimal. Nonetheless, to ensure the validity of our findings, we include a manipulation check immediately after the intervention and before respondents proceed to the outcome measures. We define the treated group as those participants who recognize that the main message of the infographic concerns the growth of the Hispanic population in the US. The exact wording used for this purpose is detailed in the Appendix.

Materials

Stimuli

The stimuli presented information on the shifting and increasing share of one demographic group in the U.S. as per procedures in germane studies. Our challenge was to strike the right balance between providing data that was seen as credible and accurate without overstating or embellishing it with visuals. To keep these contextual features constant and maximize external validity, we styled our intervention to look like an infographic prepared by official government agencies. As illustrated by Figure 10, the image features various examples and units of this demographic growth. A potential complication with this approach is that it might bumble the treatment, blurring the weight of each metric.

Prior Opinion

The core of our analysis and interpretations is the evolving nature of Latino identity. Whereas ethnic attrition describes a gradual process that occurs over time, we do not aim to replicate this dynamic in our current study. Instead, our focus is on the conditions under which our decisions and opinions are anchored to our identity. We regard differences in response as



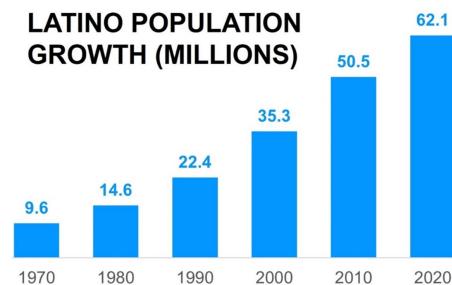
6X

MORE LATINOS IN
THE U.S. TODAY
THAN 50 YRS AGO

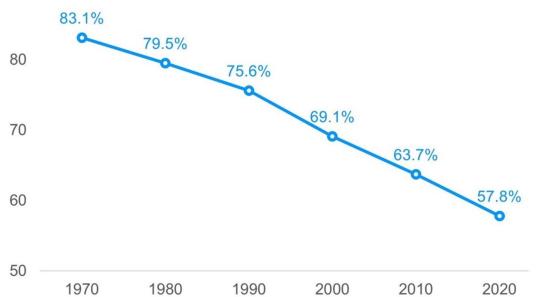


1 IN 4 CHILDREN
26% (18.8 MILLION)
IN U.S. ARE LATINO

LATINO POPULATION GROWTH (MILLIONS)



DECREASE IN THE SHARE OF THE WHITE POPULATION BETWEEN 1970 - 2020



25%
DROP
IN THE
WHITE
POPULATION

Figure 10: Data based on real-world estimates from the U.S. Census

indicators of the diverse levels of connection to or value assigned to that ethnic label. Making sense of this attitudinal variation is particularly complex. As previously outlined, identity is deeply rooted in our life experiences, complicating the task of pinpointing clear markers that capture those individual histories. Informed by interdisciplinary work on immigrant acculturation and supported by initial findings from Studies 1-3, we examine three key elements that may constrain why or how someone comes to identify as Latino—facets which

