The Allocation of Talent and U.S. Economic Growth

Chang-Tai Hsieh
Chicago Booth and NBER

Erik Hurst
Chicago Booth and NBER

Charles I. Jones
Stanford GSB and NBER

Peter J. Klenow [*]
Stanford University and NBER

October 10, 2012 – Version 1.0

Abstract

Over the last 50 years, there has been a remarkable convergence in the occupational distribution between white men and women and blacks. We measure the macroeconomic consequences of this convergence through the prism of a Roy model of occupational choice, where women and blacks face frictions in labor markets and in the accumulation of human capital. We find that the changing frictions implied by the observed convergence in occupational choice can explain 15 to 20 percent of aggregate wage growth during the last fifty years.

[*]We are grateful to seminar participants at Boston College, Boston University, Brown, Chicago, Columbia, Duke, Harvard, Michigan, MIT, Penn, Princeton, Stanford, UBC, UCLA, USC and Yale for helpful comments, and to Huiyu Li and Gabriel Ulyssea for excellent research assistance.

1. Introduction

In 1960, 94 percent of doctors and lawyers were white men. By 2008, the fraction was just 62 percent. Similar changes occurred throughout the economy over the last fifty years, particularly among highly-skilled occupations. During this time, there has been a remarkable convergence in the occupational distribution between white men and women and blacks.[1]

This paper measures the macroeconomic consequences of the changes in the labor market outcomes of white women, black men, and black women over the last fifty years through the prism of a Roy (1951) model of occupational choice. We assume that every person is born with a range of talents across all possible occupations and chooses the occupation with the highest return. Differences in the occupational distribution between men and women can be driven by differences in the distribution of talent between groups. Rendall (2010), for example, shows that brawn-intensive occupations (such as construction) in the U.S. are dominated by men, and that changes in the returns to brawn vs. brain intensive occupations can explain changes in the occupational distribution of women vs. men since the 1960s. Relatedly, Goldin and Katz (2002) and Bertrand, Goldin and Katz (2010) provide evidence that innovations in contraception and increased labor market flexibility for women had important effects on the occupational choices of women.

However, it seems likely that other forces may also play an important role. Consider the world that Supreme Court Justice Sandra Day O'Connor faced when she graduated from Stanford Law School in 1952. Despite being ranked third in her class, the only private sector job she could get immediately after graduating was as a legal secretary (Biskupic, 2006). Such barriers might explain why white men dominated the legal profession at that time. And the fact that private law firms are now more open to hiring talented female lawyers might explain why the share of women in the legal profession has increased dramatically over the last fifty years. Similarly,

[1]These statistics are based on the 1960 Census and a pooled sample of the 2006-2008 American Community Survey. We discuss the sample in more detail below. A large literature that we will not attempt to survey provides more extensive documentation of these facts. See Blau (1998), Goldin (1990), and Smith and Welch (1989) for assessments of this evidence.

the Civil Rights movement of the 1960s is surely important in explaining the change in the occupational distribution of blacks.[2]

To capture these forces, we make several changes to the canonical Roy framework. First, we allow for the possibility that each group faces different occupational frictions in the labor market. We model these frictions as a group/occupation-specific "tax" on earnings that drives a wedge between a group's marginal product in an occupation and their take home pay. One interpretation of these "taxes" is that they represent preference-based discrimination as in Becker (1957). For example, one reason why private law firms would not hire Justice O'Connor is that the law firms' partners (or their customers) viewed the otherwise identical legal services provided by female lawyers as somehow less valuable.[3]

Second, we allow for frictions in the acquisition of human capital. We model these frictions as a group-specific tax for each occupation on the inputs into human capital production. These human capital frictions could represent the fact that some groups were restricted from elite higher education institutions, that black public schools are underfunded relative to white public schools, that there are differences in prenatal or early life health investments across groups, or that social forces steered certain groups towards certain occupations.[4]

Finally, we allow for changes in the returns to skill across occupations. If these changes are common to all groups, then they will not affect the occupational distribution for men and women differently. However, some technological changes may

[2]See Donohue and Heckman (1991) for an assessment of the effect of federal civil rights policy on the economic welfare of blacks.

[3]Consistent with the Becker (1957) interpretation of labor market frictions, Charles and Guryan (2008) show that relative black wages are lower in states where the marginal white person is more prejudiced (against blacks).

[4]Here is an incomplete list of the enormous literature on these forces. Karabel (2005) documents how Harvard, Princeton, and Yale systematically discriminated against blacks, women, and Jews in admissions until the late 1960s. Card and Krueger (1992) documents that public schools for blacks in the U.S. South in the 1950s were underfunded relative to schools for white children. See Chay, Guryan and Mazumder (2009) for evidence on the importance of improved access to health care for blacks. See Fernandez (2012) and Fernandez, Fogli and Olivetti (2004) on the role of social forces in women's occupational choice. Goldin and Katz (2002), Bailey (2006), Bertrand, Goldin and Katz (2010), and Bailey, Hershbein and Milleri (2012) document that innovations related to contraception had important consequences for female labor market outcomes and educational attainment. Fernandez and Wong (2011) stress rising divorce rates as a force behind women's rising labor force participation and educational attainment.

be group specific. The innovations related to contraception mentioned earlier are a prime example.

In our augmented Roy model, all three forces — differences in barriers to occupational choice (either in the human capital market or labor market), differences in relative ability across occupations, and differences in the relative returns to occupational skills — will affect the occupational distribution between groups. To make progress analytically, we follow McFadden (1974) and Eaton and Kortum (2002) and assume that the distribution of talent follows an extreme value distribution. This assumption gives us two key results. First, we get a closed-form expression relating the share of a group in an occupation to the occupation-specific frictions faced by the group. Second, the average wage gap between groups turns out to be the same in all occupations: smaller barriers in an occupation lead to a selection effect in which less talented people choose that occupation, and these two forces exactly net out in our model. As a result, frictions show up occupation by occupation in quantities rather than prices. Taken together, these two results allow us to back out the occupation-specific frictions for each group from data on occupational shares and average wages. Using data from the decadal U.S. Censuses and the American Community Surveys, we find that the dispersion and the mean of the occupational frictions faced by women and blacks decreased over the last fifty years.

