CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of economic studies

Bachelor‘s thesis

2012 **Jan Svoboda**

CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of economic studies

Jan Svoboda

Position of Minorities in the U.S. Labor Market

Bachelor's thesis

Prague 2012

Author: Jan Svoboda
Supervisor: Ing. Barbara Pertold-Gebicka, PhD
Academic year: 2011/2012

Bibliographic entry

SVOBODA, Jan. *Position of Minorities in the U.S. Labor Market.* Prague, 2012. 94 pp. Bachelor's thesis, Charles University, Faculty of Social Sciences, Institute of economic studies. Supervisor: Ing. Barbara Pertold-Gebicka, PhD

Abstrakt

Tato práce se zabývá fenoménem rasové diskriminace na americkém trhu práce. První část přibližuje existující teorie diskriminace, historii pobytu imigrantů v USA, přináší několik statistik z amerického trhu práce a konečně výsledky několika různých výzkumů z této oblasti. Druhá část práce je věnována empirické analýze mzdové diskriminace v USA na základě dat z American Community Survey 1980 a 2007. Pomocí Oaxacovy-Blinderovy dekompozice jsme rozložili rozdíl v logaritmech mezd mezi bělochy a ostatními rasami na dvě části. První část, vysvětlená, je ta část rozdílu ve mzdách, která lze vysvětlit pomocí rozdílu v průměrných charakteristikách jednotlivých ras. Druhá část, nevysvětlená, je ta část rozdílu ve mzdách, která takto vysvětlit nelze. Ta je pak označena za důsledek diskriminace. Zjistili jsme, že zatímco v roce 1980 měli asiati v průměru o 5% vyšší mzdu, v roce 2007 to bylo již o 15%. 15% - 21% z tohoto rozdílu ale zůstalo nevysvětleno. Opačná situace je v případě hispánců a afroameričanů. Zatímco v roce 1980 měli v průměru o 10% nižší mzdy než běloši, v roce 2007 to bylo už o 30%. V případě afroameričanů zůstalo 50% - 60% nevysvětleno, v případě hispánců 46%. Zdá se, že zaměstnavatelé na americkém trhu práce mají silný sklon diskriminovat lidský kapitál menšin.

Klíčová slova

Trh práce, diskriminace, menšina, rasa, Oaxaca, rozklad, dekompozice, mzda, rozdíl

Rozsah práce

126 015 znaků

JEL klasifikace

J31, J71

Abstract

This thesis deals with the phenomenon of racial discrimination in the U.S. labor market. The first part outlines the existing theories of discrimination and history of immigrants' residence in the U.S. It also brings some statistics from the U.S. labor market and finally reviews some literature in the field of labor market discrimination. The second part of the thesis is devoted to the empirical analysis of wage discrimination in the U.S. using the American Community Survey 1980 and 2007. With the help of Oaxaca-Blinder decomposition we break down the log-wage differential between whites and other racial groups into two parts. The first, explained, part is the part of the differential that can be explained by differences in average race characteristics. The other, unexplained, part is the part of the differential that remained unexplained. That part is then assigned as a result of discrimination. We find that while in 1980 Asians earned on average 5% more than whites, by 2007 this difference reached 15%. 15% - 21% of this difference remained unexplained. The opposite situation is observed in the case of Hispanics and Blacks. While in 1980 both, Hispanics and Afro-Americans, had on average 10% lower wage than whites, by 2007 this difference reached as much as 30%. In case of Afro-Americans, 50% - 60% remained unexplained, in case of Hispanics 46% remained unexplained. It seems that employers in the U.S. labor market have strong tendency to discriminate against human capital of minorities.

Keywords

Labor market, discrimination, minority, race, Oaxaca, decomposition, wage, differential, gap

Length of the thesis

126 015 characters

JEL classification

J31; J71

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature. The author also declares that he has not used this thesis to acquire another academic degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

In Prague, 15 May 2012 Jan Svoboda

Acknowledgements

I am greatly indebted to Ing. Barbara Pertold-Gebicka, PhD for her valuable comments, suggestions, language correction, and also for encouraging me.

Bachelor Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author: Jan Svoboda
Supervisor: Ing. Barbara Pertold-Gebicka, PhD
Specialization: Labor Economics, Applied Microeconometrics
Defense Planned June 2012

Proposed Topic

Position of Minorities in the U.S. Labor Market

Preliminary outline in English

This thesis will consist of two parts. In the first one I will characterize the American labor market and today's theories related to the labor market discrimination. After that I will deal with the synthesis of works already written on this topic.

The beginning of the second part will be devoted to the methodology. Later, based on data and using econometric tools, I will examine the U.S. labor market for discrimination against ethnic minorities. With the help of Oaxaca decomposition I will be able to distinguish between the discrimination itself and the characteristics of single ethnicities. Then I will compare the results. To deal with the problem completely I will also compare the rate of discrimination over time.

My contribution will be grounded in summarizing the current state of literature on the topic and what's more in using the recent data, which will enable me to evaluate whether the position of minorities became better, or not.

Preliminary outline in Czech

Tato práce bude mít dvě části. V té první přiblížím americký trh práce a soudobé teorie týkající se diskriminace na trhu práce. Dále se budu věnovat syntéze již napsaných prací věnovaných tématu.

Začátek druhé části bude věnován metodologii. Následně budu na základě dat pomocí ekonometrických metod zjišťovat, zda na americkém trhu práce dochází k diskriminaci etnických menšin. Pomocí Oaxacovy dekompozice odliším vliv charakteristik jednotlivých etnik od samotné diskriminace a výsledky posléze mezi jednotlivými

etniky porovnám. Celým problémem se budu zabývat komplexně, a srovnám míru diskriminace s předchozími lety.
Můj přínos bude spočívat jednak ve shrnutí do této doby k tématu napsaných prací, a jednak v použití aktuálních dat, díky čemuž budu schopen zhodnotit, zda se postavení menšin zlepšilo, či nikoli.

Core Bibliography

[1] Altonji, J. G., and Blank, R. M. 1999, 'Race and Gender In the Labor Market' in Orley Ashenfelter and David Card, editors, *Handbook* of *Labor Economics,* Volume 3C, Elsevier, Amsterdam, pp. 3143-3259.

[2] Becker, Gary S. 1971, *The Economics of Discrimination,* 2nd edition. The University of Chicago Press, Chicago, IL.

[3] Oaxaca, Ronald L. 1973, 'Male-female wage differentials in urban labor markets', *International Economic Review,* Vol. 14, No. 3, pp. 693-709.

[4] Oaxaca, Ronald L., and Ransom, Michael R. 1994, 'On discrimination and the decomposition of wage differentials', *Journal of Econometrics*, Vol. 61, No. 1, pp. 5-21.

In Prague, 3 June 2011

Student's signature

Supervisor's signature

Contents

List of Figures and Tables

Introduction

As Arrow (1998) mentions, when there is a discrimination present in the society, it permeates through the whole society, in all of its aspects. Before the Title VII of the Civil Rights Act of 1964 "*the presence of racial discrimination throughout American society was, to use the words of Samuel Johnson, a fact "too evident for detection and too gross for aggravation*" (Arrow, 1998, pg. 92). This situation is likely to change only slowly in the society as the phenomenon of discrimination is rooted very deeply.

These words of Kenneth Arrow constitute the point of departure for our analysis. We are interested whether and how the situation of different races has changed since then, therefore we decided to investigate the available data and conduct an empirical research. Most of the literature on the topic of discrimination deals with only one of the ethnic minorities in the U.S. We are, however, interested in comparing the positions of three largest of them. To thess ends, we take advantage of the American Community Survey data, which contain 1% sample of the U.S. population, for the years 1980 and 2007. We are mainly interested whether the observed wage difference between white workers and workers of another race, namely Afro-Americans, Hispanics and Asians, can be sufficiently explained by the differences in their average characteristics. To decompose the observed wage differentials we use the Oaxaca-Blinder decomposition. The part of the gap that remains unexplained is then assigned to discrimination against the particular group. We are also interested in relative positions of the minority groups and how these positions have improved during the era of "information society."

The thesis consists of two parts. In the first part, in chapter one, we introduce different theories of discrimination, namely the taste based theory of discrimination and the statistical theory of discrimination. The second chapter deals with the history of residence in the U.S. concerning the Afro-American, Hispanic and Asian minorities. It also delivers some statistics from the labor market based on the data available at the U.S. Bureau of Labor Statistics. Investigating these statistics naturally leads us to formulating the research questions and hypotheses, which can be found in chapter three. Chapter four contains the literature review, where we present some outcomes of other researchers. These outcomes encouraged us even more concerning our own empirical research. Our own empirical analysis is included in the second part of the thesis, in chapter five. Conclusion follows as the last, sixth, chapter of this work.

1. Theories of Labor Market Discrimination

Although discrimination was a huge problem in the United States for decades, and not only in the labor market (Arrow, 1998, mentions legal barriers, prices paid, credit extended, etc.), the first economic theory of discrimination was not developed till 1957, when Gary S. Becker published his dissertation *The Economics of Discrimination*. As Becker (1971) recalls on the first pages, the process of publishing was problematic and the text was not easily accepted. By that time, there was a general consensus that discrimination is more a phenomenon which should be studied by sociologists or psychologists, even though it has a great economic impact on both micro- and macroeconomic level. Not surprisingly, no economic research or any interest seemed to be concentrated in this field for another 5 years. Since then economists have understood they have what to offer in the field of discrimination and research on this topic has grown extensively, bringing several theories explaining this phenomenon.

The theory Becker developed is called "taste-based discrimination" and as the first type will be introduced in the following part. The second type of theories is known under the name "statistical discrimination," and has been worked on mainly by Phelps (1972), Arrow (1972a, 1972b, 1973), Spence (1973, 1974), McCall (1972a, 1972b) and Reder (1972), and will be introduced after that.

1.1 Taste-based Theory of Discrimination

Before we go in for the theory concerning different economic agents we need to define a few phenomena and make an introduction to this part, mainly to present some basic notion and concepts of Becker's theory.

We would find many different definitions of discrimination in the literature. But not all the fields define it so intuitively. Becker (1971) mentions definitions that can be found in the socio-psychological literature, namely *"one individual is said to discriminate against (or in favor of)*[1] *another if his behavior toward the latter is not motivated by an "objective" consideration of fact"* (pg. 13). He adds that "*it is difficult to use this definition in distinguishing a violation of objective facts from an expression of tastes or values*" (pg. 13). This brings us to Becker's definition of discrimination in the market place and his concept of "taste for discrimination," the core idea behind his

[1] Positive discrimination, so called "nepotism".

theory. The "taste for discrimination" is defined as the willingness "*to pay something either directly or in the form of reduced income, to be associated with some persons instead of others*" (pg. 14).

According to Becker's theory, the taste for discrimination is to a great extent determined by the position of the group discriminated against, i.e. by its economic and quantitative importance in the society; by "*the social and physical distance between the groups and by their relative socioeconomic status*" (pg. 16). As economists, we would like to quantify the taste for discrimination. For these means Becker introduces, the concept of a discrimination coefficient (from now on defined as DC). A DC measures the percentage by which money costs or returns from certain economic action differ from the perceived ones; therefore it measures the non-monetary aspects of transactions. To give an example: assume that an employer discriminates against an individual, whose money wage rate is w, then the employer behaves as if the net wage rate was $w(1+d)$, where d is the employer's DC measuring his taste for discrimination. Similar examples can be formulated for a discriminating employee or a discriminating customer, as will be shown later. A DC is positive in case of a discrimination against the individual dealt with and as such means disutility for the discriminator.

Another concept we need to introduce is the market discrimination coefficient (from now on defined as MDC). Let us assume two different groups, W and N, which are imperfect substitutes in production, and a perfectly competitive labor market. In the context of my work N would be considered as the minority and W as the majority. Then the MDC is defined as the difference between the ratio of these groups' wage rates with and without discrimination:

$$MDC = \frac{w_w}{w_n} - \frac{w_w^0}{w_n^0}, \tag{1}$$

where index 0 denotes wage rates without discrimination. If we alter our assumption to W and N being perfect substitutes in production, the wage rates under no discrimination would be equal, turning us to a simpler definition of MDC:

$$MDC = \frac{w_w - w_n}{w_n}. \tag{2}$$

The MDC partly depends on the individual differences in tastes, therefore each individual's DC, or to be more specific, the whole distribution of DC's in the market, need to be known, as well as the weights of individual DC's. These are partly

determined by the relative importance of monopoly and competition in both, labor and product market.

1.1.1 Employer Discrimination

While talking about employer discrimination let us assume there is no discrimination from any other source but the employer. In the definition mentioned above is discrimination defined as "no objective consideration of facts." In the labor market, it is the productivity which is to be considered. Thus an employer is not discriminating by refusing to hire a worker with his marginal product being lower than the marginal costs. Such behavior is reasonable. On the other hand, refusing to hire a worker with his marginal product being greater than the marginal costs would be considered discriminatory.

Let us assume that an individual has money wage rate w. A discriminating employer perceives the wage paid to this worker as $w(1+d)$, where d is the employer's DC measuring his taste for discrimination. This DC might to a great extent depend on the degree of employer's contact with employees. He will probably discriminate less against those he hardly meets.

Let us assume that labor is the only factor of production. A rational profit maximizing employer tries to find the optimal combination of inputs for every level of output. In the classical theory with competitive markets, this means finding such combination of factors that marginal product of each factor equals its marginal cost, i.e.

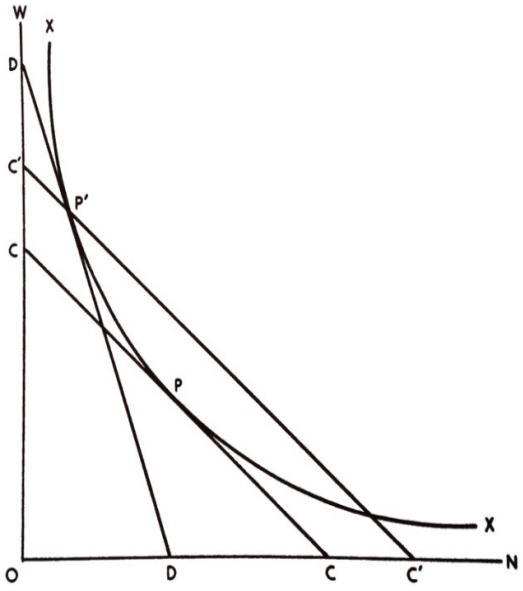
$$\frac{MP_i}{MP_j} = \frac{w_i}{w_j}, i,j = (1,2 \dots n). \tag{3}$$

With no discriminatory behavior of the employer the wage is the net cost of production at the same time. When we employ discrimination into our assumptions, nothing changes on the employer's effort to minimize net costs, but this time net costs are $w(1+d)$. Thus

$$\frac{MP_i}{MP_j} = \frac{w_i(1+d^i)}{w_j(1+d^j)}, i,j = (1,2 \dots n). \tag{4}$$

It is clear that the equilibrium combination would not be the same as in the classical case. This situation is explained in Figure 1. Horizontal axis measures the number of N workers employed and vertical axis measures the number of W workers. The XX curve is the production isoquant and the CC line is the relevant price line with no discrimination. To produce X, corresponding coordinates of P are the amounts of W and N employed in the optimum, OC being the minimum costs. Now, let us assume that the employer has a taste for discrimination represented by some d. Then the DD line is the relevant price line taking into account net costs $w(1 + d^N)$. As the net costs of hiring those discriminated against are higher the demand for them would be lower, resulting in P' being the number of W and N employed. Finally, considering that X is to be produced and P' is the optimum employment of W and N, the money costs of production (represented by C'C' line) would be higher than the minimum.

Figure 1: The effect of discrimination by a firm on its costs and employment



Source: Becker (1971)

1.1.2 Employee Discrimination

While talking about employee discrimination let us assume there is no discrimination from any other source but the co-workers. Let us also assume that a W employee has a money wage rate w_n but working together with N brings him disutility and therefore he perceives his wage rate as if it were $w_n(1 - d)$, where d is the employee's DC measuring his taste for discrimination. If the employee could get a higher wage for working with N rather than with W only, then the cost from discriminating can be measured by $c = \frac{w_n - w_w}{w_n}$. Thanks to comparing c with d, the employee can decide rationally whether to work with N even though he has a taste for discrimination. If

- $c > d$, the real loss of wage, i.e. the money cost of working with W only, is greater than the net cost stemming from the taste for discrimination, the employee would decide to work with N,
- $c < d$, the cost of working with W only is smaller than the disutility brought by working with N while having the taste for discrimination, the employee would choose to work with W only,
- $c = d$, the employee is indifferent between working with N and working with W only, because both cost are equal.

As Becker (1971) adds, "*an increase in* w_n *relative to* w_w *or a decrease in* d *increases the employee's desire to work with N*" (pg. 56).

At this point we need to define the notion "market segregation," so that we are able to examine what would be the outcome in the market place where employee discrimination is present. Market segregation in a region exists if race group members *"are employed with each other to a significantly greater extent than would result from their random distribution*" (Becker, 1971, pp. 56-57). Let us now assume that W and N are perfect substitutes in production. If a W worker with taste to discriminate against N is supposed to work with N, his employer would have to pay him a higher wage. It would be reasonable for the employer in order to maximize his income to employ workers of one group only. He would hire W if W's wage rate is lower than N's, and N in the opposite case. He would be indifferent between them only in the case when their wage rates are equal. This setup would end up as market segregation with firms hiring teams of W or teams of N only. [2]

What if we ease the assumption of perfect substitutes? As known from the basic microeconomics, the degree of competition decreases as the substitutability decreases. It is similar with tastes for discrimination. As mentioned above, the tastes for discrimination are partly determined by the economic and quantitative importance of the minority, it follows that the "*...tastes for discrimination against N are largest among factors that are the best substitutes for N*" (Becker, 1971, pg. 61).

---

[2] "*...market segregation and market discrimination are separate concepts referring to separate phenomena. Market discrimination refers to the incomes received by different groups and ignores their distribution in employment; market segregation refers to their distribution in employment and ignores their incomes*" (Becker, 1971, pp. 57-58). They can occur separately but more often occur together. Market discrimination is caused by taste for discrimination, while market segregation is caused by differences in tastes (Becker, 1971).

We also introduce discrimination by trade unions, though very briefly because the research in this area was not as extensive as in others. The DC of trade union against a group N is determined by its members' DC's against this group. There are generally two ways the trade unions can behave:

- decisions are made by majority rule. Then the trade union's DC would equal the median DC in the distribution of members' DC's,
- a union leader and decision maker is elected. Then the union's DC would on average equal the median DC in the distribution of members' DC's.

The outcomes are in general the same. If the trade unions have DC against non-union N, these are excluded from the union. A market discrimination against N may occur in this way even under the assumption of competitive labor market and W and N being perfect substitutes in production.

1.1.3 Consumers and Government Discrimination

When a customer decides whether to do the shopping in a particular retail store, he usually considers firstly prices but also the speed of its personnel and reliability. In addition to that a customer with taste for discrimination would also take into account the sex, race, religion, and personality of the personnel. Let us consider a consumer with taste for discrimination against members of group N. Such consumer would perceive the price P_n of a good or service sold by N as if it were $P_n(1+d)$, where d is again the DC of this consumer.

At this point we need to emphasize that Welch (1967) argues that intermediation in the products' market is sufficient to prevent from discrimination and therefore the costs born. The taste for discrimination would appear at the time the product is bought, therefore the producer himself is not important. It suffices to find an intermediary with DC lower than the consumers' DC's. Such person would serve for the purpose of absorbing the difference in the tastes for discrimination. This is evidently not possible in personal services as consumers are in direct contact with the sellers. Welch (1967) adds further that the case of supplier-producer relation, with the supplier having taste for discrimination against the producer, is pretty much the same story.

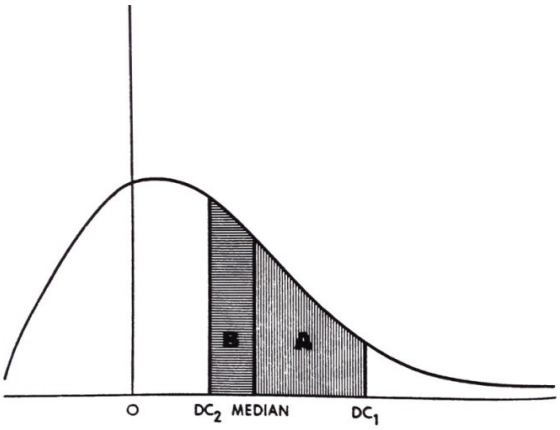
Back to Becker (1971) and the outcome in the market. Let us assume that all consumers have the same DC, d, and m units of labor supplied by either N or W are needed the produce or sell one unit of output. The MDC against N would then equal:

$$MDC = \frac{P_n d}{m w_n}. \tag{5}$$

Let us now ease the assumption of all consumers having the same DC. The d would then equal the DC of the consumer who is indifferent between buying from N and buying from W. For him holds $P_n(1 + d) = P_w$.

Turning to the government discrimination, consider the situation when there are only two political parties taking part in the election. People are deciding whom to vote only according to the promised government policy (DC) toward the groups N and W. Each individual votes for the party with promised DC closer to his own. The following figure shows the distribution of DC's among voters. Becker (1971, pg. 82) comments this:"*Clearly, a promise of DC_1 (to the right of the median) could not be an equilibrium position, since a promise of any DC in region A must receive more votes; likewise, a promise of DC_2 (to the left of the median) could not be an equilibrium one, since any DC in region B must receive more votes. Therefore, the median DC is the only possible equilibrium position*".

Figure 2: The distribution of tastes for discrimination among the electorate



Source: Becker (1971)

As it is usually not the case that electorates vote only on the basis of this one particular issue, the minority groups often have greater influence on the governmental policy concerning this issue than would correspond to their quantitative importance. Let us assume two groups, W and N, and two issues, amount of discrimination against N and another one. If W is more interested in the second issue, even though it would prefer more discrimination, and there are two groups W_1 and W_2 which differ to a great extent in their attitudes toward the second issue, N has a great opportunity to make their position concerning discrimination the governmental policy.

1.1.4 Taste-based Theory Conclusion

So far, we have taken into account tastes for discrimination of different agents one by one. In the real world, however, it is not that easy. We need to consider the situation in the market place under such circumstances when more or even all the

economic agents behave discriminatory. Becker (1971) shows that, "*the actual market discrimination is a summation of the market discrimination caused by each group considered separately*" (pg. 84). Therefore the wage differences as well as the segregation by occupation we observe in data are the outcomes (at least to some extent) of discriminatory behavior of all the different agents mentioned above. The degree of discrimination in the market place is not influenced only by the economic agents' tastes for discrimination but also by the relative economic importance of the group N. Tastes being given, "*an increase in N's relative economic importance would decrease the market discrimination against them*" (pg. 85).

Becker (1971) demonstrates that if discrimination occurs in the economy, volume of trade between the groups (and thus their net incomes) is decreased, and falls further more with all increases in discrimination. If the tastes for discrimination were so large that it would no longer be worth trading, the two groups would be in complete economic isolation including complete economic segregation. Becker (1971) points out that "*complete segregation does not avoid the bad economic effects of the discrimination but only multiplies them*" (pg. 24).

1.2 Statistical Theory of Discrimination

In the previous part we explained, how the concept of taste discrimination works, i.e. how it is possible that workers of the same ability while being of another race (sex) receive different rewards for their equal work depending on their environment. In this part we will explain how it is possible to come across racial wage differentials between two equally skilled workers even in the absence of tastes and prejudice. To do so, we will use the basic Phelps' model of statistical discrimination (Phelps, 1972, after Aigner and Cain, 1977).