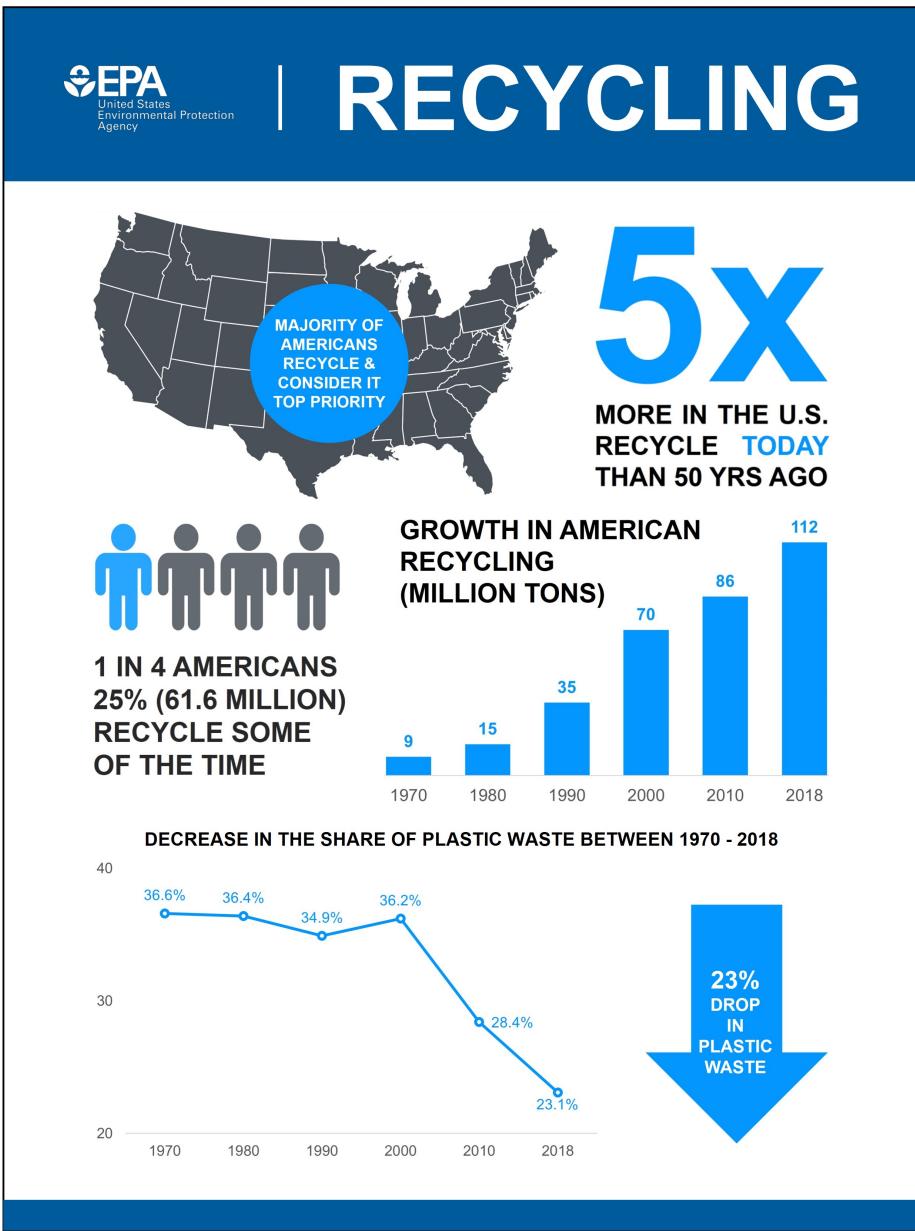


Figure 11: Data based on real-world estimates from the Environmental Protection Agency (EPA)

calibrate if that identity aligns with one's self-perception, and the tangible benefits it offers. These moderators are not comprehensive, and we acknowledge that other aspects may also play a role.

Generational Status

The most intuitive of the three that we investigate as critical to the formation and

preservation of *Latinidad* is generational status—whether a respondent is born in or outside the U.S. The perceived need, importance, and utility of ethnic identity are fundamentally context dependent, and it is these differences in referent conditions that have consistently accounted for diminished ties to that identity. The changes in cultural norms and values with each generation removed from the original arriving cohort is well documented in research on immigrant families. That this denotes reduced identification is based on the notion that valued traditions would be maintained. Loss of ethnic distinction is often cited to explain trends such as interracial marriage and declining language proficiencies.

This path is not inevitable. Recent studies have highlighted the continuous flow of immigrants from Latin America as a factor in maintaining cultural affinities among U.S. Hispanics, a phenomenon not seen in earlier waves of Italian, Irish entrants. Building on this insight, we ask respondents about their generational status, coding them as either U.S.-born or foreign-born based on the following questions:

(**gen_status1**) Were you born in the United States, on the Island of Puerto Rico, or another country?

(**gen_status2**) When did you first arrive to live in the U.S.?

(**gen_status3**) Which of these statements describes you?

[I am an immigrant to the United States and am a naturalized citizen]

[I am an immigrant to the United States but not a citizen]

[I was born in the United States but at least one of my parents was not]

[I was born in the United States but both my parents were not]

[My parents and I were born in the United States but at least one of my grandparents were not]

[My parents, grandparents, and I were all born in the United States]

We expect these different perspectives to result in varied interpretations of demographic information. Specifically, we anticipate that U.S.-born individuals will view ethnic identity differently compared to those born outside the U.S., reflecting their distinct experiences and contexts.

Religion

To capture the nuanced context of ethnic identity, we delve into the networks that reinforce Latino cultural markers, particularly focusing on community building through religion. Catholicism is the predominant religion throughout Latin America, and the Catholic Church's organizational reach means it has seamlessly extended its influence to the U.S., where the majority of Hispanics are Catholic. Places of worship in immigrant neighborhoods provide cultural continuity, spaces to gather with kin (often in their native language). The fusion of religious and cultural practices within these spaces means that engaging in one reinforces the other. Historically, this blend has enhanced civic interest among Latinos, with church networks offering essential support and mobilization. Recognizing such a role, we pose the following question:

(religion_1) What is your present religion, if any?

(religion_2) Were you part of a religious community growing up?

We then categorize the responses into Catholic and Non-Catholic groups. We anticipate that those with a Catholic affiliation will have distinct viewpoints on immigration and the presence of Latinos in the U.S., an extension of that religious community's influence.