We then embed the Roy model in general equilibrium. This allows us to estimate the effect of our three forces (occupational barriers, talent distribution, occupation-specific technical change) on aggregate productivity. We find that changes in occupational barriers facing blacks and women can explain 15 to 20 percent of aggregate wage growth between 1960 and 2008. Furthermore, essentially all of the gain is driven by the movement of women into high-skilled occupations. We also entertain the possibility that occupational frictions are unchanged and instead some group-specific technological changes drive the changing occupational allocation. Under this scenario, the aggregate wage gains from the changing occupational allocation are similar, but the allocation itself may be efficient.

In addition, we find that real wages increased from our highlighted mechanisms by roughly 40% for white women, roughly 60% for black women, and roughly 45%

for black men, but *fell* by about 5% for white men. The reduction in frictions can thus account for 90 to 95 percent of the narrowing of the wage gap between blacks and women vs. white men. Also, we find that about 75 percent of the rise in women's labor force participation is attributable to the decline in occupational frictions.

The paper proceeds as follows. The next section lays out the basic model of occupational choice. We then use this framework to measure the frictions in occupational choice between blacks and women versus white men in Section 3. Section 4 embeds the occupational choice framework into general equilibrium. In Section 5, we explore the macroeconomic consequences of the changes in occupational frictions across groups. We offer some closing thoughts in the final section.[5]

2. Occupational Sorting and Aggregate Productivity

We start with the occupational choice decision. The economy consists of a continuum of people working in N possible occupations, one of which is the home sector. Each person possesses an idiosyncratic ability in each occupation — some people are good economists while others are good nurses. The basic allocation to be determined in this economy is how to match workers with occupations.

2.1. People

Individuals are members of different groups, such as race and gender, indexed by g. A person with consumption c and leisure time $1-s$ gets utility

$$U = c^{\beta}(1-s) \tag{1}$$

where s represents time spent on schooling, and β parameterizes the tradeoff between consumption and schooling.

[5] Several other recent papers are worth noting for related contributions. Ellison and Swanson (2010) show that the highest-achieving girls in elite mathematical competitions are much more geographically concentrated than the highest-achieving boys, suggesting that many girls with the ability to reach these elite levels are not doing so. Cavalcanti and Tavares (2007) use differences in wage gaps across countries in a macro model to measure the overall costs of gender discrimination and find that it is large. Dupuy (2012) studies the evolution of gender gaps in world record performances in sport.

Each person works one unit of time in an occupation indexed by i. Another unit of time — think "when young" — is divided between leisure and schooling. A person's human capital is produced by combining time s and goods e. The production function for human capital in occupation i is

$$h(e,s) = \bar{h}_{ig} s^{\phi_i} e^{\eta}. \tag{2}$$

Note that we will omit subscripts on individual-specific variables (such as variables s and e in this case) to keep the notation clean.

The parameter $\bar{h}_{ig}$ allows for potentially different efficiency in human capital accumulation across groups. We have two potential interpretations in mind for $\bar{h}_{ig}$. One is that family background (e.g., nutrition and health care) could differ across groups, thereby affecting the human capital payoff to investments in schooling quantity (s) and quality (e). Cunha, Heckman and Schennach (2010) provide evidence of such complementarity between early and later human capital investments. A second interpretation is that women's childbearing may disrupt human capital investment, a force that could change over time with fertility and technology. For example, Goldin and Katz (2002) and Bertrand, Goldin and Katz (2010) provide evidence that innovations in contraception had important effects on the timing of childbearing and, in turn, the education and occupational choices of women.

In addition, we allow for two frictions that affect the human capital and occupational choice decisions of individuals. The first friction affects human capital choices. We model this friction as a "tax" τ^h_{ig} that is applied to the goods e invested in human capital and that varies across both occupations and groups. We think of this tax as representing forces that affect the cost of acquiring human capital for different groups in different occupations. For example, τ^h_{ig} might represent discrimination against blacks or women in admission to universities, or differential allocation of resources to public schools attended by black vs. white children, or parental liquidity constraints that affect children's education. Additionally, it can represent the differential investments made toward building up math and science skills in boys relative to girls.

The second friction we consider can be thought of as a friction in the labor market. A person in occupation i and group g is paid a wage equal to $(1-\tau^w_{ig})w_i$ where w_i denotes the wage per efficiency unit of labor paid by the firm. One interpretation of τ^w_{ig} is that it represents preference-based discrimination by the employer or customers as in Becker (1957).

Consumption is equal to labor income less expenditures on education, incorporating both frictions:

$$c = (1-\tau^w_{ig})w\epsilon h(e,s) - e(1+\tau^h_{ig}). \tag{3}$$

Note that pre-distortion labor income is the product of the wage received per efficiency unit of labor, the idiosyncratic talent draw ϵ in the worker's chosen occupation, and the individual's acquired human capital h.

Given an occupational choice, the occupational wage w_i, and idiosyncratic ability ϵ in the occupation, each individual chooses c, e, s to maximize utility:

$$U(\tau^w, \tau^h, w, \epsilon) = \max_{c,e,s} \,(1-s)c^\beta \quad s.t. \quad c = (1-\tau^w_{ig})w\epsilon h(e,s) - e(1+\tau^h_{ig}). \tag{4}$$

This yields the following expressions for the amount of time and goods spent on human capital:

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

$$e^*_{ig}(\epsilon) = \left(\frac{\eta(1-\tau^w_{ig})w_i\bar{h}_{ig}s_i^{\phi_i}\epsilon}{1+\tau^h_{ig}}\right)^{\frac{1}{1-\eta}}$$

