Alexander-Williams-Von-Neumann Realism and Dynamic Gaps in Returns to Productive Characteristics Among College Educated Workers

Job Market Paper (AEFP 2025 and IAFFE 2025)

Justin M. Adekanbi-Coger*

March 13, 2025

Abstract

This research utilizes the work of Classical, Neoclassical, and Keynesian economists, feminist theorists, and statisticians to study differentials in salaries, wage growth and returns to effort and self-reported positive attitudes toward hard work among college educated workers in the United States. The concept of intersectionality pioneered by Kimberle Crenshaw and extended to economics by bell hooks, Rhonda Williams, and William Darity Jr. is combined with the contributions of Marx, Adam Smith, and W. Arthur Lewis on technologically advanced stratified economies, statisticians on subjective probability under conditions of uncertainty, machine learning variable selection, and statistical inference techniques. I find evidence of substantial unexplained gaps in salaries and returns to effort and self-reported positive attitudes toward hard work applied to knowledge acquisition between white and non-white college graduates from selective colleges. Another contribution of this paper is to transition beyond a transactional understanding of wage differentials to an understanding that the maintenance of discrimination in the labor market and alienation of those outside a white phenotype serves a useful purpose for those wishing to maintain dominion over resources in a world of Knightian uncertainty.

*I want to thank Olaide Adekanbi, my Public Allies colleagues, Bonnie Meszaros, Liz Nash-Farley Ripple, Evangelos Falaris, Charles Link, Saul Hoffman, David Black, Barton Willage, Madalyn Romberger, Huriya Jabbar, Rosalie Rolon Dow, Shari Watkins, James E. Davis, Jamila Lyn, Stanley J. Collins, Michael Payne II, Dilek Azgin, the participants of the START statistical training workshop in Mexico City at CIDE, Jocelyn Delgado Miranda, Gregory Price, John Handy as a brilliant instructor and theorist, Michelle Holder, Bob Putnam, Angelino Viceizcas, Juliet Elu, Mona Ray, David Poyer, Nancy Jordan, Henry May, William Darity Jr, Bradley Hardy, Karl Boulware, Adrienne Lucas, Olga Gorbachev, Maria Rosales-Rueda, James Butkiewicz, James Mulligan, Kent Messer, Kevin Guidry, Mona Ray, Nikki Lynn-Llloyd, Emmanuel Balogun, Bryan Coger, Michelle Coger, Bernice Handy Coger, William Coger, Will E. Coger, my brothers and sisters, Marilyn Allen, Joseph Allen, Erskine Tucker, my Godparents, Olodumare and the Orishas (specifically Orunmila and Osanyin), the historical Jesus Christ of Nazareth, my Mohawk, West African, and European ancestors, the fauna, flora, and water that sustains us, and our common humanity and collective struggle.

Keywords: Intersectional Wage Differentials, Classical Competition, Bayesian Belief Formation, LASSO, Oligopoly, Double Gaps, Lexiographic Preferences

JEL Codes: B5-Current Heterodox Approaches; B11-Bayesian Analysis: General; C4-Econometric and Statistical Methods: Special Topics; C6-Mathematical Methods; C71-Cooperative Games; D25-Intertemporal Firm Choice: Investment, Capacity, and Financing; D41-Perfect Competition; D42-Monopoly; D43-Oligopoly and Other Forms of Market Imperfection; D63-Equity, Justice, Inequality, and Other Normative Criteria and Measurement; D81-Criteria for Decision-Making under Risk and Uncertainty; E12-General Aggregative Models: Keynesian; E71-Role and Effects of Psychological, Emotional, Social, and Cognitive Factors on the Macro Economy; F6-Economic Impacts of Globalization; I3-Welfare, Well-Being, and Poverty; J00-General: Labor and Demographic Economics

1 Introduction

Selective undergraduate institutions are often presented in policy discussions around improving the lives of disproportionately Black and Latino low-income citizens as "equalizers" for graduates from marginalized groups. However, it is certainly plausible that their hard work and graduation from a selective school do not exclude these college graduates from racial wage differentials. There is an old Black proverb that one must work twice as hard to get half as far as a white person with similar aspirations, due to persistent racism and discrimination.

That proverb is the inspiration behind the research questions explored. I estimate the effects of college selectivity and positive attitudes toward hard work on the wages of white and marginalized college graduates to understand how these effects might influence the racial wage differential between college graduates across various majors in the NLSY97. The main research questions addressed are the following:

1. How do the combined effects of college selectivity and positive attitudes toward hard work on the wages of white and marginalized graduates differ? In addition, what are the wage costs to the marginalized group and the wage benefits for the white graduates?

2. How do these differences in the combined effects of college selectivity and positive attitudes toward hard work affect the explained and unexplained component of the racial wage differential experienced by graduates of each major from marginalized groups?

This paper is also motivated by the idea that we need to understand which investments in education improve the living conditions of the poor and marginalized. The first essay of my PhD thesis seeks to understand how high school STEM education can change labor market outcomes, and this paper seeks to understand the ways in which even the most

exceptional Black, Latino, and mixed-race college graduates get less renumeration for their hard work and dedication to scholastic pursuits. Ideally, one could walk away from reading the thesis and understand more about the labor market outcomes for an individual from a non-white, marginalized group that decides to develop STEM skills in high school, has positive attitudes towards hard work, and attends a selective undergraduate institution. These three things are regularly touted as some of the essential decisions to succeed for working-class, non-white youth in the United States. Ultimately, there is considerable reason to question this prescription for prosperity, as these recommendations fail to consider some of the fundamental, systemic features of our economy and society that prevent this advice from being more fruitful.

I study the effects of an interaction between college selectivity and attitude toward hard work on racial wage differentials between college graduates. I found that the hard work and college selectivity interaction accounts for the most significant portion of the unexplained gap in wages between the marginalized and white graduates when controlling for attainment of a graduate degree, ASVAB percentile score, years of experience working, whether health limits the amount of work one can perform, and majoring in business, marketing, or management. In comparison with graduates that are white men, non-white women graduates experience an unexplained wage gap of 25.60%, where 14.11% is a wage cost of being a marginalized woman and 10.07% is the wage benefit of being a white man. Non-white graduates experience an unexplained differential of 15.48% in comparison with graduates that are white men, with 6.82% a wage cost for being non-white graduates and the remaining percentage a wage benefit to being white men.

1.1 Intergenerational Transfers and Income Distribution

A model by Loury (1981b) explains the persistence of income and education quality inequities without reference to marginalized identities, strictly based on unequal income distribution and random ability endowments. Loury presents a model of intergenerational transfers of ability and training that is dependent on parents income. The model is of the overlapping generations (OLG) variety, with three features which characterize the world as presented: the production function, ability endowments, and the utility function. The unit of analysis in the model is families, with a family guardian making the decisions regarding consumption and investment in children. Income is divided each period between consumption and education of the child. Each person begins life with a random endowment of innate aptitudes that can be developed through investment in education and training. Loury assumes that the level of the endowment becomes known during maturity, an assumption most likely made for simplicity but which does not quite reflect reality (children are 'tracked' by observed presumed ability early during their K-12 careers).

Each mature person in the model has output that is a function of ability and training, with each household presumed to have access to the same training. This assumption of equality in schooling quality is also made for the sake of tractability and clarity, but does not reflect the lived experience of students in the United States. Access to high-quality schooling is quite dependent on location and the distribution of earnings and wealth. Introducing the realism of actual wealth and income distributions and racial dynamics will likely exacerbate the differences generated from this model. For example, there is considerable causal evidence of lighter complexioned children receiving more investment in their education and development than their darker complexioned siblings in Brazil (Dondero and Marteleto, 2016). Also, many guardians still hold harmful views about the capabilities or role of women in society and raise their daughters accordingly. The production function is expressed as:

$$x = h(\alpha, e)$$

where x is individual output, $h(..)$ is technology assumed to be available to all households, α is 'ability' endowment, and e is training (or education) (Loury, 1981b). The earnings of children is assumed to be a random variable that depends on individual output and the random distribution of the 'ability' endowment. Guardians relate their children's prospective income to the child's happiness through a function Loury calls the indirect utility function, V. This function is consistent if and only if $V*$ solves :

$$[\forall y \geq 0, V^*(y) = max_{0 \leq c \leq y} E_\alpha U(c, V^*(h(a, y - c)))]$$

where $U(c, v)$ is parents utility which is a function of consumption and indirect utility, where E_α is the expectation operator under the distribution of 'ability' endowment. Loury theorizes that the earnings of an adult (x) depends on ones random ability and parent's earnings (y) in the following way: $x = h(a, e * (y))$. This relationship is a stochastic link between parent's and offspring's earnings (Loury, 1981b). Moreover, the utility obtained from having a particular level of income, $V * (y)$, depends on the extent of social mobility in the society in the question (the United States in our context). Concerning production we assume diminishing returns to training, a strictly positive marginal product of ability, that the net marginal return to training eventually becomes negative for all individuals, and the net average return is always negative for those with cognitive functioning not best suited for the prevalent schooling system (think those with learning disabilities). We are talking some inheritable aspect of measured verbal, mathematical, or visual spatial cognitive functioning. Intelligence is proven to be fluid, improvable, and not completely understood. We are discussing the odds of say, someone able to divide six digit numbers in their head with atypical memory recall having a biological child with similar traits. It is crucial to

distinguish that which is being said to discourage pseudo-scientific Eugenicist arguments. Importantly, Loury assumes an atomless distribution of ability with connected and bounded support, ruling out auto correlation in innate endowments, and makes extreme values of ability sufficiently likely that guardians will always want to invest some income in their children regardless of endowment (Loury, 1981b).

This model quickly illustrates the dilemma regarding public investments in the education system in the United States. Even if assuming no relationship between guardians propensity for scholastic pursuits and their childs, those with higher-income guardians will continue to benefit from the greater investments in their education which their guardians can afford. Over generations, these disadvantages compound without intervention of some form through the government, whether that be investment in the public schooling system or the adoption of various market-based solutions (which we are seeing now with vouchers and school choice movements). With regard to STEM reform in an increasingly STEM skills intensive economy, it quickly becomes apparent that stark differences in the income and social network of parents will result in stark differences in exposure to rigorous and high quality advanced high school mathematics coursework and persistent gaps in outcomes.

1.2 Persistent Beckerian Taste-Based Discrimination and Learning-by-Doing

One of the puzzles of the neoclassical economics of race has been persistence of racial wage discrimination in the long-run. Such an outcome is deemed inefficient and bound to fail any firm operating in this manner. Andrews (Andrews, 1999a) presents a model with a convincing explanation of how discrimination in who gets to develop their skill-set can result in persistent racial differences in economic conditions that are not the result of "cultural" differences between two groups. Suppose there is a one-commodity economy in which firms

produce one composite good in an environment with perfect competition. The economy is open and capital is perfectly mobile. Rate of return on capital is uniform across the world at a rate of $r*$. The firms are small and owner-managed with two inputs, labor (L) and capital (K) and produce output according to the following production function:

$$Y^j = min[K^j, hE^j]$$

for firm j, where E^j is the level of employment offered by firm j and h is the efficiency of labor. Imagine that this is right after the end of Reconstruction and the vast majority of enterprises are owned by whites, some who are "color-blind" and others who only hire and work with other whites. If we are assuming a situation in which this is surplus labor, the colorblind enterprise owner owns $0 < \lambda < 1$ percent of capital and produce $Y^c = \lambda K$ units of output. Racist and color-blind firm owners deem it profitable if the rate of profit in local markets is at or above the world rate of return on capital (hence there is a clear scenario where outsourcing occurs and marginalized and white workers both suffer).

Assume that with a real wage for white workers w, the rate of profit for the racist enterprise owner is $r^w = r^w = 1 - \frac{w}{h^w}$ where h^w is the efficiency of white labor. Color-blind firm owners distribute job offers to marginalized (Black, mixed-race, and Latino) and white workers so that the rate of profit from hiring marginalized labor is equal to the rate associated with hiring white labor, $r^b = 1 - \frac{\beta}{h^b}$ and β is the real wage earned by marginalized workers and h^b is the efficiency of marginalized labor.

Firms pay efficiency wages to prevent lack of effort in the workplace. The utility of work is the difference between real wages the cost of work effort. Employment rates for marginalized and white workers depend on the proportions of colorblind and racist firm owners, the fraction of the work force that is marginalized, and the racial composition of employment offered by the colorblind enterprise owner. Colorblind enterprise owners offer jobs to marginalized

workers to equate returns from using each type of worker. In addition, racist and colorblind firm owners move capital between the local markets and world economies until the rates of return are equal to the world rate of return. If marginalized labor is more efficient that the economy-wide average degree of labor efficiency, the fraction jobs offered to marginalized workers is greater than or equal to the proportion of marginalized workers in the labor force (Andrews, 1999a).

Many models of neoclassical and Post-Keynesian economic growth now assume that labor efficiency is the result of on-the-job training and experience and conscious, deliberate investments in education. Andrews models this insight by assuming that the growth rate of labor efficiency for workers in each racial group depends on the employment rate for the group:

$$Dh^i = (\epsilon e^i - \delta)h^i, i = m, w$$

where $D = d/dt$ is the differential operator, $0 > \delta > 1$ is depreciation of labor skills and $\epsilon > 0$ is the skill development or learning coefficient associated with employment. Andrews presents a model with equations for the employment rates in the short run, the medium-term levels of the capital-output ratio and racial composition of employment for colorblind firm owners, and the growth rate of labor efficiency dependent on opportunities to receive additional training. The result is a two-dimensional nonlinear system in labor efficiency units studied using dynamical methods (Andrews, 1999a). The results indicate that if work experience matters for further development of productive capabilities and increases in wages, competitive capitalism is not sufficient to guarantee the closure of gaps produced by discrimination.

