8-2012

Cracks in the Melting Pot: Immigration, School Choice, and Segregation

Elizabeth U U. Cascio
Dartmouth College

Ethan G. Lewis
Dartmouth College

Cracks in the Melting Pot:
Immigration, School Choice, and Segregation *

Elizabeth U. Cascio
Dartmouth College and NBER

Ethan G. Lewis
Dartmouth College

August 14, 2010

Abstract

This paper examines whether the large wave of Mexican immigration to the United States since 1970 has lowered non-Hispanic demand for public education. Our analysis focuses on California, where many of these immigrants settled, accounts for endogeneity of immigrant inflows using established settlement patterns, and uses relative outflows of children from a district to identify shifts in district choice working through schools. We find that between 1970 and 2000, the average metropolitan school district in California lost at least 12 non-Hispanic children to other school districts and two to private school within district for every ten additional low-English Hispanic arrivals in its public schools. These responses are similar in magnitude to "white flight" from school districts court-ordered to desegregate in the 1960s and 1970s.

* By mail: Department of Economics at Dartmouth College, 6106 Rockefeller Center, Hanover, New Hampshire 03755. By email: elizabeth.u.cascio@dartmouth.edu and ethan.g.lewis@dartmouth.edu. We thank David Card, Aimee Chin, Bill Fischel, Nora Gordon, Jennifer Hunt, Sarah Reber, Doug Staiger, Tara Watson, and seminar participants at Brown University, Columbia University, Cornell University, Dartmouth College, Stanford University School of Education, RAND, the Federal Reserve Bank of San Francisco, the NBER Summer Institute, and the American Economic Association Annual Meetings for very helpful comments on an earlier draft of this paper. We also gratefully acknowledge the research assistance of Christopher Bachand-Parente and Kory Hirak and funding from Dartmouth College. All errors are our own.

I. Introduction

The boom in immigration from Mexico over recent decades has led to a striking increase in the presence of children with limited English skills in American public schools. In fact, nearly all public school enrollment growth in the U.S. over the past decade can be accounted for by so-called "English learners" (ELs) (Park, 2009).[1] This trend is both likely to continue, given the relatively high fertility of these immigrants, and to touch many school systems across the country, as they begin to spread from their traditional destinations (Card and Lewis, 2007). It therefore has large potential consequences for public education into the foreseeable future. On one hand, there may be real impacts on school quality for native students: while ethnic and cultural diversity may enhance learning, school output also may be diminished by reductions in resources, overcrowding, or negative classroom spillovers. On the other hand, tastes regarding ethnic or socioeconomic diversity in education may affect native demand for public schools even absent any effect on school quality.

We address this issue by examining how increases in the presence of low-English Hispanics in public schools have shaped decisions over school district of residence for non-Hispanic households with children since 1970.[2] We center our analysis on California, which has received more recent Mexican immigrants than any other state: as shown in Figure I, first- and second-generation Mexicans accounted for over a quarter of children in California by 2000, compared to 7.5 percent of children nationwide.[3] By focusing on demand for schools revealed through residential choices (Tiebout, 1956), our analysis provides insight into a compelling potential cause of neighborhood segregation of recent immigrants (see, for example, Cutler, Glaeser, and Vigdor, 2008a, 2008b; Saiz and Wachter, 2006) that has not previously been explored.

---

[1] Throughout the paper, we use the terms "low-English" and "English learner" (EL) interchangeably.

[2] We focus on low-English Hispanics and non-Hispanics, rather than Mexican immigrants and natives, because of constraints on data available at the school district level. Data constraints also preclude us from examining sorting across schools within district (Alesina, Baqir, and Hoxby, 2004; Kane, Riegg, and Staiger, 2006; Weinstein, 2009).

[3] We focus on one state to hold constant a number of institutions that are critical to interpretation. About 53 percent of Mexican children living in the U.S. between 1976 and 1990 resided in California. This share fell by 2000, but remained high at 44 percent. (Authors' calculations from the 1976 Survey of Income and Education and the Decennial Censuses.)

We face two identification problems in our analysis. First, immigrant settlement is not random. As a result, correlations between low-English Hispanic inflows and non-Hispanic outflows will not identify a causal effect, an issue established in studies of the native migration response to labor market competition from immigration (e.g., Card and DiNardo, 2000; Card, 2001; Card, 2007). Second, one cannot immediately infer that rising low-English Hispanic enrollments affect non-Hispanic choice over district of residence through schools. The immigration driving this increase may affect location decisions independently, through either tastes or changes in other amenities (Saiz and Wachter, 2006) and may raise rents if the supply of housing is not perfectly elastic (Saiz, 2003, 2007). Some non-Hispanic households with children may find it optimal to locate in a different school district for these reasons alone.

We approach the first of these identification problems by constructing an instrument for the growth in low-English Hispanic enrollment in a district based on the spatial distribution of Mexican immigrants in 1970. The idea is that existing Mexican communities should have made some districts relatively attractive for new immigrant settlement irrespective of their other attributes, including their schools. California provides fertile ground for using this approach: Employment both in urban areas and in agriculture meant that the existing stock of foreign-born Mexicans was spread throughout the state, not concentrated in few places where they would have already been affecting schools for natives. Along these lines, 1970 Mexican settlement patterns are not related to proxies for existing school quality, such as per-pupil property tax revenues and expenditures, or to demand for local public schools revealed by existing levels and prior trends in the distribution of non-Hispanic children across school districts. They are, however, strongly related to subsequent growth in the presence of low-English Hispanic students in a district, as predicted.

In an attempt to isolate the contribution of schools to district choice, we embed this instrumental variables strategy in a model that compares changes over time in a district's stocks of young non-Hispanic households with and without children. That is, we test whether the arrival of low-English Hispanics to a district's schools as a result of Mexican immigration led to *relatively* large departures of non-Hispanic households with children. The assumption is that young non-Hispanic households without children valued schools no more than non-Hispanic households with children, but on average reacted similarly to the presence of Mexican immigrants more generally within a community. Supporting this idea, young non-Hispanic householders express similar sentiments regarding immigration regardless of the presence of children in survey data.[4]

We implement this identification strategy using district-level population and enrollment data from the U.S. Census and other sources. Our estimates imply that between 1970 and 2000, the average metropolitan school district in California lost at least 12 non-Hispanic children for every 10 additional low-English Hispanic arrivals in its public schools. We show that these effects are too large to be accounted for by an inelastic supply of housing suitable for households with children. Districts that received more low-English Hispanics through immigration also saw relatively large increases in the private school enrollment rates of their non-Hispanic residents.[5] On balance, our findings are thus consistent with low-English Hispanic enrollment growth reducing non-Hispanic demand for public schools.

Unfortunately, our approach does not allow us to pinpoint how low-English Hispanic students changed school quality for non-Hispanic children, if at all. To the extent that there were real impacts, however, our focus on California – while a limitation on generalizability – helps to rule out several channels. California maintained bilingual education under a highly egalitarian and potentially under-funded school finance system for most of the period of study.[6] This suggests that

[4] We omitted older households both because they express stronger anti-immigrant views, and because Proposition 13 constrains the residential mobility of Californians who owned homes in the 1970s. See below.
[5] This complements existing evidence of such an effect at the metropolitan area level (Betts and Fairlie, 2003).
[6] See Brunner and Sonstelie (2006) for a discussion of school finance in California. Under California's system of bilingual education, students could spend an extended period in separate classes taught in Spanish.

peer effects within classrooms are an unlikely explanation for our findings, as is a reduction in local school funding. It is more likely that Mexican immigration exacerbated crowding in schools or generated reductions in already limited resources for native students.[7] Further research on the impacts on educational outcomes for native students would be instructive in this regard.[8]

Still, our findings speak to the possible importance of demand for local public schools in residential and school segregation for a large share of the country's largest immigrant group. In broader historical perspective, the effects that we estimate are substantial, in percent terms roughly on the same order of magnitude as the white public school enrollment and population losses from urban school districts court-ordered to integrate blacks in the 1960s and 1970s (Reber, 2005; Baum-Snow and Lutz, 2009; Boustan, 2010a). This suggests that findings in the school desegregation literature may be relevant not only historically, but also to the modern wave of Hispanic immigration that school districts throughout the country will likely grapple with for decades to come.

II. Immigration and District Choice in Theory

Our analysis uses changes in the district choices of non-Hispanic families with children to infer whether rising low-English Hispanic enrollments have lowered non-Hispanic demand for public schools. The framework presented in this section illustrates the ways in which Mexican immigration may affect location decisions more generally and highlights the conditions under which we might plausibility isolate school-influenced district choice.

Suppose that households are exogenously assigned to different metropolitan areas by job opportunities, and within metropolitan area, choose a school district in which to reside.[9] There are

---

[7] Using data on close school bond elections and on house prices, Cellini, Ferreira, and Rothstein (2010) find evidence of underinvestment in school infrastructure in California during the period of interest. ELs were also viewed by Californians as being costlier to educate than natives even after bilingual education ended with passage of Proposition 227 in 1998 (Baldassare, 2005).

[8] See Betts (1998) and Gould, Lavy, and Paserman (2009).

[9] Metropolitan areas are by construction intended to capture labor markets. It is common to assume that metropolitan areas also define markets for schools (e.g., Hoxby, 2001; Urquiola, 2005; Rothstein, 2006). The assumption of exogeneity in metropolitan area of settlement is a simplification and is not required for identification in our analysis.

two types of households: those with children (j=1) and those without children (j=0).[10] In equilibrium, the (indirect) utility, V, associated with the school district of settlement for a household of type j must be at least as large as the highest available to it elsewhere in the metropolitan area, v:

$$V_j(p, q, k, i) \geq v_j. \tag{1}$$

Utility is decreasing in housing costs, p, and (weakly) increasing in public school output, q.[11] Utility is also (weakly) decreasing in both the low-English Hispanic share in public school enrollment, k, and the overall Mexican immigrant share in the population, i. Both capture distaste for diversity, but on different dimensions – in public schools (k) and elsewhere in the community (i).[12] k is increasing in i.

Now suppose that a district receives an influx of Mexican immigrants, all else constant. Consider first the settlement decision of a non-Hispanic household with children. Its demand for the district's public schools may decline, by way of distaste for diversity in schools ($\partial V_1/\partial k < 0$), reductions in school output ($(\partial V_1/\partial q)(\partial q/\partial k) < 0$), or both. But its demand for residence in the district may also fall if Mexican neighbors are undesirable for other reasons ($\partial V_1/\partial i < 0$). Immigration will also lower demand by raising housing costs if the housing supply is not perfectly elastic ($(\partial V_1/\partial p)(\partial p/\partial i) < 0$).

We are interested in isolating reductions in the probability of choosing the district working through schools. To do so, consider the location decision of an otherwise comparable non-Hispanic household without children after the same influx of Mexican immigrants. Being childless, this household values what happens in schools no more than its counterpart with children ($\partial V_1/\partial k \leq \partial V_0/\partial k$ and $\partial V_1/\partial q \geq \partial V_0/\partial q$) but has the same distaste for Mexican neighbors

---

[10] There is also heterogeneity in preferences within household types. We suppress this additional notation here.

[11] Housing costs include both the (rental) price of housing and taxes. Utility would generally be modeled as increasing in income less these housing costs, or potential consumption. We abstract from the effects of income here since cross-district moves within a labor market are not likely to change it.

[12] We partition the amenity space into two orthogonal parts – one potentially affected by immigration (captured in i) and one not (i.e., distance to the central business district). We abstract from the latter since its influence on district choice is accounted for by our identification strategy.

($\partial V_1/\partial i = \partial V_0/\partial i$) and the same disutility of housing costs ($\partial V_1/\partial p = \partial V_0/\partial p$).[13] This suggests that if the analysis could be limited to otherwise similar non-Hispanic households with and without children, the difference in their probabilities of choosing a district after it receives Mexican immigrants will reveal whether non-Hispanic demand for local public schools has fallen. Further, it will provide a lower bound on the magnitude of the effect, since it may be the case that households without children derive utility from local public schools. It will not, however, identify the channel(s) – preferences and/or real school quality effects – through which declines in demand for public schools are occurring.

