SKIN COLOR DISCRIMINATION AND IMMIGRANT PAY

Joni Hersch[*]

INTRODUCTION

My article, *Profiling the New Immigrant Worker: The Effects of Skin Color and Height*,[1] presents strong evidence that darker skin color is associated with lower wages for new legal immigrants to the United States. Taking into account education, English language proficiency, occupation in the source country, and family background, as well as Hispanic ethnicity, race, and country of birth, I found that immigrants with the lightest skin color earn, on average, 17% higher wages than comparable immigrants with the darkest skin color. This Essay addresses to what extent this pay disparity associated with skin color is evidence of employment discrimination.

Multiple regression analysis is the standard empirical methodology used to identify whether pay disparities between groups of workers may be due to discrimination. In multiple regression analysis, discrimination is a residual inference drawn after taking into account legitimate productivity-related characteristics. Pay disparities between groups of workers that remain after taking such characteristics into account in the regression analysis are frequently attributed to discrimination. However, such unexplained disparities may instead arise from omitted productivity characteristics. For example, pay disparities on the basis of race that remain after taking into account education may be due to unobserved differences in school quality or to neighborhood effects.

In this Essay, I demonstrate that the negative effect of darker skin color on wages is not due to omitted productivity characteristics. In contrast to characteristics such as race or sex that are dichotomous, skin color varies within race, country of birth, and even families. It is thereby unlikely that an

---

[*] Professor of Law and Economics, Vanderbilt Law School.

[1] Joni Hersch, *Profiling the New Immigrant Worker: The Effects of Skin Color and Height*, 26 J. LAB. ECON. 345 (2008).

omitted productivity factor such as school quality will be correlated with skin color, after accounting for other productivity characteristics. Empirically, I demonstrate the invariance of the magnitude of the skin color effect as extensive productivity-related characteristics are sequentially taken into account in the regression analysis.

Title VII of the Civil Rights Act of 1964 prohibits discrimination on the basis of color, as well as on the basis of race, religion, sex, and national origin.[2] My results showing that darker skin color adversely affects earnings among otherwise comparable legal immigrants provide evidence of discrimination on the basis of a protected characteristic. Furthermore, my results demonstrate that color discrimination is a form of discrimination separate from, and in addition to, any discrimination based on race or national origin. Thus, my analysis contributes generally to understanding the ethnic and racial gap in pay observed in the United States.[3] In addition, my analysis contributes to understanding current widespread opposition to immigrants in the United States. Specifically, the independent effect of skin color on wages of immigrants to the United States demonstrates that one source of discrimination is based on appearance.

Part I provides an overview of the multiple regression approach used to examine whether pay gaps between groups of workers may be due to discrimination. In that Part, I discuss how the possibility of omitted variables bias affects whether unexplained pay gaps can be interpreted as caused by discrimination. To demonstrate the possible importance of omitted variables bias, I provide an example that shows how the non-Hispanic-to-Hispanic pay gap changes as additional variables are included in a regression equation. Part II explains how analysis of gradations of skin color can be used to counter problems of omitted variables bias. Drawing on my 2008 article, I present empirical evidence that darker skin color is associated with lower wages among immigrants and demonstrate that this finding is unlikely to be due to omitted variables bias. Part III concludes with a discussion of the relevance of my research to employment discrimination litigation.

---

[2] 42 U.S.C § 2000e (2000).

[3] For statistical evidence on the ethnic and racial gap in the United States, see U.S. CENSUS BUREAU, STATISTICAL ABSTRACT OF THE UNITED STATES: 2008, at tbl.627 (2008), *available at* http://www.census.gov/compendia/statab/tables/08s0627.pdf.

I. MULTIPLE REGRESSION APPROACH TO IDENTIFYING EMPLOYMENT DISCRIMINATION

Statistical evidence of discrimination in pay between groups of workers can be provided using multiple regression analysis. Multiple regression analysis takes into account the effect on pay of differences in individual worker characteristics and allows us to isolate the contribution of each characteristic.

Discrimination is a residual inference made after all relevant measurable variables have been included in the analysis.[4] A frequent defense in employment discrimination cases is that not all factors have been included and that the omitted variables account for the observed pay gap.[5] However, it is impossible in a multiple regression analysis to control for every possible factor that may affect pay. Indeed, doing so is neither necessary nor desirable. For the most part, omitted variables are included in the random error term that is part of every multiple regression model. Furthermore, some variables are properly excluded from a regression analysis as they may themselves be the outcome of the same discriminatory process under consideration.[6]

Exclusion of relevant variables from a regression equation can result in omitted variables bias under certain conditions. If an omitted variable is correlated with the variable of interest (e.g., sex or race), and if this omitted variable is an important determinant of the outcome (so that if this variable is included in the multiple regression model, its coefficient would be statistically significant), then the coefficient on the variable of interest will reflect both the direct effect of this factor as well as the indirect effect of the omitted variable. Failure to control for this variable results in a biased estimate of the effect of the variable of interest on the outcome.

It is important to recognize that omitted variables bias does not simply arise because a variable is left out of the equation. If the omitted variables are not correlated with the included variables, or if the omitted variables are correlated

---

[4] *See, e.g.*, Daniel L. Rubinfeld, *Reference Guide on Multiple Regression*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 179, 181–83 (Federal Judicial Center, 2d ed. 2000), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/sciman00.pdf/$file/sciman00.pdf.

