Occupational Linguistic Niches and the Wage Growth of Latino Immigrants

Ted Mouw, *University of North Carolina, Chapel Hill*
Sergio Chavez, *Rice University*

Does the concentration of recent Latino immigrants into *occupational linguistic niches*—occupations with large numbers of other Spanish speakers—restrict their wage growth? On the one hand, it is possible that Latino immigrants who are concentrated in jobs with large numbers of Spanish speakers may have less on-the-job exposure to English, which may isolate them socially and linguistically and limit their subsequent economic mobility. On the other hand, working in linguistic niches can also be beneficial for upwardly mobile immigrants if it allows them to gain a foothold in the United States while they improve their English skills and develop labor market experience. Using data from the 1996, 2001 and 2004 panels of the Survey of Income and Program Participation (SIPP), we test for the effect of working in occupational linguistic niches on wages and wage growth. The results show that while workers in linguistic niche occupations earn lower wages on average, they do not experience lower rates of wage growth over time. Moreover, we find that about 20 percent of workers who start the 4-year SIPP panel in linguistic niches experience occupational mobility that reduces the percentage of workers speaking Spanish in their occupation by over 10 percent over the course of the study, and these "movers" have higher levels of wage growth than other workers in the sample.

Introduction

A central question in the public and academic debate on immigration focuses on the economic assimilation of recent immigrants. While conventional models of assimilation treat the low wages of recent immigrants as the first step on a ladder to upward mobility, proponents of the "segmented assimilation" perspective argue that reduced opportunities for less-educated workers in a postindustrial economy combined with phenotype discrimination may result in the downward assimilation of less-educated, darker skinned immigrants (Bean, Leach and Lowell

We thank Jackie Hagan for her feedback on earlier versions of the paper. We would also like to thank the Social Forces reviewers who offered comments that ultimately made the paper better, as well as current editor, Arne Kalleberg, and former editor Francois Nielsen. We are also indebted to scholars who provided constructive feedback at the 2008 Population Association of America.

Social Forces 91(2) 423–452, December 2012
doi: 10.1093/sf/sos186
Advance Access publication on November 6, 2012

2004; Portes and Rumbaut 2006). As the largest group of post-1965 immigrants, with relatively low-education levels and the possibility of social and labor market discrimination, Latino immigrants present an important test case for these contrasting perspectives on contemporary immigration in the United States[1].

Empirical evidence on the economic mobility of Latino immigrants paints a mixed picture. Using repeated cross-sections of Census data, researchers have documented that Latino immigrants – with the exception of Cubans – earn lower wages, on average, than their native counterparts throughout their working lives (Borjas 1982, 1985, 1995; Trejo 1997). Borjas and Katz (2005) show that the largest group of Latinos, Mexicans, lags the farthest behind the native-born in terms of wages and education, and argue that this disadvantage is transmitted across generations. Lubotsky (2007) uses longitudinal data on earnings and finds that Latino immigrants have lower rates of wage growth than other immigrant groups, and that wage convergence with native-born workers stalls after 10 years in the United States. In contrast, Smith (2003, 2006) argues that there is considerable evidence of intergenerational educational and earnings gains among Latino immigrants and that concerns about a lack of assimilation are unwarranted. Similarly, Bean, Leach and Lowell (2004) and Hall and Farkas (2008) argue that there is considerably more upward occupational and wage mobility among recent immigrants than one might expect given the expectations of racial stratification and segmented labor market theories. Overall, while there is widespread agreement that a substantial portion of the wage gap between Latino immigrants and native-born workers is due to differences in education and English-language ability at the time of immigration (e.g., see Catanzarite and Aguilera 2002), there is considerable disagreement as to the explanation of the remaining wage gap, the degree to which it persists over time, and what this portends for the future.

One explanation for the residual wage gap and the apparent lack of wage convergence for Latino immigrants focuses on occupational segregation. Catanzarite (2000) argues that Latino immigrants are crowded into "brown-collar" occupations that have been typecast as immigrant jobs where they receive low wages and have limited prospects for upward mobility. Catanzarite and Aguilera (2002:118) find that working in jobsites with co-ethnic Latinos is associated with lower wages for Latino workers, concluding that "Working at a predominately Latino jobsite lowers pay by a factor equivalent to *seven* or *eight* years of education" (italics in original). Kmec (2003:54) argues that "individuals with mostly white coworkers have an unmistakable advantage over those with mostly black or Latino co-workers." Chiswick and Miller (2005) argue that working alongside co-ethnics who speak a minority language has a feedback effect that slows the rate of economic assimilation for immigrant workers.

Without discounting the potential role that discrimination and job labeling might play in reducing the wages of workers in brown-collar occupations, we contend that the existing literature is misleading because it relies on occupation or job-level analyses with cross-sectional data and, as a result, misses the degree of upward mobility experienced by individual workers over time. In particular, we argue that recent Latino immigrants with poor English skills and limited U.S.

labor market experience may initially sort into "linguistic niche" occupations, in which English proficiency is not a prerequisite and there are large numbers of Spanish speakers. While these niche occupations pay low wages, on average, they also may provide immigrant workers with a foothold in the U.S. labor market, allowing them to accumulate job skills and improve their English-language ability. Over time, some workers may move out of these occupations to higher paying and stable employment. As a result, rather than being a dead end, these niche occupations could actually be "stepping stones" towards upward mobility, at least for a subset of workers. Based on cross-sectional data, however, it is impossible to evaluate the effect of working in these niche occupations on wage growth over time.

In this article, we use longitudinal data from the 1996, 2001 and 2004 panels of the Survey of Income and Program Participation (SIPP), to test for the effect of working in linguistic-niche occupations on wage growth for Latino immigrants. We merge our SIPP data with state-level data on Spanish language use within three-digit Census occupations from the Public Use Micro Sample (PUMS) of the 2000 Census and the 2001-2006 waves of the American Community Survey (ACS). This technique allows us to define occupational linguistic niches based on geographically specific information on the occupational distribution of recent immigrants. We compare cross-sectional and longitudinal estimates on the effect of working in linguistic niches on wages. In particular, we test whether working in occupations with a substantial proportion of Spanish speakers affects the subsequent economic mobility of Latino immigrants.

Working with Co-Ethnics: Spatial, Occupation and Firm-Level Effects on Wages

As discussed in the introduction, one prominent explanation for the Latino/non-Latino wage gap is the crowding of Latino immigrants into segregated jobs and occupations that restricts their social and economic assimilation. In this section, we review the literature on the economic effect of spatial, occupation and firm-level segregation for Latino workers before moving to our model of occupational sorting and mobility among immigrant workers.

Spatial/Enclave Effects

We begin with a discussion of the ethnic enclave debate, because it provides a useful parallel to our reconsideration of the literature on brown-collar occupations. Portes and Bach (1985), define an "ethnic enclave" as the spatial concentration of immigrant-owned firms with a predominately co-ethnic workforce. The central issue in this literature is whether ethnic enclave employment promotes or impedes the economic and social assimilation of immigrant workers (Sanders and Nee 1987; Wilson and Portes 1980; Zhou and Logan 1989). Portes and Bach (1985) and Wilson and Portes (1980) argue that ethnic enclaves allow recent immigrants with poor English and limited U.S. labor market experience to adjust to new labor market conditions. According to this perspective, ethnic enclaves can be seen as "training systems" in which ethnic networks help reduce the costs

of skill development for both workers and employers (Bailey and Waldinger 1991). The skills that recent immigrants learn in ethnic enclaves can serve as the basis for acquiring managerial-level positions or starting a business within the ethnic economy. Wilson and Portes (1980), for example, argue that enclave employment allowed earlier waves of Cubans to earn wages and receive a rate of return to human capital comparable to those employed in the primary labor market sector. A crucial component of this hypothesis is the idea that an ethnic enclave may provide an alternative to employment in the secondary sector and, as such, shelter recent immigrants from direct competition with native workers.

However, not everyone agrees that enclave employment is beneficial to immigrant workers. Employment in an ethnic enclave can be seen as detrimental to worker well-being if it slows the acquisition of English skills and prevents immigrants from developing networks that provide access to mainstream labor markets (Alba and Nee 2003; Aguilera 2009). Sanders and Nee (1987:746) argue that the enclave hypothesis assumes that "despite the social isolation of the enclave, there is no cost to segregation" and suggest that the main beneficiaries are Cuban immigrant bosses who exploit the labor of immigrant workers for their own profit. Gilbertson and Gurak (1993) find that Dominican and Columbian men employed in enclaves receive less employment benefits (e.g., health insurance and retirement benefits) when compared with workers employed in secondary labor markets.