Time spent on accumulating human capital is increasing in ϕ_i. Individuals in high ϕ_i occupations acquire more schooling and have higher wages to compensate them for the time spent on schooling. Forces such as w_i, τ^h_{ig}, and τ^h_{ig} do not affect s because they have the same effect on the return and on the cost of time. In contrast, these forces change the returns of investment in *goods* in human capital (relative to the cost) with an elasticity that is increasing in η. After substituting the expression for human capital into the utility function, we get the following expression for indirect utility in occupation i:

$$U(\tau_{ig}, w_i, \epsilon_i) = \left(\frac{w_i \bar{h}_{ig} s_i^{\phi_i} (1-s_i)^{\frac{1-\eta}{\beta}} \epsilon_i \cdot \eta^\eta (1-\eta)^{1-\eta}}{\tau_{ig}} \right)^{\frac{\beta}{1-\eta}} \tag{5}$$

Here, we define τ_{ig} as a "gross" tax rate that summarizes the two frictions:

$$\tau_{ig} \equiv \frac{(1+\tau_{ig}^h)^\eta}{1-\tau_{ig}^w}. \tag{6}$$

2.2. Occupational Skills

Turning to the worker's idiosyncratic talent, we borrow from McFadden (1974)'s and Eaton and Kortum (2002)'s formulation of the discrete choice problem. We assume each person gets an iid skill draw ϵ_i from a Fréchet extreme value distribution for each occupation:

$$F_{ig}(\epsilon) = \exp(-T_{ig}\epsilon^{-\theta}). \tag{7}$$

The parameter θ governs the dispersion of skills, with a higher value of θ corresponding to *smaller* dispersion. We assume that θ is common across occupations and groups. The parameter T_{ig}, however, can potentially differ. Across occupations, differences in T's are easy to understand. For example, talent is easy to come by in some occupations and scarce in others. The way we formulate the model, the differences in T's across occupations (for all groups) will be isomorphic to the sector specific productivities which we introduce below. In other words, when we observe very few individuals (across all groups) in a given occupation it could because talent for this occupation is scarce on average or because this occupation is relatively less productive relative to other occupations.

More important for our purposes is the potential that the T's differ across groups within a given occupation. We allow for this possibility between men and women but not between blacks and whites. Specifically, in some occupations, brawn may be a desirable attribute. If men are physically stronger than women on average, then one would expect to observe more men in occupations requiring more physical strength, such as firefighting or construction. To account for this, T_{ig} may be higher in these occupations for white and black men relative to white and black women.

2.3. Occupational choice

The occupational choice problem reduces to picking the occupation that delivers the highest value of U_{ig}. The assumption that the talent draws are iid and come from an extreme value distribution delivers the result that the highest utility can also be characterized by an extreme value distribution, a result reminiscent of McFadden (1974). The overall occupational share can then be obtained by aggregating the optimal choice across people, as we show in the next proposition.

Proposition 1 (Occupational Choice): *Let p_{ig} denote the fraction of people in group g that work in occupation i. Aggregating across people, the solution to the individual's occupational choice problem leads to*

$$p_{ig} = \frac{\tilde{w}_{ig}^{\theta}}{\sum_{s=1}^{N} \tilde{w}_{sg}^{\theta}} \quad \text{where} \quad \tilde{w}_{ig} \equiv \frac{T_{ig}^{1/\theta} w_i \bar{h}_{ig} s_i^{\phi_i} (1-s_i)^{\frac{1-\eta}{\beta}}}{\tau_{ig}}. \tag{8}$$

Equation (8) says that occupational sorting depends on $\tilde{w}_{ig}$, which is the overall reward that someone from group g with the mean talent obtains by working in occupation i. This reward depends on mean talent, the post-friction wage per efficiency unit in the occupation, and the amount of time spent accumulating human capital by a person in that occupation. (Notice that human capital enters twice, once as a direct effect on efficiency units and once indirectly, capturing the fact that a person who gets a lot of education has lower leisure.) Technological changes affect occupational choice through the price per unit of skill, w_i. For example, technological innovations in the home sector emphasized by Greenwood, Seshadri and Yorukoglu (2005) can be viewed as a decline in w_i in the home sector.

Given the above, the sorting model generates the average quality of workers in an occupation for each group. We show this in the following proposition:

Proposition 2 (Average Quality of Workers): *For a given group, the average quality of workers in each occupation, including both human capital and talent, is*

$$\mathbb{E}\left[h_i\epsilon_i\right] = \gamma \left[\eta^\eta \bar{h}_{ig} s_i^{\phi_i} \left(\frac{w_i(1-\tau_i^w)}{1+\tau_i^h}\right)^\eta \left(\frac{T_i}{p_{ig}}\right)^{\frac{1}{\theta}}\right]^{\frac{1}{1-\eta}} \tag{9}$$

where $\gamma \equiv \Gamma(1 - \frac{1}{\theta} \cdot \frac{1}{1-\eta})$ *is related to the mean of the Fréchet distribution for abilities.*

Notice that average quality is inversely related to the share of the group in the occupation p_{ig}. This captures the selection effect. For example, the model predicts that only the most talented female lawyers (such as Sandra Day O'Connor) would have chosen to be lawyers in 1960. And as the barriers faced by female lawyers declined after 1960, less talented female lawyers moved into the legal profession and thus lowered the average quality of female lawyers.

Next, we compute the average wage for a given group in a given occupation. This term is analogous to the average wage per group in each occupation observed in the data.

