# Statistical Discrimination, Productivity and the Height of Immigrants*

Shing-Yi Wang

New York University

September 30, 2010

**Abstract**

Building on the economic research that demonstrates the positive relationship between height and worker ability, this paper considers whether employers use height as a tool for statistical discrimination. The analysis focuses on immigrants and native-born individuals because employers are likely to have less reliable signals of productivity for an immigrant than a native-born individual. Using multiple data sets, the paper presents a robust empirical finding that the wage gains associated with height are almost twice as large for immigrants than for native-born individuals. This result is consistent with two hypotheses. First, in the relative absence of other sources of information about immigrants, employers place more weight on height for immigrants than for native-born individuals. Second, height is more correlated with productivity for immigrants than for native-born individuals. The empirical results provide strong support for the hypothesis that the productivity gap between tall and short immigrants is greater than the productivity gap between tall and short native-born workers. The hypothesis of statistical discrimination based on height is rejected. These results have implications for our understanding of the process of economic assimilation of immigrants.

*Email: shingyi.wang@nyu.edu. This paper has benefited from conversations with Santosh Anagol and Nicola Persico. A previous version has benefited from comments from Joe Altonji, Hanming Fang, Fabian Lange, T. Paul Schultz and Chris Udry. April Collaku provided excellent research assistance. All errors are my own.

# 1 Introduction

A large amount of empirical evidence demonstrates a positive correlation between height and earnings throughout the world. In the context of developing countries, the focus of this analysis has been on the relationship between health and nutrition inputs and height (Bozzoli, Deaton and Quintana-Domeque 2009, Deaton 2008, Steckel 1995, Strauss and Thomas 1998). This is not surprising given that physical size and health are likely to be important for manual labor in developing countries (Glick and Sahn 1998). However, sizable wage gains associated with height persist in rich countries such as the United States and Britain where the importance of physical strength is likely to play a smaller role in the labor market. Taste-based discrimination against short people is a possible explanation (Kuhn and Shen 2009).[1] More convincing explanations are that the returns to height in developed countries are explained by the relationship between height and cognitive ability (Case and Paxson 2008), and non-cognitive ability such as social skills (Persico, Postlewaite and Silverman 2004).

Given that height is easy to observe and strongly correlated with unobserved aspects of worker productivity, it is possible that the wage returns on height reflect, at least in part, statistical discrimination by employers. In the absence of other information about worker productivity, employers may use height to infer differences in productivity across workers. While other empirical papers on statistical discrimination have focused on race and gender, this paper introduces height as a possible mechanism of employer statistical discrimination.[2]

I examine this question by comparing immigrants and native-born individuals in the United States and in the United Kingdom. The comparison of immigrants and native-born individuals is particularly useful for this exercise because it is plausible that employers face substantial information differences in comparing the expected productivity of immigrants and native-born individuals. Employers may have uncertainty about the academic degree system, the curriculum or the quality of schools in other countries. Furthermore, language barriers may generate or exacerbate noise in employers' assessment of productivity signals from immigrants. The impact of information asymmetries on labor market outcomes of immigrants has been analyzed in the context of theoretical models of brain drain where it is assumed that host country employers have less information than employers in the originating country (Chau and Stark 1999, Kwok and Leland 1982). Rather than analyzing the impact of asymmetric information on labor market opportunities in across countries, this paper considers the

[1]This hypothesis is consistent with the findings on the returns to beauty (Hamermesh and Biddle 1994) and weight (Averett and Korenman 1996).

[2]The statistical use of height has been considered by Mankiw and Weinzierl (2009). Their theoretical paper shows that government taxation of height, which is correlated with productivity but not affected by effort, is consistent with a standard utilitarian framework.

effects of information asymmetries between immigrants and native-born individuals within a country. To my knowledge, this is the first paper that attempts to empirically examine the role of statistical discrimination on immigrant outcomes. The results of this paper contribute to our understanding of the process of economic assimilation of immigrants and the individual decision regarding whether to immigrate and whether to stay in the host country.

The previous theoretical and empirical literature on statistical discrimination has focused on employers use of average outcomes by race and gender (Altonji and Pierret 2001, Coate and Loury 1993, Farber and Gibbons 1996). A different strand of theoretical literature on statistical discrimination focuses on the amount of *uncertainty* around the information available to employers (Aigner and Cain 1977, Phelps 1972, Lundberg and Startz 1983, Oettinger 1996). In these models, employers have an observable, continuous signal of productivity, but the quality of this information is different across groups. Phelps (1972) and Aigner and Cain (1977) show that the expected productivity (and hence wages) will be flatter for the group for which there is greater uncertainty in the signal. Lundberg and Startz (1983) demonstrate that this type of statistical discrimination can lead to an equilibrium in which there is lower investment in skills in the group that has more noise in the signal of productivity even in the absence of differences in underlying ability.

My paper emphasizes differences in the precision of information that employers have about immigrants as compared with native-born individuals; thus, the main framework used in this paper builds on these latter models of statistical discrimination. I extend the model to a context where there are two signals of productivity, height and education, and there is more uncertainty regarding the signal of education for immigrants than for native-born individuals. A key prediction of the model is that the wage returns to height will be higher for the group for which the quality of other signals is worse. In other words, a model of statistical discrimination suggests that employers will place more weight on height and less weight on education for immigrants relative to native-born individuals. Using several data sets, I presents a robust empirical finding that the wage gains associated with height are almost twice as large for immigrants than for native-born individuals. In addition, the returns to education are slightly lower for immigrants. While this empirical result is consistent with the model of statistical discrimination, it is also consistent with an alternative explanation in which there is no statistical discrimination by employers but the underlying mapping of height and education into productivity is different for immigrants than for native-born individuals.

To disentangle these two hypotheses, I use additional predictions of the model. To analyze the first hypothesis of statistical discrimination, I examine the idea that as uncertainty about immigrant signals is reduced, the returns to height and education of immigrants should move to be more similar to those of native-born individuals. To analyze the alternative hypothesis, I use measures of worker productivity that are available in the data but not to employers to test whether height is more correlated with these measures of productivity for immigrants than for native-born individuals.

The results of the paper do not support the hypothesis that employers use height to statistically discriminate against immigrants in the relative absence of other good signals about their productivity. Instead, the results suggest that that the productivity gap between tall and short immigrants is greater than the productivity gap between tall and short native-born workers. The differences in the mapping between height and productivity is consistent with the idea that health and nutrition inputs vary considerably in developing countries and have long-run consequences for productivity and adult height. While height contains more information about productivity for immigrants, this information is not used by employers in the U.S. or in the U.K.

# 2 Conceptual Framework

The classic model of statistical discrimination is based on the an observable, continuous measure of skill (Aigner and Cain 1977, Phelps 1972). This skill measure has been conceptualized as a test score such as on a college entrance exam or an employer administered exam. The economic literature on test scores and statistical discrimination of groups in labor markets has been almost entirely theoretical. This may reflect that the reality that very few employers administer exams as part of their hiring practices or even ask about standardized test scores. The framework presented in this section builds on these existing theoretical models with height representing the observable, continuous measure of skill. One of the advantage of the focus on height is that it is plausibly observed by employers.

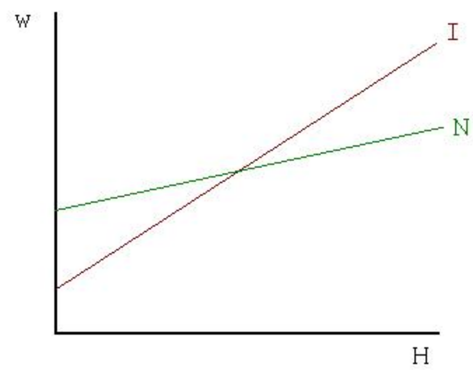
## 2.1 Statistical Discrimination

In the classical model of statistical discrimination, in making decisions regarding hiring and assignment of workers, employers use a measure, $H$, that is correlated with the worker's true marginal productivity, $P$. The relationship is given by:

$$H_i = P_i + \epsilon_i \tag{1}$$

where $\epsilon$ is a normally distribution error term with mean zero and a constant variance that is independent of $P$. While $H$ is observable to employers, $P$ is not. Thus, employers want to estimate marginal

Figure 1: Relationship Between Wages and $H$




productivity which is given by:

$$\widehat{P}_i = (1-\gamma)\alpha + \gamma H_i \tag{2}$$

where $\widehat{P}_i$ denotes predicted marginal productivity, $\alpha$ is the group mean of $H$ and

$$\gamma = \frac{Var(P)}{Var(P) + Var(\epsilon)}. \tag{3}$$

Assuming that workers are paid their marginal product, an individual's equilibrium wage will be a weighted average of mean productivity and the individual signal of productivity, $H_i$.

Consider two groups, denoted by $I$ and $N$, where $H$ is a more reliable indicator for one group than the other. In other words,

$$H_i^I = P_i + \epsilon_i^I; \quad H_j^N = P_j + \epsilon_j^N \tag{4}$$

and $Var(\epsilon^N) > Var(\epsilon^I)$. In this case, employer statistical discrimination will lead to the slopes $\gamma$ differing for the two groups with $\gamma^I > \gamma^N$, as shown in Figure 1. Tall immigrants will be paid more than tall native-born individuals but the reverse is true for short immigrants.

There are a few possible reasons that height may be a more reliable signal of ability and productivity for immigrants than for native-born individuals is plausible. One possible explanation is that there is more variance (perhaps genetic) in the height of Americans and Britons than in other groups that is not reflective of ability. Another potential (and more likely) explanation is that height is a more reliable signal of productivity for immigrants than native-born individuals *conditional* on other worker characteristics that are observable to the employer. In this case, height is correlated with something, such as educational attainment, that is observed with less noise for native-born individuals than for immigrants. Thus, employers place less weight on educational attainment for immigrants than native-born individuals because the signal of human capital has more noise for immigrants, and relatively more weight on height which is clearly observable.

To see this formally, consider the case where the true relationship determining marginal productivity, $P^*$, is given by

$$P_i^* = \alpha + H_i^* \beta + X_i^* \delta + \epsilon_i \tag{5}$$

where $H^*$ is perfectly observable by employers. True human capital, denoted by $X^*$, is observed with error:

$$X_i = X_i^* + \zeta_i. \tag{6}$$

I assume that $\zeta_i$ is uncorrelated $X_i^*$ and $H_i^*$.

The estimated returns to $H$, $\hat{\beta}$, is given by

$$\hat{\beta} = \frac{Cov(H_i^* \beta + X_i^* \delta, H_i^* - X_i \hat{\pi}_{xh})}{Var(H_i^* - X_i \hat{\pi}_{xh})} \tag{7}$$

$$\tag{8}$$

where $\hat{\pi}_{xh} = \frac{Cov(X_i, H_i^*)}{Var(X_i)}$.