The motivation of employers to incline to the statistical discrimination is clear-cut. Employers are willing to pay workers according to their productivity. However, when an employer hires a new worker and decides about her wage he does not know her true productivity. Therefore he is looking for some kind of a measure; let us call it y, which would indicate worker's true level of skills, q. To construct such measure an employer usually has a résumé at hand, makes an interview, possibly some kind of an assessment test but those are still only proxies for the applicant's productivity and as such are not exact. This is why the employer also draws information from other

available sources such as the average productivity of workers belonging to the same socio-economic group as the applicant.

To illustrate this situation let us assume that a score of the assessment test is the only aspect measured by y and that the assessment test measures worker's productivity with some error, u. Then the measurement equation is:

$$y = q + u, \tag{6}$$

where u is $i.i.d.$ with $u \sim N(0; \sigma^2)$ and $q \sim N(\alpha; \sigma_q^2)$. We assume that the measurement error has zero mean for all workers and the average true productivity, α, differs across workers coming from different social groups. The employer is interested in y only to that extent as it predicts applicant's skills, q, that implies employers are interested in the expected value of worker's productivity, $\hat{q}$, given the observed test score:

$$\hat{q} = E(q|y) = (1 - \gamma)\alpha + \gamma y, \tag{7}$$

where

$$\gamma = \frac{Var(q)}{Var(q) + Var(u)} = \frac{Cov(q, y)}{Var(y)} = \frac{Cov(q, y)^2}{Var(q)Var(y)} = r^2, \tag{8}$$

r^2 being the squared coefficient of correlation between q and y; clearly $0 < r^2 < 1$. At this point we need to stress out that γ is the core of the statistical theory of discrimination because it represents the reliability of a test score, y, as a measure of the true skills, q. It follows from the least squares regression equation (7) that q is expressed in terms of a group effect [(1 — y)α] and an individual effect (γy) using γ as weights. *"It is useful to think of equation (7) as a conditional expectation from a linear population regression function:*

$$q = (1 — y)\alpha + \gamma y + u', \tag{9}$$

where u' is the usual well-behaved error term" (Aigner and Cain, 1977, pg. 176).

The employers often feel they still don't have enough information about the workers' true productivity after passing the assessment test. As Borjas (1998) points out, assuming that the test score would be perfectly correlated with the productivity and therefore predict the true marginal product of the workers would be very unrealistic. To patch this information gap, gain more information about the possible risks or rewards stemming from hiring this or that applicant, the employers use statistics about the

relevant groups' average performance (depending on race, gender, sex, education level, social background, etc. of the particular applicant) to predict the workers' productivity.

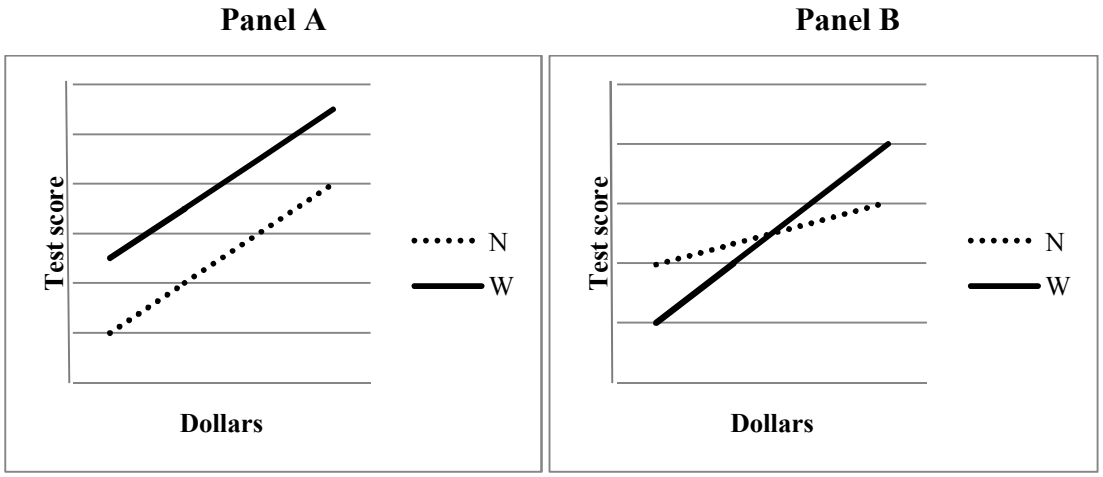
At this point, consider two groups of workers, W and N, which are differentiable. The employer then hires a worker and pays the worker the amount $\hat{q}$ according to the information available about both, the worker and the group he belongs to:

$$\hat{q}^W = (1-\gamma^W)\alpha^W + \gamma^W y^W \qquad (10)$$
$$\hat{q}^N = (1-\gamma^N)\alpha^N + \gamma^N y^N.$$

"*The slope, γ, will generally differ for the two groups if the variances of q and u do*" (Aigner, Cain, 1977, pg. 176). As mentioned earlier, γ is the reliability of the test score. Borjas (1998) interests himself in its different cases. Letting γ approach 1 gives us the case of the test score perfectly predicting the productivity and therefore the employer not being interested in the group mean at all. On the other hand, the other extreme case of γ approaching 0 means that the test score is totally irrelevant for the employer and he considers only the group mean in the process of hiring and estimating the wage.

Now, consider two cases (see Figure 3 below). In the first case slope coefficient γ being equal for groups, W and N, but group N having lower mean of the level of skills. When an employer comes across two workers with the same test score, each from one group, he will choose/pay more the one from group W, because, even though he cannot observe it, he expects the W worker being more productive according to the statistics. In the second case consider the mean of the level of skills being equal but group N having lower reliability of the test score, γ, which is usually driven by higher variance of u. Borjas (1998) notes that this situation may occur due to a "cultural bias," which is caused by the way the tests are made. Those are usually prepared by white-male academics and reflect their environment, therefore may be imprecise concerning other groups. The outcome of this case is workers from group N having relatively the same productivity and hence paid similar wages (low reliability of the score), on the opposite workers from group W rewarded accordingly to their test scores (high reliability).

Figure 3: The relationship between test score and wages with statistical discrimination



Note: In Panel A workers from both groups have the same slope test reliability but different average skill level; in Panel B workers from both groups have the same average skill level but different test reliability.

Source: Borjas (1998)

To sum up this part, employers hardly ever have all the information needed to assess perfectly worker's productivity, even though they have résumés and often also results of some assessment tests at hand. To overcome this information asymmetry and obtain some more information, employers use statistics about average performance of groups workers belong to. However, this piece of information need not necessarily follow particular worker's profile and thus may lead to his advantage or on the other hand disadvantage. As Borjas (1998) concludes, as a result of using the average group performance, the workers belonging to a high-productivity group usually benefit from their membership, whereas the workers belonging to the low-productivity group usually lose.

2 The U.S. Population and Labor Market

In the last chapter we showed, that the presence of discrimination is responsible mainly for wage differences among groups, but also that discrimination often goes hand in hand with segregation. In this chapter, we will briefly present the USA, Title VII of the Civil Rights Act of 1964, and later its labor market from the point of view of single races and the history of their migration and residence. We will also document the comparative situation of different races in the U.S. labor market using data from the U.S. Bureau of Labor Statistics.

The United States of America, situated in North America, is the third largest country in the world. "*It has the largest and technologically most powerful economy in the world*" (CIA, 20 April 2012). It is also the third most populated country in the world, the population projected to July 2012 comprises of 313 847 465 people (CIA, 20 April 2012). The American population is a great mixture of different cultures, races and ethnicities; namely for two reasons, firstly for the historical ones and secondly for the reason of economic conditions. The largest racial groups in the U.S. are whites, African-Americans (further referred to as Black), Hispanics and Asians, in order of frequency. Concerning Hispanics, estimation of their share in the U.S. population is a bit tricky as U.S. Census Bureau defines them as persons of Hispanic origin but still of any race.

2.1 Story of blacks

In 1492, when Christopher Columbus reached the coast of the Bahama Islands, the colonization of the New World slowly began. As of the early 1600s the emigration from Europe reached quite large volume. Focusing on the area of today's United States, the emigrants were mainly from England. They had usually two motives, profit or, more often, religious freedom. As the costs of transport to America were high, people often signed contracts to become for limited time indentured servants and pay the costs with their work. As the amount of new immigrants started to fall in the second half of the seventeenth century and by the time the indentured servants started to be skillful, it was time for them to leave their "lords," the settlers started to bring Black slaves from Africa.[3] At the beginning of the nineteenth century when the international slave trade was abolished, the Republican's leaders expected the slavery to fade out. Instead, with the coming Industrial Revolution it became more profitable than ever before. The situation ended up with slavery abolished in one half of the Union's states and supported in the other. This resulted in the Civil War and later the abolishment of slavery (U.S. Diplomatic Mission to Germany, 2012). Despite that, the position of Blacks remained unequal, especially in the South. People do not start to look at their slaves as at equal beings over night. This stereotype was further preserved by the "slave-like" behavior of Blacks after the slavery was abolished. It took time for them to get used to their new role in the society.

[3] Wright, 2000: The tradition of slavery in Africa dates back to the Egyptian times, Europeans only joint the slave-trade.

2.2 The Asian Story

The history of Asian influx can be characterized mainly by the example of Chinese and Japanese, who followed them. Yet, the stories are not much alike.

2.2.1 Chinese

The first large-scale in-migration of Asians to today's U.S., dated to the second half of the nineteenth century, is related to both, the Gold Rush and the economic distress in China under British dominance after the First Opium War (for details see Schaffer Library of Drug Policy, 2012). The Chinese came either to try their luck and become wealthy or as contract workers to the sugarcane plantations of Hawaii. They often became small merchants, gardeners, laundry workers, farmers and later railroad workers during the Transcontinental Railroad construction. After the Railroad was finished, the Chinese started to be perceived as an economic threat and anti-Chinese movement went on. The situation culminated in 1882 with the Chinese Exclusion Act, which prevented Chinese from becoming U.S. citizens and led to their segregation. Chinese had no other choice than found the so-called Chinatowns. Despite their exclusion they grappled for becoming *sui juris* Americans (Le, 2012a). This situation changed during the World War 2, when China was an ally of the U.S. The Chinese Americans received the right to become American citizens (Le, 2012b).

2.2.2 Japanese

At the very end of the nineteenth century, the Japanese workers were hired as cheaper replacement of the Chinese to work in Hawaii and U.S. At first, the Japanese were under the supervision of their government to make sure they are treated well. When the Japanese were exposed to the same discriminatory behavior as Chinese which culminated in 1907 with the Gentlemen's Agreement, Japan stopped issuing passports for its workers to go to the U.S. This did not decrease the number of emigrant workers. The Japanese were followed by Koreans, Filipinos and others, of which all followed the same story. They were "fighting" for their rights even more than the Chinese alone before, with no results. When the Japanese attacked Pearl Harbor, the life of Japanese in the U.S. became a nightmare. The majority of them were locked up in prison camps. Despite that they were eager to prove their loyalty to the U.S. and some even fought under the U.S. flag. After the war during the rebuilt of Japan under the baton of the

U.S., the Japanese Americans were given the right to become American citizens. In the late eighties, the Japanese population of the U.S. got an official apology for the imprisonment during the World War 2 (Le, 2012b).

After the 1965 Immigration and Nationality Act was passed and massive immigration from Asia began despite all expectations. The Asian Americans were taking advantage of the right to bring their relatives from their homelands. This had a great impact on the U.S. demographic structure and caused economic and cultural shifts. On the other hand, Europeans who were expected to immigrate to the U.S. were experiencing a post-war economic boom and therefore did not perceive the new immigration law as a great opportunity (Le, 2012c).

2.3 Hispanics

The area of today's American Southwest was since the sixteenth century colonized by Spain; the Spanish mixed with the Native Indians. Until the Mexican War (for details see Lee, 2012) it was a Spanish-speaking territory with its own customs. After the war the southwest territories joined the U.S. By that time the oppression of Hispanics by Anglo-Americans began. People living in Southwest were called Mexican Americans. Their number increased because of immigration from Mexico. The Mexicans were migrating back and forth according to the economic situation. During the World War 2 farm workers were scarce, therefore the U.S. and Mexican governments agreed on the so-called Mexican Farm Labor Supply Program, which was in force for more than twenty years and brought far more Mexicans to the States.

The U.S.-Mexican border is the world's longest border separating poverty from prosperity thus illegal immigration is a huge issue. Other immigrants come mainly from Cuba, Puerto Rico and smaller number from Latin America. Mexicans and Puerto Ricans (U.S. citizens since 1917) come to the U.S. mainly because they see it as an economic opportunity. Cubans have come to the U.S. starting from 1959 mainly for the reason of political persecution as Fidel Castro took over the governance (Guisepi (ed.), 2012).

2.4 Legal Position of Minorities

"*Title VII of the Civil Rights Act of 1964 is a federal law that prohibits discrimination in employment on the basis of sex, race, color, national origin, and religion*" (AAUW, 2012). It was initially called for by President Kennedy after a series

of Afro-American protests for emancipation, which climaxed with Birmingham Campaign in 1963 (for details see Encyclopedia of Alabama, 2007). After President Kennedy was killed, his successor, President Johnson, finished the work on this piece of legislation (United States History, 2012). Before that the social and economic conditions of Blacks were poor in the South, mainly for historical reasons. Title VII was a big step on the way to human rights protection. It guaranteed everyone equal treatment on the labor market, i.e. equal employment opportunity, job training, wages, working conditions, work assignment etc. Equal Employment Opportunity Commission was also established in order to control obeying of this federal law (U.S. Equal Employment Opportunity Commission, 2012).

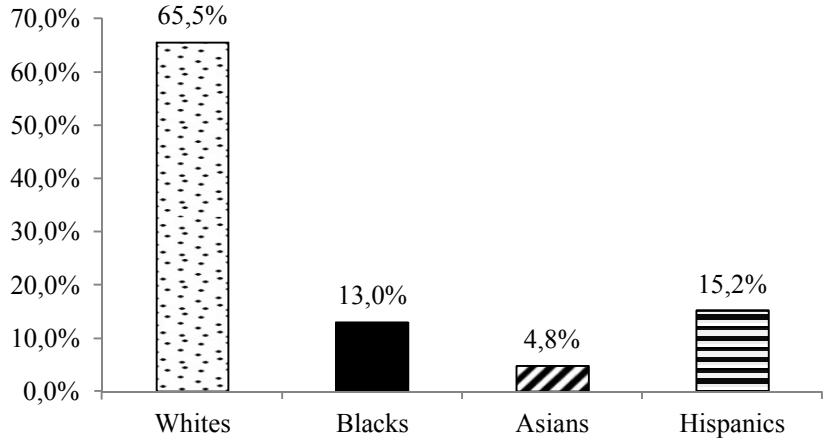
According to Donohue and Heckman (1991) there was a significant increase in Blacks' wages relative to whites' from mid-sixties. Also desegregation of firms in the South occurred in high volume. During one decade only, Blacks' economic conditions experienced great improvement which caused the outflow of Blacks from the South to stop and even led to their migration back into the region during seventies. The southern region accounted for more than two thirds of the Blacks' economic status rise.

2.5 The US Labor Market

In this part we deal with the U.S. labor market. We present several graphs which describe the market in terms of position of single races. This section will serve us as a starting point and motivation for our research. Therefore at the end of this section, after considering the races' positions from different perspectives, we will be able to formulate our research hypotheses.

According to the U.S. Census Bureau definition (mentioned above), the 2007 estimate of share of Hispanics in the U.S. population was 15.1% (CIA, 2012). In the 2010 Census they already accounted for 16.3% of the total population. In this chapter we will have to content ourselves with this interpretation but later, for the purposes of our own research, we will define Hispanics as persons of Hispanic origin and white race at the same time. Under this definition, the Hispanic population accounts for 8,3% of the entire U.S. population; whites 76.2%, Blacks 14.6% and Asians 5.6% (U.S. Census Bureau, 2011). Figure 4 represents the share of races in the U.S. working-age population in 2009.

Figure 4: Share of Races in the U.S. population in 2009, age 16-64

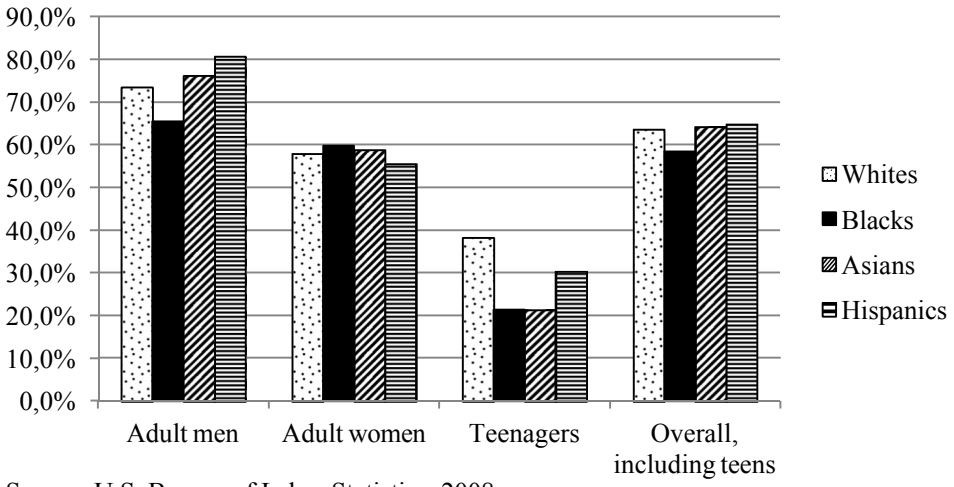


Source: U. S. Census Bureau , 2010

As we can see, whites form with 65.5% total majority of the U.S. population. The important point here is that they form not only the quantitative majority but also the economic one. The history of all three main minority groups in the U.S., namely Blacks, Hispanics and Asians, resulted in 13%, 15.2% and 4.8% of the U.S. working-age population, respectively. About one half of the Hispanics is white, the other one mostly half-caste (U.S. Census Bureau, Census 2010).

Figure 5 represents the employment-population ratio (the share of population that is employed) among the racial groups for 4 categories, namely adult men and women (age 20+), teenagers (age 16-19), and also the overall population (age 16+).

Figure 5: Employment-population ratio in 2007



Source: U.S. Bureau of Labor Statistics, 2008

Let us first assume that the position of whites is considered to be a "standard"[4] as it is the majority group in the U.S. From the figure above is evident that the overall employment-population ratio is almost the same for whites, Asians and Hispanics whereas Blacks are evidently less employed. This might be caused by their overall lower level of human capital and thus productivity, as these are generally assumed to be interlinked. Another explanation would be that it is their choice to stay out of employment. Last reasoning for Blacks' lower employment rate that comes to our mind is discrimination against them, which raises costs of hiring them, as was shown above. Thus the employers are less likely willing to hire Blacks.

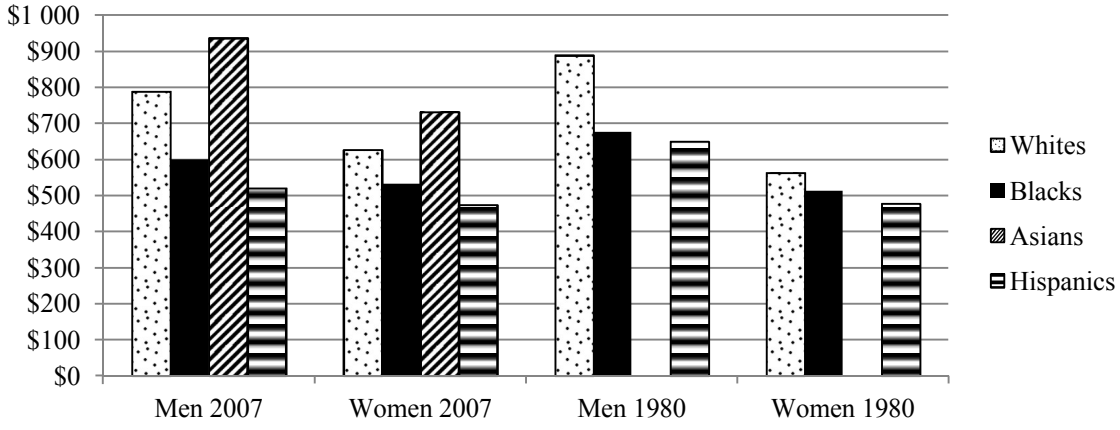
Among adult men, whites and Asians are comparably employed, between 73-76%. It is the Blacks again who are the least employed. On the other hand the Hispanics are the most employed with the employment-population ratio slightly above 80%. This might by partly caused by the fact that their wages are the lowest among these groups (see Figure 6). Hispanics have larger families than other groups (Landale, Oropesa, Bradatan, 2006), therefore it might be possible that they are willing to work for lower wage than others and also start working in lower age (quite high employment-population ration for teens compared to Blacks and Asians) in order to have enough money. Women are more or less equally employed with Blacks being the most and Hispanics the least. Concerning women, it might be partly their choice not to work in order to take care of the household, therefore this ratio is generally lower than for men.