1.3 How the Denial of Actualization of Talent and Effort for Non-White Americans Hinders the Economy

In their paper *One Giant Leap: Emancipation and Aggregate Economic Gains* Richard Hornbeck and Trevon Logan (Hornbeck and Logan, 2022) analyze the economic impact of emancipation in the United States, focusing on how the abolition of slavery led to significant aggregate economic gains despite a reduction in output. They argue that the coercive nature of slavery imposed immense costs on enslaved individuals, resulting in a misallocation of labor that, when corrected through emancipation, enhanced overall economic efficiency.

Emancipation in the United States led to significant economic gains by eliminating the inefficiencies of slavery. This document explores the mathematical framework used by Hornbeck and Logan to estimate the aggregate economic gains from emancipation. The authors' analysis challenges the conventional view that slavery was economically productive by centering the costs imposed on enslaved individuals. Their findings suggest that emancipation not only addressed a profound moral failing but also corrected a significant market failure, leading to immediate and substantial improvements in aggregate economic performance.

The core of their methodology involves comparing the economic value extracted by enslavers with the costs imposed on enslaved individuals. This is expressed mathematically as:

$$\text{Total Economic Gain} = \sum_i (V_i - C_i) \tag{1}$$

where:

- V_i represents the value of output produced by an enslaved individual i,
- C_i represents the cost imposed on that individual due to coercion and lack of freedom.

to estimate these costs, they model the difference between forced and free labor compen-

sation:

$$W_f = \alpha W_s \tag{2}$$

where W_f is the wage of free labor, W_s is the subsistence value provided to enslaved individuals, and $\alpha > 1$ represents the relative inefficiency of slavery. They estimate that the overall economic gain can be expressed in terms of national productivity increase:

$$\Delta P = \beta P_0 \tag{3}$$

where:

- ΔP is the increase in national productivity,
- P_0 is the initial productivity under slavery,
- β is a factor estimated between 0.04 and 0.35, representing the increase as a percentage of national productivity.

By eliminating forced labor and allowing free labor markets to operate, the economy saw gains equivalent to decades of technological advancement. The methodology demonstrates how correcting labor misallocation leads to significant aggregate economic improvements.

1.4 Adam Smith on the Limitations of Human Empathy and Rhonda Williams on Racial Earnings Inequality

In this section, I suggest an alternative explanation for the persistent inequities experienced by those with multiple marginalized socially significant identities. Here, the key is to cast aside neoclassical conceptions of somewhat benign, socially advantageous self-interested competition and adopt a view of competition as rivalrous and beneficial to the social status

and well-being of one group at the expense of the other. This treatment is closer to that of classical economists such as Adam Smith, David Ricardo, and Karl Marx.

Since the often cited *The Wealth of Nations*, economists have worked from a view of competition in capitalist systems as the result of a natural tendency to compete, consume, and accumulate channeled into a productive, socially helpful activity. However, Smith himself was weary of the limits of mankind for other regarding behavior (Smith, 1776). In his *The Theory of Moral Sentiments*, Smith discussed in detail how human beings attempt to understand the motivations and needs of others through attempting to consider their perspective and place themselves in their situation (empathy) (Smith, 1759). It should be noted that this imperfect attempt at understanding others in a market economy is complicated by the fact that our own biases and motives influence what we think motivates others. For example, a selfish man may assume that another person will be non-cooperative in a given situation, clouded by his own tendencies. The implication of negative consequences for the efficiency of transactions and coordination problems produced by such prevalent misunderstandings is evident.

Empathy in artificial agents can be modeled using Bayesian inference, where the agent updates its beliefs about another agent's emotional state based on observed actions. This document outlines the mathematical foundation behind such a Bayesian learning process.

The agent maintains a belief distribution over possible emotional states E of another agent. Given an observed action A, the belief update follows Bayes' theorem:

$$P(E|A) = \frac{P(A|E)P(E)}{P(A)} \tag{4}$$

where:

- $P(E|A)$ is the posterior probability of an emotional state E given action A,
- $P(A|E)$ is the likelihood of observing action A given emotional state E,

- $P(E)$ is the prior probability of emotional state E,
- $P(A)$ is the normalizing constant ensuring probabilities sum to one.

To refine its empathy model, the agent iteratively updates its belief distribution over multiple interactions. Given a sequence of observed actions $A_1, A_2, ..., A_n$, the agent applies:

$$P(E|A_1, ..., A_n) \propto P(A_n|E)P(E|A_1, ..., A_{n-1}) \tag{5}$$

This recursive update allows the agent to learn and refine its understanding of emotions over time. Once the agent has inferred an emotional state, it can choose actions that align with empathetic behavior by maximizing an expected utility function:

$$A^* = \arg\max_A \sum_E U(A, E)P(E|A) \tag{6}$$

where $U(A, E)$ represents the utility of taking action A when the emotional state is E. This Bayesian framework provides a principled approach for an agent to learn and adapt its empathetic responses based on observed behaviors. Future work can extend this model to incorporate deeper hierarchical structures and multi-agent interactions.

The similarities between the model dynamics presented by Swinton (see Swinton (1974), Swinton (1977), and Swinton (1978)) and the proposed adoption of a classical notion of competition by Williams (1987) are numerous. One difference is Swinton explicitly saying that while his model assumes classical competition, the emphasis on the white workers role in their own wage undervaluation (the decades long disconnect between productivity and real wages is well documented at this point) is decidedly non-Marxian. However, Williams makes the same claim that white workers benefit from the marginalization of non-whites but does not find this contradictory to Marx's notions of economic relations under capitalism.

Williams argues that for Marx, monopoly is not the antithesis of perfect competition

as is conceptualized in neoclassical theory. In traditional neoclassical conceptions, in an ideal world an infinite number of firms exist to drive prices down to marginal costs. In contrast, Marxian notions of competition are that the logical end result of the motives of capitalist accumulation are monopoly or oligopoly with a few firms in each industry. Market concentration still occurs on a spectrum, and the system lends itself to a tendency for firms to move along the spectrum rather than remain engaged in endless dynamic Bertrand and Cournot competition that lowers their profits. Rather, oligopoly and monopoly are the natural end result of a process of accumulation and consolidation of market power via mergers, acquisitions, and the actual rivalrous strategies of profit maximizing behavior.

I shall quote here in full from the article, as I think the specific language here is most appropriate and prone to misinterpretation if paraphrased or summarized:

"White numerical dominance of high productivity sectors and black concentration in low productivity sectors thus becomes a sufficient condition for the existence of racial earnings inequality. Wage differentials persist between equally able, yet racially distinct, workers because productivity is determined by sectoral social relations and techniques of production. Whereas neoclassical competition presumes labor mobility will equalize wages across sectors, classical competition reconciles long-term market discrimination and competition because it definitionally permits longstanding impediments to intersectoral mobility which are anathema to neoclassical competition."

2 Literature Review

2.1 Kitagawa-Oaxaca-Blinder Decomposition

The mean decomposition technique popularized in economics by Oaxaca (1973) and Blinder (1973) originated in the sociology literature with Evelyn Kitagawa's 1955 article

Components of a Difference Between Two Rates (I will refer to the method as the KOB or Kitagawa-Oaxaca-Blinder decomposition). In the article, Kitagawa explains that the technique is meant to "explain the difference between the two rates of two groups in terms of differences in their specific rates and differences in their composition" (Kitagawa, 1955). Kitagawa begins by detailing the method with two factors that may be quantitative or non-quantitative, and then extends the method to deal with three or more factors (Kitagawa, 1955).

Over time numerous studies attributed much of the unexplained racial wage differential to differences in Armed Forces Qualifying Test scores (AFQT), an imprecise proxy for cognitive ability. Despite Rodgers and Spriggs (1996) finding that AFQT composite scores produced biased estimates of discrimination in the labor market, their study has been widely overlooked. Using NLSY91, an AFQT test score model, and KOB decompositions, Rodgers III and Spriggs (1996) found that psychological variables capture any racial differences in self-confidence from less supportive background environments and psychological effects of pre-labor market discrimination. A two-step measure reduces the mean square error of the wage equation and provides an unbiased estimate of labor market discrimination (Rodgers and Spriggs, 1996). As an example, a study by Blackburn (2004) still uses the AFQT and finds that while AFQT scores account for some of the racial differential, they only account for a small part of the gender differences in wages.

My study uses personality scale variables measuring attitudes toward hard work to determine the combined effects of an orientation toward hard work and the selectivity of one's undergraduate institution on racial wage differentials across different majors. William Darity Jr., Arthur Goldsmith, and Jonathan Veum (1997, 1998) are two examples of studies I have found besides Rodgers III and Spriggs that use psychological capital measures in the study of racial wage differentials. Their studies examine one's self-esteem and locus of control and found that even when including AFQT composite scores and the higher scores on

self-esteem exhibited by black workers in their study, the effect of being black still had a significant negative effect on real wages (see Goldsmith et al. (1997) and Goldsmith et al. (1998)).

Mason (2007) also uses psychological characteristics such as motivation to study interracial differences in economic outcomes between white and Black people in the U.S. I recognize that the self-reported personality scales in the NLSY97 may be biased because respondents may wish to appear as if they have more positive attitudes towards hard work than they have. In our society, laziness is frowned upon regardless of racial group. However, these measures of laziness are the best proxy for attitudes towards hard work available with the NLSY97 data.

For my study, I decided to combine Black and Latino workers into one group, called the marginalized group, and white workers into another group due to the relatively small sample sizes in the NLSY97. This decision does have some theoretical grounding in the literature, as there is evidence that Black and Latino workers experience wage discrimination that differs by skin shade. Goldsmith et al. (2006) found that by the early 1990s, skin shade had an important effect on wages with darker skinned black men losing at least 10 percent in wages relative to white men, whereas the effects were indistinguishable from a slightly negative value near zero for light skin complexioned black men. A more recent study by Kreisman and Rangel (2015) found that skin color plays a significant role in earnings and employment for black men. Wage gaps between darker-skinned and lighter-skinned workers were shown to widen with experience, with gaps between light-skinned workers and white workers remaining constant (Kreisman and Rangel, 2015).

Although gender differentials are not examined in the dissertation for the sake of time and brevity, there is also troubling evidence that the negative effects of unemployment on mental health increase for black women in the US as skin shade darkens (Diette et al., 2015). These skin shade wage penalties have also been found for Latin American immigrants using the

New Immigrant Survey (Rosenblum et al., 2015). There is also evidence that darker skinned Brazilian women experience significant wage discrimination in the Brazilian labor market, a sign that these skin shade penalties are transcontinental (Fernandes, 2015). These results reinforce arguments that the further one is from a white phenotype, the worse the wage penalty. Unfortunately, our NLSY97 data do not include a skin shade variable to discern whether the effects of attitudes toward hard work and selectivity of one's undergraduate institution on racial wage penalties across different majors differ by skin shade.

2.2 Monitoring Decisions and Disproportionate Punishment of Black Workers

African Americans experience shorter employment duration's in the labor force than Whites with similar productive characteristics. Cavounidis and co-authors (Cavounidis et al., 2023) advanced a theory and empirical evidence that employers discriminate in acquiring or acting on ability-relevant information. Essentially, Black workers are monitored more closely than white workers and this their errors are more likely to be caught and penalized (either placed on probationary systems or eventually fired). This reduces firms' beliefs about ability, and incentivizes ever more discriminatory monitoring. Their empirical evidence confirm that layoffs are initially higher for Black than non-Black workers, but that they converge with seniority and decline more with the Armed Forces Qualification Test as a proxy for unobserved 'aptitude' of Black workers (Cavounidis et al., 2023).

2.3 Two Sided Labor Market Statistical Discrimination

For those unconvinced by the economic theory presented above to explain the persistence and magnitude of racial wage differentials between marginalized and white college graduates with other similar observable productive traits, there is a compelling explanation of two-

sided statistical discrimination. Craig and Fryer (Craig and Fryer, 2018). They contribute a model that makes clear consequences associated with a labor market coordination problem in which employers use stereotypes about a group to guess their productivity and workers try to guess whether firms will have an environment conducive to their success (no workplace discrimination, opportunity for advancement, on-the-job training, and a truly welcoming company culture).

They propose an insurance system in which the government hires workers whom they believe to be competent but are not offered employment. Curiously, this insurance system is similar in theory to the universal job guarantee proposals by some non-neoclassical economists. The universal job guarantee is not dependent on the government determining workers are being discriminated against, hence the 'universal' distinction in comparison to the proposed insurance scheme which is based on some government imposed criteria. Think the difference between means tested social safety net programs and universal programs. Both policies effectively place a floor on adverse consequences for workers experiencing undervaluation of their marginal revenue product of labor (productivity and contribution of value) and higher unexplained spells of unemployment for Black workers, though through different means and likely with different welfare consequences.

2.4 Alternatives to the Neoclassical Explanation Matching Evidence with Realism

There is a long history of social scientists developing alternative approaches to the neoclassical explanations of racial gaps that have been presented by those operating from outside and within this paradigm (Francis et al., 2022). A clear example of this is work by Rhonda Williams (Williams, 1987) to make these points and emphasize the need for more realistic theory. This is further emphasized or clarified by contributions to economic theory by

the Stratification Economics school (Darity Jr., 2001)(Chelwa et al., 2022) and ideas about racial conflict in technologically advanced societies put forward by Nobel Laureate and St. Lucian economist W. Arthur Lewis (Lewis, 1985). With regard to the importance of social networks, location, investment in education, and social standing for how these gaps evolve and persist, see (Loury, 1976) and (Loury, 1981a), (Andrews, 1999b) and more recent work by , (Andrews et al., 2017), , (Chetty et al., 2019), (Okafor, 2022), (Brundage and Tavani, 2024),(Pereira Serra, 2024), among many others (ex. see (Alexis, 1973),(Handy, 1998),(Malveaux, 1980), (Jaynes, 1982), (Datcher-Loury et al., 1980), (Mason, 1997)(Darity Jr. and Myers Jr., 1999), (Marglin, 2021).