After a little additional algebra, we get

$$\hat{\beta} = \frac{\beta Var(H_i^*)[1 - \frac{Cov(X_i, H_i^*)^2}{Var(X_i)Var(H_i^*)}] + \delta \frac{Cov(X_i^*, H_i^*)Var(\zeta_i)}{Var(X_i^*) + Var(\zeta_i)}}{Var(H_i^*) - \frac{Cov(X_i, H_i^*)}{Var(X_i)}} \tag{9}$$

$$= \beta + \frac{\frac{Cov(X_i^*, H_i^*)Var(\zeta_i)}{Var(X_i^*) + Var(\zeta_i)}}{Var(H_i^*) - \frac{Cov(X_i, H_i^*)}{Var(X_i)}} \delta \tag{10}$$

$$= \beta + \frac{\frac{Cov(X_i^*, H_i^*)Var(\zeta_i)}{Var(X_i^*) + Var(\zeta_i)}}{Var(H_i^*)(1 - R_{xh}^2)} \delta \tag{11}$$

where $R_{xh}^2$ is the R-squared of a regression of $X$ on $H^*$. The sign of the fraction preceding $\delta$ in equation 11 is determined by the direction of the correlation between $H^*$ and $X^*$. If $H^*$ and $X^*$ are positively correlated and $\delta > 0$, then error in the employers' observations of $X^*$ (in other words, $Var(\zeta_i) > 0$) and leads to an overestimate of the returns to $H$. Furthermore, if the differences across the two groups are such that $Var(\zeta_i^I) > Var(\zeta_i^N)$, then all else equal, statistical discrimination by employers implies that $\hat{\beta}^I > \hat{\beta}_N$.

The estimated returns to $X$ are given by

$$\hat{\delta} = \delta\left[1 - \frac{Var(\zeta_i)}{(1-R_{xh}^2)(Var(X_i^*) + Var(\zeta_i))}\right]. \tag{12}$$

Thus, under statistical discrimination, the returns paid by employers for human capital are attenuated by the noise associated with the signal. Greater the noise in the signal leads to a lower relationship between wages and observed human capital.

In the data, this hypothesis suggests that the wage gains associated with height to be greater for immigrants than for native-born individuals and the wage gains associated with education to be greater for native-born individuals than for immigrants. Furthermore, if uncertainty in immigrants' signals of productivity is reduced (either through time in the host country or through human capital acquisition in the host country), the model of statistical discrimination implies that the gaps between the two groups in returns should close.

## 2.2 Differences in the Relationship between Height and Productivity

The pattern of larger returns to height for immigrants than for native-born individuals is consistent with a model of statistical discrimination but it is also consistent with a model where the relationship between individual productivity and height are different across groups. In other words, it may be the case that employers do not use height to statistically discriminate among workers but

$$H_i^I = b_I P_i + \epsilon_i^I; \quad H_j^N = b_N P_j + \epsilon_j^N \tag{13}$$

and $b_I < b_N$ and $\epsilon_i^I = \epsilon_i^N$. In this case, we also get $\gamma_N < \gamma_I$.

There are three possible explanations that height and productivity may have a different relationship for immigrants than for native-born individuals. First, there may be variation in returns to height across types of jobs, and immigrants sort into jobs where height has greater returns. For example, it may be the case that height increases productivity for certain types of physical labor such as fruit picking or construction, and immigrants tend to work in these types of jobs. If this is true, the gap in the returns to height should disappear with the inclusion of controls for industry and occupation. Second, a different relationship between height and productivity may be explained by the selection of the types of individuals who choose to immigrate to the U.S. and the U.K. Third, there may be a stronger relationship between height and ability for immigrants due to the mapping of height and nutrition, cognitive ability or non-cognitive skills. If Americans and Britons experience less variation in nutrition and health inputs during the key stages of their development than individuals from poor countries, then immigrant height may reflect more information about health and cognitive development than native-born height. If either of the last two explanations is correct, we expect that the empirical relationship between height and health or ability to be very similar to the relationship between height and wages.

# 3 Data

The four main data sets used in this analysis are the National Health Interview Survey (NHIS), the Health Survey of England (HSE), the Health and Retirement Survey (HRS) and the New Immigrant Survey (NIS). These four household-level data sets contain the necessary information on height, immigrant status and labor market outcomes, and include a substantial number of immigrants.

The NHIS is a repeated cross-sectional survey conducted by the U.S. National Center for Health Statistics and the Centers for Disease Control Prevention. It is the principal source of data on the health of the civilian population in the U.S. In this paper, I use the waves from 2000 to 2007. While the annual survey began in 1989, only the waves starting after 2000 contain information on the area of birth of survey respondents that were born outside of the U.S.

The HSE is the only British data set in this analysis. It is a representative sample of adults in private households in Britain conducted by the Social Survey Division of the ONS National Statistics. The repeated cross-sectional data was collected beginning in 1991. I use the waves from 1997-1999 and 2004 because these rounds contain information about country of birth and thus allow for identification of immigrants. Immigrants were over-sampled in the 1999 and 2004 rounds and comprise over 30% of survey respondents in those two years.

Conducted by the University of Michigan, the HRS is a panel of Americans over the age of 50 that occurs every two years. Given that the focus of this paper is on labor market experiences rather than the transition into retirement, I use only the 1992 wave. I construct a pseudo-panel with retrospective questions about past labor market experiences.[3] The average age associated with the information is substantially higher than the other data sets.

The NIS is a nationally representative sample of legal immigrants drawn from U.S. government

[3] In addition to current labor market information, the survey covers job information immediately before retirement for retired respondents and work prior to the most recent job. For each of these jobs, the survey asks for both the starting and ending (or most recent) wage information.

records on admission to legal permanent residence between 1996 and 2003 (doubling checking these dates with NIS). In this paper, I use the full adult and spouse sample which occurred in 2003. While the sample of the NIS almost entirely excludes native-born Americans, the data set offers the advantage of rich retrospective information about the pre-immigration characteristics and experiences of survey respondents. This data set differs from the NHIS and HRS in that the immigrants are relatively recent arrivals and legally admitted into the U.S.

In all data sets, I restrict the sample to adults between the ages of 20 and 60. Immigrant status is defined by country of birth. Thus, individuals born in the U.S. who lived in another country before returning to the U.S. would not be classified as an immigrant. Specific country of birth is only available in the HSE and NIS; the NHIS has information on region of birth while the HRS only identifies whether the individual was born in the U.S. or not. Height is a self-reported measure in the NIS, NHIS and HRS, but it is measured by the interviewer in the HSE. Respondents are allowed to report their height in either the metric or U.S. customary units in the NIS. I drop a handful extreme outliers for adult height that are in the NIS.

The measure of earnings from the NIS is the individual's reported salary in 2003. Similarly, I use the individual's reported annual earnings in the NHIS. In the 1993 wave of the HRS, I use self-reported earnings for the respondent's current job if employed, the most recent job if retired and one additional long-term job for all respondents. In the HRS pseudo-panel, the median year of employment data is 1986 and the earliest year of data is 1938.[4] Because the NHIS and HRS data span several years, I use a deflator to convert the earnings data into 2004 dollars.

In contrast to the other data sets, the key disadvantage of the HSE data is that income is not reported at the individual level. For the HSE data, I construct an individual level measure using joint annual income reported at the couple level. In the majority of cases, the assignment is simple for the households where an individual is not married or is the only person in the household working. In other cases, the individuals' share of joint income is weighted by whether they work full-time or part-time.[5] The measure of income in the HSE is converted into 2004 pounds using a GDP deflator from the U.K. Office of National Statistics.

Table 1 displays summary statistics for the four data sets, broken down by whether the individual was an immigrant or native-born. On average, native-born individuals are taller than immigrants by

[4]To address concern regarding recall bias in past wages, I examined all of the results with only recent information on current job and the most recent job for retirees. The results are robust to this truncation and available upon request.

[5]For example, if both members are working full-time, the individual measure of income evenly divides their joint income. If one member works full-time and the other part-time, the member who works full-time is assigned three-quarters of the joint income and the remaining one-quarter is assigned to the part-time worker.

about two inches for men and one inch for women. The gap in the earnings between immigrants and native-born individuals varies across samples, and cannot be explained by the gap in human capital accumulation reflected in years of schooling.

Conditional on employment, American immigrants in the NHIS are quite similar to those in the NIS along most observable characteristics. Male NIS immigrants earn slightly more and are more likely to be in a white collar job than NHIS. This pattern is reversed for women with female NIS immigrants earning slightly less than female NHIS immigrants. These difference may reflect either that NIS sampling does not include illegal immigrants or the differences in the time periods covered. Table 1 indicates that HRS immigrants have lower earnings and lower quality jobs. This is likely explained by the older cohorts from which HRS samples.

Panel A of Table 2 shows characteristics of immigrants in the four main data sets. The average NHIS immigrant in my analysis entered the U.S. at age 20 and has lived in the U.S. for over 17 years.[6] The numbers are fairly similar for HSE immigrants; on average, they entered after age 19 and have lived in the U.K. for just over 20 years. The average characteristics for NIS and HRS immigrants are quite different, and this reflects the unique sampling approaches of the NIS, which includes recent, legal immigrants, and the HRS, which includes older adults. The average NIS immigrant entered in their late twenties and have resided in their host country for 6 to 7 years. The average HRS immigrant entered in their late twenties and have resided in the their host country for about 19 years. Host country education refers to whether the individual completed any education in the host country.[7] This is constructed from direct information on pre-immigration education in the NIS. However, the other data sets lack specific information about the location of a respondent's schooling; the variable is constructed to equal one if the number of years of schooling plus five is greater than the age of immigration. The share of immigrants that have any schooling in the host country varies substantially across the samples. This variation corresponds directly with differences in the average age of immigration.

The distribution of region of birth of immigrants is in Panel B of Table 2. The majority of immigrants in the NHIS are from Mexico or other areas of Central or South America (66% of male immigrants and 68% of female immigrants). In contrast, in the NIS sample of recent legal immigrants, more immigrants are from Asia than from Central and South America. The majority of immigrants in the U.K. were born in South Asia. Specific country or area of origin is not available for immigrants in the HRS.

[6]NHIS does not collect information on the precise time of arrival of the immigrant. The averages are constructed from the categories for time of arrival which are less than 1 year ago, from 1 to less than 5 years, 5 to less than 10 years, 10 to less than 15 years and over 15 years.

[7]The host country is the U.K. for the HSE sample and the U.S. for the other samples.

## 4 Immigrant and Native-Born Returns to Height

The basic framework to examine the impact of height on earnings is estimated using the following equation:

$$\log w_i = \alpha_0 + \alpha_1 H_i + \beta X_i + \epsilon_i \tag{14}$$

where $w_i$ is the wage of individual $i$, $H$ is height, $X$ is a vector of covariates and $\epsilon$ is an error term. The errors allow for clustering at the household level. The covariates included in $X$ vary by specifications. In the most parsimonious specification, $X$ includes a quadratic in age, indicators for region of residence in the U.S. or the U.K. and for year.

The results for the sample of native-born individuals are presented in column 1 of Table 3. The corresponding results over a sample of immigrants are in column 4, and the results from the NIS are in Table 4. Among native-born individuals, the coefficients suggest that an additional inch of height translates to a 1-2% increase in wages. The corresponding estimates for immigrants range between 2-4%. The returns to height for immigrants are 45-85% higher than the corresponding returns to height for native-born individuals.

The regressions in columns 2 and 5 also control for years of education. For men, while the returns to height decreases slightly with the inclusion of the additional control, the height premium for male immigrants is not eliminated. The gap remains such that each additional inch of height yields about twice more wage gains immigrants than for native-born individuals. In contrast, the returns to height for immigrant and native-born women converge to be quite similar in the NHIS data set. The large gap in the coefficient on height remains only for women in the HSE sample. This is consistent with some previous evidence in the literature that the returns to height are not as robust for women as for men. Selection of women out of the labor force is most likely driving the gender differences in the results.

Furthermore, the returns to education are generally lower for immigrants than for native-born individuals. These results are consistent with the prediction of the model of statistical discrimination where immigrant height is given more weight by employers because the signals of human capital for immigrants is observed by employers with error. The education signal for immigrants may be observed with less reliability for many reasons. The mapping between a foreign degree and the American or British system may be unclear to employers. The quality of the schools may be difficult to determine for immigrants than for native-born individuals. However, these results may be also be consistent with an alternative story in which the mapping between years of education and productivity in other countries is less steep due to lower quality schools.

