# Discrimination Against Hispanic Workers and U.S. Economic Growth

Melissa Carleton

Submitted in partial fulfillment of the
Prerequisite for Honors in Economics



May 26, 2022

**Abstract**

In this paper I examine whether occupational discrimination against Hispanic workers has changed from the years 1970 to 2010 and whether changing patterns in ethnic discrimination can explain changes in U.S. economic growth. Using a structural model derived from a Roy model of occupational choice, I estimate the human capital and labor market barriers faced by Hispanic workers. Despite convergence in the occupational distribution of Hispanic workers toward white workers, I find little change in ethnic discrimination against Hispanic workers over the years in my sample. I find that much of the decrease in overall discrimination against Hispanic women is explained by gender, rather than ethnicity. Finally, I find that improved occupational sorting across the entire economy led to a 10% higher level of aggregate output, while improved occupational sorting among Hispanic workers has led to a 4% higher level of aggregate output.

**Acknowledgements**

I am incredibly grateful to my thesis advisor Prof. Kyung Park for his mentorship and feedback throughout the process of writing this thesis. I am also incredibly grateful to Prof. Akila Weerapana for his extensive feedback on drafts of this manuscript and thoughtful comments on my work. I would like to thank Zahra Thabet, Anushka Kelshikar, Cindy Zhao, and Tiffany Chu for their thoughtful feedback on my results throughout this process. I am also grateful for receiving financial support from the National Science Graduate Research Fellowship. I would like to thank Prof. Casey Rothschild, Prof. Daniel Sichel, and Prof. Olga Shurchkov for helpful discussions and their comments and feedback on my writing throughout this process. I would like to thank Prof. Helena de Bres for serving on my honors committee. I would like to thank all of the faculty above, along with Prof. Stanley Chang, for their consistent support in helping me prepare for Ph.D. studies and for their contributions to my intellectual growth as a student. Finally, I am grateful to my family for always supporting me as I pursue a career in Economics research.

# 1 Introduction

There is a growing body of literature showing evidence of persistent ethnic and racial discrimination over time. While much of this literature focuses on African American workers, there are relatively few studies of discrimination against Hispanic workers. While labor market barriers against both minority groups undoubtedly exist, there are a number of barriers that may be unique to Hispanic workers. For example, Hispanic workers have been the most likely to face threats of deportation due to discriminatory immigration policies or questions of citizenship. In light of these challenges, it is not obvious whether discrimination against Hispanic workers follows similar paths over time as discrimination against workers of other minority groups. Many studies point to earnings gaps to provide evidence of potential discrimination, though no known studies explicitly quantify discrimination.

Studying patterns of discrimination against Hispanic workers is called for since these workers represent a large and growing share of the labor force in the United States. The number of Hispanic workers in the labor force has grown from 10.7 million in 1990 to 29.0 million in 2020 and is projected to reach 35.9 million in 2030.[1] However, despite these numbers, Hispanic workers continue to face significant labor market disparities. Hispanic workers are disproportionately represented in lower-wage industries, such as transportation and warehousing, as well as leisure and hospitality. They also face a 7.4% unemployment rate, compared to a 5.2% unemployment rate of their white counterparts, despite being even more likely to actively look for work.[2] Such facts suggest that disparities in labor market outcomes between Hispanic and white workers could be due to systemic barriers rather than Hispanic workers actively choosing to work less. Accordingly, it is valuable to understand how discrimination, as a labor market barrier, against these workers has evolved over the past decades. These changes can profoundly influence Hispanic workers' occupational choices and resultant contributions to U.S. economic growth.

Some anecdotal evidence suggests that discrimination against Hispanic workers may have improved over time. Consider the 2009 appointment of chief Supreme Court justice Sonia Sotomayor, who is Hispanic. Sotomayor graduated summa cum laude from Princeton, attended Yale law school, trained under well-known judges, and won increasingly high-profile

[1] Dubina, Kevin, "Hispanics in the Labor Force: 5 Facts," U.S. Department of Labor Blog, Sept. 2021.

[2] Kimball, Spencer, and Rattner, Nate."Unemployment Rate Falls for Every Group but is Higher for Black and Hispanic Workers." CNBC, 6 Aug. 2021.

cases to achieve this rank. Although she faced challenges in her upbringing and racial discrimination at various stages in her career, she overcame these obstacles. Her story raises the question of whether equally talented Hispanic individuals born in earlier decades could have had access to the educational and professional opportunities necessary for achieving this credential. It also suggests that barriers to entry for Hispanic workers could have declined to some extent in recent decades.

However, examples like Sotomayor remain far, and few between, as ethnic minorities are underrepresented in the law profession and similar occupations. Hispanics, who are 18 percent of the U.S. population, still represent only 4 percent of workers in the legal profession, according to recent data. Such disparities are problematic since when any group is under represented in a particular occupation, it suggests they might not have the same opportunity available to pursue their talents. More representative groups also design fairer policies that benefit workers across the economy, which could ultimately generate positive and widespread effects on economic growth. In light of these statistics, it would be interesting to examine to what extent barriers to entry for Hispanic workers have either lowered or persisted across occupations and whether any changes have generated an increase in the productive capacity of the U.S. economy.

If structural barriers exist that prevent Hispanic workers from reaching their occupational potential, then several economic questions naturally arise. First, to what extent are differences in labor market disparities between white and Hispanic workers explained by discrimination rather than by a group choosing to work in an occupation because it enjoys working in that occupation? In this paper, I use the terminology "frictions" and "discrimination" somewhat interchangeably. "Frictions" refer to the structural parameter in my model that represents discrimination, while "discrimination" refers to the phenomenon I attempt to measure. Second, how have these frictions evolved over time? Finally, how have changes in the occupational distribution of Hispanic workers relative to white workers contributed to growth in U.S. aggregate output over the past decades since 1970?

Most existing literature on the labor market prospects of Hispanics or other minority groups in the U.S. highlights stylized facts and empirical trends in educational attainment, labor market participation, and earnings. Modeling these changing patterns of discrimination against Hispanic workers and the effect on aggregate output over time will require more than a descriptive analysis. Since discrimination is not directly observable, I use a structural

model to conduct counterfactual experiments to analyze the effect of falling discrimination on growth. In particular, I analyze the growth trajectory of the U.S. economy had discrimination remained at its levels in the earliest year in my sample and compare this presumably lower trajectory to the actual trajectory of economic growth.

I address this paper's research question through a structural model of occupational choice adapted from Hsieh et al. (2019). In this model, workers choose a market occupation or the home sector, taking into account their individual and group-specific tastes and preferences, earnings potential, along with the supply and demand-side frictions against members of their ethnic group or gender in a given occupation. Each occupation differs in its supply-side "tax", or firm-level employer discrimination, against members of a minority group. For example, a firm in the finance sector may prefer to hire a white worker rather than a slightly more talented Hispanic worker, while a firm in the agricultural sector could be less likely to discriminate in its hiring process.

The crux of my model is an equation that relates a composite measure of discrimination (occupational frictions) to the proportion of workers of a minority group in an occupation relative to a baseline group, controlling for other observable characteristics. This composite measure is a function of the propensity of a group to work in a given occupation, the wage gap in that occupation, and the returns to experience in that occupation. An increase in the propensity of a group to work in an occupation, an increase in the returns to experience of a group in that occupation, or a decrease in the wage gap in that occupation implies a decrease in occupational frictions. Since wage gaps control for group tastes and preferences in an occupation (how much a group "likes" an occupation), I ensure that my model does not attribute to frictions what, in reality, should be attributed to patterns in tastes and preferences.

I use this equation to estimate frictions against Hispanic workers using data from the U.S. decennial census from 1970 to 2010. My dataset contains information on individuals' education levels, occupation, Hispanic ethnicity, nativity status, and earnings. I then use these estimates to conduct a counterfactual experiment, determining how economic growth would have been different, had these frictions remained at their 1970 levels. If individual Hispanic workers experience frictions in an occupation, either due to labor market discrimination or to the lack of opportunity to accumulate necessary skills, they may not choose an occupation where they are most talented when discrimination in that occupation is present.

Aggregating these effects across workers can dampen aggregate output, as these occupations would have a lower pool of Hispanic talent to draw from. Conversely, declining barriers should have a positive effect on economic growth in the years of my sample, and I would expect aggregate output to increase if I find that frictions against Hispanic workers have declined.

In my framework, workers of a minority group face different human capital barriers to entry in different occupations. For example, a Hispanic worker who would otherwise succeed in the finance sector may not have access to funds or resources that would help them obtain the necessary skillset, so they enter a sector with lower barriers to entry. Put differently, discrimination taxes, or frictions, within an occupation artifically dampen the output produced by Hispanic workers since a discriminating firm treats a unit of Hispanic output as less profitable than a unit of output produced by a white worker. If discrimination taxes on the firm side are large enough, the firm might forego a large, talented pool of Hispanic workers in favor of a less productive pool of white workers. On the worker's side, discrimination taxes also increase the cost of obtaining the human capital required for success in that occupation. We can define the socially optimal outcome as the productivity-maximizing sorting of workers into occupations that results when firms and workers do not face taxes. From a theoretical perspective, large enough frictions can create distortions from the socially optimal outcome.

In my analysis, I focus on U.S.-born Hispanic workers and exclude immigrants from my sample. Several concerns arise when using cross-sectional data to analyze questions related to immigration. Using cross-sectional data is a valid approach to the extent that the distribution of innate talent in each occupation for the Hispanic immigrant population remains constant over time relative to the baseline group of white men. However, Hu (2000) has found that using cross-sectional data to study questions related to immigration may lead to bias. Observing different individuals at different points in time for a given "synthetic cohort" does not account for patterns of selective migration. Selective migration is not likely an important confounding factor for the populations that Hsieh et al. study since the shares of African-American workers in the United States have remained relatively constant. However, selective migration could likely confound using this type of approach to estimating changes in labor market discrimination against immigrants. For instance, suppose that the share of doctors in the Hispanic migrant pool becomes higher as more innately talented Hispanic

doctors decide to immigrate to the U.S. at a particular point in time. Then, even if discrimination remains the same, one would observe rising Hispanic immigrant shares in the U.S. labor market for doctors. Hsieh et al.'s methodology would incorrectly attribute this rising share to reductions in discrimination rather than to selective migration.

While adapting this framework to analyze the outcomes of immigrants could be a valuable direction for future research, my paper exclusively focuses on U.S.-born Hispanic workers. U.S.-born Hispanics accounted for nearly 90 percent of the growth in the Hispanic population in the United States between 2010 and 2017. The proportion of U.S.-born Hispanic workers to total Hispanic workers may have been smaller in the years of my sample which end in 2010. However, since this group is a substantial fraction of the Hispanic population today, any results I uncover may only be magnified in the future if the U.S.-born Hispanic population continues to grow and barriers continue to decrease. They can also shed light on questions of upward mobility since many U.S.-born Hispanics are children or grandchildren of immigrants.

My paper is organized as follows. I first present some stylized facts on occupational choices of Hispanic workers over time and I then present the general equilibrium Roy model of occupational choice that I adapt from Hsieh et al. (2019) and explain the differences in my setup. I then present my structural estimates of the model's key parameter that captures discrimination, showing that discrimination against both Hispanic men and Hispanic women due to ethnicity has changed little throughout the years in my sample.

Despite the occupational distribution of Hispanic women converging toward white men, I find that most of this convergence is due to changes in gender discrimination, rather than due to changes in ethnic discrimination. I also perform some robustness checks to assess the validity of my structural model. If my model aims to assess discrimination, then it should not attribute Hispanic workers leaving one occupation for a more profitable occupation to increases in discrimination in the first occupation. Indeed, I do not find an increase in discrimination for occupations that employed the majority of Hispanic workers in 1970. I also find that group tastes and preferences over occupations of Hispanic workers have changed little over time. I explain how this fact suggests that the differences in the occupational distribution of Hispanic workers relative to white workers are driven by discrimination, rather than by social norms of a given group, or by how much that group enjoys working in that occupation. Finally, I show that improved occupational sorting for all workers in the econ-

omy led to an approximately 10% higher output level than it would have, had discrimination remained entrenched at its levels in 1970. I find that changes in occupational sorting for Hispanic workers based on ethnicity alone led to an approximately 4% higher output level. However, I show that this growth reflects decreases in the variance of occupational discrimination across occupations rather than widespread improvements of the overall levels of occupational discrimination. The fact that decreasing gender discrimination has substantially improved growth in aggregate output should suggest that reductions in ethnic discrimination would lead the occupational distribution of Hispanic workers to further converge, increasing output to an even greater extent. I conclude by addressing the broader context of my model and its shortcomings, and I suggest valuable directions for future research.