Proposition 3 (Occupational Wage Gaps): *Let* $\overline{wage}_{ig}$ *denote the average earnings in occupation* i *by group* g*. Its value satisfies*

$$\overline{wage}_{ig} \equiv (1-\tau_{ig}^w) w_i \, \mathbb{E}\left[h_i\epsilon_i\right] = (1-s_i)^{-1/\beta} \gamma \bar{\eta} \left(\sum_{s=1}^{N} \tilde{w}_{sg}^\theta\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}}. \tag{10}$$

In turn, the occupational wage gap between any two groups is the <u>same</u> *across all occupations. For example,*

$$\frac{\overline{wage}_{ig}}{\overline{wage}_{i,wm}} = \left(\frac{\sum_s \tilde{w}_{sg}^\theta}{\sum_s \tilde{w}_{s,wm}^\theta}\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}}. \tag{11}$$

Equation (10) states that average earnings for a given group only differs across occupations because of the first term, $(1-s_i)^{-1/\beta}$. Occupations in which schooling is especially productive (a high ϕ_i and therefore a high s_i) will have higher average earnings, and that is the only reason for earnings differences across occupations in the model. Average earnings are no higher in occupations where a group faces less discrimination or a better talent pool or a higher wage per efficiency unit. The reason is that each of these factors leads lower quality workers to enter those jobs. This composition effect exactly offsets the direct effect on earnings when the distribution of talent is Fréchet. Because of this selection force, the wage gap between two groups is the *same* for all occupations.

Equation (11) states that the wage gap is only a function of a weighted average of the occupational distortions where the weights are a function of the occupational talent and the price of an efficiency unit of skill in each occupation. In the empirical section, we will decompose the contribution of changes in the occupational distortion vs. changes in the price of occupational skills to the narrowing of the wage gap between blacks and women and white men.

Finally, putting together the equations for the occupational share and the wage gap, we get the following expression for the relative propensity of a group to work in an occupation:

$$\frac{p_{ig}}{p_{i,wm}} = \frac{T_{ig}}{T_{i,wm}} \cdot \left(\frac{\bar{h}_{ig}}{\bar{h}_{i,wm}}\right)^{\theta} \left(\frac{\tau_{ig}}{\tau_{i,wm}}\right)^{-\theta} \left(\frac{\overline{\text{wage}}_g}{\overline{\text{wage}}_{wm}}\right)^{\theta(1-\eta)} \tag{12}$$

Equation (12) states that the propensity of a group to work in an occupation (relative to white men) depends on four terms: the relative mean talent in the occupation (arguably equal to one for many occupations), the relative efficiency in accumulating human capital, the relative occupational friction, and the average wage gap between the groups. From Proposition 2, the wage gap itself is a function of the distortions faced by the group and the prices of skills in *all* occupations. With data on occupational shares and wages, we can measure a composite of the relative mean talent, human capital efficiency, and occupational frictions between groups. This will be the key equation we take to the data.

2.4. Aggregate Productivity

To aggregate the heterogeneous individual outcomes in our model, we assume that a representative firm produces aggregate output Y from labor in N occupations:

$$Y = \left(\sum_{i=1}^{N} (A_i H_i)^{\rho}\right)^{1/\rho} \tag{13}$$

where H_i denotes the total efficiency units of labor and A_i is the exogenously-given productivity in occupation i. In turn, H_i is defined as

$$H_i = \sum_{g=1}^{G} q_g p_{ig} \cdot \ \mathbb{E}\left[h_{ig}\epsilon_{ig} \,|\, \text{Person chooses } i\right]. \tag{14}$$

That is, H_i equals the average human capital of the people who choose occupation i multiplied by the number of people in the occupation (and summed over groups).

That completes the setup of the model. We can now define an equilibrium and then start exploring the model's implications.

2.5. Equilibrium

A competitive equilibrium in this economy consists of individual choices $\{c, e, s\}$, an occupational choice by each person, total efficiency units of labor in each occupation H_i, final output Y, and an efficiency wage w_i in each occupation such that

1. Given an occupational choice, the occupational wage w_i, and idiosyncratic ability ϵ in that occupation, each individual chooses c, e, s to maximize utility:

$$U(\tau^w, \tau^h, w, \epsilon) = \max_{c,e,s} \, (1-s)c^{\beta} \ \ s.t. \ \ c = (1-\tau^w_{ig})w\epsilon h(e,s) - e(1+\tau^h_{ig}). \tag{15}$$

2. Each individual chooses the occupation that maximizes his or her utility: $i^* = \arg\max_i U(\tau^w_{ig}, \tau^h_{ig}, w_i, \epsilon_i)$, taking $\{\tau^w_{ig}, \tau^h_{ig}, w_i, \epsilon_i\}$ as given.

3. A representative firm chooses labor input in each occupation, H_i, to maximize

profits:

$$\max_{\{H_i\}} \left(\sum_{i=1}^{N}(A_i H_i)^{\rho}\right)^{1/\rho} - \sum_{i=1}^{N} w_i H_i \tag{16}$$

4. The occupational wage w_i clears the labor market for each occupation:

$$H_i = \sum_{g=1}^{G} q_g p_{ig} \cdot \mathbb{E}\left[h_{ig}\epsilon_{ig} \,|\, \text{Person chooses } i\right]. \tag{17}$$

5. Total output is given by the production function in equation (13).

The equations characterizing the general equilibrium are then given in the next result.

Proposition 4 (Solving the General Equilibrium): *The general equilibrium of the model is* $\{p_{ig}, H_i^{supply}, H_i^{demand}, w_i\}$ *and* Y *such that*

1. p_{ig} *satisfies equation* (8).
2. H_i^{supply} *aggregates the individual choices:*

$$H_i^{supply} = \gamma\bar{\eta} w_i^{\theta-1}(1-s_i)^{(\theta(1-\eta)-1)/\beta} s_i^{\theta\phi_i} \sum_g q_g T_{ig} \bar{h}_{ig}^{\theta} \frac{(1-\tau_{ig}^{w})^{\theta-1}}{(1+\tau_{ig}^{h})^{\eta\theta}} \left(\sum_{s=1}^{N} \tilde{w}_{sg}^{\theta}\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}-1} \tag{18}$$

3. H_i^{demand} *satisfies firm profit maximization:*

$$H_i^{demand} = \left(\frac{A_i^{\rho}}{w_i}\right)^{\frac{1}{1-\rho}} Y \tag{19}$$

4. w_i *clears each occupational labor market:* $H_i^{supply} = H_i^{demand}$.
5. *Total output is given by the production function in equation* (13).