Finally, columns 3 and 6 of Table 3 include one-digit industry and occupation fixed effects. By looking within job categories, we can evaluate the hypothesis that the height premium for immigrants is due to sorting into specific types of jobs where height has stronger effects on worker output. The results indicate that occupational sorting does not explain the higher returns to height for immigrant men over native-born men.

Table 4 displays the estimates for immigrant men and women in the NIS sample. The results for NIS women are similar to HRS immigrant women; the magnitude of the wage returns to height for women are small and not statistically different from zero in any of the specifications that include years of education. The returns to height for NIS men are slightly lower than the other immigrant samples in the parsimonious specifications, and the estimates in the full specification with industry and occupation fixed effects are quite similar to the American immigrant men in the NHIS and HRS.

# 5 Specification and Robustness Checks

## 5.1 Selection of Immigrants

This section considers the idea that the observed relationship between height and wages of immigrants is explained by heterogeneity in the selection process across immigrants. For example, there may be negative selection of illegal immigrants from Central America, where the average height is relatively low, and positive selection of immigrants from other areas due to immigration policies.[8] Under the assumption that selection effects vary across countries rather than within countries, a specification that includes country fixed effects should remove the effects of selection. Furthermore, this specification will also address other possible explanations that depend on differences in characteristics across countries of origin. The NIS and HSE include information on country of birth of immigrants, but the NHIS only has region of birth of immigrants. The HRS does not share any information about place of origin of immigrants, and is excluded from the analysis in this section.

The results are presented in Table 5. The odd columns correspond with the specification presented in column 5 in Table 3 and columns 2 and 5 of Table 4 with the addition of country (or region) fixed effects. The results displayed in the even columns include additional controls for country, industry, occupation and years in the U.S. or U.K. For American immigrants in the NHIS and the NIS, the

[8] For analysis on the determinants of determinants of negative or positive selection of immigrants, see Borjas (1987) and Rosenzweig and Jasso (1986).

inclusion of country fixed effects does not have much effect on the estimates of the returns to height and to education. For British immigrants, the inclusion of country fixed effects slightly decreases the returns to height for men but increases the returns to height for women. Overall though, the returns to height remain substantially higher than those of native-born Britons. Thus, the results suggest that the returns to height are not solely driven by differences across countries, but also hold when comparing tall and short immigrants from the same country.

### 5.2 Nonlinearities in the Returns to Height

The results presented in Section 4 assume that the relationship between height and the logarithm of wages is linear. This specification follows the standard in the bulk of the literature on the wage returns to height. Nonparametric estimates of the returns to height provide support for the linearity assumption (Strauss and Thomas 1998). However, given that immigrants are on average several inches shorter than native-born individuals, this assumption could be problematic for the analysis of this paper if the actual relationship between height and earnings is concave. This section demonstrates that the stronger relationship between height and wages for immigrants is not driven by the functional form of the estimating equation.

I examine two alternative specifications of the relationship between height and wages. First, I estimate the relationship with a quadratic in the height of the individual. Second, I include the logarithm of height rather than the level of height in inches. The results are presented in Table 6 and are comparable to the results in column 3 of Table 3. Columns 1-4 of Table 6 demonstrate that the returns to height are still almost twice as large for immigrant men than for native-born American under the quadratic specification (Panel A) and under the logarithmic specification (Panel B). This holds for both the NHIS and the HRS data for Americans as well as for the HSE data for Britons.

For women, the nonlinear estimates of the returns to height are similar to the linear estimates. Overall, the significance of the relationship between height and wages remains weaker for women. The NHIS and the HRS results do not support the idea that immigrant women in the U.S. have higher returns to height than American-born women. The HSE results suggest that immigrant women in Britain do experience greater increases in wages for each additional unit of height.

### 5.3 Measurement Error in Height

Another potential concern is that systematic differences in reporting error for height between immigrants and native born individuals could bias the coefficient estimates and generate the observed, larger returns to height for immigrants. While height in the NHIS and NIS are self-reported, height is measured by trained interviewers in the HSE. Given that the relationship between the ratio of the returns to height for immigrants and native-born individuals are similar for the HSE and the NHIS, it is unlikely that the larger returns to height for immigrants are explained by measurement error in height. Height is self-reported in the 1992 wave of the HRS used in this analysis, as well as in all subsequent waves; in 2006, height was measured by trained staff and the average reporting error was very low at around 1-2% with no significant differences by racial or ethnic subgroups (Meng, He and Dixon 2010).

A method for addressing systematic reporting error in height was suggested by Lee and Sepanski (1995) and Bound, Brown and Mathiowetz (1999). They use an independent source of data that contains both the true and the reported values of the variable. By estimating the true value of the variable as a function of its noisy reported value and other observable characteristics, one can derive a relationship between the reported and the true values. Assuming that the relationship between the reported and the measured values are the same in both data sets, the estimated relationship from the validation data can be used to calculate the true value of height from the reported value in the primary data set.

Respondents in the Third National Health and Nutrition Examination Survey (NHANES III) from the U.S. Department of Health and Human Services reported their own estimates of height and were professionally measured four weeks later. Using this data set to implement the correction for reporting error in height separately for immigrants and native-born individuals does not remove the large gap in the returns to height for immigrants and for native-born individuals in the NHIS and NIS.[9]

# 6 Testing for Statistical Discrimination

The following sections examine whether there is evidence that employers use height as a tool of statistical discrimination by testing whether changes in signal reliability alter the returns to height and to education in ways predicted by the model of statistical discrimination. If employers statistically discriminate based on immigrant height in the absence of high quality information on other characteristics that are available for native-born individuals, then the returns to the perfectly observable characteristic for immigrants should decline with improvements in other sources of information. Fur-

[9] I use the NHANES III rather than the HRS for this exercise because the age distribution of the NHANES III sample is more similar to the age distributions of the NHIS and NIS data. These results are available from the author upon request.

thermore, assuming that employers in the immigrant's country of origin have better signals of quality than host country employers, the effects of statistical discrimination on the returns to height and education should not be observed in pre-immigration wage data. In Appendix A, I examine another type of model of statistical discrimination that does not rely on differences in the quality of information signals but rather on differences in the priors that employers have about average productivity.

### 6.1 Cross-Sectional Variation in Signal Reliability

Over a sample of immigrants, I estimate the following equation:

$$logw_i = \beta_0 + \beta_1 H_i + \beta_2 H_i * Q_i + \beta_3 S_i + \beta_4 S_i * Q_i + \beta_5 Q_i + \beta_5 X_i + \epsilon_i \tag{15}$$

where $S$ is total years of schooling and $Q$ is a measure of signal quality. If signal quality is increasing in $Q$, the model of statistical discrimination predicts that $\beta_2 < 0$ and $\beta_4 > 0$. In other words, as the reliability of the signal of $S$ improves, employers place more weight on $S$ and less weight on the perfectly observable characteristic, $H$. This relies on plausible assumptions that height is observed perfectly by employers but $S$ is observed with more error for immigrants than for native-born individuals.

I consider two measures of $Q$. The first measure is years since immigration. As an immigrant spends more time in the host country, the quality of productivity signals is likely to improve. This may occur because communication becomes easier either through improved language ability or cultural assimilation, or because immigrants accumulate labor market experience in the host country that demonstrates their true level of human capital. However, years since immigration may capture variation in worker ability and productivity in addition to variation in signal reliability. Cultural assimilation or improved English language abilities may increase worker productivity directly in addition to reducing the noise in the signal of productivity. Furthermore, over time some immigrants chose to leave the host country and this selection may generate a correlation between years in the host country and individual ability. If high ability immigrants remain in the U.S. or if productivity increases directly with the amount of time in the host country, then we would expect $\beta_2 > 0$ and $\beta_4 > 0$. If selection is such that low ability immigrants are more likely to remain in the U.S., then we would expect $\beta_2 < 0$ and $\beta_4 < 0$

The second measure of $Q$ is an indicator for whether the immigrant completed any education in the host country.[10] The quality of the signal of human capital is plausibly improved when an immigrant attends school in the host country. For example, if an individual has a graduate degree

[10] The host country is the U.K. for the HSE sample or the U.S. for the other samples.

from an American university in addition to a foreign degree, the noise in the signal for employers is plausibly lower than if the individual had a similar graduate degree from an unfamiliar foreign university. However, as with the previous measure of $Q$, host country education may be correlated with individuals characteristics, such as ability, or reflect direct differences in productivity in addition to variation in information quality. If immigrants with host country education tend to have higher ability due to admissions policies and immigration rules, or if productivity directly improves as the result of any education in the host country, then $\beta_2 > 0$ and $\beta_4 > 0$.

The results are presented in Table 7 for male immigrants and Table 8 for female immigrants. For men, the evidence generally suggests that the returns to education increase over time in the U.S. or U.K. This is consistent with the idea that information about education is improving as the immigrant remains in the U.S. However, years since immigration generally has a positive effect on the returns to height rather than the negative effect predicted by the model of statistical discrimination. In fact, the effect for each additional decade in the host country is extremely small in magnitude and not statistically different from zero. The results in the even columns where $Q$ also reject the predictions of statistical discrimination. The magnitude and significance of the estimates of the interaction between height and education in the host country suggest that there is no impact of host country education on the returns to height. Overall, there is not strong support for the hypothesis of statistical discrimination by employers against immigrants. The results are also not consistent with a combination of statistical discrimination and a positive correlation between $Q$ and productivity. This scenario would suggest that the interaction of $Q$ and years of education to be strongly positive.

The results for female immigrants displayed in Table 8 are somewhat different from the results for men. The coefficients on $\beta_2$ and $\beta_4$ are mostly consistent with the model of statistical discrimination when $Q$ is years since immigration. However, in the results in which $Q$ is host country education, the sign of the coefficients support the idea that $Q$ is positively correlated with ability. However, the coefficients are rarely significant at standard levels.

### 6.2 Variation in Signal Reliability and Panel Data

The NIS asks retrospective information on the labor market experiences of immigrants in the year that they immigrated to the U.S. Assuming that the reliability of the signal of human capital is lower for employers in the host country than for employers in the country of origin, pre-immigration labor market information offers another test of the model of statistical discrimination.

Over a sample that pools pre- and post-immigration labor market experiences of individuals in the NIS, I estimate the following equation:

$$logw_{it} = \gamma_0 + \gamma_1 H_i + \gamma_2 H_i * PreImmig_{it} + \gamma_3 S_{it} + \gamma_4 S_{it} * PreImmig_{it} + \gamma_5 X_{it} + \upsilon_{it} \tag{16}$$

where $PreImmig$ is an indicator that equals one if the data refer to a period prior to immigration to the U.S., and $X$ includes a quadratic in age, and indicators for country of origin and year. The panel data set includes two observations for every individual, one observation prior to immigration and one observation after immigration.[11] Age and years of education are adjusted appropriately in the pre-immigration data.[12] While the returns to height and education may vary in different countries, I include country fixed effects so the key estimates of interest, $\gamma_2$ and $\gamma_4$, yield the difference between the pre- and post-immigration wage returns of individuals originating from the same country.