## 2 Literature Review

This paper contributes to two existing areas of Economics research. First, this paper contributes to the literature on discrimination in a firm's hiring decision. Becker (1957) was the first researcher to introduce a supply-side "tax" or "taste for discrimination" against minority workers when making hiring decisions. This tax is modeled as a loss in utility that artificially lowers the productive value of these groups' output. Goldberg (1982), rather than assuming that workers have a taste for discrimination against minority workers, instead considers a model of "nepotism," or a preference for white workers that shows in the form of an increase in firm utility from hiring them. More recently, Charles and Guryan (2007) adopt the technique first developed by Becker to investigate whether a Becker-style "tax" affects minority workers' wages in the context of a general equilibrium model. Hsieh et al. (2019) also assume that firms have such a "tax" against minority workers, and also that workers themselves face a "tax" that makes it harder for them to accumulate skills in occupations. They examine patterns of occupational discrimination against women and black men from 1960 to 2010 and their effects on aggregate output through the prism of a Roy model of occupational choice.

Second, this paper's analysis of discrimination against Hispanic workers contributes to the general literature of discrimination against marginalized groups. To test hiring discrimination against black workers, Bertrand and Mullainathan (2004) conducted an audit study, sending resumes with black and white-sounding names to various firms. They find a lower callback rate for black applicants than their equally qualified white counterparts. Blau, Brummund, and Liu (2013) find that reductions in occupational segregation by gender over 1970-2009 could either be explained by reductions in discrimination (analogous to wage discrimination in my model) or better preparation (analogous to human capital in my model). Park (2017) uses an empirical approach to test for taste-based discrimination in judicial sentences against black convicts in Kansas. Though no significant taste-based discrimination is found, Park finds sentencing methods consistent with judges prosecuting black workers at higher rates than white workers who committed similar crimes.

Some papers address discrimination against Hispanics specifically, the population I focus on in my study. Doleac and Hansen (2016) find that banning the box, or removing the check-box on job applications that asks applicants to indicate whether they had a criminal record,

reduced the employment of low-skill young Hispanic men by 2.3 percentage points. Presumably, employers would assume these individuals are more likely to have a criminal record and discriminate on that basis. Yinger (1998) finds discrimination against Hispanics and other minority groups in the marketing of homes and cars. Fryer (2019) finds that police are more likely to use force against Hispanics and blacks in interactions. Lang and Spitzer (2020) find that police officers are more likely to statistically discriminate against Hispanics than whites. Kuka, Shenhav, and Shih (2020) find that the DACA act, a policy implemented in 2012 that grants Hispanic children who immigrated early in childhood, substantially improved Hispanic educational attainment. This study suggests that in the years in my sample, which are prior to implementation of DACA, Hispanic workers experienced barriers to accumulating the necessary human capital to succeed in a variety of occupations. Given the abundance of reduced-form evidence suggests discrimination against Hispanic workers, I should expect that my structural model will also predict that Hispanic workers experience discrimination in the labor market.

My paper will extend the work of Hsieh et al., who estimate the effect of reductions in discrimination against women and African American men on U.S. economic growth. Hsieh et al. use a Roy (1951) model of occupational choice to model workers' labor supply decisions and firms' labor demand. They find that between 20 and 40 percent of the growth in aggregate market output from 1960 to 2010 can be explained by falling discrimination against the groups they study. Given that Hispanic workers are becoming an increasingly large share of the U.S. labor force, it would not be surprising if any potential reductions in barriers found could explain a significant proportion of economic growth in recent decades. Moreover, the reductions in labor market frictions can promote the individual Hispanic worker to sort on her comparative advantage, rather than avoiding some occupations where she might, in earlier decades, have experienced discrimination. Aggregating this effect across individuals and firms can substantially increase U.S. economic growth, so I should expect to see similar patterns to Hsieh et al. (2019) in my study.

# 3 Stylized Facts

## 3.1 Data

I begin by describing my data source and presenting some stylized facts that motivate my question of whether declining barriers against Hispanic workers have led to convergence in the occupational distribution over time between these workers and white men.

I use data from the decennial Census from the years 1970 to 2010. I restrict my sample to four groups: white men, white women, Hispanic men, and Hispanic women, and I exclude workers who are identified as both white and Hispanic. This affects my results little since white and Hispanic individuals represent a tiny fraction of my sample. I include individuals who are of working-age, as defined by ages 25-54. Each individual is measured at one of three points in the life cycle. Individuals from 25-34 are classified as "young"; individuals from 35-44 are classified as "middle-aged"; and individuals from 45-54 are classified as "old age." I exclude individuals who are unemployed and searching for work, and I also exclude individuals on military duty. As discussed, I also exclude immigrants from my sample to avoid issues of selective migration, focusing exclusively on native-born workers.

My data is cross-sectional, so I create synthetic cohorts that are similar to panels in longitudinal data. For example, a certain cohort would be "young" in 1970, "middle-aged" in 1980, and "old age" in 1990. My data contains seven cohorts in total, the oldest being "old" in 1970, and the youngest being "young" in 2010. The cohorts who are young in 1970, 1980, or 1990 are able to be measured at all three points in the life cycle.

Individuals who worked less than 10 hours a week or who are not working and not actively searching for work are classified as in the "Home Sector". Individuals who worked between 10 and 30 hours a week are classified as "part time" and are weighted equally between the home and the sector they work in. Finally, individuals who work 30 hours a week or more are classified by one of 66 broad occupation categories listed in the Census occupational codes.

Earnings are given by the sum of labor, business, and farm incomes, and all earnings-related variables are calculated only with workers who work 48 weeks a year or more. I also convert all earnings data to chained real 2010 dollars.

Individual workers are indexed by $(i, g, c)$, i.e. an individual who is part of group $g$ (one of the four ethnic and gender groups) works in an occupation $i$ (one of the 66 market occupations) and belongs to cohort $c$ (one of the three ages in the life cycle). When calculating

composite frictions against workers in groups, averages are taken across occupations and groups.

## 3.2 Wage Gaps

To motivate my question of whether barriers against Hispanic workers have declined over time, I first present data on the earnings gaps between Hispanic men, Hispanic women, and white women, relative to white men. Disparities in wages are often used as reduced-form evidence of employer discrimination. Thus, differences in earnings profiles over time can motivate the question of whether occupational discrimination has declined. Figure 1 displays the average of the natural logarithm of the average wage gap between Hispanic men and white men across the 66 market occupations; Figure 2 displays this relationship between white women and white men; and Figure 3 displays this relationship between Hispanic women and white men. All three figures display the wage gap by cohort in a given year. As explained earlier, cohorts 0, 1, and 2 consist of workers that are ages 25-34, 35-44, and 45-54, respectively. Wage gaps relative to white men are shown for each cohort throughout time. A value of zero for the natural logarithm of the wage gap indicates that workers experience earnings parity with white men, on average. Upward convergence toward zero indicates declines in earnings inequality.

Figure 1 displays an overall upward trend in the life cycle profiles of wage gaps across cohorts. This trend indicates that the wage gaps of Hispanic men have steadily improved over time. In most years, young Hispanic men entering the labor force saw improved earnings prospects relative to white men in their cohort than older Hispanic men. For example, in the year 1990, the red line (old age) is below the purple line (middle age), which is below the green line (young age). Moreover, for most decades with the exception of 1970, within cohorts they have worsened at the middle point in the life cycle, and returned to their original level at the last life cycle point.

Figures 2 and 3 show declining wage disparities for both white women and Hispanic women. For white women, earnings disparities have worsened from the young to middle age points in the life cycle to return to their original levels in old age. For Hispanic women, patterns have not followed as consistent of a pattern throughout the life cycle. For example,

in 1980, young Hispanic women experienced greater earnings parity relative to white men in their cohort than older Hispanic women did in their cohort. However, in 2000, older Hispanic women experienced greater earnings parity relative to their young white male counterparts than middle-aged Hispanic women to their counterparts. Unlike white women, successive generations of Hispanic women occasionally experienced greater earnings disparities than the previous generations.

The patterns in wage gaps for Hispanic workers suggest that despite overall increases in earnings parity across cohorts, the evolution of barriers against them is not steady over time. Unlike white women, earnings disparities for Hispanic workers have not always reached their most pronounced point in middle age or substantially improved in old age. Furthermore, and perhaps most importantly in the context of this paper's focus on ethnic discrimination, Hispanic wage gaps have not always declined for the incoming younger generations at single points in time. If Hispanic wage gaps partly reflect ethnic discrimination, the patterns discussed should suggest that barriers against Hispanic workers have likely not showed a steady and consistent decline over time. The results presented in the Empirical section of this paper confirm this hypothesis.

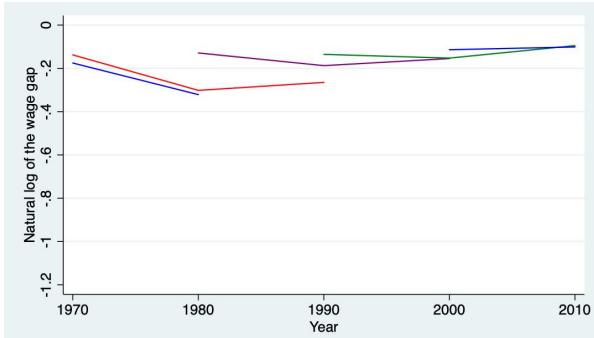



Figure 1: Wage Gaps (Hispanic Men)

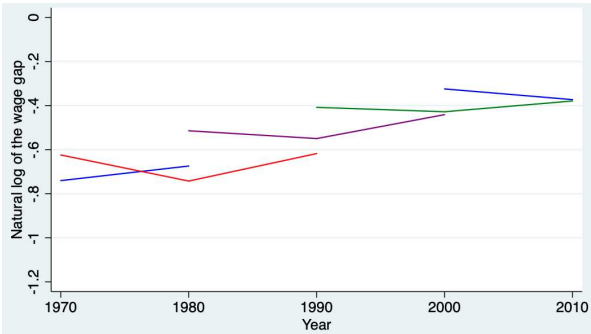



Figure 2: Wage Gaps (White Women)

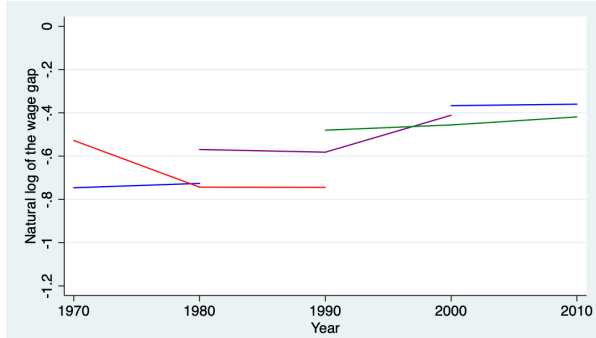



Figure 3: Wage Gaps (Hispanic Women)

These figures show the natural logarithm of wage gaps for a group by cohort relative to white men, averaging across all occupations.

## 3.3 Occupational Mobility

The relative propensity of a group $g$ to work in occupation $i$, relative to white men, is given by $\hat{p} = \frac{p_{ig}}{p_{i,wm}}$, where $p_{ig}$ is the proportion of the number of the group in the occupation and $p_{i,wm}$ is the proportion of the number of white men working in the occupation. For example, if 3% of Hispanic women work in occupation A and 5% of white men work in occupation A, then $\frac{p_{ig}}{p_{i,wm}} = .03/.05 = 0.6$. The standard deviation of $\hat{p}$, then, represents the dispersion of the propensities. If the relative propensities are roughly the same for each occupation, then all occupations in the market will be more representative of the general population. For example, if 1 out of 20 workers in group $g$ work in occupation $i$ and 1 out of 20 white men work in occupation $i$, then $\hat{p} = 1$. If $\hat{p}$ is close to 1 for all occupations, then there will be less variance in $\hat{p}$. That is, a high dispersion of $\hat{p}$ would indicate that a group

is overrepresented in some occupations and underrepresented in others.

Figure 4 plots the earnings-weighted natural logarithm of the standard deviation of $\hat{p}$ across occupations relative to white men. A value of zero for the standard deviation suggests that the likelihood that a group works in a given occupation, relative to white men, is equal across occupations. For white women, the standard deviation of $\hat{p}$ was slightly below 2 in 1970 and was about 1.5 in 2010. For Hispanic men, the standard deviation of $\hat{p}$ was nearly 1 in 1970 but dropped to .5 in 2010. For Hispanic women, the standard deviation of $\hat{p}$ was slightly above 1.5 in 1970, increased slightly in 1980, and decreased onward to reach a value below 1.5 in 2010. With the exception of the year 1970, the dispersion in the propensity of Hispanic women has closely tracked that of white women over time, suggesting that much of the improved occupational sorting of Hispanic women relative to white men is due to gender, rather than ethnicity.

