Preliminary
Do Not Cite Without Permission

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

Elizabeth U. Cascio
Dartmouth College, IZA, and NBER

Ethan G. Lewis
Dartmouth College

March 23, 2010

Abstract

This paper examines whether the large wave of low-English Hispanic immigration to the United States since 1970 has lowered native demand for public schooling. Our analysis focuses on California – where many of these immigrants settled – accounts for possible endogeneity of immigrant inflows using established settlement patterns, and uses relative outflows of the school-aged population to identify relocation in response to immigration-induced changes in school quality. We find that between 1970 and 2000, the average metropolitan school district in California lost five non-Hispanic students – three to relocation to another school district and two to private school within district – for every ten additional low-English Hispanic arrivals in its public schools. Districts that initially had fewer public schools lost more non-Hispanics to relocation, while private school flight was concentrated in districts with an initial presence of private schools. Our estimates suggest that the decline over this period in non-Hispanic public enrollment share in the average low-English Hispanic child's school district would have been 25 percent smaller in the absence of flight.

---

* 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 would like to thank Bill Fischel, Doug Staiger, Tara Watson and seminar participants at Dartmouth College, Cornell University, and the 2010 American Economic Association Annual Meetings for useful comments. We also gratefully acknowledge the research assistance of Christopher Bachand-Parente and funding from Dartmouth College. All errors are our own.

I. Introduction

Low-skilled immigration has arguably brought about the most striking change to school demographics in the United States since the federal intervention to desegregate schools. In 2000, there were 4.1 million children of school age with limited English proficiency, representing 6.7 percent of the school-aged population nationwide and a four-fold increase in the population from its 1976 level (Figure 1). This growth has been fueled not only by the immigration of children, but also by the formation of relatively large families among those who immigrate as adults. First-generation immigrants in fact accounted for less than half (39 percent) of the limited English proficient (LEP, or "low English") school-aged population by 2000.[1] Nearly all growth in public school enrollment in the last decade can be accounted for by LEP students (Park, 2009).

Like school desegregation, this most recent wave of immigration has fundamentally altered the education production function, and in doing so, potentially lowered native demand for public schooling where immigrants have chosen to settle.[2] Such "native flight" would limit the capacity of the public school system to assimilate LEP schoolchildren, and when manifested as residential change – as opposed to rising private school enrollment – would isolate their parents from natives as well. If it prompts stronger responses on the part of native white families, immigration may also be contributing to school "resegregation," reducing exposure of blacks to whites in public schools.

This paper examines whether immigration has in fact lowered native demand for public schooling and how these declines in enrollment have been manifested – as population losses from a school district or as increases in private school enrollment within its boundaries. We center our

[1] Statistics in this paragraph (and those on California below) are the authors' calculations from the 5 percent Public Use Microdata Sample (PUMS) of the 2000 Census and from the 1976 Survey of Income and Education (SIE). We classify as "low English" any individual who "does not speak English," "yes [speaks English], but not well," or "Yes [speaks English], speaks well," since this definition generates shares low-English in the school-aged population that are comparable in the SIE and the administrative data source used in our analysis.

[2] While the present study is the first to examine flight both to other school districts and to private schools in response to immigration, a more extensive literature examines white flight in response to racial desegregation of schools. Reber (2005) uncovers long-term declines in white public school enrollment after court-ordered desegregation in metropolitan areas across the country. More recently, Baum-Snow and Lutz (2009) draw from the same sample of districts (Welch and Light, 1987) to dissect these declines into increases in private school enrollment and relocation to the suburbs. Relatedly, Boustan (2009) finds declines in house prices and rents in non-Southern center city districts after desegregation plans were implemented.

analysis on the responses of non-Hispanics to the large wave of low-English Hispanic immigration to California since the 1970s. We restrict attention to one state to hold constant the institutional environment. California institutions may work against finding an effect. Since *Serrano v. Priest*, for example, overall per-pupil spending has been equalized across California districts, and tax rates have in effect been equalized since passage of Proposition 13 in 1978. As a result, immigration should not have prompted flight by lowering per-pupil spending. And until passage of Proposition 227 in 1998,[3] California schools offered bilingual instruction, limiting direct interactions between Hispanic LEP and non-Hispanic schoolchildren. As suggested by Figure 1, California has also received larger immigration inflows than any other state: By 2000, the state was home to 36.2 percent of all school-aged low-English Hispanics nationwide, and LEP Hispanics represented 12.6 percent of the state's school-aged population.[4]

We face two identification problems in our analysis. First, immigrant settlement is not random. Lower housing costs may make a declining school district attractive for an immigrant family, or unobserved factors might make a school district more or less attractive for natives and immigrants alike. Either way, simple correlations will yield biased estimates of the effect of immigration on native demand for local public schools. Second, we are interested in identifying relocation in response to immigration-induced changes in public school quality. However, immigration may affect other (non-school) amenities associated with residence in a school district and may raise housing costs. While this should not matter for private schooling decisions, native families may find it optimal to relocate in response to these developments alone.

We approach the first of these identification problems by constructing an "enclave-based" instrument for the influx of low-English Hispanics to a district, the intuition being that school districts with pre-existing Hispanic immigrant enclaves – even if quite small – should have been

[3] Standards for English development were not established until July 1999.
[4] LEP Hispanics represented three percent of the state's school-aged population in 1976. As of 2000, California had 12.7 percent of the country's school-aged population overall, and 41.9 percent of California's school-aged population was Hispanic (authors' calculations from the 1976 SIE and 2000 PUMS).

attractive for new immigrant families but otherwise similar in their propensity for subsequent non-Hispanic flight. We then embed this instrumental variables strategy in a differences-in-differences model where we compare outflows of the school aged population – where families arguably place greater value on the quality of public schools – to those of slightly younger and slightly older groups from the same district. That is, we test whether low-English Hispanic immigration led to *relatively* large outflows of non-Hispanics of school age from a district.

Using this identification strategy and district-level population and enrollment data from the U.S. Census and other sources, we find that between 1970 and 2000, the average metropolitan school district in California lost five non-Hispanic students – three to other school districts and two to private schools within district – for every ten additional LEP Hispanic arrivals in its schools.[5] Differences across school districts in the extent of non-Hispanic flight arise *within* the suburbs of metropolitan areas, suggesting that our estimates are not confounded by some common shock to the quality of public schools in a metropolitan area, or by suburbanization. Supporting causal interpretation of our estimates, we find non-Hispanic relocation to be strongest for the age groups a district serves and in districts with fewer public schools *ex ante*, which offer less scope for within-district sorting. We also find greater private school flight in districts where an existing presence of private schools presumably reduced the costs of enrollment.

These findings suggest that native flight has contributed substantially to rising ethnic segregation in public schools in recent years. A back-of-envelope calculation based on our estimates suggests that non-Hispanic relocation and private school enrollment can account for 27 percent (5.7 percentage points) of the 19 percentage point decline over the period of study in the non-Hispanic public enrollment share in the district attended by the typical low-English Hispanic child. On the other hand, we find no evidence that blacks had a lower propensity for flight than other non-Hispanics, suggesting that immigration may not account for rising segregation along racial lines.

[5] Our estimates for flight to private schools are comparable to those found (for high school aged students only) by Betts and Fairlie (2003), who estimate the effects of first-generation immigrants on private school enrollment at the metropolitan area level during the 1980s. We reconcile our findings with theirs below.

However, our estimates may understate immigration-induced reductions in exposure of both LEP Hispanics and blacks to (other) non-Hispanics, since they do not capture what could be substantial residential sorting across school attendance areas within school districts.

The paper proceeds as follows. The next section gives theoretical background that motivates our empirical strategy. Section III outlines our data and key variables. Sections IV and V present the findings for non-Hispanic relocation and private school enrollment, respectively. Section VI describes the implications of our estimates, and Section VII concludes.

II. Theoretical Framework

The objective of our empirical analysis is to test whether immigration-induced changes in public school quality have been an important mediator of the residential and private school enrollment choices of natives over the past few decades.[6] The framework presented in this section, drawing from the insights of Tiebout (1956), illustrates the ways in which immigration may affect these choices more generally and highlights the conditions under which we might plausibility isolate the contribution of immigration-induced shocks to school quality.

II.A. Isolating the Location Response to Changes in Public School Quality

For simplicity, we begin by assuming that there is no private school sector. Let the indirect utility, V, associated with residence in a particular school district be a function of (public) school quality, q, all other local amenities, g, and housing costs, p. A household of type j will choose to reside in a particular school district provided that the resulting utility is at least as large as that associated with residence in the best alternative school district, v:

$$V^{j}\left(p,q,g\right)\geq v. \tag{1}$$

For all groups j, V is (weakly) decreasing in p and (weakly) increasing in both q and g.

Now suppose that a district receives an influx of immigrant schoolchildren, ∂I. In general, a native household with school-aged children will choose to move to another district if V falls below

[6] For ease of exposition, we refer to low-English Hispanics as immigrants and non-Hispanics as natives in this section.

v. Immigration may prompt such a move by reducing school quality. However, immigration may induce cross-district moves for other reasons. For example, immigration may reduce other amenities associated with living in the district, potentially reducing V. If the housing stock is not perfectly elastic, the increase in population will also raise housing costs, all else constant, again potentially reducing V.[7] The housing market returns to equilibrium when p adjusts sufficiently to restore (1) for all groups remaining in the district.

Thus, the reduced-form relationship between population flows of school-aged immigrants and natives across school district lines does not immediately reveal that natives are fleeing from immigration-induced changes in q. To isolate the contribution of school quality, we make the additional assumption that households with school-aged children (j=1) value school quality more than those without school-aged children (j=0). If $\partial V^1 / \partial q > \partial V^0 / \partial q$, and if the disutility associated with increases in p and reductions in g does not vary systematically with the presence of school-aged children, there will be relatively large outflows of native households with school-aged children only if $\left(\partial V^1 / \partial q\right)\left(\partial q / \partial I\right) < 0$, or only if immigration reduces school quality. Our baseline empirical model is thus designed to test whether school districts that experienced more immigration also experienced *relatively* large declines in the native population of school age – that is, relative to declines in the district's native population at younger and slightly older ages.[8]

We examine the internal validity of the resulting estimates in several ways. All else constant, flight should be stronger from districts where the immigration-induced shock to school quality is greater (i.e., where $\partial q / \partial I$ is larger in magnitude). On this front, districts with fewer public schools offer less scope for sorting across school attendance zones within district boundaries (Urquiola, 2005), limiting the extent to which the effects of immigration on school quality can be contained.[9]

---

[7] For example, Saiz (2003, 2007) uncovers evidence that immigration drives up rents in destination metropolitan areas.

[8] Baum-Snow and Lutz (2009) and Boustan (2009) implicitly make a similar assumption in studying the effects of court-ordered desegregation, though we offer a tighter definition of school age. See below.

[9] Similarly, v is theoretically (weakly) increasing in the availability of alternative school districts in the household's choice set. This choice set has been defined by metropolitan area (e.g., a labor market) in previous applications (e.g., Hoxby,

Thus, we expect to find stronger outflows of the school-aged population in response to a given immigrant influx from districts with fewer schools. Second, since California has mix of unified, elementary, and secondary districts, we can examine whether the ages of any outflows vary appropriately with district type, potentially falsifying the assumption that parents of school-aged children place more value on school quality.

II.B. Private Schools

We incorporate the private sector into the model as follows. Assume that there is available a private school of quality q^{priv}, but there is a cost, c, of attending it.[10] The cost includes both any tuition and time and transportation costs. Thus, if the supply of private schools is geographically uneven, the cost of attending private school will depend in part on district of residence. For simplicity, we assume that immigration affects neither q^{priv} nor c.