The key assumption of equation 16 is that employers in the immigrants' country of origin observe signals of productivity that are less noisy than the signals observed by American employers. Statistical discrimination based on the observable characteristic height by American employers would yield $\gamma_2 < 0$ and $\gamma_4 > 0$. If employer statistical discrimination on height occurs in the absence of other reliable sources of information, then we expect that employers' reliance on height to be less strong for immigrants in their country of origin than in the U.S. In other words, the wage returns to education are higher prior to immigration when the signal is clearer. The weight placed on height is lower given the availability of other information on productivity.

The NIS pseudo-panel data offers additional predictions based on the measures of signal quality, $Q$, discussed in the previous section. I estimate the following regression:

$$\begin{aligned} logw_{it} = \gamma_0 + \gamma_1 H_i + \gamma_2 PreImmig_{it} * H_i + \gamma_3 S_{it} + \gamma_4 PreImmig_{it} * S_{it} + \\ \gamma_5 PreImmig + \gamma_6 H_i * Q_i + \gamma_7 S_{it} * Q_i + \gamma_8 PreImmig_{it} * Q_i + \\ \gamma_9 H_i * PreImmig_{it} * Q_i + \gamma_{10} S_{it} * PreImmig_{it} * Q_i + \gamma_{11} Q_i + \gamma_{12} X_{it} + \upsilon_{it} \end{aligned} \tag{17}$$

where $Q$ is measured as years since immigration to the U.S. divided by ten or whether the individual has any education in the U.S. The measures of $Q$ are time-invariant in this equation to allow us to determine whether $Q$ is measuring post-immigration statistical discrimination or time-invariant unobservable ability. The post-immigration interactions of height and $Q$ would be as previously

[11] One of the key limitations of the panel results is that the sample in this section only includes a selected group of individuals that worked both before and after immigration. For example, individuals that immigrate to the U.S. for education and never worked in their origin country would not be included in this analysis.

[12] The raw NIS pre-immigration wage data are converted into real 2004 local currency using the Penn World Tables, and then converted in the 2004 U.S. dollars using OANDA exchange rate data.

discussed ($\gamma_6 < 0$ and $\gamma_7 > 0$) because as the signal of education improves less weight is placed on height and more on education. Furthermore, under statistical discrimination, the net effect of pre-immigration interactions should be zero ($\gamma_6 + \gamma_9 = 0$ and $\gamma_7 + \gamma_{10} = 0$) because subsequent American education or tenure in the U.S. should not affect signal reliability *before* immigration. In contrast, if the effect of $Q$ is driven by a correlation with unobserved ability, we should see positive returns to the interactions of $Q$ with height and education both before and after immigration ($\gamma_6 > 0$, $\gamma_7 > 0$, $\gamma_6 + \gamma_9 > 0$ and $\gamma_7 + \gamma_{10} > 0$).

The results of equations 16 and 17 are presented in Table 9. Columns 1 and 4 corresponds to equation 16 for men and women, respectively. The signs on the interactions are opposite to the predictions of statistical discrimination for men, and they are both positive for women. The estimated signs are not consistent with statistical discrimination for either men or women. However, we cannot statistically reject the hypothesis because none of the estimated interactions are significantly different from zero at the 10% level on their own or jointly.

Columns 2 and 5 present the results where $Q$ is the amount of time that the immigrant has spent in the U.S. (divided by 10). For men, $\gamma_6 > 0$ and $\gamma_7 < 0$ which is not consistent with either statistical discrimination or $Q$ reflecting ability, but these estimates are not significantly at the standard levels. However, we can reject the prediction of the model of statistical discrimination that $\gamma_6 + \gamma_9 = 0$ at the 1% level. For women, the results indicate that the post-immigration returns to height is decreasing in years in the U.S. while the post-immigration returns to education are increasing in years after immigration. However, the standard errors are very large and the pre-immigration returns to education and height are both negative. Finally, the results where $Q$ is a dummy variable for American education is displayed in columns 3 and 6 of Table 9. The two key predictions of the model of statistical discrimination are rejected more at the 5% level for men. For women, the estimates are too noisy to be conclusive but the signs of the coefficients are not supportive of either the hypothesis of statistical discrimination or $Q$ reflecting unobserved ability.

Overall, the results do not support the model of statistical discrimination using height given variation in signal reliability across groups for men. The evidence is weaker for immigrant women in that the predictions cannot be rejected statistically at standard levels. Appendix A also shows that there is no evidence for statistical discrimination using height where there is no differences in the reliability of the signal but there are differences in average outcomes across groups. The next section explores an alternative explanation for the robust empirical finding that the wage returns to height are much higher for immigrants than for native-born individuals.

# 7 Productivity Differences in the Height Signal

The evidence in the previous sections suggests that statistical discrimination cannot explain why immigrants experience higher wage returns to height than native-born individuals. The following sections consider the alternative hypothesis that the slope of the relationship between height and productivity differs between immigrants and native born individuals. The previous literature has demonstrated evidence for the linkage between height and health (Strauss and Thomas 1998, Steckel 1995), cognitive skills (Case and Paxson 2008) and non-cognitive skills (Persico, Postelwaite and Silverman 2004). It is plausible that the larger impact that each additional unit of height has on immigrant wages over native-born wages results from non-linearities in the mapping between nutritional inputs and health and cognitive development. For example, the returns to increasing investment in health and nutrition can have higher returns in both height and productivity at low levels of investment. I test this hypothesis in two ways. First, I examine whether the higher returns to height for immigrants are driven by immigrants from poorer regions of the world. Second, I directly test whether height is more correlated with measures of productivity for immigrants than for native-born individuals.

## 7.1 Returns to Height by Income of Country of Origin

First, I examine whether the returns to height for immigrants vary by the average income of their country of origin. The following wage regression is implemented over a sample of immigrants:

$$\log w_{ij} = \alpha_0 + \alpha_1 H_{ij} + \sum_{k=2}^{4} \alpha_k GDPN_{j\in k} * H_{ij} + \beta X_{ij} + \gamma_j + \epsilon_{ij} \tag{18}$$

where $GDPN_{k\in j}$ is an indicator variable for whether the real per capita GDP of the individual's country of origin $j$ is in quartile $k$ in the year of immigration across all immigrants in the sample.[13] The specification includes country fixed effects, $\gamma_j$. The estimate of $\alpha_1$ yields the within-country returns to height for immigrants from countries in poorest quartile of the immigrant sample. The estimate of $\alpha_k$ indicates whether the within-country returns to height for immigrants from countries in the $k$th poorest quartile are different from those in the poorest quartile.

If the difference in the relationship between height and productivity for immigrants and native-born Americans and Britons is driven by higher productivity returns to nutritional and health inputs at low levels of investment, then we expect the wage returns to height to be largest for immigrants from poor countries relative to others from the same country. In other words, the productivity hypothesis

[13] Data on real GDP per capita in the country of origin across years is the Penn World Tables Laspeyres series with a reference year of 1996.

suggests that the coefficient estimate of $\alpha_1$ to be positive and large, and the coefficient estimates of $\alpha_k$ to be negative and decreasing in $k$. This is a weak test of the productivity hypothesis. If the described pattern in the coefficients is not observed, then the productivity hypothesis is rejected; however, if the pattern in the coefficients is observed, the results are consistent with the productivity story but also consistent with a model of statistical discrimination if the reliability of the signal of height is decreasing in immigrants' country of origin.[14]

These equations are estimated using the NIS and HSE samples which contain information on the specific country of origin of immigrants. The distribution of the immigrants' origins are quite different across these samples (see Appendix C of Table 2); thus, it is not surprising that the distribution of GDP per capita is very different across the samples. The quartiles are constructed within the NIS and HSE so the categories refer to different levels of GDP per capita for the samples.[15] I also implement a specification in which the interactions with the sample-specific quartiles are replaced by an interaction between height and an indicator for whether the GDP per capita of the country of origin is over USD$1600. This allows for direct comparison across the NIS and HSE samples.

Table 10 displays the results. In all of the specifications, the estimated coefficient on height is positive and large in magnitude. In the estimation of equation 18 presented in odd columns, the coefficient estimates on the interactions are generally negative. Furthermore, the coefficients corresponding to the fourth quartile are significantly different the reference category at the 10% level or higher in all four specifications and lower in magnitude than the coefficients corresponding to the second and third quartile. Similarly, the results presented in the even columns confirm the estimates of equation 18; the wage gains are substantially smaller for immigrants from countries where the GDP per capita is greater than $1600 than for immigrants where the GDP per capita is less than $1600. The gap in the wages associated to a one-inch difference in height for two male immigrants in the U.S. who are from a poor country like Ethiopia will be 2.5 to 3% but the corresponding gap would only be around 1.5% for two male immigrants from a rich country like the U.K.

These results demonstrate that the within-country slope of the relationship between height and productivity is decreasing in the level of development of immigrants' country of origin. Thus, the empirical results is consistent with the hypothesis that the larger wage returns to height for immigrants are explained by a different relationship between height and productivity for immigrants than

[14]A pattern of an inverse relationship between the magnitude of the returns to height and the level of development of the country of origin is necessary but not sufficient support for the productivity hypothesis. While the pattern is consistent with statistical discrimination, it is neither necessary nor sufficient.

[15]The cutoffs for the quartiles for the HSE are USD$1252, $1502 and $1908. In the NIS, they are $1872, $5031 and $10,545.

for native-born individuals. However, as previously mentioned, these results are necessary but not sufficient evidence for the productivity hypothesis because they can also be explained by the mechanism of statistical discrimination under some assumptions. The next section presents a stronger test of the productivity hypothesis.

### 7.2 Height and Direct Measures of Ability

In the second test, I directly examine whether height is more correlated with measures of productivity for immigrants than for native-born individuals. This hypothesis is reflected in equation 4 of the theoretical framework and tested with the following regression over a sample that includes both immigrants and native-born individuals :

$$P_i = \beta_0 + \beta_1 H_i + \beta_2 H_i * I_i + \beta_3 I_i + \beta_4 X_i + \epsilon_i \tag{19}$$

where $I_i$ is an indicator that equals 1 if individual $i$ is an immigrant. The dependent variable, $P$, is health status or cognitive ability. If the gap in the returns to height reflect differences in the relationship between height and productivity for immigrants and for native-born individuals, then we expect the coefficients $\beta_1$ and $\beta_2$ to have the same sign and the magnitude of $\beta_2$ relative to $\beta_1$ to be similar to the gap in the returns to height for immigrants relative to native-born individuals displayed in Table 3.

The results are presented in Table 11. In the first three columns, the dependent variable is individuals' self-reported health status where 1 refers to excellent health and 5 poor health. For both men (in Panel A) and women (in Panel B) in all three samples, taller individuals are also healthier, and these estimates are significant at the 1% level. Furthermore, the evidence suggests that each additional inch of height corresponds to a larger improvement in health for immigrants than for native-born individuals. The effects are strongest in the HSE sample where a ten inch change in height corresponds with one-quarter of a standard deviation of better health for native-born men and women and with one-half of a standard deviation of better health for immigrant men and women. The results for the other samples are generally not statistically significant, but in the majority of cases the sign of coefficient indicates that height is more strongly correlated with health for immigrants than for native-born individuals.

The last three columns of Table 11 correspond to equation 19 with the dependent variable as a measure of cognitive ability. Of the main data sets used in this analysis, only HRS implements a

WAIS test, which is an IQ test where the score is increasing in cognitive ability.[16] I supplement the analysis with data from the Third National Health and Nutrition Examination Survey (NHANES III), which contains information on immigration status, height and several measures of cognitive ability.[17] The symbol-digit substitution test (SDST) is one of the tests included in the WAIS and measures coding speed. Individuals are presented with pairings of digits and symbols and are asked to enter the corresponding digit for a series of the symbols as quickly as possible. Five trials were conducted and the score used is the error-corrected speed. A lower value corresponds to faster responses and higher cognition. In addition, the NHANES includes a serial digit learning test (SDLT), which measures learning and recall. Individuals are presented with a sequence of digits. Afterwards, the individual is asked to enter the entire sequence of numbers in the order presented. A smaller number represents fewer mistakes and higher cognition.