“Street Race”

To further interrogate how ethnic identity shapes daily life, we examine its utility based on recognition and validation by others. Prototypical traits often define group membership, and these boundaries are continually reinforced through social comparison—both in how individuals perceive themselves and how they are perceived by others. When our self-concept aligns with others' perceptions, our identity feels meaningful. If there's a mismatch, however, the identity can become less salient. We assess this alignment or discrepancy by asking respondents to reflect on their social interactions:

(**street_race**) How would most people describe you, if, for example, they walked past you on the street? Would they say you are ...

From their responses, we categorize subjects into two groups: those who feel recognized as Hispanic and those who do not. We anticipate that individuals who perceive their ethnicity as frequently questioned or scrutinized will exhibit a weaker attachment to their ethnic identity.

Dependent Variables

Our outcome measures capture variations in the prominence of ethnic identity. Each metric situates identity within a political landscape, allowing us to infer the frame of mind that driving behavior.

Identity centrality

We utilize a framework that evaluates subjective ratings of how much individuals feel their ethnic identity is central to self-concept. Importance assigned to changes in group size reflects the personal investment in that identity. The following questions allow us to empirically map this frame of mind:

(**centrality_1**) Do you agree or disagree with this statement: “In general, being Hispanic or Latino is NOT an important part of my self-image.” *[4-Point Likert Scale: Agree - Disagree]*

(**centrality_2**) Do you agree or disagree with this statement: “If I were to describe myself to someone, one of the first things that I would NOT say is that I’m Latino or Hispanic.” *[4-Point Likert Scale: Agree - Disagree]*

(**centrality_3**) Do you agree or disagree with this statement: “Overall, being Latino or Hispanic has very little to do with how I feel about myself.” *[4-Point Likert Scale: Agree - Disagree]*

(**linked_fate**) “How much do you think that what happens generally to Latinos in this country will affect what happens in your life?” *[4-Point Likert Scale: A lot - Not at all]*

We will construct an index based on centrality_1, centrality_2, and centrality_3, ensuring its conceptual coherence. Responses will be standardized so that lower identity centrality is coded as 0 and 1 as higher.

Pro-Group Politics

Exploring this cognitive and affective shift in more concrete terms, we zero in on how ethnic identity translates into political decisions and choices. Ethnic attachment in this conceptualization is often gauged by support for pro-group politics. We therefore ask respondents:

(progroup_1) Do you agree or disagree with this statement: “Latinos should always vote for Latino candidates when they run.” *[4-Point Likert Scale: Agree - Disagree]*

(progroup_2) Do you agree or disagree with this statement: “Public officials in the U.S. should make all government materials available in both English and Spanish.” *[4-Point Likert Scale: Agree - Disagree]*

(progroup_3) Do you agree or disagree with this statement: “Latinos should shop in Latino stores whenever possible.” *[4-Point Likert Scale: Agree - Disagree]*

As before, we will generate an aggregate response, standardized with 0 indicating low support and 1 indicating high pro-group support.

Ancillary Analyses

Immigration Policy

Because our primary interest lies in downstream political effects, we are considering incorporating more direct policy measures. Previous work has often interpreted positions on U.S. immigration policy among Latinos as proxies for the strength of Latino identity. We would propose the following questions:

(immigrationpolicy_1) Do you agree or disagree with this statement: “Illegal immigrants are a drain on American society.” *[4-Point Likert Scale: Agree - Disagree]*

(immigrationpolicy_2) Do you agree or disagree with this statement: “Undocumented or illegal immigrants living in the U.S. for an extended period of time should be allowed to have an opportunity to eventually become U.S. citizens.” *[4-Point Likert Scale: Agree - Disagree]*

(immigrationpolicy_3) Do you agree or disagree with this statement: “Local police should be able to question anyone about their immigration status, particularly if they think they may be in the country illegally.” *[4-Point Likert Scale: Agree - Disagree]*

Responses to this question set will be summarized into a metric, recoded so that 0 signifies a position typically associated with white conservatives and 1 indicates a more liberal position.

Discussion

To conclude, this project argues that coethnic population growth can mitigate attrition in Hispanic/Latino identity, in a way that has ramifications for political behavior, and does so especially for individuals most prone to attriting (the “marginal Latinos”). Three distinct but complementary observational tests generally provide support for our theoretical expectations, and we lay out plans for a final survey experimental test to buttress causal claims, shed more light on mechanisms, and map our dynamic of interest closer onto political implications. While our core argument is not entirely novel (e.g., Jiménez, 2008), we fill a gap in quantitatively testing causes of ethnic identity attrition and maintenance, our empirics theoretically flesh out how and why coethnic presence feeds into identity choice, and we draw out political consequences as well.

We hope to receive feedback on any parts of this project, but are particularly interested in the following areas:

1. Cohesiveness of our theoretical framework, its contribution relative to existing literature,

and how well it connects to empirics across the studies.

2. Other things with the observational analyses worth exploring or different directions here.
3. Rounding out the survey experimental approach, which is the least developed part of the project.

