Ethnicity, Neighborhoods, and Human-Capital Externalities

By George J. Borjas*

The socioeconomic performance of today's workers depends not only on parental skills, but also on the average skills of the ethnic group in the parents' generation (or ethnic capital). This paper investigates the link between the ethnic externality and ethnic neighborhoods. The evidence indicates that residential segregation and the external effect of ethnicity are linked, partly because ethnic capital summarizes the socioeconomic background of the neighborhood where the children were raised. Ethnicity has an external effect, even among persons who grow up in the same neighborhood, when children are exposed frequently to persons who share the same ethnic background. (*JEL* J24, J62)

Ethnic neighborhoods have long been a dominant feature of American cities (and of cities in many other countries). In fact, segregation by race and ethnicity often defines the invisible line that creates a neighborhood. These neighborhoods insulate people of similar backgrounds and foster a set of cultural attitudes, social contacts, and economic opportunities that affect workers throughout their lives.

In earlier work (Borjas, 1992, 1994), I have argued that ethnicity has an external effect on the human-capital accumulation process.[1] Persons raised in advantageous ethnic environments will be exposed to social and economic factors that increase their productivity, and the larger or more frequent the amount of this exposure, the higher the resulting "quality" of the worker.

As with the models that dominate the new growth literature, sufficiently strong ethnic externalities may delay the convergence of ethnic differentials indefinitely. My earlier empirical work indicated that the earnings of children are affected strongly not only by parental earnings as in the usual models of intergenerational income mobility, but also by the mean earnings of the ethnic group in the parents' generation (which I called "ethnic capital"). As a result, the ethnic spillover effect retards intergenerational improvement for relatively disadvantaged ethnic groups and slows down the deterioration of skills (i.e., the regression toward the mean) among the more advantaged groups.

The process through which the ethnic externalities are transmitted, however, is not well understood. This paper investigates one possible mechanism, ethnic neighborhoods. The insight that human-capital externalities and geography are linked is not new. In his pathbreaking work, Robert E. Lucas (1988) cites the crowding of similarly skilled workers into a small number of city blocks as a key determinant of the economic development of cities. Similarly, William Julius Wilson's (1987) influential work on the creation and growth of the underclass argues

*Department of Economics 0508, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, and National Bureau of Economic Research. I am grateful to Julian Betts, Thomas MaCurdy, James Rauch, Glenn Sueyoshi, Stephen Trejo, and Finis Welch for helpful comments, and to the National Science Foundation and the Russell Sage Foundation for financial support.

[1]The importance of human-capital externalities in intergenerational mobility was stressed in the early work of John Conlisk (1977) and Glenn C. Loury (1977), who uses the concept of "social capital" to analyze how racial discrimination influences the social mobility of blacks. Shelly Lundberg and Richard Startz (1992) investigate how human-capital externalities may alter the impact of antidiscrimination programs on social mobility.

that blacks who live in poor neighborhoods are not exposed to "mainstream" role models, thus hampering the economic mobility of blacks.

This paper presents an empirical study of the link between geography and ethnic externalities. The analysis uses the 1/100 Neighborhood File of the 1970 Public Use Sample of the U.S. Census and a specially designed version of the National Longitudinal Surveys of Youth (NLSY). The Census data group workers into one of over 40,000 neighborhoods, while the NLSY file groups workers into one of 1,978 zip codes. Hence it is possible to determine the extent to which ethnic groups segregate in particular neighborhoods and the impact of this segregation on the process of human-capital accumulation and intergenerational mobility.[2]

The main finding of the analysis is that residential segregation and the influence of ethnic capital on the process of intergenerational mobility are intimately linked. In particular, the impact of ethnic capital on the skills of the next generation arises partly because the ethnic-capital variable is an excellent proxy for the socioeconomic background of the neighborhood where the children were raised, and these neighborhood characteristics influence intergenerational mobility. In other words, the ethnic-capital model provides an alternative way of capturing neighborhood effects. Ethnic capital, however, plays an additional role in intergenerational mobility. Even among persons who grow up in the same neighborhood, ethnic capital matters when children are exposed frequently to other persons who share the same ethnic background.

[2]The role played by neighborhood effects in determining socioeconomic outcomes is currently the subject of intensive research; see, for instance, the survey of Christopher Jencks and Susan E. Meyer (1990) and the critical appraisal by Charles F. Manski (1993). Empirical evidence linking neighborhood effects to teenage pregnancy, criminal behavior, educational attainment, and human-capital accumulation is given by Anne C. Case and Lawrence F. Katz (1991), Jonathan Crane (1991), Mary Corcoran et al. (1992), and James E. Rauch (1993).

I. Ethnicity and Neighborhoods

Because little is known about the residential clustering of many of the ethnic groups used in the empirical analysis below, it is useful first to document the link between ethnicity and residential segregation.[3] The descriptive analysis is based initially on data drawn from the 1/100 Neighborhood File of the 1970 U.S. Census (15-percent questionnaire). These data not only contain the individual-level demographic variables typically available in Census files, but also group individuals into one of 42,950 "neighborhoods." Neighborhoods are contiguous and relatively compact (roughly the size of a Census tract), and they have an average population of 4,000 persons (U.S. Bureau of the Census, 1973). Although the specific geographic location of a neighborhood cannot be determined (other than its location in one of the nine Census regions), the data file reports a number of demographic characteristics describing the neighborhood (such as the fraction of persons who are either first- or second-generation Americans, and the fraction of persons who are college graduates).

I restrict the analysis to persons aged 18–64. I begin by documenting the residential segregation of immigrants and second-generation Americans, and the extent to which residential segregation changes across generations. A person is an immigrant if he or she was born outside the United States (or its possessions); a person is a second-generation American if either parent was born outside the United States. All other persons are grouped and labeled "third-generation" Americans, although this sample obviously includes higher-order generations. The 1970 Census does not provide any information on the ethnic ancestry of persons in the "third" generation.

[3]A large literature documents the extent of residential segregation among blacks and Hispanics; see Frank D. Bean and Marta Tienda (1987), Nancy A. Denton and Douglas S. Massey (1989), Scott McKinney and Ann B. Schnare (1989), and Mark Alan Hughes and Janice Fanning Madden (1991).

As noted above, the neighborhood file reports the proportion of the population in each neighborhood that is either first- or second-generation. This statistic was calculated by the Bureau of the Census using all available observations in the neighborhood (i.e., the 15-percent sample of respondents who filled out the relevant questionnaire). I use these data to estimate the fraction of persons in the neighborhood who are either first- or second-generation for the average person in a number of demographic groups.

Table 1 summarizes the extent of residential segregation. The first row reports that the average immigrant resided in a neighborhood where 32.7 percent of the population was either first- or second-generation. This pattern of residential location differs significantly from what one would expect if immigrants were randomly allocated across neighborhoods. The 1970 Census indicates that only 16.6 percent of the population was first- or second-generation.

Because the aggregate characteristics reported in the neighborhood file do not include the proportion of the neighborhood's population that is foreign-born, I calculate this statistic by combining the birthplace data reported in each individual's record with the aggregate neighborhood characteristics provided by the Census Bureau.[4] The typical immigrant lives in a neighborhood that is 15.3-percent immigrant, even though only 4.8 percent of the population was foreign-born.

Residential segregation persists into the second generation. As the second row of Table 1 shows, the average second-generation American resides in a neighborhood that is 28.2-percent first- or second-generation.

The 1970 Census does not provide any information on ancestry past the second generation. As a result, I cannot determine how the pattern of residential segregation changes beyond the second generation for most groups. Intergenerational changes in residential segregation, however, can be documented for the subpopulation of Hispanics, the vast majority of whom are foreign-born or have parents or grandparents who are foreign-born.[5] Table 1 indicates that there is very little movement of Hispanics out of Hispanic neighborhoods even in the third generation. The average Hispanic immigrant lives in a neighborhood that is 35-percent Hispanic; the average second-generation Hispanic lives in one that is 33-percent Hispanic; and the typical third-generation Hispanic lives in one that is 29-percent Hispanic. The fraction of Hispanics in the population is only 4.4 percent. The clustering of Hispanics into Hispanic neighborhoods, therefore, is prevalent and persistent.[6]

In addition to the clustering of first- and second-generation persons into certain neighborhoods, there is substantial segregation by ethnic group. To document the differences across national-origin groups, I focus on the 39 largest groups in the data. These 39 groups include 83.7 percent of all first-generation Americans, and over 95 percent of all second-generation Americans. The national origin of immigrants is, of course, determined by their country of birth. The national origin of a second-generation person is determined by the father's birthplace (unless only the mother was foreign-born, in which case it is determined by the mother's birthplace). Table 2 lists the 39 national-origin groups used in the analysis.

[4]In particular, I take the Census Bureau estimate of the proportion of persons in the neighborhood who are first- or second-generation to be the population proportion. I then multiply this number by the sample estimate of the proportion of the first- and second-generation individuals in the neighborhood who are foreign-born.

[5]Calculations from the General Social Surveys indicate that over 90 percent of persons who classify themselves as Hispanic are foreign-born, have parents who are foreign-born, or have grandparents who are foreign-born.

[6]Although the residential segregation found among these ethnic groups is substantial, it is not nearly as striking as that found among blacks. Table 1 reports that the average black lives in a neighborhood that is 54.7 percent black.

TABLE 1—RESIDENTIAL SEGREGATION IN THE 1970 CENSUS

Neighborhood characteristics of average person in:	Percentage of population in neighborhood that is:				
	First generation	First or second generation	Black	Hispanic	Sample size
First generation	15.3	32.7	6.9	10.2	63,099
Second generation	6.7	28.2	4.3	5.2	156,134
Third generation	3.8	13.8	11.7	3.9	905,213
Hispanics:					
First generation	22.2	36.7	6.5	35.0	10,713
Second generation	9.4	27.3	5.1	33.0	10,801
Third generation	8.9	21.9	11.4	28.8	25,202
Third generation:					
Blacks	3.1	8.0	54.7	3.7	109,533
Whites	3.7	14.4	5.6	3.1	771,359

Notes: The "white" sample includes all nonblack, non-Hispanic third-generation workers. The population proportions are as follows: immigrants, 4.8 percent; first or second generation, 16.6 percent; blacks, 11.1 percent; and Hispanics, 4.4 percent.

I first calculated the proportion of the population who are either first- or second-generation and who have a particular ethnic ancestry. This number is reported in the first column of the table and represents the probability that a first- or second-generation person from that group will be found in a particular neighborhood *if* the ethnic group is distributed randomly across neighborhoods. Most of the groups make up relatively small fractions of the population: only 0.8 percent of the population, for instance, is first- or second-generation Irish.

Table 2 reveals that immigrants and their children, regardless of national origin, cluster in neighborhoods that have large numbers of first- or second-generation Americans. The typical second-generation person of English ancestry resides in a neighborhood that is 23.5-percent first- or second-generation; the respective statistic for Irish persons is 31.3 percent, for Italians 32.0 percent, and for Mexicans 27.8 percent. There is little evidence, therefore, that only economically disadvantaged groups are crowded into ethnic neighborhoods.