The results demonstrate that for all three measures, taller men and women also have higher cognitive ability. This is consistent with the results of Case and Paxson (2008). This analysis also indicates that the correlation between height and cognition is stronger for immigrants than for native-born individuals. The difference is statistically large in magnitude and significant for the NHANES sample but not statistically significant at the 10% level for the HRS samples. The NHANES results suggest that each additional inch of height corresponds to more than twice as large an increase in cognition for immigrants than for native-born individuals. Overall, the results provide evidence in support of the hypothesis that the greater wage returns to height experienced by immigrants reflects a stronger mapping between height and productivity.

# 8 Conclusion

Using several different data sets, this paper presents a very robust empirical finding that the returns to height are much larger for immigrant men in the U.S. and the U.K. than they are for native-born men in those countries. The theoretical framework demonstrates that this finding is consistent with two hypotheses. First, it is consistent with a model of statistical discrimination whereby employers weigh a characteristic that is perfectly observable, height, more for immigrants than for native-born individuals because other signals of productivity are not reliable for immigrants. Second, the baseline results are also consistent with a model in which the mapping between productivity and height is

[16] More details about the WAIS are discussed in Appendix A.

[17] The NHANES III spans 1988-1994 and was designed to obtain nationally representative information on health and nutrition of individuals in the U.S. This data isn't used in the other analyses of the paper because it lacks information on the income of respondents.

different for immigrants.

The empirical evidence suggests that there is a stronger relationship between height and unobserved components of productivity, including health and cognitive ability, for immigrants than for native-born Americans or Britons. This suggests a concave relationship between health and nutritional inputs during early life and long-run outcomes such as adult height and productivity. This research contributes to two strands of the large and growing economic literature on height. One strand of the literature uses height as an outcome to compare individuals within countries as well as across countries. Another strand of the literature examines height as an input into an individual-level production function.

In addition, this paper contributes to the literature that tests for employer statistical discrimination. The paper is the first to present an empirical analysis that focuses on height. Given how it is as easy to observe as race and gender, this physical characteristic is simple for employers to use. The distinction between immigrants and native-born individuals presents plausible groups for whom there is a discrepancy in the reliability of other signals of productivity, such as education. While the results suggest that height offers information about productivity that is otherwise not directly observed, the empirical evidence indicates that employers do not use height as a tool of statistical discrimination. This finding is similar to previous results that suggest that employers do not use race to statistically discriminate among workers despite the differences in average outcomes by race (Altonji and Pierret 2001).

These results have important implications for our understanding of the immigration decisions of individuals as well as the process of assimilation of immigrants. The empirical findings of this paper do not support previous theoretical hypotheses that the anticipation of statistical discrimination may influence migration and human capital decisions of individuals in developing countries. Furthermore, statistical discrimination on the basis of height does not play a role in the convergence over time wages among immigrants in the U.S. or U.K.

# A Average Differences Across Groups

This section steps outside of the focus of the conceptual framework on the reliability of information and considers whether there is other evidence for statistical discrimination based on height. This section implements an empirical test developed by Altonji and Pierret (2001), hereafter referred to as AP, that considers differences in employers' priors on the average productivity of two groups. The model of AP does not assume or allow for differences in the noise-to-signal ratio for the two groups.

Their hypothesis of employer learning with statistical discrimination (EL-SD) posits that firms rely on observable correlates of productivity, such as gender, race and education, to distinguish the quality of workers in early stages of their careers. Over time, employers learn about worker productivity directly and the initial information becomes redundant.

Following AP, I implement the following regressions:

$$logw_{it} = \beta_s s_i + \beta_z z_i + G(t_i) + \beta_{es} E_{it} * s_i + \beta_x X_{it} + \epsilon_{it} \tag{20}$$

$$logw_{it} = \beta_s s_i + \beta_z z_i + G(t_i) + \beta_{es} E_{it} * s_i + \beta_{ez} E_{it} * z_i + \beta_x X_{it} + \eta_{it} \tag{21}$$

where $z_i$ is a correlate of productivity of person $i$ that is observed by the econometrician but not by the employer. $s_i$ is an easy-to-observe correlate of productivity; AP consider race and years of schooling. $G(t_i)$ represents the experience profile of the individual's productivity. In these specifications, $G(t_i)$ includes a cubic in potential experience.[18] Experience, denoted by $E$, equals potential experience divided by ten. This normalization represents the change in wages slope associated with ten additional years of experience.

Employers have limited information about new entrants into the labor force and statistically discriminate based on easy-to-observe variables such as education. As employees gain labor market experience, wages become more strongly related with variables that are correlated with productivity but are hard-to-observe. This implies that one prediction of EL-SD is that $\beta_{ez} > 0$ in equation 21. A stronger prediction of EL-SD is that the coefficient on the interaction between experience and education, $\beta_{es}$, will be greater when estimated by equation 20 than when estimated by equation 21. In other words, when conditioning the experience profile of earnings on *both* the easy- and the hard-to-observe variable leads to the partial effect of the easy-to-observe variable declining with experience. The main intuition is that because $s$ and $z$ are correlated with each other, $b_{es}$ captures the experience-earnings profile of $z$ as well as $s$ in equation 20; however, when the interaction, $E_{it} * z_i$, is included in equation 21, the experience-earnings profile of $s$ declines because $s$ is known at the time of hire and $z$ component is learned over time.[19]

I estimate equations 20 and 21 using the HRS and the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). The NLSY79 is a panel survey covering a nationally representative sample of 12,686 Americans who were between the ages of 14 and 22 in 1979. Used in the empirical analysis of AP, the NLSY79 contains key information on hard-to-observe cognitive ability in the form of the

[18] Potential experience = age - years of education - 5.

[19] See AP for the technical details.

Armed Forces Qualification Test (AFQT), administered to respondents in 1981. While AP used data up through 1992, the analysis in this paper uses the waves through 2002. I use the NLSY79 to examine whether there is evidence that employers use height as a tool of statistical discrimination in the general U.S. population without distinguishing between immigrants and native-born individuals.[20] I am able to examine the predictions separately for immigrants and native-born individuals in the HRS.

For each sample, Appendix Table 12 presents three regressions. First, following AP, I estimate equation 20 with $s$ as years of schooling. This first regression allows for direct comparison with the original EL-SD results based on education. Second, I estimate equation 20 with two easy-to-observe variables, height and years of schooling. Finally, I estimate equation 21. In the NLSY results for men, the coefficient on $Ability * Exp/10$ is positive and large and the coefficient in column 3. Furthermore, $Education * Exp/10$ falls significantly with the inclusion of $Ability * Exp/10$ (comparing column 1 or 2 with column 3). This confirms the prior results of AP on statistical discrimination based on education. In contrast to education, for NLSY men, the coefficient on $Height * Exp/10$ *rises* with the inclusion of $Ability * Exp/10$. This suggests that the wage returns to height increases with experience, and that employers do not use height in their initial assessment of worker productivity.

Of the main data sets used in this analysis, only the HRS has a direct measure of cognitive ability of adults. HRS adults are administered the Wechsler Adult Intelligence Scale (WAIS) test, which is the primary instrument used to measure the intelligence quotient (IQ) of adults and adolescents.[21] A higher score of the test corresponds to higher IQ. The results for native-born and immigrant men in the HRS are very similar to the NLSY. The estimates of $Ability * Exp/10$ suggest that the earnings profile of ability is increasing with experience. The coefficient on $Education * Exp/10$ drops with the inclusion of $Ability * Exp/10$, but the coefficient on $Height * Exp/10$ remains the same. While the pattern in the coefficients is similar for the HRS immigrant sample, the coefficients are not statistically significant at the 10% level. This may be explained by the dramatically smaller sample size. If pre-migration employer learning is lost at the time of immigration, another explanation for the weakness of the results for immigrants may be that the measure of potential experience as all years since individuals finished school is inappropriate for this group.

While the results for women do not support the idea that employer statistically discriminate on the basis of education, they also do not support statistical discrimination based on height. Overall, the results confirm the findings of statistical discrimination by education for men. The main hypothesis of interest in this paper regarding statistical discrimination based on height is rejected. While there

[20]The publicly available version of the NLSY79 does not identify immigrant status. Moreover, I do not use the NLSY79 for any comparisons of native-born and immigrants because the sample only include a couple of hundred immigrants.

[21]The WAIS covers verbal comprehension, memory, perceptual organization and processing speed.

may be information about productivity contained in height, employers do not use this to determine the quality of workers. Similarly, AP find that employers do not statistically discriminate on the basis of race; while race is correlated with productivity and easy-to-observe, the empirical results testing EL-SD find that it behaves more similarly to an unobserved characteristic.

## References


[1] Aigner, Dennis and Glen Cain, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 1977.

[2] Altonji, Joseph and Charles Pierret, "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 2001.

[3] Averett, Susan and Sanders Korenman, "Economic Reality of the Beauty Myth," *Journal of Human Resources*, 1996.

[4] Bound, John, Charles Brown and Nancy Mathiowetz, "Measurement Error in Survey Data," *Handbook of Econometrics*, volume 5, eds. James Heckman and Ed Leamer. New York: Springer-Verlag, 2002.

[5] Borjas, George, "Self-Selection and the Earnings of Immigrants," *American Economic Review*, 1987.

[6] Bozzoli, Carlos, Angus Deaton and Climent Quintana-Domeque, "Adult Height and Childhood Disease," *Demongraphy*, 2009.

[7] Case, Anne and Christine Paxson, "Stature and Status: Height, Ability, and Labor Makert Outcomes," *Journal of Political Economy*, 116(2008).

[8] Chau, Nancy and Oded Stark, "Human-Capital Formation, Asymmetric Information, and the Dynamics of International Migration," in *Economics of Globalization*, eds Assaf Razin and Efraim Sadka, 1999.

[9] Coate, Stephen and Glen Loury, "Will Affirmative Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 1993.

[10] Deaton, Angus, "Height, Health and Inequality: the Distribution of Adult Heights in India," *American Economic Review*, 2008.

[11] Farber, Henry and Robert Gibbons, "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 1996.

[12] Foster, Andrew and Mark Rosenzweig, "Information, Learning and Wage Rates in Low-Income Rural Areas," *Journal of Human Resources*, 1993.

[13] Glick, Peter and David Sahn, "Health and Productivity in a Hetergeneous Urban Labor Market," *Applied Economics*, 1998.

[14] Hamermesh, Daniel and Jeffrey Biddle, "Beauty and the Labor Market," *American Economic Review*, 1994.

[15] Hersch, Joni, "Profiling the New Immigrant Worker: The Effects of Skin Color and Height," *Journal of Labor Economics*, 2008.

[16] Kuhn, Peter and Kailing Shen, "Employers' Preferences for Gender, Age, Height and Beauty: Direct Evidence," NBER Working Paper 15564, 2009.

[17] Kwok, Viem and Hayne Leland, "An Economic Model of the Brain Drain," *American Economic Review*, 1982.

[18] Lee, Lung-Fei and J. Sepanski, "Estimation of Linear and Nonlinear Errors in Variables Models Using Validation Data," *Journal of the American Statistical Association*, 1995.

[19] Lundberg, Shelley and Richard Startz, "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review*, 1983.