2.5 Racialized Ability Tracking and Isolation

There is convincing work by Darity and Francis on the legacy of ability grouping in K-12 education and a proposed theoretical explanation as an alternative to the 'acting white' hypothesis that Black people in the United States adopt dysfunctional behaviors in opposition to forced conformation to white cultural norms for success (Darity Jr and Francis, 2020). They contend that what is actually being observed is an anti-intellectual effect that is present for high-achieving students regardless of race, and that this is intensified by the legacy of ability tracking failing to identify Black or marginalized students for talented and gifted programs (Darity Jr and Francis, 2021). This has clear implications for the majors chosen and career paths of college graduates.

2.6 Returns to Major and Educational Attainment

Choice of a college major has been shown to affect earnings premiums from college and predictably affects the racial wage disparities experienced in the labor market. Barrow and Rouse (2005) found little evidence of differences in return to schooling across racial and ethnic

groups when controlling for ability and measurement error biases. They concluded that policies increasing educational attainment for marginalized groups have a good possibility of increasing their well-being and reducing inequality. Urzua (2008) supports the conclusion that even after controlling for respondents' abilities, there are significant racial labor market gaps and the results suggest that the standard practice of equating observed test scores may overcompensate for differences in ability. Dickson (2010) found significant gender and racial differences in college major choice persist even when controlling for SAT scores and high school class ranks, where gender differences were even greater than racial and ethnic disparities. Nonwhite women were more likely than men to switch away from an initial major in engineering. This result was not true for white women (Dickson, 2010).

Webber (2014) presents a recent, high-quality examination of lifetime college earnings premiums and finds that there is a significant range in lifetime college earnings premiums, from $\$700,000$ for arts and humanities graduates to \$1.5 million for STEM graduates. These differentials are even larger when allowing for differential unemployment probabilities across majors (Webber, 2014). The Caroline Hoxby and Sarah Turner (2015) study on an intervention to get low-income, high-achieving college students to apply to more selective colleges shows that providing these students with information on cost of college, availability of curricula and peers they seek, and the different types of colleges available to them does increase their application and acceptance to selective institutions. This result has potentially large implications for high-achieving students from marginalized groups who are disproportionately more likely to come from low-income households (Hoxby and Turner, 2015). McClough and Benedict (2017) found that academic majors and first education institutions have significant effects on one's final occupation, which influences racial salary disparities. They recommend that public policy should build awareness around high school students from marginalized groups aspiring for college to prepare for majors with earnings premiums (McClough and Benedict, 2017).

2.7 College Selectivity and Earnings

We can establish some hypotheses on how the inclusion of the college selectivity variable in the wage equation may affect marginalized and white groups differently based on a small number of studies examining college selectivity and earnings. Alsalam et al. (1989) found one should choose a selective private eastern college to maximize earnings. Loury and Garman (1995) found that previous studies overstate the earnings premiums of college selectivity for whites and understate it for blacks by omitting measures of college performance like grade point average. They also note that the larger black earnings gain is offset by students whose own SAT scores are below the median of the college they attend (Loury and Garman, 1995).

Dickerson and Jacobs (2006) found that the system of higher education in the United States is segregated, with white graduates much more likely to attend selective schools. They conclude that Blacks under representation in STEM fields that characterize highly selective schools may explain some of the racial disparity in selectivity of schools attended (Dickerson and Jacobs, 2006). Dale and Krueger (2011) found that the estimates of returns to selectivity are essentially zero when controlling for the average SAT score of the schools to which students have applied. The return to college selectivity remains large for Black and Latino students from less educated families (Dale and Krueger, 2011). Another potentially significant finding by Hoxby and Avery (2013) is that more affluent high achieving students apply to several peer schools where their achievement match the typical applicant, a few reach schools where their achievements are below that of the typical applicant, and a couple safety colleges where their achievements are above the typical applicant. Low-income-typical students typically do not apply to selective colleges (Hoxby and Avery, 2013).

2.8 Underemployment and Educational Attainment

Georgetown University's Center on Education and the Workforce released a 2019 report by Campbell et al. (2019) entitled *The Unequal Race for Good Jobs: How Whites Made Outsized Gains in Education and Good Jobs Compared to Blacks and Latinos.* The study details the inequities between white and Black and Latino workers in access to what they categorize as 'good jobs' (those that pay at least $\$35,000$ per year, at least $\$45,000$ for workers 45 and older and $\$65,000$ in median earnings in 2016 using data from the U.S). Their data is from the U.S. Census Bureau and the U.S. Bureau of Labor Statistics, Current Population Survey 1992-2017. At every level of education that they examined (high school diploma or less, middle skills, and college degree or higher), Black and Latino workers 2016 median earnings were lower in good jobs (Campbell et al., 2019).

Of interest to my study is the finding that for workers with a bachelors degree or higher, the median 2016 earnings for white workers was $\$75,000$ whereas the 2016 median earnings for Black and Latino workers was $\$65,000$. Campbell et al. (2019) calculated that the total annual earnings advantage of white workers from good jobs was $554 billion in 2016, compared with a situation in which jobs were equitably distributed. These findings indicate that we would expect to find evidence of wage penalties for marginalized workers and wage benefits for white workers using our NLSY97 data on college graduates, even after controlling for the typical and relevant variables in our wage equations. Another study by Williams and Wilson (2019) found that nearly 40% of Black workers with a college or advanced degree are in a job that typically does not require a college degree, compared to 31% of white graduates. These rather stark differences in underemployment and access to good jobs most likely contribute to any significant racial wage differentials I find among the college graduates.

2.9 Contribution to Literature

This is the first study to quantify the contribution of different returns to positive attitudes towards hard work and selective collegiate institutions to racial wage differentials. No other study I have found has studied the effects of an interaction between attitudes towards hard work and college selectivity on wages using the NLSY97. I have also not found studies that match NLSY97 youth to their colleges, then calculate a college selectivity variable using the codes for schools from which they graduated. This study provides a useful update by using the newer NLSY97 cohort rather than the NLSY79 cohort to study the racial wage differentials among college graduates.

This study is also novel in applying the intersectional identity wage differential framework of Darity Jr. et al. (2022) using the Cotton (1988) method of weighting the coefficients of the demographic groups by the group proportions to establish the wage that would prevail without discrimination. There are numerous ways to calculate this non discriminatory wage, and results from another approach by Neumark (1988) using the coefficients from a pooled regression are also presented.

The use of the least absolute shrinkage and selection operator (lasso), a machine learning method to minimize error in the wage equations is also a novel contribution. After detailing the empirical strategy, I propose an extension of these methods to panel data to study racial earnings inequality using causal inference. First I detail the interventionist difference-in-differences approach with Kitagawa-Oaxaca-Blinder decomposition and the Cotton method, then I explain how synthetic control difference-in-differences quasi experimental methods can be used to study these wage differentials. I use some work from behavioral economics literature to explain how improved perceived ability and updating of beliefs modelled using Bayes theorem help us to understand racial earnings inequality, and discuss some problems associated with longitudinal wage equations and two-way causality.

3 Theoretical Framework

3.1 Microfoundations

3.1.1 Utilization of Deviation from Phenotypical Whiteness, Masculinity, and Heterosexuality as Tools for Alienation

One of the main contributions of this paper is to propose a clear transition away from more transactional explanations of the dynamics of systemically unequal economies to one in which domineering and rivalrous behavior is assumed. The concept of alienation is useful here, as the depiction of non-white workers as an antagonistic "other" serves a useful function when attempting to maintain dominion over a materially deprived population experiencing unnecessary scarcity and subjective uncertainty. Deviations from phenotypical "whiteness", defined as a skin shade further along a gradient and defined as fair or light complexioned, traditional ideas surrounding gender roles and heterosexuality can be utilized to discourage coalitions between working-class workers. This coincides with the idea of whiteness as property advanced by scholars of racial inequality.

3.1.2 Marginal Revenue Product of Labor and Measurement of Exploitation in a Neoclassical and Analytical Marxian Sense

In neoclassical labor economics, exploitation can be mathematically understood as the difference between the marginal revenue product of labor (MRPL) and the wage paid to workers. This section explains the mathematical framework that describes this phenomenon. The MRPL is defined as the additional revenue generated by employing one more unit of labor. Formally,

$$MRPL = \frac{\partial TR}{\partial L} \tag{7}$$

where:

- TR is the total revenue of the firm,
- L is the quantity of labor employed.

Under conditions of perfect competition in the output market, the firm takes price P as given, and the MRPL can be rewritten as:

$$MRPL = P \cdot MPL \tag{8}$$

where:

- P is the price of the output,
- $MPL = \frac{\partial Q}{\partial L}$ is the marginal product of labor.

Exploitation arises when workers are paid a wage (w) that is lower than their marginal revenue product:

$$MRPL - w > 0. \tag{9}$$

This can be rewritten as:

$$E = MRPL - w = P \cdot MPL - w, \tag{10}$$

where E represents the extent of exploitation. The wage being lower than MRPL can occur due to several factors:

- **Monopsony Power**: Firms with market power in hiring set wages below competitive levels.

- **Bargaining Power Asymmetry**: Workers lack negotiating power due to unemployment, lack of union representation, or other constraints.
- **Institutional Factors**: Government policies, minimum wage laws, and labor regulations impact wage levels.

In a neoclassical framework, exploitation can be quantified as the difference between MRPL and the wage. This difference arises due to market imperfections and power dynamics that prevent wages from equaling the value of workers' contributions to revenue. Roemer formalizes exploitation in terms of endowments and class positions rather than purely wage relations. An agent is considered exploited if their labor contribution exceeds their proportional share of total output in an economy.

Consider an economy with N agents indexed by $i = 1, \ldots, N$. Each agent possesses an endowment of capital K_i and labor power L_i. The total capital and labor in the economy are:

$$K = \sum_{i=1}^{N} K_i, \quad L = \sum_{i=1}^{N} L_i. \tag{11}$$

The production function of the economy is given by:

$$Q = F(K, L), \tag{12}$$

where Q is the total output.

Roemer defines an agent as exploited if they would be better off in an alternative economic arrangement known as the *independent producer economy* (IPE), where all individuals have equal access to capital and production proceeds are distributed based on labor contribution alone.

Denoting the income of agent i in the actual economy as Y_i and their hypothetical income

in the IPE as Y_i^*, the exploitation condition is:

$$Y_i < Y_i^*. \tag{13}$$

This implies that the agent receives less than they would under an equal redistribution of capital, signifying their exploitation. Roemer identifies class positions based on ownership of productive assets:

- **Capitalist class**: Owns capital and hires labor.
- **Proletariat**: Owns only labor and sells it to capitalists.
- **Middle class**: Owns both capital and labor.

Exploitation arises when a class's access to productive assets leads to income differentials that systematically favor one group over another. A key implication of Roemer's framework is that exploitation can exist even in the absence of wage labor, as it is fundamentally tied to inequalities in resource ownership rather than labor relations alone. This is captured in the class exploitation correspondence principle:

$$E_i = \max(0, Y_i^* - Y_i). \tag{14}$$

Roemer's general theory of exploitation extends beyond Marxist labor exploitation by incorporating asset ownership and alternative economic arrangements. This mathematical framework provides a robust basis for analyzing class structures and economic inequalities in capitalist societies.

3.1.3 Bayesian Belief Formation and Skill Acquisition Among Working Class and Non-White Workers

Filippin and Paccagnella (2011) model self-confidence as a Bayesian updating process in

which individuals adjust their beliefs about their own abilities based on past experiences. This document extends their framework by incorporating a diminishing returns function to effort, where effort is influenced by nutrition and physical health. Let θ represent an individual's true ability, and let b_t denote their belief about θ at time t. The belief updating process follows:

$$b_{t+1} = \frac{P(X_t|\theta)b_t}{P(X_t)} \tag{15}$$

where X_t represents observed performance outcomes. We introduce effort e as a function of nutrition (N) and physical health (H):

$$e = f(N, H) \tag{16}$$

To account for diminishing returns, we model effort as:

$$f(N, H) = \alpha \log(1 + \beta N + \gamma H), \quad \alpha, \beta, \gamma > 0. \tag{17}$$

This functional form ensures that effort increases with nutrition and health but at a decreasing rate. Effort influences performance outcomes X_t, modifying the likelihood function:

$$P(X_t|\theta, e) = g(\theta, e) + \epsilon_t, \tag{18}$$

where $g(\theta, e)$ is an increasing function of both ability and effort, and ϵ_t represents noise. The belief updating equation is then adjusted to:

$$b_{t+1} = \frac{P(X_t|\theta, e)b_t}{P(X_t|e)}. \tag{19}$$

This extension incorporates effort into the Bayesian belief formation model and accounts

for diminishing returns driven by nutrition and health. Future work could explore dynamic feedback loops where belief influences effort allocation. Imagine students from marginalized groups with ability θ that is unknown to them at the beginning of their schooling and professional careers. Suppose that our student has to decide an academic track in high school, an academic major in college, and an occupation post-college graduate according to their own perceived ability, a function of beliefs about their ability represented by a density function.

An upper middle class white male child may overestimate ability and an under-confident woman from a working class background underestimates their ability The non-white girl then woman choosing classes in high school, college major, or occupation post-undergraduate studies selects academic preparation, path, or career ψ. Some paths are understandably more difficult with regard to the rigor of the course and the complexity of ideas and effort needed to succeed. The probability of success with preparation or occupation is a function of the ability and difficulty of the path or occupation. This allows a scenario in which one may not naturally process information as quickly or efficiently as another, but can achieve extraordinary outcomes by applying themselves. Essentially, intelligence is liquid and improvable. So we have probability of success is a function of ability, difficulty of path, and effort.