Figure 6 represents median usual weekly earning of full time wage and salary workers for years 2007 and 1980 (data for these years are used in our research). There were no data for Asians available for 1980. 1980 U.S. dollars are adjusted to 2007 level using consumer price index (Minnesota Population Center, 2012). From the figure below it is apparent that in 1980 whites, both men and women, had the highest wages compared to other racial groups. The wages of Blacks and Hispanics were similar. Concerning the data for 2007, the wage and salary distribution has changed a lot since 1980 (Card, Lemieux, 1993), mainly for men. The class of high-income men is narrow whereas the class of low-income people has widened a lot. It is obvious that the median

[4] Becker, 1971: „*In estimating the quantitative importance of economic discrimination against members of a minority group, one needs to know what their earnings would be in the absence of different kinds of discrimination. Unfortunately, not much evidence is usually available, and somewhat arbitrary assumptions have to be made.*" (Pg. 4).

earnings must have fallen since 1980. Despite that, the relative situation has not changed much concerning women, or Black men. On the other hand, the white-Hispanic gap has increased by about 8%. Asians, both men and women, were the highest paid group within their gender; men therefore absolutely. On the other hand, Hispanics were the least paid group; again this holds for both, men and women. Black men were paid approximately the same as white women and Hispanic men approximately the same as Black women. The wage gap in white-black median weekly earnings was close to 25% of the white men's earning, the white-Hispanic gap was close to 35%. In women's case, these gaps were close to 15% and 25%, respectively. The Asians' median weekly earnings are on average 17.5% higher compared to whites'. It is evident that the median earnings differ a lot among races

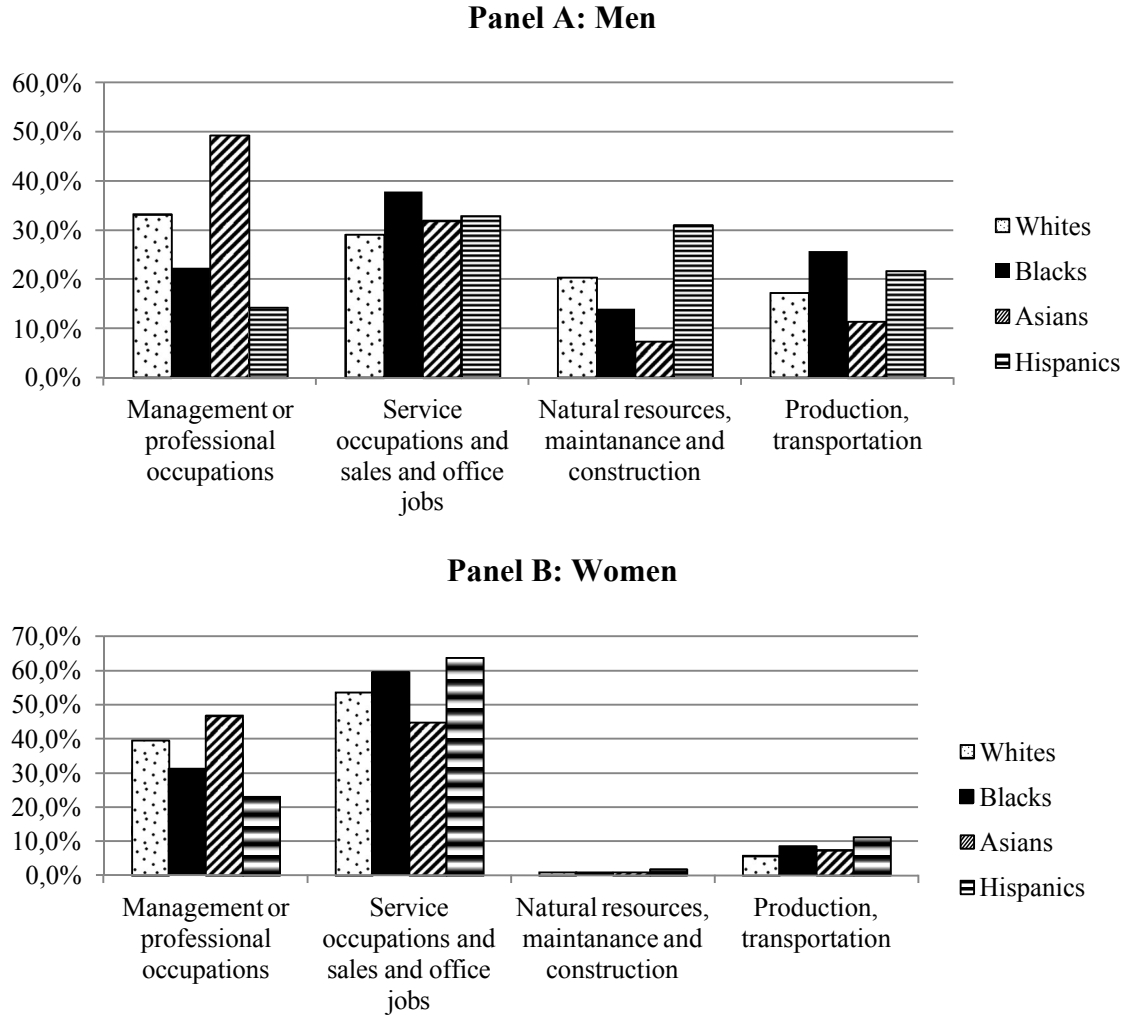
Figure 6: Median usual weekly earnings of full time wage and salary workers 2007, 1980



Source: U.S. Bureau of Labor Statistics, 2008

Figure 7 shows the employment by sector among racial groups, where men are in Panel A whereas women in Panel B. We have chosen the high-paid sector of management or professionals, the middle-paid sector of services, and sales and office jobs and on the other hand the low-paid sectors of natural resources, maintenance, construction, production and transportation in order to demonstrate the partial segregation of particular groups into low- and high-profile jobs. As we can see in Panel A below, Hispanics and Blacks are employed in middle- and low-paid jobs while Asians are employed in middle- and mainly high-paid jobs. Whites are more or less evenly distributed with about 60% employed together in middle- and high-paid jobs.

Figure 7: Employment by sectors in 2007



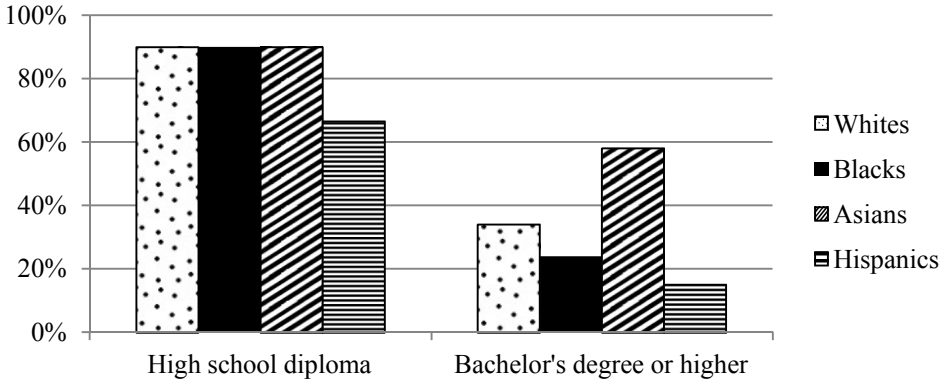
Source: U.S. Bureau of Labor Statistics, 2008

Concerning women in Panel B, the majority of them is employed in middle- and high-paid jobs, as the low-paid jobs are considered to be more of men-jobs. Nevertheless, taking into account the small percentage of women that work in low-paid jobs, mainly production and transportation, the highest share in this sector belongs to Hispanic women and after them to Black women. These two races also have the highest share in the middle-paid sector of services and sales and office jobs. On the other hand white and mainly Asian women are in majority employed in high-profile jobs. The trend is pretty much the same for both gender-groups.

Figure 8 represents the level of education among the four racial groups. As we can see, less than 70% of Hispanics finish their high school education and only 15% reach a university degree, which are the lowest figures among all the racial groups. This would to a great extent explain their employment in low-paid jobs as was shown in

Figure 8 and thus their earnings. Similar explanation is at hand for Asians but the story is totally opposite. Asians have the highest share of university degrees which would again explain their largest proportion of employment in high-paid jobs and therefore their high incomes.

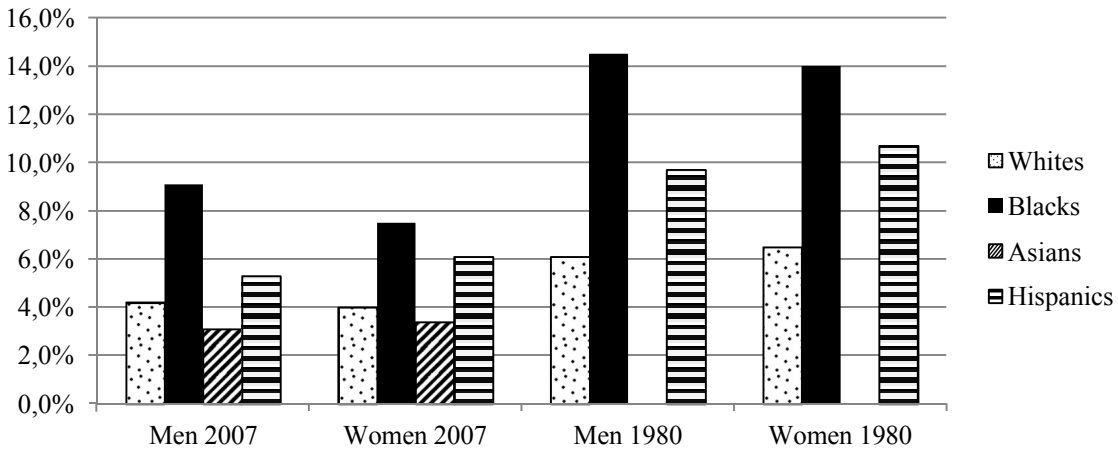
Figure 8: Levels of education in 2007



Source: U.S. Bureau of Labor Statistics, 2008

Figure 9 represents the unemployment rates among racial groups in both 1980 and 2007. While in 1980 the unemployment rates did not differ between sexes, they were generally pretty high. Blacks had the highest unemployment rate, breaking the level of 14%, after that Hispanics around 10%, whites around 6% and there are no data for Asians. In 2007 the situation was much better, the economy was at its peak with unemployment rates held low, but still the unemployment of Blacks was twice as high as that of whites. The Asians affirming their great position in the market had the lowest unemployment rate.

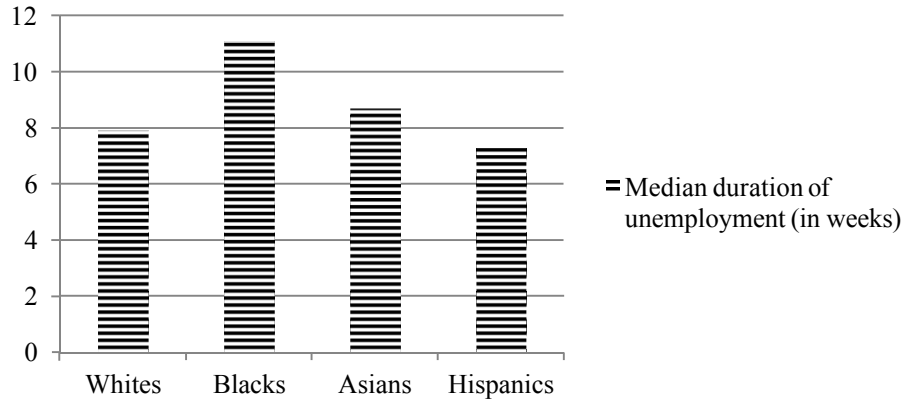
Figure 9: Unemployment rates 2007, 1980



Source: U.S. Bureau of Labor Statistics, 2008

Figure 10 represents median duration of unemployment in 2007. It shows that not only Blacks had the largest unemployment rate among all race groups but it took them on average the longest time to find another job, 11 weeks. On the second place are surprisingly Asians with 9 weeks but it might be the case that they are waiting for a job they would be satisfied with. Whites and Hispanics have the median duration 8 and 7 weeks, respectively.

Figure 10: Median duration of unemployment in 2007



Source: U.S. Bureau of Labor Statistics, 2008

To sum up this chapter, Asians have on average the highest weekly earnings. It is hypothetically so because of their higher education, which is the reason why Asians occupy much more likely high-paid jobs than other racial groups. There might also be some cultural influence but that will not be possible to scientifically account for in our research.

Hispanics have on average the lowest weekly earnings. We believe that their large families together with low level of education (which might also be one of the consequences of large families) could be the main reason why they occupy primarily low-paid jobs. Another reason for occupying mainly low-paid jobs might be that they are illegal immigrants, which is quite a big problem in the U.S.

Blacks have on average the second lowest weekly earnings, much more closer to that of Hispanics rather than to that of whites. We do not think that in their case lower level of education is the only reason. We believe that it is more a matter of their segregation into low-paid jobs which results from discrimination against Blacks and their high unemployment.

In our research we will try to explain the racial wage gaps with the help of these several factors outlined above and also some more. We suppose that these factors will

not allow us to explain all of the variance in wages among races. The rest of it will be assigned as a result of discrimination.

3 Research Question and Hypotheses

In this chapter, we go in short through our research questions and formulate our hypotheses. First of all, we are interested in finding whether the minorities are discriminated against or whether the wage gaps observed in data can be sufficiently explained by differences in the observed characteristics, such as the industry people work in, by the fact that they work in a metropolitan area, etc., but mainly the level of skills. In case we find out that there is a large part of the gap unexplained, is it so for all the minority groups? What are their relative positions? We are also interested how the situation has improved since the 1980s. Discrimination used to be a well discussed issue those days. Has the era of "information society" brought some change in the perception of minorities and improved their position in the labor market?

Our hypotheses follow from the data presented in the previous chapter. We believe that Asians have higher earnings for the reason of higher education. There might also be higher returns to their experience and on-the-job training, probably because of the stereotype prevailing about their attitude. We do not think that there will be much of the gap remaining unexplained. Although we do not know the data for Asians from the 1980s (concerning the information included in previous chapter), we think that their position improved over the years.

Similar story would hold for Hispanics but the other way round. They have the lowest educational attainment of all the racial groups. This harms them in the form of reduced income. We also expect improvement in terms of explained/unexplained ratio compared to 1980.

Concerning Blacks, we expect that the largest portion of the wage gap would remain unexplained, i.e. Blacks experience discrimination more than any other racial group. Large part of the explained portion of the gap would be accounted for by relatively low education and also working outside of the metropolitan areas. Despite that, we believe that their position in the market has improved since the 1980s.

4 Literature Review

In this chapter we go through the works of other researchers focused on the issue of racial discrimination. Firstly, we focus on researches examining wage differentials in the U.S. labor market using different survey data. Most of these past works are interested in the wage gap between white workers and Blacks; there is only little work done concerning the Hispanics and even less concerning the Asians. After that we focus on experimental researches from different countries examining equal job opportunities for minorities. These researches use the résumé audit methodology.

Oaxaca and Ransom (1994) examined the wage differentials between races and genders, namely white and Black men in the case of racial wage gap, and white men and white women in the case of gender wage gap. The sample obtained from 1988 Current Population Survey was restricted to respondents above the age of 25. As dependent variable they used the logarithm of hourly wage and controlled for potential experience, squared potential experience, years of schooling, squared years of schooling, government employment, occupation, industry, regional location, and city size. Concerning the racial differential, they detected a 22% gross wage differential which can be decomposed into productivity differential and discriminatory differential. The decomposition differs based on the assumption of competitive wage structure when there is no discrimination in the labor market. Under the assumption that Black's wage structure is the competitive one, 53.3% of the differential is assigned to discrimination and the rest to productivity differential. Taking white's wage structure as the competitive, the share assigned to discrimination is even 60.6%; the rest again to productivity differential. Assuming any other wage structure using some weights (thus a linear combination) as being the non-discriminatory one, would yield a result of the share of discrimination being between 53.3% and 60,6%. Going in this analysis further, the market discrimination coefficient was estimated to be 0.1253 and 0.1456, respectively, depending on the competitive wage structure. This brings us to the conclusion that taking the Black's wage structure as the competitive, white workers are overpaid by almost 13% and their productivity advantage is about 11%. On the other hand, taking the white's wage structure as the competitive, Blacks are underpaid by nearly 15%, whereas the productivity advantage of whites is about 9%. Concerning the gender wage gap, a gross wage differential of nearly 30% was detected. Using the same methodology as in the case of racial wage gap, the share of discrimination under the

male wage structure was estimated to nearly 77%, the rest is again assigned to productivity differential. Under the female wage structure the share of discrimination was estimated to nearly 93%. The market discrimination coefficient under male and female wage structure was estimated to 0.2575 and 0.3194, respectively. This yields the result of women being underpaid by nearly 26%, or on the other hand men being overpaid by nearly 32%.

Oaxaca (1973) examined the gender wage differentials for people of age 16 or more living in urban areas. He did so separately for Blacks and whites using the 1967 Survey of Economic Opportunity data. As usually the dependent variable used for the ordinary least squares estimation was logarithm of (hourly) wage, the independent variables were: experience (linear and squared), education (linear and squared), class of worker, industry of the worker, occupation, part-time job, health problems, migration, marital status, number of children, region and size of urban area. The wage gap for whites is 43%, for Blacks nearly 40%. Concerning whites, using the male wage structure as the competitive one the share of the gap accounted for by discrimination is 53% with the discrimination coefficient of 0.25, using the female wage structure 64% with the discrimination coefficient of 0.32. For Blacks the shares are 50% and 61%, respectively, and the discrimination coefficients 0.22 and 0.27, respectively. Leaving out the class of worker, industry and occupation variables, the estimates of share of discrimination as part of the wage gap increased. In case of whites to 77.1% under the female wage structure as the competitive one, and to 78.4% in case of male wage structure. Concerning Blacks, the increase is much larger. 98.5% of the gap is assigned to discrimination under the female wage structure and 88.7% under the male wage structure.

Reimers (1983) examined the wage differentials between white males and Black males, and between white males and Hispanic males (for different areas of origin) of age 14+ using the 1976 Survey of Income and Education data. As explained variable she used logarithm of wage (adjusted for local price levels), as explanatory variables she used education, potential work experience (linear and squared), nativity and date of immigration, fluency in English, U.S. military service, health disability, race, government employment and inverse Mills ratio as the results were adjusted for selection bias using the Heckman method (probit and after that ordinary least squares estimation). To decompose the observed logarithm wage differential she used the

Oaxaca decomposition. As the competitive wage structure she used for comparison several of them, the Blacks' wage structure, whites' wage structure and also their linear combination with the coefficient of 0.5. The logarithm wage differential, corrected for selection bias, between white males and Black males was estimated to be 22.9%, of which 57.6% - 62% are assigned to discrimination depending on the chosen competitive wage structure. Concerning Hispanics, in the worst position, in the sense of observed wage differential, are Mexicans with the logarithm wage differential, corrected for selection bias, of 33.9%, of which 15% – 22.4% are attributed to discrimination. In the case of Puerto Ricans is the estimated logarithm wage gap, corrected for selection bias, 33.2% of which 53.3% is assigned to discrimination under all three competitive wage structures. That is the greatest share among Hispanic groups. The number of respondents originating from Cuba and South and Central America was so small that the results are uncertain therefore we do not include them here. Concerning other Hispanics, the wage gap, corrected for selection bias, was estimated to be 22.5%, of which 47% - 59% are attributed to discrimination.

Cotton (1988) examined the wage differentials among white males, Black males, white females and Black females using the 1980 Census data (the same as we use) for people of age 16 or more. For the Oaxaca decomposition he assumed as the competitive wage structure a structure yielded by linear combination of the two particular groups' wage structures (he is the one to introduce this method). The explained variable is logarithm of hourly wage. As explanatory variables he used years of schooling, potential experience (linear and squared), marital status, number of marriages, products of potential experience (linear and squared) and both marital dummies (i.e. 4 variables), urban residence, region, set of industry dummies, birth place, language spoken, veteran status and government employee. He discovered a wage gap of 46% between white male and Black female, of which 41.1% are attributed to skill differential, 14.4% to advantageous treatment of white males and 45.5% to discrimination. The share for white male-black male differential were estimated as follows: the differential itself was estimated to 17.4%, of which 73.8% are assigned to skill differential, 35.3% to advantageous treatment of whites and -9.1% to disadvantageous treatment of Blacks, i.e. they are paid 9.1% more than it would correspond to their skills, yet still there is a 26% disadvantage compared to whites. Concerning the white female-Black female gap, the estimates are as follows: 11.2% gap observed, 51.6% of the gap attributed to skill

differential, -137.9% assigned to advantageous treatment of white women ("-" means that the "advantage" is more of a disadvantage) and 186.3% attributed to disadvantageous treatment of Black women, i.e. both, white and Black women, are treated worse than it would correspond to their skills, yet still Black women 50% worse. Concerning gender gap, the observed white male-female logarithm wage gap is nearly 35%. 37.7% of this gap is attributed to skill differential, nearly 18% to the advantageous treatment of males and more than 44% to disadvantageous treatment of females. In the case of Blacks, the gap is almost 29% of which is only 21% assigned to skill differential. 5.6% is attributed to preferential treatment of males and the rest, more than 73%, to disadvantageous treatment of females. To conclude Cotton's work, there is 60% of the white male-Black female logarithm wage gap unexplained with the distribution depending only on the competitive wage structure assumed. Again, the same number concerning the white male-Black male logarithm wage gap is 26%; concerning white female-Black female gap, nearly 50%; for white male-female gap is the unexplained share of differential more than 62%, for Blacks nearly 79% . The situation is therefore much worse for women.

Bertrand and Mullainathan (2004) ran the first experimental research we deal with. The methodology is pretty much the same for the rest of them so we explain it only once. It is called a résumé audit methodology and it was developed to assess the existence of equal job opportunity for minorities. The idea behind it (as it is difficult to run an economic experiment with ordinary people in their real life) is to create a so called résumé bank, i.e. hundreds of résumés that are pretty much the same in terms of skills, education and other work requirements but differ in some attribute that we want to test whether it is discriminated against (mostly gender or race). These résumés are then sent to different employers according to help-wanted ads in newspapers and the rate of call-back for interview is measured. The results are then compared between groups.

Bertrand and Mullainathan (2004) ran their experiment in Boston and Chicago, and altogether sent about 5000 résumés; some of higher quality, some of lower quality. As race is not stated in résumés, they assigned randomly some résumés strictly white-sounding names and others strictly Black-sounding names. A statistically significant call-back gap was observed. While whites need to send about 10 résumés for one call-back, Blacks need to send about 15 of them. The gap widens with résumé quality. On

the other hand, the call-back rate increases equally with wealthier, whiter and better educated neighborhood. Return to white-sounding name was estimated to be equivalent to another 8 years of experience.

Carlsson and Rooth (2007) ran similar experiment in Sweden, where surveys among public showed negative attitude toward immigrants. The worst attitude is toward Middle Eastern immigrants which copy their unemployment rates. Those are several times higher than that of natives. The question is whether it is the effect of employers' discrimination or it is given by immigrants' productivity. The researchers sent résumés to about 1550 companies in the area of Stockholm and Gothenburg. They sent always two of them, one with male Swedish name, and the other with male foreign name. The Swedish-name résumés had on average 50% higher call-back rate. Surprisingly, low-income jobs with majority of immigrant employees had higher call-back for Swedish-name résumé than for the foreign one. Probit regression analysis showed, that the difference in call-backs depends to great extent on number of employees in the company, i.e. its size, and on gender of the recruiter.

Banerjee, Bertrand, Datta and Mullainathan (2009) ran similar research in Delhi, India with focus on possible discrimination against non-upper-caste workers (51% of the population) and mainly Muslims in fast growing sectors of software and call-centers. The caste/religion was assigned to a résumé randomly by a specific last name. Similarly to Bertrand and Mullainathan (2004), there were high- and low-quality résumés. Altogether were these résumés sent as a response to 371 job ads of which 265 were software jobs and 106 call-center jobs. No evidence was found in case of non-upper-caste for lower call-back rate for software jobs compared to equally skilled upper-caste applicants (these were called back slightly less often). Compared to upper-caste applicants, there was also found no evidence in case of Muslims for lower call-back rate for any of the two types of job. The only difference found, what is more, large (62%) and statistically significant, was in case of non-upper-caste's applications for call-center jobs. These jobs are demanding for high skill of English, sociable behavior, knowledge of American culture and other soft skills, which are expected to be found mostly within the upper-caste, and surprisingly Muslims. There was also found a difference in call-back rate between genders concerning call-center jobs. Men were more likely to be called back than women. Nevertheless, the results concerning call-centers may be misleading as the sample size was rather small.

Kaas and Manger (2011) ran similar experiment in Germany focusing on student internships. In Germany is a large Turkish minority (3% of the total population) and the researchers were interested whether Turkish have to deal with employers' discrimination against them. The researchers sent more than 1000 application to different firms offering internships to students of economics and management. Again, the applications sent to a firm were pretty much the same but one with a Turkish name and photograph and the other with a German name and a photograph. The "Turkish applicant" was born and raised in Germany. He does not prove to be a Turkish apart from the name. Similarly to the Bertrand and Mullainathan (2004) experiment where there were résumés of higher and lower quality, in this case some of the applications contained a reference letter. On average an application with a German name received 14% more call-backs than a Turkish-name application. This number was 10 percentage points higher concerning small companies (less than 50 employees). The presence of a reference letter in the application made a big difference, the call-back rate differed only by 0.5% in favor of German-name applications.

To sum up this chapter, we can see in the literature that there is discrimination against minorities both in terms of wages and willingness to employ them.