References

- Agadjanian, Alexander (2022). “How Many Americans Change Their Racial Identification over Time?” In: *Socius* 8, p. 23780231221098547.
- Agadjanian, Alexander and Dean Lacy (2021). “Changing votes, changing identities? Racial fluidity and vote switching in the 2012–2016 US Presidential Elections”. In: *Public Opinion Quarterly* 85.3, pp. 737–752.
- Alba, Richard (2020). *The great demographic illusion: Majority, minority, and the expanding American mainstream*. Princeton University Press.
- Alba, Richard and Tariqul Islam (2009). “The case of the disappearing Mexican Americans: An ethnic-identity mystery”. In: *Population Research and Policy Review* 28, pp. 109–121.
- Alvira-Hammond, Marta (2019). “Hispanic women are helping drive the recent decline in the US fertility rate”. In: *National Research Center on Hispanic Children and Families*, Bethesda, MD.
- Ansolabehere, Stephen and Eitan Hersh (2012). “Validation: What big data reveal about survey misreporting and the real electorate”. In: *Political Analysis* 20.4, pp. 437–459.
- Antman, Francisca, Brian Duncan, and Stephen J Trejo (2016). “Ethnic attrition and the observed health of later-generation Mexican Americans”. In: *American Economic Review* 106.5, pp. 467–471.
- Antman, Francisca M, Brian Duncan, and Stephen J Trejo (2020). “Ethnic attrition, assimilation, and the measured health outcomes of Mexican Americans”. In: *Journal of Population Economics* 33, pp. 1499–1522.
- Barreto, Matt A, Tyler Reny, and Bryan Wilcox-Archuleta (2017). “Survey methodology and the latina/o vote: Why a bilingual, bicultural, latino-centered approach matters”. In: *Aztlan: A Journal of Chicano Studies* 42.2, pp. 211–227.
- Bedolla, Lisa Garcia (2003). “The identity paradox: Latino language, politics and selective dissociation”. In: *Latino Studies* 1.2, pp. 264–283.

- Bedolla, Lisa Garcia (2005). *Fluid borders: Latino power, identity, and politics in Los Angeles*. Univ of California Press.
- Calvillo, Jonathan E and Stanley R Bailey (2015). “Latino religious affiliation and ethnic identity”. In: *Journal for the Scientific Study of Religion* 54.1, pp. 57–78.
- Duncan, Brian and Stephen J Trejo (2010). “Who remains Mexican? Selective ethnic attrition and the intergenerational progress of Mexican Americans”. In: *Latinos and the economy: Integration and impact in schools, labor markets, and beyond*. Springer, pp. 285–320.
- (2011a). “Intermarriage and the intergenerational transmission of ethnic identity and human capital for Mexican Americans”. In: *Journal of labor economics* 29.2, pp. 195–227.
- (2011b). “Tracking intergenerational progress for immigrant groups: The problem of ethnic attrition”. In: *American Economic Review* 101.3, pp. 603–608.
- (2018). “Identifying the later-generation descendants of US immigrants: Issues arising from selective ethnic attrition”. In: *The ANNALS of the American Academy of Political and Social Science* 677.1, pp. 131–138.
- Egan, Patrick J (2020). “Identity as dependent variable: How Americans shift their identities to align with their politics”. In: *American Journal of Political Science* 64.3, pp. 699–716.
- Emeka, Amon (2019). ““Just black” or not “just black?” ethnic attrition in the Nigerian-American second generation”. In: *Ethnic and Racial Studies* 42.2, pp. 272–290.
- Emeka, Amon and Jody Agius Vallejo (2011). “Non-Hispanics with Latin American ancestry: Assimilation, race, and identity among Latin American descendants in the US”. In: *Social Science Research* 40.6, pp. 1547–1563.
- Fernández, Leticia et al. (2018). *Factors that influence change in Hispanic identification: Evidence from linked decennial census and American Community Survey data*. Tech. rep.
- Fraga, Bernard L (2018). *The turnout gap: Race, ethnicity, and political inequality in a diversifying America*. Cambridge University Press.