Our empirical model generalizes this idea, testing whether school districts that experienced relatively large increases in *k* for their metropolitan areas also saw relatively larger declines in their representation of the metropolitan area's non-Hispanic households with children than in their representation of comparison households without children. Our use of district-level tabulations limits how we are able to define the comparison group. We do our best to meet the conditions described above, dropping households with older heads (specifically, those over age 50), since older individuals express relatively negative views about immigration. By contrast, households with younger heads express similar views toward immigration regardless of the presence of children.[14]

---

[13] We assume one housing market, which returns to equilibrium when *p* adjusts sufficiently to restore (1) for both household types *j*. If Mexican immigrants and non-Hispanic households with children tend compete for the same types of houses (e.g., detached single-family residences), immigration may raise house prices relatively more for type 1 households, prompting greater losses of non-Hispanic households with children even absent changes in schools. We explore this alternative hypothesis below.

[14] The General Social Survey (Davis and Smith, 2009) asks a variety of questions about views on immigration. We examined the following questions, available in the 1996 and 2004 waves: "How much do you agree or disagree with the statement[s]:" (1) "Immigrants take jobs away from people who were born in America," (2) "Immigrants increase crime rates," (3) "America should take stronger measures to exclude illegal immigrants," and (4) "Immigrants are generally good for America's economy." We created a variable that attempted to summarize negative views of immigration, the sum of dummies for the "agree" and "strongly agree" responses for (1)-(2), "strongly agree" for (3), and "disagree" and "strongly disagree" for (4), each of which is one for 30-40 percent of respondents. The mean of this sum, conditional on a year effect, is 0.25 higher (standard error = 0.054) for those over 50 than those under 50. It shows a significant positive linear relationship with age for those above but not below 50. Unrestricted age group (5-year bands) effects, conditional on a year effect, are jointly significant at the 5 percent level among those over 50 but not under 50. All three facts are true (with slight changes in the coefficients) whether one uses age as of the survey year or age as of 2000 (i.e. cohort) as the measure of "age." Finally, the first principal component of the 16 dummies for all possible responses

Dropping older households is also appropriate since their mobility is constrained as a result of California's Proposition 13 and subsequent ballot initiatives.[15]

An additional complication for estimation is that the settlement of immigrants is not random: the same basic model given in (1) underlies their residential choices. Mexican immigrants may be attracted to declining school districts by lower rents, in which case non-Hispanic departures from a district might generate increases in *i* (and *k*), not vice versa.[16] Or a positive shock might attract households with children, regardless of origin, possibly generating a positive correlation between *i* (and *k*) and the presence of non-Hispanic children in a district. We approach this identification problem using the instrumental variables approach described below, the basic idea behind which is that new immigrants value an existing Mexican presence when choosing a residence.

III. Empirical Strategy

III.A. Basic Model

A linear version of the model of residential choice presented above is given by:

$$Y_{idmt}^{j} = \gamma_{dt} + \beta_j k_{dt} + \eta_{idmt}^{j}, \tag{2}$$

where Y_{idmt}^{j} = 1 if non-Hispanic household *i* of type *j* resides in school district *d*, which is located in metropolitan area *m*, at time *t*.[17] This decision is a function of the share of public school enrollees who are low-English Hispanic in *d* at time *t*, k_{dt}, the response to which varies by household type,

("neither agree nor disagree" excluded) to the four questions follows a similar age and statistical significance pattern. If we add a dummy variable for the presence of children in the household to any of these models, its coefficient is small and statistically insignificant. This is also true specifically in the sample under 50.

15 Proposition 13, passed in 1978, effectively locked in property taxes for existing homeowners by establishing a statewide property tax rate of 1 percent and setting assessed valuations of property at 1975 levels, with a maximum increase of 2 percent per year and no re-assessment. Propositions 60 and 90 in 1986 and 1988, respectively, allowed individuals aged 55 and over to transfer this tax benefit to a new home of equal or lesser value. These measures generated a sharp increase in residential mobility at age 55 (Ferreira, 2009).

16 The foreign-born do appear attracted to places with lower rents. For example, Boustan (2010b) shows that the foreign-born were attracted to center cities that whites had earlier fled in response to black in-migration.

17 *Y* in equation (2) might also be interpreted as the latent propensity to live in district *d*, with the probability of living in the district a non-linear transformation of it. If this transformation is a logistic CDF, then equation (3) below will be the same except that the share of type *j* households in *d* will be replaced with the natural log of the odds that a type *j* household is in *d* (equal to ln(share/[1-share])). We further motivate and estimate this model below.

with j=1 for the treatment group (non-Hispanic households with children), and j=0 for the comparison group (other young non-Hispanic households without children). This decision is also a function of other district characteristics at time t not differentially valued across types, captured by γ_{dt}. Conceptually, the γ_{dt} incorporate common shocks coming from both immigration and other (unobserved) factors. The parameter of interest is $\theta \equiv \beta_1 - \beta_0$, the difference in the sensitivity of location decisions to k across household types. If the conditions for a valid comparison group are met, θ captures whether there is a reduction in non-Hispanic demand for public schools.

We are not able to estimate (2) given a lack of household-level data with sufficient geographic detail. Instead, consider summing (2) across all non-Hispanic households i of type j in metro area m at time t, then dividing by the total number of non-Hispanic households of type j in m at time t. Stacking across districts generates a model for the proportion of non-Hispanic households of type j in m residing in d at time t:

$$\frac{N^j_{dmt}}{N^j_{mt}} = \gamma_{dt} + \beta_j\, k_{dt} + \varepsilon^j_{dmt}\,.$$

Differencing over time within household type, then across types, eliminates the effects of all common factors affecting location at a point in time (the γ_{dt}) and generates our model of interest:

$$(3)\qquad \Delta\frac{N^1_{dm}}{N^1_m} - \Delta\frac{N^0_{dm}}{N^0_m} = \theta\,\Delta k_d + \left(\Delta\varepsilon^1_{dm} - \Delta\varepsilon^0_{dm}\right).$$

The dependent variable is the treatment-comparison difference in $\Delta\left(N_{dm}/N_m\right)$. Restated, θ therefore captures *how much larger* a proportion of a metropolitan area's non-Hispanic households with children than young non-Hispanic households without children departed a district with a rise in the low-English Hispanic public enrollment share.

Unfortunately, the available district-level counts of households by presence of children are not broken out by the age of the householder, so we cannot estimate equation (3). Instead, we use available data on non-Hispanic population by age at the school district level. Our baseline model considers the treatment group to be non-Hispanic 0 to 19 year olds ("children") and the comparison group to be 20 to 49 year olds ("young adults"). That is, our estimating equation is:

$$(3') \qquad \Delta \frac{N_{dm}^{0-19}}{N_{m}^{0-19}} - \Delta \frac{N_{dm}^{20-49}}{N_{m}^{20-49}} = \tilde{\theta}\, \Delta k_d + \left(\Delta u_{dm}^{0-19} - \Delta u_{dm}^{20-49} \right),$$

where $\Delta\left(N_{dm}^{a} / N_{m}^{a}\right)$ represents the change over time in the share of non-Hispanics in age group *a* in metropolitan area *m* living in district *d*. The parameter $\tilde{\theta}$ is similar in interpretation to θ, but now gives how much larger a proportion of a metropolitan area's non-Hispanic *children* than young adults departed a district with a rise in the low-English Hispanic public enrollment share.

Model (3') has the obvious drawback of including in the comparison group individuals who are parents, but has the advantage of simplicity. Further, the fact that the comparison group is partially treated biases us against finding an effect. In the Appendix, we show that $\tilde{\theta} \approx \left(1 - \varphi^{20-49}\right)\theta$ where φ^{20-49} represents the fraction of non-Hispanic 20 to 49 year olds who live in households with at least one 0 to 19 year old. Thus, $\tilde{\theta}$ in equation (3') understates θ in equation (3) by a factor proportional to the fraction of the comparison group with children. In 2000 Census microdata for metropolitan California, $\varphi^{20-49} \approx 0.55$, suggesting that our estimates should be scaled up by about 2.2 (=1/0.45) to represent the outflows of non-Hispanic households with children.

III.C. Instrument

For ordinary least squares (OLS) estimates of (3') to produce consistent estimates, it must be the case that Δk_d is unrelated to all unobserved determinants of differences in the district choices of the treatment and comparison groups. As described above, the driving force behind rising low-English shares in public school enrollment in California is Mexican immigration, and Mexican immigrants may be attracted to districts that non-Hispanic households with children are already departing, or families with children may be attracted to the same places, regardless of ethnic background. There are other potential sources of bias as well: departures of non-Hispanic children will inflate Δk_d, and any measurement error in Δk_d will lead to attenuation bias.

We address biases from reverse causation and omitted variables by constructing a prediction of future Mexican immigrant settlement based on existing Mexican immigrant settlement – an approach previously used to examine the impacts of immigration at the metropolitan-area level (e.g., Card and Dinardo, 2000; Card, 2001). In practice, Δk_d is calculated over 1976 to 2000, so we obtain a prediction of low-English Mexican arrivals to a district's public schools over this period:

$$(4) \qquad \hat{\Delta} I_d = \frac{M_d^{1970}}{M^{1970}} \Delta I .$$

M_d^{1970} / M^{1970} is the share of the Mexican-born population (of all ages) residing in district *d* in 1970, based on Census school district tabulations , and ΔI represents the number of low-English children of school age in the 2000 Census Public Use Microdata Samples (PUMS) (Ruggles, et al., 2010) who were either born in Mexico or born in the U.S. to at least one Mexican-born parent (or if both parents are foreign-born, to a Mexican-born mother) who arrived in the U.S. in 1976 or later. We classify as low-English those who are neither fluent nor speak English very well, since this definition close to reproduces EL shares in public enrollment in California in our administrative data.[18] Thus, $\hat{\Delta} I_d$ apportions low-English Mexican public school students to school districts as if the settlement patterns of Mexican immigrants in the United States had not changed since 1970.

We address the mechanical source of bias described above in transforming the prediction in equation (4) into an instrument for Δk_d. Here, we assume that $\hat{\Delta} I_d$ represents the only source of enrollment change in a district. Our instrument, Z_d, is thus:

[18] Our TSLS estimates are not sensitive to how "low-English" is defined or to using all Mexican children; this is to be expected given that ΔI depends in no way on the district. Our choice of low-English classification serves to facilitate interpretation of the estimates. For the same reason, using 1976 data to calculate ΔI has no impact on our estimates.

(5) $$Z_d = \frac{I_d^{1976} + \hat{\Delta} I_d}{Enr_d^{1976} + \hat{\Delta} I_d} - \frac{I_d^{1976}}{Enr_d^{1976}},$$

where I_d^{1976} and Enr_d^{1976} represent the initial (1976) enrollment of low-English Hispanics and total enrollment, respectively, in district *d*. The second component of Z_d is thus the initial share low-English Hispanic in the district, while the first represents what that share would have been in 2000 had nothing else changed. Considering that Z_d is another noisy estimate of growth in the EL Hispanic enrollment share, using Z_d as an instrument may allow us to remove attenuation bias from measurement error as well.

Appendix Table I gives the top ten districts in our sample ranked by both the share of Mexicans in the United States residing in the district in 1970 (M_d^{1970} / M^{1970}) and the value of the instrument; data sources are described below. Unsurprisingly, larger districts rank high on the first measure: Los Angeles Unified alone has home to over 16 percent of Mexicans in the U.S. in 1970.[19] However, the largest districts are not those with the largest predicted increases in low-English Hispanic shares. And while districts with the highest values of the instrument may appear disproportionately drawn from the Central Valley, nearly 80 percent of the variation in Z_d in our sample is within, not across, metropolitan areas. The implied dispersion of existing Mexicans across California is consistent with the historical antecedents to the modern wave of Mexican immigration, which pulled Mexican immigrants to the countryside of the state as well as to urban areas.[20] In general, the instrument is not related to district size or location, as shown below.