5 Empirical Project

In previous chapter, we went through the works of several other researchers on the topic of wage gaps and discrimination, mainly concerning the Afro-American minority. This chapter deals with the same issue but in broader context. Not only that we take into account also Asian and Hispanic minorities but we are also interested in the possible change that occurred in perception of minorities over the last nearly thirty years. This change might be the greater the faster the life is in these days of "information society".

As indicated above, we are interested in the wage gaps as of years 2007 and 1980, and the change that occurred between them. The year 2007 was chosen on purpose of prevention from possible bias in our results due to the global financial crises of 2007-2012, which has certainly affected also the labor market. The year 1980 was chosen as an appropriate point of time, because it was long enough after the 1963-1965 protests and legal changes to take effect in the U.S and far enough in the past at the same time. We use the American Community Survey data downloaded from IPUMS

(Ruggles and collective, 2010) that are available for 1% of the U.S. population. We restricted it to 16-64 years old people. It is considered to be a random sample. The data for years 2007 and 1980 are fully comparable. To be able to generalize our results for the whole U.S. population we use probability weights in our models.

5.1 Conceptual Framework and the Data

In this part, we suggest variables to use in our models and give some economic reasoning for them. Unfortunately, not all of them are available therefore we have to manage without some of them.

From discussion presented in section 2.5 it follows that, in the market place, discrimination against a minority group occurs from our point of view if and only if there is a part of the observed wage gap that cannot be explained by differences in various features. Therefore as our dependent variable we choose logarithm of weekly wage. The 1980 wages are adjusted to 2007 level using consumer price index (Minnesota Population Center, 2012). Wages for both years, 1980 and 2007, are top coded. The 2007 wages at the level of 99.5 quantile in each state; the 1980 wages at the level of $75000. We use flag available in the dataset to tell apart those wages that are top coded from those that are not and finally, we adjust the top coded wages by multiplying them by coefficient 1,4 (Card and DiNardo, 2002).

As also mentioned above, we are interested in the situation of several minorities, namely Blacks, Asians and Hispanics[5], therefore we include dummy variables for them. We also include dummy variables for gender, namely female, marital status, and different years.

Human capital theory claims that "*differences in investments in education and on-the-job training are the major causes of wage differences*" (Cotton, 1988). In other words, education and experience are the main determinants for productivity and performance. That is the reason why we employ in our models dummy variables for different levels of schooling (primary school being the benchmark) and also a constructed variable for potential years of experience[6] as proxies for human capital. The experience is included also in quadratic form as it is expected that it has decreasing rate of return. On the other hand, performance and productivity are expected to decrease

[5] Defined as people with Hispanic origin but of white race only.

[6] Cotton, 1988: Experience = Age – Years of Schooling - 6.

with any kind of disability affecting one's work. For that reason we include also a disability dummy variable.

Another variable we include is a dummy for metropolitan area. It is expected that people working in large cities and adjacent areas earn more because that is where large and strong companies usually have a seat. These companies prefer to employ qualified work force and also to keep them employed at themselves to lower their training costs. Another set of variables is the set of dummy variables for industries people work in. Profits of companies differ depending on the industry they belong to and also on the position of their industry within a business cycle. The higher profit the more money can be distributed among workers in the form of wage[7] (Flek, 2012a). Concerning this set of variables, finance and real estate industry is used as a benchmark.

In our research we face the danger of selection bias. Not all the people in our sample have a wage; however, taking into account only those who do would bias our results. To deal with this problem we use the Heckman method (Heckman, 1976, after Wooldridge, 2002) and define an auxiliary variable (called work) equal to 1 for all people who have a wage, work at least one week a year and belong to wage and salary sector. We will then explain this variable and use the outcome further, i.e. adjust for the reasons why some people don't have a wage. Aside from the variables described above we furthermore use two other variables in the selection model; number of children and a variable equal to 1 for a person who is a head of family. The idea behind these two variables is as follows: if a person is a head of family we expect him to earn most of the family income. We at the same time expect that the larger number of children the lower probability for women to work. For men would hold the opposite, therefore we use this variable twice – once alone and once in interaction with the gender dummy variable.

Concerning variables we would like to employ into our models but are not available; first that comes to our mind is a trade union membership. Such people usually have higher wages thanks to the bargaining power of unions and thus have a comparative advantage compared to workers from non-unionized industries (Flek, 2012b). This variable was not available in our dataset at all. Another variable we would like to use is a dummy variable equal to 1 for every person from the southern states of the United States (former Confederacy). The idea behind this is that people from the South might be more likely to discriminate against Blacks for historical reasons. In our

[7] Different bonuses are not taken into account.

dataset only 15% of respondents stated the state of their residence and as such was this variable useless.

Tables of summary statistics and tables of coefficients of correlation for our variables are available in Appendix A.

5.2 Econometric Models and Estimation Methods

In previous part we went through a brainstorming and economic justification concerning the possible variables. We also slightly touched the problem of selection bias we need to deal with. In this part we go into details concerning this issue. We also suggest several possible models explaining the wage variance in the whole U.S. population. After considering these several models we pick the one that explains best the wage variance. This model is the point of departure for our further work, the Oaxaca decomposition.

5.2.1 Selection Bias

Let us use the following notation: bold signs will be used to mark vectors, signs in italics to mark variables. Consider a population model explaining wage and satisfying the CLM assumptions:

$$\log(wage) = \boldsymbol{X\beta} + u\,, \tag{11}$$

where $\boldsymbol{X}$ is the vector of variables, $\boldsymbol{\beta}$ is the vector of coefficients and $u \sim N(0, \sigma^2)$ is the disturbance term. As it is usually the case, not all the people in our sample have a positive wage. The problem is that $log(wage)$ for $wage = 0$ is not defined and therefore we need to impose a condition $wage > 0$. However, when taking into account only the observations with nonzero wage, a selection bias might occur. Employed people might have higher wages than those out of labor force would be offered. We need to deal with incidental truncation here. According to Heckman (1976, after Wooldridge, 2002) we need to add a selection equation:

$$work = 1[\boldsymbol{z\alpha} + v \geq 0]\,, \tag{12}$$

where $work = 1$, if $wage > 0$, and zero otherwise, $v \sim N(0, \sigma_v^2)$,

$$\boldsymbol{z\alpha} = \boldsymbol{X\beta} + \boldsymbol{w\gamma}\,. \tag{13}$$

In this context v is a random disturbance of the true wage equation. $\boldsymbol{z}$ is a set of variables $\boldsymbol{X}$ from the log $(wage)$ equation and variables $\boldsymbol{w}$ added to the seletion equation. Thus $\boldsymbol{z}$ is richer than $\boldsymbol{X}$. Let us assume that $E(u, v|\boldsymbol{z}) = 0$, then

$$E(\log(\text{wage})|\boldsymbol{z}, v) = \boldsymbol{X\beta} + E(u|\boldsymbol{z}, v) = \boldsymbol{X\beta} + \rho v\,, \tag{14}$$

where $\rho v = E(u|v)$ for some parameter ρ. As v is related to $work$ by equation (12), it holds that

$$E(\log(wage)\,|\boldsymbol{z}, work = 1) = \boldsymbol{X\beta} + \rho\lambda(\boldsymbol{z\alpha})\,, \tag{15}$$

where $\lambda(\boldsymbol{z\alpha}) = \frac{\phi(\boldsymbol{z\alpha})}{\Phi(\boldsymbol{z\alpha})}$ is the inverse Mill's ratio evaluated at $\boldsymbol{z\alpha}$, when $work = 1$. We don't know $\boldsymbol{\alpha}$ but we can estimate it using the entire sample, because $work$ given $\boldsymbol{z}$ follows under our assumptions a probit model

$$P(work = 1|\boldsymbol{z}) = \boldsymbol{\Phi}(\boldsymbol{z\alpha})\,. \tag{16}$$

If $\rho = 0, u$ and v are uncorrelated, and the population model can be consistently estimated using only the selected sample. Otherwise the selection bias occurs because we omitted the variable $\lambda(\boldsymbol{z\alpha})$.

5.2.2 Inference of the Model

Using the variables suggested above we need to arrive at the most appropriate and powerful model possible. We propose three different models. The results of our maximum likelihood estimation[8] using the 2007 and 1980 data are reported in Table 1 and Table 2 below, respectively (results for 1980 data together with those for 2007 can be found also in Appendix B). The estimates are already adjusted for the selection bias using the Heckman method described above[9]. The estimates of coefficients are reported in the first line beside each variable. Their significances are reported using stars; three stars for significance at 1% level, two stars for significance at 5% level, and one star for significance at 10% level. Robust standard errors are reported in parentheses under the

[8] It is more efficient than OLS estimation.

[9] The selection equation was estimated in the form: $\text{work} = \text{F}\big(\alpha + \beta_1\text{black} + \beta_2\text{asian} + \beta_3\text{hispanic} + \beta_4\text{female} + \beta_5\text{disability} + \beta_6\text{experience} + \beta_7\text{experience}^2 + \beta_8(\text{secondary school}) + \beta_9\text{college} + \beta_{10}\text{married} + \beta_{11}(\text{number of children}) + \beta_{12}\text{female} * (\text{number of children}) + \beta_{13}(\text{head of family})\big)$, where F is a probabilistic function.

corresponding estimate of a coefficient. This convention will be used for reporting also in the rest of this work. Probability weights are used for each model estimation.

Table 1: Estimates of possible population models using 2007 data

	Model 1	Model 2	Model 3
Black	-0.011***	-0.025***	-0.019***
	(0.0028)	(0.0028)	(0.0027)
Asian	0.059***	0.031***	0.031***
	(0.0042)	(0.0042)	(0.0041)
Hispanic	-0.029***	-0.051***	-0.055***
	(0.0026)	(0.0026)	(0.0025)
Female	-0.326***	-0.323***	-0.241***
	(0.0017)	(0.0017)	(0.0018)
Experience	0.069***	0.067***	0.062***
	(0.0003)	(0.0003)	(0.0003)
Experience^2	-0.001***	-0.001***	-0.001***
	($6.29e^{-6}$)	($6.39e^{-6}$)	($6.4e^{-6}$)
Secondary School	0.272***	0.267***	0.262***
	(0.0033)	(0.0034)	(0.0034)
College	0.865***	0.838***	0.846***
	(0.0037)	(0.0037)	(0.0038)
Married	0.135***	0.139***	0.124***
	(0.0018)	(0.0018)	(0.0018)
Disability	0,084***	0.109***	0,08***
	(0,0087)	(0.0089)	(0.0091)
Metropolitan		0.188***	0.181***
		(0.0018)	(0.0017)
Professional			-0.117***
			(0.0039)
Information and Communications			-0.068***
			(0.0057)
Retail Trade			-0.381***
			(0.0038)
Agriculture			-0.308***
			(0.0086)
Utilities			0.101***
			(0.0067)
Manufacturing			-0.056***
			(0.0036)
Wholesale Trade			-0.09***
			(0.0049)
Services			-0.379***
			(0.0032)
Army			0.007
			(0.0077)
Construction			-0.042***
			(0.0042)
Mining			0.305***
			(0.0107)
Transportation			-0.104***
			(0.0045)
Public Administration			-0.046***
			(0.0041)

Constant	5.507***	5.402***	5.619***
	(0.0046)	(0.005)	(0.0061)
Observations[10]	1346041	1346041	1346041
Log pseudolikelihood	-2.64e^{8}	-2.63e^{8}	-2.6e^{8}

Note: The estimates of coefficients are reported in the first line for each variable. Their significance is reported using stars; three stars for significance at 1% level, two stars for significance at 5% level, and one star for significance at 10% level. Robust standard errors are reported in the parentheses under the corresponding estimate of coefficient. Probability weights are used for the estimation.

The first, simplest, model can be thought of as a model consisting of only individual characteristics, therefore it would reflect pure discrimination. The estimated coefficients, all significant at 1% level, show that while other characteristics being equal, being Black means on average a 1%[11] lower income compared to a white worker. In case of Hispanics the gap is larger. They have on average nearly 3% lower incomes than whites. On the other hand, the estimate for Asians corresponds with the statistics shown above in section 2.5; they have *ceteris paribus* on average 6% higher incomes compared to whites. What is alarming is the estimate for gender. According to that a woman with the same characteristics as a man has more than 28% lower wage. Interestingly, the estimate of coefficient for variable $disability$ is positive and nearly 0,085, meaning that people with disability, yet working, have on average 8.7% higher wages than comparable people without any disability. This could be positive selection to work; only the smartest disabled would work.

In the second model we additionally included the variable $metropolitan$, which brings into the model the effects of agglomeration. Plugging this variable into the model changed quite a lot the estimated coefficients for different race groups, besides the return to education, mainly college, and disability. The estimated gap between whites and Asians declined to 3%, whereas the gap between Blacks and whites, and Hispanics and whites increased to 2.4% and 5%, respectively. The variable $metropolitan$ was clearly a missing endogenous variable in model 1, it biased the results up.

In the third, most complicated, model we included moreover the set of industry dummy variables in order to bring in the industry effects. Employing these variables again changed the estimates. The coefficient for Asians remained at its previous level of

[10] For selection equation was used 1905194

[11] Arrived at as $e^{coefficient} - 1$.

3%, but the coefficients for Blacks and Hispanics changed; the Blacks' decreased back to 1.9% whereas the Hispanics' increased furthermore to 5.3%. The coefficient for women decreased significantly from 28% to 21%. Also the return to being married decreased by 1.7%, and the coefficient for disability returned back to 8% from the first model. This is a sign of strong industry sorting. Concerning the industry dummy variables (finance and real estate being the benchmark), return to being employed in any of the industries apart from utilities and mining industry are negative. The remaining two might have straightforward explanation for their higher return. Mining for the reason that it is dangerous industry and thus there must be some premium for the workers. Utilities as it is network industry with increasing return to scale and thus large profits, which can be further divided among the employees.

Table 2: Estimates of possible population models using 1980 data

	Model 1	**Model 2**	**Model 3**
Black	-0.005*	-0.005*	-0.014***
	(0.0029)	(0.0029)	(0.0028)
Asian	0.016***	0.014**	0.024***
	(0.0062)	(0.0062)	(0.0060)
Hispanic	-0.017***	-0.019***	-0.02***
	(0.0036)	(0.0036)	(0.0036)
Female	-0.391***	-0.391***	-0.32***
	(0.0017)	(0.0017)	(0.0018)
Experience	0.048***	0.048***	0.043***
	(0.0002)	(0.0002)	(0.0002)
Experience^2	-0.001***	-0.001***	-0.001***
	($5.42e^{-6}$)	($5.41e^{-6}$)	($5.28e^{-6}$)
Secondary School	0.189***	0,187***	0.179***
	(0.0022)	(0.0022)	(0.0022)
College	0.576***	0.571***	0.599***
	(0.0028)	(0.0028)	(0.0029)
Married	0.146***	0.147***	0.124***
	(0.0018)	(0.0018)	(0.0018)
Disability	0.157***	0.159***	0.155***
	(0.0046)	(0.0046)	(0.0045)
Metropolitan		0.048***	0.048***
		(0.0014)	(0.0014)
Professional			-0.109***
			(0.0029)
Information and Communications			0.285***
			(0.0051)
Retail Trade			-0.227***
			(0.003)
Agriculture			-0.229***
			(0.0066)
Utilities			0.145***
			(0.0054)

Manufacturing			0.132***
			(0.0027)
Wholesale Trade			0.066***
			(0.0039)
Services			-0.188***
			(0.0037)
Army			0.159***
			(0.0071)
Construction			0.119***
			(0.0038)
Mining			0.322***
			(0.0065)
Transportation			-0.072***
			(0.0054)
Public Administration			0.026***
			(0.0038)
Constant	5.948***	5.937***	5.96***
	(0.0029)	(0.0029)	(0.0038)
Observations[12]	999131	999131	999131
Log pseudolikelihood	-1.86e^{8}	-1.86e^{8}	-1.84e^{8}

Note: The estimates of coefficients are reported in the first line for each variable. Their significance is reported using stars; three stars for significance at 1% level, two stars for significance at 5% level, and one star for significance at 10% level. Robust standard errors are reported in the parentheses under the corresponding estimate of coefficient. Probability weights are used for the estimation.

Concerning the estimates for year 1980, the coefficients for race dummy variables are smaller in magnitude compared to year 2007 for all three models. They are also quite stable across these three models. On the other hand, the coefficient for gender dummy variables is much larger in magnitude compared to year 2007 (it is the smallest in magnitude in model 3). This suggests that the gender wage gap has narrowed since the year 1980, but racial wage gaps have widened. The coefficients for experience and education are much smaller in magnitude compared to 2007. This means that during those 27 years, the return to experience and education increased a lot, namely by 2 percentage points for experience, by 10 percentage points for high school education and by 50 percentage points for university education. Also the return to working in metropolitan area increased quite a lot, by 14 percentage points. This could have been expected. On the other hand, return of being disabled (if we can call it so in this case) has decreased by nearly 50%. The returns of working in industries other than finance and real estate have decreased significantly but this was expected following the role finance has these days.

[12] For the selection equation was used 1442342 observations

While deciding which of the proposed models to use further for the Oaxaca decomposition, we followed mainly the log pseudolikelihood measure. This led us to the third, most complicated, model, which besides individual characteristics, so to speak, takes the advantage of explaining the wage variance using the industry an individual works in. The Inverse Mills ratio was significant in all the models therefore we had to estimate it using the Heckman method to avoid the danger of selection bias. It is apparent from the results for both years that there is a wage gap between white workers and other races (positive for Blacks and Hispanics and negative for Asians), yet not large. Nevertheless, it may get larger while estimating the model separately for each race as we did not use interactions in the models above. Detailed comments will be delivered in the last subchapter after the Oaxaca decomposition.

5.2.3 Oaxaca Decomposition

At this point we come to the core of our research, which is the so called Oaxaca decomposition. This is a method developed by Ronald L. Oaxaca in the seventies, which allows us to decompose the wage gap found in the data into two parts; the part explained by variables in our model and the part remaining unexplained. The unexplained part is then assigned to be the expression of discrimination.

To quantify the share of observed wage gap that is caused by discrimination against a minority, Becker (1971) points out that the wage structure which would be present in the absence of discrimination need to be known. He also adds that not much evidence is usually available concerning the non-discriminatory wage structure. Since that time a great progress has been made. Oaxaca and Ransom (1994) discuss several approaches that are often used and furthermore suggest a weighting matrix to deal with this issue. This method is, however, quite advanced and therefore we assume in our research that the wage structure of white employees is the non-discriminatory wage structure and also that white workers are not given any preferential treatment. Under these assumptions, it would hold that minorities would have the same wage structure as whites if there were no discrimination against them. Despite that, a wage gap can still be present.

Oaxaca (1973) picks up the threads of Becker's work (see chapter 1) and introduces generalized concept of discrimination coefficient (from now on referred to as D):

$$D = \frac{\frac{W_w}{W_m} - \left(\frac{W_w}{W_m}\right)^0}{\left(\frac{W_w}{W_m}\right)^0}, \tag{17}$$

where $\frac{W_w}{W_m}$ is the observed majority-minority wage ratio with w representing the majority and m the minority, and $\left(\frac{W_w}{W_m}\right)^0$ is the majority-minority wage ratio in the absence of discrimination. An equivalent expression in natural logarithms is:

$$\log(D+1) = \log\left(\frac{W_w}{W_m}\right) - \log\left(\frac{W_w}{W_m}\right)^0. \tag{18}$$

As D is unknown, estimating it is essentially the same as estimating $\left(\frac{W_w}{W_m}\right)^0$. So what we need to do is estimate the final model:

$$\log(W) = \boldsymbol{X\beta} + u \tag{19}$$

that we arrived at in previous part with OLS estimation method, separately for each race. The wage differential can be decomposed into two parts; the differential resulting from differences in characteristics (**X**) and the other part resulting from discrimination (differences in $\boldsymbol{\beta}$ among races). Let

$$G = \frac{\bar{W}_w - \bar{W}_m}{\bar{W}_m}, \tag{20}$$

where $\bar{W}_w$ and $\bar{W}_m$ are the average weekly wages of the majority and the minority, respectively. Then

$$\log(G+1) = \log(\bar{W}_w) - \log(\bar{W}_m). \tag{21}$$

From the properties of the OLS estimation it follows that

$$\log(\bar{W}_w) = \bar{\boldsymbol{X}}_{\boldsymbol{w}}\hat{\boldsymbol{\beta}}_{\boldsymbol{w}} \tag{22}$$

$$\log(\bar{W}_m) = \bar{\boldsymbol{X}}_{\boldsymbol{m}}\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}, \tag{23}$$

where $\bar{\boldsymbol{X}}_{\boldsymbol{w}}$ and $\bar{\boldsymbol{X}}_{\boldsymbol{m}}$ are the vectors of mean values of the independent variables for majority and minority, respectively, and $\hat{\boldsymbol{\beta}}_{\boldsymbol{w}}$ and $\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}$ are the vectors of corresponding estimated coefficients.

Substituting equations (22) and (23) into equation (21) gives us

$$\log(G+1) = \bar{\boldsymbol{X}}_{\boldsymbol{m}}\hat{\boldsymbol{\beta}}_{\boldsymbol{m}} - \bar{\boldsymbol{X}}_{\boldsymbol{w}}\hat{\boldsymbol{\beta}}_{\boldsymbol{w}}. \tag{24}$$

Let us define

$$\Delta\bar{\boldsymbol{X}} = \bar{\boldsymbol{X}}_{\boldsymbol{w}} - \bar{\boldsymbol{X}}_{\boldsymbol{m}} \tag{25}$$

$$\Delta\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\boldsymbol{m}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{w}}. \tag{26}$$

Substituting equations (25) and (26) into equation (24) drives us to

$$\log(G+1) = \Delta\overline{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{w}} - \overline{\boldsymbol{X}}_{\boldsymbol{m}}\Delta\widehat{\boldsymbol{\beta}}\,. \tag{27}$$

Under the assumption of white workers' wage structure being the non-discriminatory one from equation (18) follows that

$$log\left(\frac{\widehat{W_w}}{W_m}\right)^0 = \Delta\overline{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{w}} \tag{28}$$

$$\log\left(\widehat{D+1}\right) = -\overline{\boldsymbol{X}}_{\boldsymbol{m}}\Delta\widehat{\boldsymbol{\beta}}\,. \tag{29}$$

"*Thus expressions (28) and (29) represent the decomposition of the wage differential into the estimated effect of differences in individual characteristics and the estimated effect of discrimination, respectively*" (Oaxaca, 1973, pg. 696).

5.3 Results

In this subchapter we deliver comments on the estimates of model 3 from section 5.2.2, which we estimated by maximum likelihood method separately for each of the racial groups, whites, Blacks, Asians and Hispanics, therefore the race dummy variables from model 3 are left out. After that, we present output of the Oaxaca decomposition for which we used the estimated models.

Concerning the separate models, for all of them, i.e. for all four races and both years, the coefficient for Inverse Mill's ratio was significantly different from zero and therefore we had to use the Heckman method in order to avoid the selection bias. All the estimations are made using probability weights. Tables containing estimates of all eight models (4 races x 2 years) can be found in Appendix C, sorted by year and also by race in order to facilitate the reader comparing them. Tables containing detailed Oaxaca decompositions can be found in Appendix D.