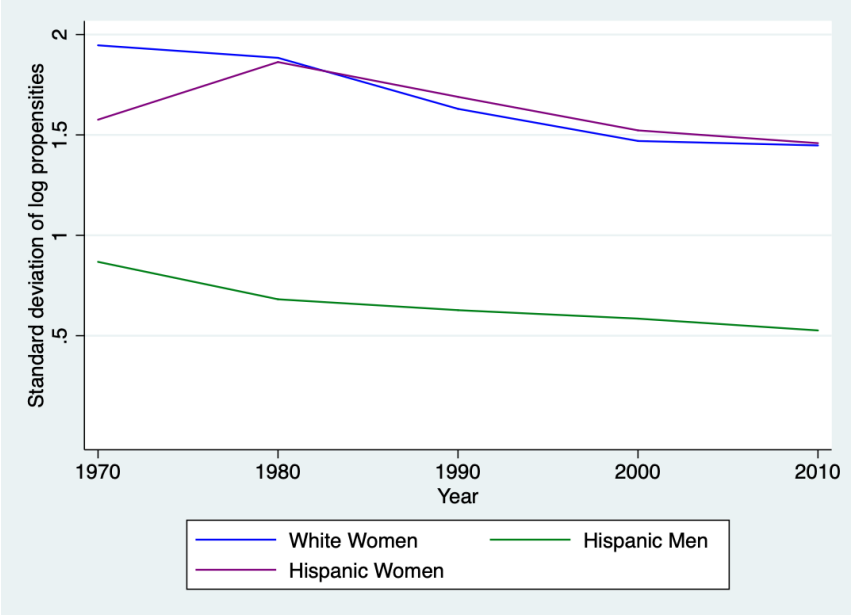



Figure 4: Convergence of Occupational Propensities

This figure displays the earnings-weighted log of the standard deviation of the propensities of a group to work in an occupation over time.

The tables below illustrate the five occupations that saw the five highest and five lowest propensities of workers of a given group relative to their white counterparts in the years 1970 and 2010. Although convergence toward the occupational distribution of white men over time suggests declines in overall barriers, I must analyze the propensities of Hispanic women relative to white women to assess the affect of ethnicity that is not confounded by gender.

The ratios in the tables can be interpreted as follows. If $\hat{p}_{ig} = .05$, then a white man is $1/.05 = 20$ times more likely than a member of that group to work in a given occupation. The table also displays the 2010 position of the highest and lowest position occupations in 1970 to show whether their ranks changed over time, as well as the 1970 position of the highest and lowest position occupations in 2010 to show the differences between the current position and the previous position.

The patterns in the tables below indicate shifts in the relative odds of the propensities over time in the tail ends of the occupational distribution. For example, Hispanic women were much more likely to work as guards than white women, both in 1970 and 2010. In the occupations where they were under-represented in 2010, they were so under-represented in 1970 that there is not sufficient data to find the relative rank of those occupations. However, the numbers in the "2010 position" column indicate that for most occupations where Hispanic women were over-represented in 1970, they were no longer over-represented in 2010. Similarly, in many the occupations where Hispanic women were over-represented in 2010, they were so under-represented in these occupations in 1970 that there was not sufficient data to determine the relative rank of that occupation at that time. Similarly, in the occupations where Hispanic women were most under-represented in 1970, they were no longer so under-represented in 2010. The tables suggest that Hispanic men were over-represented in brawn-heavy occupations and underrepresented in teaching and professional-type occupations such as law in both 1970 and 2010. However, the extent to which they were over-represented in the top five propensity occupations has declined dramatically, suggesting convergence in the occupational distribution. For example, the highest propensity occupation in 1970 ("Private Household Occupations") had a propensity of 17.04, and the highest propensity occupation in 2010 ("Communications") had a propensity of 2.678. These patterns suggest that the occupational distribution of Hispanic workers relative to white workers have not remained stagnant and motivates the question of whether occupational barriers have changed over time.

Table 1: Hispanic Men to White Men Propensities

| rank in 1970 | occupation (1970) | $\hat{p}_{ig}$ (1970) | 2010 position | rank in 2010 | occupation (2010) | $\hat{p}_{ig}$ (2010) | 1970 position |
|---|---|---|---|---|---|---|---|
| lowest | Farm Operators | .126 | lowest | lowest | Farm Operators | .168 | lowest |
| 2 | Teachers | .153 | $5^{th}$ lowest | 2 | Natural Scientists | .253 | no data |
| 3 | Engineering | .200 | $6^{th}$ lowest | 3 | Health Diagnosing | .279 | $8^{th}$ lowest |
| 4 | Management | .297 | $14^{th}$ lowest | 4 | Lawyers and Judges | .301 | no data |
| $5^{th}$ lowest | Sales | .307 | $21^{st}$ lowest | $5^{th}$ lowest | Teachers | .308 | no data |
| highest | Private Household | 17.04 | $4^{th}$ highest | highest | Communications | 2.678 | no data |
| 2 | Office Machine | 6.882 | $3^{rd}$ highest | 2 | Guards | 2.554 | $15^{th}$ highest |
| 3 | Health Technologists | 5.889 | $26^{th}$ highest | 3 | Personal Services | 1.978 | $33^{rd}$ highest |
| 4 | Precision Workers | 4.875 | $10^{th}$ highest | 4 | Private Household | 1.976 | highest |
| $5^{th}$ highest | Cleaning and Building | 3.852 | $6^{th}$ highest | $5^{th}$ highest | Health Service | 1.944 | no data |

The top half of this table lists the five occupations that Hispanic men were least likely to select in the years 1970 and 2010 and the odds of Hispanic men, relative to white men, of working in that occupation. The fourth column indicates the position in 2010 of the occupations that had the lowest propensities in 1970. The last column indicates the position in 1970 of the occupations that had the lowest propensities in 2010. The bottom half of this table displays all of the above for the five occupations that Hispanic men were most likely to work in, relative to white men, in the years 1970 and 2010.

Table 2: Hispanic Women to White Women Propensities

| rank in 1970 | occupation (1970) | $\hat{p}_{ig}$ (1970) | 2010 position | rank in 2010 | occupation (2010) | $\hat{p}_{ig}$ (2010) | 1970 position |
|---|---|---|---|---|---|---|---|
| lowest | Sales Occupations | .209 | $36^{th}$ lowest | lowest | Farm Operators | .434 | no data |
| 2 | Financial Records | .211 | $40^{th}$ lowest | 2 | Librarians | .211 | no data |
| 3 | Production Inspectors | .434 | $48^{th}$ lowest | 3 | Health Diagnosing | .209 | no data |
| 4 | Teachers | .539 | $11^{th}$ lowest | 4 | Health Assessment | .568 | $6^{th}$ lowest |
| $5^{th}$ lowest | Secretaries | .568 | $37^{th}$ lowest | $5^{th}$ lowest | Precision Woodworking | .539 | no data |
| highest | Construction | 10.53 | $22^{nd}$ highest | highest | Communications Equipment | 2.09 | $12^{th}$ highest |
| 2 | Guards | 8.369 | $3^{rd}$ highest | 2 | Information Clerks | 2.048 | $27^{th}$ highest |
| 3 | Food Production | 6.367 | $8^{th}$ highest | 3 | Guards | 2.018 | $39^{th}$ highest |
| 4 | Urban Planners | 5.772 | $61^{st}$ highest | 4 | Private Household | 1.97 | $45^{th}$ highest |
| $5^{th}$ highest | Private Household | 5.316 | $4^{th}$ highest | $5^{th}$ highest | Police and Detectives | 1.929 | no data |

The top half of this table lists the five occupations that Hispanic women were least likely to select in the years 1970 and 2010 and the odds of Hispanic women, relative to white women, of working in that occupation. The fourth column indicates the position in 2010 of the occupations that had the lowest propensities in 1970. The last column indicates the position in 1970 of the occupations that had the lowest propensities in 2010. The bottom half of this table displays all of the above for the five occupations that Hispanic women were most likely to work in, relative to white women, in the years 1970 and 2010.

Under the assumption that innate talent for each occupation is the same across all groups, the decrease in the standard deviation of $\hat{p}$ across time for white women, Hispanic women, and Hispanic men, as shown in Figure 4, indicates that workers in these groups have better sorted to their comparative advantage over time. However, it is not clear to what extent these patterns are driven by declining barriers versus other factors that affect occupational choice. In my following analysis, I examine this question in detail.

# 4 Model

I first introduce the setup of the theoretical model of Hsieh et al. (2019). The model contains numerous details, so I focus on the parts that are most relevant to address my research question of whether declining barriers for Hispanic workers in various occupations has led to greater occupational convergence over time. The full model can be found in the Online Appendix of Hsieh et al.'s paper.

The model consists of a continuum of individuals who belong to an ethnic and gender group $g$ and cohort $c$. The workers' problem is to choose an occupation $i$, a choice that will remain fixed over the lifetime. There are $M$ occupations in total, including the home sector. First, workers are born, and they must choose a profession. Before making this choice, they observe both their intrinsic ability $\epsilon_i$ in each profession, and the frictions $\tau_{ig}^w$ and $\tau_{ig}^h$ that their group will face when working in or preparing for work in that occupation. Having chosen a profession, they must make a decision about how much to invest in schooling ($s_i^*$) and how much to invest in human capital $e_{ig}^*$. They then start working, and they use their earnings for consumption over their lifetime. Earnings can evolve over the lifetime due to changing patterns in employer discrimination or differences in returns to experience across the life cycle. However, their level of investment in human capital is determined before workers begin work, and it remains fixed throughout the lifetime. Workers spend the entirety of their earnings, net of expenditures on schooling, on consumption. Utility is increasing in consumption and decreasing in the amount required for education spending $e_{ig}^*$ and the time spent in school $s_i^*$. The individuals choose an occupation to maximize their expected lifetime discounted present values of utility.

The workers' choice problem begins as follows. First, nature draws workers' idiosyncratic talent $\epsilon_i$ for each occupation. A high draw of $\epsilon_i$ indicates a higher level of talent in a given occupation. Workers' talent draw in each occupation is drawn from a multivariate Frechet distribution, and these draws are independently identically distributed across occupations:

$$F_g(\epsilon_1, \cdots, \epsilon_M) = \exp[-\sum_{i=1}^{M} \epsilon_i^{-\theta}],$$

where the parameter $\theta$ governs the dispersion of talent, and a higher value of $\theta$ indicates a smaller dispersion.

The frictions that the workers will face in each occupation $\tau_{ig}^h$ and $\tau_{ig}^w$ are determined exogenously, and these parameters are occupation specific. The parameter $\tau_{ig}^w$ captures employer discrimination against workers of a minority group. The parameter $\tau_{ig}^h$ captures human capital differences across groups and represents all differences in childhood environments that later affect human capital accumulation. Such barriers include access to quality schooling. They also include differences in access to skill-building in different occupations.

Before choosing an occupation, workers observe their individual-specific draws of $\epsilon_i$ in each occupation and the group-specific values of $\tau_{ig}^w$ and $\tau_{ig}^h$ in each occupation before beginning the working life-cycle. Given these observed values, individuals (in the pre-period) then choose the optimal levels of schooling $s_i^*$ and investment in human capital $e_{ig}^*$ to maximize lifetime utility. Workers compute these hypothetical values of $s_i^*$ and $e_{ig}^*$ for *each* of the $M$ occupations. That is, for each occupation they decide what their levels of schooling and human capital investment would be *if* they chose that occupation.

The lifetime utility of an individual worker from group $g$ and cohort $c$ who chooses occupation $i$ is given by

$$\log U_i = \beta[\sum_{t=c}^{c+2} \log C_i(c,t)] + \log[1 - s_i(c)] + \log z_{ig}(c), \tag{1}$$

subject to the constraints

$$C_i(c,t) = [1 - \tau_{ig}^w] w_i(t) \epsilon_i h_{ig}(c,t) - e_{ig}(c,t)[1 + \tau_{ig}^h(c)], \tag{2}$$

and

$$h_{ig}(c,t) = \gamma(t-c) s_i(c)^{\phi_i} e_{ig}(c)^{\eta}. \tag{3}$$

Here, $C_i(c,t)$ is the per-period consumption of cohort $c$ in year $t$, and utility in occupation $i$ depends on consumption over all three periods of the life cycle. Since utility is increasing in consumption, workers spend all of their net earnings minus the expenditures on education on consumption. Consumption is influenced by the choice variables of $e_{ig}^*$ and the following parameters that affect net earnings and education spending. Net earnings in an occupation are the product of a person's efficiency wage, or marginal revenue product $w_i(t)$, their draw of $\epsilon_i$, and their human capital $h_{ig}$. Net earnings are deflated by a factor of $[1 - \tau_{ig}^w]$, which

affects all cohorts within a group $g$ equally at a given point in time. Individuals must borrow $e(c)[1+\tau_{ig}^h(c)]$ in the first period to purchase $e(c)$ units of human capital, a loan that is repaid over the lifetime subject to the lifetime budget constraint $e(c) = \sum_{t=c}^{c+2} e(c,t)$.