To document how type-j ethnics cluster in specific neighborhoods, I calculate the fraction of the neighborhood's population that has the same ethnicity as the average type-j person. The Census Bureau does not report the fraction of the population in each neighborhood that belongs to each of the groups. Hence I calculated this statistic from within the 1/100 sample. Because the family members of a type-j ethnic are likely to be type-j ethnics, and because the 1/100 Census File is a random sample of households, the stratified sampling scheme introduces an upward bias in the calculation of the fraction of the neighborhood's population that is type j. I choose a conservative index of within-group residential segregation and calculate (for each person in the data) the proportion of persons in the neighborhood who reside outside the household unit and who are type-j ethnics.[7] Table 2 reports the average of this statistic for each of the groups. In view of the relatively small sample size available for each neighborhood (the mean and median number of observations in a neighborhood is 26, and the interquartile range is 9, from 21 to 30), some caution is required in the interpretation of the data.

The probability that type-j ethnics live near other type-j ethnics is much higher

[7]This methodology does not entirely solve the problem, because extended-family members are also likely to be type-j ethnics and to live in the same neighborhood (but as part of a different household unit).

TABLE 2—RESIDENTIAL SEGREGATION IN THE 1970 CENSUS, BY NATIONAL-ORIGIN GROUP

		First generation			Second generation		
		Percentage of population in neighborhood that is:			Percentage of population in neighborhood that is:		
National origin	Percentage of population in group	First or second generation	Same ethnicity	Sample size	First or second generation	Same ethnicity	Sample size
Austria	0.6	34.5	2.0	883	30.1	2.1	6,007
Azores	0.04	37.1	8.0	184	30.5	3.5	320
Belgium	0.07	28.9	0.4	250	21.8	0.7	573
British West Indies	0.03	24.8	0.8	175	24.0	0.8	188
Canada	1.8	25.7	6.2	6,843	24.8	7.4	13,085
Cuba	0.3	48.7	21.3	3,119	27.6	4.7	270
China	0.2	38.5	9.2	1,617	33.5	6.2	635
Czechoslovakia	0.5	34.6	2.3	797	25.6	2.9	4,571
Denmark	0.2	24.9	0.5	289	20.2	0.9	1,608
England	0.8	24.3	1.5	3,113	23.5	1.5	6,367
Finland	0.1	29.1	1.5	194	25.5	3.9	1,200
France	0.2	28.7	0.4	811	23.8	0.3	1,184
Germany	1.7	27.2	2.9	5,930	21.9	3.2	13,089
Greece	0.3	38.3	2.6	1,147	28.3	1.1	1,913
Hungary	0.4	34.3	2.6	1,020	28.0	1.9	3,472
Ireland	0.8	36.2	4.6	1,434	31.3	3.3	7,137
Italy	2.8	37.7	15.3	5,193	32.0	12.1	26,476
Jamaica	0.06	28.4	2.2	507	22.3	1.5	163
Japan	0.2	26.1	3.2	1,020	33.7	12.6	1,716
Latvia	0.04	27.0	0.2	245	33.1	0.1	260
Lebanon	0.05	27.0	0.3	118	23.7	0.4	476
Lithuania	0.2	36.2	3.7	325	30.6	1.5	2,128
Mexico	1.3	35.4	22.6	5,746	27.8	18.1	8,412
Netherlands	0.2	23.9	1.8	689	21.5	3.9	1,725
Northern Ireland	0.1	29.4	0.3	233	28.1	0.2	573
Norway	0.3	28.5	1.8	422	22.1	3.0	3,203
Other West Indies	0.04	28.8	2.5	254	25.5	1.3	250
Philippines	0.2	31.0	5.9	1,477	30.1	6.5	606
Poland	1.7	40.2	9.1	2,846	32.0	7.8	15,182
Portugal	0.1	40.9	11.2	654	32.7	6.8	1,030
Romania	0.1	38.6	0.8	373	34.5	0.7	1,150
Scotland	0.3	27.5	0.7	1,013	24.4	0.7	2,517
Sweden	0.4	29.1	1.4	445	22.3	1.7	4,284
Switzerland	0.1	26.7	0.6	315	20.3	0.8	947
Syria	0.04	30.9	1.7	103	27.9	0.8	387
Turkey	0.06	36.6	0.2	251	33.0	0.3	459
USSR	1.2	38.8	7.0	1,738	34.9	7.8	12,067
Wales	0.1	23.7	0.1	99	21.3	0.3	529
Yugoslavia	0.3	31.9	2.7	930	25.1	2.4	2,309
Sample of 39 countries	—	32.9	8.3	52,802	28.3	6.6	148,468

Note: The residential-segregation measures give the percentage of the population in the neighborhood that belongs to the specified ethnic group for the average person in the sample.

than one would expect if type-j ethnics were randomly distributed across neighborhoods. Among second-generation workers, the typical person of Irish ancestry lives in a neighborhood that is 3.3-percent Irish, although first- and second-generation Irish make up only 0.8 percent of the population; the typical Italian lives in a neighborhood that is 12.1-percent Italian, although Italians make up only 2.8 percent of the population; and the typical Mexican lives in a neighborhood that is 18.1-percent Mexican, although Mexicans make up only 1.3 percent of the population. Among the 39 national-origin groups,

TABLE 3—RESIDENTIAL SEGREGATION IN THE 1970 CENSUS, BY DEMOGRAPHIC CHARACTERISTICS

Neighborhood characteristics of average person	Percentage of population in neighborhood that is: First or second generation	Black	Hispanic	Same ethnicity	Sample size
First generation:					
Age:					
18–34	31.7	7.9	11.9	8.7	21,532
35–64	33.2	6.3	9.3	8.2	41,567
Education:					
Less than 12 years	35.5	7.3	13.3	11.7	30,590
12 years	30.8	6.6	7.8	5.8	17,000
13–15 years	29.5	6.1	7.4	4.6	7,959
16 or more years	29.1	6.4	5.7	3.4	7,550
Year moved to house:					
Before 1960	32.5	6.3	7.4	7.7	13,623
1960–1966	33.9	6.3	10.0	8.8	18,690
1967–1970	32.1	7.4	11.6	8.3	30,786
Second generation:					
Age:					
18–34	27.4	4.9	7.6	6.6	31,824
35–64	28.5	4.1	4.6	6.5	124,310
Education:					
Less than 12 years	28.7	4.8	6.7	8.4	61,896
12 years	28.1	3.9	4.4	6.0	56,725
13–15 years	27.4	3.9	4.6	4.7	19,311
16 or more years	27.8	3.8	3.3	4.2	18,212
Year moved to house:					
Before 1960	29.4	4.4	4.3	7.4	65,585
1960–1966	28.5	3.9	5.4	6.5	45,926
1967–1970	26.3	4.3	6.2	5.4	44,623
Third generation:					
Age:					
18–34	14.5	11.4	4.1	—	425,477
35–64	13.2	11.9	3.7	—	479,736
Education:					
Less than 12 years	11.5	16.5	4.3	—	346,392
12 years	14.5	9.4	3.6	—	334,888
13–15 years	15.9	8.1	3.9	—	129,884
16 or more years	17.0	6.7	3.3	—	94,049
Year moved to house:					
Before 1960	13.9	11.4	3.3	—	242,945
1960–1966	13.6	12.3	3.8	—	255,798
1967–1970	13.9	11.4	4.3	—	406,470

the typical immigrant lives in a neighborhood in which 8.3 percent of the population shares the same ethnic background, and the typical second-generation person lives in a neighborhood in which 6.6 percent of the population shares the same background.

I conclude the descriptive analysis of the Census data by documenting that ethnic residential segregation exists across a number of demographic and skill groups. Table 3 shows that there is little difference in ethnic residential segregation across age groups. The typical second-generation person aged 18–34 resides in a neighborhood that is 27.4-percent first- or second-generation, while the respective statistic for an older person is 28.5 percent. In addition, the differences in residential segregation across education groups are often small. The typical high-school dropout in the second generation lives in a neighborhood that is 28.7-percent first- or second-generation, while the respective statistic for the typical college graduate is 27.8 percent. Finally, the data indicate that internal migration decisions among first- and second-generation Americans do not seem to alter the ethnic composition of their residential environments. Second-generation persons who have lived in the same house for over 10 years live in a neighborhood that is 29.4-percent first- or second-generation, while the respective statistic for persons who have lived in the house fewer than three years is 26.3 percent.

The NLSY reveals even stronger patterns of residential segregation. The analysis uses a version of the NLSY that identifies the subset of persons who resided in the same zip code in 1979, at the time the survey of young persons (aged 14–22) began. Hence it is possible to determine whether NLSY respondents live near other NLSY respondents who share the same ethnic background.[8]

Ethnicity is determined from the response to the question: "What is your origin or descent?" Although most persons in the NLSY gave only one response to the question, about one-third of the respondents gave multiple answers. In these cases, I used the main ethnic background (as identified by the respondent) to classify people into ethnic categories.

For each person in the data, I calculated the probability that other NLSY respondents in the zip code had the same ethnic background. The NLSY, however, surveyed other persons in the family unit who were in the "correct" age range (i.e., 14–22 in 1979). As a result, there are large numbers of siblings in the data: 27 percent of the respondents have one sibling, and an additional 19 percent have at least two siblings in the data. To avoid the bias introduced by this sampling scheme, I calculated the residential-segregation measures on the sample of nonrelated persons who reside outside the household unit.[9] Moreover, because the NLSY oversampled blacks and other minorities, I used the sampling weights in the calculations.

The segregation indexes are reported in Table 4 for the 25 ethnic groups identifiable in the NLSY.[10] There is strong evidence of residential segregation. The average black lived in a neighborhood that was 63.4 percent black, while the average Mexican lived in a neighborhood that was 50.3 percent Mexican. Overall, the typical NLSY respondent lived in a neighborhood where 30.4 percent of other nonrelated respondents shared a common ethnic background.[11]

Note that this statistic is much larger than the respective statistic in the Census data, where only 7–8 percent of a neighborhood's population belonged to the same group. The Census results, however, underestimate the extent of residential segregation because all third-generation workers are classified as nonethnics (because no information is provided on the ethnic background of third-generation persons). As a result, even though the typical immigrant in the Census lives in a neighborhood where 8.3 percent of the population is composed of first- or second-generation persons who belong to the same ethnic group, a much larger frac-

[8] The numbering system used to identify zip codes in the NLSY file differs from that used by the Postal Service. Although the data indicate subsets of NLSY respondents who live in the same postal area, it is impossible to locate the zip code within a particular metropolitan area. Because the zip code refers to the 1979 residence, many of the respondents were still living in the parental household. As a result, the residential-segregation measures in the NLSY tend to reflect the ethnic environment in which the respondents were raised.

[9] To reduce costs, the NLSY also sampled households which resided geographically close to each other. This sampling strategy suggests that the measures of residential segregation calculated in these data probably overstate the true extent of segregation.

[10] Of the 12,686 observations in the 1979 wave of the NLSY, I deleted two persons because they had invalid zip codes, and 939 persons because they had invalid ethnic classifications.