- Fraga, Bernard L, Yamil R Velez, and Emily A West (2024). “Reversion to the Mean, or Their Version of the Dream? Latino Voting in an Age of Populism”. In: *American Political Science Review*, pp. 1–9.
- Goldman, Liran and Michael A Hogg (2016). “Going to extremes for one’s group: The role of prototypicality and group acceptance”. In: *Journal of Applied Social Psychology* 46.9, pp. 544–553.
- Hadah, Hussain (2023). “The Effect of Racial and Ethnic Attitudes on Hispanic Identity in the US”. In: *Available at SSRN 4566232*.
- Haner, Joanne and Mark Hugo Lopez (2023). “8 facts about recent Latino immigrants to the US”. In.
- Hill, Seth J, Daniel J Hopkins, and Gregory A Huber (2019). “Local demographic changes and US presidential voting, 2012 to 2016”. In: *Proceedings of the National Academy of Sciences* 116.50, pp. 25023–25028.
- Hopkins, Daniel J (2011). “National debates, local responses: The origins of local concern about immigration in Britain and the United States”. In: *British Journal of Political Science* 41.3, pp. 499–524.
- Ichino, Nahomi and Noah L Nathan (2013). “Crossing the line: Local ethnic geography and voting in Ghana”. In: *American Political Science Review* 107.2, pp. 344–361.
- Imai, Kosuke and Kabir Khanna (2016). “Improving ecological inference by predicting individual ethnicity from voter registration records”. In: *Political Analysis* 24.2, pp. 263–272.
- Jiménez, Tomás R (2008). “Mexican immigrant replenishment and the continuing significance of ethnicity and race”. In: *American journal of sociology* 113.6, pp. 1527–1567.
- Jiménez, Tomás Roberto (2010). *Replenished ethnicity: Mexican Americans, immigration, and identity*. Univ of California Press.

- Jones, Bradford, Kristina Flores Victor, and David Vannette (2019). “Alien citizens and the canonical immigrant: do Stigmatized attributes affect Latina/o judgment about discrimination?” In: *American Journal of Political Science* 63.4, pp. 740–757.
- Junn, Jane, Tali Mendelberg, and Erica Czaja (2012). “Race and the Group Bases of Public Opinion: Jane Junn, Tali Mendelberg, and Erica Czaja”. In: *New directions in public opinion*. Routledge, pp. 131–150.
- Kaiser, Caspar and Andrew J Oswald (2022). “The scientific value of numerical measures of human feelings”. In: *Proceedings of the National Academy of Sciences* 119.42, e2210412119.
- Kinder, Donald R and Nicholas Winter (2001). “Exploring the racial divide: Blacks, whites, and opinion on national policy”. In: *American Journal of Political Science*, pp. 439–456.
- Livingstone, Gretchen and Anna Brown (2017). *Intermarriage in the US: 50 years after loving V. Virginia*. Pew Research Center Washington, DC.
- Lopez, Mark Hugo, Ana Gonzalez-Barrera, and Gustavo López (2017). “Hispanic identity fades across generations as immigrant connections fall away”. In.
- Manrique, Ruy and Ted Mouw (2023). *Shift or replenishment? Reassessing the prospect of stable Spanish bilingualism across contexts of ethnic change*. Tech. rep.
- Mora, G Cristina (2014). *Making Hispanics: How activists, bureaucrats, and media constructed a new American*. University of Chicago Press.
- Morgan, LaKendra Beard et al. (2024). *Black Race Matters in the Latino Population*.
- Mumm, Jesse and Carolina Sternberg (2023). “Mapping racial capital: Gentrification, race and value in three Chicago neighborhoods”. In: *Urban Affairs Review* 59.3, pp. 793–831.
- Newman, Benjamin, Marcel F Roman, and Tyler Reny (2023). *The Immigration-to-Reproduction Shift: Latino Population Growth and White Support for Legal Abortion*.
- Newman, Benjamin J (2013). “Acculturating contexts and Anglo opposition to immigration in the United States”. In: *American Journal of Political Science* 57.2, pp. 374–390.
- Newman, Benjamin J and Yamil Velez (2014). “Group size versus change? Assessing Americans’ perception of local immigration”. In: *Political Research Quarterly* 67.2, pp. 293–303.

- Pepinsky, Thomas B, Sara Wallace Goodman, and Conrad Ziller (2024). “Modeling spatial heterogeneity and historical persistence: Nazi concentration camps and contemporary intolerance”. In: *American Political Science Review* 118.1, pp. 519–528.
- Robertson, Crystal and Marcel F Roman (2023). “The Wages of Latinidad: How Immigration Enforcement Mitigates Anti-Black Assimilation”. In: *Political Behavior*, pp. 1–23.
- Roman, Marcel F (2023). “Living in the Shadow of Deportation: How Immigration Enforcement Forestalls Political Assimilation”. In: *Political Research Quarterly* 76.3, pp. 1460–1474.
- Tajfel, Henri et al. (1979). “An integrative theory of intergroup conflict”. In: *Organizational identity: A reader* 56.65, pp. 9780203505984–16.
- Telles, Edward and Christina A Sue (2019). *Durable ethnicity: Mexican Americans and the ethnic core*. Oxford University Press.
- Vitriol, Joseph A, Erik Gahner Larsen, and Steven G Ludeke (2019). “The generalizability of personality effects in politics”. In: *European Journal of Personality* 33.6, pp. 631–641.
- White, Ismail K and Chryl N Laird (2020). *Steadfast democrats: How social forces shape Black political behavior*. Princeton University Press.
- Wilcox-Archuleta, Bryan (2018). “Local origins: Context, group identity, and politics of place”. In: *Political research quarterly* 71.4, pp. 960–974.