The parameter $\beta$ captures the tradeoff between lifetime consumption and time spent accumulating human capital. The common utility benefit $z_{ig}(c)$ is the group-specific tastes and preferences for a given occupation. Social norms could lead members of a group to derive different utility levels from working in different occupations. For example, if a group dislikes construction occupations, then we would expect to see a low value of $z_{ig}(c)$ in the construction occupation. The value $\gamma(t-c)$ parameterizes the returns to experience across occupations at a given point in the life cycle. A higher value of $\gamma(t-c)$ has a positive effect on human capital accumulation $h_{ig}$. The parameter $\phi_i$ is occupation-specific, and is the return to time investments in human capital. The parameter $\eta$ is the elasticity of human capital with respect to education spending. A higher value of $\eta$ indicates that education spending has a larger positive effect on human capital accumulation.

The optimal solutions to the worker's problem for each occupation $i$ are given by

$$s_i^* = \frac{1}{1 + \frac{1-\eta}{3\beta\phi_i}},$$

$$e_{ig}^* = \left( \frac{\eta(1 - \tau_{ig}^w w_i \overline{\gamma} s_i^{\phi_i} \epsilon_i}{1 + \tau_{ig}^h} \right)^{\frac{1}{1-\eta}},$$

where $\overline{\gamma} \equiv 1 + \gamma(1) + \gamma(2)$ is the sum of the returns to experience over the life cycle. In the first period of the life cycle, $\gamma(0)$ is normalized to 1. Time spent accumulating human capital increases in $\phi_i$ since individuals in higher $\phi_i$ occupations earn higher wages as compensation for time spent on schooling and thus acquire more schooling in these occupations. These optimal solutions illustrate why both time and human capital investment are needed in this model. Human capital investment ensures that distortions to human capital accumulation matter, while time is needed to explain average wage differences across occupations.

Given these optimal solutions for time and goods, workers then calculate their indirect utility for each occupation $i$, which may be written as

$$U_{ig}^* = [\overline{\gamma}_{ig} w_i s_i^{\phi_i} (1 - s_i)^{\frac{1-\eta}{3\beta}} \tilde{z}_{ig} \epsilon_i / \tau_{ig}]^{\frac{3\beta}{1-\eta}}, \tag{4}$$

where $\tau_{ig} = \frac{(1+\tau_{ig}^h)^\eta}{1-\tau_{ig}^w}$ and $\tilde{z}_{ig} \equiv z_{ig}^{\frac{1-\eta}{3\beta}}$. The parameter $\tau_{ig}$ represents composite occupational frictions and is the cumulative effect of $\tau_{ig}^w$ and $\tau_{ig}^h$.

Given their hypothetical optimal choices of $s_i^*$ and $e_{ig}^*$ for each occupation, individuals when they are born compare their hypothetical indirect utilities across *all* of the $M$ occupations, and they choose the occupation that delivers the highest indirect utility. The values of $s_i^*$ and $e_{ig}^*$ that are realized are those in the occupation that the worker has chosen. That is, given their optimal occupational choice, individuals spends $s_i^*$ years in school and invest $e_{ig}^*$ in human capital needed for that occupation before beginning work. Finally, individuals start working in their optimally-chosen occupation when young and continue working in that occupation throughout the life cycle. This occupation is the workers' choice that is realized in the data.

# 5 Empirical Results

## 5.1 Empirical Strategy

In this section, I outline how I estimate the composite frictions against Hispanic workers in each occupation. Workers' occupational choices are observed in the data, and using the probability of workers to choose an occupation, I outline my procedure for analyzing whether discrimination against women and Hispanic workers has declined over time. I then estimate to what extent this decline was due to declines in ethnic discrimination, rather than gender discrimination, in the case of Hispanic women. In the context of my research question, "improved occupational sorting" indicates that a group's occupational distribution has converged toward that of white men. If occupational barriers against Hispanic women relative to white women remain entrenched, then we would observe no decline in barriers against Hispanic women relative to white women, even if their occupational sorting has converged toward white men. Similarly, analyzing the barriers against Hispanic men relative to white men provides insight into whether ethnic discrimination has decreased.

I first estimate occupational frictions against each group relative to white men, which will then allow me to calculate frictions against white women relative to Hispanic women. I follow the empirical model in Hsieh et al. (2019) to derive a similar model used for estimating discrimination against Hispanic workers. Equation 5 gives a closed-form solution for occupational barriers and shows that frictions are related to three terms: the propensity of a group to work in a given occupation, the wage gap in that occupation, and the returns to experience in that occupation. The composite frictions against a group $g$ in an occupation $i$ may be written as

$$\tau_{ig} = \hat{p}_{ig}^{-\frac{1}{\theta}} \cdot \widehat{\overline{\text{wage}}}_{ig}^{-(1-\eta)} \cdot \hat{\overline{\gamma}}_{ig}^{\eta}, \tag{5}$$

where a hat denotes the value relative to white men, $p_{ig}$ is the propensity of a group to work in a given occupation, $\widehat{\overline{\text{wage}}}_{ig}$ is the geometric average of earnings weighted equally across cohorts for a group in a given occupation, $\overline{\gamma}_{ig}$ is the sum of the returns to experience over the life cycle for a group in a given occupation, $\eta$ is the elasticity of human capital with respect to education spending, and $\theta$ is the dispersion of skills in a given occupation.

In light of the above equation and the stylized facts presented earlier on the convergence

of white women and Hispanic men's propensities converging toward those of white men over time, the composite frictions $\tau$ for these groups should have decreased over time, relative to white men. The value of $\tau$ is inversely proportional to propensities, so workers could be better sorting to their comparative advantage due to declines in these barriers.

With data on occupational shares, geometric earnings, and talent in a given occupation, Equation 5 says that we can estimate the composite effect of human capital and labor market discrimination. There could be concern that the model does not control for differences in group tastes and preferences in a given occupation. What the model interprets as high occupational frictions could be instead represent a group's low tastes and preferences for working in a given occupation that are unrelated to discrimination. However, this is not the case. As shown in Hsieh et al. (2019), the wage gap in an occupation itself is a function of the frictions against a group in a given occupation, the average talent of a group in that occupation, and the cost of skills in all occupations. Therefore, the wage gap already controls for differences in group tastes and preferences in an occupation.

In the following section, I outline how I infer each of the five parameters that enter this equation to eventually estimate $\tau_{ig}$. Table 3 summarizes each relevant parameter and its empirical target. The procedure for estimating $\theta$ and $\eta$ is described in the Appendix.

Table 3: Estimating Tau Empirical Parameters and Target

| Parameter | Definition | Empirical Target |
|---|---|---|
| $p_{ig}(c)$ | occupational propensity | probability to work in an occupation |
| $\overline{\text{wage}}_{ig}$ | geometric earnings | data on arithmetic earnings adjusted |
| $\overline{\gamma}_{ig}$ | returns to experience | function of change in wage gaps |
| $\eta$ | elasticity of human capital to education spending | unexplained differences in wages by $(i, g, c)$ |
| $\theta$ | dispersion of talent in an occupation | unexplained differences in wages by $(i, g, c)$ |

## 5.2 The Returns to Experience

I first develop a procedure for estimating the returns to experience $\gamma(t - c)$ for a group. I derive a closed-form solution for $\gamma$ as a function of wage gaps and other parameters that a researcher can easily program using maximum likelihood methods with data on wages. In the Appendix, I outline how I make this methodological contribution to Hsieh et al.'s empirical strategy. While the authors program a model by guessing an initial value for the

parameter $\gamma$, my solution allows for a direct estimation of $\gamma$.

I estimate the value $\gamma(t-c)$ for each (group, cohort) pair, averaging across occupations. I assume that $\gamma(t-c)$ does not vary across occupations for a given group and cohort. This is a reasonable assumption since the values by occupation for the years where $\gamma(t-c)$ could be computed varied little. However, I do allow $\gamma(t-c)$ to vary across groups and cohorts. The sum of the returns to experience over the life cycle is given by $\overline{\gamma} = 1 + \gamma(1) + \gamma(2)$, where $\gamma(0)$, the return to experience for the young cohort is normalized to 1. I obtain $\gamma(1)$ by comparing the young cohort of a group at time $t = 0$ to the middle-age cohort of a group during the following life cycle point ten years later at $t = 1$. I obtain $\gamma(2)$ by comparing the middle-age cohort of that group at time $t = 1$ to the old-age cohort of that group at time $t = 2$, ten years later still.

I make use of the equations given in the Appendix of Hsieh et al. (2019) to write a simple equation for calculating the value of $\gamma$ for each group that can be estimated using maximum likelihood methods. As given in the Appendix of the original paper, the returns to experience $\gamma(t-c)$ can be estimated from the equation

$$\frac{\overline{wage}_{i,wm}(c,t)}{\overline{wage}_{i,wm}(c,c)} = \frac{w_i(t)(\gamma_{ig}(t-c))s_i^{\phi_i(c)}}{w_i(c)s_i^{\phi_i(c)}}. \tag{6}$$

The above equation says that the returns to experience is positively related to the wage gaps across successive years. Intuitively, if average earnings increases from one year to the next, this should reflect a higher value of experience in a given occupation. Given data on wage gaps and schooling, I first find the value of the efficiency wage $w_i(t)$ to fit this equation, which depends on other parameters in the model. After computing an expression for $w_i(t)$, I then use a nonlinear least squares procedure to find the value of $\gamma$ that fits this equation.

Table 4 displays the values of $\gamma(t-c)$ and $\overline{\gamma}$ for each group in my data. For white men, I find that the value rises slightly in middle age and falls slightly in old age. For white women, the value follows a very similar pattern. However, because $\overline{\gamma}$ is relatively similar to white men for all groups, the ratio of $\overline{\gamma}$ for that group to that of white men is very close to 1. When viewed through the lens of 5 this means that differences in the returns to experience are not consequential in explaining patterns of $\tau$ over time. The Appendix reports results when $\gamma$ is allowed to vary by both cohort and group for years which it can be estimated.

However, I find those values to be not much different from those displayed in Table 4, so taking averages across cohorts should not introduce substantial error into my results.

Table 4: Estimates of Gamma

| Group | $\gamma(0)$ | $\gamma(1)$ | $\gamma(2)$ | $\overline{\gamma} = 1 + \gamma(1) + \gamma(2)$ | $\hat{\overline{\gamma}} = \overline{\gamma}_g / \overline{\gamma}_{wm}$ |
|---|---|---|---|---|---|
| white men | 1 | 1.126 | .815 | 2.941 | 1 |
| white women | 1 | 1.140 | .845 | 2.985 | 1.015 |
| Hispanic men | 1 | .868 | 1.120 | 2.988 | 1.016 |
| Hispanic women | 1 | 1.096 | .816 | 2.912 | .990 |

## 5.3 Occupational Preferences

Although the occupational distribution of Hispanic workers has converged toward that of white men over time, this convergence does not necessarily suggest declines in discrimination against Hispanic workers. Occupational propensities are related to both group preferences over occupations and occupational frictions. Thus converging propensities could potentially be explained by changes in group preferences, changes in frictions, or both. Therefore, to evaluate the relative impact of frictions on occupational sorting, I must also examine differences in group preferences $\tilde{z}_{ig}$ over time.

In this section, I analyze patterns in occupational preferences $\tilde{z}_{ig}$ over time. The equation used for estimating $\tilde{z_{ig}}$ is given in the Appendix where I outline the procedure for estimating the returns to experience $\gamma(t-c)$. Since the return to market work $m_g$ is a function of $\tilde{z_{ig}}$ and $\gamma(t-c)$, I can use the estimates of $\gamma(t-c)$ and the estimates of $m_g$ to back out $\tilde{z_{ig}}$ for each group. The value of $\tilde{z_{ig}}$ is unitless, although a higher value indicates that a group derives a higher utility benefit from working in that occupation, relative to white men. The value of $\tilde{z_{ig}}$ is normalized to 1 for white men so that a value of $\tilde{z_{ig}}$ equal to 1 indicates that a group derives the same utility benefit from working in that occupation as white men.