[11] As with the Census data, the NLSY residential-segregation measures should be interpreted with caution. There are fewer than 100 observations for 11 of the 25 ethnic groups.

TABLE 4—RESIDENTIAL SEGREGATION IN THE NLSY, BY NATIONAL-ORIGIN GROUP

Ethnicity	Percentage of population in group	Percentage of population in neighborhood with same ethnic background	Sample size
American	7.6	18.2	743
American Indian	5.9	12.9	624
Asian Indian	0.2	2.0	22
Black	14.9	63.4	3,055
Chinese	0.2	3.5	26
Cuban	0.4	33.3	117
English	18.9	23.9	1,587
Filipino	0.4	5.0	44
French	3.5	5.6	316
German	17.4	25.7	1,420
Greek	0.4	7.2	31
Hawaiian	0.1	0.2	20
Irish	11.0	14.3	956
Italian	6.2	16.3	498
Japanese	0.2	0.0	20
Korean	0.1	0.0	6
Mexican	4.1	50.3	1,174
Other Hispanic	0.9	9.3	214
Polish	3.1	12.8	242
Portuguese	0.6	19.7	97
Puerto Rican	1.2	29.8	328
Russian	0.6	0.3	47
Scottish	1.5	4.6	122
Vietnamese	0.0	0.0	1
Welsh	0.5	1.0	35
All	—	30.4	11,745

tion of the neighborhood's population might be composed of third-generation workers who also belong to the same ethnic group. The NLSY avoids this problem because all persons in the data (regardless of generation) report their ancestry.

II. Econometric Framework

My objective is to determine the relationship between ethnic externalities and neighborhood effects in the intergenerational transmission process. The econometric model underlying the analysis is given by

$$y_{ij} = \beta_1 x_{ij} + \beta_2 \bar{x}_j + \varepsilon_{ij} \tag{1}$$

where y_{ij} measures the skills (such as educational attainment or the log wage) of person i in ethnic group j; x_{ij} gives the skills of his father; and $\bar{x}_j$ gives the average skills of the ethnic group in the father's generation (which I call ethnic capital). Note that $\bar{x}_j$ takes on the same value for all persons in group j. All variables are measured in deviations from the mean.

Equation (1) can be derived from a model in which utility-maximizing parents invest in their children, and in which ethnicity has an external effect on the production of children's skills (Borjas, 1992). As a result of the ethnic spillover, the human capital of children depends not only on parental inputs (as measured by the exogenous human capital of the parents), but also on the external effect of ethnicity, as summarized by the average skills of the ethnic group.

The spillover effects underlying the ethnic-capital model have much in common with the human-capital externalities that are at the heart of the recent literature on economic growth (Paul M. Romer, 1986; Lucas 1988), as well as with the notions of social capital and neighborhood effects that are stressed routinely in the sociology literature (James S. Coleman, 1988, 1990; Wilson, 1987). If the ethnic externality is sufficiently strong, skill differentials observed among ethnic groups can persist for many generations and may never disappear. Note that the expected skills of the son of the average father in ethnic group j are given by

$$E(y_{ij}) = (\beta_1 + \beta_2)\bar{x}_j. \tag{2}$$

The sum $\beta_1 + \beta_2$, therefore, determines whether the mean skills of ethnic groups converge across generations; hence $\beta_1 + \beta_2$ is an inverse measure of the rate of "mean convergence."[12] If the sum of coefficients is less than 1, ethnic differences converge over time; if it is greater than 1, ethnic differences diverge across generations.

As I have shown above, ethnic groups cluster in particular neighborhoods. This clustering suggests that part of the ethnic-capital effect in equation (1) may be capturing the influence (if there is one) of the neighborhood's socioeconomic background on intergenerational mobility. Suppose, for example, that ethnic groups are completely segregated so that there is one ethnic group per neighborhood. The ethnic-capital variable $\bar{x}_j$ would then also represent the mean skills of the neighborhood, and the coefficient β_2 in (1) would capture the total impact of the ethnic spillover and of the neighborhood's socioeconomic background. The coefficient of ethnic capital would be significant even if ethnicity did not have a direct impact on intergenerational mobility, but neighborhood characteristics mattered.

The data do not exhibit this extreme type of segregation. Ethnic groups, however, are likely to cluster by skill level, so that unskilled ethnic groups live together in low-income neighborhoods and skilled ethnic groups live in high-income neighborhoods. The ethnic-capital variable would again be correlated with the skill level of the neighborhood, and the ethnic-capital coefficient could be capturing neighborhood effects (i.e., the impact of the neighborhood's overall socioeconomic background), rather than the direct effect of ethnicity. In effect, the ethnic-capital model "works" because ethnic capital proxies for the relevant neighborhood characteristics that influence the intergenerational transmission process. If ethnicity did not have a direct impact on intergenerational mobility, controlling for the relevant neighborhood characteristics (such as mean income and education) would drive the ethnic-capital coefficient down to zero.

Ethnic capital might still matter, *above and beyond neighborhood effects*, if intragroup contacts within a neighborhood are more frequent or are more influential than intergroup contacts.[13] Children who belong to ethnic group j are then exposed to a different set of values, social contacts, and economic opportunities than children who belong to other ethnic groups but who grow up in the same neighborhood. In effect, the aggregate socioeconomic characteristics of the neighborhood are not a sufficient statistic summarizing the environment facing type-j persons. As a result, ethnic capital influences the intergenerational-mobility process even after controlling for neighborhood effects. Ethnicity per se has an impact on intergenerational mobility.

The empirical work presented in this paper decomposes the impact of the ethnic-capital coefficient into neighborhood effects (the extent to which the ethnic-capital variable proxies for neighborhood characteris-

[12] Robert J. Barro and Xavier Sala-i-Martin (1992) provide a discussion of alternative concepts of convergence in the context of growth models.

[13] Richard D. Alba's (1990) study of social contacts among U.S.-born white ethnics indicates that half of all nonrelated childhood friends belong to the same ethnic group. Harry J. Holzer (1988) has shown that friends are a key source of information about job opportunities, so that intragroup referrals play a major role in the job-search process and might explain the concentration of some ethnic groups in narrowly defined occupations.

tics that influence all persons who reside in the same neighborhood, regardless of ethnic background) and into an ethnic effect. A simple way of determining the extent to which the impact of ethnic capital (i.e., the coefficient β_2) operates through neighborhood effects is to expand the model in (1) to include a vector of neighborhood fixed effects:

$$(3) \quad y_{ij} = \delta_1 x_{ij} + \delta_2 \bar{x}_j + \sum_k \theta_k D_{ij}^k + \varepsilon'_{ij}$$

where D_{ij}^k is a dummy variable set to unity if person i in ethnic group j resides in neighborhood k. The parameter vector $(\theta_1, \ldots, \theta_K)$ gives the neighborhood fixed effects, which are assumed to be exogenous.[14] The coefficients δ_1 and δ_2 measure the within-neighborhood impact of parental skills and of ethnic capital. As long as neighborhoods matter in the transmission of skills, the "net" rate of mean-convergence (i.e., net of neighborhood effects) implied by the fixed-effects model, $\delta_1 + \delta_2$, is conceptually different from the "gross" rate implied by equation (1), $\beta_1 + \beta_2$.

Equations (1) and (3) can be estimated directly in the NLSY data discussed above. It is unusual, however, to come across data that contain all the requisite information: ethnicity, the skills of two generations of workers, and neighborhood of residence. Nevertheless, a relatively complete analysis of the relationship between ethnic capital and neighborhood effects can be conducted even if the data do not provide any information on parental background (as is the case with the 1970 Census neighborhood file). In particular, suppose *mean* parental skills in the group, $\bar{x}_j$, are observed even if parental skills are not (the source of the data on $\bar{x}_j$ will be discussed below). The individual-level data available for second-generation workers in the 1970 Census can then be used to estimate the following regression models:

$$(4) \quad y_{ij} = \beta \bar{x}_j + \omega_{ij}$$

$$(5) \quad y_{ij} = \delta \bar{x}_j + \sum_k \theta_k D_{ij}^k + \omega'_{ij}.$$

Because equations (4) and (5) regress individual-level data on an aggregate measure of ethnic skills, I call this type of model a "semi-aggregate" regression. It is easy to show that the following proposition holds.

PROPOSITION 1:

$$E(\hat{\beta}) = \beta_1 + \beta_2.$$

Data on parental skills, therefore, are not required to estimate the gross rate of mean convergence. Because the mean skills of the ethnic group instrument for parental skills, the omitted-variable bias introduced by leaving out parental skills is simply the "recovery" of the coefficient β_1. It would now be useful to determine whether $E(\hat{\delta}) = \delta_1 + \delta_2$, so that the net rate of mean convergence can also be estimated without information on parental skills. I proceed to show that this is indeed the case in an important special case and that the difference between the gross and net rates of mean convergence is attributable solely to the change in the ethnic-capital coefficient.

Consider first how the coefficients of parental and ethnic capital in equation (1) change when neighborhood fixed effects are introduced into the model. The probability limits of the estimated coefficients in (1) when the true model is given by (3) are given by equations (6) and (7), at the top of the following page, where $\Delta = \mathrm{Var}(x_{ij})\mathrm{Var}(\bar{x}_j) - \mathrm{Var}(\bar{x}_j)^2$; p_k is the fraction of the population that lives in neighborhood k; $E(x_{ij}|k)$ is the mean value of skills among parents who live in neighborhood k, where the expectation is evaluated over all i and j; and $E(\bar{x}_j|k)$ is the mean value of the ethnic-capital variable among all persons who live in neighborhood k,

[14] It would be interesting to analyze how parents choose the type and intensity of "ethnicity" that they wish to expose to their children. William N. Evans et al. (1992) show that endogenizing the "peer group" effects greatly weakens the relationship between outcomes and neighborhood characteristics.

$$\text{(6)} \qquad \operatorname{plim} \hat{\beta}_1 = \delta_1 + \frac{\operatorname{Var}(x_{ij})}{\Delta} \left\{ \sum_{\ell > k}\sum p_\ell p_k (\theta_\ell - \theta_k) \Big(\big[E(x_{ij}|\ell) - E(\bar{x}_j|\ell) \big] - \big[E(x_{ij}|k) - E(\bar{x}_j|k) \big] \Big) \right\}$$

$$\text{(7)} \qquad \operatorname{plim} \hat{\beta}_2 = \delta_2 + \frac{\operatorname{Var}(x_{ij})}{\Delta} \left\{ \sum_{\ell > k}\sum p_\ell p_k (\theta_\ell - \theta_k) \big[E(\bar{x}_j|\ell) - E(\bar{x}_j|k) \big] \right\} - \frac{\operatorname{Var}(\bar{x}_j)}{\Delta} \left\{ \sum_{\ell > k}\sum p_\ell p_k (\theta_\ell - \theta_k) \big[E(x_{ij}|\ell) - E(x_{ij}|k) \big] \right\}$$

where the expectation is again evaluated over all i and j.