Two-stage least squares (TSLS) estimates of (3') are identified if the predicted change in low-English Hispanic share defined in (5) is unrelated to unobserved shocks to the desirability of a

[19] School districts outside of California that rank high on this measure include Chicago City (4.4 percent), El Paso ISD (3.1 percent), San Antonio ISD (2.3 percent), and Houston ISD (1.6 percent). Outside of California, however, foreign-born Mexicans were much more concentrated in 1970, making application of our identification strategy less compelling.
[20] For example, the Bracero Program brought an estimated 5 million agricultural guest workers to the United States (many of these to California) between 1942 and 1964.

district for non-Hispanic children. In particular, it must be the case that established settlement patterns of Mexican immigrants do not predict subsequent (unobserved) shocks to schools. It is impossible for us to test this assumption directly. However, we can test whether Z_d is correlated with relative outflows of non-Hispanic children from a district over the 1960s or with an uneven distribution of children and young adults in 1970 – before the big wave of Mexican immigration, but when the district may already have been in decline. In addition, we will test whether the instrument is correlated with other district characteristics observed in the early 1970s, some of which may predict whether a district will later become less attractive for non-Hispanic children.

IV. Data

IV.A. Sources, Key Variables, and Sample

Our data sources on district of residence are the 1970 and 2000 school-district tabulations of the U.S. Census (hereafter referred to as the School District Tabulations, or SDT). The SDT provides information on the distribution of foreign-born Mexicans across school districts in 1970, which we use to predict future Mexican arrivals to the district (see equation (4)). It also provides non-Hispanic population counts by age group, which we use to construct our main dependent variable – the 1970 to 2000 change in the proportion of a metropolitan area's non-Hispanic children (0 to 19 year olds) residing in a district, less the 1970 to 2000 change in the proportion of a metropolitan area's non-Hispanic young adults (20 to 49 year olds) residing in a district (see equation (3')).[21] We use information on school enrollment by type (public/private) and ethnicity in the SDT to construct the 1970 to 2000 change in the non-Hispanic private school enrollment rate.

Data on public school enrollment by ethnicity and English proficiency at the district level were drawn from the 1976 and 2000 Elementary and Secondary School Civil Rights Surveys,

[21] Because school districts are sometimes missing in these data sets (see below), we obtain non-Hispanic population at the MSA level using separately-reported county level aggregates downloaded from the National Historical Geographic Information System (Minnesota Population Center, 2004), rather than by aggregating across school districts.

conducted by the Office for Civil Rights in the Department of Health, Education, and Welfare (HEW, later the Department of Education) to monitor compliance with federal civil rights law (hereafter referred to as the OCR). Non-native English speakers became protected under federal law as a result of the Equal Educational Opportunity Act of 1974, and HEW set forth guidelines for accommodation and began monitoring district compliance in 1975. The first year in which data are available for all districts is 1976.[22] We use district enrollment counts reported by OCR to create the explanatory variable of interest – the 1976 to 2000 change in the EL Hispanic share in enrollment – and in construction of the instrument (see equation (5)).

We restrict our analysis to school districts in the 23 standard metropolitan statistical areas (by their 1990 definition; MSAs) in California. For our analysis, we define "districts" so as to have constant boundaries between 1970 and 2000, and aggregate key variables accordingly.[23] We lose "aggregated" school districts for several reasons. First, while both the SDT and the OCR are in principle censuses of school districts, the 1970 SDT does not include the smallest school districts in the country (those with under 300 students), and there was some non-response to the 2000 OCR survey. We also limit attention to districts where data quality is high in both years of the OCR (see Appendix). By and large, most sample drops on these grounds occur because a district is missing in the 1970 SDT.

We make several additional exclusions to arrive at our estimation sample. There are three district types in California: unified districts, which operate schools at all levels; secondary districts, which operate high schools; and elementary districts, which operate primary and middle schools and

[22] Most of the 1970 to 2000 increase in California's Mexican population appears to have occurred after 1976 (Figure I), so this likely has little effect on our findings. The 1980 SDT lacks sufficient disaggregation of population counts by age and ethnicity to apply our empirical strategy.

[23] For example, if districts A and B in 1970 merge to form C by 2000, we aggregate A and B to create an observation for C in 1970. We identify school district reorganizations using data from the Elementary and Secondary Education General Information System and the Common Core of Data Public Agency Universe and internet searches. We drop "aggregated" districts involved in reorganizations over the period if any of the component districts are not present in years they should be or, in a few cases, if we were not able to ascertain the nature of the reorganization that occurred.

feed into secondary districts. Including both secondary and elementary districts in our analysis would be redundant, as they cover the same geographic area. Secondary districts are more comparable in geographic scope to unified districts and, in the SDT, represent elementary districts that are not directly observed due to their small size. We therefore use secondary district boundaries to define observations for our main analysis and aggregate the OCR data accordingly. Observations for "secondary districts" therefore represent a combination of elementary and secondary districts.[24] In our main sample there are a total of 40 combined elementary-secondary districts and 193 unified districts, for a total of 233 observations.

IV.B. Descriptive Statistics

Table I shows summary statistics for our estimation sample. Panel A gives statistics on the key variables in our analysis. Consistent with the rise in first and second-generation Mexican children in California (Figure I), the EL Hispanic share in public school enrollment rose a substantial 10.7 percentage points between 1976 and 2000, on a base of 2.6 percent (Panel C). The standard deviation on this variable is also 10 percentage points, suggesting that there was quite a bit of variation in the growth in immigrant share across districts. The instrument has a comparable mean and standard deviation.

The components of the main outcome – the 1970 to 2000 changes in the proportions of an MSA's non-Hispanic children and young adults in a district – are on average close to zero. This is expected by definition; deviations from zero come about because we are missing districts at times for the reasons described above. The private school enrollment rate of non-Hispanics rose by a considerable 10.3 percentage points between 1970 and 2000, more than a seven-fold increase over its 1970 level of 1.4 percent (Panel B). While suggestive of non-Hispanic "flight" from public schools, this could be a consequence of school finance equalization (e.g., Downes and Schoeman,

[24] Our estimates are robust to dropping these observations, as shown below.

1998) rather than a response to rising low-English shares in public enrollment. School finance equalization may also have affected the income heterogeneity of populations within districts (Aaronson, 1999). In general, Mexican settlement could be correlated with the effects of school finance reform on a district, making equalization a potential confounding factor that we hope to rule out through our identification strategy.

In an effort to do this and to demonstrate the credibility of our identification strategy more generally, we have collected information on district property tax revenues and expenditures per pupil prior to the implementation of the *Serrano* decisions (for 1971-72 from the 1972 Census of Governments), as well as information on a number of other district characteristics circa 1970 from various historical sources (see Appendix). Descriptive statistics on these variables are shown in Panel C of Table I. The average district in our sample enrolled about 14,849 students and raised $636 per-pupil in revenues through the property tax and spent $1084 per student, with a standard deviation of around $300 in each. About 8 percent of districts in our sample are in center cities, and 71 percent had a private school serving any grade level as of 1976, the first year in which such information is available.

V. Evaluating the Identification Strategy

Our identification strategy relies on the assumptions that the instrument is related to changes in the actual presence of low-English Hispanics in a district's public schools, but is unrelated to other factors that might have compelled non-Hispanic households with children to locate elsewhere, like a loss in the district's fiscal autonomy as a result of school finance equalization or declining school quality for other reasons. Before turning to our main estimates, it is useful to examine the credibility of these assumptions.

The first row of Table II speaks to the former, showing that there is a strong first-stage relationship between the instrument and the change in a district's EL Hispanic share in our baseline specification, which includes fixed effects for MSA by district type.[25] Figure II shows this relationship graphically. A coefficient of one on the instrument would be expected if predicted arrivals matched actual arrivals, on average, and enrollments were otherwise changing little. The actual first-stage coefficient is statistically distinguishable from zero, at the one percent level, but also significantly less than one. Fewer than the expected number of EL Hispanic students actually show up in districts, possibly due in part to the fact that Mexican immigrants have spread out geographically over time, especially since 1990 (Card and Lewis, 2007).

As a preliminary exploration of the second identifying assumption, the remainder of Table II shows slope estimates from comparable reduced-form regressions of many potential correlates of outcomes on the instrument. By and large, the findings support the credibility of our research design. Consider first the estimates for initial district observables, shown in Panel C. While the instrument is negatively correlated with the likelihood of having a private school by 1976, it is unrelated to myriad other district characteristics. It is not significantly related to initial enrollment (in logs and levels, not shown) or to center city status. It is also not correlated with the density of public schools within a district, which we capture with a dummy for whether a school district had an above average number of schools (in 1972) given its enrollment (in the 1976 OCR) and land area.[26] Finally, relationships between the instrument and the school finance variables are statistically insignificant and small, amounting to, for example, only about $20 differences in per-pupil property tax revenues and per-pupil expenditures for a one standard deviation (0.10) change in the

---

[25] We include MSA by district type fixed effects because both the instrument and private school enrollment rates vary across MSAs and district types. Districts are also sometimes missing from the sample for reasons described above, so that the sum of district shares of the MSA non-Hispanic population do not sum to one within MSA. Standard errors are clustered on metropolitan area, and t-statistics are compared to critical values in a t-distribution with 21 degrees of freedom (the number of clusters less two, a rule of thumb explored in Cameron et al., 2008).

[26] This could also be interpreted as showing that the instrument is uncorrelated the median voter's preferences for segregation. We arrive at this prediction by regressing the natural log number of schools on the natural log in enrollment, the natural log of land area, their interaction, and all of the other pre-existing district characteristics listed in Table I, Panel B (except the initial EL Hispanic enrollment share), and classifying as "high choice" those with non-negative residuals. We explored alternative regression models (e.g., in levels, not logs), as well as considered schools per square mile and schools per child enrolled. All yielded an insignificant relationship with the instrument.

instrument. Still, confidence intervals on these estimates (and some of the others discussed above) are fairly wide, so we include all of these characteristics as controls in some specifications below.

The instrument is nevertheless significantly related to the EL Hispanic share in public school enrollment in 1976. Mexican immigration may have already been changing the demographics of children in California in the early 1970s (see Figure I). Even if this were not the case, a relationship would not be surprising to see here given that the instrument derives from the district's Mexican-born population in 1970. The initial share of enrollment that was EL Hispanic was small, however, compared to the change that occurred over the next 30 years (Table I), making any changes in district choice among non-Hispanics plausibly related to the predicted change rather than the initial level. Put differently, if the initial EL Hispanic share represents a meaningful source of variation in the instrument, then the instrument should predict initial levels (and prior trends) in demand for schools, invalidating our preferred interpretation.

Table II Panel B shows that this does not appear to be the case, as the instrument is unrelated to 1970 levels of our outcome variables; it also unrelated to prior trends, shown in a separate table below. Despite having fewer private schools by 1976, the 1970 private school enrollment rate of non-Hispanics was not significantly lower in districts with higher values of the instrument. And while the instrument is negatively related, though not significantly so, to the share of an MSA's non-Hispanics residing in a district in 1970, the *difference* between the estimates for children and young adults – which represents the 1970 level version of our main dependent variable – is not statistically significant, and fairly precisely estimated.

VI. Choice of School District

VI.A. Reduced-Form Estimates by Group

To make our identification strategy transparent, we begin the presentation of our findings by decomposing the reduced-form specification of model (3) into its constituent parts – separate estimates, for the treatment group (0 to 19 year olds) and the comparison group (20 to 49 year olds), of the relationship between the instrument and the 1970 to 2000 change in the district's share of the metropolitan non-Hispanic population. We do this in the first two columns of Table III, again conditioning on metropolitan area by district type fixed effects. The difference between these estimates, presented in column (3), is the reduced-form parameter of interest.

Panel A presents these estimates for the full sample. The instrument is negatively related to the 1970 to 2000 change in the proportion of the MSA's non-Hispanic children residing in a district: a 0.1 increase in the expected EL Hispanic share – again, roughly the standard deviation of the instrument (Table I, Panel A) – is on average associated with a population loss amounting to 0.729 percent of an MSA's 0 to 19 year olds (column (1)). For a district of average initial size, this figure represents about 9.2 percent of non-Hispanic children in 1970 (0.729/7.9), a sizable effect.[27]

On the other hand, not all of this effect may represent reductions in demand for public schools: a one standard deviation increase in the instrument is associated with a loss of about 0.425 percent of an MSA's non-Hispanic young adults as well (column (2)). We assume that the estimate for the comparison group accounts for all other reasons that immigration might prompt moves from a district. The parameter of interest is therefore the difference in these coefficients, which is -0.0304, as shown in column (3). Using the metric described above, this suggests that a district of average initial size would have lost about 3.8 percent (0.3/7.9) of its non-Hispanic children to location decisions on the basis of schools for every standard deviation increase in the instrument. Since many 20 to 49 year olds are parents, this figure understates the loss of non-Hispanic *households* with children as a result of Mexican immigration, as discussed above.