[20] Mankiw, Gregory and Mathew Weinzierl, "Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution," NBER working paper 14976, 2009.

[21] Meng, Hondgao, Xiaoxing He, and Denise Dixon, "Self-Reported Versus Measured Height and Weight in the Health and Retirement Study," *Journal of American Geriatrics Society*, 2010.

[22] Oettinger, Gerald, "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap," *Journal of Labor Economics*, 1996.

[23] Persico, Nicola, Andrew Postlewaite and Daniel Silverman, "Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height," *Journal of Political Economy*, 2004.

[24] Phelps, Edmund, "Statistical Theory of Racism and Sexism," *American Economic Review*, 1972.

[25] Steckel, Richard, "Stature and the Standard of Living," *Journal of Economic Literature*, 1995.

[26] Strauss, John and Duncan Thomas, "Health, Nutrition and Economic Development," *Journal of Economic Literature*, 1998.

Table 1: Summary Statistics

| | NHIS | | HSE | | HRS | | NIS |
|---|---|---|---|---|---|---|---|
| | Native | Immigrant | Native | Immigrant | Native | Immigrant | Immigrant |
| **Panel A: Men** | | | | | | | |
| Height (Inches) | 70.35 | 68.12 | 69.23 | 67.58 | 70.10 | 67.95 | 68.29 |
| | (2.66) | (2.84) | (2.66) | (2.76) | (2.70) | (3.34) | (3.33) |
| Age | 39.40 | 37.72 | 40.34 | 40.17 | 45.04 | 46.77 | 36.11 |
| | (10.87) | (10.03) | (10.18) | (9.58) | (11.20) | (9.63) | (8.90) |
| Earnings | 43399 | 33522 | 26209 | 24861 | 24171 | 25378 | 36163 |
| | (25423) | (23705) | (21078) | (23884) | (23242) | (55128) | (49761) |
| White Collar | 0.34 | 0.24 | 0.42 | 0.41 | 0.33 | 0.26 | 0.33 |
| | (0.47) | (0.43) | (0.49) | (0.49) | (0.47) | (0.44) | (0.47) |
| Education | 13.79 | 12.42 | 11.99 | 12.92 | 12.57 | 10.76 | 14.02 |
| | (2.24) | (3.22) | (2.13) | (2.92) | (2.94) | (5.05) | (4.58) |
| White | 0.84 | 0.68 | 0.93 | 0.21 | 0.83 | 0.64 | 0.52 |
| | (0.36) | (0.47) | (0.25) | (0.41) | (0.37) | (0.48) | (0.50) |
| Health Status | 1.96 | 2.01 | 1.79 | 1.91 | 2.58 | 2.52 | 1.85 |
| | (0.90) | (0.94) | (0.78) | (0.83) | (1.03) | (1.05) | (0.91) |
| Observations | 44512 | 10555 | 8534 | 2224 | 9726 | 934 | 1989 |
| **Panel B: Women** | | | | | | | |
| Height (Inches) | 64.61 | 63.46 | 63.99 | 62.87 | 64.53 | 63.17 | 63.44 |
| | (2.55) | (2.43) | (2.45) | (2.71) | (2.56) | (2.63) | (3.06) |
| Age | 39.30 | 38.54 | 39.83 | 40.87 | 45.20 | 45.94 | 35.54 |
| | (11.01) | (9.98) | (10.14) | (9.79) | (9.90) | (9.09) | (9.19) |
| Earnings | 30547 | 25155 | 18574 | 20870 | 14535 | 13626 | 22995 |
| | (21062) | (19911) | (17461) | (22238) | (23818) | (12564) | (29829) |
| White Collar | 0.43 | 0.31 | 0.32 | 0.39 | 0.31 | 0.18 | 0.28 |
| | (0.50) | (0.46) | (0.47) | (0.49) | (0.46) | (0.46) | (0.45) |
| Education | 13.91 | 12.80 | 11.95 | 12.81 | 12.77 | 10.83 | 13.60 |
| | (2.13) | (2.98) | (1.99) | (2.69) | (2.43) | (4.32) | (4.48) |
| White | 0.79 | 0.64 | 0.91 | 0.29 | 0.78 | 0.68 | 0.52 |
| | (0.41) | (0.48) | (0.28) | (0.45) | (0.41) | (0..46) | (0.50) |
| Health Status | 2.02 | 2.09 | 1.84 | 2.00 | 2.45 | 2.68 | 2.03 |
| | (0.92) | (0.97) | (0.82) | (0.86) | (1.01) | (1.13) | (0.94) |
| Observations | 44487 | 8531 | 8765 | 1859 | 10543 | 1071 | 1347 |

Notes: Standard deviations in parentheses. Earnings are displayed in real US dollars for the NHIS and NIS samples and in real British sterling for the HSE sample. HS degree is a dummy for completion of secondary education.

Table 2: Summary Statistics of Immigrants

| | NHIS | | HSE | | HRS | | NIS | |
|---|---|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female | Male | Female |
| **Panel A: Mean Characteristics** | | | | | | | | |
| Age of Immigration | 20.2 | 20.2 | 19.1 | 19.7 | 29.6 | 26.9 | 28.9 | 29.9 |
| | (9.34) | (9.54) | (9.92) | (9.93) | (11.7) | (12.0) | (10.8) | (11.6) |
| Years in U.S. or U.K. | 17.7 | 17.9 | 20.7 | 20.0 | 18.6 | 19.7 | 7.6 | 6.5 |
| | (11.3) | (11.3) | (12.9) | (12.1) | (11.9) | (11.7) | (8.2) | (6.5) |
| Host Country Education | 0.42 | 0.48 | 0.40 | 0.41 | 0.14 | 0.13 | 0.22 | 0.24 |
| | (0.49) | (0.50) | (0.49) | (0.49) | (0.37) | (0.36) | (0.42) | (0.43) |
| **Panel B: Distribution of Region of Origin** | | | | | | | | |
| Central & South America | 66.1 | 68.0 | | | | | 24.7 | 23.7 |
| Europe & Central Asia | 10.1 | 10.4 | 5.9 | 5.4 | | | 17.4 | 16.9 |
| Africa & Middle East | 5.9 | 4.2 | 16.0 | 17.0 | | | 14.4 | 9.8 |
| Asia | 15.3 | 15.0 | 62.1 | 63.2 | | | 32.7 | 35.5 |
| Other | 2.5 | 2.5 | 15.0 | 14.4 | | | 10.8 | 14.2 |
| **Panel C: Immigrant Height by Region of Origin** | | | | | | | | |
| Central & South America | 67.7 | 63.2 | | | | | 67.5 | 63.2 |
| | (2.72) | (2.32) | | | | | (3.40) | (3.16) |
| Europe & Central Asia | 70.2 | 64.6 | 68.3 | 63.7 | | | 70.2 | 65.1 |
| | (2.75) | (2.47) | (2.44) | (2.32) | | | (2.95) | (2.70) |
| Africa & Middle East | 69.4 | 64.3 | 67.9 | 63.0 | | | 69.0 | 64.5 |
| | (2.74) | (2.38) | (2.70) | (2.75) | | | (3.79) | (3.25) |
| Asia | 67.8 | 62.8 | 66.9 | 61.6 | | | 67.4 | 62.4 |
| | (2.55) | (2.26) | (2.72) | (2.57) | | | (2.89) | (2.54) |
| Other | 70.3 | 64.5 | 67.9 | 62.9 | | | 68.1 | 63.0 |
| | (2.85) | (2.70) | (2.88) | (2.65) | | | (3.53) | (3.58) |

Notes: In panel A, standard deviations in parentheses. For the U.S. data, "Other" is comprised mainly of Canada and Oceania. For the British data, "Other" is comprised mainly of North America.

Table 3: Baseline Returns to Height for Natives and Immigrants

| | Native Born | | | Immigrant | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: U.S. Men (NHIS)** | | | | | | |
| Height | 0.017 | 0.009 | 0.007 | 0.045 | 0.019 | 0.014 |
| | [0.001]** | [0.001]** | [0.001]** | [0.002]** | [0.002]** | [0.002]** |
| Education | | 0.087 | 0.064 | | 0.083 | 0.048 |
| | | [0.001]** | [0.002]** | | [0.002]** | [0.003]** |
| Observations | 42086 | 42051 | 41562 | 10064 | 10014 | 9843 |
| Adjusted $R^2$ | 0.14 | 0.23 | 0.30 | 0.12 | 0.28 | 0.36 |
| **Panel B: U.K. Men (HSE)** | | | | | | |
| Height | 0.029 | 0.021 | 0.012 | 0.041 | 0.035 | 0.028 |
| | [0.003]** | [0.003]** | [0.003]** | [0.007]** | [0.007]** | [0.007]** |
| Education | | 0.081 | 0.036 | | 0.047 | 0.009 |
| | | [0.005]** | [0.005]** | | [0.008]** | [0.007] |
| Observations | 6827 | 6712 | 6240 | 2005 | 1961 | 1746 |
| Adjusted $R^2$ | 0.08 | 0.15 | 0.28 | 0.05 | 0.11 | 0.27 |
| **Panel C: U.S. Men (HRS)** | | | | | | |
| Height | 0.026 | 0.013 | 0.009 | 0.041 | 0.022 | 0.014 |
| | [0.003]** | [0.003]** | [0.003]** | [0.010]** | [0.009]* | [0.008]+ |
| Education | | 0.072 | 0.065 | | 0.057 | 0.046 |
| | | [0.003]** | [0.003]** | | [0.006]** | [0.007]** |
| Observations | 9689 | 9689 | 9096 | 945 | 945 | 897 |
| Adjusted $R^2$ | 0.05 | 0.13 | 0.16 | 0.17 | 0.26 | 0.34 |
| **Panel D: U.S. Women (NHIS)** | | | | | | |
| Height | 0.020 | 0.009 | 0.008 | 0.027 | 0.008 | 0.006 |
| | [0.001]** | [0.001]** | [0.001]** | [0.004]** | [0.003]* | [0.003]* |
| Education | | 0.114 | 0.082 | | 0.108 | 0.063 |
| | | [0.001]** | [0.002]** | | [0.003]** | [0.003]** |
| Observations | 44428 | 44393 | 43831 | 7718 | 7700 | 7555 |
| Adjusted $R^2$ | 0.09 | 0.22 | 0.28 | 0.05 | 0.26 | 0.34 |
| **Panel E: U.K. Women (HSE)** | | | | | | |
| Height | 0.021 | 0.013 | 0.007 | 0.034 | 0.027 | 0.021 |
| | [0.004]** | [0.004]** | [0.004]* | [0.008]** | [0.008]** | [0.009]* |
| Education | | 0.114 | 0.056 | | 0.054 | 0.007 |
| | | [0.006]** | [0.006]** | | [0.007]** | [0.009] |
| Observations | 7103 | 6999 | 5926 | 1765 | 1721 | 1393 |
| Adjusted $R^2$ | 0.06 | 0.14 | 0.28 | 0.03 | 0.09 | 0.21 |
| **Panel F: U.S. Women (HRS)** | | | | | | |
| Height | 0.023 | 0.011 | 0.011 | 0.012 | -0.002 | 0.001 |
| | [0.003]** | [0.003]** | [0.003]** | [0.008] | [0.008] | [0.008] |
| Education | | 0.092 | 0.074 | | 0.049 | 0.032 |
| | | [0.003]** | [0.003]** | | [0.006]** | [0.007]** |
| Observations | 10606 | 10606 | 10117 | 1076 | 1076 | 1023 |
| Adjusted $R^2$ | 0.1 | 0.21 | 0.25 | 0.11 | 0.2 | 0.26 |
| Ind & Occ FE | No | No | Yes | No | No | Yes |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. Height is in inches. All regressions include a quadratic in age, indicators for year and for region, and a constant term. Columns 3 and 6 include industry and occupation indicators.