In general, the introduction of neighborhood effects affects the coefficients of both parental skills and ethnic capital. Suppose, however, that type-τ ethnics residing in region k are a random sample of the population of type-τ ethnics, so that the skill distribution of type-τ ethnics in region k is the same as their skill distribution in the population. This assumption implies

$$\text{(8)} \qquad E(x_{i\tau}|k) = \bar{x}_\tau$$

where the expectation in the left-hand side is taken over all i in group τ in neighborhood k, while the right-hand side simply gives the level of ethnic capital for group τ. I refer to (8) as the "skill-invariance" assumption. Equation (8) implies $E(x_{ij}|k) = E(\bar{x}_j|k)$, $\forall k$, so that the bracketed term in (6) vanishes. It is useful to summarize this result as the following proposition.

PROPOSITION 2: *If the distribution of type-j ethnics across neighborhoods is skill-invariant,* $\operatorname{plim} \hat{\beta}_1 = \delta_1$, *so that the coefficient of parental skills is unaffected by the introduction of neighborhood fixed effects.*

Note that the skill-invariant geographic assignment of type-j workers is distinct from and weaker than assuming that type-j ethnics are distributed randomly across neighborhoods.

The skill-invariance assumption is also useful in determining the relationship between the estimator $\hat{\delta}$ [from equation (5)] and the net rate of mean convergence. This relationship is summarized by the next proposition.

PROPOSITION 3: *If the distribution of type-j ethnics across neighborhoods is skill-invariant, then* $\operatorname{plim} \hat{\delta} = \delta_1 + \delta_2$.

As before, it is unnecessary to have information on parental skills in order to estimate the rate of mean convergence (net of neighborhood effects).

The results can now be used to determine why the two rates of mean convergence estimable in Census data might differ. Because the coefficient of parental skills is unaffected by the introduction of neighborhood fixed effects, the difference between the "gross" and "net" rates of mean convergence is attributable *entirely* to the change in the coefficient of the ethnic-capital variable (assuming the skill-invariance assumption holds). Therefore, the inclusion of neighborhood fixed effects into semi-aggregate regressions can be used to assess the relationship between ethnic capital and geography. I summarize this result in the following proposition.

PROPOSITION 4: *Suppose the distribution of type-j ethnics is skill-invariant. The difference in the estimated rates of mean convergence* $\hat{\beta}$ *and* $\hat{\delta}$ *gives the impact of neighborhood effects on the ethnic-capital coefficient.*

Because of the practical importance of these results, it is worth stressing that the skill-invariance assumption is unlikely to hold strictly in the data. The analysis of the Census data presented below uses two alternative measures of skills (educational attainment and log wages) to estimate the rate of mean convergence. Even if there were no skill differentials among type-j workers residing in different neighborhoods, the restriction in (8) would be violated if there exist neighborhood wage differentials that are independent of skills. These differentials imply that the mean wage of type-j parents in a particular neighborhood differs from the measure of ethnic capital for group j. Therefore, the analysis must control for regional wage differentials prior to applying the results presented above. The construction of an index of regional wage differentials at the neighborhood level is discussed below.

A more difficult problem with the skill-invariance assumption is simply that the skill distribution of type-j ethnics probably does differ across neighborhoods.[15] I will show below, however, that the restriction implied by skill invariance is not grossly inconsistent with the geographic sorting of type-j ethnics.

Finally, the discussion has assumed that the ethnic-capital effect is constant across neighborhoods and persons. This need not be the case. In fact, the ethnic-capital model implies that the spillover effects of ethnicity should be larger for persons who are more frequently exposed to an ethnic environment. Put differently, the ethnic-capital effect should be larger for those children who grow up in neighborhoods where many of the residents share the same ethnic background. The empirical analysis presented below investigates the extent to which the ethnic-capital effect depends on the ethnic composition of the neighborhood.

III. Results

I initially use the sample of second-generation workers in the 1970 Census file (U.S. Bureau of the Census, 1973) to analyze the relationship between ethnic externalities and neighborhood effects. I restrict the analysis to second-generation men aged 18–64, who worked in the civilian sector in the year prior to the Census, who were not enrolled in school, and who were not self-employed. As before, the ethnic group of the second-generation worker is defined in terms of the father's birthplace (unless only the mother was foreign-born, in which case it is defined in terms of the mother's birthplace). I use two alternative measures of the worker's skills: educational attainment and log wage rates.

Because Census data do not directly link the skills of second-generation Americans with the skills of their immigrant parents, I use the 1/100 Public Use Sample of the 1940 Census to estimate the mean skills of the national-origin group in the parent's generation. It is likely that (adult) second-generation persons enumerated in the 1970 Census are the children of the immigrants who arrived in the period prior to 1940.[16]

[15]It is easy to determine how the coefficients of parental skills and ethnic capital change when neighborhood effects are introduced into the model and the skill invariance assumption does not hold. Suppose that highly skilled type-j workers move into wealthy neighborhoods, and unskilled type-j workers move into poor neighborhoods. This implies that $E(x_{ij}|\ell) > E(\bar{x}_j|\ell)$ in wealthy neighborhoods (ℓ) and that $E(x_{ij}|k) < E(\bar{x}_j|k)$ in low-income neighborhoods (k). It follows from equation (6) that $E(\hat{\beta}_1) = \delta_1 + \varphi$, where $\varphi > 0$. Thus the nonrandom sorting of skilled workers into "good" neighborhoods magnifies the impact of the parental contribution to the children's skills. As a result, the inclusion of neighborhood effects will reduce the coefficient of parental skills in the intergenerational-transmission equation. It is also easy to show that this type of nonrandom sorting leads to a smaller ethnic-capital coefficient in models that omit the neighborhood fixed effects.

[16]Borjas (1993) discusses the methodology of intercensal comparisons that underlie the empirical analysis using the Census data. The intercensal linkage between parents and children can be improved by focusing on workers in specific age groups. For example, the children of immigrants aged 25–44 in 1940 are likely to be relatively young in 1970. I experimented with a number of alternative age breakdowns and obtained similar results.

TABLE 5—SKILLS OF IMMIGRANT AND SECOND-GENERATION WORKERS

	Immigrants in 1940 Census			Second generation in 1970 Census		
Country of origin	Educational attainment	Log wage	Sample size	Educational attainment	Log wage	Sample size
Austria	6.7	−0.349	1,210	11.9	1.550	2,134
Azores	5.0	−0.672	63	9.6	1.232	104
Belgium	7.8	−0.483	138	11.4	1.475	197
British West Indies	8.1	−0.810	58	12.2	1.368	68
Canada	9.2	−0.427	2,741	12.0	1.431	4,720
China	6.3	−1.176	139	13.6	1.447	206
Cuba	8.6	−0.655	42	12.0	1.372	82
Czechoslovakia	6.8	−0.345	817	11.3	1.453	1,749
Denmark	9.2	−0.392	327	12.0	1.405	553
England	9.5	−0.313	1,656	12.5	1.508	2,255
Finland	6.5	−0.539	244	11.3	1.457	390
France	9.0	−0.430	248	12.3	1.450	381
Germany	8.8	−0.467	2,943	11.7	1.463	4,558
Greece	6.9	−0.737	518	12.7	1.484	694
Hungary	7.1	−0.378	809	11.7	1.509	1,298
Ireland	8.3	−0.445	1,326	12.3	1.508	2,645
Italy	5.4	−0.475	4,784	11.2	1.454	10,148
Japan	9.5	−0.849	141	12.5	1.476	662
Lithuania	4.5	−0.479	451	12.0	1.511	766
Mexico	4.4	−1.120	1,192	9.2	1.133	2,959
Netherlands	8.8	−0.557	292	11.7	1.487	623
Northern Ireland	8.3	−0.401	280	12.8	1.533	200
Norway	8.6	−0.441	606	12.0	1.457	987
Other West Indies	8.3	−0.821	119	11.9	1.353	87
Philippines	7.8	−1.009	233	11.9	1.268	188
Poland	5.4	−0.407	2,610	11.3	1.492	5,769
Portugal	4.7	−0.577	212	10.2	1.357	383
Romania	7.4	−0.339	300	13.2	1.647	428
Scotland	9.6	−0.326	862	12.4	1.511	901
Sweden	8.6	−0.378	1,038	12.3	1.503	1,534
Switzerland	9.5	−0.461	242	12.0	1.488	329
Syria	6.7	−0.547	105	12.5	1.576	131
Turkey	7.2	−0.523	211	13.7	1.644	144
USSR	7.0	−0.363	2,418	13.1	1.654	4,313
Wales	9.4	−0.426	100	12.4	1.441	189
Yugoslavia	5.4	−0.340	512	11.7	1.499	928

Table 5 reports the average educational attainment and log wages for the 36 ethnic groups that can be identified in both the 1940 and 1970 Censuses with sufficiently large numbers of observations. These 36 ethnic groups make up 97.4 percent of working immigrant men in 1940, and 95.5 percent of the second-generation working men in 1970. There is substantial dispersion in skills and wages across national-origin groups, and there is a strong positive correlation between the skills of the immigrant group in 1940 and the skills of the corresponding second-generation group in 1970.

To calculate the variable measuring mean skills in the parent's generation (i.e., the empirical measure of ethnic capital), I pool the sample of immigrant and native men in the 1940 Census (for a total of 231,606 observations) and estimate the following regression model:

$$x_{ij} = \mathbf{Z}_{ij}\boldsymbol{\alpha} + \sum_j \gamma_j G_{ij} + \varepsilon_{ij} \tag{9}$$

where x_{ij} gives the skills of person i in national-origin group j; $\mathbf{Z}_{ij}$ is a vector of socioeconomic characteristics including age, age squared, and region of residence; and

TABLE 6—ESTIMATES OF INTERGENERATIONAL CORRELATION IN 1970 CENSUS

	Regressions using neighborhood file				Regressions using county group file		
Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Education:							
Mean of group in 1940	0.3649	—	0.1707	0.2670	0.3628	—	0.3316
	(0.0828)		(0.0457)	(0.0557)	(0.0833)		(0.0709)
Includes neighborhood fixed effects	no	—	yes	no	—	—	—
Includes county fixed effects	—	—	—	—	no	—	yes
Includes neighborhood characteristics	no	—	no	yes	no	—	no
Log wage:							
Mean of group in 1940	0.4549	0.3974	0.2191	0.2474	0.4607	0.3710	0.3938
	(0.0781)	(0.0662)	(0.0578)	(0.0362)	(0.0874)	(0.0694)	(0.0772)
Includes skill-adjusted wage level	no	yes	no	yes	no	yes	no
Includes neighborhood fixed effects	no	no	yes	no	—	—	—
Includes county fixed effects	—	—	—	—	no	no	yes
Includes neighborhood characteristics	no	no	no	yes	no	no	no
Log wage, adjusted for education:							
Mean of group in 1940	0.2038	0.1767	0.1101	0.1020	0.2132	0.1589	0.1701
	(0.0400)	(0.0321)	(0.0413)	(0.0193)	(0.0511)	(0.0352)	(0.0440)
Includes skill-adjusted wage level	no	yes	no	yes	no	yes	no
Includes neighborhood fixed effects	no	no	yes	no	—	—	—
Includes county fixed effects	—	—	—	—	no	no	yes
Includes neighborhood characteristics	no	no	no	yes	no	no	no

Notes: Standard errors are reported in parentheses; the sample size is 53,703. All regressions include a second-order polynomial in the worker's age. The neighborhood characteristics included in column (iv) are the fraction of persons in the neighborhood with at least 12 years of schooling, the fraction with at least 16 years of schooling, the labor-force participation rates of men and women, the unemployment rate, the fraction of persons working in professional occupations, the fraction of families below the poverty line, and the fraction of families that earn at least \$15,000 annually. The regressions use a random-effects estimator.