In this model, a family with school-aged children will reside in a school district but attend private school if:

(2a) $$V^1\left(p+c, q^{priv}, g\right) \geq v, V^1\left(p, q, g\right),$$

and they will attend the district's public schools if:

(2b) $$V^1\left(p, q, g\right) \geq v, V^1\left(p+c, q^{priv}, g\right).$$

For families who choose public schooling, the cost and quality of private schooling do not enter utility directly, though in a broader model, might have some option value. We assume this option value is small relative to other sources of utility differences across districts.

The predicted effects of immigration in this model are similar to those in the previous section, but we highlight three additional implications. First, the lower the cost of private schooling, c, the larger the increase in native enrollment in private schools will be for a given immigration

---

2000; Urquiola, 2005; Rothstein, 2006). The small number of metropolitan areas in our sample and a lack of variation across these metropolitan areas in district concentration rendered uninformative our attempts to test this prediction.

[10] Note that the cost and quality of this private school might derive from a utility-maximizing choice over more than one private school. Let there be N private schools with costs $c_1, \ldots c_N$ and quality $q_1^{priv}, \ldots, q_N^{priv}$. For every private school, define $\hat{v}_i = V^1\left(p+c_i, q_i^{priv}, g\right)$. We define $c \equiv c_{i*}$ and and $q^{priv} \equiv q_{i*}^{priv}$ where i* is the private school with the largest $\hat{v}_i$.

shock, ∂I . Second, the lower the cost of private schooling, the smaller the population outflows will be for a given immigration shock. Third, flight to private schools within a district may be reduced when there are more public schools from which to choose.

Thus, as an additional internal validity check, we test whether flight to private schools was sensitive to the availability of public schools. We also test whether the relationships we estimate are sensitive to the cost of private schooling, which we proxy for with the presence of a private school inside the boundaries of the school district. There is substantial heterogeneity across districts in the availability of a private school: Thirty-five percent of districts did not have a private school within their borders in the late 1970s.[11]

II.C. Endogeneity of Immigration Flows

We face an additional complication in our analysis: endogeneity of immigration flows. For example, declining school quality for *other* reasons may induce native families to leave the public school system. Immigrant families may be attracted to these school districts through lower housing costs, particularly if they tend to place relatively less value on schools in their location decisions.[12] In this case, native flight from a district's public schools will generate immigrant inflows, not vice versa. Alternatively, a positive shock to school quality might simultaneously attract both natives and immigrants, possibly generating a positive correlation between changes in the school-aged native and immigrant populations. This highlights the importance of identifying an exogenous source of variation in immigrant inflows. We introduce such an approach in the next section.

III. Data

Due to data constraints, we cannot investigate the effects of school-aged immigration *per se*. Rather, we focus on how increases in the public school enrollment of limited English proficient

[11] This is not to say that the market for private schools could not be larger than the boundaries of a school district. However, previous findings indicate parents are generally not willing to travel far to send their kids to school (e.g., Hastings and Weinstein, 2008). If the presence of a private school is an inadequate proxy for costs, we will not find that our results are sensitive to it. Thus, we are performing a joint test of the model and the quality of our proxy.

[12] For example, Boustan (2010) shows that the foreign-born were attracted to center cities that whites had earlier fled in response to black in-migration.

Hispanics have affected the residential choices and private school enrollment of non-Hispanics. We identify these effects using the inflow of first- and second-generation LEP Hispanic immigrant schoolchildren to a district, predicted based in part on prior settlement patterns of Hispanic immigrants. This section discusses our data sources and defines our estimation sample and key variables; see the Data Appendix for more detail.

III.A. Primary Sources and Sample

The primary data sources for our analysis are the 1970 and 2000 school-district tabulations of the U.S. Census of Population (hereafter referred to as the School District Data Book, or SDDB), and the 1976 and 2000 Elementary and Secondary School Civil Rights Surveys, conducted by the Office for Civil Rights in the Department of Health, Education, and Welfare (later the Department of Education, hereafter referred to as the OCR) to monitor compliance with federal civil rights law. The SDDB provides the information on non-Hispanic population by age group and private school enrollment at the school-district level that is critical for our analysis. The OCR identifies numbers of public school students at the district level, by race, "in need of" (but not necessarily enrolled in) specialized classes for English instruction.[13],[14] The variables that we construct from these sources are discussed in detail below.

We restrict our analysis to school districts in 22 metropolitan statistical areas (MSAs) in California.[15] For our analysis, we define "districts" so as to have constant boundaries between 1970 and 2000, and aggregate key variables accordingly.[16] We lose "aggregated" school districts for

[13] The first year of the OCR in which questions on LEP students were asked of all districts was 1976. As discussed below, most of the 1970 to 2000 increase in the school-aged LEP Hispanic population occurred after 1976, so this likely has little effect on our findings. The 1980 SDDB also lacks sufficient disaggregation of population counts by age and ethnicity to apply our empirical strategy.
[14] The Equal Educational Opportunity Act of 1974 defined as a denial of equal educational opportunity "the failure by an educational agency to take appropriate action to overcome language barriers that impede equal participation by students in an instructional program." The Department of Health, Education, and Welfare set forth guidelines for accommodation of LEP students and began monitoring district compliance in 1975.
[15] We use 1990 definitions of Standard Metropolitan Statistical Areas. Our sample encompasses all but one California MSA – Fresno – where we had little confidence in our ability to track district reorganizations over time.
[16] So, 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. Similarly, if district A splits into districts B and C by 2000, we aggregate B and C to create an observation for A in 2000. We identify school district reorganizations using data from the Elementary and Secondary Education General

several reasons. First, while both the SDDB and the OCR are in principle censuses of school districts, the 1970 SDDB 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.[17] By and large, most sample drops on these grounds occur because a district is missing in the 1970 SDDB.

We make several additional exclusions to arrive at our estimation sample. As noted, 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 generally feed into secondary districts. For reasons outlined below, our estimation sample includes just elementary and unified districts, of which there are 256 and 196, respectively.[18] To avoid concerns about suburbanization driving our results, we drop center city districts for most of our analysis. In this subsample, there are 177 unified and 251 elementary districts. Because they tend to be relatively small (and so below the 300 student threshold in the 1970 SDDB), elementary districts are underrepresented.[19]

III.B. Key Variables

The "treatment" of interest in this study is the district-level influx of LEP Hispanic public school students between 1976 and 2000, $\Delta LEPH_d$, constructed using the enrollment figures from the OCR. As noted above, however, identifying the effect of this treatment is challenging, since rising enrollments of LEP Hispanics may be correlated with unobserved determinants of departures of non-Hispanics from the public school system. We deal with this identification problem by using

---

Information System (ELSEGIS) and the Common Core of Data Public Agency Universe. By and large, the district reorganizations observed over this period involve unification of elementary and secondary school districts.

[17] We drop districts for which either of the following holds in either 1976 or 2000: (1) the sum of non-LEP enrollment by race was more than 10 percent above or below reported non-LEP enrollment; or (2) the sum of enrollment by race was more than 10 percent above or below reported enrollment.

[18] Younger ages may not provide a valid comparison for secondary districts, because immigrant arrivals in secondary districts are likely to be correlated with arrivals in the elementary districts that feed them. Ideally, we would have historical information on which elementary districts fed which high schools to address this problem; in practice, such information is difficult to come by. We discuss this below in reference to Table 4.

[19] The typical metropolitan area in our sample had 34 (un-aggregated) school districts at the beginning of the period: 9 unified, 22 elementary, and 3 secondary. Since residential flight is larger for districts with fewer public schools (Table 5), the omission of small elementary districts from our sample likely biases downward the magnitude of our estimates.

an instrumental variables approach that has been previously been used to examine the impacts of immigration on native flight at the MSA level (e.g., Card and Dinardo, 2000; Card, 2001). The intuition behind this approach that new immigrants are attracted to areas where there is already a presence of their countrymen, but such areas are similar in the propensity for subsequent native flight.

In particular, our instrument, Z_d, is the *predicted* LEP Hispanic public school student inflow to district *d* based on pre-existing settlement patterns of Hispanic immigrants. This prediction of $\Delta LEPH_d$ is given by:

$$(3) \qquad Z_d = \sum_g \frac{M_{dg}^{1970}}{M_g^{1970}} \Delta LEPH_g \, ,$$

where $M_{dg}^{1970} / M_g^{1970}$ is the share of the U.S. population born in country *g* and residing in district *d* in 1970, based on tabulations from the 1970 SDDB,[20] and $\Delta LEPH_g$ represents LEP Hispanic arrivals from country *g* between 1976 and 2000 who attend public schools at grade levels served by *d*, based on calculations from the 2000 Census Public Use Microdata Sample. We include in $\Delta LEPH_g$ both children who were born in *g* and children for whom at least one parent was born in *g* and arrived in 1976 or later. As suggested above, the second generation accounted for a substantial share of LEP Hispanic schoolchildren in California in 2000.[21] We classify as LEP any child who does not speak English at all, speaks English but not very well, or who speaks English well (but not very well), since this definition generates similar shares low-English in the OCR and in survey data as of 1976.

We use the 1970 and 2000 SDDB to construct our dependent variables – district-level changes in the counts of non-Hispanic population and of non-Hispanic private school enrollees. For consistency across years and ease of interpretation in our analysis, we measure the first by aggregating district-level population counts of non-Hispanics under the age of 25 into five-year age

[20] Most Hispanic immigrants in 1970 are from the following countries (or county groups) observed in the 1970 SDDB: Mexico, other Latin American (including Caribbean), Cuba, and Southern Europe.
[21] Bleakley and Chin (2008) show that children native-born to immigrants who arrived after age 9 (or after the "critical period" for language acquisition) were themselves more likely to be low English.

bands. For unified districts, the three five-year age bands spanning ages 5 to 19 are considered "school age," while for elementary districts, ages 5 to 14 are school age; the remaining age groups are used for comparison purposes in the analysis below. Changes in district-level non-Hispanic private school enrollment pertain only to the grade levels that the district serves.

Ideally, we would be able to observe changes in non-Hispanic population and private school enrollment by race. Unfortunately, the 1970 SDDB does not report population and private school enrollment counts by both race *and* ethnicity. However, the data do permit us to construct comparable measures for blacks, the vast majority of whom are non-Hispanic.[22] This allows us to gain insight into whether the non-Hispanic flight that we estimate has reduced the exposure of blacks to non-Hispanics of other races in public schools.

III.C. Descriptive Statistics

Table 1 gives summary statistics for key variables from our main estimation sample of non-center city elementary and secondary school districts.[23] Panel A provides statistics on public school enrollment based on the OCR data. As shown in the first row, the average district in our sample experienced a five-fold increase in LEP Hispanic enrollment between 1976 and 2000, from 194 to 1,179 students. Panel B shows that the predicted change in Hispanic LEP enrollment, based on equation (3), is smaller (at 577 students) than that which the average district actually experienced. This is to be expected, as our instrument predicts only that part of the growth low-English Hispanic enrollment driven by first- and second-generation immigrants.