[5] It is routine in discrimination cases for the defense to claim the estimated disparity is due to omitted variables. *See, e.g.*, Bazemore v. Friday, 478 U.S. 385, 394–95 (1986); Smith v. Va. Commonwealth Univ., 84 F.3d 672, 675 (4th Cir. 1996); Equal Employment Opportunity Comm'n v. Sears, Roebuck & Co., 628 F. Supp. 1264, 1347–48 (N.D. Ill. 1986).

[6] *See, e.g.*, Trout v. Garrett, 780 F. Supp. 1396, 1412–13 (D.D.C. 1991).

with the included variables but are not themselves statistically significant determinants of the outcome, there is no statistical problem of bias. The regression may have less explanatory power, but it will not lead to invalid inferences about the magnitude of the coefficients on included variables. In addition, if the effect of the potentially omitted variable is already largely accounted for by other variables in the equation, inclusion of this additional variable will have little effect on the coefficient on the variable of interest or on the explanatory power of the equation. In sum, one might hypothesize that the statistical results are subject to bias, but the magnitude of the bias may be small.

To demonstrate how exclusion of variables correlated with the variable of interest affects the magnitude of any estimated pay disparity between groups of workers, consider the source of pay differences between non-Hispanic and Hispanic workers. It is widely established that education is a major determinant of earnings and that those with higher education levels have considerably higher pay.[7] In addition, Hispanics have lower education levels on average than non-Hispanics.[8] If we are interested in estimating the non-Hispanic-to-Hispanic pay disparity, but do not take into account differences in education, then the estimated pay disparity will reflect not only the effect of being Hispanic rather than non-Hispanic, but also the effect of the omitted years of education. The estimated penalty to being Hispanic will be larger in wage equations that exclude education than in wage equations that include education.

Specifically, I show below how the magnitude of the pay disparity between non-Hispanic and Hispanic workers is affected by inclusion of additional variables, using data from the 2003 Current Population Survey (CPS).[9] This example will then be contrasted with a similar analysis that examines how inclusion of additional information on individuals affects the magnitude of the skin color disparity for immigrants reported in my earlier study.

---

[7] *See, e.g.*, JENNIFER CHEESEMAN DAY & ERIC C. NEWBURGER, U.S. CENSUS BUREAU, THE BIG PAYOFF: EDUCATIONAL ATTAINMENT AND SYNTHETIC ESTIMATES OF WORK-LIFE EARNINGS (2002), http://www.census.gov/prod/2002pubs/p23-210.pdf.

[8] *See* U.S. CENSUS BUREAU, STATISTICAL ABSTRACT OF THE UNITED STATES: 2008, at tbl.217 (2008), *available at* http://www.census.gov/compendia/statab/tables/08s0217.pdf.

[9] The CPS is a monthly survey of households conducted by the U.S. Bureau of Census for the Bureau of Labor Statistics. This survey is the source of the national unemployment rate reported monthly and provides other information on the labor force. *See* Bureau of Labor Statistics, U.S. Dep't of Labor, Labor Force Statistics from the Current Population Survey, http://www.bls.gov/cps/home.htm (last visited Nov. 14, 2008).

Hispanic ethnicity is reported in the CPS separately from race. For the purposes of this example, I restrict my analysis to respondents who report their race as white, who are employed but not self-employed,[10] who are between 18 and 64 years of age, and who have an hourly wage between $1.50 and $100. The number of observations is 134,530, with 119,000 observations on white non-Hispanic workers and 15,530 on white Hispanic workers.

Table 1, panel A presents the means or percentages of selected variables for non-Hispanic and Hispanic workers in the sample. The non-Hispanic and corresponding Hispanic values of all means or percentages reported in this table are significantly different from each other at the 1% level. The difference in hourly wage is considerable, with non-Hispanic workers averaging $18.10 per hour and Hispanic workers averaging $12.96 per hour. Also notable are the large differences in education, with non-Hispanic workers averaging 13.99 years of education and Hispanic workers averaging 11.33 years of education.

Table 1, panel B reports the coefficient on the non-Hispanic indicator variable as well as the adjusted R-squared (a measure of the goodness of fit of the regression) and the percent disparity in wages. In all equations, the dependent variable is the log of the hourly wage. I start with a minimal specification and then discuss the consequence on the magnitude of the non-Hispanic-to-Hispanic disparity and on the explanatory power of the equation as additional individual and productivity-related variables are included in the equation.

The results summarized in row 1 are based on a regression controlling only for the non-Hispanic indicator. The coefficient is 0.313, which corresponds to a 36.8% difference in wages between non-Hispanic workers and Hispanic workers.[11] Note that the adjusted R-squared is 0.03, which indicates that only 3% of the variation in wages is explained by whether a worker is of Hispanic ethnicity.[12]

---

[10] Earnings information is not reported for self-employed workers in the monthly CPS.

[11] Since the dependent variable is the log of hourly wage, the percentages are calculated as 100 ($e^b - 1$) where b is the value of the estimated coefficient on the indicator variable. (In this example, b is the value of the coefficient on the indicator variable for non-Hispanic.)

[12] Technically, the R-squared provides the percent of the variation in the dependent variable explained by the independent variables in the model, while the adjusted R-squared is adjusted for degrees of freedom. Because of the large sample size, the R-squared and adjusted R-squared are almost identical in this example.