This functional form ensures that effort increases with nutrition and health but at a decreasing rate.The agent's utility from human capital acquisition depends on effort and perceived ability. Let $U(\Psi)$ represent utility from human capital accumulation:

$$U(\Psi) = \delta \log(1 + \lambda e + \mu b_t) - \frac{ce^2}{2}, \tag{20}$$

where:

- $\delta > 0$ represents the baseline benefit from human capital acquisition,

- $\lambda > 0$ captures the impact of effort on human capital,
- $\mu > 0$ represents the role of self-belief in learning,
- $c > 0$ accounts for the cost of effort, introducing diminishing returns.

Effort influences performance outcomes X_t, modifying the likelihood function:

$$P(X_t|\theta, e) = g(\theta, e) + \epsilon_t, \tag{21}$$

where $g(\theta, e)$ is an increasing function of both ability and effort, and ϵ_t represents noise. The belief updating equation is then adjusted to:

$$b_{t+1} = \frac{P(X_t|\theta, e)b_t}{P(X_t|e)}. \tag{22}$$

This extension incorporates effort into the Bayesian belief formation model and accounts for diminishing returns driven by nutrition and health. The utility function formalizes the acquisition of productivity-enhancing skills and knowledge and highlights the trade-off between effort and learning efficiency. Future work could explore dynamic feedback loops where belief influences effort allocation.

3.1.4 A Swinton Model with Classical Competition, Nash and Kantian Optimizers, and Coalitions

Following Swinton's framework, labor market discrimination affects the updating process of self-beliefs. If employer expectations introduce biased signals, the belief updating equation modifies to:

$$b_{t+1} = \frac{P(X_t|\theta, D)b_t}{P(X_t|D)}, \tag{23}$$

where D represents discriminatory constraints such that $P(X_t|\theta, D) < P(X_t|\theta)$ for discriminated groups. Firms can use intersectional identities to segment workers and prevent solidarity. Let g_i represent a worker's group identity, such as race, gender, or immigration status. Firms strategically set wages w_i to create divisions:

$$w_i = w^* - \phi(g_i), \tag{24}$$

where w^* is the competitive wage and $\phi(g_i)$ is a discriminatory wage penalty. Firms maximize profits:

$$\max_{w_i} \sum_i \left[pf(e_i) - w_i \right], \tag{25}$$

subject to labor supply constraints. By setting $\phi(g_i)$ such that different groups receive unequal pay, firms prevent coalitional bargaining and sustain wage suppression. Roemer's work on Nash and Kantian optimization introduces cooperative dynamics where firms and households optimize labor decisions collectively. Let w_i denote the wage of worker i and p the firm's price function. Under Nash optimization:

$$\max_{w} \sum_i U_i(w_i) - C(p, w), \tag{26}$$

where U_i is the worker's utility and C is firm cost. Under Kantian optimization, firms and workers coordinate so that each agent optimizes collectively:

$$\max_{w} U_i(w) \quad \text{s.t.} \quad w = \bar{w}, \tag{27}$$

where $\bar{w}$ is the wage determined by collective fairness criteria. This model integrates Bayesian learning, effort functions, labor market discrimination, and cooperative game-theoretic approaches to decision-making. The introduction of intersectional wage discrim-

ination highlights conditions for profitable exploitation and the strategic fragmentation of worker solidarity. Future work may explore empirical applications in labor market dynamics and policy interventions.

3.2 Lexiographic Preferences in a Von Neumann Model with Class and Race Dynamics

The Von Neumann growth model is a mathematical formulation of economic growth based on input-output relationships between productive sectors. It assumes an economy with n goods and m production processes, where production is represented by:

$$X = AX + BY, \tag{28}$$

where:

- X is the total output vector,
- A is the input matrix (showing how goods are used in production),
- B is the output matrix (showing how goods are produced),
- Y is the vector of production levels for each process.

Von Neumann proved that under certain conditions, there exists a balanced growth rate g such that:

$$(B - (1 + g)A)Y = 0. \tag{29}$$

This implies that production expands at a constant rate, determined by technological and resource constraints. The innovation of this model in economics was its introduction of

game-theoretic concepts to macroeconomic growth, its focus on intersectoral linkages, and its rigorous treatment of balanced growth conditions.

Let:

- n be the number of commodities,
- m be the number of production processes,
- $A = [a_{ij}]$ be an $n \times m$ input matrix where a_{ij} is the amount of commodity i consumed per unit of activity j,
- $B = [b_{ij}]$ be an $n \times m$ output matrix where b_{ij} is the amount of commodity i produced per unit of activity j,
- $x = (x_1, x_2, \ldots, x_m)$ be the activity level vector,
- $p = (p_1, p_2, \ldots, p_n)$ be the price vector,
- g be the uniform growth rate,
- λ be the discount factor (where $0 < \lambda < 1$).

3.2.1 Equilibrium Conditions

The model is based on the following fundamental constraints:

Production Feasibility (Material Balance Condition)

$$Bx \geq (1 + g)Ax \tag{30}$$

This ensures that total output meets or exceeds total input.

Profitability Condition (Nonnegative Prices)

$$pB \leq \lambda pA \tag{31}$$

This ensures that prices remain non-negative and maintain equilibrium.

3.2.2 Balanced Growth Rate

To ensure economic balance, we seek the nontrivial growth rate g such that:

$$Bx = (1 + g)Ax \tag{32}$$

and the corresponding price equation:

$$pB = \lambda pA \tag{33}$$

where g represents the maximum possible uniform growth rate and λ represents the corresponding competitive rate of return.

3.2.3 Optimization Formulation

Von Neumann formulated the equilibrium as a minimax optimization problem:

$$\max_{x} \min_{p} \frac{pBx}{pAx} \tag{34}$$

This guarantees the existence of a unique solution $(g^*, \lambda^*, x^*, p^*)$ under reasonable assumptions about production and consumption constraints.

We extend the Von Neumann growth model by incorporating Roemer's theory of exploitation with class structures based on asset ownership and labor provision. Additionally,

we introduce two races with lexicographic preferences.

Let the economy have n goods and m production processes. The production system satisfies:

$$X = AX + BY, \tag{35}$$

where asset-owners extract surplus from workers, and racial identities influence preference structures. The exploitation condition is modeled as:

$$\pi = BY - wL, \tag{36}$$

where:

- π represents surplus extraction,
- w represents wages,
- L represents labor input.

Lexicographic preferences imply:

$$U_R > U_W \Rightarrow \exists \lambda : U_R = f(X) + \lambda g(Y), \tag{37}$$

where racial dynamics and discrimination create segmented labor markets. Firms optimize:

$$\max_Y (B - (1+g)A)Y - wL, \tag{38}$$

under class and racial constraints, producing differential growth paths for workers based on asset ownership and racial identity.

3.3 A Model from the Stratification Economics School

This paper stems from a doctoral thesis written between August 2018 and December 2022. Since then, an elegant contribution to formalize and explain these dynamics in a two-period model has emerged. A recent working paper by Brundage and Tavani (Brundage and Tavani, 2024) situates the racial economic conflict of the technologically advanced U.S. economy and persistent inequality in a model inspired by W. Arthur Lewis and contributions from The Stratification Economics School. They present a model of group conflict, racial conflict, and stratification. They assume two groups, one dominant and the other marginalized. Both groups live for two periods, a pre-market period in which they investment in schooling and a market-phase in which they utilize their investments. For the sake of brevity we will not present their full, rather sophisticated model in its entirety here and will describe it with written language.

An individual in the marginalized group chooses how to invest in the pre-market phase to maximize their net material resources. Increased discriminatory efforts by the dominant group reduce the effectiveness of the pre-market investments of the marginalized group. A key feature of this theoretical framework is that as the productivity of the marginalized group increases, the dominant group must increase their discriminatory effort to reduce the material resources obtained by the marginalized group. This is a theoretical explanation of the historically witnessed "backlash" to development of "middle class" members of marginalized groups in the United States (see the history of Reconstruction and increased political participation, as well as hate group resurgence in recent years as progress was made toward racial harmony) (Brundage and Tavani, 2024).

4 Data

I started with the entire NLSY97 dataset of 8,984 individuals. I dropped cases with missing data on 2019 wage and salary income (3,572 deleted), the attitude toward hard work personality measures (403 deleted), whether or not their health limited the amount of work they could perform (3 deleted), the occupation type of their main job (157 deleted), and their marital status for 2019 (17 deleted). Next, I also dropped cases missing data on whether they live in a metropolitan statistical area (49 deleted), their highest degree obtained (16 deleted), ASVAB score percentile (818 deleted) and their household structure in 1997 (16 deleted). Next, I matched respondents to the undergraduate institution where they earned their bachelors degree (2,882 deleted). This left me with a sample of 1,051 college graduates. Finally, I created a college selectivity variable that is the total SAT score of the 75th percentile in 2016 (or the ACT score of the 75th percentile converted to an SAT score) for undergraduate colleges using data from the Integrated Postsecondary Education Data System (IPEDS). I was left with a sample of 860 college graduates since there was no ACT or SAT data for the schools of all 1,051 graduates.

The variable names and definitions for the variables used in the analyses are listed in Table 1. In the subsections following this one I present the descriptive statistics for different demographic groups analyzed in this paper. I start with the pooled sample, then describe descriptive statistics for the white college graduates and marginalized college graduates. Next I present statistics for non-white women, white women, non-white men, and white men. The National Longitudinal Survey of Youth 1997 cohort consists of 8,984 men and women born between 1980 and 1984. Respondents were interviewed annually between 1997 and 2011 then biennially since 2011. I use data from round 1 (1997-1998), round 3 (1999-2000), round 12 (2008-2009), and round 19 (2019-2020).

4.1 Pooled Sample

Of the 854 people in the sample, 54.3% are female (n=467) and 45.7% are male (n=393). With regard to racial composition, 70.58% are white (n=607) and 29.42% (n=253) are members of the non-white racial group consisting of 142 black college graduates, 102 Latino college graduates, and 9 mixed-race college graduates. With regard to major, 7% of the sample graduated with a degree in business, marketing, or management (n=63), the only major found to have a significant effect on income from wages and salary. 64.88% (n=558) of the sample was married at least one month in 2019. 2.21% (n=19) of respondents reported that their health limited the amount that they could work in 2019, and 36.74% (n=316) of the sample has attained a graduate degree.

The mean age of the pooled sample is 36.90 years and the mean ASVAB percentile score is 70.40. Out of a possible 28 points total, the mean attitude toward hard personality scale score is 24.55. The mean SAT score of the 75th percentile score for the colleges attended is 1,226.70 points. The mean number of years of experience for the pooled sample is 16.84 years. The mean income from salaries and jobs in 2019 is $83,484.07. The mean log of income from wages and salaries in 2019 is 11.07.

4.2 Non-White Graduates

With the non-white group, 59.29% (n = 150) are women, 6.32% (n = 16) graduated with a degree in business, marketing or management, 50.59% (n=128) were married in 2019, 2.37% (n=6) reported that health problems limit the amount they can work and 37.15% (n = 94) have attained a graduate degree. They had a mean income from wages and salaries of $75,249.02 in 2019.

The mean age for the non-white graduates is 36.92 years. The mean ASVAB percentile score is 58.01 and the mean attitude toward hard work personality scale score is 24.84 out

Group	Mean	St.Dev.	N
Non-White Men	11.128	.617	103
Non-White Women	10.935	.666	150
White Men	11.358	.659	288
White Women	10.848	.934	313

Table 1: Log of wage and salary income by group

of 28 total possible points. The mean SAT score for the 75th percentile for the colleges attended is 1,188.84, significantly lower than that of the white graduates which aligns with past research on the distribution of white and marginalized graduates at selective colleges. The mean number of years of experience working is 16.20 years, nearly one year less than the mean years of experience for the white graduates. The mean log of income from wages and salaries in 2019 is 11.01, signifying that the mean income of the non-white graduates is 8 percent lower than the mean income of the white graduates.

4.3 White Graduates

52.22% (n = 317) of the white graduates in the sample are female, 7.74% (n = 47) graduated with degrees in business, marketing, or management, 70.84% (n = 430) were married in 2019, 2.14% (n = 13) have health problems that limit the amount they can work and 36.57% (n = 222) have obtained a graduate degree.

Whites had a mean income from wages and salaries of $87,784.19. The mean ASVAB percentile score is 75.51. The mean number of years of experience working is 17.11 and at the mean their colleges had a 75th percentile SAT score of 1,242.08. Their mean score on the attitude toward hard work personality scale is 24.44 out of 28 points. The mean log of wage income from wages and salaries in 2019 is 11.09. The mean age of the white graduates is 36.89 years.

4.4 Baccalaureate and Beyond Data from NCES

I am also using the Baccalaureate & Beyond (B&B) Survey from the National Center for Education Statistics to illustrate the method. There is data on the typical wage equation demographic characteristics : marital status, age, years of experience, degrees attained, disability status, sexual orientation, workplace experiences, number of dependents, etc. The survey follows graduates 1, 4, and 10 years post graduation. There is also data on personal assets and debt, housing insecurity, parent/family finances and household structure. There is data on graduates SAT and/or ACT composite scores as well as high school and undergraduate academic record/achievement (GPA and coursework).

Since we are talking about measures of observed achievement and not unobserved ability, we can quickly dismiss the need to mention problematic proxies for cognitive ability such as the AFQT (see (Rodgers and Spriggs, 1996)). The usage of SAT or ACT composite (can use conversion if respondents are missing one or the other) also lends support for the rigor of the pre-post difference-in-differences design and allows one to make a reasonable assumption that we are comparing similar students, not some unobservable difference between the PWI and HBCU graduates.