The remaining rows of Panel A help to put these figures into perspective. Hispanic enrollment grew threefold in the average district of our sample over this period (from 900 to 2773). On average, more than half of this enrollment change can be accounted for by low-English Hispanics, and by our prediction, over half of this by recent immigrants. In contrast, non-Hispanic

[22] In 1976, 0.86 percent of school-aged blacks were Hispanic. In 2000, 2.1 percent of school aged blacks were Hispanic (authors' calculations from the 1976 SIE and the 5 percent 2000 Census PUMS).
[23] Statistics for the larger sample of all California districts for which we have obtained data are shown in Appendix Table 1. Unsurprisingly, including these districts raises the average district's size and the magnitude of its population losses.

enrollment hardly grew at all. Indeed, in the average district, 91 percent of public school enrollment growth is driven by Hispanics and nearly half by low-English Hispanics. By 2000, LEP Hispanics represented 16.3 percent of the average district's public school enrollment, compared to only 3.7 percent in 1976. Thus, even if there had been no LEP Hispanic schoolchildren in California in 1970, the lion's share of growth in this population between 1970 and 2000 would have occurred after 1976.[24]

As shown in Panel C, the lack of growth in non-Hispanic public school enrollment is mirrored by a lack of growth in non-Hispanic population between 1970 and 2000, especially among those of school age. However, these means mask would could be substantial movement of non-Hispanics across districts in the sample in response to changes in school quality. More suggestive of such moves is the near six-fold increase in non-Hispanic enrollment in private schools over this period. Blacks saw similarly high rates of growth in private school enrollment in the period – and increases in population – but their baseline population shares outside of the central city are lower than the state average (see Appendix Table I).

IV. Residential Flight

IV.A. Empirical Approach

Our empirical approach to estimating the population response to Hispanic LEP arrivals is similar in spirit to that used in Boustan's (2010) study of the effect of black migration on white flight from non-Southern center cities. The primary innovation is that we use non-school-aged individuals as a comparison group to isolate migration in response to immigration-induced changes in school quality.

[24] Weighted by public school enrollment, the LEP Hispanic share in public school enrollment in our sample in 1976 is lower, at 2.9 percent. This figure is quite similar to what we see for metropolitan California in the 1976 SIE (2.8 percent) when we classify as low-English those respondents who did not speak English at all, not well, or well. On the other hand, only 13.1 percent of the metropolitan California public school enrollees are LEP by the same definition in the 2000 Census, compared to a (weighted) mean of 16 percent in our sample. This suggests that our sample is weighted toward districts with more growth in LEP Hispanics over the period, or that the definition of LEP used by schools has become inclusive of non-native English speakers who speak English very well. If the latter, recall that our instrument is based on the Census definition.

Estimated on the population change data for individuals under age 25, our model is given by:

(4) $$\Delta NH_{ad} = \theta 1[a \in SA_d] \cdot \Delta LEPH_d + \gamma_d + \lambda_a + \varepsilon_{ad},$$

where ΔNH_{ad} represents the 1970 to 2000 change in the non-Hispanic population for five-year age group *a* in district *d*; $\Delta LEPH_d$ is the 1976 to 2000 change in the number of low-English Hispanic public school students in district *d*;[25] $1[a \in SA_d]$ is an indicator set to one if age group *a* is of school age for district *d*; and γ_d and λ_a represent vectors of district and age-group fixed effects. Because the model is in first differences, the γ_d account for unobserved district-level determinants of trends in the non-Hispanic population common to all age groups under observation. As suggested above, such unobservables might include changes in other amenities associated with residing in the district or changes in housing costs. However, the district fixed effects absorb the effects of *any* district observable that we might consider including as a control in the model. The age effects absorb statewide changes in the age composition of non-Hispanics.

The coefficient of interest in model (4) is θ, which gives *how many more* non-Hispanics of school age left per low-English Hispanic arrival in the public schools of the average district, relative to what would have been expected given demographic change in California. We choose comparison age groups – those slightly below and slightly above school age – so that the γ_d plausibly capture what would have happened for the population of school-aged non-Hispanics in the absence of inflows of LEP Hispanics into the local public schools. Note that if families moved in response to immigration-induced shocks to school quality when their children were not yet school age (ages 0 to 4), this will tend to bias our estimates downward. That said, we show below that there were similar

[25] To account for the fact that the school-aged population spans different age categories depending on the district, we rescale the actual inflow of LEP Hispanic students to a district by the number of age groups it spans. That is, we divide the inflow by three for unified districts (ages 5 to 9, 10 to 14, 15 to 19) and by two for elementary districts (ages 5 to 9, 10 to 14). If population counts were available by single year of age, this normalization would make $\Delta LEPH_d$ the average number of LEP Hispanic students per grade (year). Instead, it is the average number of LEP students per five-year age span, consistent with variation in the dependent variable.

population changes for the different comparison age groups in the average district; using multiple comparison groups improves the precision of our estimates.

We estimate (4) using both ordinary least squares (OLS) and two-stage least squares (TSLS), using $1[a \in SA_d] \cdot Z_d$ (with Z_d defined in equation (3)) as an instrument for $1[a \in SA_d] \cdot \Delta LEPH_d$. TSLS estimates of θ will be identified if the predicted inflows of LEP Hispanic schoolchildren to a district are otherwise unrelated to school-aged non-Hispanic departures. Intuitively, it must be the case that established Hispanic immigrant settlement patterns do not predict subsequent (unobserved) shocks to school quality.[26] Unfortunately, it is impossible for us to test this assumption directly, and the informal tests that we would ideally perform – specifically, testing whether $1[a \in SA_d] \cdot \Delta Z_d$ is correlated with *prior* outflows of school-aged non-Hispanics from a district – are not possible given the lack of population tabulations at the school district level (or by Hispanic background) prior to 1970. However, even if school districts with some Spanish-speaking immigrants were already declining in quality by 1970, the sheer magnitude of immigration in the ensuing decades was arguably unforeseeable.

IV.B. Baesline Findings

Table 2 presents TSLS estimates (in Panel A) and OLS estimates (in Panel B) of θ from model (4). Estimates are based on the pooled elementary-unified sample in columns (1) - (3) and are shown separately by district type in columns (4) and (5). The corresponding first-stage regression estimates are reported in Panel C.[27] In the first stage, the coefficients on the instrument tend to be

[26] One might also be concerned that Z_d is correlated with increases in the enrollment of other demographics, in which case our estimates would not reflect displacement driven purely by LEP Hispanics. Existing settlement patterns of foreign-born Hispanics, for example, are also likely to predict changes in enrollment of Hispanics *not* in need of English instruction, or the arrival of other immigrant groups besides Hispanics. Appendix Table 2 – which shows the "first stage" relationship between predicted low-English Hispanic arrivals (our instrument) and other demographic changes in the public schools – shows that this is the case. However, the other changes are much smaller in magnitude than the one-for-one relationship of the instrument with the change in the number of low-English Hispanics. Each predicted low-English Hispanics is associated with fewer than 0.4 (non-low English) Hispanics in all specifications, and this coefficient is not usually not precisely estimated enough to be distinguished from zero. Each predicted low-English Hispanic is also associated with the arrival of 0.2 low-English non-Hispanics. Thus, the effects we estimate are not necessarily *only* a response to the enrollment of low-English Hispanics, but they are likely to be mostly driven by that.

[27] Throughout, we cluster standard errors on MSA, and given the potentially small number of clusters (22), we report p-values under the conservative assumption that test statistics are drawn from a t-distribution with 20 degrees of freedom.

highly significantly different from zero but are rarely statistically distinguishable from one. This is what we would expect if Hispanic immigrants settled in the same districts in the same proportion over 1976 to 2000 as in 1970.

Consider first the TSLS estimates for the pooled sample. Including districts regardless of center-city status (column (1)), the TSLS estimate of θ is significant at the 1 percent level and implies that nearly one non-Hispanic of school age left a district for every four additional LEP Hispanic arrivals. This result does not appear to be contaminated by suburbanization: dropping center-city districts, in column (2), produces estimates that are slightly larger, suggesting that one non-Hispanic of school age left a district for every three additional LEP Hispanic arrivals. By contrast, panel B shows that the OLS point estimate drops by half when center city districts are dropped. The TSLS estimates are also not being driven by differences across MSAs in the age composition of non-Hispanics, accounted for with age-by-MSA fixed effects in column (3).[28] They are also quite similar, though less precisely estimated, for unified and elementary districts separately, as shown in columns (4) in (5).

In general, the TSLS estimates are larger in magnitude than their OLS counterparts, especially once center city districts are dropped. The most likely explanation for this is that the OLS estimates are downward biased because Hispanic and non-Hispanic families alike were attracted to the same districts over the period – for example, both drawn to where land was relatively plentiful and housing relatively cheap. When center city districts are included, the downward bias in OLS may be offset by the fact that native families are suburbanizing for reasons besides immigration, while immigrants continue to be attracted to center city districts in large numbers.[29]

Clustering standard errors with a small number of clusters may not produce tests of correct size (Bertrand, Duflo, and Mullainathan, 2004). Simulations presented in Cameron, Gelbach, and Miller (2008) suggest that adjusting the standard errors as described yields only slight over-rejection of the null hypothesis in applications with 20 to 25 clusters. An alternative approach, which we will explore in a future draft of the paper, is to calculate bootstrapped standard errors (Cameron, Gelbach, and Miller, 2008).

[28] Thus, low-English Hispanic immigration appears to induce sorting across district lines within metropolitan areas. Rivkin (1994), Clotfelter (1999), and Urquiola (2005) also document within-MSA segregation across school districts.

[29] Immigrants may continue to be relatively more attracted to center cities because of availability of public transportation (Cutler, Glaeser, and Vigdor, 2008a) or because past suburbanization has driven down housing costs (Bouston, 2010).

There are other possible interpretations of the difference between TSLS and OLS that we cannot rule out entirely, but seem less plausible. First, enrollment growth of LEP Hispanics may be measured with error, leading to attenuation bias in OLS. When we instead instrument for changes in LEP Hispanic enrollment with another noisy measure – changes in the Hispanic school-aged population over 1970 to 2000 (from the SDDB) – the resulting TSLS estimates (available on request) are not all that different from OLS, suggesting that measurement error might not be much of a concern. Second, if the true model is one with heterogeneous effects, the local average treatment effect (LATE) identified by TSLS would be relatively large if Hispanic families that cluster into enclaves have children that are particularly difficult to educate, or if non-Hispanics who reside in such districts have a strong distaste for immigrant-induced changes to the public school system. The sheer magnitude of the inflows of Hispanic migrants to such districts may also have driven their population shares above the point at which many of their schools "tip" toward becoming predominantly Hispanic. While we cannot rule out the first LATE interpretation of our estimates, the 1970 joint share Hispanic and black for the average district in our sample was already above recently-estimated tipping points for schools (Card, Mas, and Rothstein, 2008).

Table 3 is identical to Table 2, but the dependent variable is the 1970 to 2000 change in the black population. If school-aged blacks had the same propensity to leave a school district as the non-Hispanic population at large, we would expect the TSLS coefficients from these models to be about 4 percent of the magnitude of those presented in Table 2 (or in the range of -0.014 to –0.015), since blacks represent about 4 percent of the non-Hispanic school-aged population in our sample (see Table 1).[30] The actual estimates tend to be larger in magnitude and statistically significant, suggesting that, if anything, blacks were *more* likely to relocate in response to LEP Hispanic immigration than other non-Hispanics. Nevertheless, given imprecision in the estimates, we cannot rule out equal (or lesser) responses on the part of blacks.

[30] There is small share of blacks are also low-English Hispanics, and any such overlap biases against finding a negative relationship. However, this fraction is so small that it is highly unlikely to be having a large impact on the estimates.

IV.C. Internal Validity

In Section II, we suggested several ways to test the internal validity of these estimates. First, if our approach identifies the causal effect of immigrant arrivals in public schools on native flight, then it should be the case that the largest outflows are for age groups that the school district serves. Second, the shock to school quality associated with any given immigrant inflow – and correspondingly the decline in non-Hispanic population – should be greater for districts that *ex ante* had fewer public schools. Third, districts with more private schools should see lower declines in the school-aged population. We find evidence consistent with the first two of these hypotheses using TSLS and our identification strategy, but not using OLS (results available on request).