5.3.1 Estimates for year 1980

Concerning whites, all the estimates are significant at 1% level. The coefficient for $college$ was estimated to nearly 0.62 meaning that obtaining a university degree would on average while keeping everything else constant bring a 86% higher wage compared to finishing only an elementary school. The coefficient for $secondary\ school$ was estimated according to our expectation smaller. The return to finishing high school is 21% meaning that by this percentage would *ceteris paribus* increase one's wage compared to finishing only the elementary school. The coefficients

for $experience$ and $experience^2$ were estimated 0.045 and -0.0006[13], respectively. This means that the return to another year of experience is driven by formula $-0.0006\ experience^2 + 0.045\ experience$. When we take the derivative and put it equal to zero it brings us to finding that the return to another year of experience is increasing up to the circa 37,5th year of experience. The growth rate is $-0{,}0012\ experience + 0{,}045$. The coefficient for $female$ was estimated to -0,364 meaning that women have on average 30% lower wages than men. Marriage has on the other hand positive influence on wage, approximately 15%. This can be driven by the responsibility for family that an individual bears. What we did not expect is the estimated coefficient for the variable $disability$, which is positive and furthermore indicates, that having disability compounding work brings more than 12.7% higher wage. On the other hand, we expected that the magnitude of coefficient for the variable $metropolitan$, which was estimated to 5%, would be bigger. The coefficients for single industry dummy variables differ quite a lot, not only in magnitudes but also in signs. This is probably caused by the fact that back in the 1980s finance did not have such a big impact on the economy and society. People working in information and communications industry, utilities, surprisingly manufacturing and wholesale trade, construction, public administration, army and mining had higher wages. For exact numbers see Table C1.

5.3.1.1 Estimates and decomposition for Blacks

Concerning Blacks, all the estimates are significant at 1% level besides the coefficient for variable $married$, which is significant at 5% level. Comparing Blacks (for details see Table C3) and whites, the return to high school education is about one third of that of whites, i.e. 7%. Also the return to university degree is much lower, 53% compared to nearly 86% for whites. The estimates for $experience$ and $experience^2$ were 0.027 and -0.0004. This implies that the return to experience is increasing for the first 33.75 years of work with the growth rate of $-0.0008\ experience + 0.027$, which is lower than that for whites during the whole possible working life (0-90 years of experience). At this point we can conclude that there is large gap in return to experience and education between whites and Blacks. Concerning other variables, the gender wage gap is less than half of that of whites, only about 16%. The return to being disabled is

[13] The estimates in Appendix C are rounded up

huge compared to whites, about 35%. On the other hand, the advantage of working in metropolitan area is comparable with nearly 5%. Concerning the industry dummy variables, the signs are exactly the same compared to whites, only the magnitudes differ. The return to working as a professional is nearly 8.5 percentage points higher for Blacks (finance and real estate as benchmark), in case of retail trade and public administration is the difference (the same direction) about 5 percentage points, in information and communications even 9 percentage points. In mining industry and construction is the return about 5 percentage points lower for Blacks, in services even 10 percentage points lower, and agriculture 4 percentage points. The rest, i.e. utilities, manufacturing, wholesale trade, army and transportation, is comparable.

Proceeding to the Oaxaca decomposition (for details see Table D1), the wage differential between Blacks and whites is estimated to be 9.9%.

Table 3: Oaxaca decomposition of Black-white logarithm wage differential, 1980

Logarithm wage differential	Explained, White wage structure	Unexplained, White wage structure	Explained, Black wage structure	Unexplained, Black wage structure
0.099***	0.089***	0.01***	0.036***	0.062***
(0.0005)	(0.0002)	(0.0005)	(0.0002)	(0.0004)
As a percentage of total difference	89.9%	10.1%	36.4%	62.6%

Taking the wage structure of white workers as the competitive one that would be present in the absence of discrimination, 90% of the gap can be explained by differences in human capital endowment between the groups. This means if Blacks would on average have the same level of skills and distribution among industries, they would have 8.9% higher wage. The unexplained part accounts for the rest of the gap, i.e. 10% of the gap. This share is assigned to discrimination against Blacks. Assuming Blacks' wage structure as the competitive one would yield quite different results. The explained part of the gap accounts this time only for 36.4% of the gap while the unexplained for 63.6%. That is also the share assigned to discrimination. If Blacks were treated according to the same measures as whites they would have on average 6% higher wage than they do. Looking closely at the decomposition results, the explained part is in both cases driven mainly by the difference in university degree, yet much more under the white wage structure. On the other hand, the unexplained part is driven mainly by return to secondary school and college education, but moreover by the return to experience, therefore we can conclude that Blacks were in the 1980s discriminated against mainly in

terms of return to lower education and obtained experience, probably also on-the-job training. They significantly fell behind in terms of higher education.

5.3.1.2 Estimates and decomposition for Hispanics

Concerning Hispanics, all the estimates are significant at 1% level. The estimated coefficients for Hispanics (for details see Table C4) were compared to whites as follows; the return to high school education is with 15% by 6 percentage points lower than for whites. Similarly for college education, where the difference is about 17 percentage points (in the same direction). Concerning experience, with coefficients 0.032 and -0.00046 for $experience$ and $experience^2$, respectively, is the number of years for which is the return to education increasing equal to 34.8 years. The growth rate is $-0.00092\ experience + 0.032$, which is lower than for whites up to 46.4 years of experience. The gender wage gap is, as in case of Blacks, much smaller than for whites, only about 10% compared to more than 30% for whites. Interestingly, the return to working in a metropolitan area is negative for Hispanics, but the coefficient is with 0.2% not economically significant. There is smaller return to being married, only 9% compared to 15% for whites. This might be due to the fact that in order to maintain their large families poorer Hispanics take any jobs available. The return to being disabled is more than twice as high as for whites. Concerning the industry dummy variables, the signs are again the same as for whites but the estimates differ in magnitude. The return to being employed information and communications industry is about 5 percentage points higher than for whites, the same trend is in retail trade, where is the difference more than 8 percentage points, and also in agriculture, where is the difference more than 17 percentage points. Worth mentioning is also that whites have about 4,5 percentage points higher return to working in utilities and about 5 percentage points higher return to working in manufacturing industry. The other industries are comparable.

Turning to the Oaxaca decomposition (for details see Table D3), the estimated white-Hispanic logarithm wage gap is 9.8%, i.e. pretty much the same as the white-Black gap. Also the ratio of explained part and total gap is almost the same under the white wage structure; 86.7% can be explained by differences in characteristics. 13.3% of the gap is then assigned to discrimination. Under the Hispanic wage structure assumed as the competitive one the proportion is quite similar to the first case, compared to that of Blacks under the assumption of their wage structure as the competitive one. 75.5% can be explained using the differences in individual

Table 4: Oaxaca decomposition of Hispanic-white logarithm wage differential, 1980

Logarithm wage differential	Explained, White wage structure	Unexplained, White wage structure	Explained, Hispanic wage structure	Unexplained, Hispanic wage structure
0.098*** (0.0005)	0.085*** (0.0002)	0.012*** (0.0005)	0.074*** (0.0003)	0.024*** (0.0005)
As a percentage of total difference	86.7%	12.2%	75.5%	24.5%

characteristics, whereas 24.5% of the gap remains unexplained and as such is assigned to discrimination. Looking closely at the decomposition results, the situation is again similar to that of Blacks. The explained part is driven mainly by the differences in college education and partially also in high school education and experience. The unexplained, part, assigned to discrimination, is driven mainly by difference in return to experience, and partially high school education, marital status and also return to working in a metropolitan area. At this point we can conclude that in the 1980s, Hispanics to a great extent fell behind whites in terms of education but on the other hand were discriminated against mainly in terms of their experience. If Hispanics closed the characteristics gap they would have 7.4%-8.5% higher wage. What is more, if they were assessed according to the same criteria as whites, they would have on average 1.2%-2.4% higher wage.

5.3.1.3 Estimates and decomposition for Asians

Concerning Asians, all the estimates are significant at 1% level. Comparing the estimates for Asians (for details see Table C2) with those for whites, Asians had much smaller return to education; 57.4% for college and 11.6% for high school, compared to 85.5% for college and 21% for high school for whites. Concerning return to experience, with coefficients 0.039 for $experience$ and -0.00059 for $experience^2$ was the return to another year of experience increasing up to the 33rd year of experience and the growth rate was $0.039 - 0.00118\ experience$, which is again lower than that of whites. Yet, this time considering some measurement error, the growth rates could be the same as that of whites is $0.045 - 0.0012\ experience$. The gender gap is with 22% smaller than that of whites but the largest among minority groups. The return to being married is only about 3 percentage points lower than for whites, the return to working in a metropolitan area is less than 2% (it is expected that it sharply increases for all racial groups in 2007), on the other hand the return of being disabled is almost 3 percentage points higher than for whites. Concerning the industry dummy variables, the signs are

again corresponding to those of whites. Aside from wholesale trade, retail trade and army there are quite significant differences in the estimates. Asians had 3 percentage points higher return to working in public administration, nearly 5 percentage points higher return to working in transportation, and nearly 5 percentage points smaller return to working in mining. Their return to working in construction was more than twice that of whites, the returns to working in services and manufacturing were both about 5 percentage points higher, the same for utilities. Large difference was in the estimates of return to working in agriculture, about 15 percentage points. Asians had also about 9 percentage points higher return to working in information and communications industry and about 10.5 percentage points higher return to working as a professional.

Moving to the Oaxaca decomposition (for detailed results see Table D5), the estimated logarithm wage differential was -0.048, i.e. 4.8% in favor of Asians. Under

Table 5: Oaxaca decomposition of Asian-white logarithm wage differential, 1980

Logarithm wage differential	Explained, White wage structure	Unexplained, White wage structure	Explained, Asian wage structure	Unexplained, Asian wage structure
-0.048*** (0.0009)	-0.03*** (0.0004)	-0.018*** (0.0005)	-0.027*** (0.0004)	-0.021*** (0.001)
As a percentage of total difference	62.5%	37.5%	56.25%	43.75%

the assumption of white wage structure being the competitive one, the explained part of the differential accounts for 62,5% whereas the unexplained, assigned to discrimination (in this case nepotism), accounts for the remaining 37,5%. Under the assumption of Asian wage structure as the competitive one, the explained share of the differential is a bit smaller, 56,25%. Those remaining 43,75% of the gap is assigned to underpayment of whites. The explained part of the gap is driven mainly by differences in college education (thus productivity). On the other hand, the unexplained part is driven mainly by different treatment of female workers. To conclude the situation of Asians in the 1980s, whites seem to had higher return to their experience which was, however, in case of Asians outweighed by their return to higher education.

5.3.2 Estimates for year 2007

Concerning whites, all the estimates are significant at 1% level, for detailed results see Table C1. The return to education has according to our results significantly increased since the 1980s. The coefficients for secondary school and college are

estimated 0,266 and 0,857, respectively. The return to high school education has increased by more than 9 percentage points, while the return to university education has increased by 50 percentage points, i.e. by more than one third. The coefficient for $experience$ increased to 0.068, on the other hand, the coefficient for $experience^2$ slightly decreased from -0.0006 to -0.00107. This resulted in significant shortening of the time for which the return to experience is increasing from 37.5 years to 31.7 years. The growth rate is $-0.00214\ experience + 0.068$, which is smaller than the growth rate in the 1980s. The estimate for gender wage gap decreased by 5.5 percentage points, i.e. more than one fifth. The return to being married remained practically the same, it decreased from 15% to 14%. The return to being disabled has fallen significantly, from more than 12% to 4%. On the other hand, as we expected, the return to working in a metropolitan area has increased a lot since the 1980s, from 5% to almost 21%. Moving to the industry dummy variables, as the role of finance and real estate has grown during the years, it is not surprising, that return to working in any other industry than finance and real estate (apart from professionals) has significantly fallen. The only industries that are according to our results still attracting workers due to higher wages are utilities and mainly mining. The reasons for this trend were discussed above in section 5.2.2.

5.3.2.1 Estimates and decomposition for Blacks

Also for Blacks are all the estimates significant at 1% level; for detailed results see Table C3. The return to education has since the 1980s increased enormously, much more than for whites. The return to graduating from high school has increased from 7% to more than 32%. This resulted in closing the gap in return to high school education between Blacks and whites. Pretty much the same happened also in terms of return to university degree. It more than doubled (increased from 53% to 131%) and nearly closed the gap in return to college education between Blacks and whites (the gap remained smaller than 2 percentage points). Unlike for whites, both coefficients for experience changed in case of Blacks. Both coefficients more than doubled, the coefficient for $experience$ increased from 0.027 to 0.059, while the coefficient for $experience^2$ increased from -0.0004 to -0.0009. This led to the growth rate of $-0.0018\ experience + 0.059$ and the number of years for which is the return to experience increasing equal to 32.8. The number of years is slightly lower than in 1980 but on the other hand, the growth rate is for up to 32 years of experience higher than in 1980. The gender wage gap was estimated more than 4 percentage points smaller than

in 1980 and still is much narrower than for whites. The return to being married also slightly decreased, copying the trend for whites. The return to being disabled also copied the white trend, but went much further. It decreased from 35.5% to -7% and being disabled changed its substance, so to speak, from being an "advantage" to being a disadvantage. The return to working in a metropolitan also increased and keeps the pace with the return for whites at the level of nearly 22%. Concerning the industry dummy variables, the situation is very similar to that of whites. All the returns has decreased relatively to the return to working in finance in 1980, but there are much more industries in which Blacks on average get higher wage than in finance, such as public administration by nearly 3%, mining by 27%, army by 13.5%, utilities by more than 12%, and information and communications by 1.5%.

Moving on to the Oaxaca decomposition (for details see Table D2), the logarithm wage differential was estimated to 29.6%. Assuming the white wage structure

Table 6: Oaxaca decomposition of Black-white logarithm wage differential, 2007

Logarithm wage differential	Explained, White wage structure	Unexplained, White wage structure	Explained, Black wage structure	Unexplained, Black wage structure
0.296***	0.147***	0.149***	0.118***	0.177***
(0.0005)	(0.0002)	(0.0005)	(0.0002)	(0.0005)
As a percentage of total difference	49.7%	50.3%	39.8%	59.8%

as the competitive one yields about one half of the gap being explained by differences in characteristics. It follows that if Blacks had on average the same characteristics as whites, they would have nearly 15% higher wages. Those remaining 15%, the other half of the gap, is assigned to discrimination as we cannot explain it using the average characteristics. If Blacks were treated according to the same criteria as whites they would have about 15% higher wage. Let us now assume that Black wage structure is the competitive one, which would prevail under no discrimination. Explained part of the logarithm wage gap decreased by 10% percentage points increasing thus the share of the gap assigned to discrimination. That is now nearly 60%. Under this assumed wage structure are therefore whites overpaid by nearly 18%. The explained part of the gap is driven mainly by differences in college education and partly by differences in experience. On the other hand, the unexplained part is driven mainly by differences in return to education. So at this point we can conclude that Blacks in 2007 significantly

fell behind whites in terms of higher education but at the same time were enormously discriminated against in terms of their obtained experience, on the job training etc.

5.3.2.2 Estimates and decomposition for Hispanics

Concerning Hispanics, all the estimates are significant at 1% level besides the coefficient for $army$, which is significant at 5% level (for details see Table C4). The return to secondary school education increased since the 1980s from 15% to more than 30% and closed the existing gap in return to high school education between whites and Hispanics. However, this was not the case concerning the return to college education. Unlike Blacks, after the return has increased to 117.5% since 1980 (the return was about 68% by that time) there is still a gap of 14 percentage points in the return to college education between Hispanics and whites. Concerning the return to experience, the coefficients were estimated as follows: for $experience$ 0.046 and for $experience^2$ -0.0007. This results in the growth rate $-0.0014\ experience\ + 0.046$, which is higher than that of whites after 29 years, and up to 29 years of experience also higher than that of Hispanics back in 1980. The return to another year of experience is increasing up to 32.8 years of experience. In case of gender wage differences was the trend opposite to that of whites and Blacks. The return to being a woman decreased by nearly 7 percentage points to -18.4%. On the other hand, the return to being married increased by 2 percentage points to almost 11.5%. The returns to working in a metropolitan area and being disabled followed the changes that occurred in case of whites as well as in case of Blacks. The return to being disabled decreased from more than 30% to 10%. The change in case of return to working in a metropolitan area had the other direction. It increased from more or less zero to 12%. Concerning the industry dummy variables, apart from public administration all the coefficients has decreased compared to 1980. This means that differences between wages in finance and real estate, and in other industries widened.

Proceeding to the Oaxaca decomposition (detailed results can be found in Table

Table 7: Oaxaca decomposition of Hispanic-white logarithm wage differential, 2007

Logarithm wage differential	**Explained, White wage structure**	**Unexplained, White wage structure**	**Explained, Hispanic wage structure**	**Unexplained, Hispanic wage structure**
0.304*** (0.0006)	0.165*** (0.0002)	0.138*** (0.0006)	0.165*** (0.0002)	0.14*** (0.0006)
As a percentage of total difference	54.3%	45.4%	54.3%	46%

D4), the estimated logarithm wage white-Hispanic differential is 30.4%. Let us first assume that white wage structure is the competitive one that would prevail in the absence of discrimination. Then the part of the wage differential that can be explained by differences in average characteristics accounts for more than 54% of the gap. Nearly one half of the gap remains unexplained and is as such assigned to discrimination. Taking the Hispanic wage structure as the competitive one yields in this particular case practically the same result. It follows that Hispanics would get on average 14% higher wage if they were assessed according to the same criteria as whites, and about 16.5% higher wage if they had on average the same characteristics as whites. The explained part of the gap is in case of Hispanics driven in particular by differences in college education and their experience. The unexplained part is driven mainly by differences in return to experience. We can conclude that Hispanics are discriminated against mainly in terms of their experience.

5.3.2.3 Estimates and decomposition for Asians

The estimated coefficients for Asians were all but the one for variable $army$ significant at 1% level. The coefficient for $army$ was insignificant at all common levels. For detailed results see Table C2. The return to secondary school education increased only little, by 4 percentage points to level of 15.5% compared to 1980. It remained far behind the 1980 return to high school education for whites. Similar story holds for the college education, the return has increased since 1980 from 57.5% to 108% but there is a gap of about 27 percentage points between Asians and whites. Concerning the return to another year of experience, the coefficients for $experience$ and $experience^2$ were estimated to 0.059 and -0.00104, respectively. This implies the growth rate of $0.059 - 0.00208\ experience$, which is at any point higher than that of whites in 2007 and up to 22 years of experience also higher than that of Asians in 1980. The number of years for which is the return to another year of experience increasing is equal to 28.4 years. The return to being a woman has increased by one third since 1980, from -18.3% to -13%, thus the gender wage gap has partially narrowed. Furthermore, it is half of that for whites. The return to being married remained more or less the same, about 13%. Nor the return to being disabled has changed much, it decreased from 16% to 14% which is only little decrease compared to all other races. On the other hand, the return to working in a metropolitan area has followed the same trend and it increased by nearly 14 percentage points to more than 15.8%. Concerning the industry dummy

variables, all the coefficients decreased compared to 1980, confirming the increasing role of finance. The only industry with significantly higher wages than in finance and real estate was mining with 28% higher wages. Utilities industry fell behind this time, the wages are on average only 3% higher for Asians in the sector.

Turning to the Oaxaca decomposition (for detailed results see Table D6), the

Table 8: Oaxaca decomposition of Asian-white logarithm wage differential, 2007

Logarithm wage differential	Explained, White wage structure	Unexplained, White wage structure	Explained, Asian wage structure	Unexplained, Asian wage structure
-0.148*** (0.001)	-0.116*** (0.0002)	-0.032*** (0.001)	-0.125*** (0.0003)	-0.022*** (0.001)
As a percentage of total difference	78.4%	21.6%	84.5%	14.9%

estimated logarithm wage differential was 14.8% in favor of Asians. Assuming that the white wage structure is the competitive one, the differential can be decomposed so that more than 78% of the differential is explained by differences in average individual characteristics, whereas about 21.5% remains unexplained. This part is assigned to nepotism in favor of Asians. Assuming that the Asian wage structure is the competitive one that would prevail in absence of nepotism, 84.5% of the gap can be explained and 15% of the gap is assigned to underpayment of whites. The explained part of the gap is driven by differences mainly in university education; on the other hand, the unexplained part has several dominant factors: disadvantageous treatment of women, return to secondary school education and mainly return to experience.

Conclusion

In this work we introduce two types of theories explaining wage differentials among groups (in our case races) as a result of discrimination. First, we deal with the taste-based theory of discrimination based on Becker (1971). He presented the concept of a discrimination coefficient, which measures the non-pecuniary aspects of economic transactions, the disutility an individual perceives when employing, working with or buying from an individual from some other group. The main message of the theory is that any expression of discrimination decreases the volume of trade among groups and thus harms all of the groups. The minority groups are, however, harmed much more as they gain relatively more from the trade with the majority group. That is also the reason

why these should not retaliate the discriminatory behavior. It would further decrease the trade and might lead up to economic segregation.

The second type of theories we deal with is the so called statistical theory of discrimination. We considered the basic version of this theory by Phelps (1972). As employer never has perfect information about workers and it would be very costly to obtain them, he often uses some statistics on average performance of different groups in order to help him predict workers' productivity. In our case, the differentiating element is race. As a result, workers belonging to a group with higher productivity often benefit from this approach, no matter what their real productivity is, the opposite happens to workers from a low productivity group.

We also revealed the history of U.S. immigrants. Afro-Americans were brought to the U.S. as slaves. After the Civil War, when slavery was abolished they were, however, not treated as *sui juris* citizens, mainly in the South. It was difficult to perceive your own slave as equal over night. Unfortunately, this attitude has not fully diminished yet.

Many Hispanics originally lived in the southwest of the U.S. as this area was colonized by Spanish. The Southwest was later connected to the U.S. The number of Mexicans in U.S. fluctuated according to economic conditions in both countries. During the World War 2, Mexicans were hired to work on farms due to labor scarcity. There are also many illegal immigrants, who perceive migrating to U.S. as their economic opportunity.

Asians, starting with Chinese, came to U.S. to become rich during the Gold Rush. They later worked on farms, as laundry workers, gardeners and also on the Transcontinental Railroad construction. When this construction was finished they were no longer needed and started to be perceived as a threat. They had no rights and therefore had no other choice than living in a closed community in so called Chinatowns. Japanese came later as a cheaper replacement of Chinese. They later faced to the same treatment as Chinese. Their position further worsened after the Japanese attacked Pearl Harbor. After the World War 2, they were given the opportunity to become U.S. citizens.

Investigating the labor market data based on the Bureau of Labor Statistics (2008) we find there are quite large differences in income among the racial groups, and not only the income but furthermore in educational attainment, unemployment rates etc.

This is together with the nature of the minorities' residence history in the U.S. the main motivation for our research. The research questions then arise naturally. We are interested whether the minorities are discriminated against or whether the observed differences in income can be sufficiently explained by differences in race characteristics. Moreover, is the position of all the minorities the same or are there any fundamental differences? We are also interested how the situation has changed since the 1980s, during the era of "information society". Our hypotheses are basically these:

- We believe that Asians have higher earnings for the reason of higher education. There might also be higher returns to their experience and on-the-job training, probably because of the stereotype prevailing about their attitude. We do not think that there will be much of the gap remaining unexplained. We also believe that their position has significantly improved over the years.
- Hispanics have the lowest educational attainment of all the racial groups. We are convinced that this harms them in the form of reduced income. We also expect improvement in terms of explained/unexplained ratio compared to 1980.
- We expect that the largest portion of the wage gap would remain unexplained for Blacks, i.e. Blacks experience discrimination more than any other racial group. We think that large part of the explained portion of the gap would be accounted for by relatively low education and also working outside of the metropolitan areas. Despite that, we believe that their position in the market has improved since the 1980s.