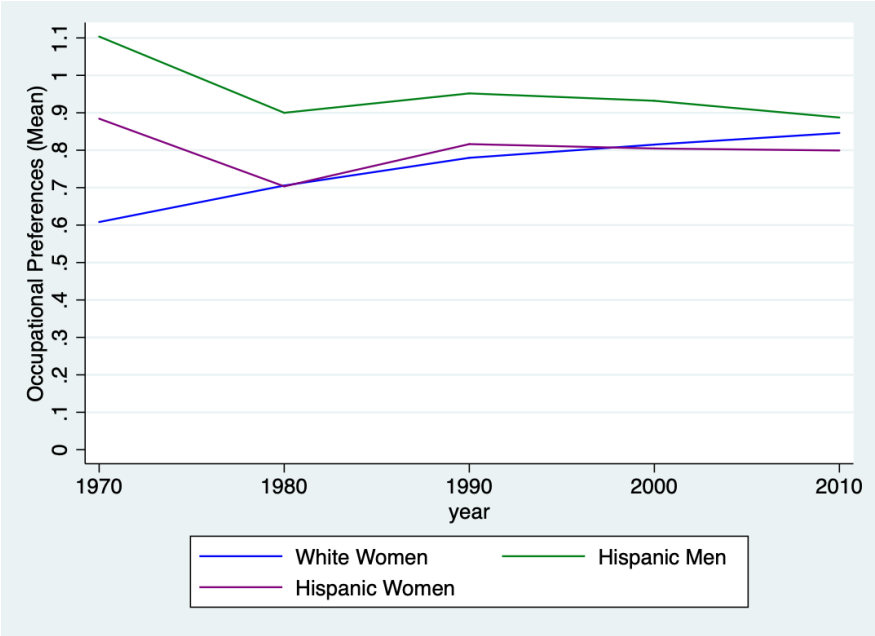



Figure 5: Group Preferences $\tilde{z}_{ig}$

Earnings weighted mean $\tilde{z}_{ig}$ of group preferences across occupations relative to white men.

The green line in Figure 5 indicates that, on average, occupational preferences of Hispanic men have been roughly similar to those of white men over time, with the exception of the years 1970 and 2010. In 1970, the value of $\tilde{z}_{ig}$ was nearly 1.1, suggesting that on average, Hispanic workers preferred to work in certain occupations slightly more than white men. In the years following, the value was very close to 1, indicating that Hispanic men had roughly similar preferences over occupations as white men. In 2010, the value was roughly 0.9, indicating that on average, Hispanic men preferred certain occupations slightly less than white men. These patterns suggest that some of the convergence in the occupational distribution of Hispanic men towards that of white men over time is due to changes in preferences of Hispanic men over time.

Figure 5 also indicates that from 1980 onward, the patterns of group tastes for Hispanic women have closely tracked those of white women, with the exception of 1970. In 1970, the value of $\tilde{z}_{ig}$ for Hispanic women was closer to 1 than the value for white women, suggesting that Hispanic women had relatively similar preferences over occupations as white men, while

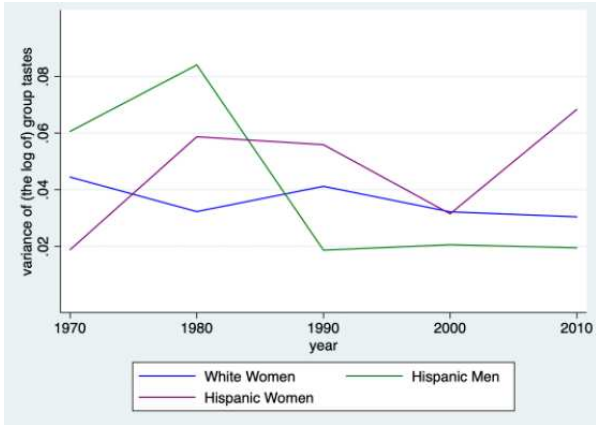



Figure 6: Preferences

Variance of occupational preferences $\tilde{z}_{ig}$. The variance across occupations has remained low over time for each group.

white women disliked certain occupations more than white men. For white women, the value of $\tilde{z}_{ig}$ has converged toward 1, suggesting that the occupational preferences of white women have converged toward those of white men. From 1980 onward, the proximity of the purple line to the blue line indicate that occupational preferences of Hispanic women have closely tracked those of white women. The similar patterns in preferences for these two groups suggest that convergence in the occupational distribution of Hispanic women toward white men was driven by gender rather than Hispanic origin. Thus, Hispanic ethnicity cannot explain differences in the occupational distribution between Hispanic women and white women.

Furthermore, Figure 6 indicates that variances in occupational preferences have not followed a clear pattern for any group. The important point is that the variance has always been below 0.1, which is considered a low value for the variance of this variable in my model since the baseline value for $\tilde{z}_{ig}$ is 1. This pattern indicates small confidence intervals for Figure 5 and that preferences across occupations within a group are similar at a single point in time.

## 5.4 Estimating the Frictions

I now turn to analyzing this paper's central question of whether occupational barriers against Hispanic workers have declined over time. With data on relative propensities, wage gaps, $\theta$, $\eta$, and $\overline{\gamma}$, I next estimate the composite frictions given by Equation 5. Figure 7 shows the earnings-weighted mean of composite occupational frictions $\tau_{ig}$ for each group. The value of $\tau_{ig}$ for white men is normalized to 1 so that $\tau_{ig}$ can be interpreted as occupational barriers relative to white men. This parameter is unitless, although a higher value indicates that a group experienced greater barriers relative to white men. A value of $\tau_{ig}$ equal to 1 indicates that a group did not experience any discrimination relative to white men.

The green line in Figure 7 indicates that for Hispanic men, composite occupational frictions have remained relatively stagnant across time and persistently above those of white men. They began slightly above 2 in 1970 and ended slightly below 2 in 2010, suggesting that little change in discrimination over time occurred. In one way, this result is not surprising since Hsieh et al. (2019) find that barriers have persisted for black men.

Figure 7 indicates that for Hispanic women, overall occupational frictions, which represent both ethnic and gender discrimination, have declined since 1970. The patterns of the lines in Figure 7 show that frictions against Hispanic women began below those against white women in 1970 but were above those against white women in 1980, and remained above white women thereafter. These patterns motivate the examination of frictions against Hispanic women relative to white women. Figure 8 displays these frictions and is consistent with the pattern shown in Figure 8. As the figure indicates, frictions against Hispanic women relative to white women were below .8 in 1970, suggesting that Hispanic women experienced fewer occupational barriers relative to white women in 1970. However, frictions were above 1 in 1980 and mostly increased thereafter, only decreasing slightly in 2010, suggesting that Hispanic women experienced more occupational barriers than white women for the majority of the years in my sample. Taken together, the patterns of frictions against both Hispanic men and Hispanic women relative to their white counterparts suggests that ethnic discrimination remains entrenched for both of these groups. Although occupational barriers against Hispanic women have converged toward that of white men over time, much of this decline was due to gender, obscuring the fact that ethnic discrimination remains persistent.

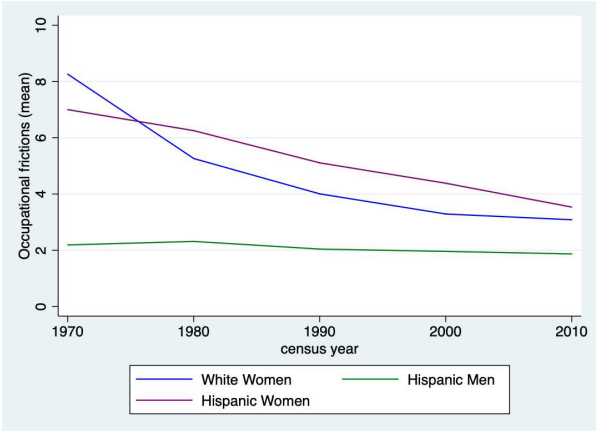



Figure 7: Groups to White Men

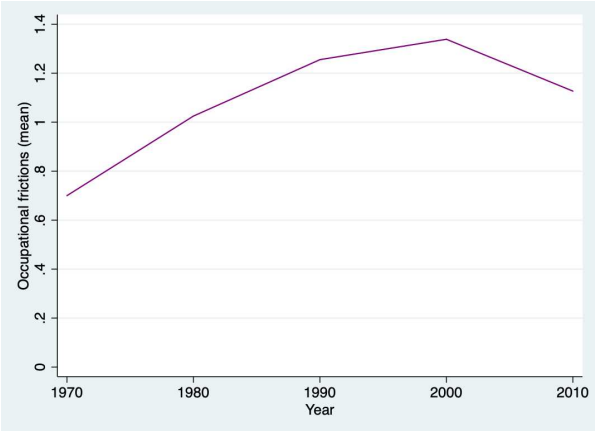



Figure 8: Hispanic Women to White Women

The figure on the left displays the earnings-weighted mean of occupational barriers relative to white men for white women, Hispanic men, and Hispanic women, all relative to white men. The figure on the right displays the ratio of occupational frictions against Hispanic women relative to white women.

Furthermore, the variance of the frictions over time for each group have decreased. Figure 9 shows the variance of frictions for each group. A value of zero for the variance would indicate that barriers are the same for a group in every occupation. Although Figure 7 indicates that on average, barriers against Hispanic men have not decreased over time, Figure 9 indicates that the variance of these barriers has decreased. Hispanic men experienced more dispersed levels of frictions across occupations, some higher and some lower, in 1970, but they experienced similar levels of frictions in 2010. However, the average value of these frictions across occupations were similar in 1970 and 2010. Convergence in the occupational distribution between Hispanic men and their white counterparts could reflect that Hispanic men have experienced more similar levels of discrimination across occupations over time, even if the overall level remains high.

The variances in the occupational frictions of white women and Hispanic women have also declined over time. In 1970, the variance of occupational frictions against Hispanic women was above that of white women, suggesting that Hispanic women experienced more variance in discrimination across occupations than white women, relative to white men. However, in the years following, the variance of frictions for both Hispanic women and white women have declined over time, suggesting that levels of discrimination across occupations against these

groups have converged. It also suggests, indicating that ethnicity, in the case of Hispanic women, is not the main determinant of differences in frictions across occupations.

The declines in the variances of frictions for both Hispanic women and Hispanic men suggest that frictions have become more stable across occupations throughout the years despite overall levels remaining high. Although occupational propensities have converged toward white men in the case of Hispanic men and have shown great movement relative to white women in the case of Hispanic women, much of this decline could be due to changes in the variance in frictions rather than decreases in the absolute levels of the frictions. Put differently, if Hispanic workers experience high barriers in some occupations and lower barriers in others, they will be more likely to choose the occupations where they experience lower barriers. However, if Hispanic workers experience the same level of barriers in all occupations, then barriers should not affect occupational sorting. Changes in propensities, then, could reflect changes in the variances of frictions and should not obscure the fact that discrimination remains high.

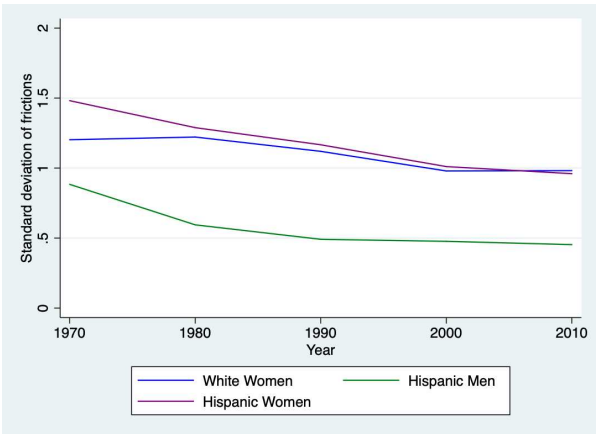



Figure 9: Frictions (variance)
Variance of composite occupational frictions over time relative to white men.

## 5.5 Estimating Human Capital and Wage Discrimination

The persistence of occupational frictions against Hispanic workers over time also raises the question of to what extent human capital barriers, versus labor market barriers, remain entrenched. The parameter $\tau_{ig}^{w}$ captures employer discrimination against workers of a mi-

nority group. It affects all cohorts within a group $g$ equally at a given point in time. The parameter $\tau_{ig}^h$ captures human capital differences across groups and represents all differences in childhood environments that later affect human capital accumulation. Such barriers include access to quality schooling. They also include differences in access to skill-building in different occupations: for example, we may expect to see a high value of $\tau_{ig}^h$ in the medical profession for Hispanic workers if they are more likely to live in under served communities where few adults are medical professionals who can provide the guidance necessary for them pursue a similar career path. The parameter $\tau_{ig}^h$ is fixed for a given cohort-group over time but it can vary across cohorts and groups. This parameter is unlike wage discrimination $\tau_{ig}^w$ which changes across time within a group but affects all cohorts within a group equally at a given point in time. These parameters are also unitless, although a higher value indicates that workers experience more barriers.

As in Hsieh et al., I write $\tau$ as the Cobb-Douglass split between $\tau_{ig}^h$ and $\tau_{ig}^w$. That is, $\tau = \tau^\alpha \tau^{1-\alpha}$, where $\tau^\alpha = \frac{1}{1-\tau_{ig}^w}$ and $\tau^{1-\alpha} = (1+\tau_{ig}^h)^\eta$. Using the procedure outlined in the Appendix of Hsieh et al. (2019), I estimate both $\tau_{ig}^h$ and $\tau_{ig}^w$. I assume that these two parameters are split 50/50 in the first year of my sample so that $\alpha = 1/2$. I then use the equation below, given frictions in the first year $\tau_{ig}^w(t)$, data on wages and schooling, as well as the returns to experience and the efficiency wage $w_i$ estimated from the previous steps, to estimate $\tau_{ig}^w(t+1)$ for the following year. From here, I use the relationship $\tau^\alpha = \frac{1}{1-\tau_{ig}^w}$ to find $\alpha$, from which I can then back out $\tau_{ig}^h$.