Table 4 investigates the first of these hypotheses, reporting TSLS estimates for non-center-city districts, by type, with unrestricted interactions of $\Delta LEPH_d$ with dummies for each of the age categories. Instruments are constructed analogously, by interacting the age dummies with Z_d. The interaction with one age category, 0 to 4 years, is excluded, as estimates would otherwise be perfectly collinear with the district fixed effects.[31]

Given the marginal significance of the findings by district type in Table 2, we unsurprisingly lack power in this exercise, though the results appear to support the validity of our approach. For example, for each additional LEP Hispanic public school entrant, non-Hispanic outflows for each of the three school-age categories in unified districts (column (1)) are greater than that for the 0 to 4 age category, though only the coefficient on the interaction with the ages 5 to 9 dummy is statistically significant. Outflows of 20 to 24 year olds are one-third as large and not statistically significant. For elementary districts (column (2)), we again only find the coefficient on the interaction with the ages 5 to 9 dummy to be statistically significant, suggesting that most immigration-driven moves on the part of non-Hispanics occur when children are young.

[31] Note that estimates in previous tables restricted the coefficients on the $\Delta LEPH_d$ by age category interactions (ages 5 to 9, 10 to 14, 15 to 19 for unified and ages 5 to 9 and 10 to 14 for elementary) to be the same. We also force the population changes for comparison groups to be identical.

The findings presented in Table 4 provide support for our analytical approach in other ways. First, regardless of district type, we cannot reject identical outflows of 0 to 4 year olds and 20 to 24 year olds with LEP Hispanic arrivals for the average district, justifying our restriction that they be identical in model (4). Further, the final column of Table 3 shows results for the 50 secondary districts outside of center cities, which we excluded from the analysis. Contrary to expectations, there are marginally significant relative outflows of 5 to 9 year olds in response to LEP Hispanic arrivals into high schools, but no significant relative outflows of the 15 to 19 year olds served by the district. Though we cannot completely rule out other explanations, we suspect that the "treatment" for secondary districts is not totally clean: high school districts with growing low-English Hispanic enrollment encompass elementary districts where this is also the case. Absent information on which elementary districts historically fed which high school districts, we cannot assess the magnitude of this contamination. Since the number of districts involved is small, we decided to drop them the estimation sample.[32]

Table 5 investigates the second of the hypotheses described above – that reductions in the school-aged population in response to increasing low-English Hispanic enrollments should be greatest in districts with fewer public schools, where there is less capacity to sort residentially within district across school attendance zones. Here, we first adapt model (4) to include the interaction between the variable of interest ($1\left[a \in SA_d\right] \cdot \Delta LEPH_d$) and a dummy for having an above median number of public schools (as of 1972) for a district of that type (four for elementary and 10 for unified districts).[33] The instrument for this new variable is constructed in an analogous way. We also include interactions between this dummy (as well as between a dummy for whether the district has a private school) and the age group indicators.

[32] Including secondary districts in Table 2 has the effect one would expect from adding 50 districts for which there is no relationship to the sample: the point estimates and t-statistics in columns (1)-(3) are somewhat smaller in magnitude.
[33] Information on the number of public schools by district comes from the 1972 ELSEGIS. Creating interactions with the number of public schools makes the estimates more sensitive to outliers.

Column (1) of Table 5 shows that non-Hispanic outflows from districts with an above median number of schools are substantially smaller than those for districts below median – on the order of six fewer non-Hispanic departures for every 10 LEP Hispanic arrivals. This estimate is marginally significant and suggests that non-Hispanic population outflows in response to Hispanic LEP arrivals are concentrated in smaller districts, as expected. Figure 2 provides graphical representation of these findings, plotting the predicted difference in 1970 to 2000 population changes between the school-aged groups and the comparison groups against predicted LEP Hispanic inflow, separately by number of schools.[34] Predicted LEP Hispanic inflows unsurprisingly vary over a more restricted range for districts with a below median number of schools for their type (Panel A). However, the proportional scaling of the graphs makes clear that this group on average sees more non-Hispanic school-aged departures for every additional LEP Hispanic school entrant than the group with an above median number of schools. The pattern is the same when the model is estimated separately by district type (columns (3) and (4)), though the standard errors are large.

Column (2) of Table 5 shows that this conclusion is unchanged when we account for the possibility that private school availability – which is positively correlated with the number of public schools in the district – reduces the need to relocate, by estimating a model with all controls interacted with dummies both for having an above-median number of public schools and for the presence of a private school within the district boundaries.[35] Indeed, there is no evidence of such a trade-off, though the estimate is quite imprecise.

V. Flight to Private Schools

So far we have examined whether non-Hispanics move across district lines in response to shocks to school quality associated with low-English Hispanic arrivals in public schools. We have

---

[34] The model underlying the predictions on the vertical axis is analogous to the reduced form of the TSLS model presented in column (2) of Table 5. This model includes district fixed effects, age fixed effects, age fixed effects interacted with whether the district has a private school, and interactions between district fixed effects and whether the age group is treated ($1[a \in SA_d]=1$); the last vector of coefficients is plotted.

[35] Information on the number of private elementary and secondary schools inside each public school district comes from the 1980 Census of Private Schools and corresponds to schools in existence between 1976 and 1980.

not yet examined whether there is also a more traditionally-studied type of "flight," that is, to private schools. Betts and Fairlie (2003) found that roughly two native-born secondary students moved to private school for every 10 new immigrant enrollees in the public schools.

To see whether we find a similar relationship in our data, we can no longer rely on our strategy of looking at the difference between population changes in school-aged and comparison age groups. Indeed, the motivation for such an approach no longer applies, as flight to private schools is highly unlikely to be driven by omitted factors (like changes in other amenities) that might prompt relocation and also be correlated with immigrant inflows. Instead, we simply look directly at the relationship between changes in the private school enrollment of non-Hispanic school-aged residents of a school district and changes in public school enrollment of low-English Hispanics in the district, conditional on controls:

$$(5) \qquad \Delta PRIVNH_{dt} = \theta^{PRIV} \Delta LEPH_{d} + X'\beta + \varepsilon_{dt} ,$$

where $\Delta PRIVNH_{dt}$ is the change in non-Hispanic private-school enrollment of children enrolled in the grades that district d serves, and X represents a vector of controls. In order to make our estimates comparable to the ones based on model (4), we include in this vector the population change of 0-4 and 20-24 year olds. This control may pick up any unobserved factors driving people away from a district's public schools. Reassuringly, our findings are not sensitive to dropping this control (results available on request).[36]

Results for private school enrollment appear in Table 6, with TSLS estimates in Panel A and OLS estimates in panel B. The first stage for this alternative specification appears in Panel C, with the instrument now simply being that from equation (3) – the unadjusted and un-interacted version of the predicted change low-English Hispanics; again, we usually cannot reject a coefficient of one, and the first-stage coefficient on the instrument is strong. Regardless of specification, estimation

[36] Applying this approach to the population change outcomes we have already examined produces similar results to those shown in Table 2. In these models, the coefficient on the comparison population change is not statistically different than one, suggesting that the treatment and comparison groups would have experienced similar population changes in the absence of any growth in LEP Hispanic public school enrollment.

method, or sample, the estimated coefficient θ^{PRIV} is around the 0.2 found by Betts and Fairlie (2003) at the MSA level for secondary students.[37] However, compared to Betts and Fairlie, we find evidence of immigration-induced flight to private schools *within* the suburbs of metropolitan areas.[38] Indeed, our baseline findings from the pooled sample (column (1)) are robust to dropping center city districts (columns (2) through (5)) and to controlling for MSA-specific trends in private school enrollment through inclusion of MSA fixed effects (column (3)).[39] We also find effects among younger students, as evidenced by the estimates for elementary districts (column (5)). And again, the model for blacks, shown in Table 7, yields a TSLS coefficient at least as large as would be expected as if blacks had the same propensity to move to private schools as other non-Hispanics in response to LEP Hispanic arrivals in the public schools.

Returning to the model for all non-Hispanics (Table 6), we again see TSLS estimates larger in magnitude than OLS, but to a lesser extent than was the case in our analysis of population flows. This supports our earlier contention the difference between OLS and TSLS is driven neither by measurement error nor by heterogeneous treatment effects, which would arguably also be present here. That the coefficient on the population change of comparison groups in the first stage hovers around zero in all specifications and is generally not statistically significant suggests that endogeneity bias might not be all that severe in this context. In particular, in the absence of LEP Hispanic immigration, existing LEP Hispanics would have not have systematically settled in places where the non-Hispanic population was already declining.

As noted in Section II, for our estimates to be given a causal interpretation, it should be the case that non-Hispanic flight to private schools in response to a given influx of LEP Hispanics is greater where it is less costly (in terms of time and transportation costs) to enroll. Table 8

---

[37] This estimate implies that roughly 37.8 percent of the rise in private school enrollment in the average district in our sample can be explained by flight from low-English Hispanics.

[38] Our study differs from theirs in ways apart from the unit of observation. For example, our estimates are identified in part off of children born in the United States to low-English immigrant parents, and we restrict attention to Hispanic immigrants to California, whereas they look at first-generation immigration of all types and across the country.

[39] We obtain similar estimates for secondary districts (TSLS coefficient of 0.23 with a standard error of 0.057).

investigates the sensitivity of private school enrollment changes to the availability of private schools in the district.[40] The specifications are similar to those in Table 5. Figure 3 shows the difference in responsiveness of non-Hispanic private school enrollment across districts with and without at least one private school as of the late 1970s.

Flight to private schools does in fact appear to be very sensitive to the local availability of private schools. As shown in Figure 3, there is an upward-sloping relationship between the (residual) instrument and the (residual) 1970 to 2000 change in private school enrollment for districts with private schools, but no relationship for districts without.[41] Across specifications in the table, the interaction with a dummy for having at least one private school is similar in magnitude to our earlier estimates from Table 6, and direct effect of $\Delta LEPH$ is negative, but generally statistically insignificant. Thus, there does not appear to be flight to private schools unless there is at least one private school located within the boundaries of the public school district. This suggests that a geographic market for private schools is potentially more localized than a metropolitan area and illustrates the value of our testing for immigration-induced changes in private school enrollment *within* – not just across – metropolitan areas.[42]

Column (2) of Table 8 explores whether flight to private schools is also sensitive to the number of public alternatives located within the district. The coefficient is negative – which is the right direction to indicate a tradeoff between public and private choices – though it is overwhelmed by the standard error. Again, we find no clear indication that relocation is a substitute for remaining within a district and attending a private school, though our test is low powered.

---

40 About half of elementary districts contain at least one private school, and about 75 percent of unified districts do.
41 Each point represents the residuals from regressions of the 1970 to 2000 private school enrollment change (vertical axis) and of the predicted inflow of LEP Hispanic children to a district (horizontal axis) on the average population change for the comparison age group, an indicator for whether the district has an above median number of public schools for its type, the interaction of these variables, an indicator for whether the district is an elementary district, and the interaction between this variable and the indicator for above median number of public schools. This is the reduced-form of a model less restrictive than but otherwise analogous to that in column (2) of Table 8.
42 This is consistent with work on intra-district public school choice, where parents seem to place a lot of value on distance to a choice school (e.g., Hastings and Weinstein, 2008).