G_{ij} is a dummy variable set to unity if person i belongs to group j (natives are the omitted group). The regression is estimated separately using educational attainment and the log wage rate as dependent variables. The parameter vector $(\gamma_1,\dots,\gamma_J)$ gives the empirical measure of ethnic capital for the J groups.

Table 6 reports the estimated rates of mean convergence. Equations (4) and (5) give the basic specification of the model, except that the regressions also control for the second-generation worker's age and age squared. The regressions use a random-effects estimator which allows for an ethnic-group-specific component in the error term.[17] Consider initially the middle panel reporting the transmission coefficients obtained in the log-wage regression model. Column (i) indicates that the rate of mean convergence (or $\beta_1+\beta_2$ in terms of the model in the previous section) is 0.45, in line with the results of earlier work (Borjas, 1992, 1993).

The next column controls for the bias introduced by regional wage differentials. As noted above, the skill-invariance assumption is violated if some ethnic groups have relatively high wage levels simply because they live in high-wage areas. To control for regional wage variation, I estimated the following regression in the sample of *third*-generation workers in the 1970 neighborhood file:

$$w_i = \mathbf{X}_i\boldsymbol{\alpha} + \sum_k \psi_k D_i^k + \varepsilon_i \qquad (10)$$

where w_i gives the log wage of person i; $\mathbf{X}$ is a vector of standardizing variables (including educational attainment, age, age squared, marital status, and dummy variables indicating whether the person is black

[17] In particular, the residual $\varepsilon_{ij}=\nu_j+u_{ij}$, where ν_j is the group component. It is well known that ignoring the group component in the error term seriously underestimates the standard error of the ethnic-capital coefficient.

or Hispanic); and D_i^k is a dummy variable indicating whether person i resides in neighborhood k. The vector $(\psi_1,\dots,\psi_K)$ gives the skill-adjusted neighborhood wage level. This wage level is included as an additional regressor in the intergenerational earnings equations, and the resulting transmission coefficient is reported in column (ii) of Table 6.[18] The transmission coefficient falls to 0.40 (a drop of about 0.05 units).

Column (iii) of the table adds the vector of neighborhood fixed effects into the regression.[19] Controlling for the neighborhood fixed effects reduces the estimated transmission parameter substantially, to about 0.2. Assuming that the skill-invariance assumption holds, the transmission coefficient changes because the estimate in column (iii) "nets out" the relationship between the ethnic externality and neighborhood effects (but leaves unchanged the impact of parental skills). It is interesting to note that the resulting coefficient of 0.2 is roughly the same as the coefficient of parental capital in my earlier work (Borjas, 1992). It seems as if neighborhood effects account for most (if not all) of the ethnic influence in the intergenerational-transmission process. Ethnic capital seems to be a very good proxy for the relevant characteristics of the neighborhood's economic and social environment which influence the intergenerational-transmission process *and* which are common to all persons living in the same neighborhood, regardless of ethnicity.[20]

The last three columns of Table 6 use the 1/100 County Group File of the 1970 Public Use Sample (15-percent questionnaire) to estimate an identical model in a sample of second-generation workers defined exactly as in columns (i)–(iii). This Census file reports the metropolitan area (instead of the neighborhood) of current residence. Persons who live outside metropolitan areas are grouped into economically similar "county groups." A total of 408 metropolitan areas and country groups are identified in the data.

Not surprisingly, the transmission coefficient reported in column (v) is almost identical to the respective statistic in column (i). To control for regional wage variation, I estimated a regression in the sample of third-generation workers similar to (10) with county-group dummies instead of neighborhood dummies. The skill-adjusted county wage level was then introduced as an additional regressor in the model. This reduced the coefficient to about 0.37, which is roughly the same as the analogous coefficient in column (ii).

The coefficient in the last column of Table 6, however, differs drastically from the respective coefficient in column (iii). Control-

[18]The coefficient of the neighborhood wage level was typically in the 0.4–0.5 range.

[19]There are 53,703 observations in the sample of second-generation working men and 23,415 neighborhoods. There are 9,522 neighborhoods with only one observation; 5,895 neighborhoods with two; 3,616 neighborhoods with three; 2,162 neighborhoods with four; and 1,161 neighborhoods with five. The remaining 5 percent of the neighborhoods have between 6 and 12 observations. Despite the fact that a sizable number of neighborhoods have only one observation, the estimated rate of mean convergence is consistent. I use a two-stage procedure to estimate the random-effects model which includes the vector of neighborhood fixed effects. The first-stage regression includes age, age squared, and a vector of ethnic fixed effects. This regression is estimated on a data set in which all variables are differenced from the respective neighborhood means. This procedure is numerically equivalent to introducing the neighborhood fixed effects. The second stage then uses a generalized least-squares estimator to estimate the relationship between the coefficients of the first-stage ethnic dummy variables and the ethnic-capital variable.

[20]The discussion assumes that the neighborhood of current residence is the same as the neighborhood where the individual was raised. Because of the misdefinition of the neighborhood, the results confound the ethnic externalities that influenced the human-capital accumulation process with externalities that arise from living in an ethnic neighborhood at the present time. This problem, however, does not seem to be very important. The results are very similar for two alternative skill variables, log wages and educational attainment (which presumably was completed at an early age). Moreover, the transmission coefficients are roughly the same regardless of how long the person has lived in his current residence. The estimated transmission parameter is 0.43 for persons who moved to the house prior to 1960, and it is 0.46 for persons who moved to the house after 1967.

ling for county fixed effects barely affects the estimated transmission coefficient; it remains at about 0.4. Put differently, the ethnic-capital variable and the vector of county fixed effects are uncorrelated. There is no evidence, therefore, that ethnic capital has anything to do with geography at the county level. At the neighborhood level, however, geography is intimately linked to the ethnic-capital effect.[21]

The top panel of Table 6 reports the transmission coefficients obtained from regressions which use the worker's educational attainment as the dependent variable. The results are virtually identical to those obtained in the log-wage regressions. Including the neighborhood fixed effects reduces the transmission coefficient from 0.36 to 0.17, while adding in the county dummies barely changes the estimated parameter (it declines to 0.33).[22]

Finally, the bottom panel of Table 6 reports the transmission coefficients obtained in a log-wage regression that also includes the educational attainment of the second-generation worker as a regressor. Although the transmission rates are much smaller (because the transmission that occurs through educational attainment is netted out), adding neighborhood fixed effects changes the estimated coefficients in exactly the same way as in the top two panels of the table.

In sum, the analysis reveals a link between ethnic capital and neighborhood effects, but it provides no information about which set of neighborhood characteristics are being proxied by the ethnic-capital variable. Column (iv) of Table 6 shows that the neighborhood fixed effects can be summarized in terms of a small number of neighborhood characteristics. The neighborhood characteristics included in the regression are: the percentage of the neighborhood's population that has at least a high-school diploma; the percentage with at least a college diploma; the labor-force participation rates of men and women; the unemployment rate; the percentage of workers employed in professional occupations; the percentage of families below the poverty level; and the percentage of families with at least \$15,000 in household income. All of these neighborhood characteristics were calculated by the Census Bureau (and are included in the Public Use Sample).

The inclusion of these aggregate neighborhood characteristics reduces the transmission coefficient from 0.36 to 0.27 in the education regressions, and from 0.4 to 0.25 in the log-wage regressions. In other words, a small vector of variables that are common to all persons living in the neighborhood, regardless of ethnic background, can explain over half of the drop in the ethnic-capital coefficient.[23]

It seems, therefore, that a large part of the impact of ethnic capital is simply disguising for neighborhood effects which have nothing to do with ethnicity. This interpretation of the results, of course, depends on the validity of the skill-invariance assumption. As shown in Section II, when the distribution of persons across neighborhoods is skill-invariant, including neighborhood effects in semi-aggregate regressions reduces the estimated rate of mean convergence solely because the ethnic-capital coefficient is "standing in" for neighborhood effects.

Parental skills are not observed in the Census data, so that it is not possible to assess directly the validity of the skill-invariance assumption. I can test, however, whether the geographic distribution of second-generation workers rejects the skill-invariance assumption. Consider the follow-

[21]A regression of education (or log wages) on a vector of county-group dummies has an R^2 of about 0.09, so that 91 percent of the variance in education and log wages is attributable to within-county variation. In contrast, only about 45 percent of the variance in these variables is attributable to within-neighborhood variation.

[22]Because education differences across neighborhoods almost entirely reflect true differences in skill levels, I did not attempt an analogous construction of a "skill-adjusted" neighborhood education level.

[23]Although I do not report or discuss the estimated coefficients of the neighborhood characteristics, it would be interesting to study how (and why) these various characteristics influence the intergenerational transmission process.

ing regression model:

$$(11) \qquad y_{ijk} = \boldsymbol{\gamma}_0 \mathbf{G} + \boldsymbol{\gamma}_1 (\mathbf{G} \times \mathbf{D}) + \varepsilon_{ijk}$$

where y_{ijk} gives the skills of second-generation worker i in group j in neighborhood k; **G** gives a vector of dummy variables indicating the worker's ethnic group; and **D** gives a vector of dummy variables indicating the worker's neighborhood. The skill-invariance assumption states that the mean skills of a worker in ethnic group j are independent of the neighborhood of residence, so that the coefficient vector $\boldsymbol{\gamma}_1$ is zero.

I calculated the analysis-of-variance decomposition implied by (11) using both the educational attainment and log wage of workers in the second generation.[24] To net out the impact of regional wage differentials on the analysis, the worker's log wage is deflated by the skill-adjusted neighborhood wage level defined earlier. Despite the very large samples used in the analysis, testing the hypothesis that the coefficient vector $\boldsymbol{\gamma}_1$ differs from zero yields F statistics that are barely above the critical value of 1; the F statistic in the educational-attainment regression was 1.21, and the F statistic in the log-wage regression was 1.18. In contrast, the F statistic testing the significance of the group effect (i.e., whether the coefficient vector $\boldsymbol{\gamma}_0$ was zero) was 17.2 in the educational-attainment regressions and 95.9 in the log-wage regressions, substantially above the critical value of 1.4.[25]

[24]The test excludes the 9,522 neighborhoods that have only one second-generation working man.