We estimated separately the same model for each race using the maximum likelihood estimation method and adjusted it for possible selection bias using the Heckman method. Then we proceeded to the Oaxaca decomposition. Reduced forms of results of the Oaxaca decompositions can be found in Table 9 below. For detailed results see Appendix D.

Table 9: Comparison of the Oaxaca decomposition by race and year

	Blacks	**Hispanics**	**Asians**
1980			
Logarithm wage gap	0.099	0,098	-0,048
Unexplained share	10.1% - 62.6%	12.2% - 24.5%	37.5% - 43.75%
2007			
Logarithm wage gap	0,296	0,304	-0,148
Unexplained share	50.3% - 59.8%	45.4% - 46%	14.9% - 21.6%

We begin with the situation of Asians. In 1980 their earnings exceeded those of whites by nearly 5% on average. In this case we cannot reject our hypotheses that it is given mainly by their high educational attainment. The Oaxaca decomposition of the logarithm wage differential shows that Asians would have on average 7% - 9.5% lower wages having the same level of college education as whites. On the other hand, their return to education is lower than that of whites. In this field they are not treated equally. The same holds for their return to experience, thus we can reject the hypotheses of higher return to experience. Concerning the year 2007, Asians had on average 15% higher wages than whites. The average Asian level of education further increased compared to that of whites, the same trend, however, occurred in the difference in return to education in favor of whites and much more in the difference in return to experience. It follows that Asians have higher wages than whites for the reason of higher level of schooling while their human capital is discriminated against and thus Asians have lower return to it than whites. Despite the fact that the wage gap widened during the years, the unexplained share of it rapidly decreased from 37.5% - 43.7% to 15% - 21.5% depending on the assumed competitive wage structure. The Asian position has certainly improved throughout the years.

Moving to Hispanics, in 1980 was the estimated wage gap on average 10% in favor of whites. We were not able to reject the hypotheses that this was partially due to their low level of educational attainment. On the other hand the unexplained part of the gap accounted for 12% - 24.5%. Hispanics were discriminated against in terms of return to their experience. The estimated wage gap has dramatically widened by 2007 (up to 30.5%) and so has the unexplained amount of the wage gap. It increased to nearly 46%. According to the Oaxaca decomposition this was mainly due to radical increase in return to experience gap between Hispanics and Whites. Also the differences in level of schooling further increased. Hispanics seem to be worse off than in the 1980s.

Turning to Blacks, the wage gap in 1980 was pretty much the same as for Hispanics, about 10%; however, the distribution depending on assumed competitive wage structure was quite uncertain. The difference in working in metropolitan area was, though statistically significant, economically insignificant. Concerning the explained part of the gap, the main drawback of Blacks was again the college education. On the other hand they were discriminated against their return to high school education and mainly experience. Following the trend for Hispanics, the wage gap has enormously

widened since 1980, up to 30%. On the other hand the share of the unexplained part of the gap converged to 50% - 60%. The situation remained more or less similar to that in the 1980s; Blacks have relatively lower far partly for the reason of low level of college education. The unexplained part of the gap is still driven by difference in return to experience but compared to 1980 the difference is half. The position of Blacks in terms of discrimination against their human capital has certainly improved (despite that the wage gap widened), yet it is still not perfect. Large portion of the wage differential is accounted for by low level of university education.

It would be interesting to investigate the impact of the global financial crisis on minorities in terms of both, wage differentials and unemployment rates.

Reference List

Aigner, Dennis J., Cain, Glen G. 1977, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review* 30, pp. 175-187.

American Association of University Women (AAUW) n.d., *Title VII of the Civil Rights Act of 1964*. Available from: <http://www.aauw.org/act/laf/library/workplaceharassment_TitleVII.cfm>. [24 April 2012]

Arrow, Kenneth J. 1972a, "Models of Job Discrimination," *Racial Discrimination in Economic Life*. Lexington Books, D. C. Heath and Co., Lexington, pp. 83-102.

Arrow, Kenneth J. 1972b, "Some Mathematical Models of Race in the Labor Market," *Racial Discrimination in Economic Life*. Lexington Books, D. C. Heath and Co., Lexington, pp. 187-204.

Arrow, Kenneth J. 1973, "The Theory of Discrimination," *Discrimination in Labor Markets*. Princeton University Press, Princeton, pp. 3-33.

Arrow, Kenneth J. 1998, "What Has Economics to Say About Racial Discrimination?" *Journal of Economic Perspectives*, Vol. 12, No. 2, pp. 91-100.

Banerjee, Abhijit, Bertrand, Marianne, Datta, Saugato, Mullainathan, Sendhil 2009, "Labor Market Discrimination in Delhi: Evidence from a field experiment," *Journal of Comparative Economics*, Vol. 37, No. 1, pp.14-27.

Becker, Gary S. 1971, *The Economics of Discrimination,* 2nd edition. The University of Chicago Press, Chicago, IL. ISBN 0226041166

Bertrand, Marianne, Mullainathan, Sendhil 2004, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *The American Economic Review*, Vol. 94, No. 4, pp. 991-1013.

Borjas, George J. 2005, *Labor Economics*, 3rd ed. McGraw-Hill/Irwin, New York, NY. ISBN 0072871776

Card, David, DiNardo, John E. 2002, "Skill-biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, Vol. 20, No. 4, pp. 733-783.

Card, David, Lemieux, Thomas 1993, "Wage Dispersion, Returns to Skill, and Black-white Wage Differentials," *NBER Working Paper Series*, No. 4365. Cambridge, Ma.

Carlsson, Magnus, Rooth, Dan-Olof 2007, "Evidence of ethnic discrimination in the Swedish labor market using experimental data," *Labour Economics*, Vol. 14, No. 4, pp.716-729.

Central Intelligence Agency (CIA) n.d., *The World Factbook – United States*. Available from: <https://www.cia.gov/library/publications/the-world-factbook/geos/us.html>. [20 April 2012]

Cotton, Jeremiah 1988, "Discrimination and Favoritism in the U.S. Labor Market: The Cost to a Wage Earner of Being Female and Black and the Benefit of Being Male and White," *American Journal of Economics and Sociology*, Vol. 47, No. 1, pp. 15-28.

Donohue, John J., III, Heckman, James 1991, "Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks," *Journal of Economic Literature*, Vol. 29, No. 4, pp. 1603-1643.

Encyclopedia of Alabama 2007, *Birmingham Campaign of 1963*. Available from: <http://www.encyclopediaofalabama.org/face/Article.jsp?id=h-1358>. [27 April 2012]

Flek, V. 2012a, lecture notes from JEB053 – Labour Economics at the Charles University, Czech Republic on 23 April 2012.

Flek, V. 2012b, lecture notes from JEB053 – Labour Economics at the Charles University, Czech Republic on 19 March 2012.

Heckman, James J. 1976, "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," *Annals of Economic and Social Measurement*, Vol. 5, No. 4, pp. 475–492.

Jann, Ben 2008, "The Blinder-Oaxaca decomposition for linear regression models," *The Stata Journal*, Vol. 8, No. 4, pp. 453-479.

Kaas, Leo, Manger, Christian 2011, "Ethnic Discrimination in Germany's Labour Market: A Field Experiment," *German Economic Review*, Vol. 13, No. 1, pp. 1-20.

Landale, Nancy S., Oropesa, R. Salvador, Bradatan, Christina 2006, "Hispanic Families in the United States: Family Structure and Process in an Era of Family Change" in Marta Tienda and Faith Mitchell, editors, *Hispanics and the Future of America*. National Academies Press, Washington. ISBN 0309100445 Available from: <http://www.ncbi.nlm.nih.gov/books/NBK19902/>. [24 April 2012]

Lee, R. n.d., "The History Guy: The Mexican-American War." Available from: <http://www.historyguy.com/Mexican-American_War.html>. [22 April 2012]

Le, Cuong N. 2012a. "The First Asian Americans," *Asian-Nation: The Landscape of Asian America*. Available from: <http://www.asian-nation.org/first.shtml>. [21 April 2012]

Le, Cuong N. 2012b. "Construction and Destruction: Japanese American Internment," *Asian-Nation: The Landscape of Asian America*. Available from: <http://www.asian-nation.org/internment.shtml>. [21 April 2012]

Le, Cuong N. 2012c. "The 1965 Immigration Act," *Asian-Nation: The Landscape of Asian America*. Availabe from: <http://www.asian-nation.org/1965-immigration-act.shtml>. [21 April 2012]

McCall, John J. 1972a, *Income Mobility, Racial Discrimination, and Economic Growth.* Lexington Books, D. C. Heath and Co., Lexington. ISBN 9780669908527

McCall, John J. 1972b, "The Simple Mathematics of Information, Job Search, and Prejudices," *Racial Discrimination in Economic Life*. Lexington Books, D. C. Heath and Co., Lexington, pp. 205-44.

Minnesota Population Center, University of Minnesota, "Note on Adjusting Dollar Amount Variables for Inflation (CPI-U)," *IPUMS CPS*. Available from: <http://cps.ipums.org/cps/cpi99.shtml>. [24 April 2012]

Oaxaca, Ronald L. 1973, "Male-Female wage Differentials in Urban Labor Markets," *International Economic Review*, Vol. 14, No. 3, pp. 693-709.

Oaxaca, Ronald L., Ransom, Michael R.1994, "On Discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics*, Vol. 61, No. 1, pp. 5-21.

Phelps, Edmund S. 1972, "The Statistical Theory of Racism and Sexism," *American Economic Review*, Vol. 62, No. 4, pp. 659-61.

Reder, Melvin W. 1972, "Human Capital and Economic Discrimination," *Human Resources and Economic Welfare; Essays in Honor of Eli Ginzberg*. Columbia University Press, New York, pp.71-88.

Reimers, Cordelia W. 1983, "Labor Market Discrimination Against Hispanic and Black Men," *Review of Economics and Statistics*, Vol. 65, No. 4, pp. 570-579.

Ruggles, Steven, Alexander, J. Trent, Genadek, Katie, Goeken, Ronald, Schroeder, Matthew B., Sobek, Matthew. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010. Available from: <http://usa.ipums.org>.

Schaffer Library of Drug Policy n.d., *A Short History of the Opium Wars*. Available from: <http://www.druglibrary.org/schaffer/heroin/opiwar1.htm>. [21 April 2012]

Spence, A. Michael 1973, "Job Market Signaling," *Quarterly Journal of Economics*, Vol. 87, No. 3, pp. 355-74.

Spence, A. Michael 1974, *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*. Harvard University Press, Cambridge. ISBN 9780674549906

United States. Bureau of Labor Statistics. 2008, *Labor Force Characteristics by Race and Ethnicity, 2007*. Available from: <www.bls.gov/cps/cpsrace2007.pdf>. [20 April 2012]

United States Census Bureau. *2010 Census*. Available from: <http://2010.census.gov/2010census>. [20 April 2012]

United States Census Bureau. Humes, Karen R., Jones, Nicholas A., Ramirez, Roberto R. 2011, *The 2010 Census Brief: Overview of Race and Hispanic Origin*. Available from: <http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf>. [20 April 2012]

United States Census Bureau 2010, *Table 10. Resident Population by Race, Hispanic origin and Age: 2000 and 2009*. Available from: <http://www.census.gov/compendia/statab/2012/tables/12s0010.pdf>. [24 April 2012]

U. S. Diplomatic Mission to Germany n.d., *History of the United States*. Available from: <http://usa.usembassy.de/>. [21 April 2012]

U.S. Equal Employment Opportunity Commission n.d., *Title VII of the Civil Rights Act of 1964*. Available from: <http://www.eeoc.gov/laws/statutes/titlevii.cfm>. [24 April 2012]

United States History n.d., *Civil Rights Act of 1964*. Available from: <http://www.u-s-history.com/pages/h3947.html>. [24 April 2012]

Guisepi, R. A., editor, n.d., "Hispanic Americans," *International World History Project*. Available from: <http://history-world.org/hispanics.htm>. [22 April 2012]

Welch, Finis 1967, "Labor-market discrimination: An Interpretation of Income Differences in the Rural South," *The Journal of Political Economy*, Vol. 75, No. 3, pp. 225-240.

Wooldridge, Jeffrey M. 2002, *Introductory Econometrics: A Modern Approach*, 2nd ed. South-western Cengage Learning, Mason. ISBN 0324113641

Wright, Donald R. 2000, *Slavery in Africa*. Available from: <http://autocww.colorado.edu/~blackmon/E64ContentFiles/AfricanHistory/SlaveryInAfrica.html>. [21 April 2012]

List of Appendix Tables

Appendix A

A1. Table of Summary Statistics, 2007, weighted

Observations: 1905194

	Mean	Standard deviation	Description
Wage	935.074	0.155	Wage in 2007 USD
log(wage)	6.389	$793 \cdot 10^{-7}$	log(wage)
White	0.691	$382 \cdot 10^{-7}$	Dummy for white race
Black	0.121	$27 \cdot 10^{-6}$	Dummy for Blacks
Asian	0.049	$18 \cdot 10^{-6}$	Dummy for Asians
Hispanic	0.137	$285 \cdot 10^{-7}$	Dummy for Hispanics
Female	0.474	$413 \cdot 10^{-7}$	Dummy for gender
Married	0.494	$413 \cdot 10^{-7}$	Dummy for marital status
Disability	0.022	$123 \cdot 10^{-7}$	Dummy for disbility affcteng work
Experience	20.433	0.0011	Potential years of experience
Experience^2	580.84	0.0468	Potential years of experience squared
Primary School	0.106	$255 \cdot 10^{-7}$	Dummy for the highest level of education acquired
Secondary School	0.609	$403 \cdot 10^{-7}$	Dummy for the highest level of education acquired
College	0.283	$373 \cdot 10^{-7}$	Dummy for the highest level of education acquired
Metropolitan area	0.674	$388 \cdot 10^{-7}$	Work in metropolitan area
Professional	0.097	$245 \cdot 10^{-7}$	Industry dummy
Finance, Real Estate	0.067	$208 \cdot 10^{-7}$	Industry dummy
Information and Communications	0.024	$129 \cdot 10^{-7}$	Industry dummy
Retail Trade	0.119	$268 \cdot 10^{-7}$	Industry dummy
Agriculture	0.011	$8.61 \cdot e^{-6}$	Industry dummy
Utilities	0.008	$7.48 \cdot e^{-6}$	Industry dummy
Wholesale	0.031	$143 \cdot 10^{-7}$	Industry dummy
Services	0.351	$395 \cdot 10^{-7}$	Industry dummy
Army	0.007	$6.99 \cdot e^{-6}$	Industry dummy
Construction	0.072	$214 \cdot 10^{-7}$	Industry dummy
Mining	0.004	$5.69 \cdot e^{-6}$	Industry dummy
Transportation	0.042	$166 \cdot 10^{-7}$	Industry dummy
Public Administration	0.046	$174 \cdot 10^{-7}$	Industry dummy
Head of family	0.484	$413 \cdot 10^{-7}$	Dummy for being head of family
Number of children	0.767	$909 \cdot 10^{-7}$	Number of own children in household
Work	0.958	$164 \cdot 10^{-7}$	Dummy for people having wage, from wage and salary sector and working at least one week a year

A2. Table of Coefficients of Correlation among variables for year 2007, weighted

	Black	Asian	Hispanic	Female	Disability	Experience	Experience^2	Secondary school	College	Married	Metropolitan area	Professional	Information and Communications	Retail trade
Black	1.0000													
Asian	-0.0897	1.0000												
Hispanic	-0.1577	-0.0949	1.0000											
Female	0.0179	0.0085	-0.0227	1.0000										
Disability	0.0404	-0.0335	-0.0350	0.0083	1.0000									
Experience	-0.0379	-0.0279	-0.0613	0.0163	0.1928	1.0000								
Experience^2	-0.0356	-0.0288	-0.0634	0.0202	0.2079	0.9614	1.0000							
Secondary school	0.0533	-0.0745	-0.0524	0.0100	0.0262	0.0474	0.0427	1.0000						
College	-0.0906	0.1038	-0.1368	0.0196	-0.0996	0.0031	-0.0442	-0.7059	1.0000					
Married	-0.1609	0.0302	-0.0429	0.0132	-0.0650	0.3489	0.2737	-0.0337	0.1713	1.0000				
Metropolitan area	-0.0181	0.0380	0.0214	-0.0825	-0.2396	0.0066	-0.0464	-0.0262	0.1687	0.0696	1.0000			
Professional	-0.0164	0.0108	-0.0062	-0.0340	-0.0417	-0.0044	-0.0205	-0.0604	0.1083	0.0189	0.0944	1.0000		
Information and Communications	-0.0040	0.0050	-0.0217	-0.0149	-0.0216	-0.0087	-0.0151	-0.0150	0.0523	0.0052	0.0497	-0.0449	1.0000	
Retail trade	-0.0113	-0.0043	-0.0142	0.0153	-0.0345	-0.0732	-0.0593	0.0868	-0.0800	-0.0514	0.0341	-0.1042	-0.0489	1.0000

A2. Table of Coefficients of Correlation among variables for year 2007, weighted, continue

	Black	Asian	Hispanic	Female	Disability	Experience	Experience^2	Secondary school	College	Married	Metropolitan area	Professional	Information and Communications	Retail trade
Agriculture	-0.0175	0.0463	-0.0543	-0.0115	0.0151	0.0156	-0.0107	-0.0333	0.0110	-0.0470	-0.0333	-0.0156	-0.0362	-0.0175
Utilities	-0.0084	-0.0145	-0.0444	-0.0136	0.0297	0.0235	0.0198	-0.0011	0.0309	0.0147	-0.0251	-0.0118	-0.0273	-0.0084
Manufacturing	0.0079	0.0038	-0.1206	-0.0355	0.0731	0.0532	0.0413	-0.0313	0.0585	0.0463	-0.1005	-0.0471	-0.1093	0.0079
Wholesale trade	0.0014	0.0024	-0.0584	-0.0252	0.0222	0.0121	0.0170	0.0018	0.0299	0.0520	-0.0499	-0.0234	-0.0543	0.0014
Services	0.0086	-0.0321	0.2183	-0.0795	-0.0666	-0.0654	-0.0459	0.1057	-0.0216	0.0863	-0.2033	-0.0953	-0.2211	0.0086
Army	-0.0037	-0.0088	-0.0546	-0.0163	-0.0563	-0.0535	0.0261	-0.0048	-0.0056	0.0210	-0.0231	-0.0108	-0.0251	-0.0037
Construction	-0.0418	0.0674	-0.2168	-0.0318	0.0124	-0.0051	0.0530	-0.0958	0.0116	0.0374	-0.0833	-0.0391	-0.0906	-0.0418
Mining	-0.0107	-0.0048	-0.0458	-0.0073	0.0088	0.0054	0.0129	-0.0114	0.0189	-0.0184	-0.0192	-0.0090	-0.0209	-0.0107
Transportation	-0.0052	-0.0049	-0.0931	-0.0168	0.0528	0.0409	0.0626	-0.0470	0.0223	0.0476	-0.0598	-0.0281	-0.0651	-0.0052
Public administration	-0.0122	-0.0311	-0.0182	-0.0285	0.0430	0.0284	-0.0007	0.0596	0.0417	0.0486	-0.0609	-0.0286	-0.0663	-0.0122
Number of children	0.0198	0.0804	0.0756	-0.0664	0.0656	-0.0421	-0.0177	0.0507	0.4172	0.0617	0.0060	-0.0059	-0.0299	0.0198
Head of family	-0.0288	-0.0587	-0.1009	0.0126	0.2557	0.2006	-0.0143	0.1342	0.0591	0.1295	0.0404	0.0252	-0.0409	-0.0288

A2. Table of Coefficients of Correlation among variables for year 2007, weighted, continue

	Agriculture	Utilities	Manufacturing	Wholesale trade	Services	Army	Construction	Mining	Transportation	Public administration	Number of children	Head of family
Agriculture	1.0000											
Utilities	-0.0087	1.0000										
Manufacturing	-0.0349	-0.0264	1.0000									
Wholesale trade	-0.0173	-0.0131	-0.0523	1.0000								
Services	-0.0706	-0.0533	-0.2131	-0.1059	1.0000							
Army	-0.0080	-0.0061	-0.0242	-0.0120	-0.0490	1.0000						
Construction	-0.0289	-0.0218	-0.0873	-0.0434	-0.1767	-0.0201	1.0000					
Mining	-0.0067	-0.0050	-0.0201	-0.0100	-0.0407	-0.0046	-0.0167	1.0000				
Transportation	-0.0208	-0.0157	-0.0627	-0.0312	-0.1269	-0.0144	-0.0520	-0.0120	1.0000			
Public administration	-0.0212	-0.0160	-0.0639	-0.0317	-0.1292	-0.0147	-0.0529	-0.0122	-0.0380	1.0000		
Number of children	0.0123	0.0079	0.0309	0.0147	0.0099	-0.0063	0.0142	0.0033	0.0107	0.0136	1.0000	
Head of family	0.0048	0.0303	0.0684	0.0292	-0.0274	-0.0011	0.0221	0.0160	0.0378	0.0595	0.1361	1.0000

A3. Table of Summary Statistics, 1980, weighted

Observations: 1442342

	Mean	Standard deviation	Description
Wage	754.677	0.088	Wage in 2007 USD
log(wage)	6.288	$881 \cdot 10^{-7}$	log(wage)
White	0.825	$374 \cdot 10^{-7}$	Dummy for white race
Black	0.102	$298 \cdot 10^{-7}$	Dummy for Blacks
Asian	0.017	$128 \cdot 10^{-7}$	Dummy for Asians
Hispanic	0.055	$225 \cdot 10^{-7}$	Dummy for Hispanics
Female	0.444	$49 \cdot 10^{-6}$	Dummy for gender
Married	0.591	$485 \cdot 10^{-7}$	Dummy for marital status
Disability	0.047	$21 \cdot 10^{-6}$	Dummy for disbility affcteng work
Experience	17.755	0.0013	Potential years of experience
Experience^2	499.612	0.0604	Potential years of experience squared
Primary School	0.244	$423 \cdot 10^{-7}$	Dummy for the highest level of education acquired
Secondary School	0.584	$486 \cdot 10^{-7}$	Dummy for the highest level of education acquired
College	0.171	$372 \cdot 10^{-7}$	Dummy for the highest level of education acquired
Metropolitan area	0.3	$452 \cdot 10^{-7}$	Work in metropolitan area
Professional	0.161	$306 \cdot 10^{-7}$	Industry dummy
Finance, Real Estate	0.046	$175 \cdot 10^{-7}$	Industry dummy
Information and Communications	0.045	$173 \cdot 10^{-7}$	Industry dummy
Retail Trade	0.145	$293 \cdot 10^{-7}$	Industry dummy
Agriculture	0.025	$13 \cdot 10^{-6}$	Industry dummy
Utilities	0.01	$8.54e^{-6}$	Industry dummy
Wholesale	0.033	$149 \cdot 10^{-7}$	Industry dummy
Services	0.072	$215 \cdot 10^{-7}$	Industry dummy
Army	0.005	$6.16e^{-6}$	Industry dummy
Construction	0.05	$183 \cdot 10^{-7}$	Industry dummy
Mining	0.008	$7.43e^{-6}$	Industry dummy
Transportation	0.034	$151 \cdot 10^{-7}$	Industry dummy
Public Administration	0.036	$156 \cdot 10^{-7}$	Industry dummy
Head of family	0.509	$493 \cdot 10^{-7}$	Dummy for being head of family
Number of children	0.922	$1222 \cdot 10^{-7}$	Number of own children in household
Work	0.97	$165 \cdot 10^{-7}$	Dummy for people having wage, from wage and salary sector and working at least one week a year