VI. Implications

Over the period of our study, California's low-English Hispanics experienced large declines in the average share non-Hispanic in public schools in their districts of residence.[43] In particular, the non-Hispanic share of public enrollees fell from 64 percent to 43 percent in the average low-English Hispanic student's district between 1976 and 2000.[44] Much of this decline is mechanical: non-Hispanic share of enrollment would have fallen without flight because of the statewide increase in Hispanic enrollment. However, our estimates indicate that flight from low-English Hispanics substantially exacerbated this decline. Based on our estimates from the first columns of Table 5 and Table 8, the decline in non-Hispanic share experienced by the average Hispanic LEP student would have been 27 percent (5.7 percentage points) smaller absent flight.[45] Given evidence of sorting within districts (Alesina, Baqir, and Hoxby, 2004; Kane, Riegg, and Staiger, 2006; Weinstein, 2009), immigration has also likely had a substantial effect on segregation within districts, which we cannot directly explore due to data constraints.

Flight and the corresponding increase in isolation may have negative implications for the education of low-English Hispanic students in California. Estimates from court-ordered desegregation, for instance, suggest that exposure to whites may increase high school completion rates of minorities (e.g., Guryan, 2004).[46] More in context, some evidence also suggests that neighborhood-level segregation of immigrant students reduces their ability to learn English and to

[43] In short, these are declines in non-Hispanic "exposure" measured at the district level. We have avoided the term exposure, however, to not give the reader the mistaken impression that we have school-level data, which is the level at which exposure indices are often measured.
[44] The percentage point decline is similar in all of California's districts (not just those in our sample), though the levels are lower in both periods.
[45] For this exercise, to each district's actual 2000 enrollment of non-Hispanics and total enrollment was added the predicted decline in non-Hispanic population based on column (1) of Table 5 (0.86*$\Delta LEPH$ in districts with a below median number of public schools and 0.239*$\Delta LEPH$ in districts with at least the median number of public schools, where $\Delta LEPH$ is the actual change in low-English Hispanic enrollment between 1976 and 2000) and the predicted increase in private school enrollment of non-Hispanics based on column (1) of Table 7 (-0.103*$\Delta LEPH$ in districts without a private school and 0.117*$\Delta LEPH$ in districts with one, where $\Delta LEPH$ is the actual change in low-English Hispanic enrollment between 1976 and 2000). Based on these additions, the counterfactual non-Hispanic share, whose averages are reported above, was constructed.
[46] Court-ordered desegregation in the context of high historical levels of segregation may, however, have a different impact on achievement than other changes in racial mixing.

"assimilate" in other ways (Cutler, Glaeser, and Vigdor, 2008b).[47] One might further speculate that this isolation is part of the reason so many second- and third-generation Hispanic schoolchildren in California are limited English proficient.

VII. Conclusion

This paper has examined whether the large wave of low-English Hispanic immigration to California since 1970 has induced non-Hispanic flight from the public schools where these immigrants have settled. Our empirical approach accounts for endogeneity of immigrant inflows using earlier immigrant settlement patterns, and when examining effects on relocation, we make within-district comparisons across school-aged and non school-aged populations to isolate the contribution of school quality.

We find that one more non-Hispanic child left the public school system for every two additional LEP Hispanic arrivals to its schools over 1970 to 2000. Sixty percent of this effect is explained by relocation to other school districts, and 40 percent by moves to private schools within district. Back-of-the-envelope calculations based on our estimates imply that the decline in non-Hispanic public school enrollment share over this period in the average low-English Hispanic child's school district would have been 25 percent lower in the absence of flight. Blacks appear to have left public schools in response to immigration in numbers proportional to their representation in the non-Hispanic population. This suggests that low-English Hispanic immigration may not have contributed to rising racial isolation of blacks, though we do not know whether non-Hispanics of different races subsequently settled or attended private school.

Our findings suggest that recently-documented increases in residential isolation of immigrants (Cutler, Glaeser, and Vigdor, 2008a, 2008b) may be driven in part by sorting in response to immigration-induced changes in school quality. While Saiz and Wachter (2006) note the

[47] There are other studies on this subject, but most are plagued by the fact that immigrants who settle in ethnically concentrated neighborhoods are severely negatively selected, which is one thing the Cutler et al. (2008b) study attempts to address. It is important to point out that there are also some benefits to segregation for immigrants – better access to job networks, etc. – and in some cases these benefits overwhelm the costs of being isolated from the native population (in particular, for highly educated immigrants according to Cutler, Glaeser, and Vigdor, 2008b).

possibility that immigrants change amenities associated with a neighborhood, broadly speaking, the importance of local public goods in general– and public schools in particular – to changes in the housing market equilibrium in response to immigration has to this point been overlooked. How these flight-induced increases in ethnic isolation have affected the well-being of the students left behind is an important question for future research.

VIII. Data Appendix
School District Level Data: Sources and Construction of Key Variables

A. 1970 Fourth Count (Population) School District Data Tapes

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. These data permit identification of all school districts in the country with at least 300 students (as of the 1969-70 school year).

Counts of school district residents by gender were originally reported for the total population and for the "Spanish Heritage" (hereafter referred to as Hispanic) population in the following age bins: under 3, 3-4, 5, 6, 7-9, 10-13, 14, 15, 16, 17, 18, and 19, 20, 21, to 22-24 (Table 17).[48] For comparability with the 2000 data, we aggregated resident counts for the total population and the Hispanic population into five-year age bins (0-4, 5-9, 10-14, 15-19, and 20-24). We computed corresponding resident counts of non-Hispanics with the difference.

The original data also report counts school district 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 28). For comparability with the 2000 data, we combine the kindergarten and elementary counts. To arrive at one private enrollment figure for the total population and for Hispanics, we also drop data on private school enrollment of individuals at levels not served by the district.[49] We compute private enrollment counts for non-Hispanics by taking the difference between the total and Hispanic figures.

The final data we draw from this source are counts of "persons of foreign stock by nativity and country of origin," for the total population and for the Spanish Heritage population (Table 22), used in construction of the instrument. Specifically, we use these data to construct $M_{dg}^{1970} / M_{g}^{1970}$ in equation (3) – the share of the U.S. Hispanic population born in country group *g* and residing in district *d* in 1970.

B. Census 2000 School District Tabulation

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 (STP2), available at <http://nces.ed.gov/surveys/sdds/downloadmain.asp>. All operating districts are included in the age-specific resident counts, but private enrollment counts are missing for districts with 49 or fewer children.

Counts of school district residents by gender were originally reported for the total population in one-year age bins through age 21 and for ages 22-24 (Table P8 for Total – Population and Households (TT)) and for the Hispanic/Latino population for the age categories 0-4, 5-9, 10-14, 15-17, 18-19, 20, 21, and 22-24 (Table 145H for TT). As in 1970, we aggregated resident counts for the total population and the Hispanic population into five-year age bins (0-4, 5-9, 10-14, 15-19, and 20-24) and computed the corresponding resident counts for non-Hispanics with the difference.

---

[48] 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.

[49] For example, if the district is an elementary district, we do not include private school enrollees at the secondary level in the district private enrollment count.

Counts of school district residents in private school, by gender, were available 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 the age categories 0-4, 5-9, 10-14, 15-17, 18-19. For comparability with the 1970 data, we aggregate across all age categories and across gender to create one private enrollment figure each for the total population and for Hispanics.

To avoid disclosure, cell values are also 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.

C. Fall 1976 Elementary and Secondary School Civil Rights Survey

For 1976, school district level data on the number of LEP students, by 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 elementary and secondary school districts in the United States.

The original data give counts of "pupils whose primary language is other than English" in total and by race/ethnicity. Our treatment variable ($\Delta LEPH_d$) is constructed using the number of Hispanics (of all races) with this designation. Race/ethnicity categories are American Indian/Alaskan Native, Asian/Pacific Islander, Black non-Hispanic, White non-Hispanic, and Hispanic.

D. 2000 Elementary and Secondary School Civil Rights Compliance Report District Summary

For 2000, school district level data on the number of LEP 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 and downloaded from <http://www.ed.gov/about/offices/list/ocr/data.html>.The 2000 OCR survey covered all elementary and secondary school districts in the United States, with tabulations rounded to the nearest 5, to avoid disclosure.

References

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

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

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119(1): 249-275.

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.

Bleakley, Hoyt and Aimee Chin (2008). "What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital." *Journal of Human Resources* 43: 267-298.

Boustan, Leah Platt (2009). "School Desegregation and Urban Change: Evidence from City Boundaries." Mimeo, UCLA.

Boustan, Leah Platt (2010). "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration." *Quarterly Journal of Economics*, forthcoming.

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

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

Clotfelter, Charles T. (1999). "Public School Segregation in Metropolitan Areas." *Land Economics* 75(4): 487-504.

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

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

Guryan, Jonathan (2004). "Desegregation and Black Dropout Rates." *American Economic Review* 94(4): September, p.. 919-943.

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.

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.

Rivkin, Steven G. (1994). "Residential Segregation and School Integration." *Sociology of Education* 67(4): 279-292.

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.

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 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.

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

Welch, Finis and Audrey Light (1987). *New Evidence on School Desegregation.* Washington, D.C.: U.S. Commission on Civil Rights.

Table 1. Descriptive Statistics on California Public School Districts

Elementary & unified public school districts outside center cities

	Base Year † Mean	2000 Mean	Difference	Std Dev. Difference
A. OCR				
Public School Enrollees				
Low-English Hispanics[X]	194	1,179	984	1,788
Not Low-English Hispanics	706	1,594	889	1,525
Non-Hispanics	4,242	4,431	189	3,379
Total	5,142	7,204	2,062	4,481
B. 1970 SDDB and 2000 Census of Population				
Predicted (Change in) Hispanic Low-English Public School Enrollment[Z]			577	1,059
C. SDDB				
Non-Hispanic Population:[Y]				
Ages 0 to 4	1,740	1,914	174	1,675
Ages 5 to 9	2,357	2,259	-98	2,357
Ages 10 to 14	2,252	2,325	72	2,091
Ages 15 to 19	2,085	2,125	40	2,059
Ages 20 to 24	1,883	1,873	-10	1,982
Black Population:[Y]				
Ages 0 to 4	70	175	105	349
Ages 5 to 9	95	221	127	435
Ages 10 to 14	87	222	135	413
Ages 15 to 19	74	185	111	333
Ages 20 to 24	66	159	93	272
Non-Hispanic Private School Enrollees[Y]	89	653	565	844
Black Private School Enrollees[Y]	10	84	74	498

Notes: The sample size in all panels is 428 districts (177 unified and 251 elementary). Data sources and †years: A: The Office for Civil Rights (OCR) Elementary and Secondary School Surveys, 1976 and 2000. B: The 2000 Census of Population, used to measure low-English Hispanic school-aged arrivals nationally between 1976 and 2000, and the 1970 School District Data Book., used to apportion them to particular districts (see below). C: The School District Data Book, 1970 and 2000. [X]Independent variable in regressions in subsequent tables. [Y]Dependent variable in regressions in subsequent tables. [Z]Instrumental variable in regressions in subsequent tables. The instrumental variable is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according to the 1970 SDDB). Four country groups have substanial low-English Hispanic immigration: Mexico, the rest of Latin America and carribbean, southern Europe, and Cuba. "Second generation" is defined here to be native-born children whose parents report arriving in the U.S. in 1976 or later.