Although they lack data on the ethnicity of firm owners, a number of recent studies have analyzed the effect of living in areas with high spatial concentrations of immigrants on labor market outcomes. Edin, Fredriksson and Åslund (2003) use initial government placement of refugees in Sweden to attempt a quasi-experimental test of the enclave hypothesis, arguing that the initial placement was independent of unobserved factors that otherwise would have influenced the location decision. They find that the earnings of less-educated refugees were 13 percent higher when the size of the ethnic enclave size was increased by one standard deviation. Damm (2009) uses a similar approach based upon a government relocation program in Denmark and finds that a standard deviation increase in the size of the ethnic enclave results in a four-percentage point increase in employment and a 21-percentage point increase in earnings. She interprets these results by arguing that ethnic enclaves provide access to ethnic networks that transmit labor market information about their host country.

In contrast to the positive effect of ethnic concentration reported by Edin, Fredriksson and Åslund (2003) and Damm (2009), Chiswick and Miller (2005) use PUMS data from the 1990 Census to test a version of the enclave hypothesis based on linguistic concentration. Using a measure of the proportion of speakers of the immigrant's language group at the state level, they find that linguistic concentration is associated with both lower levels of English proficiency and lower earnings for immigrant workers, which makes sense given the evidence that English proficiency is a critical component of the wages for immigrant workers. (See, for example, Dustmann and Fabbri 2003; Kossoudji 1988; McManus 1985; McManus Gould and Welch 1983.) Chiswick and Miller (2010) use Census PUMS data matched to occupational-level information on

English-language requirements to test a variant of the enclave hypothesis, and they find that there is a strong positive correlation between occupational English requirements and wages in the cross-section. Finally, Warman (2007) uses synthetic cohorts constructed from Canadian Census data from 1981-2001 and finds that living in an ethnic enclave is negatively associated with wage growth for immigrant groups.

Overall, the methods used in the quasi-experimental studies of Edin, Fredriksson and Åslund (2003) and Damm (2009) discussed above would seem to be preferable to those in Chiswick and Miller (2005) and Warman (2007). The latter studies are based on cross-sectional or synthetic-cohort data and, hence, are vulnerable to problems of self-selection based on English ability, as discussed in our theoretical model below. On the other hand, the relatively small size of the ethnic enclaves in the data from Sweden and Denmark that Edin, Fredriksson and Åslund (2003) and Damm (2009) use might point to a qualitatively different process among larger and more concentrated ethnic enclaves in the United States or Canada.

Overall, the debate on the ethnic enclave hypothesis – either based on employment in spatially concentrated immigrant-owned firms or ethnic concentration more generally – highlights the complexity of the immigrant economic assimilation process. Particularly with respect to language, working alongside co-ethnics may provide an entrée into the U.S. labor market for recent immigrants with poor English-language skills. As Chiswick and Miller (2005:10) note "working within a linguistic enclave is a mechanism for sheltering oneself from or mitigating the adverse labor market consequences of limited destination language proficiency."

Occupation and Firm-Level Effects

A number of papers focus specifically on the link between firm-level and occupation-level segregation among Latino workers and the Latino/white pay gap. Catanzarite (2000) uses PUMS data from Los Angeles for 1980 and 1990 to show that Latinos are highly segregated from white workers, and that their earnings are lower than whites' even after controlling for education, potential labor market experience, English ability and family composition. She uses the term brown-collar occupations to refer to those occupations with a high proportion of Latino workers, and argues that labeling them as "immigrant" jobs simultaneously makes them less desirable, lowers their status and reduces wages. Once immigrants cluster in these occupations, their networks control the labor process, everything ranging from hiring to training (Waldinger and Lichter 2003). In a follow-up study, Catanzarite (2003) uses 1990 PUMS data from 18 metropolitan areas to estimate a multilevel model of the effect of immigrant density on wages, and finds that working in an occupation with recent Latino immigrants reduces the wages of all workers – not just immigrants – though the negative effects are more pronounced for blacks and earlier Latino immigrants.

In a related study, Catanzarite (2002) uses 1980 and 1990 Census data from Los Angeles to test the relationship between earnings and Latino concentration

at the occupational level using a cross-lagged regression model. She finds a negative relationship between native worker's occupational earnings in 1990 and the percent of male immigrant Latino workers in the occupation in 1980. In contrast to Catanzarite's (2002) results, Howell and Mueller (2000) use PUMS data from the New York metropolitan area to estimate models of the change in wages on changes in the proportion immigrant in occupations from 1980-1990. They find no effect of changes in the proportion immigrant on changes in the wages of recent immigrants or Latino workers, arguing that there is no evidence of a negative causal relationship between wage growth and change in the proportion of immigrant workers at the occupational level (2000:488).

In addition to effects at the occupational level, it is possible that the negative effect of immigrant concentration on wages is more pronounced at the job or firm level. Catanzarite and Aguilera (2002) use data from the 1992 Legalized Population Survey, which includes a categorical variable asking respondents to identify the largest race/ethnic group among their coworkers. They found that, on average, legalized-Latino males tend to earn 13 percent less when employed at predominately Latino jobsites, even after controlling for occupational characteristics and an extensive set of human capital variables. Similarly, Kmec (2003) uses data on the race and ethnic composition of jobs from the employer survey of the Multi-City Study of Urban Inequality, and finds that working in jobs with predominately Latino or black coworkers reduces the wages of all workers by 18 percent and 15 percent less per hour, respectively, compared with jobs in which whites are the majority. In segregated jobsites, Latina/os experience further wage penalties when skin color and nativity status are accounted for, thus suggesting structural inequality in accounting for labor market outcomes (Morales 2009).

In perhaps the most comprehensive test of the correlation between Latino concentration and wages at the firm level, Hellerstein and Neumark (2002) use a large data set of matched employer-employee records constructed from the Decennial Census of 1990. They document substantial firm-level segregation of Latino workers by ethnicity and the degree of English proficiency. Their regression results for log wages show that the share of coworkers who are Latino reduces wages by 0.168 log points for Latino workers and .037 log points for white workers.

Although the job and occupational segregation literature argues that working with co-ethnics has a negative effect on Latino immigrants' wages, the existing evidence relies on cross-sectional data or aggregate analysis at the occupational level. While it may be true, as this literature suggests, that the crowding of recent immigrants into brown-collar occupations reduces wages in those occupations and represents a structural constraint on race and ethnic equality, the overall portrayal of Latino immigrants' economic assimilation is too static because it doesn't look at the wage trajectories of individual workers. In particular, a lack of longitudinal data does not permit an accurate assessment of the degree to which some immigrant workers move out of segregated occupations over time. If substantial numbers of successful immigrants do experience upward mobility, then any cross-sectional estimate of the effect of working in a brown-collar

occupation on wages will be biased by the negative selectivity of the remaining workers.

In the following section, we present a simple model of occupational sorting among immigrant workers. We argue that although immigrants may initially "sort" into occupational niches, this ethnic concentration is not necessarily bad during an initial period of adjustment. This sorting helps newcomers to develop language and occupational skills and is not detrimental so long as this strategy is used as a "stepping stone" to mainstream labor markets.

A Theoretical Model of Occupational Linguistic Niches

For recent Latino immigrants, a lack of English fluency and limited knowledge about opportunities represent major constraints in the U.S. labor market. Given these constraints, "occupational linguistic niches" – occupations with large numbers of Spanish-speaking workers – provide employment for Latino immigrant workers with insufficient English ability, either where the job requirements for English fluency are minimal or the presence of large numbers of Spanish speakers eases the language difficulties. The key question is whether the sheltering effect of working in linguistic niche occupations reduces wage growth by slowing the process of linguistic and social assimilation.

Equations 1 and 2 present this sorting argument more formally. In Equation 1, we depict log wages for immigrant i in occupation j at time t as a function of the degree of concentration of Spanish speakers in the occupation:

$$\ln w_{ijt} = \beta_1 \text{occ-Spanish}_j + \beta_2 \textit{English}_i + \beta_3 X_i + \alpha_i + \varepsilon_{it} \quad (1)$$

Where occ-Spanish_j is the proportion of workers in occupation j who speak Spanish, $\textit{English}_i$ is the worker's English-language proficiency, X is a set of other observed individual level control variables, ε_{it} is an error term, and α_i represents fixed unobserved factors that affect wages.