Appendices

Contents

A Study 1	1
A.1 Descriptive Statistics	1
A.1.1 What do Latinos Attrit To?	1
A.1.2 Attrition and Politics	2
A.1.3 Descriptive Relationship Between Δ % Latino and Ethnic Attrition .	3
A.2 Estimation Strategies	4
A.3 Alternative Independent Variables	6
A.3.1 Alternative Latinx Growth Variables	6
A.3.2 Growth in Other Groups	7
A.3.3 Group Growth x Group ID Tests	7
A.3.4 County-Level Metrics	8
A.4 Residential Tenure Sample Splits	9
A.5 Adjusting for Geographic FEs	9

A Study 1

A.1 Descriptive Statistics

A.1.1 What do Latinos Attrit To?

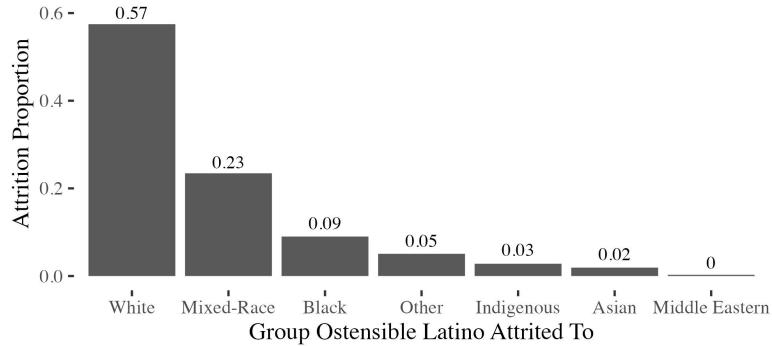


Figure A1: Latinos primarily identify as white or mixed-race if they attrit.

A.1.2 Attrition and Politics

Table A1: Ethnic Attritors are More Politically Conservative than Non-Attritors

Outcome	Non-Attritor	Attritor	DIM T-Value
Pr(Democratic ID)	0.54	0.44	20.49
Pr(Independent ID)	0.16	0.19	6.56
Pr(Republican ID)	0.22	0.30	18.02
Pr(Voted Dem '12)	0.69	0.56	6.68
Pr(Voted Dem '16)	0.67	0.56	6.55
Pr(Voted Dem '20)	0.65	0.57	4.07
Pr(Liberal)	0.26	0.26	0.92
Pr(Conservative)	0.24	0.28	10.78
Pr(Legalize Undocumented)	0.70	0.59	20.38
Pr(Oppose Border)	0.59	0.50	15.22

Estimates use stacked CES data from 2010-2022 with the exception of the vote choice outcomes, which use CES data from the year of the respective election.

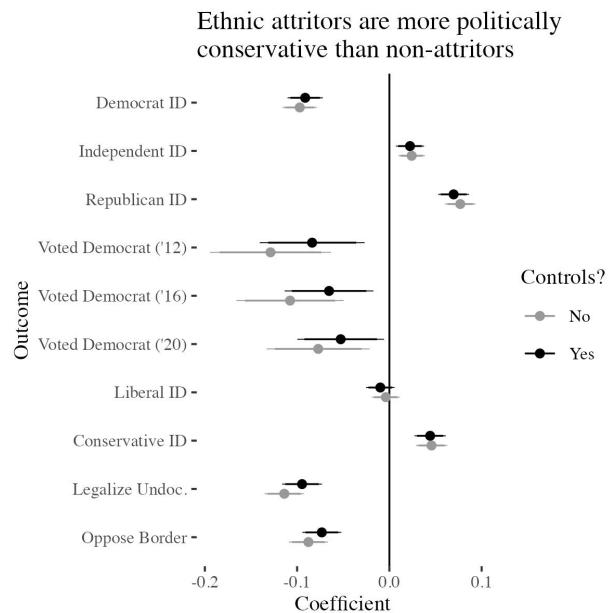


Figure A2: Ethnic attritors are more politically conservative than non-attritors.
95% CIs displayed from robust SEs.