Table 2. Change in non-Hispanic population, 1970-2000, on change in Hispanic LEP * enrollment, 1976-2000

Sample:	All	No Center City Districts (Columns 2-4)			
District Types:	Elementary and Unified			Unified	Elementary
	(1)	(2)	(3)	(4)	(5)
A. Two-Stage Least Squares (TSLS)					
ΔHispanic LEP* enrollment	-0.261	-0.346	-0.371	-0.377	-0.362
1976-2000, × school age	(0.0151)	(0.0620)	(0.0641)	(0.160)	(0.201)
P-value (t_{20})	0.000	0.000	0.000	0.029	0.086
Reduced form p-value (t_{20})	0.000	0.000	0.002	0.075	0.099
Root MSE	1051	767.1	755.1	1029	508.8
R-squared	0.967	0.881	0.890	0.886	0.864
B. Ordinary Least Squares					
ΔHispanic LEP* enrollment	-0.249	-0.127	-0.135	-0.0844	-0.173
1976-2000, × school age	(0.0154)	(0.0784)	(0.0660)	(0.123)	(0.0669)
P-value (t_{20})	0.000	0.122	0.053	0.499	0.018
Root MSE	1051	762.6	750.2	1021	505.2
R-squared	0.967	0.882	0.892	0.888	0.866
C. First Stage. Dependent Variable = ΔHispanic LEP* enrollment, 1976-2000, x school age					
Predicted ΔHispanic LEP* enroll-	1.214	1.034	0.972	0.842	1.275
ment, 1976-2000, × school age	(0.0121)	(0.113)	(0.112)	(0.0908)	(0.135)
R-squared	0.985	0.744	0.758	0.759	0.727
Fixed Effects (Trends)					
Age Group	Yes	Yes	Yes	Yes	Yes
School District	Yes	Yes	Yes	Yes	Yes
Age group x Metro Area	No	No	Yes	No	No
Observations	2290	2170	2170	895	1275
# of districts	452	428	428	177	251

Notes: The dependent variable in all columns of panels A and B is the change in the non-Hispanic population, 1970-2000, which varies by age group -- 5 year age bands between 0-4 and 20-24 -- and district. Data source : 1970 and 2000 school district data book (SDDB). *The independent variable in panels A and B, and the dependent variable in panel C, is the change in Hispanic LEP enrollment = change in the enrollment of limited English proficient (LEP) Hispanic students in this public school district ineracted with a dummy for being of school age for the district's type. School-aged is defined to be age 5-14 in elementary school districts, and age 5-19 in unified districts. The LEP count is the district's estimate of the number of students "in need of," not necessarily actually enrolled in, specialized classes for English instruction. To give the coefficient estimate the proper interpretation, this count is divided by the number of school aged age groups (3 for unified districts, 2 for elementary districts). Data source: 1976 and 2000 Elementary and Secondary School Survey. The instrumental variable in panel A, and the independent variable in Panel C, is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according to the 1970 SDDB), interacted with a dummy for being school-aged for the district's type and divided by the number of school-aged age groups in the district type. Four country groups have substanial low-English Hispanic immigration: Mexico, the rest of Latin America and carribbean, southern Europe, and Cuba. "Second generation" is defined here to be native-born children whose parents report arriving in the U.S. in 1976 or later. Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan area. There are 22 metropolitan areas in the sample.

Table 3. Change in black population, 1970-2000, on change in Hispanic LEP[*] enrollment, 1976-2000

Sample:	All	No Center City Districts (Columns 2-4)			
District Types:	Elementary and Unified			Unified	Elementary
	(1)	(2)	(3)	(4)	(5)
A. Two-Stage Least Squares (TSLS)					
ΔHispanic LEP[*] enrollment	-0.0273	-0.00888	-0.0267	0.00903	-0.0453
1976-2000, × school age	(0.00201)	(0.0107)	(0.0109)	(0.0153)	(0.0213)
P-value (t_{20})	0.000	0.417	0.024	0.561	0.046
Reduced form p-value (t_{20})	0.000	0.431	0.039	0.580	0.075
Root MSE	418.2	109.2	109.1	138.0	83.01
R-squared	0.921	0.929	0.932	0.933	0.913
B. Ordinary Least Squares					
ΔHispanic LEP[*] enrollment	-0.0245	0.0304	0.0229	0.0466	0.00846
1976-2000, × school age	(0.00110)	(0.0187)	(0.0193)	(0.0414)	(0.0323)
P-value (t_{20})	0.000	0.119	0.249	0.274	0.796
Root MSE	418.2	108.2	107.6	137.1	81.22
R-squared	0.921	0.930	0.934	0.934	0.917
Fixed Effects (Trends)					
Age Group	Yes	Yes	Yes	Yes	Yes
School District	Yes	Yes	Yes	Yes	Yes
Age group x Metro Area	No	No	Yes	No	No
Observations	2290	2170	2170	895	1275
# of districts	452	428	428	177	251

Notes: The dependent variable in all columns of panels A and B is the change in the black population, 1970-2000, which varies by age group -- 5 year age bands betweeen 0-4 and 20-24 -- and district. Data source : 1970 and 2000 school district data book (SDDB). *The independent variable in panels A and B is the change in Hispanic LEP enrollment = change in the enrollment of limited English proficient (LEP) Hispanic students in this public school district ineracted with a dummy for being of school age for the district's type. School-aged is defined to be age 5-14 in elementary school districts, and age 5-19 in unified districts. The LEP count is the district's estimate of the number of students "in need of," not necessarily actually enrolled in, specialized classes for English instruction. To give the coefficient estimate the proper interpretation, this count is divided by the number of school aged age groups (3 for unified districts, 2 for elementary districts). Data source: 1976 and 2000 Elementary and Secondary School Survey. The instrumental variable in panel A is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according to the 1970 SDDB), interacted with a dummy for being school-aged for the district's type and divided by the number of school-aged age groups in the district type. Four country groups have substanial low-English Hispanic immigration: Mexico, the rest of Latin America and carribbean, southern Europe, and Cuba. "Second generation" is defined here to be native-born children whose parents report arriving in the U.S. in 1976 or later. Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan area. There are 22 metropolitan areas in the sample.

<u>Table 4</u>. Two-stage least squares estimates of change in non-Hispanic population on change in Hispanic LEP enrollment, by age group

Sample:	No Center City Districts		
District Types:	Unified	Elementary	High
	(1)	(2)	(3)
School Age (coefficients in bold):			
ΔLEP* Hispanic enroll, 1976-2000, × age 5-9	**-0.502**	**-0.415**	-0.655
	(0.165)	**(0.0987)**	(0.353)
ΔLEP* Hispanic enroll, 1976-2000, × age 10-14	**-0.349**	**-0.109**	-0.0641
	(0.219)	**(0.145)**	(0.438)
ΔLEP* Hispanic enroll, 1976-2000, × age 15-19	**-0.560**	0.0888	**0.327**
	(0.430)	(0.111)	**(0.475)**
Comparison groups:			
ΔLEP* Hispanic enroll, 1976-2000, × age 0-4		*excluded*	
ΔLEP* Hispanic enroll, 1976-2000, × age 20-24	-0.186	0.211	0.435
	(0.274)	(0.317)	(0.654)
Fixed Effects (Trends):			
Age Group	Yes	Yes	Yes
School District	Yes	Yes	Yes
Root MSE	1029	509.7	1415
R-squared	0.886	0.864	0.906
Observations	885	1255	250
# of districts	177	251	50

Notes: * Change in Hispanic LEP enrollment = change in the enrollment of limited English proficient (LEP) Hispanic students in this public school district. This is the district's estimate of the number of students "in need of," not necessarily actually enrolled in, specialized classes for English instruction. To give the coefficient estimate the proper interpretation, this count is divided by the number of school aged age groups (3 for unified districts, 2 for elementary districts -- see below). Data source: 1976 and 2000 Elementary and Secondary School Survey. The dependent variable in all columns is the change in the non-Hispanic population between 1970 and 2000. School-age defined to be age 5-14 in elementary school districts, age 15-19 in high school districts, and age 5-19 in unified districts . Data source : 1970 and 2000 school district data book (SDDB). Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan area. There are 22 metropolitan areas in the sample.

Table 5. **Change in non-Hispanic population, 1970-2000, by number of public and private schools in the district**

Sample:	Non-Center City Districts			
District Types:	Elementary & Unified		Unified	Elementary
Estimation Method:	TSLS	TSLS	TSLS	TSLS
	(1)	(2)	(3)	(4)
ΔHispanic LEP* enrollment, 1976-2000				
× school age	-0.860	-0.839	-0.824	-0.892
	(0.359)	(0.363)	(0.682)	(0.544)
× school age × ≥median	0.621	0.664	0.529	0.667
# of public schools for type†	(0.370)	(0.354)	(0.763)	(0.610)
× school age × has a private		-0.0658		
school serving district-aged kids‡		(0.132)		
Root MSE:	762.4	762.7	1024	501.5
R-squared	0.883	0.883	0.888	0.869
Fixed Effects (Trends)				
Age Group	Yes	Yes	Yes	Yes
× ≥median # public schools†	Yes	Yes	Yes	Yes
× has a private school‡	Yes	Yes	Yes	Yes
School District	Yes	Yes	Yes	Yes
Observations	2140	2140	885	1255
Number of Districts	428	428	177	251

Notes: The dependent variable in all columns is the change in the non-Hispanic population, 1970-2000, which varies by age group -- 5 year age bands betweeen 0-4 and 20-24 -- and district. Data source : 1970 and 2000 school district data book (SDDB). * The main independent variable is the change in Hispanic LEP enrollment = change in the enrollment of limited English proficient (LEP) Hispanic students in this public school district ineracted with a dummy for being of school age (for the district type -- see below). The LEP count is the district's estimate of the number of students "in need of," not necessarily actually enrolled in, specialized classes for English instruction. To give the coefficient estimate the proper interpretation, this count is divided by the number of school aged age groups (3 for unified districts, 2 for elementary districts -- see below). Data source: 1976 and 2000 Elementary and Secondary School Survey. The instrumental variable in all columns is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according the 1970 school district data book). (Four country groups have substanial low-English Hispanic immigration: Mexico, the rest of Latin America and carribbean, southern Europe, and Cuba. "Second generation" is defined here to be native-born children whose parents report arriving in the U.S. in 1976 or later.) The main independent variable and the age group effects are also ineracted with a dummy for having an above median number of schools in a district of that type. †4 and 10 are, respectively, the median number of schools in our "aggregated" elementary school districts and unified districts according to the 1972 Elementary and Secondary General Information System (ELSEGIS). (Districts have been "aggregated" to create districts which are geographically consistent over our sample period). ‡The main independent variable, and the age groups effects, are also ineracted with a dummy for having a at least one private school which serves kids of school age for that district. School-aged for district type defined to be age 5-14 in elementary school districts and age 5-19 in unified districts. Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan area. There are 22 metropolitan areas in the sample.