Table 4: Returns to Height for Immigrants in the NIS

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Height | 0.023 | 0.015 | 0.012 | 0.011 | 0.004 | 0.004 |
| | [0.003]** | [0.003]** | [0.003]** | [0.004]** | [0.004] | [0.004] |
| Education | | 0.073 | 0.029 | | 0.077 | 0.034 |
| | | [0.006]** | [0.007]** | | [0.009]** | [0.010]** |
| Observations | 3080 | 3028 | 2997 | 2104 | 2070 | 2051 |
| Adjusted $R^2$ | 0.07 | 0.11 | 0.21 | 0.03 | 0.07 | 0.14 |
| Ind & Occ FE | No | No | Yes | No | No | Yes |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. Height is measured in inches. All regressions include a quadratic in age, indicators for year and for region, and a constant term. Columns 3 and 6 include indicators for industry and occupation.

Table 5: Within-Country Estimates of Immigrants' Returns to Height

| | U.S. (NHIS) | | U.S. (NIS) | | U.K. (HSE) | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Men** | | | | | | |
| Height | 0.017 | 0.013 | 0.014 | 0.008 | 0.024 | 0.019 |
| | [0.002]** | [0.002]** | [0.003]** | [0.003]* | [0.008]** | [0.008]* |
| Education | 0.072 | 0.044 | 0.069 | 0.032 | 0.045 | 0.015 |
| | [0.003]** | [0.003]** | [0.008]** | [0.008]** | [0.008]** | [0.007]* |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Ind & Occ FE | No | Yes | No | Yes | No | Yes |
| Observations | 10012 | 9650 | 3015 | 2930 | 1684 | 1486 |
| Adjusted $R^2$ | 0.29 | 0.38 | 0.15 | 0.23 | 0.11 | 0.27 |
| **Panel B: Women** | | | | | | |
| Height | 0.008 | 0.008 | 0.005 | 0.003 | 0.019 | 0.013 |
| | [0.003]** | [0.003]** | [0.005] | [0.005] | [0.009]* | [0.010] |
| Education | 0.099 | 0.058 | 0.080 | 0.042 | 0.049 | 0.003 |
| | [0.003]** | [0.004]** | [0.010]** | [0.010]** | [0.008]** | [0.010] |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Ind & Occ FE | No | Yes | No | Yes | No | Yes |
| Observations | 7694 | 7403 | 2064 | 2006 | 1469 | 1177 |
| Adjusted $R^2$ | 0.27 | 0.36 | 0.10 | 0.18 | 0.08 | 0.21 |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. Height is measured in inches. All regressions include a quadratic in age, indicators for year and for region of residence in the U.S. or U.K., and a constant term. The additional controls in the even columns are indicators for industry and occupation. The NIS and HSE both have country of birth of immigrants, but the NHIS only provides information on region of birth.

Table 6: Nonlinear Estimates of the Returns to Height

| | Men | | | | | | Women | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NHIS Data | | HSE Data | | HRS Data | | NHIS Data | | HSE Data | | HRS Data | |
| | Native | Immigr | Native | Immigr | Native | Immigr | Native | Immigr | Native | Immigr | Native | Immigr |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| **Panel A: Quadratic Specification** | | | | | | | | | | | | |
| Height | 0.114 | 0.217 | 0.114 | 0.234 | -0.002 | 0.287 | 0.170 | -0.071 | 0.046 | 0.245 | 0.023 | 0.050 |
| | [0.046]* | [0.086]* | [0.098] | [0.208] | [0.059] | [0.190] | [0.049]** | [0.124] | [0.123] | [0.147]+ | [0.093] | [0.095] |
| $\text{Height}^2$ | -0.001 | -0.001 | -0.001 | -0.002 | 0.000 | -0.002 | -0.001 | 0.001 | 0.000 | -0.002 | 0.000 | 0.000 |
| | [0.000]* | [0.001]* | [0.001] | [0.002] | [0.000] | [0.001] | [0.000]** | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] |
| F-stat | 22.9** | 21.6** | 6.86** | 6.18** | 4.08* | 2.8 | 29.7** | 2.27 | 1.20 | 4.21* | 6.12** | 0.22 |
| Obs | 41524 | 9832 | 6342 | 1784 | 9096 | 897 | 43805 | 7542 | 6013 | 1422 | 10117 | 1023 |
| Adj $R^2$ | 0.30 | 0.36 | 0.30 | 0.29 | 0.15 | 0.31 | 0.29 | 0.35 | 0.29 | 0.23 | 0.23 | 0.25 |
| **Panel B: Logarithmic Specification** | | | | | | | | | | | | |
| Log(Height) | 0.510 | 0.926 | 0.704 | 1.507 | 0.574 | 1.118 | 0.526 | 0.405 | 0.355 | 1.220 | 0.606 | 0.209 |
| | [0.079]** | [0.153]** | [0.199]** | [0.460]** | [0.203]** | [0.553]* | [0.076]** | [0.197]* | [0.230] | [0.515]* | [0.173]** | [0.475] |
| Obs | 41524 | 9832 | 6342 | 1784 | 9096 | 897 | 43805 | 7542 | 6013 | 1422 | 10117 | 1023 |
| Adj $R^2$ | 0.30 | 0.36 | 0.30 | 0.29 | 0.15 | 0.30 | 0.29 | 0.35 | 0.29 | 0.23 | 0.23 | 0.25 |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. The dependent variable is the logarithm of hourly wages. Height is measured in inches. All regressions include a quadratic in age, education, indicators for year, region, 1 digit industry and occupation, and a constant term. The F-statistic refers to whether height and height squared are jointly significant.

Table 7: Information Quality and the Returns to Height and Education of Male Immigrants

| | NHIS Data | | HSE Data | | HRS Data | | NIS Data | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Height | 0.012 | 0.014 | 0.027 | 0.028 | 0.017 | 0.016 | 0.011 | 0.011 |
| | [0.004]** | [0.003]** | [0.016]+ | [0.010]** | [0.013] | [0.008]+ | [0.005]** | [0.004]* |
| Height*Years Since Immigration/10 | 0.001 | | 0.000 | | -0.006 | | 0.000 | |
| | [0.002] | | [0.006] | | [0.006] | | [0.004] | |
| Education | 0.046 | 0.041 | -0.004 | 0.008 | 0.040 | 0.041 | 0.053 | 0.025 |
| | [0.004]** | [0.003]** | [0.012] | [0.007] | [0.010]** | [0.008]** | [0.010]** | [0.008]** |
| Education*Years Since Immigration/10 | 0.002 | | 0.010 | | 0.000 | | -0.018 | |
| | [0.002] | | [0.005]* | | [0.004] | | [0.007]** | |
| Years Since Immigration/10 | 0.011 | | -0.004 | | 0.548 | | 0.472 | |
| | [0.013] | | [0.039] | | [0.371] | | [0.064] | |
| Height*$I$(Educated in Host Country) | | 0.001 | | -0.005 | | -0.027 | | 0.003 |
| | | [0.004] | | [0.016] | | [0.021] | | [0.006] |
| Education*$I$(Educated in Host Country) | | 0.002 | | 0.004 | | 0.015 | | 0.005 |
| | | [0.004] | | [0.015] | | [0.025] | | [0.014] |
| $I$(Educated in Host Country) | | 0.076 | | 0.384 | | 1.814 | | 0.046 |
| | | [0.291] | | [1.041] | | [1.597] | | [1.114] |
| Observations | 9649 | 9649 | 1486 | 1486 | 842 | 885 | 2942 | 2971 |
| P-value of F-test: | | | | | | | | |
| $\beta_1 = 0$ & $\beta_2 = 0$ | 0.330 | 0.864 | 0.130 | 0.924 | 0.365 | 0.100 | 0.025 | 0.703 |
| Adjusted $R^2$ | 0.38 | 0.37 | 0.25 | 0.25 | 0.33 | 0.31 | 0.22 | 0.22 |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. The dependent variable is the logarithm of hourly wages. Height is measured in inches. All regressions include a quadratic in age, indicators for year, region, 1 digit industry and occupation, and a constant term.

Table 8: Information Quality and the Returns to Height and Education of Female Immigrants

| | NHIS Data | | HSE Data | | HRS Data | | NIS Data | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Height | 0.008 | 0.006 | 0.042 | 0.018 | 0.003 | -0.003 | 0.006 | 0.002 |
| | [0.006] | [0.004] | [0.021]* | [0.012] | [0.012] | [0.008] | [0.006] | [0.005] |
| Height*Years Since Immigration/10 | 0.000 | | -0.010 | | -0.004 | | -0.002 | |
| | [0.003] | | [0.008] | | [0.006] | | [0.005] | |
| Education | 0.041 | 0.003 | -0.016 | -0.004 | 0.008 | 0.018 | 0.038 | 0.033 |
| | [0.006]** | [0.006] | [0.021] | [0.012] | [0.009] | [0.007]* | [0.014]** | [0.013]** |
| Education*Years Since Immigration/10 | 0.013 | | 0.007 | | 0.009 | | 0.001 | |
| | [0.002]** | | [0.008] | | [0.005]+ | | [0.011] | |
| Years Since Immigration/10 | -0.068 | | 0.052 | | 0.186 | | 0.059 | |
| | [0.317] | | [0.054] | | [0.034] | | [0.085] | |
| Height*$I$(Educated in Host Country) | | 0.051 | | 0.004 | | 0.006 | | 0.007 |
| | | [0.004]** | | [0.018] | | [0.024] | | [0.008] |
| Education*$I$(Educated in Host Country) | | 0.016 | | 0.020 | | 0.048 | | 0.003 |
| | | [0.005]** | | [0.020] | | [0.020]* | | [0.017] |
| $I$(Educated in Host Country) | | -0.265 | | -0.495 | | -0.965 | | -0.96 |
| | | [0.385] | | [1.177] | | [1.506] | | [1.357] |
| P-value of F-test: | | | | | | | | |
| $\beta_1 = 0$ & $\beta_2 = 0$ | 0.000 | 0.008 | 0.332 0.572 | 0.077 | 0.068 | 0.905 | 0.587 | |
| Observations | 7404 | 7404 | 1177 | 1177 | 966 | 1006 | 2011 | 2034 |
| Adjusted $R^2$ | 0.36 | 0.35 | 0.20 | 0.20 | 0.31 | 0.29 | 0.15 | 0.15 |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. The dependent variable is the logarithm of hourly wages. Height is measured in inches. All regressions include a quadratic in age, indicators for year, region, 1 digit industry and occupation, and a constant term.