2.6. Intuition

To develop intuition, consider the following simplified version of the model. First, assume there are only two groups, men and women. Occupational choice is distorted for women but not for men ($\tau_{ig}^{w} = \tau_{ig}^{h} = 0$ for men). Second, assume $\rho = 0$

(occupations are perfect substitutes) so that $w_i = A_i$, i.e. the production technology parameter pins down the wage per unit of human capital in each occupation. When this is the case, the τ_{ig}'s affect the average wage and occupational choices of group g, but will have no effect on the other groups. Third, assume $\phi_i = 0$ (no schooling time), $T_{ig} = 1$ (mean occupational talent is the same for every group), and $\overline{h}_{ig} = 1$ (groups are equally efficient at building human capital). Aggregate output is then equal to the sum of wages paid to men and to women plus the revenues from the labor market friction:

$$Y = q_m \cdot \overline{wage}_m + q_f \cdot (1 + \overline{\tau}^w) \cdot \overline{wage}_f \tag{20}$$

where $\overline{\tau}^w \equiv \sum_{i=1}^N p_{iw}\tau_i^w$ denotes the employment-weighted average of the labor market friction facing women. The average wage of men and women, respectively, are given by:

$$\overline{wage}_m = \left(\sum_{i=1}^N A_i^\theta\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}} \tag{21}$$

$$\overline{wage}_f = \left(\sum_{i=1}^N \left(\frac{A_i\,(1-\tau_i^w)}{(1+\tau_i^h)^\eta}\right)^\theta\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}} \tag{22}$$

The average male wage is a power mean of the occupational productivity terms (the A_i's) and is not affected by the occupational distortions facing women (this is driven by the assumption that $\rho = 0$). The average wage of women is a power mean of the occupational productivities and distortions.

To see the effect of the distortions on aggregate output, consider first the case where the only distortion is in the labor market ($\tau^h = 0$). Furthermore, assume that $1 - \tau_i^w$ and A_i are jointly log-normally distributed. The average female wage is then equal to:

$$\ln \overline{wage}_f = \ln\left(\sum_{i=1}^N A_i^\theta\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}} + \frac{1}{1-\eta}\cdot\ln\left(1-\overline{\tau}^w\right) - \frac{\theta+1}{2}\cdot\frac{1}{1-\eta}\cdot \text{Var}(\ln(1-\tau_i^w)). \tag{23}$$

The first term says that, just as for the wages of men, the average female wage is increasing in the power mean of occupational productivities. The second term states that the average female wage is decreasing in the weighted average of the labor market frictions, and more so the greater the importance of acquired human capital (i.e., the higher is η). The third term says that the average female wage is decreasing in the dispersion of $1-\tau^w$ with an elasticity that is increasing in the substitutability of skills (θ) and the elasticity of human capital to school spending (η).

Now consider the case where the only distortion is in the human capital market ($\tau^w = 0$). This time, assuming that $1-\tau_i^h$ and A_i are jointly log-normally distributed, the average female wage is equal to:

$$\ln \overline{wage}_f = \ln \left(\sum_{i=1}^{N} A_i^\theta \right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}} + \frac{\eta}{1-\eta} \cdot \ln\left(1-\overline{\tau}^h\right) - \frac{\theta\eta+1}{2} \cdot \frac{\eta}{1-\eta} \cdot \text{Var}(\ln(1-\tau_i^h)). \tag{24}$$

Here $\overline{\tau}^h \equiv \sum_{i=1}^{N} \frac{E_{iw}}{E_w} \tau_i^h$ is a schooling-expenditure-weighted average of the human capital distortion. When $\eta = 0$, there is no schooling expenditure for τ^h to distort. But when η is positive, the average female wage is decreasing in both the weighted average and dispersion of τ_h.

To recap, both the mean and dispersion of τ_h and τ_w reduce productivity and lower the average female wage (in addition to the direct effect of τ_w on women's wages). The productivity losses come from two sources: underinvestment in human capital (tied to weighted mean levels of τ_h and τ_w and increasing in η) and misallocation of female talent across occupations (tied to the dispersion of τ_h and τ_w across occupations, and increasing in η and θ). A higher average distortion has no effect on the allocation of female labor supply across occupations.

3. Empirically Evaluating the Occupational Sorting Model

3.1. Data

We use data from the 1960, 1970, 1980, 1990, and 2000 Decennial Censuses as well data from the 2006-2008 American Community Surveys (ACS) for all analysis in the paper. When using the 2006-2008 ACS data, we pool all three years together and treat them as one cross section.[6] We make only four restrictions to the raw data when constructing our analysis samples. First, we restrict the analysis to include only white men (wm), white women (ww), black men (bm) and black women (bw). These will be the four groups we analyze in the paper.[7] Second, we restrict the sample to include only individuals between the ages of 25 and 55 (inclusive). This restriction helps to focus our analysis on individuals after they finish schooling and prior to considering retirement. Third, we exclude individuals on active military duty. Finally, we exclude individuals who report their labor market status as being unemployed (not working but searching for work). Our model is not well suited to capture transitory movements into and out of employment. Appendix Table A1 reports the sample size for each of our six cross sections, including the fraction of the sample comprised of our four groups.[8]

A key to our analysis is to use the Census data to create a consistent set of occupations over time. We treat the home sector as a separate occupation. Anyone in our data who is not currently employed or who is employed but usually works less than ten hours per week is considered to be working exclusively in the home sector. Those who are employed but usually work between ten and thirty hours per week are classified as being part-time workers. We split the sampling weight of part-time workers equally between the home sector and the occupation in which they are working. Individuals working more than thirty hours per week are considered to be working full-time in an occupation outside of the home sector.

For our base analysis, we define the non-home occupations using the roughly 70 occupational sub-headings from the 1990 Census occupational classification system.[9] We use the 1990 occupation codes as the basis for our occupational defini-

[6]Henceforth, we refer to the pooled 2006-2008 sample as the 2008 sample. A full description of how we process the data, including all the relevant code, is available at http://faculty.chicagobooth.edu/erik.hurst/research/chad_data.html.