Table 6. Change in non-Hispanic private school enrollment, 1970-2000

Subsample:	All	No Center City Districts (Columns 2-4)			
District Types:	Elementary & Unified			Unified	Elementary
	(1)	(2)	(3)	(4)	(5)
A. Two-Stage Least Squares (TSLS)					
ΔHispanic LEP* enrollment	0.217	0.226	0.213	0.246	0.172
1976-2000	(0.0263)	(0.0699)	(0.0786)	(0.0968)	(0.0242)
ΔNon-Hispanic Population,	0.101	0.161	0.184	0.168	0.137
age 0-4 and 20-24, 1976-2000	(0.0959)	(0.0550)	(0.0630)	(0.0693)	(0.0405)
Root MSE	1057	721.1	672.5	1051	336.6
R-squared	0.852	0.275	0.401	0.173	0.205
B. Ordinary Least Squares					
ΔHispanic LEP* enrollment	0.185	0.155	0.138	0.167	0.120
1976-2000	(0.0351)	(0.0628)	(0.0594)	(0.0829)	(0.0303)
ΔNon-Hispanic Population,	0.0112	0.153	0.178	0.159	0.129
age 0-4 and 20-24, 1976-2000	(0.123)	(0.0588)	(0.0660)	(0.0744)	(0.0414)
Root MSE	1039	710.5	660.7	1034	331.1
R-squared	0.857	0.296	0.421	0.198	0.231
C. First Stage. Dependent Variable = ΔHispanic LEP* enrollment, 1976-2000					
Predicted ΔHispanic LEP* enroll-	1.168	0.982	0.924	0.887	1.350
ment, 1976-2000	(0.0663)	(0.135)	(0.130)	(0.118)	(0.0985)
ΔNon-Hispanic Population,	-0.123	0.0395	0.0254	0.0193	0.140
age 0-4 and 20-24, 1976-2000	(0.210)	(0.0507)	(0.0486)	(0.0570)	(0.0709)
R-squared	0.968	0.357	0.395	0.282	0.475
Fixed Effects (Trends)					
District Type	Yes	Yes	Yes	No	No
Metro Area	No	No	Yes	No	No
Observations	452	428	428	177	251

Notes: The dependent variable in panels A and B is the change in the non-Hispanic private school enrollment between 1970 and 2000. Data source : 1970 and 2000 school district data book (SDDB). * The main independent variable is the change in Hispanic LEP enrollment = change in the enrollment of limited English proficient (LEP) Hispanic students in this public school district. This is the district's estimate of the number of students "in need of," not necessarily actually enrolled in, specialized classes for English instruction.. Data source: 1976 and 2000 Elementary and Secondary School Survey. The instrumental variable is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according to the 1970 SDDB). School-aged for district type defined to be age 5-14 in elementary school districts and age 5-19 in unified districts. Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan area. There are 22 metropolitan areas in the sample.

Table 7. Change in black private school enrollment, 1970-2000

Sample:	All	No Center City Districts (Columns 2-4)			
District Types:		Elementary & Unified		Unified	Elementary
	(1)	(2)	(3)	(4)	(5)
A. Two-Stage Least Squares (TSLS)					
ΔHispanic LEP* enrollment	0.0316	0.0190	0.0137	0.0221	0.0127
1976-2000	(0.00405)	(0.00699)	(0.0107)	(0.0106)	(0.00252)
ΔNon-Hispanic Population,	-0.0127	-0.00331	-0.00144	-0.00582	0.00666
age 0-4 and 20-24, 1976-2000	(0.0142)	(0.00682)	(0.00687)	(0.00842)	(0.00186)
Root MSE	194.5	139.7	141.2	212.4	35.63
R-squared	0.861	0.130	0.155	0.107	0.179
B. Ordinary Least Squares					
ΔHispanic LEP* enrollment	0.0300	0.0271	0.0267	0.0320	0.0137
1976-2000	(0.00261)	(0.00500)	(0.00610)	(0.00680)	(0.00330)
ΔNon-Hispanic Population,	-0.0171	-0.00240	-0.000300	-0.00476	0.00680
age 0-4 and 20-24, 1976-2000	(0.01000)	(0.00550)	(0.00546)	(0.00673)	(0.00219)
Root MSE	194.2	139.0	139.5	211.1	35.61
R-squared	0.862	0.139	0.175	0.117	0.180
C. First Stage. Dependent Variable = ΔHispanic LEP* enrollment, 1976-2000					
Predicted ΔHispanic LEP* enroll-	1.168	0.982	0.924	0.887	1.350
ment, 1976-2000	(0.0663)	(0.135)	(0.130)	(0.118)	(0.0985)
ΔNon-Hispanic Population,	-0.123	0.0395	0.0254	0.0193	0.140
age 0-4 and 20-24, 1976-2000	(0.210)	(0.0507)	(0.0486)	(0.0570)	(0.0709)
R-squared	0.968	0.357	0.395	0.282	0.475
Fixed Effects (Trends)					
District Type	Yes	Yes	Yes	No	No
Metro Area	No	No	Yes	No	No
Observations	452	428	428	177	251

Notes: The dependent variable in panels A and B is the change in the non-Hispanic private school enrollment between 1970 and 2000. Data source : 1970 and 2000 school district data book (SDDB). * The main independent variable is the change in Hispanic LEP enrollment = change in the enrollment of limited English proficient (LEP) Hispanic students in this public school district. This is the district's estimate of the number of students "in need of," not necessarily actually enrolled in, specialized classes for English instruction.. Data source: 1976 and 2000 Elementary and Secondary School Survey. The instrumental variable is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according to the 1970 SDDB). School-aged for district type defined to be age 5-14 in elementary school districts and age 5-19 in unified districts. Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan area. There are 22 metropolitan areas in the sample.

Table 8. Change in non-Hispanic private school enrollment, 1970-2000, by number of public and private schools in district

Sample:	Non-Center City Districts			
District Types:	Elementary & Unified		Unified	Elementary
Estimation Method:	TSLS	TSLS	TSLS	TSLS
	(1)	(2)	(3)	(4)
ΔHispanic LEP* enrollment,	-0.103	-0.0881	-0.0828	-0.0494
1976-2000	(0.0442)	(0.0673)	(0.0533)	(0.0337)
× has a private school	0.220	0.224	0.205	0.147
serving district-aged kids‡	(0.0411)	(0.0462)	(0.0638)	(0.0273)
× ≥median #of public		-0.0205		
schools for district type†		(0.0961)		
Root MSE:	630.3	629.6	931.3	281.5
R-squared	0.457	0.457	0.372	0.457
Other Controls:				
ΔNon-Hispanic non-schl aged pop**	Yes	Yes	Yes	Yes
× ≥median # public schools†	Yes	Yes	Yes	Yes
× has a private school‡	Yes	Yes	Yes	Yes
Fixed Effects (Trends)				
District Type	Yes	Yes	No	No
× ≥median # public schools†	Yes	Yes	Yes	Yes
× has a private school‡	Yes	Yes	Yes	Yes
Observations	428	428	177	251

Notes: The dependent variable in all columns is the change in the non-Hispanic private school enrollment between 1970 and 2000. Data source : 1970 and 2000 school district data book (SDDB). * The main independent variable is the change in Hispanic LEP enrollment = change in the enrollment of limited English proficient (LEP) Hispanic students in this public school district . The LEP count is the district's estimate of the number of students "in need of," not necessarily actually enrolled in, specialized classes for English instruction. Data source: 1976 and 2000 Elementary and Secondary School Survey. **All regressions control for the 1970-2000 change in the non-Hispanic population "not of school age,"defined here age 0-4 and 20-24 (summed together). Source: 1970 and 2000 SDDB. The instrumental variable in all columns is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according the 1970 school district data book). (Four country groups have substanial low-English Hispanic immigration: Mexico, the rest of Latin America and carribbean, southern Europe, and Cuba. "Second generation" is defined here to be native-born children whose parents report arriving in the U.S. in 1976 or later.) The main independent variable, dummies for district type, and the change in the non-Hispanic not-school aged population are also all ineracted with a dummy for having an above median number of schools in a district of that type. †4 and 10 are, respectively, the median number of schools in our "aggregated" elementary school districts and unified districts according to the 1972 Elementary and Secondary General Information System (ELSEGIS). (Districts have been "aggregated" to create districts which are geographically consistent over our sample period). ‡The main independent variable is also ineracted with a dummy for having a at least one private school which serves kids of school age for that district. School-aged for district type defined to be age 5-14 in elementary school districts and age 5-19 in unified districts. Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan area. There are 22 metropolitan areas in the sample.

Appendix Table 1. Descriptive Statistics on California Public School Districts
All Available Districts

	Base Year † Mean	2000 Mean	Difference	Std Dev. Difference
A. OCR				
Public School Enrollees				
Low-English Hispanics[X]	269	2,118	1,849	12,996
Not Low-English Hispanics	1,250	2,421	1,171	2,368
Non-Hispanics	6,251	5,701	-550	9,585
Total	7,770	10,240	2,470	7,191
B. 1970 SDDB and 2000 Census of Population				
Predicted (Change in) Hispanic Low-English Public School Enrollment[Z]			1,226	10,530
C. SDDB				
Non-Hispanic Population:[Y]				
Ages 0 to 4	2,995	2,862	-132	4,264
Ages 5 to 9	3,951	3,329	-622	5,863
Ages 10 to 14	3,758	3,396	-362	5,380
Ages 15 to 19	3,540	3,140	-401	5,221
Ages 20 to 24	3,338	2,978	-359	4,536
Black Population:[Y]				
Ages 0 to 4	305	357	52	1,473
Ages 5 to 9	392	449	57	1,770
Ages 10 to 14	347	443	96	1,228
Ages 15 to 19	286	368	82	977
Ages 20 to 24	241	328	87	844
Non-Hispanic Private School Enrollees[Y]	163	962	799	760
Black Private School Enrollees[Y]	3	43	40	149

Notes: The sample size in all panels is 503 districts (196 unified and 256 elementary, and 51 high school districts). Data sources and †years: A: The Office for Civil Rights' (OCR) Elementary and Secondary School Surveys, 1976 and 2000. B: The 2000 Census of Population -- used to measure low-English Hispanic school-aged arrivals nationally between 1976 and 2000, and the 1970 School District Data Book., used to apportion them to particular districts (see below) C: The School District Data Book, 1970 and 2000. C: 1972 Elementary and Secondary General Information System . [X]Independent variable in regressions in subsequent tables. [Y]Dependent variable in regressions in subsequent tables. [Z]Instrumental variable in regressions in subsequent tables. The instrumental variable is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according to the 1970 SDDB). Four country groups have substanial low-English Hispanic immigration: Mexico, the rest of Latin America and carribbean, southern Europe, and Cuba. "Second generation" is defined here to be native-born children whose parents report arriving in the U.S. in 1976 or later.)

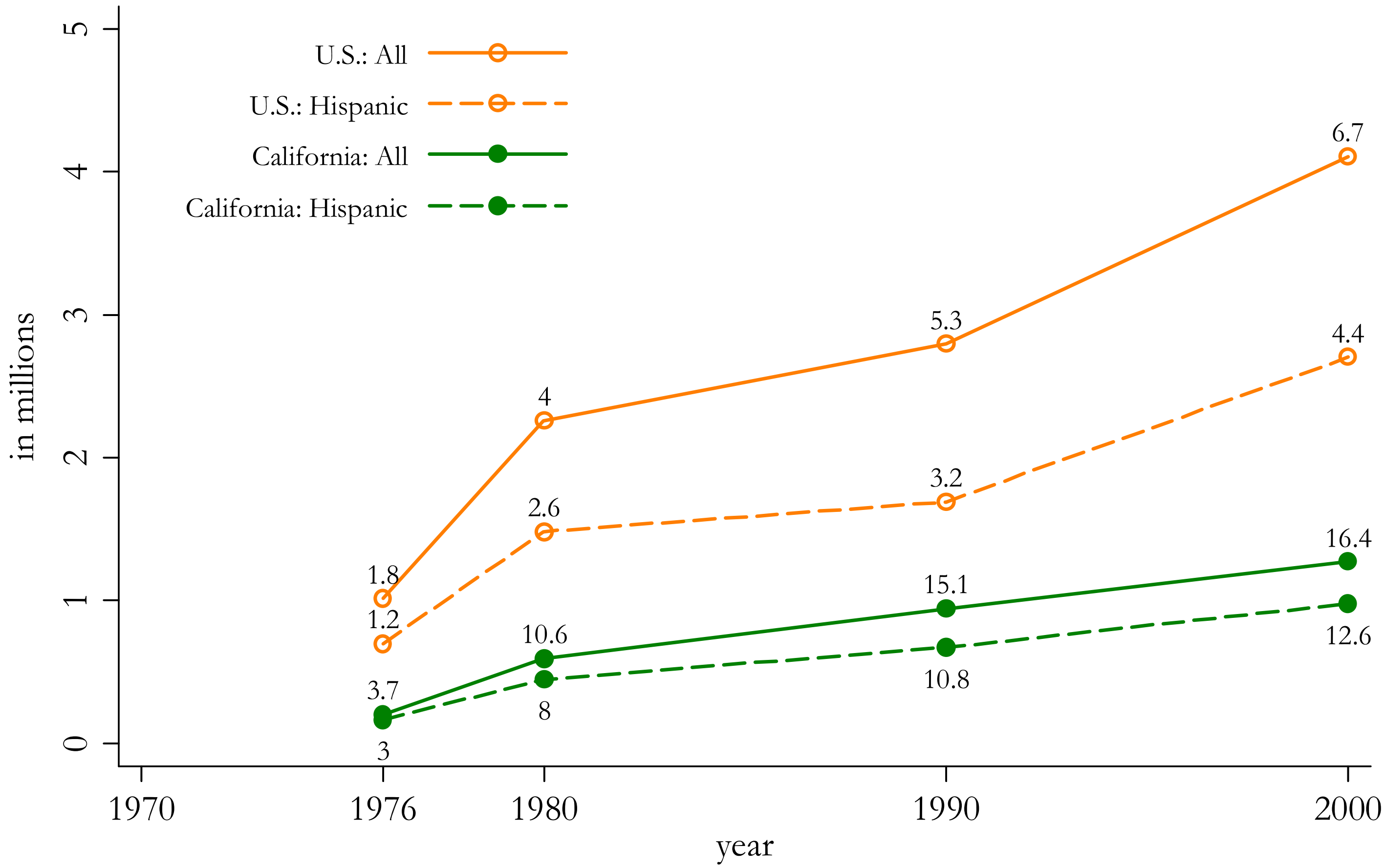
<u>Appendix Table 2</u>. Reduced form relationship of other public school demographic changes with the instrument, 1976-2000