Table 1: Pay Disparities Between Non-Hispanic and Hispanic Workers[a]

Panel A: Means or Percent[b]

Variable	Non-Hispanic	Hispanic
Hourly wage (2003$)	18.10	12.96
Male (%)	50.17	57.55
Age (years)	40.36	35.77
Northeast (%)	24.06	14.19
Midwest (%)	29.61	11.74
West (%)	21.23	44.38
South (%)	25.10	29.69
Education (years)	13.99	11.33
Observations	119,000	15,530

Panel B: Estimates of Non-Hispanic-to-Hispanic Pay Disparity in Log Wage Equations[c]

	Control Variables	Non-Hispanic Coefficient (1)	Adjusted R-squared (2)	Non-Hispanic-to-Hispanic Wage Advantage (3)
1	Non-Hispanic indicator only	0.313** (0.005)	0.03	36.8%
2	Non-Hispanic indicator, male indicator, age, age squared	0.275** (0.004)	0.18	31.7%
3	Non-Hispanic indicator, male indicator, age, age squared, education	0.065** (0.004)	0.32	6.7%
4	Non-Hispanic indicator, male indicator, age, age squared, education, occupation	0.061** (0.006)	0.38	6.3%

a. Table 1 is based on the author's calculations from the 2003 Current Population Survey.
b. All means or percentages differ significantly at the 1% level.
c. ** significant at 1% (two-sided tests). Panel B reports coefficients on the non-Hispanic indicator, with standard errors in parentheses, controlling for the indicated variables. All equations also include a constant term.

The results summarized in row 2 are based on a regression that includes the non-Hispanic indicator variable as well as information on the worker's sex and age. Inclusion of information on the worker's sex, age, and age squared substantially increases the explanatory power of the equation, raising the adjusted R-squared to 0.18. The magnitude of the coefficient on the non-Hispanic indicator declines and correspondingly shows a smaller wage disparity of 31.7% between non-Hispanic and Hispanic workers. The

reduction of the coefficient on the non-Hispanic indicator and the increase in the adjusted R-squared reflect the inclusion of variables that both influence wages (which raises the adjusted R-squared) and are correlated with Hispanic ethnicity (which affects the coefficient on the indicator for non-Hispanic). The reduction in the coefficient on the non-Hispanic indicator arises because the additional variables in the regression are correlated with Hispanic ethnicity. In particular, because wages rise with age (albeit at a decreasing rate) and non-Hispanics are on average older than Hispanics, part of the pay disparity is due to the age difference.

Turning now to the regression summarized in row 3, which also includes in the equation years of education, we find that the magnitude of the non-Hispanic-to-Hispanic pay disparity drops substantially to only 6.7%. The explanatory power of the equation also increases, with the adjusted R-squared increasing from 0.18 to 0.32. The very large impact of inclusion of years of education on the disparity and on the adjusted R-squared reflects the importance of education as a determinant of earnings, and also reflects the high correlation between years of education and Hispanic ethnicity. The regression results demonstrate that failing to take into account differences in education would misleadingly suggest a much larger pay gap due to Hispanic ethnicity than the pay gap estimated taking education into account.

Row 4 provides a final regression example for comparison. In this regression, I include indicator variables for occupation. Note that occupation may be itself the outcome of any possible discriminatory process that is under consideration. Inclusion of indicator variables for occupation yields a still smaller non-Hispanic-to-Hispanic pay disparity of 6.3%. But in comparison to the large impact of education on the pay disparity, the effect of including occupation on the pay disparity is relatively minor. This is not because occupation does not have an important influence on wages; to the contrary, occupation has an important influence on earnings. For instance, professional occupations pay much higher wages than food preparation occupations. But because education is such a critical determinant of occupation, inclusion of education accounts for much of the influence of occupation on wages.

The message here is that adding variables correlated with Hispanic ethnicity does affect the magnitude of the pay disparity. Clearly, there are other important determinants of wages that may be correlated with Hispanic ethnicity, such as English language proficiency. To the extent these other possible productivity characteristics are taken into account by the variables

already included in the equation (such as education and age), the addition of other variables will have a relatively small effect on the magnitude of the non-Hispanic-to-Hispanic disparity. However, this example indicates that it may be hard in theory to counter the charge that omitted variables bias is responsible for the remaining non-Hispanic-to-Hispanic disparity.

I now analyze the effect of skin color on wages for a sample of legal immigrants. This analysis will demonstrate the invariance of the magnitude of the skin color effect with respect to inclusion of additional explanatory variables and, in so doing, will argue against the prospect that omitted variables are responsible for the skin color effect on wages found in my earlier study.

II. SKIN COLOR DISCRIMINATION AGAINST IMMIGRANTS

In this Part, I show that, on average, immigrants with darker skin color receive lower pay than comparable immigrants with lighter skin color. Furthermore, I show that the magnitude of the skin color effect is largely invariant with respect to inclusion of additional productivity-related variables. My analysis takes into account the major factors that affect pay, and as I discuss, it is unlikely that the magnitude of the skin color effect is biased due to omission of important determinants of pay that are also correlated with skin color.

My analysis uses data from the New Immigrant Survey 2003.[13] This survey provides a large nationally representative sample, with 8,573 individual respondents, all of whom gained lawful permanent resident status in 2003. The sample is drawn from the electronic records maintained by the U.S. government. The survey contains extensive information on a wide range of topics, including health measures, pre-immigration history, family members, income, assets, transfer payments, insurance, religion, language skills, and labor market information.