Figure III gives a graphical representation of the regression in column (3). With the exception of large districts, like Los Angeles Unified, San Francisco Unified, and San Diego Unified

[27] The average district in our sample had 7.9 percent of its MSA's non-Hispanic children in 1970 (Table I, Panel B).

– which naturally account for larger shares of their respective MSA's non-Hispanics at baseline – districts are tightly clustered around the downward sloping line. Below, we show that our findings continue to hold when large center city school districts are dropped from the sample – a finding which is not surprising given that the instrument is uncorrelated with district enrollment and center city status – and under alternative specifications of the dependent variable that are less sensitive to scale. We also present TSLS estimates that correspond to the same model, explore the sensitivity of this model to inclusion of controls, and discuss ways to interpret the magnitude of the estimate.

The remainder of Table III presents a falsification exercise that complements that performed in Table II, Panel B. In particular, we test whether there was an "effect" on relocation in the 1960s – prior to the big wave of Mexican immigration.[28] If so, it would suggest that it is not the arrival of Mexicans *per se* that drove the "outflows" of non-Hispanic children, but some other factor correlated with it. Constructing similar measures for 1960 requires us to use a smaller sample, as detailed geographic tabulations are available only for a subset of counties in the 1960 Census.[29] Limiting the sample to the 13 MSAs which are observable in 1960, the 1970 to 2000 finding is still negative, significant, and quite similar in magnitude to that found for the full sample, at -0.0353 (column (3)). Nevertheless, the age group-specific relationships with the instrument (columns (1) and (2)) are different in this subsample of older cities. The sensitivity of the estimates in column (1) but not (3) reinforces the utility of using a comparison group, as differencing accounts for any unobservables that make a residence in district more or less desirable for non-Hispanics regardless of age.

The final panel shows the findings from the falsification exercise. Supporting the identifying assumptions of our model, there is little evidence that districts with larger values of the instrument were already experiencing greater losses of children in the 1960s: the coefficient in column (3) is a

[28] Mexicans were arriving California in the 1960s, but in much smaller numbers (Figure I). We cannot perform a similar exercise for private school enrollment given data limitations.
[29] In the 1960 Census, only metropolitan counties were assigned tracts, the lowest level of geography with sufficiently detailed tabulations for our analysis. Assignment of 1960 tracts to school districts is discussed in the Appendix.

small -0.0066 and is not statistically significant. Figure IV shows scatter plots that correspond to this model (Panel B) and its 1970 to 2000 counterpart for the restricted sample (Panel A). The graph shows that the lack of a pre-trend is not driven by some outlying observation. Thus, a significant negative relationship between the instrument and relative loss of children from a district appears limited to the period in which Mexicans were arriving in California in large numbers.

VI.B. Main Findings

Table IV shows our main results – TSLS (Panel A) and OLS (Panel B) estimates of model (3'). The TSLS estimate in column (1) corresponds to the reduced form estimates in column (3) of Table III, Panel A, and is thus based on a specification that includes fixed effects for MSA by district type and no further controls. The statistically significant point estimate of -0.0705 implies that a 10 percentage point increase in the share of a district's public school enrollment that is low-English Hispanic (again, roughly one standard deviation) induced 0.705 percent of an MSA's non-Hispanic children to locate elsewhere, or about 8.9 percent (0.705/7.9) of non-Hispanic children in a district of average size initially. This estimate is larger than the reduced-form estimate presented in Panel A of Table III because the corresponding first-stage coefficient is less than one (Table II, Panel A).[30]

As anticipated, the TSLS estimates change little with the addition of controls for the pre-existing district characteristics listed in Table I, Panel C (column (2)) and for the initial (1970) level of the dependent variable (column (3)).[31] In fact, they are slightly larger. By contrast, OLS estimates fall in magnitude with the addition of controls (Panel B). They are also considerably smaller in magnitude at the outset (-0.0215 in column (1)) and not statistically significant in subsequent specifications. The OLS estimates may be biased downward from measurement error, or by

[30] The first-stage coefficient estimates on the instrument are very stable across specifications (available on request).

[31] We do not include the 1976 low-English Hispanic share in public school enrollment among these controls. Adding this control has little impact on the OLS estimates, and increases the TSLS estimates in magnitude to -0.0934 (se=0.0363) in column (1), -0.0992 (se=0.0467) in column (2), -0.0984 (se=0.0473) in column (3).

unobserved positive shocks – such as changes in nearby employment opportunities or cheap housing – attracting families with children, regardless of origin, to the same districts.[32]

In the remaining columns of Table IV, we present estimates for different subsamples of the data. Column (4) drops center city districts from the sample. The TSLS point estimate is slightly smaller than that in column (1), but it remains highly significant. This finding confirms that our full-sample estimates are not driven by the sensitivity of this outcome measure to district size. Further, it shows that our main findings are not simply being driven by non-Hispanic families with children moving out of center cities; Mexican immigration appears to have shaped their location decisions within the suburbs of California metropolitan areas as well. The final columns break out the estimates by district type. Though we cannot reject that the estimates are identical, the findings suggest that our main estimates are driven entirely by unified districts (column (5)). Estimates for the combined elementary-high school districts (column (6)) are much smaller and not statistically significant.[33]

VI.C. Additional Sensitivity Analysis

Table V examines various alternative formulations of the dependent variable. For comparison purposes, row (a) of Panel A repeats the TSLS (column (1)) and OLS (column (2)) estimates that appeared in column (1) of Table IV. Row (b) replaces the district's share of the MSA's non-Hispanic population, s_{dm}, with the log odds, $\ln\left(s_{dm}/\left(1-s_{dm}\right)\right)$, which would be appropriate if the household-level model underlying district choice were logit. The dependent variable is therefore the treatment-comparison difference in the 1970 to 2000 change in log odds, or approximately the treatment-comparison difference in the *growth rates* of s_{dm}. Logit is an attractive

---

32 Using an instrumental variables with similar motivations to that used here, Saiz and Wachter (2006) also obtain OLS estimates that are much less negative than TSLS when examining the response of native and white non-Hispanic population outflows to changes in the foreign-born population at the tract level in the entire U.S.

33 The TSLS estimate for elementary districts only is slightly larger in magnitude than that in column (6) and is more precisely estimated (coefficient=-0.0193, se=0.0126).

specification for this application, where the distribution of district size (and hence, s_{dm}) is right skewed. Both OLS and TSLS estimates are statistically significant in this specification, and the marginal effects, evaluated for a district of average initial size, are larger than were estimated in the linear model.[34] Moreover, we see the same pattern as in Table IV: TSLS is larger in magnitude than OLS, and OLS is more sensitive to controls (latter not shown).

We also estimate these two models weighted by the district's 1970 population of non-Hispanic children in Panel B. Weighting provides insight into whether the effects of immigration for the typical non-Hispanic student differ from those for the typical district. In the main specification (row (a)), the weighted TSLS coefficient is larger than the unweighted TSLS coefficient.[35] Weighted estimation of the logit model may provide more insight into differences in the actual treatment of low-English Hispanic students across districts of different size, since the logit transformation is theoretically less sensitive to scale. The weighted TSLS logit coefficient (row (b)) is a bit smaller than its unweighted counterpart in Panel A. Roughly speaking, this implies that larger districts lose on average fewer non-Hispanic children as a percent of their baseline population than smaller districts. This could be because there may be greater capacity to sort within larger districts, or less centralized administration may allow larger districts to mitigate some of the potential adverse effects of immigration for native students. Still, it is again the case that the marginal effect, now evaluated using weighted means, is larger than in the linear model.

Returning to the bottom two rows of Panel A, we present (unweighted) estimates for dependent variables that may be easier to interpret and have analogs in previous literature: the difference in growth rates (change in logs) between the district's child and young adult populations (row (c)) and the change in the share of the district's non-Hispanic population aged 0 to 49 who are

[34] Thus, we calculated the marginal effect as the coefficient*0.079*(1-0.079).
[35] If the response to a given change in low-English Hispanic share is a loss of a constant percent of a district's non-Hispanic child population, this will be larger as a share of the metropolitan population in initially larger districts.

children (row (d)). In both cases, there is again a significant negative relationship: In TSLS, an increase of 0.10 in the low-English Hispanic share in enrollment is associated with a 11 percentage point decline in the relative population growth rate of non-Hispanic children, and a 2.5 percentage point decline in the child share of the non-Hispanic population under 50.

These figures are in the same neighborhood as white population losses associated with court-ordered school desegregation in the 1960s and 1970s. Baum-Snow and Lutz (2009) find that major court orders to desegregate center city districts on average led to an increase in white exposure to black peers of 0.09 and declines in white enrollment and population of 12 percent and 6 percent, respectively.[36] This is comparable to our finding that 11 percent of non-Hispanic children leave an initially average-sized district with every 10 percentage point increase in the low-English Hispanic share.[37] Boustan (2010a) estimates that court orders to desegregate by 1980 in city districts were associated with a 13 percentage point increase in white exposure to black peers and about a 1.9 percentage point decline in the school-aged share in the white population. Our estimate of the decline in non-Hispanic child share is larger for the same percentage point increase in low-English Hispanic share, though our measure is also systematically larger because it is on the base of 0 to 49 year olds, not the entire population, and includes children under age five.

VI.D. Displacement Rates

How similar or different are the estimates in Table V? One way to make them comparable to each other is to restate them as "displacement rates" – how many non-Hispanic children "left" the average district for each EL Hispanic arrival into public schools.[38] Displacement rates based on our TSLS estimates are given in column (3) of Table V. To be consistent with earlier discussion, we

[36] See also Reber (2005).
[37] A 10 percentage point increase in low-English Hispanic share will lead to a 10 percentage point increase in the exposure of non-Hispanics to low-English Hispanics if there is no sorting response within the district.
[38] Though the word "displacement" connotes a fixed capacity, it is important not to take that connotation literally. We discuss the interpretation of our estimates in the next section.

base our calculations on a district of average initial size and assume a 10 percentage point increase in the EL Hispanic public enrollment share over 1970 to 2000. The displacement rates are similar if we consider larger or smaller changes or a district of median (rather than mean) initial size.

To illustrate, it is useful to walk through calculation of the displacement rate for our main specification, given in Table V, Panel A, row (a). In a district of average initial size, it would have required the addition of about 1690 low-English Hispanics to increase the EL Hispanic enrollment share by 10 percentage points.[39] We calculated above that this 10 percentage point increase led to an 8.9 percent reduction in non-Hispanic 0 to 19 year olds – about 1946 non-Hispanic kids. So, the TSLS coefficient implies a displacement rate of 1.15 (=1946/1690), or that 11 to 12 non-Hispanic children located elsewhere for every 10 additional low-English Hispanic arrivals in a district's public schools. This is an order of magnitude larger than existing estimates of private school displacement rates (Betts and Fairlie, 2003), as well as our own estimates of this parameter, given below. Residence may be a more important form of school choice in this context than private schooling.

It is also useful to consider displacement rates in terms of households, rather than children. We present such estimates in column (4). To arrive at these numbers, we rescaled the TSLS coefficients upward, to adjust for the fact that slightly over half of young adults are in households with children (see earlier discussion and Appendix), and converted the number of low-English Hispanic public school enrollees needed to deliver a 0.1 increase the EL Hispanic share into the number of households they represent.[40] These household-level displacement rates are uniformly larger – for example, 1.6 in our main specification.

---

[39] The average district in our sample had 14,850 students in 1976, 313 (about 2 percent) of whom were low-English Hispanics. To raise the low-English Hispanic share by 10 percentage points, to 12 percent, required the addition of about 1690 low-English Hispanic students, since $(313 + 1690)/(14850 + 1690) \approx 0.12$.

[40] In metropolitan California, there were on average 1.63 low-English Hispanic kids in public schools in households that had any in the 2000 Census PUMS, though there are more kids overall in these households. In this calculation, we also use 1970 district-level counts of non-Hispanic households with children instead of non-Hispanic children.