Sample:	All	No Center City Districts (Columns 2-5)			
District Types:		Elementary and Unified		Unified	Elementary
	(1)	(2)	(3)	(4)	(5)
		Δ Hispanic non-LEP enrollment × school age			
Predicted ΔHispanic LEP* enrollment × school age	0.137	0.336	0.224	0.164	0.375
	(0.0114)	(0.190)	(0.193)	(0.139)	(0.304)
		Δ Non-hispanic LEP enrollment × school age			
Predicted ΔHispanic LEP* enrollment × school age	0.0869	0.193	0.186	0.146	0.220
	(0.00316)	(0.0460)	(0.0497)	(0.0571)	(0.0777)
Fixed Effects (Trends)					
Age Group	Yes	Yes	Yes	Yes	Yes
School District	Yes	Yes	Yes	Yes	Yes
Age group x Metro Area	No	No	Yes	No	No
Observations	2290	2170	2170	895	1275
# of districts	452	428	428	177	251

Notes: *LEP=limited English proficient. Data source: 1976 and 2000 Elementary and Secondary School Survey. The independent variable in all columns is the predicted change in Hispanic LEP enrollment based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled in 1970 (according to the 1970 school district data book). (Four country groups have substanial low-English Hispanic immigration: Mexico, the rest of Latin America and carribbean, southern Europe, and Cuba. "Second generation" is defined here to be native-born children whose parents report arriving in the U.S. in 1976 or later.) To maintain the same data structure as the rest of the analysis, each district's observation is repeated five times, corresponding to the five five-year age bands -- 0-4, 5-10, 10-14, 15-19, and 20-24 -- of the dependent variable in the main analysis. In addition, both predicted LEP enrollment and the dependent variables are interacted with a dummy for school age, which is defined to be age 5-14 in elementary school districts and age 5-19 in unified districts, and divided by three for unified districts and two for elementary school districts (each corresponds to the number of "school aged" age categories). Standard errors, in parentheses, are calculated to be robust to arbitrary error correlation within metropolitan areas.

Figure 1.

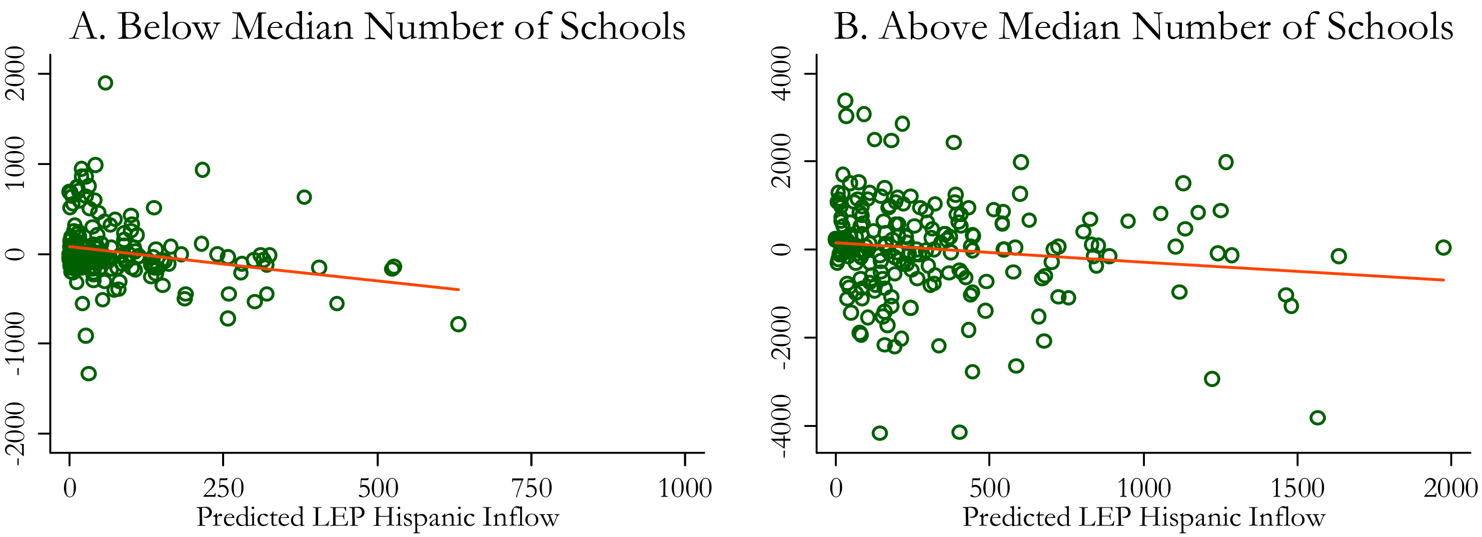
Trends in the Low-English School-Aged Population



Notes: Authors' calculations from the 1976 Survey of Income and Education and from the 5 percent Public Use Microdata Samples of the 1980, 1990, and 2000 Censuses of Population. We classify as school-aged any individual between the ages of 5 to 19 as of the time of the survey, and as "low English" any individual who "does not speak English," "yes [speaks English], but not well," or "Yes [speaks English], speaks well," since this definition generates shares low-English in the school-aged population that are comparable in the SIE and the administrative data source used in our analysis. The numbers adjacent to the orange (green) lines represent the group share in the school aged-population of the U.S. (California).

Figure 2.

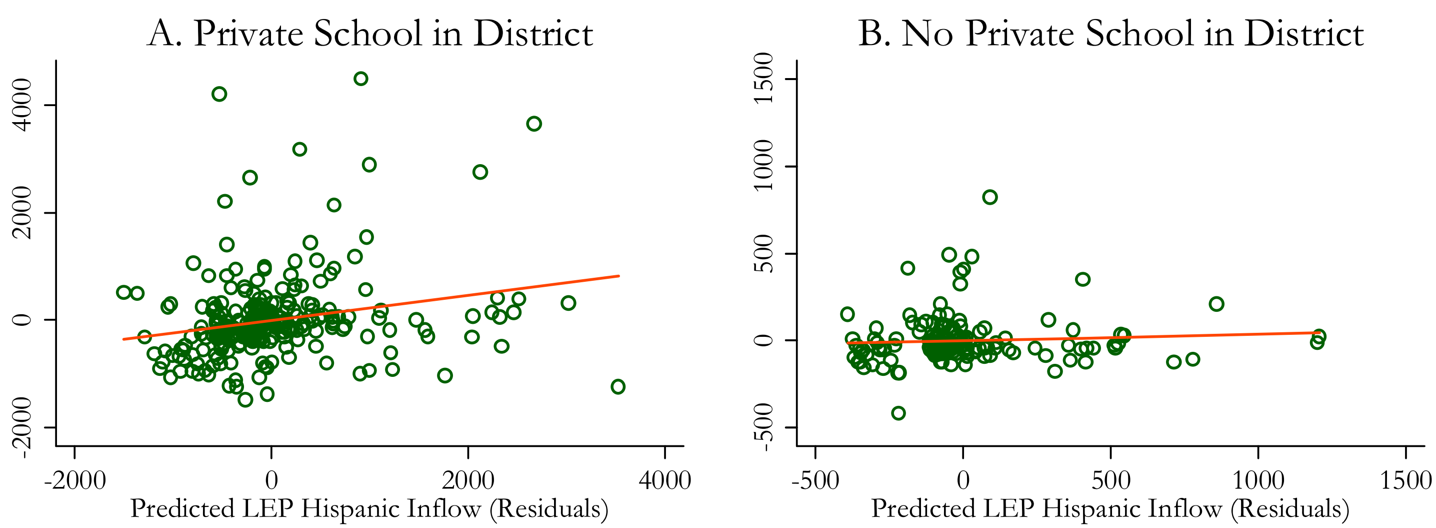
School Aged - Comparison Group Difference in 1970 to 2000 Population Change, by Predicted Change in LEP Hispanic Enrollment and District Size (Proportional Scales)



Notes: Vertical axis gives coefficients on interactions between a district indicator and whether the age group is a school-aged for the district from a regression of the 1970 to 2000 population change on these indicators, age group fixed effects, and district fixed effects. Underlying data are from the 1970 SDDB and 2000 SDDB. Horizontal axis gives the predicted inflow of LEP Hispanic public school children to a district based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled according to the 1970 SDDB, divided by the number of age groups the district serves. The slope of the line in Panel A is -0.76 (robust standard error=0.21), and the slope of the line in Panel B is -0.43 (robust standard error=0.24). The three districts with the largest predicted changes in LEP Hispanic enrollment – all of which have an above median number of schools in the early 1970s – are omitted due to scale. With these districts added, the slope of the regression line in Panel B becomes -0.27 (robust standard error=0.14).

Figure 3.

Non-Hispanic Private School Enrollment Change 1970-2000, by Predicted Change in LEP Hispanic Enrollment and Presence of a Private School as of Late 1970s (Proportional Scales)



Notes: Each point represents the residuals from regressions of the 1970 to 2000 private school enrollment change (vertical axis) and of the predicted inflow of LEP Hispanic children to a district (horizontal axis) on the average population change for the comparison age group, an indicator for whether the district has an above median number of public schools for its type, the interaction of these variables, an indicator for whether the district is an elementary district, and the interaction between this variable and the indicator for above median number of schools. The predicted inflow of LEP Hispanic public school children to a district is based on apportioning Hispanic first- and second-generation immigrants in the 2000 Census of Population who report being low English (the bottom three categories of self-reported English speaking ability) to the school districts where immigrants of all ages from the same country group settled according to the 1970 SDDB. The slope of the line in Panel A is 0.23 (robust standard error=0.10), and the slope of the line in Panel B is 0.037 (robust standard error=0.033). The three districts with the largest predicted changes in LEP Hispanic enrollment are omitted due to scale. With these districts added, the slope of the regression line in Panel A becomes 0.103 (robust standard error=0.088) and the slope of the regression line in Panel B becomes 0.006 (robust standard error=0.019).