A4. Table of Coefficients of Correlation among variables for year 1980, weighted

	Black	Asian	Hispanic	Female	Disability	Experience	Experience^2	Secondary school	College	Married	Metropolitan area	Professional	Information and Communications	Retail trade
Black	1.0000													
Asian	-0.0472	1.0000												
Hispanic	-0.0887	-0.0332	1.0000											
Female	0.0189	0.0055	-0.0033	1.0000										
Disability	0.0379	-0.0232	-0.0139	-0.0199	1.0000									
Experience	-0.0293	-0.0199	-0.0225	0.0286	0.2424	1.0000								
Experience^2	-0.0237	-0.0223	-0.0244	0.0283	0.2589	0.9634	1.0000							
Secondary school	-0.0454	-0.0233	-0.0795	0.0649	-0.0761	-0.0882	-0.1247	1.0000						
College	-0.0760	0.0554	-0.0621	-0.0762	-0.0689	-0.0872	-0.1178	-0.4585	1.0000					
Married	-0.1519	0.0003	-0.0238	0.0033	-0.0205	0.3593	0.2615	0.0588	0.0758	1.0000				
Metropolitan area	-0.0242	0.0131	0.0023	-0.0916	-0.0979	-0.0216	-0.0451	0.0393	0.0924	0.0213	1.0000			
Professional	0.0125	0.0097	-0.0334	0.1599	-0.0533	-0.0408	-0.0556	-0.0609	0.2914	0.0137	0.0637	1.0000		
Information and Communications	0.0050	-0.0054	-0.0091	-0.1006	-0.0165	0.0285	0.0142	0.0585	-0.0293	0.0335	0.0438	-0.0955	1.0000	
Retail trade	-0.0475	0.0074	-0.0139	0.0458	-0.0416	-0.1178	-0.0917	0.0368	-0.0866	-0.0872	0.0206	-0.1808	-0.0896	1.0000

A4. Table of Coefficients of Correlation among variables for year 1980, weighted, continue

	Black	Asian	Hispanic	Female	Disability	Experience	Experience^2	Secondary school	College	Married	Metropolitan area	Professional	Information and Communications	Retail trade
Agriculture	-0.0220	-0.0057	0.0297	-0.0880	-0.0016	0.0108	0.0166	-0.0249	-0.0302	0.0003	-0.0376	-0.0704	-0.0349	-0.0661
Utilities	0.0013	-0.0053	-0.0066	-0.0677	-0.0088	0.0158	0.0094	0.0226	-0.0085	0.0252	0.0166	-0.0455	-0.0225	-0.0427
Manufacturing	-0.0098	-0.0074	0.0171	-0.1431	-0.0398	0.0405	0.0207	0.0243	-0.0557	0.0627	0.0590	-0.2053	-0.1017	-0.1926
Wholesale trade	-0.0267	-0.0026	-0.0012	-0.0753	-0.0197	-0.0012	-0.0079	0.0325	-0.0040	0.0224	0.0339	-0.0813	-0.0403	-0.0763
Services	0.0162	0.0040	0.0069	0.0110	-0.0065	-0.0204	-0.0184	0.0121	-0.0223	-0.0402	0.0239	-0.1222	-0.0606	-0.1147
Army	0.0068	0.0107	0.0013	-0.0235	-0.0017	0.0192	0.0136	0.0173	0.0103	0.0148	0.0153	-0.0326	-0.0162	-0.0306
Construction	-0.0228	-0.0162	-0.0017	-0.1948	-0.0080	0.0007	-0.0053	0.0129	-0.0500	0.0201	0.0053	-0.1017	-0.0504	-0.0954
Mining	-0.0197	-0.0088	-0.0009	-0.0682	-0.0071	-0.0015	-0.0057	0.0024	-0.0039	0.0236	-0.0133	-0.0395	-0.0196	-0.0371
Transportation	0.0073	-0.0041	-0.0056	-0.1119	-0.0091	0.0342	0.0229	0.0398	-0.0328	0.0319	0.0318	-0.0825	0.8640	-0.0774
Public administration	0.0239	-0.0024	-0.0091	-0.0248	-0.0121	0.0037	-0.0051	0.0268	0.0485	0.0073	0.0304	-0.0854	-0.0423	-0.0802
Number of children	0.0247	0.0125	0.0588	0.0558	-0.0459	0.1185	-0.0108	0.0315	0.0030	0.4293	0.0064	0.0075	0.0143	-0.0577
Head of family	-0.0045	-0.0171	-0.0180	-0.4973	0.0449	0.2496	0.1961	-0.0265	0.1399	0.0971	0.1313	-0.0441	0.1003	-0.0907

A4. Table of Coefficients of Correlation among variables for year 1980, weighted, continue

	Agriculture	Utilities	Manufacturing	Wholesale trade	Services	Army	Construction	Mining	Transportation	Public administration	Number of children	Head of family
Agriculture	1.0000											
Utilities	-0.0166	1.0000										
Manufacturing	-0.0750	-0.0484	1.0000									
Wholesale trade	-0.0297	-0.0192	-0.0866	1.0000								
Services	-0.0446	-0.0289	-0.1302	-0.0516	1.0000							
Army	-0.0119	-0.0077	-0.0348	-0.0138	-0.0207	1.0000						
Construction	-0.0371	-0.0240	-0.1083	-0.0429	-0.0645	-0.0172	1.0000					
Mining	-0.0144	-0.0093	-0.0421	-0.0167	-0.0251	-0.0067	-0.0208	1.0000				
Transportation	-0.0301	-0.0195	-0.0879	-0.0348	-0.0523	-0.0140	-0.0435	-0.0169	1.0000			
Public administration	-0.0312	-0.0202	-0.0910	-0.0361	-0.0542	-0.0145	-0.0451	-0.0175	-0.0366	1.0000		
Number of children	0.0075	0.0102	0.0429	0.0037	-0.0205	0.0037	0.0089	0.0121	0.0149	-0.0006	1.0000	
Head of family	0.0312	0.0616	0.1400	0.0611	-0.0061	0.0342	0.1117	0.0593	0.0974	0.0507	0.1401	1.0000

Appendix B

The estimates of coefficients are reported in the first line for each variable. Their significance is reported using stars; three stars for significance at 1% level, two stars for significance at 5% level, and one star for significance at 10% level. Robust standard errors are reported in the parentheses under the corresponding estimate of coefficient. Probability weights are used for the estimation. This holds for all tables in Appendices B, C and D.

B1. Table containing estimates of possible population models using 1980 data

	Model 1	**Model 2**	**Model 3**
Black	-0.005*	-0.005*	-0.014***
	(0.0029)	(0.0029)	(0.0028)
Asian	0.016***	0.014**	0.024***
	(0.0062)	(0.0062)	(0.0060)
Hispanic	-0.017***	-0.019***	-0.02***
	(0.0036)	(0.0036)	(0.0036)
Female	-0.391***	-0.391***	-0.32***
	(0.0017)	(0.0017)	(0.0018)
Experience	0.048***	0.048***	0.043***
	(0.0002)	(0.0002)	(0.0002)
Experience^2	-0.001***	-0.001***	-0.001***
	($5.42e^{-6}$)	($5.41e^{-6}$)	($5.28e^{-6}$)
Secondary School	0.189***	0.187***	0.179***
	(0.0022)	(0.0022)	(0.0022)
College	0.576***	0.571***	0.599***
	(0.0028)	(0.0028)	(0.0029)
Married	0.146***	0.147***	0.124***
	(0.0018)	(0.0018)	(0.0018)
Disability	0.157***	0.159***	0,155***
	(0.0046)	(0.0046)	(0.0045)
Metropolitan		0.048***	0.048***
		(0.0014)	(0.0014)
Professional			-0.109***
			(0.0029)
Information and Communications			0.285***
			(0.0051)
Retail Trade			-0.227***
			(0.003)
Agriculture			-0.229***
			(0.0066)
Utilities			0.145***
			(0.0054)
Manufacturing			0.132***
			(0.0027)
Wholesale Trade			0.066***
			(0.0039)
Services			-0.188***
			(0.0037)

Army			0.159***
			(0.0071)
Construction			0.119***
			(0.0038)
Mining			0.322***
			(0.0065)
Transportation			-0.072***
			(0.0054)
Public Administration			0.026***
			(0.0038)
Inverse Mills Ratio	-0.725***	-0.723***	-0.689***
	(0.0025)	(0.0026)	(0.0026)
Constant	5.948***	5.937***	5.96***
	(0.0029)	(0.0029)	(0.0038)
Observations	999131	999131	999131
Log pseudolikelihood	$-1.86e^{8}$	$-1.86e^{8}$	$-1.84e^{8}$

Note: For the selection equation was used 1442342 observations

B2. Table containing estimates of possible population models using 2007 data

	Model 1	Model 2	Model 3
Black	-0.011***	-0.025***	-0.019***
	(0.0028)	(0.0028)	(0.0027)
Asian	0.059***	0.031***	0.031***
	(0.0042)	(0.0042)	(0.0041)
Hispanic	-0.029***	-0.051***	-0.055***
	(0.0026)	(0.0026)	(0.0025)
Female	-0.326***	-0.323***	-0.241***
	(0.0017)	(0.0017)	(0.0018)
Experience	0.069***	0.067***	0.062***
	(0.0003)	(0.0003)	(0.0003)
Experience^2	-0.001***	-0.001***	-0.001***
	($6.29e^{-6}$)	($6.39e^{-6}$)	($6.4e^{-6}$)
Secondary School	0.272***	0.267***	0.262***
	(0.0033)	(0.0034)	(0.0034)
College	0.865***	0.838***	0.846***
	(0.0037)	(0.0037)	(0.0038)
Married	0.135***	0.139***	0.124***
	(0.0018)	(0.0018)	(0.0018)
Disability	0.084***	0.109***	0.08***
	(0.0087)	(0.0089)	(0.0091)
Metropolitan		0.188***	0.181***
		(0.0017)	(0.0017)
Professional			-0.117***
			(0.0039)
Information and Communications			-0.068***
			(0.0057)
Retail Trade			-0.381***
			(0.0038)
Agriculture			-0.308***
			(0.0086)
Utilities			0.101***
			(0.0067)
Manufacturing			-0.056***
			(0.0035)
Wholesale Trade			-0.09***
			(0.0049)
Services			-0.379***
			(0.0032)
Army			0.007
			(0.0077)
Construction			-0.042***
			(0.0042)
Mining			0.305***
			(0.0107)
Transportation			-0.104***
			(0.0045)
Public Administration			-0.046***
			(0.0041)
Inverse Mills Ratio	-0.538***	-0.507***	-0.453***
	(0.0047)	(0.0051)	(0.0059)
Constant	5.507***	5.402***	5.619***
	(0.0046)	(0.0049)	(0.0061)
Observations	1346041	1346041	1346041
Log pseudolikelihood	$-2.64e^8$	$-2.63e^8$	$-2.6e^8$

Note: For the selection equation was used 1905194 observations

Appendix C

C1. Table of maximum likelihood estimates for whites, weighted

	1980	2007
Female	-0.364***	-0.288***
	(0.00019)	(0.00017)
Experience	0.045***	0.068***
	(0.00002)	(0.00088)
Experience^2	-0.001***	-0.001***
	($5.7e^{-7}$)	($6.21e^{-7}$)
Secondary School	0.192***	0.266***
	(0.00024)	(0.00039)
College	0.618***	0.857***
	(0.00031)	(0.00042)
Married	0.142***	0.13***
	(0.00019)	(0.00017)
Disability	0.124***	0.045***
	(0.00048)	(0.00088)
Metropolitan	0.053***	0.191***
	(0.00015)	(0.00016)
Professional	-0.122***	-0.1***
	(0.00031)	(0.00038)
Information and Communications	0.266***	-0.091***
	(0.00054)	(0.00056)
Retail Trade	-0.236***	-0.394***
	(0.00032)	(0.00037)
Agriculture	-0.256***	-0.345***
	(0.00075)	(0.00101)
Utilities	0.144***	0.097***
	(0.00057)	(0.00066)
Manufacturing	0.13***	-0.043***
	(0.00029)	(0.00035)
Wholesale Trade	0.063***	-0.075***
	(0.00042)	(0.00048)
Services	-0.178***	-0.395***
	(0.00041)	(0.00032)
Army	0.161***	-0.012***
	(0.00078)	(0.00069)
Construction	0.118***	-0.026***
	(0.0004)	(0.00041)
Mining	0.313***	0.308***
	(0.00067)	(0.00095)
Transportation	-0.064***	-0.107***
	(0.00058)	(0.00045)
Public Administration	0.012***	-0.074***
	(0.00041)	(0.00041)
Inverse Mills Ratio	-0.673***	-0.475***
	(0.00028)	(0.00054)
Constant	5.936***	5.569***
	(0.00041)	(0.0006)

C2. Table of maximum likelihood estimates for Asians, weighted

	1980	2007
Female	-0.203***	-0.141***
	(0.00124)	(0.00068)
Experience	0.039***	0.059***
	(0.00019)	(0.00011)
Experience^2	-0.001***	-0.001***
	($4.15e^{-6}$)	($2.58e^{-6}$)
Secondary School	0.109***	0.144***
	(0.00184)	(0.00142)
College	0.454***	0.735***
	(0.00208)	(0.00154)
Married	0.117***	0.126***
	(0.00138)	(0.00071)
Disability	0.151***	0.131***
	(0.00458)	(0.00326)
Metropolitan	0.019***	0.147***
	(0.00104)	(0.00085)
Professional	-0.015***	-0.006***
	(0.00199)	(0.00127)
Information and Communications	0.357***	-0.047***
	(0.00352)	(0.0019)
Retail Trade	-0.214***	-0.454***
	(0.00208)	(0.00141)
Agriculture	-0.081***	-0.337***
	(0.00439)	(0.00549)
Utilities	0.192***	0.031***
	(0.00411)	(0.00314)
Manufacturing	0.081***	-0.098***
	(0.00194)	(0.00123)
Wholesale Trade	0.054***	-0.213***
	(0.00291)	(0.00191)
Services	-0.122***	-0.395***
	(0.00245)	(0.00113)
Army	0.167***	0.002
	(0.0039)	(0.00306)
Construction	0.233***	-0.094***
	(0.0032)	(0.00203)
Mining	0.266***	0.25***
	(0.00677)	(0.00643)
Transportation	-0.017***	-0.164***
	(0.00384)	(0.00166)
Public Administration	0.043***	-0.062***
	(0.00279)	(0.00165)
Inverse Mills Ratio	-0.669***	-0.479***
	(0.00202)	(0.00201)
Constant	6.061***	5.85***
	(0.00298)	(0.00251)

C3. Table of maximum likelihood estimates for Blacks, weighted

	1980	2007
Female	-0.15***	-0.111***
	(0.00062)	(0.00038)
Experience	0.027***	0.059***
	(0.00008)	(0.00006)
Experience^2	-0.0004***	-0.001***
	($1.84e^{-6}$)	($1.45e^{-6}$)
Secondary School	0.068***	0.283***
	(0.00074)	(0.00076)
College	0.425***	0.839***
	(0.00113)	(0.0009)
Married	0.125**	0.106***
	(0.0006)	(0.00039)
Disability	0.304***	-0.069***
	(0.00158)	(0.00182)
Metropolitan	0.045***	0.197***
	(0.00051)	(0.00043)
Professional	-0.038***	-0.167***
	(0.001)	(0.00091)
Information and Communications	0.358***	0.015***
	(0.00175)	(0.00129)
Retail Trade	-0.182***	-0.339***
	(0.00112)	(0.00084)
Agriculture	-0.299***	-0.299***
	(0.00243)	(0.00351)
Utilities	0.144***	0.123***
	(0.00196)	(0.00171)
Manufacturing	0.156***	-0.038***
	(0.00095)	(0.00084)
Wholesale Trade	0.048***	-0.122***
	(0.00173)	(0.00129)
Services	-0.274***	-0.3***
	(0.00123)	(0.00071)
Army	0.169***	0.127***
	(0.00232)	(0.00152)
Construction	0.069***	-0.009***
	(0.00149)	(0.00115)
Mining	0.277***	0.241***
	(0.00375)	(0.00346)
Transportation	-0.073***	-0.058***
	(0.00187)	(0.00092)
Public Administration	0.068***	0.027***
	(0.00121)	(0.00092)
Inverse Mills Ratio	-0.741***	-0.163***
	(0.00105)	(0.00105)
Constant	6.551***	5.398***
	(0.00139)	(0.00139)

C4. Table of maximum likelihood estimates for Hispanics, weighted

	1980	2007
Female	-0.108***	-0.184***
	(0.00086)	(0.00047)
Experience	0.032***	0.046***
	(0.0001)	(0.00006)
Experience^2	-0.0004***	-0.0007***
	($2.2e^{-6}$)	($1.3e^{-6}$)
Secondary School	0.14***	0.263***
	(0.00082)	(0.00044)
College	0.523***	0.775***
	(0.00145)	(0.00068)
Married	0.091***	0.109***
	(0.00075)	(0.00034)
Disability	0.262***	0.097***
	(0.00219)	(0.00196)
Metropolitan	-0.002***	0.115***
	(0.00061)	(0.00037)
Professional	-0.111***	-0.211***
	(0.00135)	(0.00088)
Information and Communications	0.309***	-0.016***
	(0.00265)	(0.00146)
Retail Trade	-0.15***	-0.323***
	(0.00134)	(0.00086)
Agriculture	-0.08***	-0.294***
	(0.00182)	(0.00111)
Utilities	0.099***	0.072***
	(0.00288)	(0.00206)
Manufacturing	0.079***	-0.137***
	(0.00121)	(0.00081)
Wholesale Trade	0.056***	-0.126***
	(0.00178)	(0.00104)
Services	-0.171***	-0.346***
	(0.00152)	(0.00076)
Army	0.147***	0.004**
	(0.00316)	(0.00177)
Construction	0.114***	-0.101***
	(0.00171)	(0.00084)
Mining	0.307***	0.223***
	(0.00282)	(0.0024)
Transportation	-0.077***	-0.074***
	(0.00283)	(0.00101)
Public Administration	0.045***	0.072***
	(0.00181)	(0.00109)
Inverse Mills Ratio	-0.715***	-0.291***
	(0.00121)	(0.00136)
Constant	6.551***	5.72***
	(0.00139)	(0.00127)

C5. Table of maximum likelihood estimates by race, 1980, weighted

	Whites	Blacks	Hispanics	Asians
Female	-0.364***	-0.15***	-0.108***	-0.203***
	(0.00019)	(0.00062)	(0.00086)	(0.00124)
Experience	0.045***	0.027***	0.032***	0.039***
	(0.00002)	(0.00008)	(0.0001)	(0.00019)
Experience^2	-0.001***	-0.0004***	-0.0004***	-0.001***
	($5.7e^{-7}$)	($1.84e^{-6}$)	($2.2e^{-6}$)	($4.15e^{-6}$)
Secondary School	0.192***	0.068***	0.14***	0.109***
	(0.00024)	(0.00074)	(0.00082)	(0.00184)
College	0.618***	0.425***	0.523***	0.454***
	(0.00031)	(0.00113)	(0.00145)	(0.00208)
Married	0.142***	0.125**	0.091**	0.117***
	(0.00019)	(0.0006)	(0.00075)	(0.00138)
Disability	0.124***	0.304***	0.262***	0.151***
	(0.00048)	(0.00158)	(0.00219)	(0.00458)
Metropolitan	0.053***	0.045***	-0.002***	0.019***
	(0.00015)	(0.00051)	(0.00061)	(0.00104)
Professional	-0.122***	-0.038***	-0.111***	-0.015***
	(0.00031)	(0.001)	(0.00135)	(0.00199)
Information and Communications	0.266***	0.358***	0.309***	0.357***
	(0.00054)	(0.00175)	(0.00265)	(0.00352)
Retail Trade	-0.236***	-0.182***	-0.15***	-0.214***
	(0.00032)	(0.00112)	(0.00134)	(0.00208)
Agriculture	-0.256***	-0.299***	-0.08***	-0.081***
	(0.00075)	(0.00243)	(0.00182)	(0.00439)
Utilities	0.144***	0.144***	0.099***	0.192***
	(0.00057)	(0.00196)	(0.00288)	(0.00411)
Manufacturing	0.13***	0.156***	0.079***	0.081***
	(0.00029)	(0.00095)	(0.00121)	(0.00194)
Wholesale Trade	0.063***	0.048***	0.056***	0.054***
	(0.00042)	(0.00173)	(0.00178)	(0.00291)
Services	-0.178***	-0.274***	-0.171***	-0.122***
	(0.00041)	(0.00123)	(0.00152)	(0.00245)
Army	0.161***	0.169***	0.147***	0.167***
	(0.00078)	(0.00232)	(0.00316)	(0.0039)
Construction	0.118***	0.069***	0.114***	0.233***
	(0.0004)	(0.00149)	(0.00171)	(0.0032)
Mining	0.313***	0.277***	0.307***	0.266***
	(0.00067)	(0.00375)	(0.00282)	(0.00677)
Transportation	-0.064***	-0.073***	-0.077***	-0.017***
	(0.00058)	(0.00187)	(0.00283)	(0.00384)
Public Administration	0.012***	0.068***	0.045***	0.043***
	(0.00041)	(0.00121)	(0.00181)	(0.00279)
Inverse Mills Ratio	-0.673***	-0.741***	-0.715***	-0.669***
	(0.00028)	(0.00105)	(0.00121)	(0.00202)
Constant	5.936***	6.551***	6.551***	6.061***
	(0.00041)	(0.00139)	(0.00139)	(0.00298)

C6. Table of maximum likelihood estimates by race, 2007, weighted

	Whites	Blacks	Hispanics	Asians
Female	-0.288***	-0.111***	-0.184***	-0.141***
	(0.00017)	(0.00038)	(0.00047)	(0.00068)
Experience	0.045***	0.059***	0.046***	0.059***
	(0.00088)	(0.00006)	(0.00006)	(0.00011)
Experience^2	-0.001***	-0.001***	-0.0007***	-0.001***
	($6.21e^{-7}$)	($1.45e^{-6}$)	($1.3e^{-6}$)	($2.58e^{-6}$)
Secondary School	0.266***	0.283***	0.263***	0.144***
	(0.00039)	(0.00076)	(0.00044)	(0.00142)
College	0.857***	0.839***	0.775***	0.735***
	(0.00042)	(0.0009)	(0.00068)	(0.00154)
Married	0.13***	0.106***	0.109***	0.126***
	(0.00017)	(0.00039)	(0.00034)	(0.00071)
Disability	0.045***	-0.069***	0.097***	0.131***
	(0.00088)	(0.00182)	(0.00196)	(0.00326)
Metropolitan	0.191***	0.197***	0.115***	0.147***
	(0.00016)	(0.00043)	(0.00037)	(0.00085)
Professional	-0.1***	-0.167***	-0.211***	-0.006***
	(0.00038)	(0.00091)	(0.00088)	(0.00127)
Information and Communications	-0.091***	0.015***	-0.016***	-0.047***
	(0.00056)	(0.00129)	(0.00146)	(0.0019)
Retail Trade	-0.394***	-0.339***	-0.323***	-0.454***
	(0.00037)	(0.00084)	(0.00086)	(0.00141)
Agriculture	-0.345***	-0.299***	-0.294***	-0.337***
	(0.00101)	(0.00351)	(0.00111)	(0.00549)
Utilities	0.097***	0.123***	0.072***	0.031***
	(0.00066)	(0.00171)	(0.00206)	(0.00314)
Manufacturing	-0.043***	-0.038***	-0.137***	-0.098***
	(0.00035)	(0.00084)	(0.00081)	(0.00123)
Wholesale Trade	-0.075***	-0.122***	-0.126***	-0.213***
	(0.00048)	(0.00129)	(0.00104)	(0.00191)
Services	-0.395***	-0.3***	-0.346***	-0.395***
	(0.00032)	(0.00071)	(0.00076)	(0.00113)
Army	-0.012***	0.127***	0.004**	0.002
	(0.00069)	(0.00152)	(0.00177)	(0.00306)
Construction	-0.026***	-0.009***	-0.101***	-0.094***
	(0.00041)	(0.00115)	(0.00084)	(0.00203)
Mining	0.308***	0.241***	0.223***	0.25***
	(0.00095)	(0.00346)	(0.0024)	(0.00643)
Transportation	-0.107***	-0.058***	-0.074***	-0.164***
	(0.00045)	(0.00092)	(0.00101)	(0.00166)
Public Administration	-0.074***	0.027***	0.072***	-0.062***
	(0.00041)	(0.00092)	(0.00109)	(0.00165)
Inverse Mills Ratio	-0.475***	-0.163***	-0.291***	-0.479***
	(0.00054)	(0.00105)	(0.00136)	(0.00201)
Constant	5.569***	5.398***	5.72***	5.85***
	(0.0006)	(0.00139)	(0.00127)	(0.00251)