In our theoretical model, we hypothesize that α_i represents traits such as ambition and skills that are not measured on typical surveys or adequately proxied by educational credentials, but observed by employers and rewarded in the labor market. In Equation 1, we expect that working in an occupation with a high concentration of Spanish speakers is correlated with lower wages ($\beta_1 < 0$). We argue that the expected negative correlation between the percentage of Spanish speakers in an occupation and wages is not causal – in other words, there is nothing about speaking Spanish itself that reduces wages – but because occupations with large numbers of Spanish speakers tend to have lower skill requirements, as indicated by our discussion of Tables 1 and 4 below. In addition, workers with better English ability should have higher wages ($\beta_2 > 0$).

In Equation 2, we present a simple model of occupational sorting based on English-language proficiency and unobserved productivity:

$$\text{occ-Spanish}_i = \eta_1 \textit{English}_i + \eta_2 \alpha_i + \nu_{it} \quad (2)$$

Table 1. Descriptive Statistics for Latino Immigrants using the 2005-6 American Community Surveys

Panel A: English speaking ability (Latino immigrants)

English ability	Proportion Spanish Speakers in Occupation	Proportion Latino in Occupation	O*NET Occupational English Level	Average Hourly Wage	(N)
Speaks only English	.200	.228	2.260	18.68	6,879
Speaks very well	.196	.220	2.395	18.98	45,941
Speaks well	.278	.305	1.810	14.58	38,475
Not well	.354	.383	1.493	11.36	42,856
Does not speak English	.414	.443	1.395	9.93	24,263

Panel B: Years Since Immigration (Latino Immigrants)

Years Since Immigration	Proportion Spanish Speakers in Occupation	Proportion Latino in Occupation	O*NET Occupational English level	Average Hourly Wage	(N)
0-5 years	.312	.334	1.648	11.18	27,795
6-10 years	.303	.327	1.697	12.23	27,262
11-15 years	.303	.330	1.762	13.16	22,181
16-20 years	.304	.334	1.808	14.24	25,254
21+ years	.268	.297	2.078	17.90	55,922

Panel C: Linguistic Niche Category (% Spanish Speakers in Occupation, Latino Immigrants)

			Average O*NET Skill Requirements (in Years)			
	Average Wage	O*NET Occupational English Level	On the Job Training	Specific Occupational Experience	Required Education Level	(N)
A (Less than 10%)	19.26	2.48	.773	2.459	13.669	30,278
B (10-19%)	16.55	2.30	.674	2.116	13.163	39,613
C (20-30%)	12.80	1.60	.704	1.575	12.187	25,570
D (>30%)	11.46	1.35	.698	1.546	11.883	62,942

The benefit of a niche occupation with a high proportion of Spanish speakers is that it provides employment for Latino immigrants who are not fluent in English; hence we would expect a negative value for η_1. The coefficient $\eta_2$depicts the effect of unobserved factors that affect wages such as "ambition" on sorting into enclave occupations. If occupations with a high proportion of Spanish

speakers tend to be lower skilled occupations in general, or if more skilled (or ambitious) immigrants learn English more rapidly, then we would expect a negative relationship between occupational Spanish and the unobserved individual-level skills that affect wages, i.e., $\eta_2 < 0$.

Referring back to Equation 1, we can develop an intuition about how skill-based occupational sorting in Equation 2 will affect our coefficients in Equation 1. A negative correlation between the unobserved factor α_i and occupational Spanish (as hypothesized in Equation 2) will tend to result in a downward bias on the coefficient on occupational Spanish in Equation 1, as immigrants with less "ambition" or lower unobserved skills stay longer in linguistic niche occupations. If this kind of negative sorting is taking place, then regression estimates of β_1 will overstate the negative effect of working in a linguistic niche occupation.

If we are worried about the possibility that cross-sectional data may overstate the effect of occupational Spanish on wages because of sorting, an alternative approach is to use longitudinal data to model wage growth rather than wage levels. If occupational linguistic niches restrict economic assimilation by delaying English-language acquisition or other skills necessary for upward mobility, then this should result in a negative effect of niche occupations on subsequent wage growth. This is depicted in Equation 3:

$$\frac{\Delta \ln w_{ijt}}{\Delta time} = \alpha + \phi_1 \text{occ-Spanish}_{j1} + \phi_2 Z_i + \varepsilon_{it} \quad (3)$$

where the dependent variable is the change in wages over time, $occ - Spanish_{i1}$ is the level of occupational Spanish in the first wave of data, and Z_i represents a set of relevant control variables. If working in a niche occupation constrains wage growth, then we would expect that $\phi_1 < 0$. In contrast, the "sorting" argument claims that although linguistic niche occupations are associated with lower wages in the cross section – because of the sorting of workers with poor English into those occupations – they do not affect the subsequent wage growth of immigrants workers; hence $\phi_1 = 0$. In other words, the test is quite simple: do immigrants who work in occupational niches have lower rates of subsequent wage growth than other immigrants?

Data and Methods

In this article, we analyze the effect of occupational linguistic niches on the wages of foreign-born Latino workers who immigrated to the United States as adults (aged 17 years or older) using data from the SIPP. Overall, the key advantage of the SIPP, compared with other longitudinal labor market data sets such as the Panel Study of Income Dynamics and the National Longitudinal Study of Youth, is that it includes a large number of Latino workers, which allows us to estimate the effect of occupational niches on wage growth over time. Because the nativity and age restrictions reduce the sample size, we combine data from the 1996, 2001 and 2004 panels of the SIPP.[2] The 1996 and 2004 panels of the SIPP were 4-year longitudinal surveys that began with a nationally representa-

tive sample of 40,188 and 51,400 households, respectively, and the 2001 SIPP was a 3-year longitudinal survey comprising 36,700 households. In the SIPP, households are interviewed once every 4 months, resulting in 12 waves of data collection for the 1996 and 2004 panels, and nine waves of data for the 2001 panel. [3] Because of funding constraints and debate over the future of the SIPP, the sample size for the 2004 panel was reduced to 21,300 households in Wave 9 (National Academy Press 2009). Because the 1996 SIPP panel does not have data on self-reported English ability, we used Stata's ICE command (see Royston 2009) to impute the Wave 1 English ability for the 1996 panel using multiple imputations with all of the variables listed in Table 2. [4]

In a related paper, Hall and Farkas (2008) use data from the 1996 and 2001 panels of the SIPP to estimate growth-curve models of wage growth among immigrants and native workers. They find that while initial wages are considerably lower for immigrants compared with native workers, the estimates of wage growth are statistically indistinguishable among natives and different immigrant groups.

We supplement the individual-level data from the SIPP panels with aggregate occupation-level variables on Spanish language use and occupational skill requirements. First, the variable "occupational Spanish" measures the proportion of workers in a three-digit Census occupation who report speaking Spanish at home, aggregated at the state level. We construct this variable using data on employed workers older than 16 years of age from the 2000 Census PUMS and the 2001-2006 waves of the ACS. The 2000 PUMS is a 5 percent sample of the U.S. population and the 2005 and 2006 ACS are 1 percent samples of the U.S. population (the 2001-2004 samples comprise 1 / 232, 1 / 261, 1 / 236 and 1 / 239 size samples, respectively). The benefit of the combined 2000-2006 PUMS-ACS sample is that it provides a large number of cases, which allows us to calculate the proportion of Spanish speakers within each three-digit Census occupation for each state.

The state-level variation by occupation is important because it allows us to take regional variation in occupational composition into account. Overall, the aggregate state by occupation-level measure of the proportion of Spanish speakers is highly correlated (above the .9 level) with the proportion of workers in who report speaking English either not well or not at all and the proportion of Latino workers in the respondent's occupation. Therefore, our results would be very similar if, instead of the proportion of workers in the respondent's occupation who spoke Spanish, we used either the proportion who were co-ethnics or the proportion who were limited-English speakers. There are geographic differences in the degree to which certain occupations function as occupational linguistic niches; for example, the proportion of carpenters who speak Spanish may be higher in Texas and California than in South Dakota. We would like to go to a more detailed level of geography, as the 2000-2006 PUMS-ACS would allow us to go the level of a Public Use Micro Area (about the size of a county), but the smallest level of geography in the SIPP is the state. We add the variable "occupational Spanish" to our SIPP data by merging it at the state and occupational level. Overall, for the average respondent in our SIPP data, an average of 5,454 cases from the combined 2000-2006 PUMS-ACS data are used to calcu-

late the proportion of Spanish speaking workers at the occupation level within each state.