A.1.3 Descriptive Relationship Between $\Delta \%$ Latino and Ethnic Attrition

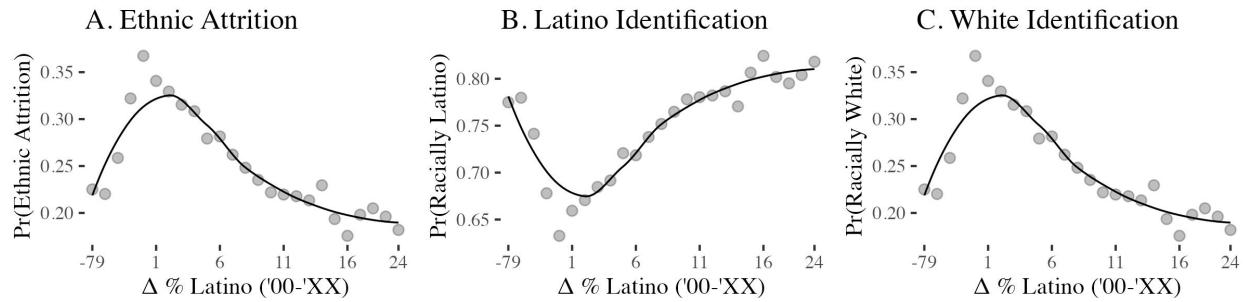


Figure A3: Descriptive statistics characterizing association between Latinx population change and ethnic attrition. X-axis is $\Delta \% \text{ Latino}$ at the zipcode-level (50 quantile bins), Y-axis is *ethnic attrition*, Latino identification, and white identification for Panels A-C respectively. Loess fit displayed on each plot.

A.2 Estimation Strategies

To test **H1**, we use the following linear model:

$$(1) \quad Attrition_i = \alpha + \beta_1 \Delta \% Latino_{iz} + \sum_{k=1}^k \beta_{k+1} X_{izs}^{k+1} + \varepsilon_z$$

Where $Attrition_i$ is the attrition indicator for respondent i , $\Delta \% Latino_{iz}$ is the change in the proportion of the zipcode (z) population that is Latino for respondent i between 2000 and the year prior to the interview year. $\sum_{k=1}^k \beta_{k+1} X_{izs}^k$ are k covariates at the individual (i), zipcode (z), and state-level (s). ε_z are robust errors clustered at the zipcode-level. If H1 is supported, we expect β_1 to be *positive*.

To test **H2**, we use the following linear models:

$$(2) \quad Attrition_i = \alpha + \beta_1 (\Delta \% Latino_{iz}^{20XX-2000} \times \% Latino_{iz}^{2000}) + \beta_2 \Delta \% Latino_{iz}^{20XX-2000} \\ + \beta_3 \% Latino_{iz}^{2000} + \sum_{k=1}^k \beta_{k+3} X_{izs}^{k+3} + \varepsilon_z$$

$$(3) \quad Attrition_i = \alpha + \beta_1 (\Delta \% Latino_{iz}^{20XX-2000} \times ImmigrantCitizen_i)$$

$$+ \beta_2 (\Delta \% Latino_{iz}^{20XX-2000} \times SecondGen_i) \\ + \beta_3 (\Delta \% Latino_{iz}^{20XX-2000} \times ThirdGen_i) \\ + \beta_4 (\Delta \% Latino_{iz}^{20XX-2000} \times FourthGen_i) \\ + \beta_5 \Delta \% Latino_{iz}^{20XX-2000} + \beta_6 ImmigrantCitizen_i \\ + \beta_7 SecondGen_i + \beta_8 ThirdGen_i \\ + \beta_9 FourthGen_i + \sum_{k=1}^k \beta_{k+9} X_{izs}^{k+9} + \varepsilon_z$$

$$(4) \quad Attrition_i = \alpha + \beta_1 (\Delta \% Latino_{iz}^{20XX-2000} \times Non-Catholic_i) + \beta_2 \Delta \% Latino_{iz}^{20XX-2000} \\ + \beta_3 \% Non-Catholic_i + \sum_{k=1}^k \beta_{k+3} X_{izs}^{k+3} + \varepsilon_z$$

For Model (2), $\% Latino_{iz}^{2000}$ is the proportion of the zipcode (z) population for respondent

(i) that is Latino at 2000. For Model (3), $ImmigrantCitizen_i$, $SecondGen_i$, $ThirdGen_i$, and $FourthGen_i$ are binary indicators for whether a respondent self-reports they are an immigrant citizen, second-generation, third-generation or fourth-generation. For Model (4), $Non-Catholic_i$ is a binary indicator for whether a respondent self-reports they are not Catholic. If **H2** is supported, we expect β_1 to be positive for Models (2-4) respectively.

To test **H3**, we use the following linear models:

$$(5) \ Attrition_i = \alpha + \beta_1 \Delta \% SpanishLatino_{iz} + \beta_2 \Delta \% EnglishLatino_{iz} + \sum_{k=1}^k \beta_{k+2} X_{izs}^{k+2} + \varepsilon_z$$

$$(6) \ Attrition_i = \alpha + \beta_1 \Delta \% ImmigrantLatino_{ic} + \beta_2 \Delta \% NonImmigrantLatino_{ic} \\ + \sum_{k=1}^k \beta_{k+3} X_{izs}^{k+3} + \varepsilon_c$$

Where $\Delta \% SpanishLatino_{iz}$ is the change in the proportion of the zipcode (z) population that is Latino and Spanish-speaking that respondent i lives in between 2000 and the year prior to being surveyed. $\Delta \% EnglishLatino_{iz}$ is the change in the proportion of the zipcode (z) population that is Latino and English-speaking that respondent i lives in between 2000 and the year prior to being surveyed. $\Delta \% ImmigrantLatino_{ic}$ is the change in the proportion of the county population that is Latino and immigrant that respondent i lives in between 2000 and the year prior to being surveyed. $\Delta \% NonImmigrantLatino_{ic}$ is the change in the proportion of the county population (c) that is Latino and non-immigrant that respondent i lives in between 2000 and the year prior to being surveyed. For model (6), ε_c are robust errors clustered at the county-level.