Appendix D

D1. Oaxaca decomposition of Black-white logarithm wage differential, 1980

925927 observations **Variables**	**1980** **White wage structure, explained**	**1980** **White wage structure, unexplained**	**1980** **Black wage structure, explained**	**1980** **Black wage structure, unexplained**
Logarithm wage differential	0.099*** (0.00046)	0.099*** (0.00046)	0.099*** (0.00046)	0.099*** (0.00046)
Female	0.017*** (0.00006)	-0.106*** (0.00032)	0.007*** (0.00003)	-0.096*** (0.00029)
Experience	-0.006*** (0.0002)	0.319*** (0.00161)	-0.003*** (0.00012)	0.316*** (0.0016)
Experience^2	-0.002*** (0.00012)	-0.099*** (0.00094)	-0.001*** (0.00008)	-0.099*** (0.00095)
Secondary School	0.008*** (0.00003)	0.069*** (0.00043)	0.002*** (0.00003)	0.074*** (0.00046)
College	0.052*** (0.00006)	0.018*** (0.00011)	0.036*** (0.0001)	0.034*** (0.00021)
Married	0.022*** (0.00004)	0.062*** (0.00028)	0.0001** (0.00009)	0.084*** (0.00037)
Disability	-0.0009*** ($9.89e^{-6}$)	-0.009*** (0.00009)	-0.002*** (0.00002)	-0.008*** (0.00008)
Metropolitan	0.0004*** ($8.01e^{-6}$)	0.002*** (0.00015)	0.0003*** ($7.86e^{-6}$)	0.002*** (0.00015)
Professional	0.004*** (0.00002)	-0.019*** (0.00023)	0.001*** (0.00003)	-0.016*** (0.0002)
Information and Communications	-0.002*** (0.00002)	-0.006*** (0.00012)	-0.003*** (0.00003)	-0.005*** (0.0001)
Retail Trade	-0.012*** (0.00003)	-0.006*** (0.00013)	-0.009*** (0.00006)	-0.009*** (0.0002)
Agriculture	-0.00005*** (0.00001)	0.0006*** (0.00004)	-0.00006*** (0.00001)	0.0006*** (0.00004)
Utilities	-0.0001*** ($5.84e^{-6}$)	$3.01e^{-6}$ (0.00003)	-0.0001*** ($6.05e^{-6}$)	$2.84e^{-6}$ (0.00003)
Manufacturing	$6.92e^{-6}$ (0.00001)	-0.006*** (0.00023)	$8.28e^{-6}$ (0.00002)	-0.006*** (0.00023)
Wholesale Trade	0.001*** ($7.73e^{-6}$)	0.0004*** (0.00004)	0.0008*** (0.00002)	0.0006*** (0.00007)
Services	0.004*** (0.00002)	0.009*** (0.00012)	0.007*** (0.00004)	0.006*** (0.00009)
Army	-0.0004*** ($5.41e^{-6}$)	-0.00007*** (0.00002)	-0.0004*** ($8.16e^{-6}$)	-0.00005*** (0.00001)
Construction	0.001*** (e^{-5})	0.002*** (0.00006)	0.0009*** (0.00002)	0.002*** (0.00009)
Mining	0.002*** ($8.83e^{-6}$)	0.0001*** (0.00001)	0.002*** (0.00002)	0.0004*** (0.00004)
Transportation	0.0005*** ($6.94e^{-6}$)	0.0004*** (0.0001)	0.0006*** (0.00001)	0.0003*** (0.00008)
Public Administration	-0.0002*** ($8.05e^{-6}$)	-0.003*** (0.00008)	-0.001*** (0.00002)	-0.002*** (0.00006)
Constant		-0.218*** (0.00145)		-0.218*** (0.00145)
Total	0.089*** (0.00015)	0.01*** (0.00046)	0.036*** (0.00019)	0.062*** (0.00041)

D2. Oaxaca decomposition of Black-white logarithm wage differential, 2007

1126497 observations **Variables**	**2007** **White wage structure, explained**	**2007** **White wage structure, unexplained**	**2007** **Black wage structure, explained**	**2007** **Black wage structure, unexplained**
Logarithm wage differential	0.29615*** (0.0005)	0.29615*** (0.0005)	0.29615*** (0.0005)	0.29615*** (0.0005)
Female	0.013*** (0.00003)	-0.094*** (0.00022)	0.005*** (0.00002)	-0.086*** (0.0002)
Experience	0.084*** (0.00022)	0.187*** (0.00135)	0.072*** (0.0002)	0.199*** (0.00143)
Experience^2	-0.069*** (0.00015)	-0.061*** (0.00084)	-0.062*** (0.00016)	-0.068*** (0.00094)
Secondary School	-0.022*** (0.00004)	-0.012*** (0.00018)	-0.023*** (0.00007)	-0.01*** (0.00053)
College	0.106*** (0.0001)	0.003*** (0.00018)	0.104*** (0.00014)	0.005*** (0.00031)
Married	0.028*** (0.00004)	0.007*** (0.00013)	0.023*** (0.00008)	0.012*** (0.00022)
Disability	-0.001*** (4.59e^{-6})	0.003*** (0.00005)	0.0003*** (9.06e^{-6})	0.002*** (0.00004)
Metropolitan	-0.01*** (0.00002)	-0.003*** (0.00032)	-0.01*** (0.00003)	-0.003*** (0.0003)
Professional	-0.0004*** (7.58e^{-6})	0.005*** (0.00008)	-0.0006*** (0.00001)	0.006*** (0.00009)
Information and Communications	-0.0001*** (3.91e^{-6})	-0.002*** (0.00003)	0.00003*** (2.8e^{-6})	-0.002*** (0.00003)
Retail Trade	-0.002*** (0.00003)	-0.006*** (0.0001)	-0.002*** (0.00002)	-0.006*** (0.00011)
Agriculture	-0.001*** (7.17e^{-6})	-0.0001*** (0.00001)	-0.001*** (0.00001)	-0.0003*** (0.00002)
Utilities	0.0002*** (2.72e^{-6})	-0.0001*** (0.00001)	0.0002*** (4.9e^{-6})	-0.0002*** (0.00001)
Manufacturing	-0.0008*** (7.67e^{-6})	-0.0005*** (0.00009)	-0.0007*** (0.00001)	-0.0006*** (0.0001)
Wholesale Trade	-0.0008*** (6.06e^{-6})	0.0009*** (0.00002)	-0.001*** (0.00001)	0.001*** (0.00004)
Services	0.017*** (0.00005)	-0.037*** (0.00031)	0.013*** (0.00005)	-0.033*** (0.00027)
Army	0.00001*** (1.04e^{-6})	-0.001*** (0.00001)	-0.0001*** (3.82e^{-6})	-0.001*** (0.00001)
Construction	-0.0007*** (0.00001)	-0.0006*** (0.00004)	-0.0002*** (0.00003)	-0.001*** (0.00007)
Mining	0.001*** (5.42e^{-6})	0.0001*** (6.98e^{-6})	0.0009*** (0.00001)	0.0003*** (0.00002)
Transportation	0.002*** (0,00001)	-0.003*** (0.00006)	0.001*** (0.00002)	-0.002*** (0.00004)
Public Administration	0.001*** (8.79e^{-6})	-0.006*** (0.00006)	-0.0005*** (0.00001)	-0.005*** (0.00005)
Constant		0.171*** (0.00151)		0.171*** (0.00151)
Total	0.147*** (0.00015)	0.149*** (0.00049)	0.118*** (0.00017)	0.177*** (0.00047)

D3. Oaxaca decomposition of Hispanic-white logarithm wage differential, 1980

877799 observations **Variables**	**1980 White wage structure, explained**	**1980 White wage structure, unexplained**	**1980 Hispanic wage structure, explained**	**1980 Hispanic wage structure, unexplained**
Logarithm wage differential	0.098*** (0.00052)	0.098*** (0.00052)	0.098*** (0.00052)	0.098*** (0.00052)
Female	-0.012*** (0.00007)	-0.106*** (0.00037)	-0.003*** (0.00003)	-0.114*** (0.00039)
Experience	0.016*** (0.00025)	0.236*** (0.00185)	0.011*** (0.00018)	0.241*** (0.00189)
Experience^2	-0.019*** (0.00015)	-0.063*** (0.00105)	-0.014*** (0.00014)	-0.067*** (0.00112)
Secondary School	0.029*** (0.00005)	0.023*** (0.00038)	0.021*** (0.00013)	0.031*** (0.00051)
College	0.067*** (0.00007)	0.006*** (0.0001)	0.056*** (0.00016)	0.017*** (0.00026)
Married	0.004*** (0.00003)	0.029*** (0.00044)	0.003*** (0.00003)	0.031*** (0.00047)
Disability	0.0001*** (0.00001)	-0.005*** (0.00008)	0.002*** (0.00003)	-0.006*** (0.0001)
Metropolitan	-0.001*** (0.00001)	0.017*** (0.0002)	0.00005*** (0.00001)	0.016*** (0.00019)
Professional	-0.006*** (0.00002)	-0.001*** (0.00019)	-0.006*** (0.00007)	-0.002*** (0.00027)
Information and Communications	0.001*** (0.00002)	-0.002*** (0.00013)	0.002*** (0.00003)	-0.002*** (0.00015)
Retail Trade	-0.005*** (0.00003)	-0.012*** (0.0002)	-0.003*** (0.00003)	-0.014*** (0.00023)
Agriculture	0.009*** (0.00003)	-0.009*** (0.0001)	0.002*** (0.00006)	-0.002*** (0.00003)
Utilities	0.0005*** (6.94e^{-6})	0.0004 (0.00003)	0.0003*** (0.00001)	0.00065 (0.00004)
Manufacturing	-0.005*** (0.00002)	0.013*** (0.00034)	-0.003*** (0.00005)	0.011*** (0.00028)
Wholesale Trade	0.00001*** (5.6e^{-6})	0.0003*** (0.00007)	0.00001*** (4.94e^{-6})	0.00032*** (0.00007)
Services	0.003*** (0.00002)	-0.0006*** (0.00014)	0.003*** (0.00003)	-0.0005*** (0.00011)
Army	-0.0001*** (6.16e^{-6})	0.0001*** (0.00002)	-0.0001*** (6.59e^{-6})	0.00009*** (0.00002)
Construction	-0.0003*** (0.00001)	0.0002** (0.0001)	-0.0003*** (0.00001)	0.0002** (0.0001)
Mining	0.0002*** (0.00001)	0.00006** (0.00003)	0.0002*** (0.00001)	0.00007** (0.00003)
Transportation	-0.0001*** (5.67e^{-6})	0.0005*** (0.00011)	-0.0002*** (9.48e^{-6})	0.0005*** (0.00012)
Public Administration	0.0007*** (2.79e^{-6})	-0.001*** (0.00007)	0.0003*** (0.00001)	-0.001*** (0.00008)
Constant		-0.113*** (0.00166)		-0.113*** (0.00166)
Total	0.085*** (0.0002)	0.012*** (0.00052)	0.074*** (0.00026)	0.024*** (0.00053)

D4. Oaxaca decomposition of Hispanic-white logarithm wage differential, 2007

1147351 observations **Variables**	**2007** **White wage structure, explained**	**2007** **White wage structure, unexplained**	**2007** **Hispanic wage structure, explained**	**2007** **Hispanic wage structure, unexplained**
Logarithm wage differential	0.304*** (0.00061)	0.304*** (0.00061)	0.304*** (0.00061)	0.304*** (0.00061)
Female	-0.021*** (0.00003)	-0.042*** (0.0002)	-0.013*** (0.00004)	-0.05*** (0.00024)
Experience	0.127*** (0.00021)	0.427*** (0.00125)	0.085*** (0.00017)	0.47*** (0.00137)
Experience^2	-0.103*** (0.00015)	-0.173*** (0.00071)	-0.07*** (0.00015)	-0.206*** (0.00085)
Secondary School	0.013*** (0.00003)	0.001*** (0.00033)	0.013*** (0.00004)	0.001*** (0.00036)
College	0.161*** (0.00011)	0.009*** (0.00009)	0.146*** (0.00014)	0.025*** (0.00025)
Married	0.012*** (0.00002)	0.008*** (0.00016)	0.01*** (0.00003)	0.01*** (0.0002)
Disability	0.0002*** ($5.97e^{-6}$)	-0.0008*** (0.00003)	0.0006*** (0.00001)	-0.001*** (0.00005)
Metropolitan	-0.016*** (0.00002)	0.055*** (0.0003)	-0.009*** (0.00003)	0.049*** (0.00026)
Professional	0.0003*** ($7.37e^{-6}$)	0.01*** (0.00009)	0.0006*** (0.00001)	0.01*** (0.00008)
Information and Communications	-0.0009*** ($6.3e^{-6}$)	-0.001*** (0.00002)	-0.0001*** (0.00001)	-0.002*** (0.00004)
Retail Trade	-0.004*** (0.00003)	-0.008*** (0.0001)	-0.003*** (0.00002)	-0.008*** (0.00011)
Agriculture	0.007*** (0.00002)	-0.001*** (0.00004)	0.006*** (0.00002)	-0.0004*** (0.00001)
Utilities	0.0004*** ($3.66e^{-6}$)	0.0001 (0.00001)	0.0003*** ($9.91e^{-6}$)	0.0002 (0.00002)
Manufacturing	0.0003*** ($4.46e^{-6}$)	0.011*** (0.00011)	0.001*** (0.00001)	0.011*** (0.0001)
Wholesale Trade	0.0001*** ($3.4e^{-6}$)	0.001*** (0.00003)	0.0001*** ($5.78e^{-6}$)	0.001*** (0.00003)
Services	-0.013*** (0.00004)	-0.015*** (0.00026)	-0.011*** (0.00004)	-0.017*** (0.00029)
Army	-0.00002*** ($1.54e^{-6}$)	-0.00008*** (0.00001)	$7.85e^{-6}$** ($3.89e^{-6}$)	-0.0001*** (0.00001)
Construction	0.001*** (0.00002)	0.009*** (0.00012)	0.006*** (0.00005)	0.004*** (0.00006)
Mining	0.0004*** ($5.29e^{-6}$)	0.0003*** (0.00001)	0.0003*** ($5.27e^{-6}$)	0.0004*** (0.00001)
Transportation	0.0001*** ($5.26e^{-6}$)	-0.001*** (0.00004)	0.0001*** ($3.92e^{-6}$)	-0.001*** (0.00004)
Public Administration	-0.001*** ($8.41e^{-6}$)	-0.004*** (0.00003)	0.001*** (0.00002)	-0.007*** (0.00005)
Constant		-0.15*** (0.00141)		-0.15*** (0.00141)
Total	0.165*** (0.00016)	0.138*** (0.00062)	0.165*** (0.00016)	0.14*** (0.0006)

D5. Oaxaca decomposition of Asian-white logarithm wage differential, 1980

838723 observations **Variables**	**1980** **White wage structure, explained**	**1980** **White wage structure, unexplained**	**1980** **Asian wage structure, explained**	**1980** **Asian wage structure, unexplained**
Logarithm wage differential	-0.048*** (0.00093)	-0.048*** (0.00093)	-0.048*** (0.00093)	-0.048*** (0.00093)
Female	0.012*** (0.00014)	-0.077*** (0.0006)	0.007*** (0.00009)	-0.072*** (0.00056)
Experience	0.053*** (0.00043)	0.1*** (0.00314)	0.046*** (0.00043)	0.107*** (0.00336)
Experience^2	-0.044*** (0.00026)	-0.005*** (0.00176)	-0.044*** (0.0004)	-0.006*** (0.00207)
Secondary School	0.022*** (0.00008)	0.04*** (0.0009)	0.013*** (0.00022)	0.05*** (0.0011)
College	-0.095*** (0.00023)	0.055*** (0.0007)	-0.07*** (0.00036)	0.03*** (0.0004)
Married	0.00003 (0.00005)	0.015*** (0.00084)	0.00003 (0.00004)	0.015*** (0.00084)
Disability	0.003*** (0.00002)	-0.0006*** (0.0001)	0.003*** (0.00011)	-0.001*** (0.0002)
Metropolitan	-0.002*** (0.00002)	0.011*** (0.00037)	-0.001*** (0.00005)	0.009*** (0.00031)
Professional	0.004*** (0.00004)	-0.024*** (0.00047)	0.0006*** (0.00007)	-0.021*** (0.0004)
Information and Communications	0.002*** (0.00004)	-0.004*** (0.00017)	0.003*** (0.00006)	-0.005*** (0.0002)
Retail Trade	0.002*** (0.00007)	-0.004*** (0.00038)	0.002*** (0.00006)	-0.004*** (0.00036)
Agriculture	-0.0001*** (0.00002)	-0.003 (0.00007)	-0.00005*** ($8.23e^{-6}$)	-0.003 (0.00007)
Utilities	0.0008*** (0.00001)	-0.0004*** (0.00003)	0.001*** (0.00003)	-0.0007*** (0.00006)
Manufacturing	0.003*** (0.00004)	0.01*** (0.0004)	0.002*** (0.00005)	0.011*** (0.0004)
Wholesale Trade	0.0005*** ($9.64e^{-6}$)	0.0003*** (0.0001)	0.0004*** (0.00002)	0.0004*** (0.0001)
Services	0.003*** (0.00004)	-0.005*** (0.00022)	0.002*** (0.00005)	-0.004*** (0.00017)
Army	-0.001*** (0.00001)	-0.0001* (0.00006)	-0.001*** (0.00004)	-0.00004* (0.00002)
Construction	0.003*** (0.00002)	-0.003*** (0.0001)	0.006*** (0.0001)	-0.007*** (0.0002)
Mining	0.002*** (0.00001)	0.0001*** (0.00002)	0.002*** (0.00006)	0.0005*** (0.00008)
Transportation	-0.0003*** ($9.86e^{-6}$)	0.004*** (0.00014)	-0.0008*** (0.00003)	0.004*** (0.00016)
Public Administration	-0.00001*** ($2.08e^{-6}$)	-0.001*** (0.00012)	0.00006*** ($7.87e^{-6}$)	-0.001*** (0.00013)
Constant		-0.124*** (0.00301)		-0.124*** (0.00301)
Total	-0.03*** (0.00037)	-0.018*** (0.00089)	-0.027*** (0.0004)	-0.021*** (0.00098)

D6. Oaxaca decomposition of Asian-white logarithm wage differential, 2007

1056643 observations **Variables**	**2007** **White wage structure, explained**	**2007** **White wage structure, unexplained**	**2007** **Asian wage structure, explained**	**2007** **Asian wage structure, unexplained**
Logarithm wage differential	-0.148*** (0.00096)	-0.148*** (0.00096)	-0.148*** (0.00096)	-0.148*** (0.00096)
Female	-0.0008*** (0.00005)	-0.071*** (0.00034)	-0.0004*** (0.00003)	-0.071*** (0.00034)
Experience	0.099*** (0.00033)	0.179*** (0.00231)	0.086*** (0.00033)	0.193*** (0.00248)
Experience^2	-0.082*** (0.00023)	-0.019*** (0.00137)	-0.08*** (0.0003)	-0.022*** (0.00158)
Secondary School	0.049*** (0.00009)	0.053*** (0.00064)	0.026*** (0.00026)	0.076*** (0.0009)
College	-0.155*** (0.00018)	0.06*** (0.00078)	-0.133*** (0.00031)	0.037*** (0.0005)
Married	-0.003*** (0.00002)	0.002*** (0.0004)	-0.003*** (0.00003)	0.002*** (0.0004)
Disability	0.0003*** (6.67e^{-6})	-0.001*** (0.00005)	0.0009*** (0.00002)	-0.002*** (0.00008)
Metropolitan	-0.029*** (0.00004)	0.035*** (0.0007)	-0.022*** (0.00013)	0.028*** (0.00056)
Professional	0.003*** (0.00001)	-0.011*** (0.00016)	0.0002*** (0.00004)	-0.008*** (0.00012)
Information and Communications	0.0004*** (6.74e^{-6})	-0.001*** (0.00006)	0.0002*** (9.09e^{-6})	-0.001*** (0.00005)
Retail Trade	-0.005*** (0.00005)	0.006*** (0.00016)	-0.006*** (0.00006)	0.007*** (0.00017)
Agriculture	-0.001*** (9.19e^{-6})	-0.00002 (0.00002)	-0.001*** (0.00002)	-0.00006 (0.00004)
Utilities	0.0004*** (4.23e^{-6})	0.0003*** (0.00002)	0.0001*** (0.00001)	0.0006*** (0.00003)
Manufacturing	0.0007*** (8.51e^{-6})	0.007*** (0.00017)	0.002*** (0.00002)	0.007*** (0.00015)
Wholesale Trade	-0.0001*** (5.14e^{-6})	0.004*** (0.00006)	-0.0004*** (0.00001)	0.004*** (0.00006)
Services	0.009*** (0.00007)	0.0003 (0.0004)	0.009*** (0.00008)	0.0003 (0.0004)
Army	-0.00001*** (1.16e^{-6})	-0.00008*** (0.00002)	3.28e^{-6} (4.85e^{-6})	-0.0001*** (0.00002)
Construction	-0.001*** (0.00001)	0.001*** (0.00005)	-0.003*** (0.00008)	0.004*** (0.00013)
Mining	0.001*** (6.45e^{-6})	0.00007*** (8.7e^{-6})	0.001*** (0.00002)	0.0003*** (0.00004)
Transportation	-0.00003*** (8.14e^{-6})	0.002*** (0.00006)	-0.00004*** (0.00001)	0.002*** (0.00006)
Public Administration	-0.001*** (7.59e^{-6})	-0.0004*** (0.00006)	-0.0008*** (0.00002)	-0.0005*** (0.00008)
Constant		-0.28*** (0.00258)		-0.28*** (0.00258)
Total	-0.116*** (0.00023)	-0.032*** (0.00094)	-0.125*** (0.00028)	-0.022*** (0.00102)