The other estimates in Panel A imply larger displacement rates. The estimated marginal effect from the logit specification (row (b)) implies displacement rates of 1.44 and 2 at the child and household levels, respectively. The 11 percent loss of non-Hispanics in the growth specification (row (c)) with a 10 percentage point increase in low-English Hispanic share implies similar displacement rates. When the dependent variable is the change in the share of 0 to 49 year olds who are children (row (d)), the displacement rates are also slightly larger than in the main specification. The weighted estimates in Panel B imply lower displacement rates, on the basis of weighted means of all key variables, but they remain above one at the household level.

VII. Schools versus Other Explanations

Inferring that the effects that we have estimated so far are working through schools rests on the assumption that we have identified a credible comparison group – one which removes the effects of all other immigration-related factors that may have induced non-Hispanic families with children to choose other districts. It is useful to explore several competing hypotheses and to provide additional evidence that EL Hispanics reduce non-Hispanic demand for public schools.

VII.A. Crowding in the Housing Market

A competing hypothesis for the effects we estimate is crowding in the housing market. To understand this, suppose that that housing costs, p, in equation (1) differ across household types j, and that Mexican immigration puts greater pressure on housing costs for households with children (i.e., by bidding up prices on detached single-family residences), or that $\partial p^1 / \partial i > \partial p^0 / \partial i$. Then even if p is valued in the same way regardless of household type, changes in housing costs resulting from immigration will have a relatively large impact on the settlement decisions of non-Hispanic households with children.

It is difficult to rule out a contribution of crowding in the housing market to our estimates, particularly since California metropolitan areas appear to have low housing supply elasticities (Saiz, 2010). Nevertheless, several pieces of evidence suggest that this cannot fully account for (and may not even account for much of) our findings. First, the household displacement rates presented in Table V are uniformly above one. Second, non-Hispanic children and low-English Hispanic children in metropolitan California live in very different types of houses. Our tabulations from the 2000 Census, for example, show that the houses in which low-English Hispanic children reside are much more likely to be rented (35 percent versus 62 percent for non-Hispanic kids), much less likely to be detached single-family homes (41 percent versus 66 percent), and on average have two fewer rooms (3.5 versus 5.5). In fact, low-English Hispanic families with children look more like non-Hispanic young adults *without* children in their housing consumption.[41] This suggests that, if anything, Mexican immigration may have driven up housing costs for the comparison group by relatively more, possibly biasing us against finding any effect.

Third, a direct investigation suggests that tight markets for family housing are not greatly affecting our estimates. Ideally, we would have some district-level measure of constraints on new construction, such as the initial share of land in a district that is developable (see, for example, Card, Mas, and Rothstein, 2008). Unfortunately, such a measure is not available for 1970. However, we can observe initial population densities at the district level, both overall and by age group, which are strongly correlated with the more sophisticated measures available in later years, but not correlated with our instrument. We find no evidence that the parameter of interest is larger in districts that initially have higher population densities. However, consistent with initial population density measuring constraints on housing supply, initially denser districts see significantly less growth in the representation of non-Hispanic children between 1970 and 2000 (results available on request).

VII.B. Other Amenities

[41] Forty-six percent of non-Hispanic 20-49 year olds without kids lived in owner occupied units and 44 percent lived in detached single family homes in 2000. The average home of this group had 4.5 bedrooms in 2000.

Another competing hypothesis for the effects we estimate is that households with children are more sensitive to immigration along other dimensions. While we justified our choice of comparison group by appealing to survey data showing that young non-Hispanics have similar views about immigration regardless of the presence of children, stated and revealed preferences may differ. For example, even though non-Hispanics with and without children have the same propensity to express that immigration raises crime rates, those with children might be more apt to consider impacts of immigration on crime when choosing a residence. While we cannot entirely rule this out, the same wave of immigration studied here appears to have had no impact on crime rates at the MSA level (Butcher and Piehl, 1998).[42] Mexican immigration may also make a community more desirable, adding to diversity in local restaurants (Mazzolari and Neumark, 2009) or reducing the costs of child care or housekeeping services (Cortes, 2008). The latter in particular should be valued relatively more by households with children and may work against us finding an effect.

VII.C. Private Schooling

Another interpretation of our estimates is that parents might be more sensitive to the presence of low-English children in local playgrounds, at the mall, etc., not in schools *per se*. We cannot rule this out, but unlike decisions about where to live, the decision to attend private school should be driven by the quality of local public schools.

Table VI presents TSLS and OLS estimates of the effect of the EL Hispanic enrollment share on the 1970 to 2000 change in the non-Hispanic private school enrollment rate; as above, the baseline model controls for MSA by district type fixed effects. Both the TSLS estimates (Panel A) and OLS estimates (Panel B) are positive. The TSLS coefficient of roughly 0.18 (column (1)) implies that a 10 percentage point increase in the EL Hispanic share in public school enrollment prompted

[42] Nevertheless, a substantial minority of young non-Hispanics (about 30 percent) express a belief that immigration raises crime. Butcher and Piehl (1998) argue this belief may come from the tendency of immigrants to settle in MSAs that already have high crime rates.

about 1.8 percent of a district's non-Hispanic children to enroll in private school.[43] This estimate suggests that Mexican immigration can account for about 17 percent of the 10.3 percentage point increase in California's non-Hispanic private school enrollment rate over 1970 to 2000. Controlling for the availability of a private school in 1976 – which was correlated with instrument (Table II, Panel C) – raises this estimate somewhat (column (2)), as might be expected, but adding the remaining controls has little impact (column (3)).[44]

These findings offer complementary evidence of an effect of rising low-English enrollment shares on demand for public schools and allow us to paint a more complete picture of how Mexican immigration may have affected non-Hispanic demand for public education in California. Taken together, our estimates suggest that non-Hispanic demand for public education may have been reflected considerably more in the housing market than in the market for private schooling.

VIII. Conclusions

This paper has examined whether the large increase in the low-English Hispanic presence in California public schools since 1970 has reduced non-Hispanic demand for public education, as revealed through changes in the distribution of non-Hispanic children across school districts through 2000. Our empirical approach accounts for endogeneity using 1970 patterns of settlement among Mexican immigrants and compares the district choices of young non-Hispanic households with and without children – who express anti-immigrant sentiments to a similar degree – in an attempt to isolate district choice working through schools. Supporting the credibility of our research design, districts predicted on the basis of 1970 Mexican settlement to receive more low-English

[43] This response is close to the effect that Betts and Fairlie (2003) found at the MSA level for secondary students – that two natives enroll in private school for every 10 immigrant arrivals. Our study differs from theirs in several ways beyond the unit of observation, so this comparison should be made cautiously.
[44] We have also estimated the model for sub-samples defined by center city status and district type. Estimates were statistically indistinguishable across groups (available on request).

Hispanics in their schools through 2000 were not already losing non-Hispanic children in the 1960s, and were comparable to other districts along many observable dimensions.

We find that districts with larger increases in their low-English Hispanic enrollment shares lost more non-Hispanic children than young adults between 1970 and 2000. Not all of this effect can be explained by crowding in the housing market, and districts with larger increases in low-English Hispanic shares saw larger increases in their non-Hispanic private school enrollment rates over 1970 to 2000 as well. Both findings support our interpretation that the loss of non-Hispanic children was brought about at least in part by reductions in demand for public schools. Our focus on California suggests that Mexican immigration may have led to overcrowding in schools or diversion of already limited resources for non-Hispanic students, but our estimates might also reflect distaste for ethnic or socioeconomic diversity in schools, leaving open the question of whether there have in fact been real negative spillovers from immigration for native students.

Regardless, our findings suggest that public education is an important determinant of the residential choices of natives in response to immigration, and in turn, the residential isolation of immigrants. Our findings are akin to “white flight” in response to court-ordered racial desegregation in the preceding years (Reber, 2005; Baum-Snow and Lutz, 2009; Boustan, 2010a) and our estimates are similar in magnitude. As with school desegregation, how this sorting plays out in human capital accumulation is an important topic for future research.

IX. Appendix

School District Level Data: Sources and Construction of Key Variables

1. 1970 and 2000 School District Tabulations

For 1970, school district level data on total population, by age and ethnicity, and private school enrollment, by ethnicity and level, were drawn from the 1970 Fourth Count (Population) School District Data Tapes (U.S. Department of Education, 1970).[45] These data permit identification of all school districts in the country with at least 300 students as of the 1969-70 school year. For 2000, school district level data on total population and private school enrollment, by age and ethnicity, were drawn from the Census 2000 School District Tabulation.[46] All operating districts are included in the age-specific resident counts, but private enrollment counts are missing for districts with 49 or fewer children.

Presentation of the data differs across years. For consistency over time in the definition of our key dependent variable, we aggregate non-Hispanic resident counts to the 0-19 and 20-49 age groups and aggregate these counts to constant secondary district boundaries.[47] To arrive at the dependent variable used in our analysis, we divide by the non-Hispanic MSA population for that age group, generated from county level Census data available at the National Historical Geographic Information System (Minnesota Population Center, 2004). We use a similar approach to calculating the non-Hispanic private school enrollment rate: we first create comparable counts of non-Hispanics enrolled in private school;[48] we next aggregate these counts to consistent secondary district boundaries; and finally, we divide by the aggregated district's 5 to 19 year old population.

From the 1970 data, we also collected information on the distribution of foreign-born Mexicans across school districts, used in construction of the instrument.

2. 1976 and 2000 Office for Civil Rights Data

For 1976, school district level data on the number of EL students, by race/ethnicity, were drawn from the Fall 1976 Elementary and Secondary School Civil Rights Survey, fielded by the Office for Civil Rights in the Department of Health, Education, and Welfare and recently decoded from binary to Stata format by Denckla and Reber (2006). The 1976 OCR survey covered all school districts in

[45] For California residents, the Spanish Heritage population includes "persons of Spanish language or persons not of Spanish language but of Spanish surname identified by matching with a list of about 8,000 such names."

[46] These data are available at <http://nces.ed.gov/surveys/sdds/downloadmain.asp>. To avoid disclosure, cell values are rounded so that exact values cannot be inferred; generally, this rounding is to the nearest 5, or to 4, when the population count is under 5. On a few occasions, rounding leads to (small) negative values.

[47] In 1970, counts of residents by gender were originally reported for the total population and for the "Spanish Heritage" (hereafter referred to as Hispanic) population in detailed age bins (Table 17). In 2000, counts of residents by age and gender were reported for the total population (Table P8 for Total – Population and Households (TT)) and for the Hispanic/Latino population (Table 145H for TT).

[48] The 1970 data reports counts of residents aged 3 to 34 enrolled in private school, by level (kindergarten, elementary, and secondary), for the total population and for the Hispanic population (Table 38). The 2000 data report counts of residents in private school, by gender, separately for all children and for Hispanic/Latino children either enrolled in or of age to be enrolled in the grades served by the district (Tables P8 and 145H for Children (CO): Relevant Children – Enrolled Private). For consistency, we limit counts of non-Hispanic private school enrollees in 1970 to the grade levels served by the district, and aggregate the 2000 figures across age and gender.

the United States. For 2000, school district level data on the number of EL students by ethnicity were drawn from the 2000 Elementary and Secondary School Civil Rights Compliance Report District Survey, fielded by the OCR in the U.S. Department of Education.[49] The 2000 OCR survey covered all school districts in the United States, with tabulations rounded to the nearest 5, to avoid disclosure.

The original data give counts of "pupils whose primary language is other than English" in total and by race/ethnicity. Our treatment variable is constructed using the number of Hispanics (of all races) with this designation and total enrollment. We drop districts for which either of the following held in either 1976 or 2000: (1) the sum of non-EL enrollment by race was more than 10 percent above or below reported non-EL enrollment; or (2) the sum of enrollment by race was more than 10 percent above or below reported enrollment.