To include information on the education, skill, and English-language requirements of an occupation, we use data from the O*NET database version 15.0, which is the most current version of the successor to the Dictionary of Occupational Titles.[5] We use the 5-point scale on the importance of English in the occupation, ranging from 1 (*not important*) to 5 (*extremely important*). In addition to English-language requirements, we calculate the average number of years of required education, on-the-job training and related occupational experience by calculating the mean value of the responses to the O*NET occupational skills survey as reported in the O*NET 15 database. We convert the O*NET data from Standard Occupational Classification codes to the Census 2000 occupational codes using the occupational crosswalk provided by the IPUMS website (http://usa.ipums.org/usa/volii/census_occtooccsoc.shtml).

To provide a formal test of the effect of linguistic niches on wage growth, we turn to growth curve and fixed effects models using our combined sample of SIPP data. The growth curve model is estimated using the command *xtmixed* in Stata by treating time as a random coefficient and including interaction terms between time and selected covariates. (See Rabe-Hesketh and Skrondal 2005 for a more complete discussion of growth curve models.) The benefit of the growth curve model is that it takes advantage of the longitudinal SIPP data to model both wage levels and wage growth. Hall and Farkas (2008) provide an example of using a growth curve models to study immigrant earnings trajectories. A basic depiction of the growth curve model we estimate is as follows: first, in Equation 4, we are modeling log wages of individual i at time t with a random intercept, β_0, and slope, β_1 where ε_{it} is a standard error term.

$$\ln w_{ijt} = \beta_0 + \beta_1(\text{time}_{it}) + \varepsilon_{it} \tag{4}$$

In Equation 5, we model the intercept as a function of sets of observed covariates X and Z, along with a person specific random effect, μ_i.

$$\beta_0 = \alpha_0 + \alpha_1 X_{it} + \alpha_2 Z_{it} + u_i \tag{5}$$

Finally, in Equation 6, we model the effect of time on wages with a constant, a subset of our observed covariates, Z, and a person specific random effect:

$$\beta_1 = \delta_0 + \delta_1 Z_{it} + \phi_i \tag{6}$$

In all of the models we estimate, we interact the initial linguistic niche category with time, which allows us to test for the effect of working in an enclave occupation on subsequent wage growth.

Results

Table 1 presents cross-sectional descriptive statistics on Latino immigrants from the 2005-2006 waves of the American Community Survey. This table is important because it clarifies the basic parameters of the debate on the effect of working in occupations with high proportions of Spanish speakers or Latinos and provides the motivation for our analysis of longitudinal data with the SIPP. In this table, the proportion of Spanish speakers and the proportion Latino in the respondent's three-digit occupation are calculated at the state level from the ACS data, and the O*NET skill requirements are averages of survey responses at the occupation level.

Panel A of Table 1 presents descriptive statistics by respondent's English ability. English ability among Latino immigrants is clearly related to sorting into occupations with more Spanish speakers. Among workers who reported not speaking English, on average, they worked in occupations with 41.4 percent Spanish speakers, in contrast to 20 percent for workers who reported only speaking English, and 27.8 percent for those who reported speaking English "well." In the next column, the proportion Latino workers by English ability is almost identical to the proportion Spanish speakers, which should not be surprising, as these two measures are correlated at the 0.98 level in the data. The O*NET occupational English level indicates the importance of English to perform the job, on a 5-point scale ranging from 1 being the lowest level of English requirement to 5 being the highest. Finally, wages are clearly linked to English ability, with an average wage of $9.93 for workers who do not speak English, and between $18-$19 for workers who speak only English or who speak it "very well." Overall, Panel A indicates that as English ability improves, the proportion of Spanish-speaking workers in the occupation goes down and wages go up.

Panel B of Table 1 presents the same statistics by years since immigration. In contrast to the findings on English ability, here we find essentially no change in the proportion of Spanish speakers or fellow Latino workers in the respondent's occupation over the first 20 years since immigration. Keeping in mind that it is perilous to infer individual trajectories for immigrants from cross-sectional data – since we are comparing across different immigration cohorts (which is why we will next turn to longitudinal data) – Panel B suggests stagnation in terms of the crowding of Latino immigrants into occupations with large numbers of Spanish speakers and Latinos over time. In contrast, however, average hourly wages are 27 percent higher for workers who immigrated 16-20 years ago compared with those who immigrated 0-5 years ago.

Finally, Panel C presents summary statistics based on our proposed "Linguistic Niche" categorization based on the percentage of Spanish speakers in the respondent's occupation at the state level (less than 10%, 10%-20%, 20%-30%, and greater than 30%). In this table, we see that, in the cross section, there are clear differences in average wages and occupational skill requirements (education and related occupational experience) across the difference niche categories. It is our contention that less than fluent English-language ability sorts immigrant workers into occupations with lower English-language requirements, and that these

occupations tend to pay lower wages because they have lower skill requirements in general (as evidenced by the education and experience levels).

The juxtaposition of Panels A, B and C provide the empirical motivation for our analysis of the wage trajectories of immigrant Latino workers using the SIPP panels. Panel A shows that workers sort into Spanish speaking occupations on the basis of English ability – consistent with our theoretical perspective – and Panel C shows that working alongside Spanish speakers is associated with lower cross-sectional wages and lower occupational skill levels. On the other hand, Panel B suggests that there is continued wage growth for Latino workers since immigration, even though the overall level of crowding into Spanish-speaking and Latino occupations holds constant. This is consistent – at an aggregate level – with the idea that skill development and wage growth is possible, even though workers may have sorted into linguistic niche occupations. Overall, the advantage of the 1996, 2001 and 2004 SIPP panels, compared with the cross-sectional ACS data, is that it follows individuals over time, which allows us to analyze trends in wage growth based on initial occupational category. When we move to our growth curve models, we want to test whether working in a linguistic niche occupation negatively affects wage growth.

In Table 2, we use our SIPP sample to present an analysis of the change in the proportion Spanish speakers in workers' occupations based on an initial classification based on the proportion of Spanish speakers in those occupations at the state level (less than 10%, 10%-19%, 20%-35% and greater than 35% Spanish speaking).[6] The columns of Table 2 indicate the worker's linguistic niche category at Wave 1 of the SIPP, and the rows indicate the degree of change by Wave

Table 2. Change in % Spanish Speakers in Respondent's Occupation Between Waves 1 and 12 (by Column)

Change in Occupation % Spanish Speakers	Occupational Linguistic Niche Category in Wave 1.			
	A (≤10%)	B (11-20%)	C (21-30%)	D (>30%)
-100 to -20%	.00	.00	1.99	11.47[a]
-20 to -10%	.00	1.57	6.62	8.68[a]
-10 to -5%	1.36	5.49	3.97	5.00[b]
-5 to -1%	10.88	8.24	5.30	3.09[b]
0	57.14	63.92	56.95	56.32
+1 to 10%	20.41	10.59	7.28	6.03
+10 to 20%	6.12	4.31	3.31	5.00
+20 to 100%	4.08	5.88	14.57	4.41
Total	100	100	100	100
(N)	(147)	(255)	(151)	(680)

[a]Group D2
[b]Group D3

12 (the final year of the 1996 and 2004 panels). Focusing on niche category D (workers in occupations with more than 30 percent Spanish speakers), we see that there is evidence of considerable mobility over the study period. Only 56.32 percent of workers experience no change in their occupational % Spanish over the study period. About 20 percent have a decline of at least 10 percent: 11.47 percent had a decline of more than 20 percent in occupational Spanish, and 8.68 percent had a decline of between 10 percent and 20 percent. For the purposes of the growth curve model presented below, we combine both of these two categories of mobility for group D into the subgroup D2, as indicated in Table 2. Next, 8.09 percent of workers in Group D had a decline of between 1 percent and 10 percent in their occupational Spanish, and we place them in subgroup D3.

In addition to mobility for workers who start in niche category D, there is also evidence of up – and down – mobility for workers who start in other occupational categories. For example, 15.3 percent of workers in group B experience a reduction in occupational Spanish (1.57 + 5.49 + 8.24), and 20.78 percent experience an increase. Overall, the results in Table 2 are important because they indicate that Latino immigrant workers who start out in occupations with high concentrations of Spanish speakers – category D in Table 2 – experience a substantial amount of mobility over time, with 20 percent reducing their occupational Spanish concentration by over 20 percent. This suggests prima facie evidence that many of these immigrant workers are not "trapped" in these linguistic niche occupations. Nonetheless, is important to qualify this finding: the majority of workers from group D do not move "up" by decreasing their occupational % Spanish, at least over the 4 years of the panel data. This is, however, consistent with our theoretical perspective as discussed above in reference to Equations 1 to 3: some workers, through a combination of ambition, skill and luck, may use these niche occupations as "stepping stones" even as the majority of immigrant Latino workers stay in these niche occupations over extended periods of time.