I conduct an analysis similar to that reported above that examines the effect on the non-Hispanic-to-Hispanic pay disparity as additional productivity-related characteristics are included in the regression equation. Specifically,

[13] For more information on the survey, including the survey overview, data, and documentation, see The New Immigrant Survey (2003), http://nis.princeton.edu.

using data from the New Immigrant Survey 2003, I estimate a series of wage equations controlling for skin color and other characteristics and consider the effect on the magnitude of the skin color coefficient as additional productivity-related characteristics are added to the wage equation. The analysis proceeds in two stages. In the first stage, I include only demographic, physical, and labor market characteristics acquired before current employment. In the second stage, I also include characteristics associated with current U.S. employment. The purpose of proceeding in these two stages is to consider separately characteristics that precede U.S. employment and are thereby unlikely to be subject to any possible discrimination in the United States from characteristics that are determined by current employment, which may be subject to discrimination in the United States. The variables used in this analysis and the expected effect on wages of these variables are discussed below in the order in which I add the variables to the regression equation.[14]

The key variable of interest is the individual's skin color. Skin color is reported by the interviewer based on the interviewer's observation and is recorded as one of eleven values ranging from 0 (albino) to 10 (the darkest possible skin color), using a color scale designed by Massey and Martin.[15] The color scale provided to the interviewers shows a series of hands with color increasing in darkness.[16] The NIS scale provides finer gradations than are available in any other data set based on a national sampling frame. All other data sets outside of the medical area report skin color as one of three to five categories.[17] Earlier studies based on ordinal measures of skin color typically

---

[14] In this Essay, I do not provide individual cites to well-known results in the economics literature. For more information on variable definitions, economic rationale for the empirical specification, and cites to supporting literature, see Hersch, *supra* note 1.

[15] Douglas S. Massey & Jennifer A. Martin, NIS Skin Color Scale (2003), http://nis.princeton.edu/downloads/NIS-Skin-Color-Scale.pdf.

[16] The Field Interviewer Manual includes the instruction:

> As you know, human beings display a wide variety of physical attributes. One of these is skin color. Unfortunately discrimination on the basis of skin color continues to be a reality in American life. Substantial evidence suggests that lighter skinned people fare better in a variety of social and economic settings than those with darker skins. In order to detect such discrimination, it is important that the NIS include a measure of skin color. We therefore ask interviewers to use the Scale of Skin Color Darkness as a guide to rate the skin color of each respondent on a scale of 0 to 10, where 0 is the lightest possible skin color (such as that of an albino) and 10 is the darkest possible skin color.

Id.

treated this skin color measure as interval-level data.[18] This often is not empirically valid, and indeed the impact of skin color on wages estimated for African Americans differs based on whether skin color is treated appropriately as ordinal data rather than as interval-level data. In contrast, as my earlier study verifies,[19] the eleven-point scale used in the NIS can appropriately be treated as interval-level data. Thus, the coefficient on skin color in the log wage equations presented in Table 2 can be interpreted in the same way that the coefficient on any other interval-level variable, such as education, is interpreted. Specifically, the coefficient on skin color is interpreted as the average percent difference in wages associated with a one-unit difference in skin color.

Because skin color is reported by interviewer observation and therefore may be somewhat subjective, an additional concern is measurement error. By comparing country averages calculated from respondents to the NIS survey to objective skin color levels for the same country measured by a reflectance spectrometer, my earlier study demonstrates that the NIS skin color measure is highly reliable.[20] Thus, systematic measurement error is not likely to be a critical concern. However, random measurement error is possible. To the

---

[17] In addition to the New Immigrant Survey 2003 used here, data sets using a sampling framework and including interviewer-reported skin tone for blacks are the Detroit Area Study (DAS) for the years 1975, 1992, and 1995; the National Survey of Black Americans 1979–1980 (NSBA); and the General Social Survey (GSS) 1982. The Multi-City Study of Urban Inequality 1992–1994 (MCSUI) includes interviewer-reported skin color for all respondents (with the exception of white respondents in Detroit). The 1990 Latino National Political Survey (LNPS) includes interviewer-reported skin color for a sample of Hispanic/Latinos. Of these data sets, only the NSBA and MCSUI include information on individual earnings, although all of the data sets include information on family income and educational attainment. The NSBA reports skin color in five categories, and the MCSUI reports skin color in three categories.

[18] For examples of studies that treat ordinal skin color measure as interval-level data, see Michael Hughes & Bradley R. Hertel, *The Significance of Color Remains: A Study of Life Chances, Mate Selection, and Ethnic Consciousness Among Black Americans*, 68 SOC. FORCES 1105 (1990); Verna M. Keith & Cedric Herring, *Skin Tone and Stratification in the Black Community*, 97 AM. J. SOC. 760 (1991). These studies are based on the NSBA and analyze outcomes such as personal income and education, but not wages. By treating the ordinal skin color measure as though it is an interval scale, these studies indicate than there is an increasing penalty to darker skin color. In contrast, see Joni Hersch, *Skin-Tone Effects Among African Americans: Perceptions and Reality*, 96 AM. ECON. REV. 251 (2006); Arthur H. Goldsmith, Darrick Hamilton, & William Darity, Jr., *From Dark to Light: Skin Color and Wages Among African-Americans*, 42 J. HUM. RESOURCES 701 (2007). These studies also use data from the NSBA to show that the wage penalty is not an increasing penalty to those with darker skin color and, therefore, treating the skin color measure in the NSBA as an interval-level measure is empirically invalid.

[19] Hersch, *supra* note 1, at 359–60.