3. Other data sources

We use a number of data sources to construct additional district-level covariates. Data on per-pupil property tax revenues and per-pupil total expenditures for 1971-72 are from *Census of Governments, 1972: Government Employment and Finance Files* (U.S. Dept. of Commerce, 1972). The indicator for having an above-average number of public schools as of 1972 given land area and 1976 enrollment (from OCR) was calculated using counts of public schools in 1972 reported in *Elementary and Secondary General Information System (ELSEGIS): Public School District Universe Data, 1972-1973* (U.S. Department of Education, 1973) and land area data information in GeoLytics Neighborhood Change Data Base. Information on center city status was also drawn from U.S. Department of Education (1973). The presence of a private school within district boundaries as of 1976 was constructed from data in *Universe of Private Schools, 1976-1980: Condensed Version* (U.S. Dept. of Education, 1981). All raw data were aggregated to consistent secondary district boundaries for 1970 to 2000 prior to construction of the variables used in the analysis.

Tract-level data on non-Hispanic population for 1960 are from NHGIS (Minnesota Population Center, 2004).

Matching of 1960 Tracts to School Districts

In many cases 1960 and 1970 tract boundaries were identical. In the cases where they were not, we used published Census tabulations of the correspondence between 1960 and 1970 tracts (Table B in US Bureau of the Census, 1972) to construct collections of tracts that could be used to identify the smallest possible identical geographic regions in each census. For example, if a 1960 tract was split into two pieces, we would use that tract in the 1960 data and the aggregate of the two corresponding tracts in the 1970 data. In some cases the overlap between tracts was more complex than this example, but it was almost always possible to construct an exact match by aggregating enough tracts in both years.[50] We then used the *School District Geographic Reference File, 1969-1970* (U.S. Department of Commerce, 1970) to determine the fraction of each tract aggregate's total population inside the borders of each school district in 1970. We apportioned non-Hispanics in each "tract aggregate" to school districts with these weights – which were mostly 0 or 1 – in 1960.

[49] We downloaded these data from <http://www.ed.gov/about/offices/list/ocr/data.html>
[50] The only exception to this was there were a handful of 1970 tracts or parts of tracts on the edge of metro areas that were untracted in 1960.

Bias in Population Model

As noted, the parameter of interest in our estimating equation, $\tilde{\theta}$ in model (3'), likely understates the true effect of interest, θ in model (3), because the comparison group (20-49 year olds) includes a substantial number of parents. To see this, sum (2) across all native households *i* of type *j* in metro area *m* at time *t* to arrive at the number of households of type *j* in district *d* at time *t:*

$$N^{j}_{dmt} = N^{j}_{mt}\left(\gamma_{dt} + \beta_{j}\, k_{dmt}\right) + \sum \eta^{j}_{idmt} .$$

Let τ^{ja}_{mt} represent the average number of individuals in age group *a* (=0-19 or 20-49) per household of type *j* in *m* at *t*. Multiplying the above equation by τ^{ja}_{mt} and summing across household types within age group then generates a model for the (approximate) number of individuals of age *a* in district *d*:

$$N^{0a}_{dmt} + N^{1a}_{dmt} = N^{0a}_{mt}\left(\gamma_{dt} + \beta_{0}\, k_{dmt}\right) + N^{1a}_{mt}\left(\gamma_{dt} + \beta_{1}\, k_{dmt}\right) + \sum\left(\eta^{0a}_{idmt} + \eta^{1a}_{idmt}\right),$$

where $N^{ja}_{dmt} \equiv \tau^{ja}_{mt} N^{j}_{dmt}$ and $N^{ja}_{mt} \equiv \tau^{ja}_{mt} N^{j}_{mt}$. Letting $N^{a}_{dmt} \equiv N^{0a}_{dmt} + N^{1a}_{dmt}$ and $N^{a}_{mt} \equiv N^{0a}_{mt} + N^{1a}_{mt}$, and noting that population aged *a* in *m* at *t* is $N^{a}_{mt} = \sum_{d\in m} N^{a}_{dmt}$, this model can be rewritten as

$$\frac{N^{a}_{dmt}}{N^{a}_{mt}} = \gamma_{dt} + \left(\beta_{0} + \varphi^{a}_{mt}\left(\beta_{1} - \beta_{0}\right)\right) k_{dmt} + u^{a}_{dmt} ,$$

where $\varphi^{a}_{mt} \equiv N^{1a}_{mt} / N^{a}_{mt}$ is the fraction of individuals aged *a* in *m* at *t* who are in households with any 0-19 year olds. Thus, by definition, $\varphi^{0-19}_{mt} = 1$ for all *m* and *t*. To simplify, further assume that there is no variation over time and across metropolitan areas in this parameter for 20-49 year olds, or that $\varphi^{20-49}_{mt} = \varphi^{20-49}$. Differencing over time within age groups, then across age groups then yields:

$$\Delta\frac{N^{0-19}_{dm}}{N^{0-19}_{m}} - \Delta\frac{N^{20-49}_{dm}}{N^{20-49}_{m}} = \left(1 - \varphi^{20-49}\right)\theta\, \Delta k_{d} + \left(\Delta u^{0-19}_{d} - \Delta u^{20-49}_{d}\right)$$

where the parameter of interest is $\theta = \beta_{1} - \beta_{0}$. Our use of population data and all 20-49 year olds as a comparison group thus generates an attenuated estimate of this parameter, with the attenuation bias proportional to the share of 20-49 year olds in households with children.

The logit and growth specifications in Table V are also biased by the factor $\left(1 - \varphi^{20-49}\right)$, and their coefficients are divided by this in the calculation of household-level displacement rates. Converting the estimate in row (d) of Table V to a household-level displacement rate is not as simple, but also depends on φ. (This and other details of our displacement calculations are available on request.)

References

Aaronson, Daniel (1999). "The Effect of School Finance Reform on Population Heterogeneity." *National Tax Journal* 52(1): 5-29.

Alesina, Alberto, Reza Baqir, and Caroline Hoxby (2004). "Political Jurisdictions in Heterogeneous Communities." *Journal of Political Economy* 112(2): 348-396.

Baldassare, Mark (2005). "PPIC Statewide Survey: Special Survey on Education, April 2005." San Francisco: Public Policy Institute of California.

Baum-Snow, Nathaniel and Byron Lutz (2009). "School Desegregation, School Choice and Changes in Residential Location Patterns by Race." Mimeo, Brown University.

Betts, Julian R. (1998). "Educational Crowding Out: Do Immigrants Affect the Educational Attainment of American Minorities?" In *Help or Hindrance? The Economic Implications of Immigration for African-Americans* (ed. D.S. Hamermesh and F.D. Bean), pp. 253-81. New York: Russell Sage Foundation.

Betts, Julian R. and Robert W. Fairlie (2003). "Does Immigration Induce 'Native Flight' from Public Schools into Private Schools?" *Journal of Public Economics* 87: 987-1012.

Boustan, Leah Platt (2010a). "School Desegregation and Urban Change: Evidence from City Boundaries." Mimeo, UCLA.

Boustan, Leah Platt (2010b). "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration." *Quarterly Journal of Economics* 125(1): 417-443.

Brunner, Eric J. and Jon Sonstelie (2006). "California's School Finance Reform: An Experiment in Fiscal Federalism." In *The Tiebout Model at Fifty: Essays in Public Economics in Honor of Wallace Oates* (Ed. William A. Fischel), pp. 55-93. Cambridge, MA: Lincoln Institute for Land Policy.

Butcher, Kristin F. and Anne Morrison Piehl (1998). "Cross-City Evidence on the Relationship between Immigration and Crime." *Journal of Policy Analysis and Management* 17(3): 457-493.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008). "Bootstrapped-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90(3): 414-427.

Card, David (2001). "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration." *Journal of Labor Economics* 19(1): 22-64.

Card, David (2007). "How Immigration Affects U.S. Cities." Centre for Research and Analysis of Migration Discussion Paper No 11/07. London: University College London.

Card, David and John DiNardo (2000). "Do Immigrant Inflows Lead to Native Outflows?" *American Economic Review* 90(2): 360-67.

Card, David and Ethan Lewis (2007). "The Diffusion of Mexican Immigrants During the 1990s." in George Borjas, ed., *Mexican Immigration to the United States*. New York: The University of Chicago Press.

Card, David, Alexandre Mas, and Jesse Rothstein (2008). "Tipping and the Dynamics of Segregation." *The Quarterly Journal of Economics* 123(1): 177-218.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein (2010). "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *The Quarterly Journal of Economics* 125(1): 215-261.

Cortes, Patricia (2008). The Effect of Low-skilled Immigration on US Prices: Evidence from CPI Data. *Journal of Political Economy* 116(3): 381-422.

Cutler, David M; Glaeser, Edward L; Vigdor, Jacob L. (2008a). "When Are Ghettos Bad? Lessons from Immigrant Segregation in the United States." *Journal of Urban Economics* 63(3): May, p. 759-774.

Cutler, David M; Glaeser, Edward L; Vigdor, Jacob L. (2008b). "Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation." *Review of Economics and Statistics*, vol. 90(3): 478-97.

Davis, James Allan and Smith, Tom W. (2009). *General social surveys, 1972-2008* [machine-readable data file] /Principal Investigator, James A. Davis; Director and Co-Principal Investigator, Tom W. Smith; Co-Principal Investigator, Peter V. Marsden; Sponsored by National Science Foundation. --NORC ed.-- Chicago: National Opinion Research Center [producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut [distributor]

Denckla, Ben and Sarah Reber (2006). "Documentation for Decoded OCR Files." Mimeo, University of California, Los Angeles.

Downes, Thomas A. and David Schoeman (1998). "School Finance Reform and Private School Enrollment: Evidence from California." *Journal of Urban Economics* 43: p. 418-443.

Ferreira, Fernando (2009). "*You Can Take It with You:* Proposition 13 Tax Benefits, Residential Mobility, and Willingness to Pay for Housing Amenities." *Journal of Public Economics*, forthcoming.

Gould, Eric D. Victor Lavy and M. Daniele Paserman (2009). "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." *Economic Journal* 119 (October): 1243-1269.

Hoxby, Caroline. 2000. "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review* 90(5): 1209-38.

Kane, Thomas J., Stephanie K. Riegg, and Douglas O. Staiger (2006). "School Quality, Neighborhoods, and Housing Prices," *American Law and Economics Review* 9: 183-212.

Mazzolari, Francesca and David Neumark (2009). "Immigration and Product Diversity." NBER Working Paper 14900. Cambridge, MA: National Bureau of Economic Research.

Minnesota Population Center (2004). National Historical Geographic Information System: Pre-release Version 0.1. Minneapolis, MN: University of Minnesota.

Park, Haeyoun. (2009). "New to English (Interactive Graphic)." *The New York Times* March 13, 2009 < http://www.nytimes.com/interactive/2009/03/13/us/ELL-students.html?ref=education > (accessed August 8, 2009).

Reber, Sarah (2005). "Court-Ordered Desegregation. Successes and Failures Integrating American Schools since *Brown versus Board of Education*." *Journal of Human Resources* 40(3): 559-90.

Rothstein, Jesse M. (2006). "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions." *American Economic Review* 96(4): 1333-1350.

Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek (2010). Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota.

Saiz, Albert (2003). "Room in the Kitchen for the Melting Pot: Immigration and Rental Prices." *Review of Economics and Statistics* 85(3): 502-521.

Saiz, Albert (2007). "Immigration and Housing Rents in American Cities." *Journal of Urban Economics* 61: 345-371.

Saiz, Albert (2010). "On the Geographic Determinants of Housing Supply." *The Quarterly Journal of Economics,* Forthcoming.

Saiz, Albert and Susan Wachter (2006). "Immigration and the Neighborhood." IZA Discussion Paper 2503. Bonn, Germany: Institute for the Study of Labor (IZA).

Tiebout, Charles M. (1956). "A Pure Theory of Local Expenditures."*Journal of Political Economy* 64(5): 416-24.

Urquiola, Miguel (2005). "Does School Choice Lead to Sorting? Evidence from Tiebout Variation." *American Economic Review* 95(4): 1310-1326.

U.S. Dept. of Commerce, Bureau of the Census (1970). *School District Geographic Reference File, 1969-1970* [Computer file]. ICPSR version. Washington, DC: U.S. Dept of Commerce, Bureau of the Census [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2002.