[7]We think an interesting extension would be to include Hispanics in the analysis. In 1960 and 1970, however, there are not enough Hispanics in the data to provide reliable estimates of occupational sorting. Such an analysis can be performed starting in 1980. We leave such an extension to future work.

[8]For all analysis in the paper, we weight our data using the sample weights available in the different surveys.

[9]http://usa.ipums.org/usa/volii/99occup.shtml.

tions because the 1990 occupation codes are available in all Census and ACS years since 1960. We start our analysis in 1960, as this is the earliest year for which the 1990 occupational crosswalk is available. Appendix Table A2 reports the 67 occupations we analyze in our main specification using the 1990 occupational subheadings. Example occupations include "Executives, Administrators, and Managers", "Engineers", "Natural Scientists", "Health Diagnostics", "Health Assessment", and "Lawyers and Judges". Appendix Table B.3 gives a more detailed description of some of these occupational categories. For example, the "Health Diagnostics" occupation includes physicians, dentists, veterinarians, optometrists, and podiatrists, and the "Health Assessment and Treating" occupations include registered nurses, pharmacists, and dieticians. For shorthand, we sometimes refer to these occupations as doctors and nurses, respectively. The way the occupations are defined ensures that each has positive mass in all years.

As seen with the examples above, there is some heterogeneity within our 67 base occupational categories. To assess the importance of such heterogeneity, we later perform robustness exercises using different levels of occupational aggregation. Specifically, we use the roughly 340 detailed 1990 occupation codes that are consistently available in 1980, 1990, 2000, and 2008. We start this in 1980 because the occupational classification system is roughly similar across the Censuses and ACS starting in 1980. We perform our main analysis using the 340 detailed occupation codes for the 1980–2008 period and show that the quantitative outcomes are very similar to what we get using our 67 base occupation codes for the same period. Additionally, we show quantitative robustness to using only 20 broad occupation categories as opposed to the roughly 67 occupation codes in our base analysis. The 20 occupation categories we use for this robustness analysis are shown in Appendix Table A4. As we show throughout, our key empirical results come from the fact that women and blacks are converging toward white male propensities of working in a handful of high skilled occupations.

Our measure of earnings throughout the paper sums together each individual's labor, business, and farm income. Incomes in the Census are from the prior year. Implicitly we assume that individuals who report working in a given occupation in the survey year also worked in that same occupation during the prior year which corresponds to their income report. When measuring earnings, we only focus on those individuals who worked at least 48 weeks during the prior year and who had at least 1000 dollars of earnings (in year 2007 dollars). We define the wage rate by dividing individual earnings from the prior year by the product of weeks worked during the prior year and the reported current usual hours worked. When computing individual wage measures, we further restrict the sample to those who usually work more than 30 hours per week.[10]

For a few of our empirical results, we need a measure of average wages in the home sector. We impute average earnings for the home sector by extrapolating the relationship between average education and average earnings for the 66 non-home occupations, taking into account group fixed effects. Using this year-specific relationship by group and the actual year-specific average education and group composition of participants in the home sector, we predict the average earnings of participants in the home sector. In our empirical work we carry out robustness checks to see if our results are sensitive to our imputation procedure for home sector wages. For example, we assess how our results change if we set home sector wages one standard deviation higher or lower in each year (based on the standard error around our imputed value). Our estimated productivity gains are quantitatively similar when using these alternative imputations for wages in the home sector.

3.2. Occupational Sorting and Wage Gaps By Group

We begin our analysis by documenting the degree of convergence in the occupational distribution between white men and the other groups over the last fifty years. To illustrate, we create a simple occupational similarity index Ψ_g, defined as:

[10] In some Census years, weeks worked during the prior year and usual hours worked are reported as categorical variables. In these instances, we use the midpoint of the range when computing the wage rate. See the full details of our data processing in the detailed online data appendix available on the authors' web sites.

$$\Psi_g \equiv 1 - \frac{1}{2}\sum_{i=1}^{N} |p_{i,wm} - p_{ig}| \tag{25}$$

We first compute the absolute value of the difference in the propensity for white men vs. a given group to be in an occupation. We then sum these differences across all occupations for the given group. For ease of interpretation, we normalize the measure so that it runs from zero (no occupational overlap between the two groups) to 1 (identical occupational propensities for the two groups). When computing Ψ_g, we exclude the home sector. However, the broad patterns are very similar — particularly the index for white women — when the home sector is included.

Panel A of Table 1 shows the measure of Ψ_g for white women, black men, and black women in 1960, 1980, and 2008. We also compare less educated individuals (those with a high school degree or less) with higher educated individuals (those with more than a high school degree). Within the educational categories, for example, we compare the occupational distribution of lower educated white women to the occupational distribution of lower educated white men.

A few things of note from Panel A of Table 1. First, each group experienced substantial occupational convergence relative to white men between 1960 and 2008. Second, the timing of the convergence occurred differentially across the groups. For example, occupational convergence occurred both from 1960–1980 and 1980–2008 period for white women and black women. For black men, however, the bulk of the convergence occurred prior to 1980. Third, there are differences in the occupational convergence by educational attainment. This is seen particularly for white women. In 1960, there were substantial occupational differences both between high educated white women and high educated white men, and between low educated white women and low educated white men. Low educated white men worked primarily in construction and manufacturing, while low educated white women worked primarily as secretaries or in low skilled services like food service. High educated white men in 1960 were spread out across many high skilled occupations, while high educated white women worked primarily as teachers and nurses. Between 1960 and

Table 1: Occupational Similarity and Conditional Wage Gaps Relative to White Men

	Panel A: Occupational Similarity Index, Relative to White Men				
Race-Sex Group	1960	1980	2008	1980-1960 Difference	2008-1980 Difference
White Women: All	0.42	0.49	0.55	0.07	0.06
White Women: High Educated	0.38	0.49	0.59	0.10	0.12
White Women: Low Educated	0.40	0.43	0.46	0.02	0.04
Black Men: All	0.56	0.72	0.76	0.16	0.04
Black Men: High Educated	0.51	0.74	0.77	0.23	0.03
Black Men: Low Educated	0.59	0.75	0.75	0.16	0.00
Black Women: All	0.28	0.42	0.50	0.14	0.08
Black Women: High Educated	0.31	0.44	0.53	0.13	0.11
Black Women: Low Educated	0.27	0.41	0.44	0.14	0.03