Table 9: Comparison of Pre- and Post-Immigration Wages of NIS Immigrants

| | Male Immigrants | | | Female Immigrants | | |
|---|---|---|---|---|---|---|
| $Q$= | | Yrs in US | US Edu | | Yrs in US | US Edu |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Height | 0.018+ | 0.009 | 0.020+ | 0.008 | 0.011 | 0.017 |
| | [0.009] | [0.014] | [0.011] | [0.017] | [0.024] | [0.019] |
| Pre-Immig*Height ($\gamma_2$) | 0.010 | -0.016 | 0.006 | 0.032 | 0.051+ | 0.031 |
| | [0.013] | [0.019] | [0.014] | [0.020] | [0.030] | [0.022] |
| Education | 0.093** | 0.100** | 0.127** | 0.087** | 0.077+ | 0.097* |
| | [0.021] | [0.030] | [0.024] | [0.027] | [0.046] | [0.041] |
| Pre-Immig*Education ($\gamma_4$) | -0.027 | -0.005 | -0.023 | 0.022 | 0.030 | -0.001 |
| | [0.030] | [0.039] | [0.030] | [0.044] | [0.064] | [0.053] |
| Pre-Immig | -2.454 | 1.594 | -1.888 | -6.371* | -9.511* | -5.829+ |
| | [2.118] | [3.143] | [2.309] | [3.176] | [4.574] | [3.413] |
| Height*Q ($\gamma_6$) | | 0.011 | -0.022 | | -0.004 | -0.067 |
| | | [0.013] | [0.020] | | [0.016] | [0.043] |
| Education*Q ($\gamma_7$) | | -0.008 | -0.109* | | 0.013 | -0.037 |
| | | [0.019] | [0.050] | | [0.034] | [0.053] |
| Pre-Immig*Q | | -5.030 | -1.436 | | 4.463 | -6.018 |
| | | [3.218] | [6.044] | | [3.681] | [10.211] |
| Pre-Immig*Height*Q ($\gamma_9$) | | 0.034+ | 0.016 | | -0.026 | 0.030 |
| | | [0.020] | [0.033] | | [0.023] | [0.062] |
| Pre-Immig*Education*Q ($\gamma_{10}$) | | -0.035 | -0.107 | | -0.011 | 0.081 |
| | | [0.035] | [0.110] | | [0.050] | [0.102] |
| Q | | -1.711 | 5.571 | | 0.627 | 11.409 |
| | | [2.204] | [3.530] | | [2.657] | [6.979] |
| P-values of F-test: | | | | | | |
| $\gamma_2 = 0$ & $\gamma_4 = 0$ | 0.593 | | | 0.183 | | |
| $\gamma_6 = 0$ & $\gamma_7 = 0$ | | 0.693 | 0.030 | | 0.916 | 0.195 |
| $\gamma_6 + \gamma_9 = 0$ | | 0.003 | 0.805 | | 0.131 | 0.427 |
| $\gamma_7 + \gamma_{10} = 0$ | | 0.158 | 0.022 | | 0.978 | 0.665 |
| Observations | 1,068 | 1,056 | 1,065 | 621 | 619 | 620 |
| Adjusted R-squared | 0.203 | 0.213 | 0.220 | 0.180 | 0.181 | 0.177 |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. The dependent variable is the logarithm of pre-immigration hourly wages at the time of immigration (in real U.S. dollars). Height is measured in inches. Years since immigration and education in host country refer to the individual's post-immigration status. All regressions include a quadratic in age, indicators for country and year of immigration, and a constant term.

Table 10: Immigrants' Returns to Height and Per Capita GDP of Country of Origin

| | Male Immigrants | | | | Female Immigrants | | | |
|---|---|---|---|---|---|---|---|---|
| | NIS Sample | | HSE Sample | | NIS Sample | | HSE Sample | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Height | 0.025 | 0.031 | 0.031 | 0.03 | 0.007 | 0.074 | 0.023 | 0.021 |
| | [0.009]** | [0.017]+ | [0.011]** | [0.010]** | [0.012] | [0.031]* | [0.016] | [0.013] |
| Height*($\frac{GDP}{N}$ quartile 2) | -0.003 | | 0.001 | | -0.005 | | -0.021 | |
| | [0.004] | | [0.021] | | [0.004] | | [0.030] | |
| Height*($\frac{GDP}{N}$ quartile 3) | -0.002 | | -0.033 | | -0.004 | | -0.003 | |
| | [0.004] | | [0.022] | | [0.004] | | [0.047] | |
| Height*($\frac{GDP}{N}$ quartile 4) | -0.011 | | -0.047 | | -0.015 | | -0.043 | |
| | [0.006]+ | | [0.027]+ | | [0.007]* | | [0.026]+ | |
| Height*($\frac{GDP}{N}$ >\$1600) | | -0.013 | | -0.046 | | -0.098 | | -0.033 |
| | | [0.023] | | [0.022]* | | [0.035]** | | [0.025] |
| Observations | 2984 | 1828 | 1011 | 1011 | 2045 | 1286 | 610 | 610 |
| Adjusted $R^2$ | 0.22 | 0.24 | 0.32 | 0.32 | 0.17 | 0.2 | 0.23 | 0.24 |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. Height is measured in inches. All regressions include a quadratic in age, indicators for year and for region, education, urban residence, and indicators for country of origin, industry and occupation, and a constant term.

Table 11: Relationship between Height, Health and Cognition

| | Health Status | | | Cognitive Ability | | |
|---|---|---|---|---|---|---|
| | NHIS | HSE | HRS | NHANES | | HRS |
| | | | | SDST | SDLT | WAIS |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Men** | | | | | | |
| Height | -0.031 | -0.018 | -0.047 | -0.014 | -0.091 | 0.126 |
| | [0.001]** | [0.003]** | [0.007]** | [0.004]** | [0.017]** | [0.019]** |
| Immigrant*Height | -0.005 | -0.027 | 0.012 | -0.031 | -0.161 | 0.002 |
| | [0.003] | [0.006]** | [0.019] | [0.007]** | [0.035]** | [0.059] |
| Immigrant | 0.297 | 1.972 | -0.924 | 5.778 | 30.103 | -0.767 |
| | [0.210] | [0.430]** | [1.299] | [1.269]** | [5.950]** | [3.976] |
| Observations | 83156 | 15631 | 4534 | 2300 | 2250 | 3871 |
| Adjusted $R^2$ | 0.06 | 0.04 | 0.01 | 0.12 | 0.13 | 0.02 |
| **Panel B: Women** | | | | | | |
| Height | -0.026 | -0.021 | -0.024 | -0.020 | -0.104 | 0.122 |
| | [0.001]** | [0.003]** | [0.006]** | [0.004]** | [0.016]** | [0.016]** |
| Immigrant*Height | -0.010 | -0.022 | -0.001 | -0.033 | -0.144 | 0.024 |
| | [0.003]** | [0.006]** | [0.017] | [0.009]** | [0.035]** | [0.048] |
| Immigrant | 0.684 | 1.599 | 0.392 | 5.993 | 25.653 | -2.671 |
| | [0.216]** | [0.365]** | [1.067] | [1.381]** | [5.543]** | [3.005] |
| Observations | 101694 | 19194 | 6324 | 2761 | 2697 | 5813 |
| Adjusted $R^2$ | 0.05 | 0.05 | 0.03 | 0.12 | 0.12 | 0.04 |

Notes: Robust standard errors clustered by household in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. Height is measured in inches. All regressions include a quadratic in age, indicators for year and a constant term. In columns 1-3, the dependent variable, health, is a self-reported measure where 1 equals excellent health and 5 equals poor health. The measure of cognition is the error-corrected speed for the symbol digit substitution test (SDST), total score in the serial digit learning test (SDLT) in column 5, and the Wechsler Adult Intelligence Scale (WAIS) score in column 6. Cognitive ability is increasing in the WAIS score, but decreasing in the other measures.

Appendix Table 12: Employer Learning and Statistical Discrimination with Height

| | NLSY Full Sample | | | HRS Native-Born Sample | | | HRS Immigrant Sample | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| **Panel A: Men** | | | | | | | | | |
| Education | 0.073 | 0.074 | 0.091 | 0.073 | 0.074 | 0.087 | 0.035 | 0.037 | 0.044 |
| | [0.002]** | [0.003]** | [0.003]** | [0.009]** | [0.009]** | [0.010]** | [0.030] | [0.030] | [0.031] |
| Height | 0.004 | 0.004 | 0.001 | 0.010 | -0.008 | -0.008 | 0.023 | 0.003 | 0.001 |
| | [0.001]** | [0.001]** | [0.001] | [0.003]** | [0.008] | [0.008] | [0.009]* | [0.029] | [0.029] |
| Ability | 0.128 | 0.129 | 0.054 | 0.015 | 0.015 | -0.009 | 0.027 | 0.027 | -0.001 |
| | [0.002]** | [0.002]** | [0.005]** | [0.003]** | [0.003]** | [0.008] | [0.011]* | [0.011]* | [0.033] |
| Education*Exp/10 | 0.004 | 0.003 | -0.013 | 0.001 | 0.001 | -0.004 | 0.005 | 0.004 | 0.002 |
| | [0.001]** | [0.002] | [0.002]** | [0.003] | [0.003] | [0.003] | [0.009] | [0.009] | [0.009] |
| Height*Exp/10 | | 0.000 | 0.003 | | 0.007 | 0.007 | | 0.006 | 0.007 |
| | | [0.000] | [0.000]** | | [0.003]** | [0.003]** | | [0.008] | [0.008] |
| Ability*Exp/10 | | | 0.069 | | | 0.009 | | | 0.009 |
| | | | [0.004]** | | | [0.003]** | | | [0.010] |
| Observations | 53279 | 53279 | 53279 | 8536 | 8536 | 8536 | 753 | 753 | 753 |
| Adjusted $R^2$ | 0.31 | 0.31 | 0.32 | 0.12 | 0.12 | 0.12 | 0.23 | 0.23 | 0.23 |
| **Panel B: Women** | | | | | | | | | |
| Education | 0.100 | 0.087 | 0.094 | 0.087 | 0.087 | 0.082 | 0.040 | 0.045 | 0.058 |
| | [0.002]** | [0.002]** | [0.002]** | [0.008]** | [0.008]** | [0.009]** | [0.022]+ | [0.022]* | [0.024]* |
| Height | 0.002 | 0.006 | 0.004 | 0.010 | 0.009 | 0.008 | -0.009 | -0.046 | -0.042 |
| | [0.001]** | [0.001]** | [0.001]** | [0.002]** | [0.007] | [0.007] | [0.007] | [0.025]+ | [0.025]+ |
| Ability | 0.118 | 0.116 | 0.078 | 0.017 | 0.017 | 0.026 | 0.045 | 0.046 | 0.005 |
| | [0.002]** | [0.002]** | [0.004]** | [0.002]** | [0.002]** | [0.007]** | [0.008]** | [0.008]** | [0.026] |
| Education*Exp/10 | -0.017 | -0.004 | -0.011 | 0.000 | 0.000 | 0.002 | -0.002 | -0.004 | -0.008 |
| | [0.001]** | [0.002]* | [0.002]** | [0.003] | [0.003] | [0.003] | [0.007] | [0.007] | [0.007] |
| Height*Exp/10 | | -0.003 | -0.002 | | 0.000 | 0.000 | | 0.012 | 0.011 |
| | | [0.000]** | [0.000]** | | [0.002] | [0.002] | | [0.007] | [0.007] |
| Ability*Exp/10 | | | 0.036 | | | -0.004 | | | 0.013 |
| | | | [0.003]** | | | [0.002] | | | [0.008]+ |
| Observations | 53790 | 53790 | 53790 | 9968 | 9968 | 9968 | 927 | 927 | 927 |
| Adjusted $R^2$ | 0.33 | 0.33 | 0.33 | 0.2 | 0.2 | 0.2 | 0.22 | 0.22 | 0.23 |

Notes: Robust standard errors in brackets. **, *, + denotes significance at the 1%, 5% and 10% level, respectively. Height is measured in inches. All regressions include a cubic in potential experience, indicators for year and a constant term. In the NLSY results, ability is measured with standardized AFQT scores. In the HRS results, ability is measured with standardized WAIS scores.