There is data on the selectivity of colleges attended, as well as reported actual hours per week spend outside the classroom on coursework which can be used as a measure of effort for the 2008-2018 cohort. There are 10 year longitudinal datasets available for the 1993-2003 cohort and 2008-2018 cohort so that we may compare results. Ideally, we want the largest dataset with a representative sample for maximum statistical power. The Opportunity Insights Data with information from the IRS is mentioned in (Price and Viceisza, 2023) and is missing some demographic information as well as the information on scholastic achievement, record, and effort. Another clear, preferred choice is data on 1,000,000 test takers from College Board and various other sources used by Edwards et al. (Edwards et al., 2023).

5 Methods/Empirical Strategy

5.1 Machine Learning Variable Selection with LASSO

Variable selection is a crucial aspect of machine learning, improving model interpretability and reducing overfitting. The Least Absolute Shrinkage and Selection Operator (LASSO) is a regression method that performs both variable selection and regularization, leading to sparse models.

5.2 Mathematical Formulation

Consider a linear regression model:

$$Y = X\beta + \epsilon, \tag{39}$$

where:

- $Y \in \mathbb{R}^n$ is the response vector,
- $X \in \mathbb{R}^{n \times p}$ is the matrix of predictors,
- $\beta \in \mathbb{R}^p$ is the vector of regression coefficients,
- $\epsilon \in \mathbb{R}^n$ is the vector of errors.

The LASSO estimator solves the optimization problem:

$$\hat{\beta} = \arg\min_{\beta} \left\{ \|Y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\}, \tag{40}$$

where $\lambda \geq 0$ is a regularization parameter controlling sparsity.

5.3 Geometric Interpretation

The L_1 norm penalty encourages sparsity by shrinking some coefficients to exactly zero. Unlike Ridge regression, which imposes an L_2 penalty, LASSO constrains the solution within a diamond-shaped feasible region, leading to feature elimination.

5.4 Solution via Coordinate Descent

LASSO can be efficiently solved using coordinate descent. For each coefficient β_j, the update rule is:

$$\hat{\beta}_j \leftarrow S\left(\frac{X_j^T(Y - X_{-j}\beta_{-j})}{\|X_j\|_2^2}, \lambda\right), \tag{41}$$

where $S(z, \lambda)$ is the soft-thresholding operator:

$$S(z, \lambda) = \text{sign}(z)\max(0, |z| - \lambda). \tag{42}$$

LASSO is a powerful technique for variable selection in high-dimensional settings. By enforcing sparsity, it enhances interpretability and reduces overfitting, making it a cornerstone of modern machine learning.

5.5 Non-Convexity and Regularization

Non-convexity in regression models can lead to multiple local minima, making optimization challenging. Regularization techniques help address this issue by imposing penalties on model complexity. Regularization helps smooth the optimization landscape, reducing sensitivity to local minima and improving generalization in high-dimensional settings. Consider a regression problem:

$$\min_{\beta} \sum_{i=1}^{n} (y_i - X_i\beta)^2 + \lambda R(\beta), \tag{43}$$

where $R(\beta)$ is a regularization function and $\lambda > 0$ is a tuning parameter. Common choices for $R(\beta)$ include:

- Lasso (L_1 regularization): $R(\beta) = \sum_j |\beta_j|$, which induces sparsity,
- Ridge (L_2 regularization): $R(\beta) = \sum_j \beta_j^2$, which shrinks coefficients,
- Elastic Net: A combination of L_1 and L_2 penalties.

The NLSY97 collects data on four personality scale variables related to attitudes toward hard work. They measure the extent to which respondents believe they work as hard as the majority of people around them, do the bare minimum work and nothing more, have high standards for their work and strive toward them, and make every effort to do more than is expected of them. I used these variables to study how they affect the Oaxaca-Blinder decompositions examined in my study.

I began by selecting variables from the NLSY97 that are typically used in estimating wage equations and studying racial gaps in wages. The variables selected were marital status in 2019, gender, age, experience in years, experience in years squared, experience in years to the fourth power (see Thomas Lemieux 2006), residence in the southern United States, ASVAB score percentile as an imperfect measure of "ability" or "intelligence", attainment of a graduate degree, a major in business,marketing, or management, a major in liberal arts, general sciences, or humanities, and whether the respondent's health limits the amount of work they can perform. To decide which variables to include in our model, I used the least absolute shrinkage and selector operator (lasso).

With a dependent variable y_i and independent variables x_{ij} where i=1,2,3,...,N and j=1,2,3,...,p, the lasso finds $\beta = \{\beta_j\}$ to minimize

$$\sum_{i=1}^{N}(y_i - \sum_j x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{44}$$

where we're minimizing the sum of squares with a constraint of the form $\sum |\beta_j| \leq s$ (Tibshirani, 2011). The λ controls the strength of the penalty and amount of shrinkage, so that as λ goes to infinity all the coefficients would be eliminated, and a λ value of 0 would mean that the lasso is equivalent to the procedure for ordinary least squares regression. The lasso does variable selection, and I used the lasso command in Stata to select the independent variables for our model from the list mentioned above. The independent variables selected by the lasso are the selectivity and hard work interaction variable of interest for my dissertation, marital status, gender, age, experience in years to the fourth power, ASVAB score percentile, attainment of a graduate degree, a major in business, marketing, or management, and whether or not one's health limits the amount of work they can perform.

First, I estimate wage equations for white and marginalized graduates from each major with an interaction term examining the combined effects of positive attitudes toward hard work as measured by our personality scale variables and a measure of college selectivity. We have two groups, White (W) and Marginalized (M), which consists of Black, Latino, and mixed-race respondents), an outcome variable, W, and independent variables selected using the lasso listed above, as well as our variable of interest, the interaction of college selectivity and attitudes toward hard work. We want to know how much of the mean outcome difference, where $E(W)$ denotes the expected value of the outcome variable wages, is explained by group differences in the predictors (with a focus on our interaction term), where:

$$D = E(W_W) - E(W_M) \tag{45}$$

We are working with a linear wage equation model:

$$W_L = X_L^{'}\beta_L + \epsilon_L \tag{46}$$

where $E(\epsilon) = 0$ and $L \in (W, M)$. X is a vector containing the predictors and a constant, β

contains the slope parameters and the intercept, and ϵ is the error term. The mean outcome difference can be expressed as the difference in the linear prediction of the group-specific means of our regressors:

$$D = E(W_W) - E(W_M) = E(X_W)'\beta_W - E(X_M)'\beta_M \tag{47}$$

We want to examine a "twofold" decomposition in which the outcome difference is separated into the following two components:

$$D = E + C \tag{48}$$

The first component is the part of the differential attributed to group differences in the predictors (the "endowments effect"):

$$E = (E(X_W) - E(X_M))'\beta_M \tag{49}$$

The second component is the part of the differential attributed to the differences between the coefficients and the intercept:

$$C = E(X_M)'(\beta_W - \beta_M) \tag{50}$$

Some economists (e.g. Cotton (1988)) argue that when there are racial wage differentials, the undervaluation of one group (wage cost) is accompanied by the overvaluation of another (wage benefit; see Figure 1). We wish to examine this alternative decomposition. The general idea is that there is a non-discriminatory coefficient vector β^* that can be used to examine the contribution of differences in predictor variables to the difference in outcomes. Following Jann (2008), the outcome difference is written as

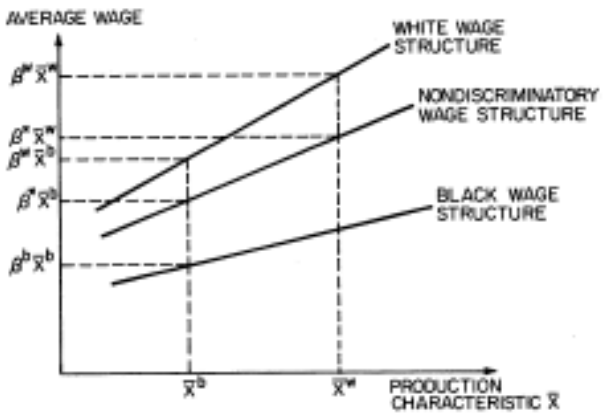


Figure 1: Source: Cotton 1988

$$D = (E(X_W) - E(X_M))'\beta^* + (E(X_W)'(\beta_W - \beta^*) + E(X_M)'(\beta^* - \beta_M)) \tag{51}$$

which is a twofold composition:

$$D = Q + U \tag{52}$$

We will follow Jann (2008) so that the first component is as follows:

$$Q = (E(X_W) - E(X_M))'\beta^* \tag{53}$$

This is the portion of the outcome differential that is explained by differences in the predictors (a "quantity effect"), where the second component

$$U = (E(X_W)'(\beta_W - \beta^*) + E(X_M)'(\beta^* - \beta_M)) \tag{54}$$

is the unexplained part of the outcome differential. To understand the benefit to one group and the cost to the other, we can use the Cotton (1988) method of weighting the coefficients by group sizes n_W and n_M which is shown as follows.

$$\widehat{\beta^*} = \frac{n_W}{n_W + n_M}\widehat{\beta_W} + \frac{n_M}{n_W + n_M}\widehat{\beta_M} \tag{55}$$

To answer research question 2, we can use a detailed decomposition method that will help us investigate how much of the explained differential between white and marginalized (Black and Latino) graduates can be attributed to the effects of our college selectivity and hard work interaction predictors. Isolating the effects of our college selectivity and hard work interaction predictor is fairly straightforward. Following (Jann 2008), the total component of the explained part is a sum over the individual contributions

$$\widehat{Q} = (\overline{X}_W - \overline{X}_M)'\widehat{\beta_W} = (\overline{X}_{1W} - \overline{X}_{1M})\widehat{\beta_{1W}} + (\overline{X}_{2W} - \overline{X}_{2M})\widehat{\beta_{2W}} + ... \tag{56}$$

where $\overline{X}_1, \overline{X}_2$...represent the means of the single regressors and $\widehat{\beta}_1, \widehat{\beta}_2$...are associated coefficients. I estimate decompositions with the pooled and Cotton methods to compare the results. It should be noted that in our particular case, the interpretation of the coefficient differences as discrimination warrant some caution. There is an advantage in the constant term of the marginalized graduates wage equations that is unexplained. I will also discuss some of the debate surrounding inclusion of the AFQT or ASVAB score in wage equations for each racial group in the results section.

5.6 Extension to Panel Data and Causal Inference

For each group of graduates, white men, non-white women, non-white men, and white women, we have an equation estimating the relationship between their salary in the initial year of the ten year period, Y_t (natural log of a continuous salary income variable measured in current years dollars), their vector of endowments in the initial year t, X_t, and a vector containing the returns to their characteristics in the initial year β_t and an intercept in the initial year. ϵ_t is an error term. Thus, we have the following equations for the salaries of

non-white women and white men at initial year t and end year s:

$$Non-white\ women\ at\ time\ t: Y_t^{BW} = X_t^{BW}\beta_t^{BW} + \upsilon_t^{BW} + \epsilon_t^{BW}, \quad E(\epsilon_t^{BW}) = 0$$

$$\mathrm{COV}(\mathrm{X}_t, \epsilon_t) = 0 \quad (57)$$

$$Non-white\ women\ at\ time\ s: Y_s^{BW} = X_s^{BW}\beta_s^{BW} + \upsilon_s^{BW} + \epsilon_s^{BW}, \quad E(\epsilon_s^{BW}) = 0$$

$$\mathrm{COV}(\mathrm{X}_s, \epsilon_s) = 0 \quad (58)$$

$$White\ men\ at\ time\ t: Y_t^{WM} = X_t^{WM}\beta_t^{WM} + \upsilon_t^{WM} + \epsilon_t^{WM}, \quad E(\epsilon_t^{WM}) = 0$$

$$\mathrm{COV}(\mathrm{X}_t, \epsilon_t) = 0 \quad (59)$$

$$White\ men\ at\ time\ s: Y_s^{WM} = X_s^{WM}\beta_s^{WM} + \upsilon_s^{WM} + \epsilon_s^{WM}, \quad E(\epsilon_s^{WM}) = 0$$

$$\mathrm{COV}(\mathrm{X}_s, \epsilon_s) = 0 \quad (60)$$

With a three-fold decomposition in a panel regression, there is the explained difference between the endowments of non-white Women and white Men at time t we call E_t, the unexplained difference between the returns to the characteristics of non-white women and white men at time t we call C_t, and the difference between the time-constant error terms of non-white women and white men at t, U_t. So we have:

$$\Delta Y_t = E_t + C_t + U_t \quad (61)$$

$$E_t = [E(X_t^{WM}) - E(X_t^{BW})]'\beta_t^{BW} \tag{62}$$

$$C_t = E(X_t^{BW})'(\beta_t^{WM} - \beta_t^{BW}) \tag{63}$$

$$U_t = E[\upsilon^{WM}] - E[\upsilon^{BW}] \tag{64}$$

Some economists (e.g., Cotton (1988)) argue that when there are racial wage differentials, the undervaluation of one group (wage cost) is accompanied by the overvaluation of another (wage benefit; see Figure 1). We wish to examine this alternative decomposition. The general idea is that there is a non-discriminatory coefficient vector β_t^* that can be used to examine the contribution of differences in endowment variables to the difference in outcomes. The salary disparity at time t is written as

$$\Delta Y_t = [E(X_t^{WM}) - E(X_t^{BW})]'\beta_t^* + (E(X_t^{WM})'(\beta_t^{WM} - \beta_t^*) + E(X_t^{BW})'(\beta_t^* - \beta_t^{BW})) + E[\upsilon^{WM}] - E[\upsilon^{BW}] \tag{65}$$

To understand the benefit to one group and the cost to the other, we can use the Cotton (1988) method of weighting the coefficients by group sizes n_{WM}, n_{WW}, n_{BM} n_{BW} which is shown as follows.

$$\widehat{\beta_t^*} = \frac{n_{WM}}{n_{WM} + n_{BW} + n_{BM} + n_{WW}}\widehat{\beta_t^{WM}} + \frac{n_{BW}}{n_{WM} + n_{BW} + n_{WW} + n_{BM}}\widehat{\beta_t^{BW}} + ... \tag{66}$$

We can use a detailed decomposition method that will help us investigate how much of

the unexplained differential between white men and Black women at times t and s can be attributed to the effects of our college selectivity, effort, degree attainment, and achievement variables. Isolating the effects of our various explanatory variables is fairly straightforward. The total component of the unexplained part is a sum over the individual contributions

$$\widehat{Q} = (\overline{X}_{WM} - \overline{X}_{BW})'\widehat{\beta_{WM}} = (\overline{X}_{1WM} - \overline{X}_{1BW})\widehat{\beta_{1WM}} + (\overline{X}_{2BW} - \overline{X}_{2BW})\widehat{\beta_{2WM}} + ... \quad (67)$$

where $\overline{X}_1, \overline{X}_2$...represent the means of the single regressors and $\widehat{\beta}_1, \widehat{\beta}_2$...are associated coefficients.