	Panel B: Conditional Log Difference in Wages, Relative to White Men				
Race-Sex Group	1960	1980	2008	1980-1960 Difference	2008-1980 Difference
White Women: All	-0.57	-0.47	-0.26	0.10	0.21
White Women: High Educated	-0.50	-0.40	-0.24	0.10	0.16
White Women: Low Educated	-0.56	-0.47	-0.27	0.09	0.20
Black Men: All	-0.38	-0.22	-0.15	0.16	0.07
Black Men: High Educated	-0.29	-0.16	-0.18	0.13	-0.02
Black Men: Low Educated	-0.38	-0.23	-0.12	0.15	0.11
Black Women: All	-0.86	-0.48	-0.31	0.38	0.17
Black Women: High Educated	-0.62	-0.39	-0.31	0.23	0.08
Black Women: Low Educated	-0.88	-0.48	-0.28	0.40	0.20

Note: Panel A of the table reports our occupational similarity index for white women, black men, and black women relative to white men in 1960, 1980, and 2008. The occupational similarity index runs from zero (no overlap with the occupational distribution for white men) and one (identical occupational distribution to white men). The index is also computed separately for higher educated and lower educated individuals in the different groups. Panel B reports the difference in log wages between the groups and white men. The entries come from a regression of log wages on group dummies and controls for potential experience and hours worked per week. The regression only includes a sample of individuals working full time.

2008, however, the occupational similarity between higher educated white men and women converged dramatically, while the occupational similarity between lower educated white men and women barely changed. Today, low skilled women still primarily work in services and office support occupations, while low skilled men still primarily work in construction and manufacturing.

One of the strong predictions of our occupational sorting model is that wage gaps relative to white men should be the same across all occupations for a given group. The reason for this is that an occupation that pays a high wage per unit of ability will attract less talented workers. As discussed above, this sorting makes the wage gap in a given occupation a poor guide to any frictions or absolute advantage in that occupation. There are, however, at least three reasons why the empirical wage gaps between groups may not be equated across all occupations. First, there is likely some measurement error in the occupational wage gap estimates due to small sample sizes, especially for some groups in some occupations. Second, although we expect sorting will help offset the effect of differences in wages per ability on the average wage in an occupation, the exact offset due to sorting is a feature of the extreme value distribution. If ability is not drawn from an extreme value distribution, the wage gaps need not be perfectly equated across occupations. Third, we focus on occupational sorting due to heterogeneity in ability, but some of the occupational sorting might be driven by other factors such as heterogeneity in preferences. High wage (per unit of ability) occupations might induce the entry of people with high disutility for an occupation rather than individuals with low ability in the occupation. All three forces will generate variation in wage gaps across occupations.

Panel B of Table 1 shows the estimated wage gap between white men and, respectively, white women, black men, and black women over time and by educational attainment. To obtain these estimates, we regress log wages of the individual on group dummies, a quadratic in potential experience, a polynomial in usual hours worked, and our base specification occupation dummies. This regression is estimated only for those individuals who are currently working more than 30 hours per week and who worked at least 48 weeks during the prior year when earnings were measured. We estimated this regression separately for 1960, 1980, and 2008.

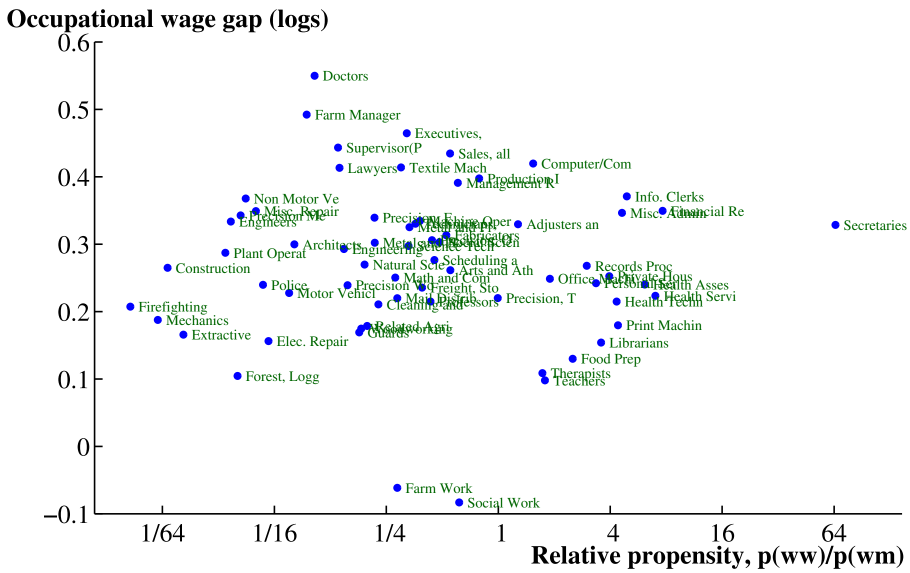
The coefficients on the race-sex dummies are shown in the table and should be interpreted as log deviations relative to white men. We also estimated the regression separately for individuals with 12 years or less of schooling and for individuals with more than 12 years of schooling.

As seen in panel B of Table 1, the wage gap for white women relative to white men is similar for those with more vs. less education in 1960, 1980, and 2008. For example, in 1960, low educated white women earned a wage that was 56 log points lower than low educated white men. The comparable number for high educated white women relative to high educated white men was 50 log points. Between 1960 and 2008, the relative wage of low educated white women narrowed by 29 log points. During the same time period, the relative wage of high educated white women narrowed by 26 log points. Despite the very different trends in occupational similarity by educational attainment for white women (as seen in Panel A), the change in the wage gap was nearly identical by educational attainment for white women. According to our model, changes in the $\hat{\tau}_{ig}$'s for white women in high ϕ occupations would generate exactly this result.