[20] *Id.* at 356–61.

extent that any measurement error is random, the skin color effects estimated in my earlier study and reported in this Essay are biased toward zero.

Hispanic ethnicity, race, and country of birth are included in all equations reported in this Essay. Individuals report whether they are of Hispanic or Latino ethnicity and report their race in detail, with multiple races recorded if reported. The racial group options are American Indian or Alaskan Native, Asian, black, white, and Native Hawaiian or other Pacific Islander. The survey identifies country of origin for the twenty-two countries with the largest number of immigrants and reports the regional groups (e.g., Latin America and the Caribbean) for the remaining sample. Because discrimination on the basis of Hispanic ethnicity and race may be present within the U.S. labor market, controlling directly for Hispanic ethnicity and race in the wage equations allows identifying the impact of skin color net of any effect on wages of Hispanic ethnicity and race. I control for country of birth because countries differ in numerous dimensions that may affect the productivity of immigrants from these countries once they are employed in the United States. Examples of country characteristics that might affect productivity in the United States include quality of education; use of English in business, government, or as a language of instruction for non-native English speakers; and familiarity with U.S. social norms (perhaps communicated through television).

It is important to note that Hispanic ethnicity, race, and country of birth are highly correlated with skin color. For example, almost all immigrants from Jamaica are racially categorized as black, and the average skin color for those from Jamaica is among the darkest on the NIS color scale.[21] By including these characteristics that are highly correlated with skin color, it is possible that multicollinearity will reduce the precision of the estimate of the skin color effect on wages. Thus, the findings of my earlier study, which demonstrate a statistically significant effect on wages of skin color even after controlling for Hispanic ethnicity, race, and country of birth, provide strong evidence of a direct effect of skin color on wages.[22]

The demographic, physical, and labor market characteristics acquired before current employment that are included in the regression analysis are age, sex, height, weight, family background, education (divided into home and

[21] Specifically, the average skin color values on the 0–10 scale for immigrants from Jamaica are 7.19 for females and 7.95 for males. *Id.* at 357.

[22] Hersch, *supra* note 1.

United States), occupation before migrating, English language skills, whether new arrival to the United States, potential U.S. work experience, and U.S. region. These are defined below.

An indicator variable for sex is included in the regression equations as numerous studies demonstrate that men have higher wages than comparable women. Age in years is calculated from the respondent's date of birth. Both age and age squared are included in the regression equations to allow for a nonlinear effect of age on wages.

Height and weight are self-reported by the respondent. I allow height to have a nonlinear effect on wages by considering the individual's height relative to the U.S. gender-specific mean height, and by allowing the effect of height below the mean to differ from the effect of height above the mean. I use information on height and weight to calculate the body mass index (BMI) for each individual and control in the regressions for BMI. There are two reasons for including these variables. First, both height and BMI may have productivity effects. Numerous studies have found height to have significant effects on wages, with taller individuals having higher wages.[23] Although excess weight may lower productivity, there is mixed evidence of the effect of BMI or obesity on wages. Second, and most relevant for my analysis of skin color, I find that height and skin color are correlated. With the exception of predominantly black countries, countries with taller individuals have on average lighter skin color. Thus, excluding height (which may have productivity effects) would raise the possibility that the skin color findings are spuriously driven by omitted variables bias.

Because skin color discrimination may exist in many of the countries from which the survey respondents migrated, it is possible that those with darker skin color experienced discriminatory treatment in their home country. Such discrimination may result in weaker unobservable labor market characteristics. One way to mitigate possible omitted variables bias arising from unobserved characteristics is to control for family background, as family background will control for economic opportunities while growing up. I use two measures to control for family background. One is father's educational attainment. The other is relative family income at age sixteen, which is reported in five

[23] My results also show that taller immigrants have higher wages. *See id.* at 346–47. Each inch of height above the U.S. mean for the individual's sex is associated with a 2% increase in wages. *Id.*

categories ranging from "childhood family income far below average" to "childhood family income far above average."

Respondents report their years of educational attainment before migrating as well as their years of education attained in the United States. There are two reasons for treating years of education attained before migrating as distinct from years of education attained within the United States. First, it is possible that U.S.-acquired education is more valuable in the United States than education acquired elsewhere, so the returns to the two categories of education may differ, with the return to education attained in the United States having a larger effect on wages. Second, and most relevant for my analysis of skin color, skin color may be correlated with education attainment in the home country. To the extent that immigrants experienced skin color discrimination in their home country, the quality of education may be lower for those with darker skin color than for those with lighter skin color, even among those with the same number of years of measured education.

Occupation in the previous job before migrating to the United States is included in the regression analysis, categorized into five occupational categories: professional and managerial, health, services, sales and administrative, and production. Not all respondents report a previous occupation, so the effect of previous occupation is estimated relative to those not reporting a previous occupation before migrating.

The measure of English language skills used in the analysis are respondents' self-reports of how well they understand spoken English, with those reporting that they understand spoken English very well or well considered more proficient in English than those who report that they understand spoken English not well or not at all.

As indicators of familiarity with the United States as well as of potential U.S. employment experience, I include two variables. The first variable is an indicator variable for whether the respondent is a new arrival immigrant or an adjustee immigrant. New arrival immigrants acquired their immigration documents abroad through the U.S. consular service in their home country. Adjustee immigrants were already in the United States at the time they reached lawful permanent resident status and include temporary workers and students, who may have performed some work under temporary visas, as well as those adjusting from an illegal to a legal status. Because adjustee immigrants have had time to become more established in the United States, they are likely to have higher wages than new arrival immigrants. The second variable indicates

time since the first U.S. job and the date of interview, which provides a measure of potential U.S. labor market experience. To allow for nonlinearities in the returns to experience, I include both potential experience and its square.