Table 3 presents summary statistics for the variables used in our empirical models. After selecting all Latino immigrant workers who migrated to the United States at 16 years of age or older, we have 29,548 observations representing 4,446 unique individuals in the combined 1996, 2001 and 2004 SIPP data. Years of potential labor market experience is calculated as age-(years of education + 6). The occupational-linguistic niche is categorized as described above in Table 2. For the growth curve analysis, we use the linguistic niche category for the first wave of the sample, because we want to analyze the effect of linguistic niches on subsequent wage growth.[7] As described above in reference to Table 2, categories D2 and D3 represent "niche movers" who started in category D but experienced a decline in their occupational Spanish concentration of more than 10 percent and 1 percent to 10 percent, respectively.[8] In addition to occupational niche, Table 3 also shows the respondent's self-reported English-language ability in Wave 1. As discussed above, this variable is not included in the 1996 panel, so we imputed it using multiple imputation with five replications.[9] Finally, Table 3 presents occupational level variables. In addition to the proportion Spanish speakers (at the occupation level by state) constructed from the 2000-2006

Table 3. Summary Statistics for Variables used in the Analysis 1996, 2001 and 2004 SIPP data

Variable	Mean	*SD*
Individual-Level Variables		
Log wages	2.39	.431
Years of potential labor market experience	23.37	11.18
Potential experience squared	671.36	589.66
Years of education	11.12	2.92
Female	.399	
Years since immigration	13.92	9.39
Initial Linguistic niche category (% Spanish speakers in Wave 1 occupation)		
A (Less than 10%)	.140	
B (10-19%)	.210	
C (20-30%)	.130	
D (>30%)	.520	
D2 (D + decline > 10%)	.096	
D3 (D + decline 1-10%)	.037	
How well speaks English		
Speaks only English	.092	
Very well	.177	
Well	.185	
Not well	.355	
Not at all	.191	
Occupation-Level Variables		
Occupation proportion Spanish speakers[a]	.340	.218
Level of importance of English-language ability[b]	1.61	.828
Required years of firm-specific training[b]	.586	.410
Required years of on-the job training[b]	.644	.507
Required years of specific occupation experience[b]	1.64	1.08
Required years of education[b]	12.29	1.204
Number of observations	29,548	
Number of individuals	4,446	

Note: SD = standard deviation.
[a]PUMS and ACS data at the occupation and state level.
[b]O*NET 15 occupational database.

PUMS-ACS data, we include measures of occupation skill requirements from the O*NET occupational data base, as discussed above in the data section.

Table 4. Regression Models of Wave 1 Wages, Occupation-Specific and Individual-Level Variables

	(1)	(2)
Variables	**Log Wages**	**Log Wages**
Occupation-Specific Variables		
Occupation % Spanish speakers	-.363***	-.110**
	(.0362)	(.0383)
Required level of firm-specific training		-.0501
		(.0340)
Required level of on-the job training		.0964***
		(.0249)
Required years of specific experience		.0537***
		(.0104)
Required years of education		.0562***
		(.00751)
Worker-Specific Variables		
Work experience	.00808***	.00686***
	(.00218)	(.00208)
Work experience^2	-.000114**	-9.60e-05**
	(4.03e-05)	(3.85e-05)
Female	-.224***	-.176***
	(.0125)	(.0129)
Ln(years in the United States)	.0924***	.0874***
	(.00993)	(.00945)
Years of education	.0251***	.0177***
	(.00247)	(.00243)
How well R speaks English (excluded category: speaks only English)		
Very well	.110**	.101**
	(.0369)	(.0347)
Well	.0288	.0443
	(.0362)	(.0347)
Not well	-.0689**	-.0470*
	(.0277)	(.0257)
Not at all	-.101**	-.0864**
	(.0369)	(.0343)
Panel: 1996	-.103***	-.109***
	(.0159)	(.0152)
2001	-.0329*	-.0382**
	(.0154)	(.0147)
Dummy variables for Census region	Yes	Yes
Constant	2.097***	1.332***
	(.0629)	(.0992)
Observations	4323	4323

***p < .01 ** p < .05 * p < .1

Note: Standard errors in parentheses.

Table 4 presents OLS models of Wave 1 log wages based on individual and occupation-specific characteristics. Model 1 shows that the occupation proportion Spanish is negatively associated with wages, with a coefficient of -0.363, after controlling for potential work experience, time since migration, education and self-reported English ability. Model 2 adds measures of occupational skill requirements from the O*NET data for firm and job-specific training, specific occupational experience and education levels. A comparison of the coefficient on occupational % Spanish across these two models shows that it declines by about 70 percent from Model 1 to Model 2 (-.363 to -.110), suggesting that much of the negative effect of occupational Spanish is a function of its association with these measures of occupational skill requirements. In other words, Table 4 indicates that, to a large extent, the negative correlation between occupational % Spanish and wages in Model 1 is not a direct causal effect of working alongside Spanish-speaking workers per se, but because, in general, these occupations tend to have lower skill levels. In Model 2, the remaining effect of occupational % Spanish on wages net of the occupation-level skill variables (-0.111) could be interpreted as the additional wage effect of crowding Latino immigrant workers into occupations with lower English-speaking requirements, above and beyond the intrinsic skill levels of those occupations.

Table 5 presents logit models of the occupational mobility out of the linguistic niches. Here, we model the probability of being in group D2 from Table 2 (a decline in % occupation Spanish of 10 percent or more over the course of the study) conditional on starting in category D (>30% occupational Spanish). Model 1 controls for gender, time since immigration, education, and Wave 1 English-language ability. Model 2 adds a variable for the change in the respondent's self-reported English-language ability over the course of the study, calculated as the final reported English-language ability minus the initial level. Model 2 is estimated only for the 2001 and 2004 panels, as we imputed only the Wave 1 English level for the 1996 data. The striking finding of Table 4 is that none of the independent variables are statistically significant at the .05 level of significance. Although Table 1 shows English-language ability is strongly associated with sorting into Spanish-language concentrated occupations in the ACS data (identical results obtain for a cross-sectional analysis with the SIPP data, available on request), self-reported language ability and the change in language ability are not associated with a greater probability of mobility out of category D. The null result for the change in language ability could be due to measurement error in the reporting of language ability, which would be more pronounced in trying to use the variable to identify changes in language ability over time. Overall, however, the results in Table 4 are consistent with the theory that mobility out of linguistic niches is based on unobserved variables such as ambition, ability or luck that are not measured well on standard labor market surveys.

Table 6 presents the results of our growth curve models for the 1996, 2001 and 2004 SIPP data. The results for each model in Table 5 are presented in two panels. The "Levels" panel presents coefficients for wage levels (the model for the intercept terms in Equation 5), while the "Slopes" panel presents coefficients for the individual slope of wage growth over time (the model for wage growth

Table 5. Logit Models of Mobility Out of Linguistic Niche Occupations

Model Number	(2)	(3)
	Mobility Out of Spanish-Intensive Wave 1 Occ.[a]	
Dependent Variable	Decrease in % Spanish Speakers of >10%	Decrease in % Spanish Speakers of >10%[b]
Female	-.0336	.178
	(.122)	(.173)
Ln(years in the United States)	-.0336	.134
	(.0727)	(.103)
Years of education	.0326	.0242
	(.0238)	(.0275)
How well speaks English (excluded category: speaks only English)		
Very well	.312	.317
	(.344)	(.372)
Well	.206	.312
	(.424)	(.358)
Not well	.0351	.0911
	(.312)	(.331)
Not at all	-.0269	.119
	(.343)	(.358)
Panel: 1996	.0433	
	(.151)	
2001	-.173	-.225
	(.159)	(.181)
Change in how well speaks English		.0350
		(.0667)
Constant	-1.825***	-2.276***
	(.432)	(.487)
Observations	2,101	1,116

*** p < .01 ** p < .05 * p < .1

Note: Standard errors in parentheses.

[a]Models 3 and 4 use only workers who started in Wave 1 in occupations with more than 30% Spanish speakers.

[b]Model 4 uses only the 2001 and 2004 panels.

depicted in Equation 6). The slope coefficients measure the effect of time and the interaction effects of selected independent variables with time.