[25]A related way of assessing the importance of the skill-invariance assumption uses the concept of the intracluster correlation (Leslie Kish, 1965; William G. Cochran, 1977). This correlation is positive if the characteristics of persons within a cluster are more closely related than those of persons randomly chosen from the population. When the cluster is defined to be the ethnic group, the intracluster correlation is about 0.1 (for both education and log wages). This correlation increases to 0.2 when the cluster is defined to be type-j ethnics living in neighborhood k. Put differently, the neighborhood provides additional information about the skill distribution of persons in a particular ethnic group.

I now use the NLSY (where parental skills are observed *and* where it is unnecessary to maintain the skill-invariance assumption) to confirm that there is a very strong link between neighborhood effects and the ethnic-capital coefficient. The analysis uses the 1990 wave of the NLSY, by which time the respondents were aged 25–33 and only about 5 percent were still enrolled in school. Equations (1) and (3) give the basic specifications of the models. The regressions also control for age, gender, whether the person is a first- or second-generation American, and whether the person was enrolled in school in 1990.

As with the analysis of Census data, I use two measures of skills: educational attainment and the log wage rate. Each NLSY respondent in 1979 reported the father's education and occupation (which was coded using the 1970 Census codes). I constructed a wage for each father by matching the father's occupation code with the average log wage in the occupation, as reported by the 1970 Census.

To obtain a measure of ethnic capital, I used the 1/100 1980 U.S. Census to calculate the mean educational attainment and mean log wage for each of the ethnic groups in the parents' generation.[26] The Census data report the ancestral background of U.S.-born residents (obtained from questions resembling the self-reported ethnic background in the NLSY). To increase the probability that the average skills of the ethnic milieu corresponded to that in which the NLSY respondents were raised, I restrict the 1980 Census sample to men aged 35–64.

Table 7 reports the summary statistics of the variables used in the analysis. There are

[26]The ethnic characteristics are calculated using a 20-percent random sample of the 5/100 A File of the 1980 Public Use Sample. I also constructed comparable ethnic characteristics from within the NLSY itself. Although the findings do not depend on which measure of ethnic capital is used, I only report the regressions that use the Census measure (which are calculated over much larger samples and contain less sampling error).

TABLE 7—SKILLS OF ETHNIC GROUPS IN THE NLSY

	Educational attainment			Log wage			
National origin	NLSY respondents	NLSY fathers	Census men	NLSY respondents	NLSY fathers	Census men	Sample size
American	12.4	10.9	11.2	2.099	1.292	1.945	480
American Indian	12.1	10.2	11.2	1.977	1.285	1.904	429
Asian Indian	14.0	11.9	16.7	1.684	1.464	2.180	7
Black	12.8	10.1	11.0	1.948	1.177	1.852	1,795
Chinese	15.2	10.0	13.8	2.403	1.146	1.955	16
Cuban	13.4	11.0	11.3	2.403	1.293	1.876	69
English	13.0	11.9	12.9	2.085	1.353	2.093	1,125
Filipino	13.9	13.0	13.8	2.471	1.388	2.009	21
French	12.8	11.7	11.7	2.074	1.329	2.123	203
German	13.4	12.2	12.9	2.167	1.317	2.115	1,009
Greek	14.3	12.4	12.8	2.330	1.404	2.081	21
Hawaiian	12.7	9.5	12.1	2.470	1.282	2.006	6
Irish	13.4	12.6	12.8	2.219	1.401	2.098	651
Italian	13.4	12.3	12.6	2.345	1.375	2.141	347
Japanese	13.4	11.9	14.1	2.093	0.907	2.194	13
Korean	15.5	13.5	14.9	1.982	1.058	2.007	4
Mexican	12.3	7.4	9.0	2.015	1.114	1.808	723
Other Hispanic	13.1	10.5	11.4	2.217	1.254	1.893	102
Polish	13.4	11.8	13.0	2.242	1.389	2.164	171
Portuguese	12.0	8.8	10.5	2.159	1.267	1.984	59
Puerto Rican	11.9	7.9	9.6	2.249	1.156	1.798	170
Russian	15.0	13.6	15.3	2.666	1.486	2.324	39
Scottish	14.4	13.5	13.8	2.224	1.458	2.158	86
Welsh	14.8	14.5	13.8	1.987	1.542	2.150	23

sizable ethnic differentials in educational attainment and log wages among NLSY respondents and their parents. Table 8 reports the estimates of the ethnic-capital model. The coefficients in the first column of the top panel reveal that the educational attainment of NLSY respondents depends on both the father's education and on the mean education of the ethnic group in the parents' generation. The estimated rate of mean convergence is 0.44. The introduction of a vector of 510 county dummies in the second column reduces both of the coefficients somewhat; the parental coefficient falls from 0.24 to 0.2, and the ethnic-capital coefficient falls from 0.20 to 0.14. Column (iii) investigates the relationship between ethnic capital and neighborhoods by introducing a vector of 1,937 dummies indicating the zip code of residence.[27] The parental coefficient declines further to 0.17, and the ethnic-capital effect evaporates (the coefficient falls to 0.04). Net of neighborhood effects, therefore, the rate of mean convergence is only 0.21, about half the size of the gross rate, and the decline is mostly due to the weakening of the ethnic-capital effect. The NLSY results, therefore, strongly confirm the implications of the analysis of the Census neighborhood data.[28]

[27]Of the 1,937 zip-code fixed effects included in the educational-attainment regressions, there are 900 zip codes with one observation, 256 with two, 168 with three, and 123 with four; the remainder have five or more observations. Of the 1,453 zip-code fixed effects included in the log-wage regressions, there are 733 zip codes with one observation, 223 with two, 140 with three, and 80 with four; the remainder have five or more observations. A regression of educational attainment (or log wages) of the NLSY respondents on a vector of zip-code dummies has an R^2 value of about 0.4, so that about 60 percent of the variance in educational attainment and log wages can be attributed to within-zip-code variation. Over 80 percent of the variance in these variables, however, can be attributed to within-county variation.

[28]It is also possible to estimate Census-type semi-aggregate regressions on the NLSY data, so that the regressions omit the worker's parental background. Using educational attainment as the dependent vari-

TABLE 8—ESTIMATES OF THE ETHNIC-CAPITAL MODEL IN THE NLSY

	Regression			
Variable	(i)	(ii)	(iii)	(iv)
Education:				
Parental skills	0.2404	0.2005	0.1745	0.1784
	(0.0666)	(0.0669)	(0.0718)	(0.0849)
Ethnic capital	0.2004	0.1356	0.0376	0.1480
	(0.0465)	(0.0301)	(0.0288)	(0.0504)
Includes county fixed effects	no	yes	no	no
Includes neighborhood fixed effects	no	no	yes	no
Includes neighborhood characteristics	no	no	no	yes
Log wage:				
Parental skills	0.3774	0.2645	0.2500	0.2460
	(0.0371)	(0.0398)	(0.0418)	(0.0480)
Ethnic capital	0.3190	0.3107	0.0458	0.0229
	(0.1559)	(0.1116)	(0.1331)	(0.1636)
Includes county fixed effects	no	yes	no	no
Includes neighborhood fixed effects	no	no	yes	no
Includes neighborhood characteristics	no	no	no	yes
Log wage, adjusted for education:				
Parental skills	0.1765	0.1158	0.1214	0.1221
	(0.0369)	(0.0394)	(0.0410)	(0.0476)
Ethnic capital	0.0759	0.1581	−0.0231	−0.0584
	(0.1571)	(0.1141)	(0.1289)	(0.1621)
Includes county fixed effects	no	yes	no	no
Includes neighborhood fixed effects	no	no	yes	no
Includes neighborhood characteristics	no	no	no	yes

Notes: Standard errors are reported in parentheses. The sample size is 7,569 for the educational-attainment regressions and 4,261 for the log-wage regressions. All regressions include variables indicating the worker's age, gender, whether the person is first-generation or second-generation, and whether the person is enrolled in school in 1990. The neighborhood characteristics included in column (iv) are the average educational attainment and the average log wage of parents in the neighborhood. The regressions use a random-effects estimator.

The remaining two panels of Table 8 reestimate the model using the (log) wage and the adjusted wage. The estimate of the rate of mean convergence using the log wage is 0.70, which is higher than the one found in the Census. The introduction of county dummies reduces the rate of mean convergence to 0.57, with the ethnic-capital coefficient remaining unchanged. Finally, the introduction of neighborhood fixed effects reduces the coefficient of ethnic capital to 0.05, which is statistically insignificant. Note, however, that the coefficient of parental capital has declined by about 0.13 units, which indicates that the geographic distribution of NLSY respondents is not consistent with the skill-invariance assumption.[29]

The last column of Table 8 shows what happens to the parental and ethnic-capital coefficients when I introduce a small vector

able, the coefficient of the ethnic-capital variable (and standard error) is 0.438 (0.047) in the model that does not include either county or neighborhood dummies; 0.329 (0.030) in the model that includes county dummies; and 0.173 (0.029) in the model that includes zip-code dummies. This pattern of coefficients closely mirrors the results documented in the Census data. A similar pattern is obtained in the log-wage regressions.

[29] In particular, highly skilled type-j ethnics tend to cluster in wealthier neighborhoods, while less-skilled type-j ethnics cluster in poorer neighborhoods.

of neighborhood characteristics (rather than zip-code dummies) to control for neighborhood effects. Because the NLSY file does not contain any population estimates of economic or social characteristics in the zip code, all neighborhood-specific variables must be calculated from within the data and contain substantial sampling error. I estimated the mean education and log wage of the parents of NLSY respondents in each zip code. Controlling for these two characteristics reduces the ethnic-capital coefficient by about 0.05 units in the education regression and by almost 0.3 units in the log-wage regression. As with the Census, a small vector of neighborhood characteristics that are common to all persons living in the neighborhood helps explain why the ethnic-capital variable matters (particularly in the log-wage regressions).[30]

IV. Ethnic Capital and the Ethnic Composition of the Neighborhood

The evidence suggests that, to a large extent, the ethnic-capital effect summarizes the impact of neighborhood characteristics (common to all the residents of the neighborhood) on the intergenerational-transmission process. In view of this result, it is worth asking whether ethnicity per se plays any role in intergenerational mobility, above and beyond the influence of parents and neighborhoods.

Ethnicity is likely to play a more important role among persons who grow up in a segregated ethnic environment. After all, these persons will probably experience (and be influenced by) more frequent social, cultural, and economic intragroup contacts. The analysis in the preceding section ignored this implication of the model because it assumed that the ethnic-capital coefficient was constant across workers. To determine whether ethnicity plays an independent role among workers raised in segregated neighborhoods, I now allow the ethnic-capital coefficient to vary according to the extent of residential segregation in the neighborhood.