Ultimately, if we consider the attendance of an HBCU or supportive, high-quality, high-mobility school to be an 'intervention' and the salaries of graduates at times t and s to be our outcomes, utilize appropriate fixed or random effects, and control for relevant observable characteristics, we can look at differences in the differences between the wages of Black graduates and white graduates between 1993 and 2003 and 2008 and 2018 and quantify 'overvaluation' and 'undervaluation' of salaries using the frameworks of Cotton (1988) and Darity Jr. et al. (2022). We have the difference in disparities in the salaries of non-white women and white men over time below:

$$\Delta Y_S - \Delta Y_t = DiD \quad (68)$$

$$\Delta Y_s = [E(X_s^{WM}) - E(X_s^{BW})]'\beta_s^* + (E(X_s^{WM})'(\beta_t^{WM} - \beta_s^*) + E(X_s^{BW})'(\beta_s^* - \beta_s^{BW})) + E[\upsilon^{WM}] - E[\upsilon^{BW}] \quad (69)$$

$$\Delta Y_t = [E(X_t^{WM}) - E(X_t^{BW})]'\beta_t^* + (E(X_t^{WM})'(\beta_t^{WM} - \beta_t^*) + E(X^{BW})'(\beta_t^* - \beta^{BW})) + E[\upsilon^{WM}] - E[\upsilon^{BW}] \quad (70)$$

5.7 Synthetic Control Difference-in-Differences and Cotton Decomposition Method

An alternative to estimating the value added of selective institutions and decomposing the contributions of different independent variables and drivers of the equalizing effects of selective school attendance on earnings is to combine synthetic control difference-in-differences, the Cotton decomposition, and theory on the intersectionality of identities when studying racial inequality. For the sake of the reader I will refer them to the seminal article on synthetic difference-in-differences by Arkhangelsky et al. (2021) rather than reproducing their explanation. A recent article by Rambachan and Roth discusses how violations of the parallel trends assumption can be treated using some new methods (Rambachan and Roth, 2023). The synthetic control difference-in-differences estimator for average treatment effects on the treated is represented by the following equation, which is merely a combination of the mathematics behind the synthetic control method and the difference-in-difference estimator:

$$(\hat{\chi}^{\text{sdid}}, \hat{\alpha}, \hat{\mu}, \hat{\beta}) = \arg\min_{\alpha,\mu,\beta} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\chi)^2 \widehat{w_i^{\text{sdid}}} \widehat{\lambda_t^{\text{sdid}}} \right\} \quad (71)$$

The same logic applies here. We can decompose the differences between non-white women and white men college graduates from HBCUs and PWIs (or Black men and white men, or hypothetically any mix of race, sex, sexual orientation and gender identities given data available), and split it into an undervaluation and overvaluation. This can be expressed via percentages or in dollar terms, with intuitive application to understanding how this could impact household wealth disparities under a counterfactual scenario in which graduates invest

these saved earnings in relatively safe, interest accruing assets.

5.8 Perceived Ability and the Two Way Causality of the Wage Equations in Systemically Unequal Economies

Now, consider a wage equation with typical demographic characteristics and some unobserved variables as co-variates: health status, marital status, educational attainment, experience, discriminatory reactions to increased status or earnings, ability, perceived ability, and strength of coalitions or trust among the marginalized, selectivity of college attended, effort, locus of control, quality of social network, and our perceived ability variable that evolves with successful experiences or setbacks. The wage is a continuous dependent variable Y, where we transform the function using natural logs to account for the skewed wage distribution.

$$Y_t = X_t\beta_t - D[Y_t, a, \Theta(a)] + \Theta(a) + \lambda_t + \pi_t + \epsilon_t \tag{72}$$

Here we see the two way causality present when estimating wage equations to calculate undervaluation and overvaluation of wages using the Kitagawa-Oaxaca-Blinder mean decomposition technique of components of differences in the rates of two variable. If we have wages of Black female and white male graduates at time t and s, we can calculate causal estimates of interventions meant to improve the human capital acquisition of marginalized graduates.

Using the method of Cotton (1988) we can understand the undervaluation of marginalized women's wages the overvaluation of white men's wages, assuming a non-discriminatory wage prevails without discrimination that is determined by weighting the coefficients of the individual groups with multiple identities by their group sizes. Increased self-image and social networks increases wages, and the success of mobility within ones field raises these co-variates. We also have to account for discriminatory reactions to gains. The net effects depend on market conditions and the discriminatory attitudes and actions prevalent in the

economy.

5.9 Standard Differences-in-Differences

The traditional DiD estimator compares treated and control units before and after treatment:

$$\hat{\tau}_{DiD} = \left(\bar{Y}_{T,post} - \bar{Y}_{T,pre}\right) - \left(\bar{Y}_{C,post} - \bar{Y}_{C,pre}\right), \tag{73}$$

where $\bar{Y}$ denotes the average outcome in treated (T) and control (C) groups over pre- and post-treatment periods. This relies on the **parallel trends assumption**, which may not always hold.

5.10 Synthetic Control Method

Synthetic Control Method (SCM) estimates a counterfactual for the treated unit by constructing a weighted combination of control units:

$$\hat{Y}_{T,t}^{N} = \sum_{i \in C} w_i Y_{i,t}, \tag{74}$$

where weights w_i are chosen to minimize:

$$\sum_{t \in pre} \left(Y_{T,t} - \sum_{i \in C} w_i Y_{i,t} \right)^2, \tag{75}$$

subject to $w_i \geq 0$ and $\sum w_i = 1$.

5.11 Bayesian Approach to Difference-in-Differences

The Bayesian framework incorporates prior beliefs and uncertainty in DiD and SCM. Let $Y_{T,t}^{(0)}$ denote the counterfactual outcome for the treated unit in the absence of treatment.

The posterior distribution for the treatment effect τ_t is:

$$P(\tau_t|Y, X) \propto P(Y|\tau_t, X)P(\tau_t), \tag{76}$$

where:

- $P(Y|\tau_t, X)$ is the likelihood given data Y and covariates X,
- $P(\tau_t)$ is the prior distribution for the treatment effect.

A common model for Bayesian DiD assumes:

$$Y_{T,t}^{(0)} \sim \mathcal{N}(\mu_t, \sigma^2), \tag{77}$$

where μ_t follows a hierarchical structure based on synthetic control weights:

$$\mu_t = \sum_{i \in C} w_i Y_{i,t}. \tag{78}$$

The posterior mean and credible intervals for τ_t can be computed using Markov Chain Monte Carlo (MCMC) sampling:

$$\tau_t = Y_{T,t} - \mu_t. \tag{79}$$

Bayesian Differences-in-Differences with Synthetic Controls improves causal inference by incorporating prior beliefs and uncertainty estimation, addressing violations of parallel trends and improving counterfactual estimation.

5.12 A Synthesis of Robust Bayesian Difference-in-Differences with Synthetic Controls and Intersectional Kitagawa-Oaxaca-Blinder Decomposition

We extend the Bayesian Difference-in-Differences (DiD) framework by integrating synthetic control methods to estimate counterfactual wage trajectories for marginalized groups facing intersectional discrimination. Let w_{it} represent the wage of worker i at time t, and define the wage equation as:

$$w_{it} = \alpha_i + \beta_t + \gamma D_{it} + \delta X_{it} + \varepsilon_{it}, \tag{80}$$

where:

- α_i represents individual fixed effects,
- β_t captures time-specific effects,
- D_{it} is a treatment indicator for discriminatory policies,
- X_{it} denotes other control variables,
- ε_{it} is an error term.

To integrate synthetic control estimation, let w_{it}^0 denote the counterfactual wage for a treated unit had it not received discrimination. We construct a synthetic control $\hat{w}_{it}^0$ as a weighted combination of control units:

$$\hat{w}_{it}^0 = \sum_{j \neq i} \omega_j w_{jt}, \tag{81}$$

where ω_j are non-negative weights summing to one, chosen to minimize the pre-treatment wage difference:

$$\min_{\omega} \sum_{t<T} (w_{it} - \sum_{j \neq i} \omega_j w_{jt})^2. \tag{82}$$

The Bayesian posterior distribution for wage discrimination effects updates prior beliefs using observed wage gaps:

$$P(\gamma|X) \propto P(X|\gamma)P(\gamma), \tag{83}$$

where prior beliefs $P(\gamma)$ reflect expected disparities before intervention. The posterior distribution enables robust inference even under unobserved heterogeneity.

5.13 Intersectional Wage Differentials and Profitable Discrimination

Firms can exploit intersectional identities to segment workers and suppress wages. Let g_i denote a worker's intersectional group identity, and define the wage-setting equation:

$$w_i = w^* - \phi(g_i), \tag{84}$$

where w^* is the competitive wage, and $\phi(g_i)$ represents an intersectional wage penalty. Firms maximize profits:

$$\max_{w_i} \sum_i \left[pf(e_i) - w_i \right], \tag{85}$$

subject to labor supply constraints. Strategic wage differentials ensure worker segmentation, preventing solidarity and collective bargaining.

5.14 Measuring Dynamic Intersectional Differentials in Wage Growth Among the College Educated

To model wage growth in early careers, we define a worker's wage vector $\mathbf{w} \in \mathbb{R}^n$ over time as a function of human capital accumulation, experience, and employer-specific factors. Let $\mathbf{x}_t$ be a vector of individual characteristics at time t, including education, job tenure, and sector affiliation. We define wage growth as:

$$\mathbf{w}_{t+1} = \mathbf{A}\mathbf{w}_t + \mathbf{B}\mathbf{x}_t + \boldsymbol{\varepsilon}_t, \tag{86}$$

where:

- $\mathbf{A}$ is a transition matrix capturing persistence in wages,
- $\mathbf{B}$ represents the impact of human capital accumulation,
- $\boldsymbol{\varepsilon}_t$ is an error term capturing shocks to wages.

The steady-state wage distribution satisfies:

$$(I - \mathbf{A})^{-1}\mathbf{B}\mathbf{x} = \mathbf{w}, \tag{87}$$

where $(I-\mathbf{A})^{-1}$ represents the long-run effects of wage dynamics. This formulation allows us to analyze how initial endowments and career trajectories influence wage disparities.

We extend the Kitagawa-Oaxaca-Blinder decomposition to analyze intersectional wage disparities over time, using longitudinal wage data and income volatility as dependent variables. Let w_{it} represent the wage of individual i at time t, and decompose the wage function as:

$$w_{it} = X_{it}\beta + \varepsilon_{it}, \tag{88}$$

where:

- X_{it} is a vector of explanatory variables (education, experience, industry, etc.),
- β is the vector of coefficients,
- ε_{it} is the error term.

To analyze wage disparities between two groups A and B across time, we define:

$$\Delta w_t = (X_{A,t} - X_{B,t})\beta_B + X_{B,t}(\beta_A - \beta_B), \tag{89}$$

where the first term represents the explained portion of the wage gap due to differences in characteristics, and the second term represents the unexplained portion due to discrimination and structural inequality.

To incorporate income volatility, we define wage variance over time as:

$$\sigma^2_{w_i} = E[(w_{it} - \bar{w_i})^2], \tag{90}$$

and decompose volatility effects by group:

$$\Delta\sigma^2_w = (\sigma^2_{X_A} - \sigma^2_{X_B})\beta_B + \sigma^2_{X_B}(\beta_A - \beta_B). \tag{91}$$

Using a linear algebraic formulation, let wage vectors be represented in matrix form:

$$W = X\beta + \varepsilon, \tag{92}$$

where wage differences and volatilities are analyzed using matrix projections:

$$\Delta W = (X_A - X_B)\beta_B + X_B(\beta_A - \beta_B). \tag{93}$$

This method allows decomposition of wage gaps in a dynamic framework where income volatility plays a central role in understanding inequality.