For all the models in Table 6, category D (greater than 30% Spanish speaking) is the excluded category. As discussed above, the key test of the paper is whether workers in niche occupations (category D) have lower rates of wage

Table 6. Growth Curve Model of Log Wages, Immigrant Latino Workers

Variables	(1)	(2)	(3)	(4)
Slopes	-.00739*	-.00560	.00135	.00670
Time (Years)	(.00419)	(.0112)	(.0148)	(.0146)
Linguistic niche category (% Spanish in occupation). Excl. category: D(>30%)				
A. 0-10%	.00192	4.65e-05	-.00156	-.00889
	(.00595)	(.00614)	(.00637)	(.00725)
B. 11-20%	.000247	-.000794	-.00184	-.00216
	(.00509)	(.00523)	(.00540)	(.00575)
C. 21-30%	.00191	.000896	.000432	-.000836
	(.00605)	(.00609)	(.00617)	(.00678)
D2. D + decrease of 10% or more	.0185***	.0181***	.0180***	.0187***
	(.00657)	(.00656)	(.00657)	(.00652)
D3. D + decrease of .1 to 10%	.0117	.0113	.0113	.0118
	(.00984)	(.00984)	(.00985)	(.00977)
E. A-C + increase of 1% or more				.00448
				(.00637)
Years of education		.000817	.000641	-.00337
		(.000679)	(.000706)	(.00310)
Female		-.00596	-.00600	-.00659
		(.00384)	(.00393)	(.00391)
Log(years in the United States)		-.00271	-.00335	.000646
		(.00303)	(.00309)	(.000699)
Panel: 1996	.0302***	.0294***	.0295***	-.109***
	(.00443)	(.00453)	(.00455)	(.0131)
2001	.00597	.00479	.00447	-.0240*
	(.00495)	(.00502)	(.00504)	(.0131)
How well speaks English	Not included	Not included	Included	Included
Levels				
Linguistic niche category A. 0-10%	.211***	.212***	.213***	.262***
	(.0191)	(.0191)	(.0192)	(.0211)
B. 11-20%	.115***	.116***	.116***	.133***
	(.0160)	(.0160)	(.0161)	(.0179)
C. 21-30%	.0492***	.0501***	.0505**	.0641***
	(.0173)	(.0174)	(.0174)	(.0189)

Table 6. Continued

Variables	(1)	(2)	(3)	(4)
D2. D + decrease of 10% or more	-.0209	-.0206	-.0205	-.0164
	(.0188)	(.0188)	(.0188)	(.0188)
D3. D + decrease of .1 to 10%	-.0455	-.0452	-.0453	-.0396
	(.0292)	(.0292)	(.0292)	(.0293)
E. A-C + increase of 1% or more				.0675***
				(.0188)
Work experience	.0100***	.00978***	.00972***	.00934***
	(.00160)	(.00161)	(.00161)	(.00160)
Work experience^2	-.000147***	-.000140***	-.000139***	-.000137***
	(2.89e-05)	(2.93e-05)	(2.93e-05)	(2.94e-05)
Years of education	.0141***	.0134***	.0136***	.0138***
	(.00150)	(.00163)	(.00164)	(.00167)
Female	-.219***	-.214***	-.214***	-.213***
	(.0103)	(.0107)	(.0107)	(.0113)
Log(years in the United States)	.0815***	.0810***	.0816***	.0827***
	(.00769)	(.00769)	(.00775)	(.00792)
How well speaks English	Included	included	included	included
Dummy variables for Census region	Yes	Yes	Yes	Yes
Constant	2.046***	2.053***	2.047***	2.053***
	(.0465)	(.0471)	(.0489)	(.0515)
Observations	29,548 (4,446)	29,548 (4,446)	29,548 (4,446)	29,548 (4,446)

*** p < .01 ** p < .05 * p < .1
Note: Standard errors in parentheses.

growth compared with other workers. To test this, we look at the slope panel in Table 6. Model 1 is a baseline model with only the linguistic niche category and indicator variables for the SIPP panel used as explanatory variables for the slopes. In Model 1 we find that there is no evidence that wage growth was any higher for workers who started in categories A, B or C compared with category D: the coefficients on these variables are close to 0 and statistically insignificant. In contrast, workers in group D2 – those who started in D but experienced a decrease in their occupational Spanish of 10 percent or more – experienced a wage growth of 1.85 percent more per year than group D.

Model 2 adds education, gender, and time since immigration as controls for the slopes of the wage trajectories. As with Model 1, there is no evidence that workers who start in category D have lower rates of wage growth than A, B or C, and niche movers (D2) have a rate of wage growth that is 1.8 percent higher per year. Finally, Model 3 allows the slope of wage growth to depend upon English proficiency, in addition to the initial occupational linguistic category, gender and education. The coefficients on the English proficiency categories are not shown (they are all statistically insignificant at the .001 level, results available on request). The inclusion of self-reported English-language proficiency has no effect on the key independent variables: again, we find no effect of categories A, B and C compared with D, and a significant effect of D2 (1.8% higher wage growth per year).

Model 4 adds an additional linguistic niche category, E, which comprises workers who started in categories A to C but experienced an *increase* of occupational % Spanish of 1 percent or more. The reason for pulling these workers out is that it provides an even stronger test of wage trajectories for workers in category D, as now we are comparing them to workers in categories A, B and C who experienced no increase in their % Spanish over the course of the study. The results for Model 4 again indicate there is no evidence that workers in A, B or C have higher rates of wage growth than workers who started in category D.

In contrast to the effect on wage growth, the levels panel of Table 6 demonstrates important differences in the Wave 1 wages among the different linguistic niche categories. Immigrant Latinos in occupations with more than 30 percent Spanish speakers (category C) earn about 21 percent less than immigrant workers in occupations with less than 10 percent Spanish speakers, and about 11 percent less than workers in occupations with 11 percent to 20 percent Spanish speakers.

In Table 7 we test for the robustness of our results for alternative ways to code the linguistic niche categories. Each row of Table 7 presents results on the key slope variables for a different cut point for category D, ranging from 20 percent to 50 percent, as shown in column 1. For all of the models, we continue to code category A as workers who start in occupations with less than 10 percent Spanish speakers, but category B combines categories B and C (including all workers from 11% to the cut point for category D). Category D2 is coded as before, including all workers who start in D but experience a decline of their % Spanish of 10 percent or more. Overall, all of the different cutpoints arrive at the same conclusion as Table 6: workers in category D do not have lower levels of wage growth. Workers in category D2 have higher rates of wage growth, although the smaller number of cases in D2 at cut points of 45 percent and 50 percent reduces the significance of this coefficient to the $p = .01$ level.

Overall, the results in Tables 6 and 7 are consistent with a sorting model of linguistic niches and wages. Although the occupational linguistic niche categories have an important effect on the levels of wages, they do not have much of an effect on wage growth. Only workers who move out of the linguistic niches experience higher than average wage growth, after controlling for individual labor market characteristics that affect the baseline level of wages. One possible objection for

Table 7. Robustness Check, Wage Growth Models With Different Categorization of Linguistic Niche Categories

	Coefficient on Slopes (Wage Growth Over Time)				
Alternative Definition of Linguistic Niche Based on Wave 1 Occupational Spanish (Category D)	Coefficient on Category A (*SE*)	Coefficient on Category B[a] (*SE*)	Coefficient on D2[b] (Niche Movers (*SE*)	Number of Wave 1 Cases in Group D	Number of Wave 1 Cases in D2
>20%	.0019 (.0057)	.00070 (.0058)	.020 (.0061)***	2,378	421
>25%	.0021 (.0058)	.0021 (.0044)	.019 (.0062)***	2,092	398
>30%	.0011 (.00586)	-.00002 (.0043)	.018 (.0064)***	1,852	376
>35%	.00017 (.0060)	-.0017 (.0043)	.019 (.0069)***	1,550	331
>40%	-.000098 (.0062)	.0011 (.0044)	.017 (.0074)**	1,286	285
>45%	.0004 (.0063)	.00056 (.0044)	.013 (.0078)*	1,115	257
>50%	.00080 (.0063)	.0012 (.0045)	.014 (.0080)*	1,066	246

*** p < .01 ** p < .05 * p < .1

Note: Standard errors in parentheses.

[a]For the purposes of this table, category B is defined as between 10% Spanish and the lower limit of category D (i.e., column 1).