There are regional differences in average pay as well as regional differences in the concentration of immigrants. Region may thereby affect individual wages through the general pay structure as well as by providing networks to new legal immigrants. To control for regional variation in wages, I control for the location to which the green card was sent by grouping reported location into the four broad Census categories of the Northeast, South, Midwest, and West.

The remaining variables included in the analysis are those associated with current U.S. employment. The analysis examines the impact on the magnitude of the skin color effect of inclusion of the major employer and job characteristics widely shown to affect wages. Because these are characteristics over which the current employer has control, it is possible that any skin color discrimination will be manifested via these characteristics. Thus, while omitting these variables does not create a "bias" in the estimate of the effect on wages of skin color, including these variables may lead to an understatement of the extent of skin color discrimination insofar as these variables reflect mechanisms by which skin color discrimination occurs. These variables and their expected effect on wages and relation to skin color are discussed next.

Recall that all survey respondents are permanent legal resident immigrants. The four broad categories of visa types are employment visa, visa as a spouse of a U.S. citizen, diversity visa, and all other permanent legal resident visa types, which include nonspousal family members of U.S. citizens, refugees, asylees, and those adjusting from an illegal to legal status. Eligibility for any of these visa types may be influenced by skin color, in part because visa applications require photos for identification. Employment visas are particularly subject to the possibility of skin color discrimination, as most employment visa holders are sponsored by an employer.

Tenure (i.e., years of experience with the current employer) is calculated using information on the start date for the current job, and both tenure and tenure squared are included in the regression to allow tenure to have a nonlinear effect on wages. It is widely established that wages rise with tenure at a decreasing rate. Thus, to the extent that employers discriminate against immigrant workers with darker skin color, immigrants with darker skin color

may have lower tenure as they are hired last among a set of applicants or change jobs more frequently to avoid discrimination.

Characteristics other than skin color that may be related to wages that are included in the analyses are indicators for governmental employer, union contract, paid hourly rate, full-time employment, and whether the individual is self-employed. Governmental employers typically pay lower wages than private employers. Jobs in which pay and working conditions are set by union contract typically have higher pay, but access to such jobs may be restricted by skin color discrimination. Jobs that are paid hourly, in contrast to salaried positions, may be of lower status and may be correlated with skin color. Full-time employment is associated with higher hourly wages, and access to full-time jobs may be correlated with skin color. Workers may choose self-employment in reaction to employer discrimination on the basis of skin color.

Because outdoor work may cause skin color to darken and may also be associated with lower pay, I also control for whether outdoor work is probable for the worker.[24] Inclusion of information on outdoor work mitigates the possibility that the skin color findings are spuriously driven by omitted variables bias. Finally, I consider the role of occupation in the United States, with occupation grouped into the same five categories used to group occupation before migration.

[24] There is no direct information on whether the worker's job requires outdoor work, so I classify whether outdoor work is probable for an occupation based on the U.S. Bureau of Labor Statistics categorization of occupations with extensive outdoor work.

Table 2: Estimates of the Skin Color Effect for New Legal Immigrants in Log Wage Equations[a]

	Control Variables[b]	Skin Color Coefficient (1)	Adjusted R-squared (2)	Lightest-to-Darkest Wage Advantage (3)
Including only pre-U.S. labor market characteristics				
1	Skin color, Hispanic ethnicity, race, country of birth	−0.017* (0.007)	0.15	17%
2	1 + male, age, height, weight, time period	−0.021** (0.007)	0.24	21%
3	2 + family background, education	−0.018** (0.007)	0.29	18%
4	3 + occupation in last job abroad	−0.018** (0.007)	0.30	18%
5	4 + English language, new arrival, potential U.S. work experience, U.S. region	−0.017** (0.006)	0.38	17%
Including current characteristics				
6	5 + visa type	−0.013* (0.006)	0.43	13%
7	6 + tenure, employer characteristics, job characteristics	−0.014* (0.006)	0.45	14%
8	7 + occupation in U.S.	−0.011+ (0.006)	0.49	11%

a. Table 2 is based on the author's calculations from the New Immigrant Survey 2003. It is adapted from Joni Hersch, *Skin Color, Immigrant Wages, and Discrimination*, *in* RACISM IN THE 21ST CENTURY: AN EMPIRICAL ANALYSIS OF SKIN COLOR 85 tbl.5.2 (Ronald E. Hall ed., 2008). The number of observations is 1,536. ** significant at 1%; * significant at 5%; + significant at 5.5% (two-sided tests). Table 2 reports coefficients on skin color, with standard errors in parentheses, controlling for the indicated variables. For additional information, see Hersch, *supra* note 1.

b. The complete list of control variables is: skin color (0–11 scale), Hispanic ethnicity, race (7 categories), country of birth (22 countries), sex, age, age squared, inches below U.S. gender average height, inches above U.S. gender average height, BMI, time period, father's education, relative family income at age 16, years of education before United States, years of U.S. education, occupation in last job before migrating to the United States, understands English very well or well, new arrival, potential U.S. work experience, potential U.S. work experience squared, region of United States, visa type (4 categories), tenure with current employer, tenure with current employer squared, governmental employer, union contract, outdoor work probable, paid hourly rate, full-time employment, self-employed, and U.S. occupation. All equations also include a constant term.