U.S. Dept. of Commerce, Bureau of the Census (1972). *Census of Governments, 1972: Government Employment and Finance Files* [Computer file]. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1993.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-9 Anaheim-Santa Ana-Garden Grove, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-18 Bakersfield, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-75 Fresno, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-117 Los Angeles-Long Beach, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-178 Sacramento, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-187 San Bernadino-Riverside-Ontario, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-188 San Diego, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-189 San Francisco, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-190 San Jose, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-191 Santa Barbara, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-208 Stockton, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census (1972). "Census of Population and Housing: 1970 Census Tracts Final Report PHC(1)-223 Vallejo-Napa, Calif, SMSA." Washington, DC: U.S. Government Printing Office.

U.S. Dept. of Commerce, Bureau of the Census (1981). *Survey of Income and Education, 1976: Rectangular File* [Computer file]. Conducted by U.S. Dept. of Commerce, Bureau of the Census, and University of Michigan, National Chicano Research Network. ICPSR07919-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor].

U.S. Dept. of Education, National Center for Education Statistics (1970). *User's Manual for 1970 Census Fourth Count (Population): School District Data Tape* [Computer file]. ICPSR version. Washington, DC: U.S. Dept. of Education, National Center for Education Statistics [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004. doi:10.3886/ICPSR03525

U.S. Dept. of Education, National Center for Education Statistics (1973). *Elementary and Secondary General Information System (ELSEGIS): Public School District Universe Data, 1972-1973* [Computer file]. ICPSR version. Washington, DC: U.S. Dept. of Education, National Center for Education Statistics [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2000.

U.S. Dept. of Education, National Center for Education Statistics (1981). *Universe of Private Schools, 1976-1980: Condensed Version [United States]* [Computer file]. ICPSR version. Washington, DC: U.S. Dept. of Education, Office of Educational Research and Improvement [producer], 1981. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2000. doi:10.3886/ICPSR02394

Weinstein, Jeffrey (2009). "The Impact of School Racial Compositions on Neighborhood Racial Compositions: Evidence from School Redistricting." Mimeo, Syracuse University.

Table I. Descriptive Statistics for California School Districts

	Mean (1)	St. Dev. (2)
A. Key Variables		
Δ public enr. share low-English Hispanic, 1976-2000	0.107	0.098
Instrument †	0.114	0.110
Δ share of MSA's non-Hispanics in district, 1970-2000:		
0-19 Year Olds (T)	-0.004	0.040
20-49 Year Olds (C)	-0.003	0.044
T-C Difference	-0.001	0.014
Δ Non-Hispanic private enrollment rate, 1970-2000	0.103	0.051
B. Pre-existing levels of outcomes		
Share of MSA's non-Hispanics in district, 1970:		
0-19 Year Olds (T)	0.079	0.120
20-49 Year Olds (C)	0.079	0.125
T-C Difference	0.000	0.013
Non-hispanic private enrollment rate, 1970	0.014	0.015
C. Other pre-existing district characteristics		
public enr. share low-English Hispanic, 1976	0.026	0.044
public school enrollment, 1976	14849	41486
per-pupil property tax revenue, 1971-72	636	292
per-pupil total expenditures, 1971-72	1084	259
=1 if elementary-high organization	0.172	
=1 if above-average #public schools|enrollment, land area, 1972	0.579	
=1 if at least one private school, 1976	0.712	
=1 if center city, 1972	0.094	
Number of Observations	233	
Number of MSAs	23	

Notes: The unit of observation is either a unified school district or a combination of school districts that serve the same geographic area and all elementary and secondary grades (one secondary district plus a number of elementary districts). The sample includes all such observations with complete data on the chacteristics listed. See text for more details on sample construction and description of data sources. † The instrument is the predicted 1976 to 2000 change in low-English Hispanic public school enrollment share arrived at by apportioning low-English Hispanic first- and second-generation immigrants from Mexico and enrolled in public school in 2000 to the school districts where immigrants of all ages from Mexico settled in 1970. An individual is classified as "low-English" if reported to speak English "not at all," "not well," or "well" and "second generation" if native-born to at least one Mexican-born parent (or if both parents are foreign-born, to a Mexican-born mother) who arrived in the U.S. in 1976 or later.

Table II. Are the Identifying Assumptions Satisfied?
The Instrument and District Observables

Dependent Variable	Coefficient (standard error) on the Instrument
A. Change in immigrant share	
Δ public enr. share low-English Hispanic, 1976-2000	0.431***
	(0.099)
B. Pre-existing levels of outcomes	
Share of MSA's non-Hispanics in district, 1970:	
0-19 Year Olds (T)	-0.238
	(0.196)
20-49 Year Olds (C)	-0.242
	(0.206)
T-C Difference	0.005
	(0.017)
Non-hispanic private enrollment rate, 1970	0.000
	(0.007)
C. Other pre-existing district characteristics	
public enr. share low-English Hispanic, 1976	0.253***
	(0.040)
ln(public school enrollment, 1976)	-1.69
	(1.95)
per-pupil property tax revenue, 1971-72	-197.2
	(188.8)
per-pupil total expenditures, 1971-72	-172.5
	(167.6)
=1 if above-average #public schools|enrollment, land area, 1972	0.103
	(0.296)
=1 if at least one private school, 1976	-1.350***
	(0.416)
=1 if center city	0.001
	(0.332)
Number of Observations	233
Number of MSAs	23

Notes: Each entry gives the coefficient (standard error) on the predicted 1976 to 2000 change in low-English Hispanic public school enrollment share arrived at by apportioning low-English Hispanic first- and second-generation immigrants from Mexico and enrolled in public school, in the 2000 Census of Population, to the school districts where immigrants of all ages from Mexico settled in 1970, according to the 1970 Fourth Count (Population) School District Data Tapes. An individual is classified as "low-English" if reported to speak English "not at all," "not well," or "well," and "second generation" if native-born to at least one Mexican-born parent (or if both parents are foreign-born, to a Mexican-born mother) who arrived in the U.S. in 1976 or later. All regressions are based on the full sample of California unified and combined elementary x high school districts (233 districts in 23 MSAs), and include fixed effects for MSA by district type. Standard errors (in parentheses) are calculated to be robust to arbitrary error correlation within MSA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table III. Reduced-Form Estimates for Non-Hispanic District Choice

	Dep. Var.: Δ Share of MSA's non-Hispanic Population in District		
	Ages 0-19 (T)	Ages 20-49 (C)	Difference (T-C)
	(1)	(2)	(3)
	A. Full Sample: 1970-2000		
Coefficient (standard error)	-0.0729	-0.0425	-0.0304**
on instrument	(0.0546)	(0.0548)	(0.0140)
Number of Districts	233	233	233
Number of MSAs	23	23	23
	B. Districts with 1960 Data: 1970-2000		
Coefficient (standard error)	-0.0041	0.0312	-0.0353*
on instrument	(0.0664)	(0.0712)	(0.0163)
Number of Districts	161	161	161
Number of MSAs	13	13	13
	C. 1960-1970		
Coefficient (standard error)	-0.0217	-0.0151	-0.0066
on instrument	(0.0344)	(0.0319)	(0.0093)
Number of Districts	161	161	161
Number of MSAs	13	13	13

Notes: The dependent variable in the first two columns is the change in the share of an MSA's non-Hispanic population residing in a district, for the age group specified in the column and over the period and for the sample of districts specified in the panel. The instrument is the predicted 1976 to 2000 change in low-English Hispanic public school enrollment share arrived at by apportioning low-English Hispanic first- and second-generation immigrants from Mexico and enrolled in public school to the school districts where immigrants of all ages from Mexico settled in 1970. An individual is classified as "low-English" if reported to speak English "not at all," "not well," or "well," and "second generation" if native-born to at least one Mexican-born parent (or if both parents are foreign-born, to a Mexican-born mother) who arrived in the U.S. in 1976 or later. All regressions are based on the based on the full sample of California unified and combined elementary x high school districts, unless otherwise noted, and include fixed effects for MSA by district type. Standard errors (in parentheses) are calculated to be robust to arbitrary error correlation within MSA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IV. TSLS and OLS Estimates for Non-Hispanic District Choice

Dependent Variable:	T-C Difference: Δ Share of MSA's non-Hispanic Population in District, 1970-2000					
Sample:	All (Unified+Combined Elem/High)			Not Center City	Unified	Elem/High
	(1)	(2)	(3)	(4)	(5)	(6)
Mean (C), 1970		0.079		0.060	0.073	0.111
			A. Two Stage Least Squares			
Coefficient (standard error) on:						
Δ public school enrollment share	-0.0705**	-0.0802*	-0.0790*	-0.0674**	-0.0764*	-0.0164
low-English Hispanic, 1976-2000	(0.0318)	(0.0441)	(0.0438)	(0.0288)	(0.0385)	(0.0206)
RMSE	0.0147	0.0147	0.0147	0.0119	0.0153	0.0101
First stage partial F-stat on instrument	19.0	25.8	27.7	18.2	12.4	14.1
			B. Ordinary Least Squares			
Coefficient (standard error) on:						
Δ public school enrollment share	-0.0215**	-0.0035	-0.0040	-0.0183*	-0.0218**	-0.0068
low-English Hispanic, 1976-2000	(0.0100)	(0.0107)	(0.0105)	(0.0101)	(0.0102)	(0.0100)
RMSE	0.014	0.0132	0.0133	0.0111	0.0144	0.0101
Number of Districts	233	233	233	211	193	40
Number of MSAs	23	23	23	23	23	19
Controls:						
MSA fixed effects	X	X	X	X	X	X
District type fixed effect	X	X	X	X		
MSA by district type fixed effects	X	X	X	X		
Pre-existing district characteristics†		X	X			
T-C diff. in dep. var., 1970			X			

Notes: "T" represents 0-19 year olds and "C" represents 20-49 year olds. The first row gives the share of an MSA's non-Hispanic 20-49 year olds residing in the average district in 1970. The instrument used in panel A is the predicted 1976 to 2000 change in low-English Hispanic enrollment share arrived at by apportioning low-English Hispanic first- and second-generation immigrants from Mexico and enrolled in public school in 2000 to the school districts where immigrants of all ages from Mexico settled in 1970. An individual is classified as "low-English" if reported to speak English "not at all," "not well," or "well," and "second generation" if native-born to at least one Mexican-born parent (or if both parents are foreign-born, to a Mexican-born mother) who arrived in the U.S. in 1976 or later. Standard errors (in parentheses) are calculated to be robust to arbitrary error correlation within MSA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively. † Pre-existing district characteristics include: natural log of 1976 total enrollment, per-pupil property tax revenue and per-pupil total expenditure (1971-72), and indicators for center city district (1969-70), above-average number of public schools given land area and enrollment (1972), and with at least one private school (1976).