[b]Category D2 is defined as starting in category D and having a decline of 10% or more in occupational Spanish.

separating the respondents who start in niche occupations into "movers" (group E) and "stayers" (group D) is that it may be selective; the workers who are able to switch occupational categories are not a random sample of the workers who start the panel working in niche occupations. However, that is precisely our argument as described above in the theoretical model: occupational linguistic niches may provide a temporary "safe haven" to work alongside co-ethnics or other Spanish speakers while adjusting to the U.S. labor market, and workers who move out of niche occupations are likely to be positively selected on the basis of unobserved characteristics such as ambition and ability (α_i in Equations 1 and 2 above). For these upwardly mobile workers, linguistic niche occupations provide a "stepping stone" to the U.S. labor market even if they do not pay high wages. Over time, a substantial minority of workers may move out of these niche occupations and to mainstream jobs with less co-ethnic concentration.

Overall, the results presented in Tables 2 and 6 point to an important divergence in results. An analysis of the effect of occupational niches based on the levels panel of Table 6 suggests that working in occupations with a large number of poor English speakers reduces wages, even after controlling for a large number of individual-level variables. In contrast, Table 2 indicates that there is substantial mobility out of niche occupations over time, and the slopes panel of Table 6 shows that working in an occupational linguistic niche does not result in lower levels of wage growth, even among workers who stay in niche occupations (category D). Finally, those workers who start out in niche occupations but subsequently switch categories (category D2) have higher rates of wage growth than workers in all other categories.

At the same time, of course, these niche occupations continue to pay low wages to workers who remain in them. Although Table 6 shows that the level of wage growth is not any higher in other occupational categories, after controlling for individual-level human capital variables, the typical worker in a linguistic niche occupation will not catch up to the wage level of workers in other categories unless he or she moves out of the niche occupation. While 20 percent of our initial sample of workers in linguistic niches moves out of these occupations over the course of the SIPP panel, our argument about occupational sorting suggests that the transition rate will decline over time for this SIPP cohort as the most ambitious and upwardly mobile workers move to non-niche occupations first, leaving behind a set of workers less likely to make the transition.

Discussion and Conclusions

This study uses longitudinal data from the 1996, 2001 and 2004 panels of the SIPP to analyze the effect of working in occupations with large numbers of limited English speakers on the wages and wage growth of Latino immigrant workers. To measure Spanish language use at the occupational level, we aggregated data from the 2000 Census and the 2001 to 2006 waves of the ACS to the occupational and state level and then merged this onto our individual level SIPP data. As noted above, the aggregate state-level measure of the proportion of Spanish is highly correlated (above the .9 level) with the proportion of limited English speakers and Latinos in the respondent's occupation. Overall, our findings point to a crucial distinction between wage levels and wage growth: although the proportion of workers in the respondent's occupation who speak Spanish was negatively associated with wages in the cross-section, it had no effect on wage growth, based on our analysis of longitudinal data from the SIPP. In addition, 20 percent of workers who start in niche occupations move out over the course of the 4-year panel, and these "movers" experience substantially higher rates of wage growth than other workers in the sample.

The results in this paper have important implications for understanding the process of economic assimilation for Latino immigrants. While recent studies on Latino occupational segregation have argued that the crowding of immigrant Latinos workers into brown-collar occupations and segregated jobs reduces their wages (Catanzarite 2000; Catanzarite and Aguilera 2002; Kmec 2003),

our results indicate that this literature paints an overly pessimistic picture of the effect of working with co-ethnics because it relies upon cross-sectional data and, as a result, misses upwardly mobile workers who move to other occupations over time. Instead, we offer an alternative explanation based on the ethnic enclave hypothesis, which stresses the "sheltering" effect of working in an ethnically based economy for immigrant workers (e.g., Wilson and Portes 1980, Evans 2004, Bailey and Waldinger 1991). We recast brown-collar occupations as linguistic niches – occupations with substantial numbers of Spanish speakers and/or workers with limited English ability – and argue that these occupations provide immigrants with employment opportunities while they adjust to new labor market conditions, learn English, and acquire U.S.-based human capital.

In contrast to the ethnic enclave hypothesis, which suggests that the benefits of working in the enclave economy should continue to accrue to immigrant workers, we maintain that the benefit of these occupational niches is temporary; for upwardly mobile immigrants, they are stepping stones to better jobs, a means to an end rather than an end themselves. Moreover, for immigrant workers who don't move on, the opposite is true: the relatively low average pay of these occupations indicates that they are not desirable jobs to end up with.

In interpreting these results it is important to point out that we are not claiming that occupation-level discrimination and crowding do not have potentially important implications at a structural level for race and ethnic inequality. Instead, our goal is to point out the complexity of the process of economic mobility for immigrant Latinos, many of whom arrive in the United States with limited English ability and only partial knowledge of the U.S. labor market. Indeed, it is quite possible that the two effects could coexist simultaneously: a structural effect could constrain immigrant economic assimilation at the aggregate level even as a steady flow of upwardly mobile immigrants make a successful transition from brown-collar occupations to the mainstream labor market.

In conclusion, we would like to stress the importance of dynamic models of immigrant economic mobility that can capture both the heterogeneity of wage trajectories experienced by different workers as well as larger structural factors, such as those identified by the brown-collar occupations literature, which may be operating at the same time. The persistently large wage gap between white and Latino workers, combined with evidence of lower overall levels of wage growth among Latino immigrants (Lubotsky 2007), means that it would be naïve to expect the upward mobility of successful Latino immigrants to eliminate ethnic stratification between Latinos and other groups in the U.S. labor market.

At the same time, however, we argue that advising recent immigrants to avoid working with other co-ethnics and/or in niche occupations – e.g., advice based upon a literal interpretation of the brown-collar occupations literature – would be misleading, particularly if immigrants are constrained by a lack of fluency in English and limited U.S. labor market skills. Ideally, future research on this topic would attempt to go beyond our categorical classification of niche occupations to identify occupation-specific effects on wage growth, under the assumption that some immigrant occupations may promote mobility while others may impede it, as well as attempt to extend these SIPP panels by linking them to Social Security data (i.e.,

similar to the approach adopted in Lubotsky 2007) to see if the results we have documented here with relatively short SIPP panels hold up over longer periods of time.

Notes

1. In 1965, the U.S. government ended the national-origins quotas and instituted a new immigration system based primarily on family reunification. Countries from the Eastern Hemisphere were capped at 20,000 visas annually. Quotas on Western Hemispheric countries were not instituted until the Immigration Act of 1976. Although immigration quotas were placed on the Eastern and Western Hemispheres, close family relatives who entered through family reunification where not counted. This, along with continued undocumented migration, diversified cities and labor markets throughout the U.S. (see Alba and Nee 2003:174–84).
2. In addition to the age and nativity restrictions, 83 individuals who moved to a different state during the study period were dropped from the data on the suggestion of an anonymous reviewer, out of concern that their change in occupational % Spanish would reflect their migration to a different state rather than occupational mobility per se.
3. Although the SIPP includes retrospective monthly employment information since the previous wave, our sample comprises one observation per wave using the respondent's current occupation and wage at the time of the interview. Our sample includes all eligible workers with valid wage and occupation data. In addition, 403 cases with wages that are below $3 per hour or greater than $200 per hour in 2005 dollars are recoded as missing.
4. We use five imputations of the English ability variable for the 1996 panel. Excluding this variable and using only nonimputed data do not affect the substantive conclusions of the models.
5. Further information about the O*NET database, including links to downloads, is provided at http://www.onetcenter.org/database.html
6. The sorting into these categories is strongly associated with Wave 1 English ability for the 2001 and 2004 samples: 64.93 percent of workers who report not speaking English are in category D (>30% Spanish speaking) versus 28.92 percent who speak English "very well" and 42.23 percent who speak English "well."
7. Respondents who are not working in Wave 1 but subsequently enter our sample are classified according the linguistic niche category of the first observation that they are in the workforce.
8. Note that we will use "occupation % Spanish" to refer to the proportion of Spanish-speaking workers in the occupation.
9. The summary statistics on English ability in Table 2 are for the nonimputed data.