Table 2 reports the effect on the skin color coefficient in alternative wage equation specifications. I start with a minimal specification controlling initially only for skin color, Hispanic ethnicity, race, and country of birth, and then sequentially consider the effect of adding additional variables on the skin color coefficient, the adjusted R-squared, and the percent disparity in wages for the lightest to darkest immigrants. In all equations, the dependent variable is the log of the hourly wage. I analyze the sample of workers with hourly wages between $1.50 and $100 per hour, who have skin color reported, and for whom individual country is identified. The sample size is 1,536.[25]

The first set of results summarized in rows 1–5 of Table 2 include only demographic, physical, and labor market characteristics acquired before current employment, in addition to skin color, Hispanic ethnicity, race, and country of birth. The second set of results, summarized in rows 6–8, additionally include characteristics associated with current U.S. employment.

The first row of Table 2 reports the coefficient on skin color and the adjusted R-squared controlling only for Hispanic ethnicity, race, and country of birth. The coefficient on skin color is −0.017 and is statistically significant at the 5% level. The adjusted R-squared is 0.15. The coefficient on skin color indicates that without controlling for any other variables, a one-unit increase in skin color value lowers wages by 1.7% on average. Considering the difference between those with the lightest skin color and those with the darkest skin color, the coefficient of −0.017 indicates that without controlling for any variables other than Hispanic ethnicity, race, and country of birth, those with the lightest skin color have hourly wages that are 17% higher than those with the darkest skin color. Note that because this equation (as well as all others discussed in this Essay) controls for Hispanic ethnicity and race, the independent effect of skin color on wages is isolated. This is a very important point. It is widely established that Hispanics and African Americans in the United States have lower wages than whites. Because of the high correlation between skin color and Hispanic ethnicity and race, without also controlling for Hispanic ethnicity

[25] *See* Hersch, *supra* note 1, at 368–79 (providing complete information on sample composition). I note that about half of the original sample is not included in the analysis because about half of the interviews were conducted entirely by phone, so interviewers did not observe skin color. In addition, as the analysis examines the influence of skin color on pay, the sample analyzed is further restricted to those reporting pay, which eliminates about half of the remaining sample. Complete regression results for the sample with individual country reported and for the sample with either individual country or region reported are provided in my earlier study. *Id.* at 370–71, 376–77.

and race what we observe as an independent skin color effect may instead be the effect of Hispanic ethnicity and race on wages.

The remaining rows in Table 2 present the coefficient on skin color and the adjusted R-squared as we add additional explanatory variables. Start by considering the results summarized in rows 2–5, which include only demographic, physical, and labor market characteristics acquired before current employment (in addition to skin color, Hispanic ethnicity, race, and country of birth). Row 2 adds variables for sex, age, height, and time period (to account for price-level changes). Row 3 adds to the variables included in the equation summarized in row 2 variables for family background and education. Row 4 adds variables for occupation in last job before migrating to the United States to the equation summarized in row 3. Row 5 adds variables for English language proficiency, whether a new arrival or adjustee immigrant, potential U.S. work experience, and region of the United States to the equation summarized in row 4.

Notably, the coefficient on skin color in rows 1–5 ranges from −0.017 to −0.021. Specifically, compare the coefficient on skin color of −0.017 in row 1 to the coefficient on skin color in row 5. This value too is −0.017 and is statistically significant at the 1% level. Thus, the magnitude of the effect of skin color is not reduced at all by inclusion of extensive characteristics that affect wages. This is not because these additional variables lack explanatory power. To the contrary, the adjusted R-squared of 0.38 in row 5 shows a substantial increase in explanatory power relative to row 1, which has an adjusted R-squared of 0.15. The invariance or even increase in the skin color coefficient as additional variables are added demonstrates that the correlation between skin color and these other characteristics that affect wages is low, and that the skin color effect observed in row 1 is not likely to be a consequence of omitted variables bias.

The stability of the coefficient on skin color as additional variables are added to the regression equations suggests that even if there are other relevant productivity variables not associated with current U.S. employment that could be included in the regression, it is unlikely that the magnitude of the skin color effect would change greatly. The equation summarized in row 5 controls extensively for the most important productivity characteristics acquired before current U.S. employment. For the skin color effect to be driven by omitted variables bias, two conditions must be met. First, any omitted variables would need to be correlated with both skin color and market productivity, and those

with lighter skin color would need to have more of the productivity-related characteristics. If those with darker skin color instead have more of the productivity-related characteristics, the true darkest-to-lightest skin color penalty would be larger than 17%. Second, any omitted variables would need to have a low correlation with the variables already included in the analysis. Otherwise, the effect of any such omitted variables is accounted for by inclusion of the variables in the equation. Because the equation summarized in row 5 controls for extensive characteristics, it becomes difficult to posit additional variables that may influence wages that are not themselves correlated with variables already included in the equations.[26]

Rows 6–8 add variables associated with current employment in the United States. The magnitude of the skin color effect is smaller once characteristics associated with current employment are included in the wage equation. The magnitude of the skin color effect reported in row 8, which includes the most comprehensive set of characteristics, indicates a darkest-to-lightest skin color penalty of 11%, statistically significant at the 5.5% level. Since these characteristics are under the control of employers in the United States, the decline in the magnitude of the skin color effect demonstrates that skin color discrimination is partly manifested indirectly through work characteristics, particularly with respect to access to employment visas and occupation.