In particular, I interact both the ethnic-capital variable and the parental-skills variable (when available) with dummies indicating the proportion of persons in the neighborhood who share the same ethnic background. The regression model also includes the dummy variables indicating the proportion of the neighborhood's population who belong to the respondent's ethnic group (so as to allow for different constant terms). Finally, I estimate the models both with and without neighborhood fixed effects.[31]

The evidence is summarized in Table 9. Consider initially the results obtained from the 1970 Census file. Even after controlling for neighborhood effects, both the education and log-wage regressions show that the rate of mean convergence is larger among

[30] It is of interest to note that the results do not change substantially when the model is estimated on the subsample of NLSY respondents who were 14–18 years old at the time of the initial interview in 1979. The residential location decision for these young persons was probably made by their parents, so that the neighborhood fixed effects are less likely to be endogenous. In the educational-attainment regressions which do not include neighborhood fixed effects, the parental coefficient (and standard error) was 0.235 (0.009), and the ethnic-capital coefficient was 0.097 (0.026). The inclusion of neighborhood fixed effects changed the coefficients to 0.170 (0.007) and 0.017 (0.032), respectively. In the log-wage regressions which do not include neighborhood fixed effects, the parental and ethnic-capital coefficients were 0.343 (0.048) and 0.498 (0.111). Including neighborhood fixed effects changed these coefficients to 0.202 (0.043) and 0.054 (0.160).

[31] I did not interact the neighborhood effects with the dummy variables describing the proportion of persons in the neighborhood who have the same ethnic background as the worker. This restriction helps to isolate the impact of ethnic capital among persons who live in the same neighborhood (and hence who were exposed to the same overall neighborhood characteristics). I also estimated the models by simply interacting the fraction of persons in a neighborhood who have the same ethnicity with the relevant variables and obtained qualitatively similar results. Table 9 indicates, however, that there are strong nonlinearities in the relationship between the ethnic-capital coefficient and the extent of residential segregation. Moreover, there is a great deal of sampling error in the residential-segregation statistics. As a result, I prefer the specification that clusters persons into a small number of neighborhood types.

TABLE 9—ESTIMATES OF INTERGENERATIONAL CORRELATION, BY ETHNIC COMPOSITION OF NEIGHBORHOOD

	Education				Log wage			
	(i)		(ii)		(i)		(ii)	
Ethnic composition of neighborhood	Parental skills	Ethnic capital	Parental skills	Ethnic capital	Parental skills	Ethnic capital	Parental skills	Ethnic capital
A. *1970 Census*								
Percentage with same ethnicity:								
0 percent	—	0.2458	—	0.1467	—	0.2567	—	0.1322
		(0.1195)		(0.0781)		(0.1020)		(0.0447)
Between 0 percent and 15 percent	—	0.3206	—	0.2261	—	0.4702	—	0.2920
		(0.1410)		(0.0930)		(0.1320)		(0.0653)
More than 15 percent	—	0.5325	—	0.2711	—	0.6769	—	0.3782
		(0.2338)		(0.2166)		(0.1496)		(0.1091)
B. *NLSY*								
Percentage with same ethnicity:								
Less than 5 percent	0.2748	0.1482	0.2071	0.0491	0.4636	0.1850	0.3178	0.0290
	(0.0126)	(0.0791)	(0.0131)	(0.0257)	(0.0719)	(0.2085)	(0.0758)	(0.1422)
Between 5 percent and 33 percent	0.2933	0.2699	0.2014	0.0439	0.4198	0.2189	0.3292	0.0152
	(0.0116)	(0.0863)	(0.0125)	(0.0267)	(0.0654)	(0.2092)	(0.0737)	(0.1440)
More than 33 percent	0.1965	0.2998	0.1311	0.1188	0.3828	0.2958	0.2586	0.1429
	(0.0105)	(0.0848)	(0.0105)	(0.0268)	(0.0575)	(0.2094)	(0.0618)	(0.1253)
Includes neighborhood fixed effects?	no		yes		no		yes	

Notes: Standard errors are reported in parentheses. The Census regressions include a second-order polynomial in the worker's age. The NLSY regressions control for the worker's age, gender, whether the person is first- or second-generation, and whether the person is enrolled in school in 1990. The Census regressions have 53,703 observations; the NLSY education regressions have 7,569 observations, and the NLSY log-wage regressions have 4,261 observations. The regressions use a random-effects estimator.

persons who live in highly segregated neighborhoods. The education regressions, for example, indicate that the net rate of mean convergence is 0.15 for those who live in neighborhoods where none of the neighbors share the same ethnic background; 0.23 for those who live in neighborhoods where at most 15 percent of the population share the same ethnic background; and 0.27 for those who live in neighborhoods where over 15 percent of the population has the same ethnic background.[32] In the log-wage regressions, the respective statistics are 0.13, 0.29, and 0.38.

It is worth stressing that these estimates of the rate of mean convergence net out neighborhood effects. If the impact of parental skills is constant across neighborhoods, the evidence suggests that ethnicity might be playing an important role for persons who live in segregated neighborhoods, above and beyond the influence of parents and neighborhoods.

This implication is partially confirmed by the analysis of the NLSY data, where the rate of mean convergence can be decomposed into the parental and ethnic effects. The educational-attainment regressions, for instance, show that (even after controlling for neighborhood fixed effects) the ethnic-capital coefficient increases from 0.05 for

[32]The results are not sensitive to the particular definition of residential segregation. This particular breakdown, as well as the breakdown of neighborhoods in the NLSY data, was chosen because it provided a reasonable number of observations for each type of neighborhood. In the Census data, there were 27,006 persons who lived in the most integrated neighborhoods, 18,676 who lived in the "mixed" neighborhoods, and 8,021 who lived in the most segregated neighborhoods. In the NLSY education regressions, the respective numbers of observations are 1,999, 2,506, and 3,064, while in the NLSY log-wage regressions, they are 1,189, 1,428, and 1,644.

children who grew up in areas where fewer than 5 percent of the nonrelated neighbors have the same ethnic background to 0.12 for children who grew up in areas where at least 33 percent of the neighbors share the same ethnicity. Similarly, the ethnic-capital coefficient in the log-wage regressions rises from 0.03 for those who grew up in "integrated" neighborhoods to 0.14 for the children raised in the most "segregated" neighborhoods (although many of these coefficients have large standard errors).

The NLSY results suggest that not only does the ethnic-capital coefficient increase as the neighborhood becomes more segregated, but also the coefficient of parental skills decreases. The log-wage regressions, for instance, indicate that the parental coefficient (net of neighborhood effects) declines from 0.32 for persons raised in the most integrated neighborhoods to 0.26 for persons raised in the most segregated neighborhoods. The relative unimportance of parental skills for persons raised in segregated neighborhoods might indicate that group influences "take over" as the neighborhood becomes more segregated.

Because the coefficients of parental skills and ethnic capital move in different directions as persons are raised in more segregated neighborhoods, the rate of mean convergence (net of neighborhood effects) only increases slightly in the NLSY log-wage regressions, from 0.35 for persons living in integrated neighborhoods to 0.40 for persons raised in segregated neighborhoods. In the educational-attainment regressions, however, the net rate of mean convergence is roughly the same (around 0.25) across the various types of neighborhoods.

V. Ethnic Capital and Measurement Error

Many of the results presented in this paper are consistent with a different interpretation of the ethnic-capital effect. Suppose that parental skills are measured with error. The ethnic mean then provides a very good instrument for parental skills. As a result, part of the parental influence on intergenerational mobility would be captured by the coefficient of the ethnic-capital variable, even if ethnic capital did not enter the model (see Borjas [1992] for a formal derivation of the biases introduced by measurement error). The greater the noise-to-signal ratio in parental skills, the greater the ethnic-capital coefficient.

This interpretation of the results is particularly important in light of recent evidence that measurement error in parental skills imparts a sizable downward bias on the correlation between the earnings of fathers and sons (Joseph G. Altonji and Thomas A. Dunn, 1991; Gary R. Solon, 1992; David J. Zimmerman, 1992). Prior to these studies, it was generally believed that the coefficient of parental skills in an intergenerational-transmission equation was on the order of 0.2 (see, for example, the survey by Gary S. Becker and Nigel Tomes [1986]). The recent studies, which typically use panel data to average parental earnings over a number of years (and thus "wash out" the measurement error introduced by transitory changes in earnings), report much higher coefficients, on the order of 0.3–0.4.

The empirical results presented in this paper suggest transmission coefficients (as defined by the rate of mean convergence) that are typically above 0.4. In fact, the rate of mean convergence was roughly 0.5–0.7 for children raised in segregated neighborhoods. Taken at face value, therefore, the evidence suggests that ethnic capital might play an important role even if the intergenerational correlation between parents and children was as high as 0.3–0.4.

The NLSY data permit a more detailed analysis of some of the biases introduced by measurement error. As noted earlier, there are large numbers of siblings in the data, with each sibling independently reporting ethnic background, as well as the parent's education and occupation. The correlation among the siblings' responses is high, but it is far from unity. For example, the correlation between a sibling's report of the father's education and the father's average educational attainment as reported by all other siblings is 0.9; the respective statistic for the father's occupational earnings is 0.8; and nearly 30 percent of the respondents identify most with an ethnic background that differs from the "main" ancestry reported by at least one other sibling (although typically the other siblings report the alternative ancestry as a second or third ethnic background). The availability of other sources of information on parental skills and ethnic background suggests that these alternative measures of the variables can be used as instruments in the intergenerational-transmission equation. The instrumental-variables (IV) estimates of the transmission parameter can then be used to assess the practical importance of the bias introduced by measurement error in parental skills and ethnic background.[33]

I restrict the analysis to NLSY respondents who have at least one sibling in the data. For those who have only one sibling (58 percent of the sample), the instruments are given by the sibling's response. For those who have more than one sibling, the instruments are defined as the average response of all other siblings. The instruments are the average skills of the father (either educational attainment or log occupational wage) as reported by the other siblings in the data and a set of dummy variables indicating the ethnic background of the other siblings.[34] The regressions use the IV random-effects estimator proposed by Jerry A. Hausman and William E. Taylor (1981).[35]

A comparison of the IV estimates in the first row of Table 10 with the corresponding ordinary least-squares (OLS) estimates in Table 8 indicates that the coefficient of parental skills increases both in the education regression (from 0.24 to 0.28) and in the log-wage regression (from 0.38 to 0.48). The results also indicate that the IV estimates of the ethnic-capital coefficient remain sizable and significant. In particular, the coefficients are 0.18 and 0.30 in the education and log-wage regressions, respectively, only slightly below the OLS estimates reported in Table 8. It is evident, therefore, that measurement error in parental skills or in ethnic background cannot account for the results.

The remaining rows of the table interact the measures of parental skills and ethnic capital with dummy variables indicating the proportion of persons in the zip code who share the same ethnic background as the worker. The coefficients in Table 10 resemble those reported earlier, particularly in the education regressions. The coefficient of parental skills is smaller and the ethnic-capital coefficient is larger among workers who grew up in segregated neighborhoods (even after controlling for neighborhood effects). The impact of parental education, for instance, declines from 0.29 to 0.14 (in an IV model which includes neighborhood effects) for workers who live in more segregated neighborhoods, while the ethnic-capital coefficient rises from 0.01 to 0.17. In view of the small sample sizes and large standard errors, however, many of these differences are not statistically significant.[36]

[33]Orley Ashenfelter and Alan B. Krueger (1994) use this methodology to analyze the impact of measurement error in educational attainment on estimates of the rate of return to schooling. Their analysis suggests that measurement error imparts a sizable downward bias on estimates of the rate of return to schooling.