6 Results

6.1 Non-White Graduates and White Graduates

6.1.1 Wage Costs and Wage Benefits

Using Cottons' decomposition method, I found that the unexplained wage gap between white graduates and non-white graduates in my sample is .103 log points, or approximately 10.84 percentage points if we compute $(e^{.103} - 1) * 100$. The wage cost affecting the wages of the non-white graduates is .075 log points or approximately 7.78 percentage points. The wage benefit accruing to the white graduates is .028 log points or approximately 2.83 percentage points. Splitting the penalty and wage benefit using pooled methodology produces significantly different results. The unexplained gap is smaller at .082 log points or approximately 8.54 percentage points, with a wage penalty experienced by non-white graduates of .058 log points (5.97 percentage points) and a wage benefit accruing to white graduates of .024 log points (2.43 percentage points). Although a 10.84% gap may seem insignificant, we can consider a case with two graduates possessing the same characteristics. The marginalized graduate earns $100,000 per year while the white graduate earns $110,000 per year. The loss in earnings over a 10 year period is $100,000. As is typically the case with high-income earners, we can assume that the lost $100,000 may have been invested in a mutual fund during those years. Quickly we see the stark difference in wealth that such unexplained gaps in wages and salaries can produce, even if considering the unlikely case that the graduates from the two racial groups have the same amount of wealth inherited from previous generations.

6.1.2 Detailed Decomposition Results

The gap between non-white graduates and white graduates' logarithmic wages is 8.11%. For the non-white graduate-white graduate decomposition's, the controls that were selected explain more than 100 percent of the gap in log wages. If non-white graduates had the same characteristics as white graduates, their wages would increase by 19.48%. This indicates that the non-white graduates actually have a higher wage structure than the white graduates at the beginning of their wage function. This seems to be due to an unexplained benefit of group membership that increases the y-intercept (constant term) of their wage function higher. However, the slope is less steep, indicating lower returns to their characteristics. Put simply, their wage functions intersect at a low level of characteristics and for the rest of the wage structure, white graduates have a wage advantage due to differences in coefficients.The unexplained portion of the differential is 10.84%, indicating that rather than having the advantage of their higher constant term and higher coefficients throughout the wage structure, there is an unexplained penalty due to their coefficients.

The characteristics that have the largest effects on the explained part of the differential are the selectivity and hard work interaction, ASVAB percentile, experience to the fourth power, and the marital status dummy. The marginalized graduates attended less selective schools on average, but had slightly higher scores on the attitudes towards hard work scale, according to the summary statistics. This explains the significance of the interaction on the explained part of the differential. The ASVAB (or AFQT) is a controversial and imperfect measure of ability, so with the large effects of the scores on the explained differential it is plausible that the ASVAB is a biased indicator for racial wage gaps, explaining more of the effects of discrimination with its inclusion.

A higher percentage of non-white graduates are female compared to white graduates (52.22 percent vs. 59.29 percent), and the non-white women in our sample have larger

unexplained gaps in their wages compared to the white men, so if the women in the non-white group were men, more of the differential would be explained. The significance of experience in explaining some of the gap makes sense, as non-white workers are the last hired and first fired, a structural characteristic of the economy that perpetuates gaps in resumes.

The characteristics that have the largest effects on the unexplained part of the wage differential are college selectivity and attitudes toward hard work interaction, work experience in years, ASVAB percentile, and the constant term. As explained earlier, the difference in constant terms cannot be analyzed, given that it represents unexplained differences in the wage structures of the two groups strictly due to group membership. The large differences in coefficients for the selectivity and hard work interaction term may be measuring what this study set out to understand, differences in the returns to hard work and attending a selective school. The difference in coefficients for the experience in years characteristic could be explained by the research on the smaller number of marginalized graduates in 'good jobs', which seems to indicate less promotions for non-white employees, which would result in smaller returns to years of experience in an occupation. A theoretical model by Costas Cavounidis and Kevin Lang (2015) which posits that non-white workers are more closely monitored and penalized for minor errors compared to their white colleagues could also help to make sense of these results. In their model, black employees are more closely monitored than white employees, leading to poorer performance reviews due to relatively minor errors, which would also lead to slower advancement and smaller returns to experience. With regard to ASVAB percentile, it makes sense that the "natural ability" of non-white workers is undervalued if we assume that there is widespread discrimination in wages (Cavounidis and Lang, 2015).

Now that we have analyzed the source of gaps in non-white and white graduates' wages, I will turn toward more detailed analysis by looking at main sources of racial gaps by gender.

I look at the gaps between non-white women and white men and non-white graduates and white men.

6.2 Non-White Graduates and White Men

6.2.1 Wage Costs and Wage Benefits

The economic advantage of white men over the non-white graduates is obscured by the lower wages of the white women in the sample. I estimated my results using white men and non-white graduates (both men and women). I found that the unexplained wage gap is 15.48% using the Cotton method. The wage cost experienced by non-white graduates is 6.82%. A wage benefit of 8.11% is experienced by the white men. The results for using pooled coefficients instead of Cotton's methodology produce an unexplained wage gap of 12.29%, with a 6.39% wage penalty for non-white graduates and a 5.54% wage benefit for the white men. The results are presented in table 5.

6.2.2 Detailed Decomposition Results

The detailed decompositions for white men graduates and non-white graduates are presented in table 6. The total wage gap between the wages of non-white graduates and white men is 41.05%. The Cotton decomposition produces an explained wage gap of 21.65% and an unexplained wage gap of 15.48%. The results of the pooled decomposition indicate a smaller unexplained part of the gap. These decomposition results appear to confirm that the gap between non-white graduates and white graduates that I calculated is significantly reduced by the inclusion of white women, who have lower wages than both the marginalized men and women in my sample. With the explained part of the gap, ASVAB percentile score is once again the largest contribution when looking at the detailed decomposition. Again, the verdict is still out on whether including the ASVAB in racial wage differential equations

does not introduce bias due to biases in the testing and preparation for the testing resulting from one's environment (those with higher ASVAB scores may just have had higher socioeconomic status). Marital status is important but far less influential on the explained gap than ASVAB percentile, which is important considering mass incarceration and recent research into how that has affected marriage rates in non-white communities (essentially a large pool of potential partners is incarcerated). The remaining independent variables contribute small percentages to the overall explained gap, with our college selectivity and hard work score interaction term the next most significant contribution to the explained part of the gap.

The unexplained part of the wage differential is once again impacted significantly by the college selectivity and attitude toward hard work personality score interaction term. This contributes the most to the unexplained differential by far. Next is the contribution of the constant term, which in this scenario is lowering the unexplained gap between the two groups. The constant term is the unexplained part of the gap that we have no explanation for, so there is no insight to be gained. The ASVAB percentile score variable detracts from the unexplained part of the differential because the non-white graduates actually have a larger coefficient for the effect of ASVAB percentile on wages. Experience and experience to the fourth power have opposite but significant effects on the unexplained part of the differential. The non-white graduates also do not have the same return to marital status as the white graduates, and a graduate degree is rewarded slightly more when a non-white graduate has one (they may have to face higher standards when hired or evaluated due to discrimination).

6.3 Non-White Women and White Men

6.3.1 Wage Costs and Wage Benefits

As shown in table 7, the Cotton methodology of calculating the wage benefit and wage penalty associated with differentials reveals that the unexplained gap between the non-white

women graduates and white men is 25.60%. There is a wage cost experienced by non-white women of 14.11%. The wage benefit accruing to white men is 10.07%. Again, the results are different when using pooled coefficients, with an unexplained wage gap of 18.88%. With the pooled methodology the wage cost is 12.07% and the wage benefit is 6.07%.

6.3.2 Detailed Decomposition Results

The detailed decomposition results for the white men and non-white women graduates are presented in table 8. The wage gap between non-white women and white men is 52.65%. Using Cotton decomposition methods, of the 52.65% total gap, 18.29% is explained by differences in the control characteristics that we used and 25.60% was unexplained. This gap is the largest of the race by gender gaps I analyzed, most likely reflecting what Holder (2020) has named the "double gap" wage penalty experienced by black women. The double gap reflects the fact that there is evidence of a wage penalty on Black women for being women and an additional penalty for being Black. I think this may also hold true for our aggregate non-white women group in my sample consisting of Latina, Black, and mixed-race women.

The difference in ASVAB percentile characteristics contributes the most to the explained part of the gap between the non-white women and white men graduates. As discussed before, the AFQT, which the ASVAB is based on, has been questioned as a biased indicator for racial wage gaps, so that its importance in the explained part of the wage gap may be measuring some type of racial bias or discrimination. Our selectivity and hard work interaction term the explains the rest of the gap, again reflecting the differences in selectivity for colleges attended between marginalized and white graduates. Marital status is crucial here, as educated black women tend to have lower marriage rates than other demographics. The explained part of the gap would be larger if not for the higher educational attainment of the marginalized women in the sample, the most educated demographic out the four groups in this sample

(see data section for summary statistics for each demographic).

Here, we see the importance and contribution of the college selectivity and attitudes toward hard work interaction term to the unexplained part of the wage gap between non-white women and white men. Marital status is also salient, with non-white women not getting the same return to marriage as white men in the sample. ASVAB percentile is also significant, although as discussed before, there is serious debate about what exactly these armed forces tests truly measure. The difference in experience in years coefficients contributes some to the unexplained gap, but differences in the returns to college selectivity and attitudes towards hard work are the main effects impacting the unexplained part of the differential.

It seems that the non-white women college graduates in our sample are at a significant disadvantage when we examine whether their pursuit of high quality education and high work ethic pays off in the same way as white men in our sample. This result confirms our hypothesis that non-white graduates have to be more exceptional than their white peers with comparable attitude toward hard work personality scores and college selectivity in order to achieve the same wages. It should also be noted that although Black women, Latina women, and mixed-race women have become increasingly educated in terms of graduate degrees and bachelor's degrees attained, these results indicate that this achievement is not being reflected in their income from wages and salaries.

7 Conclusions and Extensions

Moving forward, a larger dataset with information on labor market outcomes, college selectivity, and a set of rich demographic characteristic independent variables would help to understand the complexities in differences in returns to college selectivity and positive attitudes toward hard work or high motivation influence racial wage differentials. The main

conjectures were proven to be true, and the same differences in returns to college selectivity and positive attitudes towards hard work may hold true with an analysis by gender.

Another potential extension would be to look at differences in the returns to these characteristics for immigrants from marginalized groups in comparison with white American graduates. For example, sometimes immigrant non-white graduates are viewed more favorably and as more hard working than people from marginalized groups born in the U.S. This difference in perception may be reflected in the differences in their returns to our interaction term in comparison with native born graduates from marginalized groups. There is also the possibility of calculating wage gaps by considering differences in the returns to these characteristics through an intersectional lens where people hold multiple marginalized identities (for example a disabled woman that is also queer and Black, Latino, or Native American) (Darity Jr. et al., 2022). Previous research in this area suggests that the combination of two more socially marginalized identities does not necessarily interact in additive ways. These research questions warrant further investigation with a larger data set so that results can be generalized to the U.S. population, shedding light on the wage penalties and disparate rewards for achievement between marginalized and white college graduates.

8 Appendix: Mass Incarceration as a Destructive Policy Choice Instead of Equitable Education and Health Equity

Andrews (1993) begins his examination of the emergence of an educational underclass and the attempts to control such a 'surplus' population by explaining the features of the economy. There is one produced commodity used for either investment or consumption. Investment is either increases in capital deepening or expenditures on education to increase 'human

capital'. The population is composed of four groups: educated workers (N), students (S), unemployable members of the underclass (U), and prisoners (P). Workers own all capital and receive all profits. Firms produce output through a fixed coefficients production function with educated workers (N) and capital goods (K)

$$Y^S = min[aN, K/s] \tag{94}$$

where a is level of output per worker (a measure of productivity) and x is the capital-output ratio. Andrews assumes that output is constrained by the supply of educated workers.

An 'educational underclass" can develop if the supply of high quality education is less than the number of students that require training (see Blair and Smetters (2021) for evidence of this with elite colleges artificially keeping supply low and welfare loss consequences). Alternatively, an educational underclass can develop if a significant segment of the population cannot complete the required training to become an educated worker for whatever reasons.

Andrews assumes a rate of population growth (g) for all classes, and the number of new students in need of training at any time is gT. A fraction of the society's existing population will fail and become unemployable ($0 < f(b) < 1$). If education system capacity is sufficient for population growth and needs, the growth of the student population is

$$DS = gT - hS \tag{95}$$

where $0 < h < 1$ is the graduation rate. The increase in the size of the underclass at any time is

$$DU = f(b)S \tag{96}$$

where be is the level of education spending per student and the failure rate $f(b)$ is nega-

tively related to b. If education supplied is insufficient $S = E$, then

$$DS = DE - hS \tag{97}$$

and

$$DU = gT - DE + f(b)S \tag{98}$$

There is a tax rate β, and a government budget constraint expressed as

$$\beta Y = bS + vDE + rU \tag{99}$$

where v is the cost of building a unit of educational system capacity and r is the level of poor relief per member of the educational underclass. If one assumes a balanced budget, the increase in the capacity of the education system is described by the following equation

$$DE = v^{-1}[\beta a(N/T) - b(S/T) - r(U/T)] \tag{100}$$

where S/T and U/T are the fraction of the total population in school (schooling ratio) and fraction in the underclass (underclass rate). After examining the mathematical properties of the system, Andrews finds that society has a stable (though potentially high) underclass rate if choosing to spend more on education per student than poor relief. Jails are introduced into the model as the government in this scenario enforces laws of contract and public safety on behalf of an educated population. Police and prisons are used for controlling crime and detaining members of the evolving underclass.

He finds that the larger the apprehension rate, the smaller the equilibrium underclass rate and the larger is the equilibrium incarceration rate. In short, the strategy of imprisoning the under educated rather than spending the money to educate them simply moves those

in the unemployable underclass from deprivation in the general population to deprivation of a different kind in jails. In the end, poverty is not alleviated and this theoretical exercise confirms the fears of many criminal justice reform advocates: incarceration is an extremely costly reaction to a symptom caused by poverty and lack of investment in human beings mental capabilities and primary functionings (health, education, productive employment, housing security, and civil rights/dignity).