Note that my analysis allows us to identify the source of skin color discrimination as arising within the U.S. labor market. Although there is much evidence of discrimination on the basis of skin color in a number of countries, new lawful immigrants to the United States will not have experienced historic differential treatment within the United States on the basis of skin color, in contrast to African Americans. For example, consider someone from Brazil, who has light skin relative to other Brazilians. He or she is likely to have experienced preferential treatment while living in Brazil. But after migrating to the United States, this person may have darker skin relative to the U.S.

[26] For example, Richard Lynn (among others) has proposed that skin color is an index of white ancestry and that whites are more intelligent than those of other races. Under this view, intelligence is then an omitted variable, and inclusion of a measure of intelligence would eliminate the negative effect on wages of darker skin color. However, if white ancestry is actually correlated with superior intellectual ability (and there is simply no scientific basis for this view) then the effect of intelligence would already be captured by the observed variables such as education, previous occupation, and family background that are included in the wage equation. *See* Richard Lynn, *Skin Color and Intelligence in African Americans*, 23 POPULATION & ENV'T 365 (2002).

population. This means that discriminatory treatment on the basis of skin color is arising within the U.S. labor market rather than reflecting a legacy of treatment in the originating country.

III. SKIN COLOR AND EMPLOYMENT DISCRIMINATION LITIGATION

Skin color discrimination claims[27] have been on the rise. While still a small share of the 85,000 charges filed annually, the Equal Employment Opportunity Commission (EEOC) reports that allegations of skin color discrimination have been rising, with 413 charges filed in 1994, and 1,382 in 2002.[28] In recognition of ongoing concerns about race and color discrimination, in 2007 the EEOC launched the E-RACE (Eradicating Racism and Colorism from Employment) Initiative.[29]

The role of color as distinct from race or national origin warrants some discussion. Color has served as a proxy for race or nationality in discrimination claims.[30] For example, in *Vigil v. City & County of Denver*, the court found that claims of discrimination against Mexican Americans based on color are permissible because Mexican Americans are often identified on the

---

[27] Title VII of the Civil Rights Act of 1964 states (in part):

> It shall be an unlawful employment practice for an employer—
> (1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin . . . .

42 U.S.C. § 2000e-2(a) (2000). Color discrimination claims are also brought under 42 U.S.C. § 1981, which guarantees all persons the same rights as white citizens. (Section 1981 makes no reference to the words "race" or "color.")

[28] Press Release, EEOC, EEOC Settles Color Harassment Lawsuit with Applebee's Neighborhood Bar and Grill (Aug. 7, 2003), http://www.eeoc.gov/press/8-07-03.html [hereinafter Press Release].

[29] EEOC, The E-RACE Initiative (2008), http://www.eeoc.gov/initiatives/e-race/index.html. The home page for the E-RACE Initiative states:

> The E-RACE Initiative is designed to improve EEOC's efforts to ensure workplaces are free of race and color discrimination. Specifically, the EEOC will identify issues, criteria and barriers that contribute to race and color discrimination, explore strategies to improve the administrative processing and the litigation of race and color discrimination claims, and enhance public awareness of race and color discrimination in employment.

Id.

[30] For an overview of major color discrimination lawsuits, see Taunya Lovell Banks, *Colorism: A Darker Shade of Pale*, 47 UCLA L. REV. 1705 (2000). For an overview of the history of color discrimination relevant to African Americans, see Trina Jones, *Shades of Brown: The Law of Skin Color*, 49 DUKE L.J. 1487 (2000).

basis of their skin tone.[31] Color discrimination claims have also been brought by individuals against individuals of their own race. For example, in 2003, a dark-skinned black waiter at an Applebee's restaurant received $40,000 to settle a claim that his light-skinned black supervisor was discriminating against him.[32]

The empirical evidence presented in this Essay suggests that observed opposition to immigrants arises in part from ethnic discrimination as characterized by outward appearance, with immigrants who have lighter skin faring better than their counterparts who are darker, even after accounting for race and country of origin. Projected population trends indicate that the United States is becoming less white. Non-Hispanic whites comprised 66% of the total population in 2008.[33] By 2050, this share is projected to drop to 46%.[34] Such trends suggest that color discrimination lawsuits may continue to increase.

Within the context of such litigation, there will be a debate as to whether observed differences in pay based on skin color reflect labor market discrimination or are reflective of legitimate productivity differences. This Essay demonstrates that statistical analysis can be used to disentangle these effects. From a broad societal standpoint, there is skin color discrimination that cannot be accounted for by differences in labor market productivity.

---

[31] Vigil v. City & County of Denver, 1977 WL 41, at *1 (D. Colo. May 23, 1977) ("The key factor in determining whether § 1981 should apply is whether a motivation for the discrimination was the victim's color. Plaintiff is a Mexican-American. Although skin color may vary significantly among those individuals who are considered Mexican-Americans, skin color may be a basis for discrimination against them. We note that skin color may vary significantly among individuals who are considered 'blacks' or 'whites'; both these groups are protected by § 1981, and § 1981 is properly asserted where discrimination on the basis of color is alleged.").

[32] Press Release, *supra* note 28.

[33] U.S. Census Bureau, An Older and More Diverse Nation by Midcentury, U.S. Census Bureau News (2008), http://www.census.gov/Press-Release/www/releases/archives/population/012496.html.

[34] *Id.*