[34]I created a vector of dummy variables indicating the ethnic group reported by each sibling in the data. The instrument is formed by averaging this vector over all other siblings, so that it can be interpreted as the probability that the other siblings report a particular ethnic background.

[35]The model is estimated in two stages. In the first stage, the children's skills are regressed on the father's skills, other explanatory variables (age, gender, etc.), and a vector of dummy variables indicating the self-reported ethnic background. The first-stage model is estimated using instrumental variables. The second stage consists of a generalized least-squares regression in which the estimated coefficients of the ethnic dummy variables are regressed on the ethnic-capital variable. The regressions that control for neighborhood effects use a data set which has been differenced from the within-zip-code means in the first stage.

[36]Although the evidence is not consistent with an explanation that stresses classical measurement error in parental skills or ethnic background, there are other measurement problems which may account for some of the results. I have focused on a one-factor model in which one particular type of skills (either educational attainment or the log wage) is transmitted across generations. There is evidence that this one-factor approach does not provide a satisfactory explanation of the process of intergenerational mobility. Altonji and Dunn (1991) report that the correlation in earnings among siblings is larger than would be expected given the size of the correlation between parents and children. This result suggests that perhaps a vector of traits is being transmitted, so that the ethnic-capital variable could be proxying for an aggregate measure of these traits.

TABLE 10—INSTRUMENTAL-VARIABLE ESTIMATES OF INTERGENERATIONAL CORRELATION IN THE NLSY

	Education				Log wage			
	(i)		(ii)		(i)		(ii)	
Model	Parental skills	Ethnic capital	Parental skills	Ethnic capital	Parental skills	Ethnic capital	Parental skills	Ethnic capital
All workers	0.2781	0.1772	0.1984	0.0885	0.4776	0.3000	0.1366	0.4433
	(0.0111)	(0.0658)	(0.0129)	(0.0510)	(0.0764)	(0.2879)	(0.0978)	(0.2723)
Interactions with percentage of population that has same ethnicity:								
Less than 5 percent	0.3360	0.1230	0.2912	0.0090	0.5384	0.2516	0.1460	0.3955
	(0.0210)	(0.0675)	(0.0242)	(0.0546)	(0.1435)	(0.3010)	(0.1623)	(0.2852)
Between 5 percent and 33 percent	0.3378	0.1076	0.2387	0.0765	0.4209	0.2794	0.1439	0.5579
	(0.0202)	(0.0670)	(0.0224)	(0.0533)	(0.1379)	(0.2977)	(0.1785)	(0.2465)
Greater than 33 percent	0.1963	0.2357	0.1350	0.1677	0.4744	0.3248	0.2848	0.3436
	(0.0168)	(0.0660)	(0.0176)	(0.0532)	(0.1154)	(0.2929)	(0.1354)	(0.2805)
Includes neighborhood fixed effects?	no		yes		no		yes	

Notes: Standard errors are reported in parentheses. The Census regressions include a second-order polynomial in the worker's age. The NLSY regressions control for the worker's age, gender, whether the person is first- or second-generation, and whether the person is enrolled in school in 1990. The instruments used in the regression include the average skills of the father (either educational attainment or the log occupational wage) as reported by the other siblings in the data; and the average of a set of dummy variables indicating the ethnic background reported by the other siblings. The NLSY education regressions have 3,157 observations; the NLSY log-wage regressions have 1,978 observations. The regressions use a random-effects estimator.

VI. Summary

It is increasingly evident that ethnic skill differentials tend to persist from generation to generation. Part of the correlation arises because of the linkage between parental skills and the skills of children. Even if ethnicity did not matter, the children of skilled parents are likely to have above-average skills. This correlation, however, is not sufficiently high to account for the sluggish rate at which the mean skills of ethnic groups converge over time. To explain the slow rate of convergence, recent work borrows from the new growth literature and stresses the importance of ethnic externalities in the human-capital accumulation process. This ethnic spillover implies that the skills of ethnic children depend not only on parental skills, but also on the mean skills of the ethnic group in the parents' generation. The intergenerational transmission of this ethnic fixed effect explains why it takes a relatively long time for ethnic skill differentials to converge.

This paper investigates the nature of the ethnic externality. The study focuses on one possible channel through which the ethnic externality might operate, the ethnic neighborhood. Using the Neighborhood File of the 1970 U.S. Census and the National Longitudinal Surveys of Youth, I documented substantial residential segregation by ethnicity. Even though only 16.6 percent of the population in 1970 was first- or second-generation, the typical immigrant resided in a neighborhood that was 32.7-percent first- or second-generation, and the respective statistic for second-generation workers was 28.2 percent. In addition, there was a strong likelihood that persons belonging to a particular ethnic group reside in a neighborhood where a relatively high number of persons share the same ethnic background.

The empirical analysis indicated that the rate of mean convergence in the skills of ethnic groups was significantly reduced after controlling for neighborhood fixed effects. This finding indicates that much of the ethnic-capital effect works through the fact that low-income ethnic groups cluster in low-income neighborhoods, and these neighborhood effects influence intergenerational mobility. The analysis, however, also revealed that neighborhood effects cannot account for the entire impact of ethnicity on intergenerational mobility, particularly for persons residing in ethnically segregated neighborhoods. Ethnicity has an impact above and beyond both parental and neighborhood effects for persons who are frequently exposed to a particular ethnic environment.

There are many related issues and questions that are not addressed in this paper. For instance, what happens to the nature and impact of ethnic externalities as the groups intermarry? How do the different ethnic influences clash when disparate ethnic and racial groups cluster in the same neighborhoods? What are the policy implications of the interactions among ethnic externalities, residential segregation, and intergenerational mobility? Because of the underlying significance of these questions, the study of the links between race or ethnicity and human-capital externalities is sure to remain a fertile ground for future research

REFERENCES

Alba, Richard D. *Ethnic identity: The transformation of white America.* New Haven, CT: Yale University Press, 1990.

Altonji, Joseph G. and Dunn, Thomas A. "Relationships Among the Family Incomes and Labor Market Outcomes of Relatives," in Ronald Ehrenberg, ed., *Research in labor economics*, Vol. 12. Greenwich, CT: JAI Press, 1991, pp. 269–310.

Ashenfelter, Orley and Krueger, Alan B. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *American Economic Review*, December 1994, *84*(5), pp. 1157–73.

Barro, Robert J. and Sala-i-Martin, Xavier. "Convergence." *Journal of Political Economy*, April 1992, *100*(2), pp. 223–51.

Bean, Frank D. and Tienda, Marta. *The Hispanic population of the United States.* New York: Russell Sage Foundation, 1987.

Becker, Gary S. and Tomes, Nigel. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, July 1986, *4*(3), Part 2, pp. S1–S39.

Borjas, George J. "Ethnic Capital and Intergenerational Mobility." *Quarterly Journal of Economics*, February 1992, *107*(1), pp. 123–50.

______. "The Intergenerational Mobility of Immigrants." *Journal of Labor Economics*, January 1993, *11*(1), Part 1, pp. 113–35.

______. "Long-Run Convergence of Ethnic Skill Differentials: The Children and Grandchildren of the Great Migration." *Industrial and Labor Relations Review*, July 1994, *47*(4), pp. 553–73.

Case, Anne C. and Katz, Lawrence F. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." Unpublished manuscript, National Bureau of Economic Research, Cambridge, MA, May 1991.

Cochran, William G. *Sampling techniques.* New York: Wiley, 1977.

Coleman, James S. "Social Capital in the Creation of Human Capital." *American Journal of Sociology*, Supplement 1988, *94*, pp. S95–S120.

______. *Foundations of social theory.* Cambridge, MA: Harvard University Press, 1990.

Conlisk, John. "An Exploratory Model of the Size Distribution of Income." *Economic Inquiry*, July 1977, *15*(3), pp. 345–66.

Corcoran, Mary; Gordon, Roger; Laren, Deborah and Solon, Gary. "The Association Between Men's Economic Status and Their Family and Community Origins." *Journal of Human Resources*, Fall 1992, *27*(4), pp. 575–601.

Crane, Jonathan. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." *American Journal of Sociology*, March 1991, *96*(5), pp. 1226–59.

Denton, Nancy A. and Massey, Douglas S. "Hypersegregation in United States Metropolitan Areas—Black and Hispanic Segregation Along Five Dimensions." *Demography*, August 1989, *26*(3), pp. 373–91.

Evans, William N.; Oates, Wallace E. and Schwab, Robert M. "Measuring Peer Group Effects: A Study of Teenage Behavior." *Journal of Political Economy*, October

1992, *100*(5), pp. 966–91.

Hausman, Jerry A. and Taylor, William E. "Panel Data and Unobservable Individual Effects." *Econometrica*, November 1981, *49*(6), pp. 1377–98.

Holzer, Harry J. "Search Method Use by Unemployed Youth." *Journal of Labor Economics*, January 1988, *6*(1), pp. 1–20.

Hughes, Mark Alan and Madden, Janice Fanning. "Residential Segregation and the Economic Status of Black Workers: New Evidence for an Old Debate." *Journal of Urban Economics*, January 1991, *29*(1), pp. 28–49.

Jencks, Christopher and Meyer, Susan E. "The Social Consequences of Growing Up in a Poor Neighborhood: A Review," in Laurance E. Lynn and Michael G. H. McGeary, eds., *Inner city poverty in the United States*. Washington, DC: National Academy Press, 1990, pp. 111–86.

Kish, Leslie. *Survey sampling*. New York: Wiley, 1965.

Loury, Glenn C. "A Dynamic Theory of Racial Income Differences," in Phyllis A. Wallace and Annette A. LaMond, eds., *Women, minorities, and employment discrimination*. Lexington, MA: Lexington Books, 1977, pp. 153–86.

Lucas, Robert E. "On the Mechanics of Economic Development." *Journal of Monetary Economics*, July 1988, *22*(1), pp. 3–42.

Lundberg, Shelly and Startz, Richard. "On the Persistence of Racial Inequality." Unpublished manuscript, University of Washington, 1992.

Manski, Charles F. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, July 1993, *60*(3), pp. 531–42.

McKinney, Scott and Schnare, Ann B. "Trends in Residential Segregation by Race: 1960–1980." *Journal of Urban Economics*, November 1989, *26*(3), pp. 269–80.

Rauch, James E. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics*, November 1993, *34*(3), pp. 380–400.

Romer, Paul M. "Increasing Returns and Long-Run Growth." *Journal of Political Economy*, October 1986, *94*(5), pp. 1002–37.

Solon, Gary R. "Intergenerational Income Mobility in the United States." *American Economic Review*, June 1992, *82*(3), pp. 393–408.

U.S. Bureau of the Census. *Supplement no. 1 to public use samples of basic records from the 1970 Census: Description and technical documentation*. Washington, DC: U.S. Government Printing Office, 1973.

Wilson, William Julius. *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press, 1987.

Zimmerman, David J. "Regression Toward Mediocrity in Economic Stature." *American Economic Review*, June 1992, *82*(3), pp. 409–29.