A.3 Alternative Independent Variables

A.3.1 Alternative Latinx Growth Variables

Table A2: Association Between Alternative Latinx Growth Independent Variables and *Ethnic Attrition*

	Ethnic Attrition					
	(1)	(2)	(3)	(4)	(5)	(6)
> Median Δ % Latino	-0.04*** (0.00)					
2nd Tercile Δ % Latino		-0.02*** (0.01)				
3rd Tercile Δ % Latino			-0.06*** (0.01)			
2nd Quartile Δ % Latino				-0.02** (0.01)		
3rd Quartile Δ % Latino					-0.04*** (0.01)	
4th Quartile Δ % Latino					-0.07*** (0.01)	
% Latino ('00)	-0.25*** (0.01)	-0.25*** (0.01)	-0.25*** (0.01)	-0.25*** (0.01)	-0.25*** (0.01)	-0.24*** (0.01)
Percent Change % Latino ('00-'XX)					-0.29*** (0.05)	
Δ Latino Pop.						-0.24*** (0.06)
Δ Total Pop.						-0.01 (0.06)
Controls?	Y	Y	Y	Y	Y	Y
R ²	0.10	0.11	0.11	0.10	0.10	0.10
Num. obs.	61134	61134	61134	61134	61134	61134
N Clusters	10479	10479	10479	10479	10479	10479

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

A.3.2 Growth in Other Groups

Table A3: Association Between Growth in Other Ethno-Racial Groups and *Ethnic Attrition*

	Ethnic Attrition			
	(1)	(2)	(3)	(4)
Δ % White ('00-'XX)	0.32*** (0.04)			0.08 (0.05)
Δ % Black ('00-'XX)		-0.03 (0.07)		
Δ % Asian ('00-'XX)			-0.04 (0.05)	
Δ % Latino ('00-'XX)				-0.43*** (0.07)
% Latino ('00)	-0.26*** (0.01)	-0.25*** (0.01)	-0.25*** (0.01)	-0.26*** (0.01)
Controls?	Y	Y	Y	Y
R ²	0.10	0.10	0.10	0.11
N	61134	61134	61134	61134
N Clusters	10479	10479	10479	10479

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

A.3.3 Group Growth x Group ID Tests

Table A4: Association Between Alternative Group Growth Independent Variables and Ethno-Racial Identity among Ostensible Latinxs

	Latino ID (1)	White ID (2)	Black ID (3)	Asian ID (4)
Δ % Latino ('00-'XX)	0.50*** (0.05)			
Δ % White ('00-'XX)		0.25*** (0.03)		
Δ % Black ('00-'XX)			0.14*** (0.04)	
Δ % Asian ('00-'XX)				0.02 (0.01)
Controls?	Y	Y	Y	Y
R ²	0.11	0.06	0.04	0.01
Num. obs.	61134	61134	61134	61134
N Clusters	10479	10479	10479	10479

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

A.3.4 County-Level Metrics

Table A5: Re-analyzing main results using county-level measures of Latino population growth

Ethnic Attrition (1)	
Δ % Latino ('00-'XX)	-0.36*** (0.05)
% Latino ('00)	-0.23*** (0.03)
Controls?	Y
R ²	0.10
Num. obs.	61026
N Clusters	1964

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

A.4 Residential Tenure Sample Splits

Table A6: Association Between Latino Growth and *Ethnic Attrition*, Sample Splits by Residential Tenure

	Ethnic Attrition		
	(1)	(2)	(3)
$\Delta \%$ Latino ('00-'XX)	-0.62*** (0.06)	-0.48*** (0.06)	-0.41*** (0.08)
% Latino ('00)	-0.24*** (0.02)	-0.26*** (0.02)	-0.29*** (0.02)
Tenure Sample Controls?	Below Median Y	Above Median Y	Maximum Tenure Y
R ²	0.08	0.09	0.09
Num. obs.	30579	30538	16851
N Clusters	8277	8037	5845

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

A.5 Adjusting for Geographic FEs

Table A7: Adjusting for Geographic FEs

	Ethnic Attrition	
	(1)	(2)
$\Delta \%$ Latino ('00-'XX)	-0.45*** (0.05)	-0.33*** (0.05)
% Latino ('00)	-0.22*** (0.01)	-0.21*** (0.02)
Controls?	Y	Y
County FE?	N	Y
State FE?	Y	N
R ²	0.11	0.16
Num. obs.	61134	61134
N Clusters	10479	10479

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$