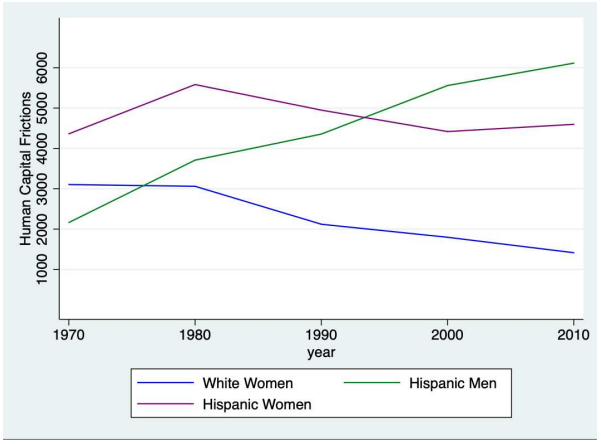
$$\frac{\overline{\text{wage}}_{ig}(c,t+1)}{\overline{\text{wage}}_{ig}(c,t)} = \frac{1-\tau_{ig}^w(t+1)}{1-\tau_{ig}^w(t)} \cdot \frac{w_i(t+1)}{w_i(t)} \cdot \frac{\gamma_{ig}(t+1-c)}{\gamma_{ig}(t-c)} \cdot \frac{s_i(c)^{\phi_i(t+1)}}{s_i(c)^{\phi_i(t)}} \tag{7}$$

Figure 10 displays the earnings-weighted human capital barriers $\tau_{ig}^h$ in each occupation. Relative to white men, human capital barriers against white women have fallen from 1970 to 2010. The human capital barriers against Hispanic women have risen, fallen, and then risen slightly over the years in the sample. The level in 2010 is similar to that in 1970, indicating that overall human capital discrimination against Hispanic women has not improved throughout this time period. Figure 11 indicates that the barriers faced by Hispanic women relative to white women have *increased* over time. This pattern can be attributed to the fact that barriers of white women due to gender have converged while barriers against Hispanic women due to both ethnicity and gender have stagnated over time, magnifying the

effect of entrenched ethnic human capital barriers. Figure 10 indicates that human capital barriers relative to white men have *risen* throughout this time period for Hispanic men. The patterns of these barriers for both Hispanic women and Hispanic men relative to their white counterparts suggest that lack of opportunities to invest in human capital due to ethnicity alone have not improved over time.

Figure 12 displays the earnings-weighted labor market frictions, or wage discrimination, $\tau_{ig}^{w}$ in each occupation. For white women, these frictions across occupations have remained relatively constant throughout the years in the sample relative to white men. The figure also indicates that wage discrimination against Hispanic women relative to white men has declined. These patterns suggest that wage discrimination has declined due to ethnicity rather than gender. Figure 11 shows that indeed wage discrimination against Hispanic women relative to white women has declined. Examining the patterns of wage discrimination against Hispanic men relative to white men in 12 also provides evidence that employer discrimination against Hispanic workers due to ethnicity has declined.

Examining the patterns of both human capital and wage discrimination against Hispanic workers relative to their white counterparts suggests that while employer discrimination against Hispanic workers has declined over time, barriers to obtaining the necessary human capital for work in given occupations have increased. As discussed in the previous section, overall ethnic discrimination has changed little over time. Therefore, the patterns of $\tau_{ig}^{h}$ and $\tau_{ig}^{w}$ observed show that the combined effects of lowering wage discrimination and rising human capital frictions against Hispanic workers approximately offset each other, leading to little change in overall occupational barriers.



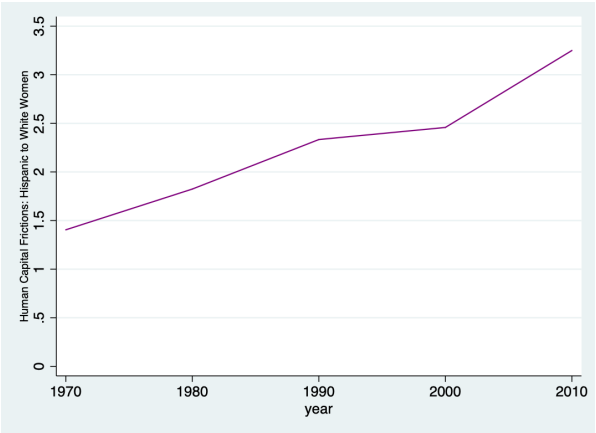
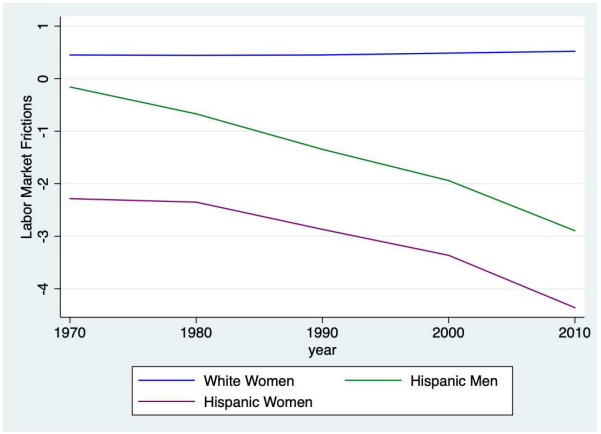




Figure 10: Human Capital Frictions $\tau_{ig}^{h}$ Figure 11: Hispanic Women to White Women

The figure on the left displays the mean of the human capital barriers $\tau_{ig}^{h}$ across occupations for each group, relative to white men. The figure on the right displays the mean of $\tau_{ig}^{h}$ of Hispanic women relative to white women.



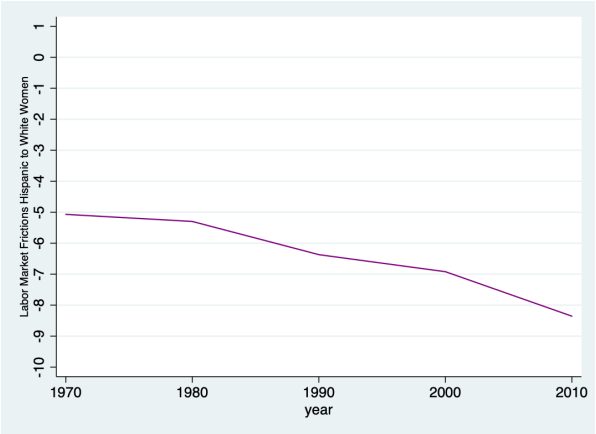




Figure 12: Labor Market Frictions $\tau_{ig}^{w}$ Figure 13: Hispanic Women to White Women

The figure on the left displays the mean of the labor market discrimination $\tau_{ig}^{w}$ across occupations for each group, relative to white men. The figure on the right displays the mean of $\tau_{ig}^{w}$ of Hispanic women relative to white women.

## 5.6 Robustness Check

To analyze the robustness of my model, I analyze the frictions for a subset of occupations to ensure that declines in propensities do not indicate that an *increase* in overall discrimination occurred in these occupations. For example, if many white and Hispanic women leave the Secretarial and Teaching occupations between 1970 and 2010, the model should not attribute this shift to an increase in discrimination in these occupations. Rather, the model should indicate a decrease in discrimination in *other* occupations. This robustness check is critical since the solution for the parameter $\tau_{ig}$ is heavily dependent on occupational propensities.

Figures 14 and 15 display the frictions against white women and Hispanic women for a select subset of occupations. This exercise should make a point in favor of the validity of my structural model. For both groups, frictions have remained constant in the occupations "Secretaries" and "Teachers," despite the declining propensities in these occupations over the years. For both groups, overall frictions due to both ethnicity and gender have declined in the law and medical occupations, and this effect is even more pronounced for Hispanic women, suggesting declines in ethnic barriers. In the construction occupation, they have declined or stagnated but have always remained high. In the year 1970, there were too few Hispanic women in the data to estimate frictions. This could indicate that barriers to entry for Hispanic women were exceptionally high in these occupations. The main takeaway from this robustness check is that my model does not incorrectly attribute declining occupational propensities across ethnicity and gender to increasing occupational discrimination in these occupations when it should point to decreasing discrimination in other occupations.

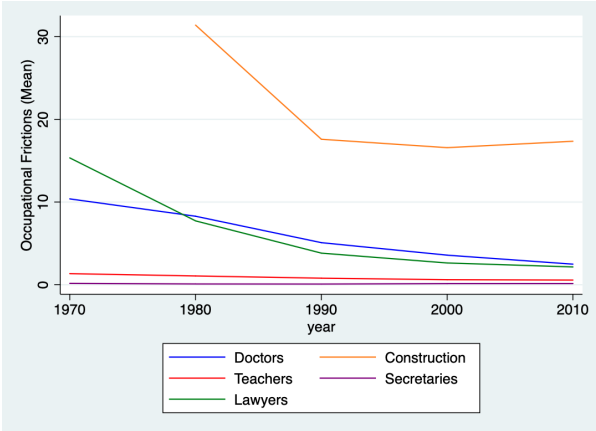



Figure 14: Frictions (White Women)

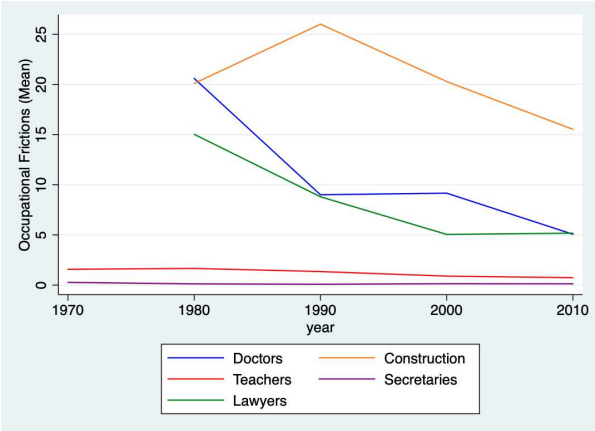



Figure 15: Frictions (Hispanic Women)

These figures show the level of occupational frictions relative to white men for white women (left) and Hispanic women (right). The figures indicate that occupations that discrimination has decreased where this pattern would be expected but has not increased in the occupations that see workers leave.

# 6 Growth in Aggregate Output Over Time

## 6.1 Model Specification

Finally, I turn to addressing the question of whether changes in the occupational distribution of Hispanic workers have affected growth in aggregate output over time. Since Hsieh et al. (2019) found that the improved allocation of talent over time among women and black men can explain a large portion of U.S. economic growth, I would expect to see that converging propensities of Hispanic workers to white workers can also explain U.S. economic growth. However, in light of the above discussion regarding frictions remaining entrenched, gains in aggregate output due to changes in the occupational distribution could reflect declines in the variances of occupational frictions over time rather than average levels of occupational frictions improving.

Since the occupational distributions of both Hispanic women and white women have converged toward that of white men over time due to changes in the variances in occupational barriers, I expect to find growth in aggregate output over time, even if the mean levels of barriers have not improved. I first provide an intuitive discussion of how the presence of variances in occupational barriers against Hispanic workers can affect aggregate output, as follows. Discrimination taxes on the firm side $\tau_w$ in a single occupation would lead a firm to hire a white worker rather than more talented Hispanic worker if the tax is large enough to offset the effect of talent on productivity. I describe a simple model as follows to illustrate this point. From the discriminating firm's perspective, aggregate productivity of an individual Hispanic worker is given by $(1 - \tau_w)q_{\text{Hispanic}}$, where $p$ represents the price of output, and $q_{\text{Hispanic}}$ represents the quantity of units produced by Hispanic workers in a given unit of time. Aggregate productivity of an individual white worker is given by $q_{\text{white}}$. Even if $q_{\text{Hispanic}} > q_{\text{white}}$, meaning that the Hispanic worker is more productive, the discrimination tax will, from the firm's perspective, make it appear "as if" the Hispanic worker is producing fewer units, when the opposite is the case. If such a tax did not exist ($\tau_w = 0$), the firm would view the Hispanic workers as producing $q_{\text{Hispanic}}$ units of output, choose to hire the Hispanic worker, and ultimately produce a higher value of real output. Hispanic workers would then receive worse compensation from employers and be less likely to choose occupations where $\tau_w$ is higher. In occupations with higher values of $\tau_w$, firms would experience worsened productivity due to overlooking talented Hispanic workers, and aggregating across the economy

would dampen aggregate output. However, if the values of $\tau_w$ are roughly the same across occupations, then employer discrimination will not negatively affect occupational sorting or aggregate output.