Table V. Sensitivity of the Estimates for Non-Hispanic District Choice
and Implied Displacement Rates

Dependent Variable:	TSLS (1)	OLS (2)	Displacement Rate (TSLS) Children (3)	HHs (4)
		A. Choice of Outcome		
(a) T-C Diff: Δs_{dm} (Share of MSA non-Hispanic Pop. in District), 1970-2000	-0.0705***	-0.0215**	1.15	1.60
	(0.0318)	(0.0100)		
RMSE	0.0147	0.0140		
(b) T-C Diff: $\Delta \ln(s_{dm}/1-s_{dm})$, 1970-2000	-1.213***	-0.281**	1.44	2.00
	(0.357)	(0.110)		
Marginal Effect (at mean s_{dm},1970)	*-0.0880*	*-0.0204*		
RMSE	0.209	0.191		
(c) T-C Diff: $\Delta \ln(\text{pop}_d)$, 1970-2000	-1.077***	-0.242**	1.39	1.92
	(0.319)	(0.098)		
RMSE	0.195	0.18		
(d) $\Delta \text{pop}_{d,T}/\text{pop}_d$ 1970-2000	-0.250***	-0.062**	1.24	1.67
	(0.077)	(0.024)		
RMSE	0.0445	0.0409		
		B. Weighting		
	Weight: 1970 District Population of non-Hispanic 0-19 Year Olds			
(a) T-C Diff: Δs_{dm} (Share of MSA non-Hispanic Pop. in District), 1970-2000	-0.138***	-0.074**	0.90	1.28
	(0.047)	(0.028)		
RMSE	0.0195	0.0184		
(b) T-C Diff: $\Delta \ln(s_{dm}/1-s_{dm})$, 1970-2000	-0.940***	-0.351***	0.99	1.39
	(0.244)	(0.106)		
Marginal Effect (at weighted mean s_{dm},1970)	*-0.152*	*-0.057*		
RMSE	0.161	0.15		

Notes: "T" represents 0-19 year olds and "C" represents 20-49 year olds. Each entry in columns (1) and (2) represents a different regression. All regressions are based on the full sample of California unified and combined elementary-high school districts (233 districts in 23 MSAs). Column (3) gives an estimate of the number of non-Hispanic child departures for every low-English Hispanic arrival over 1970 to 2000 to a district of average initial (1970) size, based on the TSLS estimate in the row and the assumption that the low-English Hispanic share increased by 0.1. Column (4) gives an estimate of the number of departures of non-Hispanic *households* with children with the arrival of a *household* with at least one low-English child enrolled in public schools. To calculate this figure, we rescale the TSLS coefficients upward - to adjust for the fact that slightly over half of the comparison group of 20-49 year olds are in households with children (see text and Appendix) - and convert the number of low-English Hispanic public school enrollees needed to deliver a 0.1 increase in their public enrollment share into the number of households they represent. Specifications in Panel A are unweighted, and specifications in Panel B are weighted by the 1970 non-Hispanic population of 0-19 year olds. Throughout, the explanatory variable of interest is the 1976 to 2000 change in the share Hispanic low-English in public school enrollment, instrumented (in column (1)) with the predicted change arrived at by apportioning low-English Hispanic first- and second-generation immigrants from Mexico and enrolled in public school in 2000 to the school districts where immigrants of all ages from Mexico settled in 1970. (See notes to earlier tables for more details.) Marginal effects (row (b) of both panels) are calculated by multiplying the regression coefficient by s_{nh}*(1-s_{nh}). Standard errors (in parentheses) are clustered on MSA. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table VI. TSLS and OLS Estimates for Non-Hispanic Private School Enrollment

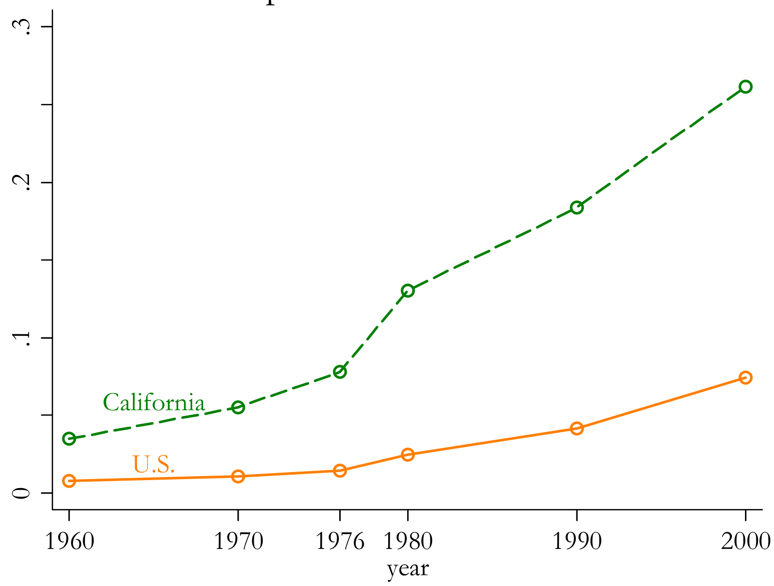
	(1)	(2)	(3)
Dependent Variable:	Δ Share of Non-Hispanics Enrolled in Private School, 1970-2000		
Sample:	All (Unified+Combined Elem/High)		
Mean, 1970		0.0137	
		A. Two Stage Least Squares	
Coefficient (standard error) on:			
Δ public school enrollment share low-English Hispanic, 1976-2000	0.184*	0.260**	0.274**
	(0.089)	(0.100)	(0.104)
RMSE	0.0418	0.0433	0.0427
First stage partial F-stat on instrument	19.0	17.3	27.7
		B. Ordinary Least Squares	
Coefficient (standard error) on:			
Δ public school enrollment share low-English Hispanic, 1976-2000	0.061***	0.074***	0.069**
	(0.028)	(0.029)	(0.031)
RMSE	0.0403	0.0398	0.0391
Number of Districts	233	233	233
Number of MSAs	23	23	23
Controls:			
MSA by district type fixed effects	X	X	X
Has private school, 1976 (=1)		X	X
Other pre-existing district characteristics†			X

Notes: The first row gives the share non-Hispanics of school age attending private school in the average district in 1970. The dependent variable in Panels A and B is the 1970 to 2000 change in the district's non-Hispanic private school enrollment rate. The instrument used in panel A is the predicted 1976 to 2000 change in low-English Hispanic enrollment share arrived at by apportioning low-English Hispanic first- and second-generation immigrants from Mexico and enrolled in public school in 2000 to the school districts where immigrants of all ages from Mexico settled in 1970. See notes to earlier tables and text for more details. Standard errors (in parentheses) are calculated to be robust to arbitrary error correlation within MSA (23 MSAs). ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively. † Other pre-existing district characteristics include: the 1970 non-Hispanic private school enrollment rate, the natural log of 1976 total enrollment, per-pupil property tax revenue and per-pupil total expenditure (1971-72), and indicators for center city district (1969-70) and above-average number of public schools given land area and enrollment (1972).

Appendix Table I. Construction of the Instrument

Share of Mexicans in 1970: Rank	Value	District	MSA	Instrument: Rank	Value
		A. CA Districts Ranked by Share of U.S. Mexican-Born Population, 1970			
1	0.1626	LOS ANGELES UNIF	Los Angeles, CA	27	0.2539
2	0.0124	SAN DIEGO CITY UNIF	San Diego, CA	84	0.1136
3	0.0094	SWEETWATER UNION HIGH	San Diego, CA	13	0.3028
4	0.0091	SAN FRANCISCO UNIF	San Francisco, CA	67	0.1435
5	0.0066	SANTA ANA UNIF	Anaheim, CA	42	0.2013
6	0.0064	EL MONTE UNION HIGH	Los Angeles, CA	4	0.5145
7	0.0058	ALHAMBRA CITY ELEM-HIGH	Los Angeles, CA	14	0.3015
8	0.0051	EAST SIDE UNION HIGH	San Jose, CA	33	0.2385
9	0.0050	SAN JOSE UNIF	San Jose, CA	69	0.1406
10	0.0045	OAKLAND CITY UNIF	Oakland, CA	95	0.0961
		B. CA Districts Ranked by Predicted Change in Hispanic low-English Share in Public School Enrollment, 1976-2000			
95	0.0006	LE GRAND UNION HIGH	Merced, CA	1	0.6363
26	0.0023	SANTA PAULA UNION HIGH	Oxnard-Ventura, CA	2	0.6147
36	0.0018	DELANO JOINT UNION HIGH	Bakersfield, CA	3	0.5401
6	0.0064	EL MONTE UNION HIGH	Los Angeles, CA	4	0.5145
77	0.0007	LIBERTY UNION HIGH	Oakland, CA	5	0.4169
61	0.0010	FALLBROOK UNION HIGH	San Diego, CA	6	0.3987
44	0.0015	DINUBA UNIFIED	Visalia, CA	7	0.3759
55	0.0012	CUTLER-OROSI UNIF	Visalia, CA	8	0.3701
127	0.0003	KINGSBURG JT UNION HIGH	Fresno, CA	9	0.3363
54	0.0012	FILLMORE UNIF	Oxnard-Ventura, CA	10	0.3311

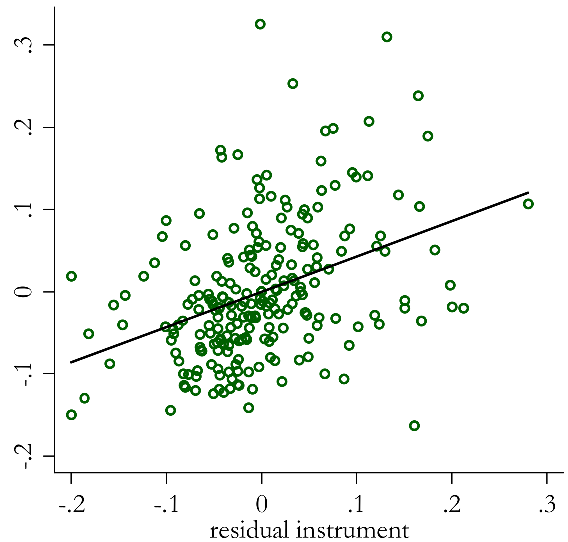
Figure I. Trends in the Mexican Share of the Child Population in the United States



Note: Sources are SIE (1976) and Census PUMS. Sample includes 0-17 year olds and Mexicans include both the Mexican-born and the children of Mexican-born parents.

Figure II. First-Stage Relationship

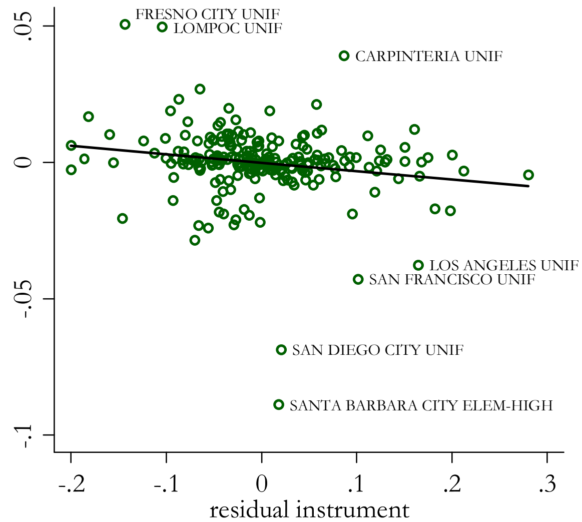
1970 to 2000 Change in ELL Hispanic Pub. Enr. Share



Notes: Regression adjusted for MSA x district type fixed effects. Sample includes unified and combined HS-elem districts in 23 MSAs in CA.

Figure III. Reduced-Form Relationship

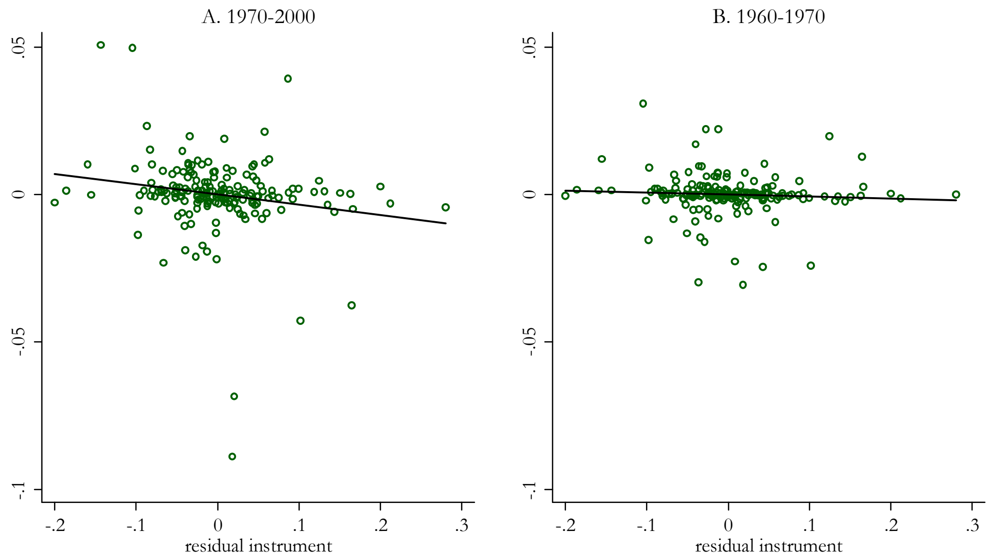
T-C: 1970-2000 Change in Share of MSA's non-Hispanics in District



Notes: T=0-19 year olds. C=20-49 year olds. Regression adjusted for MSA x district type fixed effects. Sample includes unified and combined HS-elem districts in 23 MSAs in CA.

Figure IV. Pre-Trends in Outcomes?

T-C Diff: Change in Share of MSA's non-Hispanics in District



Notes: T=0-19 year olds. C=20-49 year olds. Regression adjusted for MSA x district type fixed effects. Sample includes 13 MSAs in CA with 1960 data available