The Opportunity and Inclusive Growth Institute models mass incarceration as a dynamic system influenced by economic incentives, labor market discrimination, and policy interventions. Let P_t represent the prison population at time t, evolving according to:

$$P_{t+1} = \lambda P_t + \eta E_t - \mu R_t, \tag{101}$$

where:

- λ represents the persistence rate of incarceration,
- η captures entry rates into the carceral system, influenced by economic conditions and discrimination,
- E_t represents new entrants to the prison system,
- μ represents the reintegration success rate,
- R_t represents the number of released individuals.

The entry function is modeled as:

$$E_t = \alpha + \beta U_t + \gamma D_t + \delta C_t + \varepsilon_t, \tag{102}$$

where:

- U_t represents the unemployment rate,

- D_t represents demographic and racial disparities in policing,
- C_t represents crime rates,
- ε_t is an error term capturing unobserved shocks.

Reintegration is governed by:

$$R_t = \rho P_t + \sigma L_t - \theta B_t, \tag{103}$$

where:

- L_t represents labor market reintegration policies,
- B_t represents barriers to reentry (e.g., employer discrimination, lack of social support).

The steady-state incarceration rate satisfies:

$$P^* = \frac{\eta E^*}{1 - \lambda + \mu}. \tag{104}$$

This model highlights the economic feedback loops that reinforce incarceration and suggests policy interventions targeting η, μ, and ρ to reduce incarceration rates and improve the results of reintegration.

References

Alexis, Marcus (1973) "A Theory of Labor Market Discrimination with Interdependent Utilities," *The American Economic Review*.

Alsalam, Nabeel, Joseph Conety, Estelle James, and Duc-Le To (1989) "College Quality and Future Earnings: Where Should You Send Your Child to College?" *The American Economic Review*, 79, 247–252, http://www.jstor.org/stable/1827765.

Andrews, Marcellus (1993) "Schools, jails, and dynamics of an educational underclass: some dreadful social arithmetic," *Journal of Economic Behavior and Organization*.

——— (1999a) *The Political Economy of Hope and Fear: Capitalism and Black Condition in America*: Yale University Press.

——— (1999b) *The Political Economy of Hope and Fear: Capitlism and the Black Condition in America*: New York University Press.

Andrews, Rodney, Marcus Casey, Bradley Hardy, and Trevon Logan (2017) "Location matters: Historical racial segregation and intergenerational mobility," *Economic Letters*.

Arkhangelsky, Dmitry, Susan Athey, David Hirshberg, Guido Imbens, and Stefan Wager (2021) "Synthetic Difference-in-Differences," *American Economic Review*.

Barrow, Lisa and Cecilia Elena Rouse (2005) "Do Returns to Schooling Differ by Race and Ethnicity?" *American Economic Review*, 95, 83–87, 10.1257/000282805774670130.

Blackburn, McKinley L. (2004) "The role of test scores in explaining race and gender differences in wages," *Economics of Education Review*, 23, 555–576, 10.1016/j.econedurev.2004.04.005.

Blair, Peter Q. and Kent Smetters (2021) "Why Don't Elite Colleges Expand Supply?" *National Bureua of Economic Research*.

Blinder, Alan S. (1973) "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources*, 8, 436, 10.2307/144855.

Brundage, Brendan C. and Daniele Tavani (2024) "Group Conflict, Racial Inequality, and Stratification," *The Review of Black Political Economy*.

Campbell, Kathryn, Anthony Carnevale, Artem Gulish, Jeff Strohl, and Martin Van Der Werf (2019) "The Unequal Race for Good Jobs: How Whites Made Outsized Gains in Education and Good Jobs Compared to Latinos and Blacks," https://cew.georgetown.edu/cew-reports/raceandgoodjobs/.

Cavounidis, Costas and Kevin Lang (2015) "Discrimination and Worker Evaluation," 10, https://www.nber.org/papers/w21612.

Cavounidis, Costas, Kevin Lang, and Russell Weinstein (2023) "Boss is Watching: How Monitoring Decisions Hurt Black Workers," *The Economic Journal*.

Chelwa, G, Darrick Hamilton, and James Stewart (2022) "Stratification Economics: Core Constructs and Policy Implications," *Journal of Economic Literature*.

Chetty, Raj, Nathaniel Hendren, Maggie Jones, and Sonya R. Porter (2019) "Race and Economic Opportunity in the United States: an Intergenerational Perspective," *The Quarterly Journal of Economics*.

Cotton, Jeremiah (1988) "On the Decomposition of Wage Differentials," *The Review of Economics and Statistics*, 70, 236, 10.2307/1928307.

Craig, Ashley and Roland Fryer (2018) "Complementarity Bias: A Model of Two-Sided Statistical Discrimination," *Harvard University Working Paper*.

Dale, Stacy and Alan Krueger (2011) "Estimating the Returns to College Selectivity Over the Career Using Administrative Earnings Data," 06, https://www.nber.org/papers/w17159.

Darity Jr., William A. (2001) "The Functionality of Market Based Discrimination," *International Journal of Social Economics*.

Darity Jr, William and Dania Francis (2020) "Isolation: an Alternative to the "Acting White" Hypothesis in Explaining Black Under-Enrollment in Advanced Courses," *Journal of Economics, Race, and Policy*.

——— (2021) "Separate and Unequal Under One Roof: How the Legacy of Racialized Tracking Perpetuates Within-School Segregation," *The Russell Sage Foundation Journal of the Social Sciences*.

Darity Jr., William and Samuel Myers Jr. (1999) *Persistent Disparity: Race and Economic Inequality in the United States since 1945*: Edward Elgar Publishing.

Darity Jr., William, Mark Paul, and Khaing Zaw (2022) "Returns in the Labor Market: A Nuanced View of Penalties at the Intersection of Race and Gender in the U.S.," *Feminist Economics*.

Datcher-Loury, Linda, Julianne Malveaux, and Phyllis Wallace (1980) *Black Women in the Labor Force*: MIT Press.

Dickerson, Niki T. and Jerry A. Jacobs (2006) "Race differentials in college selectivity, 1981–2000," *Journal of African American Studies*, 10, 3–18, 10.1007/s12111-006-1009-0.

Dickson, Lisa (2010) "Race and Gender Differences in College Major Choice," *The ANNALS of the American Academy of Political and Social Science*, 627, 108–124, 10.1177/0002716209348747.

Diette, Timothy M., Arthur H. Goldsmith, Darrick Hamilton, and William Darity (2015) "Skin Shade Stratification and the Psychological Cost of Unemployment: Is there a Gradient for Black Females?" *The Review of Black Political Economy*, 42, 155–177, 10.1007/s12114-014-9192-z.

Dondero, Molly and Leticia Marteleto (2016) "Racial Inequality in Education in Brazil: A Twins Fixed-Effects Approach," *Demography*.

Edwards, Ashley, Justin Ortagus, Jonathan Smith, and Andria Smythe (2023) "HBCU Enrollment and Longer Term Outcomes," *Washington Center for Equitable Growth*.

Fernandes, Gustavo Andrey De Almeida Lopes (2015) "Brazilian Female Labor Market: Racial-Skin Color Discrimination and Inefficiency," *Economia Aplicada*, 19, 241–259, 10.1590/1413-8050/ea85456.

Filippin, A. and M Paccagnella (2011) "Family background, self-confidence, and economic outcomes," *IZA Discussion Paper Series*.

Francis, Dania, Bradley Hardy, and Damon Jones (2022) "Black Economists on Race and Policy: Contributions to Education, Poverty and Mobility, and Public Finance," *Journal of Economic Literature*.

Goldsmith, Arthur H., William Darity, and Jonathan R. Veum (1998) "Race, Cognitive Skills, Psychological Capital and Wages," *The Review of Black Political Economy*, 26, 9–21, 10.1007/s12114-998-1001-0.

Goldsmith, Arthur H, Darrick Hamilton, and William Darity (2006) "Shades of Discrimination: Skin Tone and Wages," *American Economic Review*, 96, 242–245, 10.1257/000282806777212152.

Goldsmith, Arthur H., Jonathan R. Veum, and William Darity Jr. (1997) "The Impact of Psychological and Human Capital on Wages," *Economic Inquiry*, 35, 815–829, 10.1111/j.1465-7295.1997.tb01966.x.

Handy, John (1998) "The Shadow of the Future and the Complexity of Cooperation," *The Review of Black Political Economy*.

Holder, Michelle (2020) "The "Double Gap" and the Bottom Line: African American Womens' Wage Gap and Corporate Profits," 03, https://rooseveltinstitute.org/publications/the-double-gap-and-the-bottom-line-african-american-womens-wage-gap-and-corporate-profits/.

Hornbeck, Richard and Trevon D. Logan (2022) "The Great Leap: The Decline of Agriculture and the Rise of the American Dream," *Journal of Economic History*, 82 (1), 147–186, 10.1017/S0022050721000579.

Hoxby, Caroline and Christopher Avery (2013) "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity*, 2013, 1–65, 10.1353/eca.2013.0000.

Hoxby, Caroline M. and Sarah Turner (2015) "What High-Achieving Low-Income Students Know About College," *American Economic Review*, 105, 514–517, 10.1257/aer.p20151027.

Jann, Ben (2008) "The Blinder–Oaxaca Decomposition for Linear Regression Models," *The Stata Journal: Promoting communications on statistics and Stata*, 8, 453–479, 10.1177/1536867x0800800401.

Jaynes, Gerald (1982) "Production and Distribution in Agrarian Economies," *Oxford Economic Papers*.

Kitagawa, Evelyn M. (1955) "Components of a Difference Between Two Rates," *Journal of the American Statistical Association*, 50, 1168, 10.2307/2281213.

Kreisman, Daniel and Marcos A. Rangel (2015) "On the Blurring of the Color Line: Wages and Employment for Black Males of Different Skin Tones," *Review of Economics and Statistics*, 97, 1–13, 10.1162/rest_a_00464.

Lewis, W. Arthur (1985) *Racial Conflict and Economic Development*: Harvard University Press.

Loury, Glenn C. (1976) "A Dynamic Theory of Income Differences," *Northwestern University Discussion Paper*.

——— (1981a) "Intergenerational transfers and the distribution of earnings," *Econometrica*.

——— (1981b) "Intergenerational Transfers and the Distribution of Income," *Econometrica*.

Loury, Linda Datcher and David Garman (1995) "College Selectivity and Earnings," *Journal of Labor Economics*, 13, 289–308, 10.1086/298375.

Malveaux, Julianne (1980) "Unemployment Differentials by Race and Occupation," *Massachussetts Institute of Technology Ph.D thesis*.

Marglin, Stephen (2021) *Raising Keynes: A Twenty-First Century General Theory*: Harvard University Press.

Mason, Patrick (1997) "Male Interracial Wage Differentials: Competing Explanations," *Cambridge Journal of Economics*.

——— (2007) "Intergenerational Mobility and Interracial Inequality: The Return to Family Values," *Industrial Relations*.

McClough, David and Mary Ellen Benedict (2017) "Not All Education Is Created Equal: How Choice of Academic Major Affects the Racial Salary Gap," *The American Economist*, 62, 184–205, 10.1177/0569434516683029.

Neumark, David (1988) "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *The Journal of Human Resources*, 23, 279–295, 10.2307/145830.

Oaxaca, Ronald (1973) "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14, 693, 10.2307/2525981.

Okafor, Chika O. (2022) "Social Networks as a Mechanism for Discrimination," *arXiv*.

Pereira Serra, Gustavo (2024) "(Trying to) Catch Up with the Higher-Skilled Joneses: Student loans in a segmented educational market from a Post-Keynesian perspective," *New School for Social Research Economics Working Paper Series*.

Price, Gregory N. and C.G. Viceisza (2023) "What Can Historically Black Colleges and Universities Teach about Improving Higher Education Outcomes for Black Students?" *Journal of Economic Perspectives*.

Rambachan, Ashesh and Jonathan Roth (2023) "A More Credible Approach to Paralell Trends," *Review of Economic Studies*.

Rodgers, William M. and William E. Spriggs (1996) "What Does the Afqt Really Measure: Race, Wages, Schooling and the Afqt Score," *The Review of Black Political Economy*, 24, 13–46, 10.1007/bf02690041.

Rosenblum, Alexis, William Darity, Angel L. Harris, and Tod G. Hamilton (2015) "Looking through the Shades," *Sociology of Race and Ethnicity*, 2, 87–105, 10.1177/2332649215600718.

Smith, Adam (1759) *The Theory of Moral Sentiments*: Scotland.

——— (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*: W. Strahan and T. Cadell, London.

Swinton, David (1977) "A Labor Force Competition Theory of Discrimination in the Labor Market," *American Economic Review*.

Swinton, David H (1974) *The Logic of Economic Discrimination in Non-Competitive Labor Markets* Ph.D. dissertation, Harvard University.

——— (1978) "A Labor Force Competition Model of Racial Discrimination in the Labor Market," *The Review of Black Political Economy*.

Tibshirani, Robert (2011) "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73, 273–282, 10.1111/j.1467-9868.2011.00771.x.

Urzua, Sergio (2008) "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices," *Journal of Human Resources*, 43, 919–971, 10.1353/jhr.2008.0011.

Webber, Douglas A. (2014) "The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors," *Labour Economics*, 28, 14–23, 10.1016/j.labeco.2014.03.009.

Williams, Jhacova and Valerie Wilson (2019) "Black Workers Endure Persistent Disparities in Employment Outcomes," https://www.epi.org/publication/labor-day-2019-racial-disparities-in-employment/.

Williams, Rhonda (1987) "Capital, Competition, and Discrimination: A Reconsideration of Racial Earnings Inequality," *Review of Radical Political Economics*.