Discrimination taxes on the worker's side $\tau_h$ would lead a Hispanic worker with a high enough value of $\tau_h$ for one occupation relative to an occupation with a lower value of $\tau_h$ to work in the second occupation, even if that worker is more innately talented in the first occupation. Intuitively, workers' consumption utility is negatively related to the amount of "goods" invested in accumulating the necessary skills in a given occupation. In occupations with high human capital frictions, obtaining these goods is more costly for the worker, since these frictions artificially inflate the price of goods. If these frictions are higher in some occupations than in others, then a Hispanic worker who is equally talented as a white worker in a high $\tau^h$ occupation would not choose that occupation when the white worker would. Aggregating over all individuals in the economy in such cases, a large fraction of Hispanic workers would be excluded from a given occupation. As a result, the talent pool in the occupation would be much smaller, which would negatively affect aggregate productivity. However, if values of $\tau^h$ vary little across occupations, even if they remain high, then this parameter should not negatively affect occupational sorting or aggregate output.

To address my question of how convergence in occupational sorting affects aggregate output over time, I develop an expression that relates aggregate output to occupational frictions. I outline how I make a methodological contribution to Hsieh et al. (2019) by finding a closed-form solution for aggregate output in the general equilibrium model given in the Appendix of this paper. In the model, the supply aggregates workers' individual occupational choices in the presence of human capital frictions, and the discriminating firm's demand function satisfies profit maximization in each occupation. The efficiency wage, the profit-maximizing wage that the firm selects, clears the labor market so that the total labor supply equals labor demand in each occupation. I derive an expression for aggregate output that can be estimated using observed data. This approach contrasts with that of Hsieh et al., who begin by "guessing" initial values of parameters in the model, and is useful for calculating aggregate output when data on all relevant parameters is readily available.

After solving the model, I find that aggregate output across the $M$ occupations equals

$$Y(t) = \sum_{i=1}^{M} [(c_i \cdot w_{it}^{\sigma})^{\frac{1}{\sigma-1}} \cdot c_i]^{\frac{\sigma-1}{\sigma}} \tag{8}$$

where $c_i = \sum_g \sum_c q_g(c) p_{ig}(c) E[\epsilon_{ig} h_{ig} | \text{choose occupation } i]$ and, as given earlier, $h_{ig}(c,t) = \gamma(t-c) s_i(c)^{\phi_i} e_{ig}(c)^{\eta}$, where $e_{ig}(c)$ is the function of $\tau_{ig}^h$ and $\tau_{ig}^w$ described in the Model section. The value $\sigma$ is the elasticity of substitution across occupations. I use the value $\sigma = 3$, though the model is robust to different values of $\sigma$. With data on the quantity of workers in a cohort-group in total $q_g(c)$, their propensity to work in a given occupation $p_{ig}(c)$, the residuals from the cross-sectional regression used to fit the maximum likelihood expression in Section 1.4, and $w_i$, I can calculate aggregate output, summing across all occupations. Aggregate output is positively related to the number of workers in a group-cohort in a given occupation (the product of $q_g(c)$ and $p_{ig}(c)$), the efficiency wage units in that occupation $w_i$, and the expected talent of a worker in an occupation *given that* the worker chooses that occupation. The table below displays the relevant parameters in the model and their empirical target.

Table 5: Empirical Parameters and Target

| Parameter | Definition | Empirical Target |
|---|---|---|
| $q_g(c)$ | quantity of workers | number of workers in a cohort-group |
| $p_{ig}(c)$ | occupational propensity | likelihood relative to white men |
| $w_{it}$ | efficiency wage units | given by Equation 11 |
| $\epsilon_{ig}$ | talent in an occupation | residuals from regression in the Appendix |
| $\sigma$ | elasticity of substitution across occupations | $\sigma = 3$ by assumption |
| $\tau_{ig}^w$ | wage discrimination | above procedure splitting $\tau$ |
| $\tau_{ig}^h$ | human capital discrimination | above procedure splitting $\tau$ |
| $s_i$ | years of schooling | given in the data |
| $e_{ig}$ | education spending | estimate $e^*$ from the "Model" section |
| $\gamma(t-c)$ | returns to experience | discussed previously |

With data on these parameters, I can estimate the structural model to analyze the counterfactual growth trajectory of the U.S. economy, had discrimination remained at its levels in 1970. For this counterfactual experiment, I hold $\tau_{ig}^h$ and $\tau_{ig}^w$ at their 1970 levels in each occupation. I implement this procedure for three specifications. I also must calculate the hypothetical propensities, had discrimination remained constant. I use Equation 5 to back

out the hypothetical propensities, given the counterfactual values of $\tau_{ig}$ and actual values of $\gamma(t-c)$ wages. First, I allow discrimination against white women, Hispanic men, and Hispanic women to evolve, representing the actual growth trajectory of the U.S. economy. I compare this trajectory to the baseline counterfactual scenario where discrimination against these groups remains constant at its levels in 1970. This exercise addresses the question of whether improved occupational sorting due to either ethnicity or gender have improved growth in aggregate output. In the second experiment, I compare the actual growth trajectory of the economy to to a counterfactual scenario where against Hispanic women are held at their 1970 levels but barriers against other groups evolve. This experiment can assess the effects of improved occupational sorting of Hispanic women, but it does not address the relative contributions of improved occupational sorting due to ethnicity when confounded with gender. Therefore, I conduct a third counterfactual experiment. I compare the actual growth trajectory of the economy to to a counterfactual scenario where against Hispanic men are held at their 1970 levels but barriers against other groups evolve. Since men do not experience gender disparities in my model, this experiment will allow me to isolate the effect of ethnicity-related occupational sorting on aggregate output.

Figure 16 illustrates the results from each of the three experiments. The black line illustrates the actual growth trajectory of the U.S. economy, taking into account declining barriers for all three groups. The blue line illustrates the counterfactual where barriers against all three groups remain entrenched at their 1970 levels. As the figure indicates, the growth trajectory of the U.S. economy is noticeably higher than it would have been if both gender and ethnic discrimination had remained at their 1970 levels. The purple line illustrates the counterfactual where barriers against Hispanic women remain entrenched. The position of the purple line under the black line indicates that improvements in occupational sorting of Hispanic women can explain some of the growth in aggregate output. Finally, the green line illustrates the counterfactual where barriers against Hispanic men remain entrenched at their 1970 levels. The line is obscured by the purple line, indicating that the effects of changing barriers against Hispanic workers on aggregate output are roughly the same for Hispanic men and Hispanic women. In the first experiment, I find that as a result of falling barriers against Hispanic workers and women, aggregate output in the year 2010 is 9.89% higher than it would have been, had discrimination remained constant over this time frame. In the second experiment, I find that aggregate output is about 4.29% higher than it

would have been, had discrimination against Hispanic women only remained constant over this time frame. Finally, in the third experiment I find that aggregate output is also about 4.29% higher than it would have been, had discrimination against Hispanic men only across occupations remained constant over this time frame.

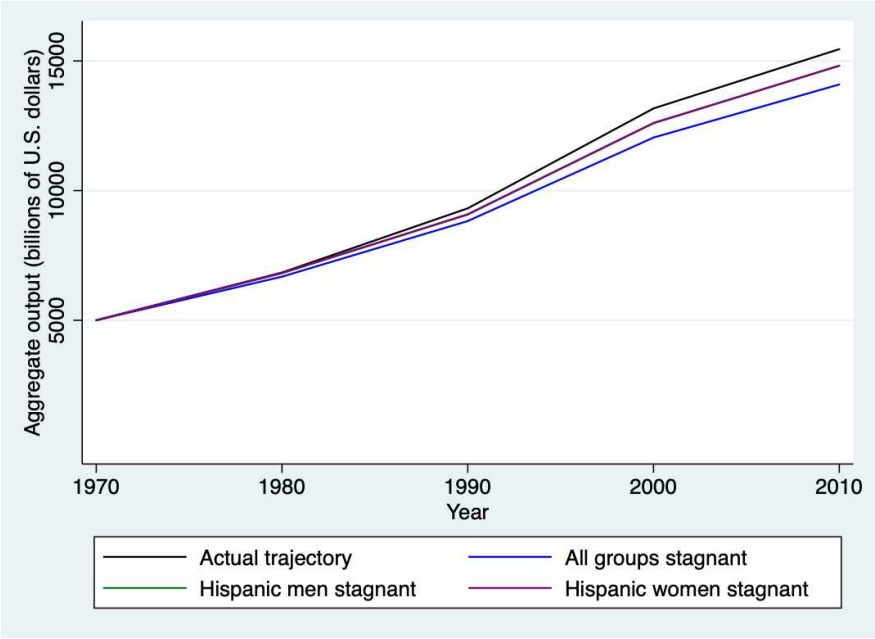



Figure 16: Aggregate Output

As the three counterfactual experiments indicate, some of the improvements in aggregate output over time are due to convergence in the occupational distribution by gender and some is driven by Hispanic ethnicity. An important conceptual point is to be made here. Improved occupational sorting and the resultant growth in aggregate output can be explained by declining gender discrimination. However, improved occupational sorting and growth in aggregate output can sometimes be explained by convergence in the variances over time of the levels of ethnic discrimination, rather than improvements in the mean levels. In Section 5.4, I find persistently high averages in ethnic barriers over time, but I find decreases in differences in barriers across individual occupations over time. As discussed, if little variance in occupational barriers exists, then barriers will not influence occupational propensities, even

if they remain high overall. If the *levels* of discrimination against Hispanic workers were to further decline in the occupations where Hispanic workers are most underrepresented, occupational propensities would likely converge further and generate more growth in aggregate output.

# 7 Discussion and Conclusion

This paper investigates the question of whether discrimination against Hispanic workers in the United States has declined from the years 1970 to 2010. It also investigates the extent to which improvements in the growth trajectory of the United States economy are due to changing patterns in ethnic discrimination against Hispanic workers.

In response to the question of whether ethnic discrimination has decreased, I find that discrimination against Hispanic workers has not declined since 1970. When examining the patterns of discrimination against Hispanic women relative to white men, I find that Hispanic women have experienced overall improvements in occupational opportunities, just as white women have. However, seeing that discrimination against Hispanic women has not converged toward that of white women over time allows me to conclude that much of these improvements were driven by declining gender discrimination rather than declining ethnic discrimination. Similarly, my results suggest that occupational discrimination against Hispanic men has shown little change since 1970. The patterns of occupational frictions for both Hispanic men and Hispanic women suggest that ethnic discrimination against Hispanic workers has remained entrenched over the last decades.

I also examine the patterns of occupational preferences over time. I find that for the years in my sample, group preferences of Hispanic workers have closely tracked those of their white counterparts. Since the probability of working in a given occupation is related to both group preferences and discrimination, these patterns suggest that differences in the occupational distribution between Hispanic workers and white workers are primarily due to discrimination rather than tastes and preferences. That is, my results indicate that Hispanic workers are not choosing certain occupations since they enjoy working in those occupations more than others or avoiding certain occupations due to a dislike for these occupations. Instead, discrimination is preventing them from working in the occupations where their talent would best be allocated. Since disparities in labor market outcomes between Hispanic workers and white workers are due to discrimination rather than group choices or cultural norms, policies to address these disparities are called for.

Decomposing discrimination against Hispanic workers into human capital barriers and labor market discrimination can provide insight into what policy recommendations would be most effective for reducing these disparities. When I conduct this decomposition ex-

ercise, I find that labor market discrimination against Hispanic workers has declined, but human capital barriers have worsened over time. That is, employers were less likely to discriminate against a Hispanic worker in 2010, relative to 1970. However, disparities in the early childhood environments or opportunities to invest in skills between Hispanic and white individuals have increased over time. Since overall discrimination against Hispanic workers has remained relatively stagnant over time, these two effects approximately offset each other. This exercise should make the conceptual point that analyzing movement in overall discrimination is not enough to examine where these patterns come from. Although I find that overall discrimination against Hispanic workers shows little movement, I find that both labor market barriers and human capital barriers show substantial movement in opposite directions. This exercise should also suggest that policies to reduce discrimination against Hispanic workers should focus on providing them with more opportunities for building skills in a variety of occupations. Such policies could include improvements in the public school systems of predominantly Hispanic neighborhoods or better access to internships and extracurricular opportunities for children in these environments.

In response to the question of whether improvements in U.S. economic growth are due to declines in ethnic discrimination against Hispanic workers, I find that improvements in U.S. economic growth due to the improved allocation of talent is driven by declines in the variances of frictions rather than improvements in the overall levels of frictions. I find that the growth trajectory of the economy is higher than it would have been if discrimination in individual occupations against women and Hispanic workers had remained at their levels in 1970. This suggests that the allocation of talent in the United States economy has shown substantial improvements and that workers of discriminated-against groups have increasingly contributed their talents to occupations that would not have otherwise pursued. Given that declines in gender discrimination have enhanced the level of aggregate output, one might expect that declines in overall levels of ethnic discrimination against Hispanic workers are also what explains growth in aggregate output to ethnicity. However, in the case of Hispanic workers, much of the growth in aggregate output due to convergence in propensities can be explained by changes in the variances of occupational frictions rather than improvements in the absolute levels. Furthermore, since the occupational distribution of Hispanic workers has converged toward that of white workers but not fully, any remaining differences suggest that reductions in discrimination are called for to further improve talent allocation and the

growth trajectory of the U.S. economy.