Table A1. Occupations With High Concentrations of Spanish Speakers, 2005 and 2006 American Community Survey

	Occupation	Occ. Code[a]	Average Wage[b]	Proportion Limited English[c]	Proportion Spanish Speaking[d]	Number of Cases[e]
1	Plasterers and stucco masons	646	13.79	.332	.503	753
2	Graders and sorters, agricultural products	604	10.70	.347	.453	600
3	Drywall installers, ceiling tile installers, and tapers	633	14.79	.261	.419	3282
4	Miscellaneous agricultural workers	605	8.74	.298	.390	14692
5	Cement masons, concrete finishers, and terrazzo workers	625	15.29	.213	.354	1519
6	Roofers	651	13.28	.235	.351	3479
7	Helpers, construction trades	660	11.32	.219	.351	1483
8	Packers and packagers, hand	964	9.60	.263	.349	6365
9	Pressers, textile, garment and related materials	831	8.52	.303	.344	1012
10	Miscellaneous media and communications workers	286	17.19	.013	.331	1242
11	Maids and housekeeping cleaners	423	8.92	.226	.318	21488
12	Packaging and filling machine operators and tenders	880	10.72	.244	.317	4644
13	Grounds maintenance workers	425	10.58	.194	.303	18937
14	Carpet, floor and tile	624	13.80	.16	.294	3777
15	Painters, construction and maintenance	642	12.82	.177	.288	9765
16	Construction laborers	626	13.30	.181	.286	26716
17	Butchers and other meat, poultry, and fish processing workers	781	11.68	.204	.277	4119
18	Sewing machine operators	832	9.05	.303	.274	4740
19	Brickmasons, blockmasons, and stonemasons	622	17.07	.188	.270	3602
20	Textile cutting machine setters, operators, and tenders	840	10.06	.185	.265	313

21	Insulation workers	640	14.73	.126	.264	728
22	Helpers–production workers	895	10.56	.181	.254	840
23	Shoe machine operators	834	10.84	.212	.253	99
24	First-line supervisors/ managers	600	15.09	.101	.251	1216
25	Helpers: installation, maintenance, and repair workers	761	10.70	.150	.249	393
26	Fence erectors	671	12.18	.115	.243	523
27	Dishwashers	414	7.12	.205	.242	4319
28	Reinforcing iron and rebar workers	650	19.03	.108	.241	158
29	Cleaners of vehicles and equipment	961	9.98	.142	.238	5324
30	Laundry and dry cleaning workers	830	9.03	.186	.237	3368
31	Food cooking machine operators	785	11.22	.170	.235	200

[a]3-digit code, 2000 census occupations
[b]Current Population Survey data, all workers in occupation.
[c]Proportion of workers in occupation who report speaking English poorly or not at all.
[d]Proportion of workers who report speaking Spanish at home.
[e]Number of Hispanic workers in the combined 2005-2006 ACS.
Source: 2005 and 2006 ACS.

References

Aguilera, Michael B. 2009. "Ethnic Enclaves and the Earnings of Self-Employed Latinos." *Small Business Economics* 33:413-25.

Alba, Richard, and Victor Nee. 2003. *Remaking the American Mainstream: Assimilation and Contemporary Immigration*. Cambridge, MA: Harvard University Press.

Bailey, Thomas, and Roger Waldinger. 1991. "Primary, Secondary, and Enclave Labor Markets: A Training Systems Approach." *Annual Review of Sociology* 56:432-45.

Bean, Frank D., Mark Leach, and Lindsay B. Lowell. 2004. "Immigrant Job Quality and Mobility in the United States." *Work and Occupations* 31:499-518.

Borjas, George. 1982. "The Earnings of Male Hispanic Immigrants In the United States." *Industrial and Labor Relations Review* 35(3):343-53.

_____. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." Journal of Labor Economics 3(4):463-89.

_____. 1995. "The Economic Benefits from Immigration." *Journal of Economic Perspectives* 9:3-22.

Borjas, George J., and Lawrence Katz. 2005. "The Evolution of the Mexican- Born Workforce in the United States." Cambridge, MA: National Bureau of Economic Research. Working Paper 11281. Available at: http://www.nber.org/papers/w11281

Catanzarite, Lisa. 2000. "Brown-Collar Jobs: Occupational Segregation and Earnings of Recent-Immigrant Latinos." *Sociological Perspectives* 43(1):45-75.

_____. 2002. "Dynamics of Segregation and Earnings in Brown-collar Occupations." *Work and Occupations* 29(3):300-45.

_____. 2003. "Race-Gender Composition and Occupational Pay Degradation." *Social Problems* 50(1):14-37.

Catanzarite, Lisa, and Michael B. Aguilera. 2002. "Working with Co-Ethnics: Earnings Penalities for Latino Immigrants at Latino Job Sites." *Social Problems* 49 (1):101-27.

Chiswick, Barry R., and Paul W. Miller. 2005. "Do Enclaves Matter in Immigrant Adjustment?" *City and Community* 4:5-35.

_____. 2010. "Occupational Language Requirements and the Value of English in the U.S. Labor Market." *Journal of Population Economics* 23:353-72.

Damm, Anna Piil. 2009. "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence." *Journal of Labor Economics* 27(2):281-314.

Dustmann, Christian, and Francesca Fabbri. 2003. "Language Proficiency and Labour Market Performance of Immigrants in the UK." *The Economic Journal* 113 (489):695-717.

Edin, Pers-Anders, Peter Fredriksson, and Olof Åslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment." *Quarterly Journal of Economics* 118:329-57.

Evans, M.D.R. 2004. "Do Ethnic Enclaves Benefit or Harm Linguistically Isolated Employees?" *Research in Social Stratification and Mobility* 22:281-318.

Gilbertson, Greta A., and Douglas T. Gurak. 1993. "Broadening the Enclave Debate: The Labor Market Experiences of Dominican and Columbian Men in New York City." *Sociological Forum* 8:205-20.

Hall, Matthew and George Farkas. 2008. "Does Human Capital Raise Earnings for Immigrants in the Low Skill Labor Market?" *Demography* 45(3):619-39.

Hellerstein, Judith, and David Neumark, 2002 "Ethnicity, Language, and Workplace Segregation: Evidence from a New Matched Employer-Employee Data Set," NBER Working Papers 9037, National Bureau of Economic Research, Inc.

Howell, David R., and Elizabeth J. Mueller. 2000. "Immigration and Native-Born Male Earnings: A Jobs-Level Analysis of the New York City Metropolitan Area Labour Market, 1980-90." *Journal of Ethnic and Migration Studies* 26(3):469-93.

Kmec, Julie. 2003. "Minority Job Concentration and Wages." *Social Problems*, 51(1):38-59.

Kossoudji, Sherrie A. 1988. "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men." *Journal of Labor Economics* 6:205-28.

Lubotsky, Darren. 2007. "Chutes or Ladders? A Longitudnal Analysis of Immigrant Earnings. *Journal of Political Economy* 115(5):820-67.

McManus, Walter S. 1985. "Labor Market Costs of Language Disparity: An Interpretation of Hispanic Earnings Differences." *The American Economic Review* 75:818-27.

McManus, Walter S., William Gould, and Finis Welch. 1983. "Earnings of Hispanic Men: The Role of English Language Proficiency." *Journal of Labor Economics* 1:101-30.

Morales, Maria Cristina. 2009. "Ethnic-Controlled Economy or Segregation? Exploring Inequality in Latina/o Co-ethnic Jobsites." *Sociological Forum* 24:589-610.

National Research Council. 2009. Reengineering the Survey of Income and Program Participation. Washington, DC: National Academy Press.

Portes, Alejandro, and Robert L. Bach. 1985. *Latin Journey: Cuban and Mexican Immigrants in the United States*. Berkeley, CA: University of California Press.

Portes, Alejandro, and Rubén G. Rumbaut. 2006. *Immigrant America: A Portrait*. (3rd ed.). Berkeley, CA: University of California Press.

Rabe-Hesketh, Sophia. and Anders Skrondal. 2005. Multilevel *and Longitudinal Modeling Using Stata*. Stata Press.

Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. *Stata Journal* 9(3):466-77.

Sanders, Jimy M., and Victor Nee. 1987. "Limits of Ethnic Solidarity in the Enclave Economy." *American Sociological Review* 52:745-73.

Smith, James P. 2003. "Assimilation Across the Latino Generations." American Economic Review, 93(2):315-19.

_____. 2006. "Immigrants and the Labor Market." *Journal of Labor Economics* 24(2):203-33

Trejo, Stephen. 1997. "Why do Mexican Americans Earn Low Wages?" *Journal of Political Economy* 105(6):1235-68.

Waldinger, Roger, and Michael I. Lichter. 2003. *How the Other Half Works: Immigration and the Social Organization of Labor*. Berkeley and Los Angeles, CA: University of California Press.

Warman, Casey. 2007. "Ethnic Enclaves and Immigrant Earnings Growth." *Canadian Journal of Economics* 40(2):410-22.

Zhou, Min, and John R. Logan. 1989. "Returns on Human Capital in Ethnic Enclaves: New York City's Chinatown." *American Sociological Review* 54:809-20.