I now explain the limitations of my study and suggest valuable directions for future research. First, I assume that workers' talent draws across occupations are uncorrelated. While this assumption allows for tractability in my model, this assumption may not hold in reality. It does not account for the fact that workers may be similarly skilled at similar occupations. For example, a worker skilled in the finance occupation could possess analytical capabilities that would also make them a successful engineer. Future research could develop a model to analyze discrimination that relaxes the assumption that skill draws are independent.

Second, and most substantially, I do not include immigrants in my study due to the issue discussed earlier of selective migration potentially confounding my results. Future research could investigate whether occupational discrimination against Hispanic immigrants has declined over time and whether these potential declines have improved occupational sorting and U.S. economic growth. Such research would need to develop a methodological approach for ensuring that selective migration does not confound results. That is, it must ensure that rising shares of Hispanic immigrants in given occupations are not attributed to declines in discrimination when they are in fact due to the particular skillsets of the immigrants who come to the United States to pursue work in these occupations. Both of these potential improvements to my study could provide more robust insight into how declines in barriers against minority groups in the United States could improve the productive capacity of the U.S. economy.

# References


Becker, H. and B. Geer (1957). Participant observation and interviewing: a comparison. *Human Organization 16* (3), 28–32.

Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review 94* (4), 991–1013.

Blau, F. D., P. Brummund, and A. Y.-H. Liu (2013). Trends in occupational segregation by gender 1970–2009: Adjusting for the impact of changes in the occupational coding system. *Demography 50*(2), 471–492.

Doleac, J. L. and B. Hansen (2016). Does "ban the box" help or hurt low-skilled workers? statistical discrimination and employment outcomes when criminal histories are hidden. Technical report, National Bureau of Economic Research.

Goldberg, M. S. (1982). Discrimination, nepotism, and long-run wage differentials. *The Quarterly Journal of Economics 97*(2), 307–319.

Guryan, J. and K. K. Charles (2007). *Prejudice and the Economics of Discrimination.* National Bureau of Economic Research.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019). The allocation of talent and U.S. economic growth. *Econometrica 87* (5), 1439–1474.

Hu, W.-Y. (2000). Immigrant earnings assimilation: estimates from longitudinal data. *American Economic Review 90*(2), 368–372.

Kuka, E., N. Shenhav, and K. Shih (2020). Do human capital decisions respond to the returns to education? Evidence from DACA. *American Economic Journal: Economic Policy 12* (1), 293–324.

Lang, K. and A. Kahn-Lang Spitzer (2020). Race discrimination: an economic perspective. *Journal of Economic Perspectives 34*(2), 68–89.

Park, K. H. (2017). Do judges have tastes for discrimination? Evidence from criminal courts. *Review of Economics and Statistics 99*(5), 810–823.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers 3*(2), 135–146.

Yinger, J. (1998). Evidence on discrimination in consumer markets. *Journal of Economic Perspectives 12*(2), 23–40.

# 8 Appendix

## 8.1 Estimates of Relevant Shape Parameters

To structurally estimate the frictions, I must obtain the parameter $\theta$, which refers to the distribution of talent in an occupation, and the parameter $\eta$, which refers to the elasticity of human capital with respect to education spending. A high value of $\theta$ corresponds to a smaller dispersion of talent within an occupation. A higher value of $\eta$ would indicate that human capital accumulation is more sensitive to education spending. I take $\eta$ to be .103, as in Hsieh et al. (2019). In the following section, I show how I develop a procedure of estimating these procedures, given data on income that is top coded.

Following Hsieh et al., I estimate $\theta(1-\eta)$ by fitting a distribution of residuals $x_i$ from the following equation, which is a regression of hourly wages on $66 \times 4 \times 3$ occupation-group-year fixed effects in each year

$$\ln(\text{wage})_{kigc} = \gamma_{igc}\{\text{occupation } i{=}1\} \times \{\text{group } g{=}1\} \times \{\text{cohort } c{=}1\} + x_{kigc}, \tag{9}$$

where $k$ indexes the individual in occupation $i$ group $g$ and cohort $c$.

An individuals' residual $x_i$, or the part of their earnings that cannot be explained by group, occupation, or cohort, follows a Type II i.i.d. Frechet distribution with a shape parameter of $\theta(1-\eta)$. That is, we have $F(wage_i|x_i, \theta(1-\eta)) = \exp(-x_i^{-\theta(1-\eta)})$. The distribution of these residuals has a shape parameter of $\theta(1-\eta)$, where $\theta$ represents the dispersion of skills within an occupation. In the extreme case that $\theta = 0$, the dispersion of residuals is entirely governed by the elasticity of substitution with respect to education spending. The parameter $(1-\eta)$ indicates that a higher value of $\eta$ corresponds to a higher dispersion of these residuals. That is, as the elasticity of human capital with respect to education spending increases, unexplained differences in wage between individuals become less dispersed.

Figure 17 confirms that the residuals in the data indeed follow this distribution. I maximize the log likelihood of observing the wage, given the values of the skill draw. I perform this procedure five times, once for each year in my sample and take the average value of $\theta(1-\eta)$ across these years.

I now explain how I develop a procedure for estimating this parameter using maximum likelihood methods given earnings data with observations that are top coded. The earnings of individuals with incomes greater than certain values are replaced with a maximum value. Since earnings for individuals in 2010 are given by state, I use the state fip code to identify which observations are top coded. I find that in the data, 0.4% of the observations in total are top coded. I must use an alternative log likelihood formula to fit a given distribution of residuals that include top coded observations. I maximize the likelihood function using the following formula for top coded data:

$$\ln(L_n(\theta(1-\eta)) = \sum_{i=1}^{n}\{(1-d_i)\ln f^*(\text{wage}_i|x_i,\theta(1-\eta)) + d_i(1-\ln F^*(\text{wage}_i|x_i,\theta(1-\eta))\},$$

where $n$ is the number of individuals in the data, $d_i$ is a dummy variable that takes on the value of "1" if that observation is topcoded, $F^*$ is the probability density function, and $f^*$ is the cumulative density function, the derivative of the probability density function. By the above formula, the problem that I take to the data becomes

$$\max_{\theta(1-\eta)} \ln(L_n(\theta(1-\eta)) = \sum_{i=1}^{n}\{(1-d_i)[\ln(\theta(1-\eta))-x_i^{-\theta(1-\eta)}+(-\theta(1-\eta)-1)\ln(x_i)]+d_i[1+x_i^{-\theta(1-\eta)}]\}.$$

The estimate that I obtain for $\theta(1-\eta)$ each year varies little and has an average value of 1.299 across the years in the data, implying a value of $\theta = 1.488$. With these values of $\theta$, $\gamma$, and $\eta$, I can now estimate my model to infer the returns to experience $\gamma$, and eventually the $\tau$'s and $\tilde{z}$'s across occupations.

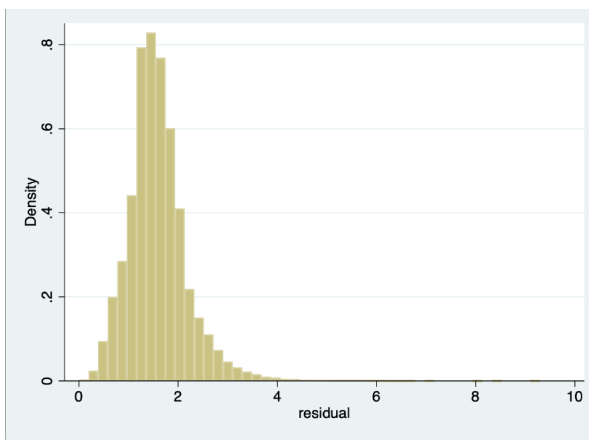



Figure 17: Distribution of Residuals

## 8.2 Estimating the Returns to Experience

To develop a computational procedure for estimating the returns to experience, I make use of the equations given in the Appendix of Hsieh et al. (2019) to write a simple equation for calculating the value of $\gamma$. As stated, the returns to experience $\gamma(t-c)$ can be estimated from the equation

$$\frac{\overline{wage}_{i,wm}(c,t)}{\overline{wage}_{i,wm}(c,c)} = \frac{w_i(t)(\gamma_{ig}(t-c))s_i^{\phi_i(c)}}{w_i(c)s_i^{\phi_i(c)}}. \tag{10}$$

In the above equation, the occupational efficiency wage $w_i(t)$ is not directly observed in the data. Therefore, to derive a closed-form solution that will allow me to back out $\gamma(t-c)$, I must express $w_i(t)$ in terms of observed data. I now show how I perform this procedure.

The pre-market period is assumed to be 25 years long so that $s_i = \frac{\text{Years of Education}}{25}$. Then $\phi_i$, the return to time investments in human capital, is given by

$$\phi_i = \frac{1-\eta}{3\beta} \cdot \frac{s_i}{1-s_i},$$

where I take the value $\beta = .231$, as given by Hsieh et al.

Since 67 values of $z_{ig}$ along with $m_{ig}$ are to be recovered and the data contains information on wages for only 66 market occupations, we must make two assumptions to pin down each parameter. First, the value of $\tilde{z}_{home,wm}$ is normalized to 1 and average earnings in the home sector for young white men are assumed to be the same as the occupation "Secretaries." Then the group-specific productivity in an occupation is given by

$$m_{wm} = \left( \frac{\overline{wage}_{i,wm} \tilde{z}_{i,wm} (1-s_i)^{\frac{1}{3\beta}}}{\overline{\gamma}\eta} \cdot \frac{\overline{\gamma}_{i,wm}}{\gamma_{i,wm}} \right)^{\theta(1-\eta)}. \tag{11}$$

Since $\tilde{z}_{home,wm} = 1$ and $m_{wm}$ does not vary across occupation, we may write

$$m_{wm} = \left( \frac{\overline{wage}_{\text{home},wm} (1-s_{\text{home},wm})^{\frac{1}{3\beta}}}{\overline{\gamma}\eta} \cdot \frac{\overline{\gamma}_{\text{home},wm}}{\gamma_{\text{home},wm}} \right)^{\theta(1-\eta)}. \tag{12}$$

I use the above equations to begin to derive a closed-form solution relating $\gamma$ to the wage gap and other parameters observed in the data. I first substitute the equation for $m_{wm}$, given by Equation 11, into Equation 12 to obtain an expression for the ratio of $\tilde{z}_{i,wm}$ in terms of the $\gamma$'s and observed data. After some algebra, I show that we may write

$$\frac{\tilde{z}_{i,wm}(t-c=1)}{\tilde{z}_{i,wm}(t-c=0)} = \left( \frac{1}{C} \cdot \gamma(1)^{\theta(1-\eta)} \cdot \frac{p_{i,wm}(t=1)}{p_{i,wm}(t=0)} \right)^{\frac{1}{\theta(1-\eta)}}, \tag{13}$$

where

$$C = \left( \frac{\overline{wage}_{home,wm}(1)}{\overline{wage}_{home,wm}(0)} \cdot \frac{(1-s_i(1))^{\frac{1}{3\beta}}}{(1-s_i(0))^{\frac{1}{3\beta}}} \right)^{\theta(1-\eta)}.$$

Then $w_i$ is estimated from the data on occupational shares across occupations, and is given by

$$w_i = \frac{[p_{i,wm} \cdot m_{wm}]^{\frac{1}{\theta}}}{\overline{\gamma_{i,wm}} \cdot s_i^{\phi_i} [(1 - s_i) z_{i,wm}]^{\frac{1-\eta}{3\beta}}}. \tag{14}$$

Substituting the expression for $m_{wm}$ the ratio of $\tilde{z}$ into the equation above, we obtain the ratio $w_i(c, t = 1)/w_i(c, t = 0)$ expressed entirely in terms of the parameter $\gamma$ and observed data. I substitute this ratio into Equation 10 to express the wage gap in terms of $\gamma$ and observed data. After some algebra, I may write the wage gap as given by Equation 15.

$$\frac{\overline{wage}_{i,wm}(c,t)}{\overline{wage}_{i,wm}(c,c)} = C^{(3\beta\theta-1+\eta)/(\theta(1-\eta))} * \gamma^{\eta} * \left(\frac{p_{i,wm}(c,t=0)}{p_{i,wm}(c,t=1)}\right)^{(3\beta\theta-1+\eta)/(\theta(1-\eta)(1-3\beta))}. \tag{15}$$

This is an equation that can be taken to the data. Since the value of $\gamma(t - c)$ is fixed across occupations for a given group and cohort while the wage gap varies across occupations for a given group and cohort, I use a nonlinear least squares procedure to find the value of $\gamma(t - c)$ that minimizes the squared distance between the wage gap for each occupation and its theoretical equivalent given by the right side of the last equation.