For black men, the wage gap evolved similarly for less vs. more educated individuals from 1960 to 1980. After 1980, however, there was little change in occupational similarity for either high or low skilled black men, and there was no change in the wage gap for high skilled black men. The wage gap for low skilled black men, however, continued to narrow after 1980. This may be due to the rapid decline in labor market participation of low skilled black men during the last thirty years, if it was not random. As currently formulated, our model would not predict such a trend. However, as we discuss in Section 5, the change in labor market outcomes for black men between 1980 and 2008 do not materially affect our estimates of aggregate productivity gains.

A further test of the plausibility of our framework is to examine, occupation by occupation, whether the change in the wage gap is related to the change in relative propensities across occupations. Our model suggests that the two should be unrelated, with no variation in the wage gaps. For white women vs. white men in 1980, Figure 1 plots the (log) wage gap in an occupation against the relative propen-

Figure 1: Occupational Wage Gaps for White Women in 1980



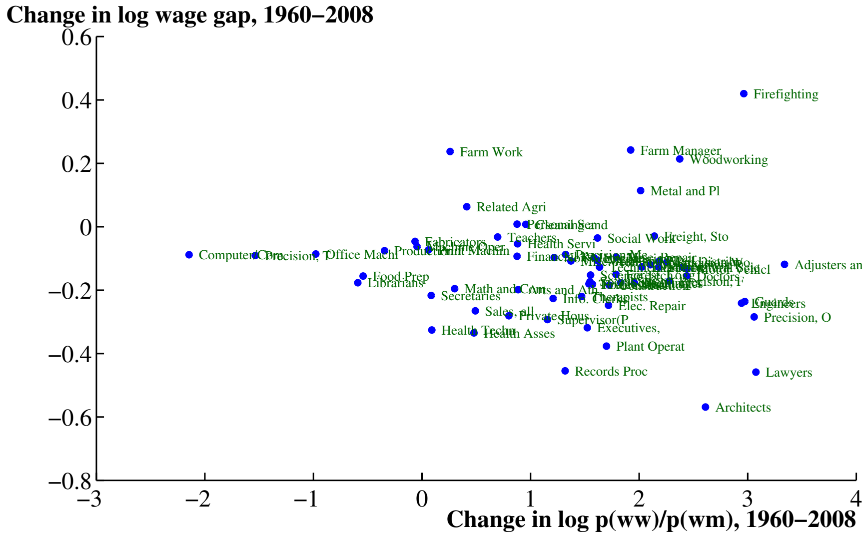
Note: The figure shows the relationship between the (log) occupational wage gap for white women compared to white men and the relative propensity to work in the occupation between white women and white men, $p_{i,ww}/p_{i,wm}$.

sity to work in that occupation $p_{i,ww}/p_{i,wm}$. As an example, in 1980, a white woman was 65 times more likely than a white man to work as a secretary, but only 1/7 as likely to work as a lawyer. Given this enormous variation, the difference in the wage gaps between these two occupations seems remarkably small. White women secretaries earned about 33 percent less than white men secretaries in 1980, while the gap was 41 percent for lawyers. Fitting a regression line through the points in the figure would reveal no systematic relationship between the wage gap and the relative propensity in that occupation. [11] The patterns in other years and for other groups were quite similar.[12] From the perspective of the model, the weak relationship be-

[11] The coefficient from a regression of the occupational wage gap on log $p_{i,ww}/p_{i,wm}$ was 0.002 with a standard error of 0.008 and an adjusted R-squared of essentially zero. For interpretation, the standard deviation of the independent variable was 1.96 and the mean of the dependent variable was -0.31. The regression weighted occupations by the share of all workers (across all groups) in the occupation.

[12] In additional work, we explored the relationship between occupational wage gaps and the average earnings of individuals in those occupations. On average, high income occupations tended to have larger wage gaps. Nonetheless, the magnitude of this correlation was almost always small. For example, in 2006-2008, white working women had about a 3 percentage point larger wage gap relative to white men in response to a one-standard deviation increase in occupational log income. As seen

Figure 2: Change in Occupational Wage Gaps for White Women,1960–2008



Note: The figure shows the relationship between the change in (log) occupational wage gap for white women compared to white men between 1960 and 2008 and the change in the relative propensity to work in the occupation between white women and white men, $p_{i,ww}/p_{i,wm}$, over the same time period.

tween wages gaps and propensities is not surprising. Within the model, the relative propensity, not the wage gap, reveals frictions facing a group in that occupation.

Our productivity gains in the subsequent sections are based on the change in the occupational distribution over time. Figure 2 shows that the change in log $p_{i,ww}/p_{i,wm}$ between 1960 and 2008 is also uncorrelated with the change in the wage gap between white women and white men from 1960 to 2008. The relative fraction of white women who are doctors increased by 144 percent between 1960 and 2008. For nurses, in contrast, the relative fraction who are white women decreased by 52 percent. Yet the relative wage gap between white men and white women narrowed by 20 to 30 log points in both occupations. Our model predicts that the change in the wage gaps should be uncorrelated with the change in the occupational sorting.

from Table 1, the average wage gap was 26 percentage points.

4. Estimating the Frictions

Motivated by our model, we now use data on the difference in occupational propensities across groups as well as the average wage gaps to infer a composite measure of the occupation-specific frictions. Specifically, given equations (12) and (11), we can define the composite friction measure for each group (relative to white men) in each occupation as:

$$\hat{\tau}_{ig} \equiv \frac{\tau_{ig}}{\tau_{i,wm}} \cdot \frac{\bar{h}_{ig}}{\bar{h}_{i,wm}} \left(\frac{T_{i,wm}}{T_{i,g}}\right)^{\frac{1}{\theta}} = \left(\frac{p_{ig}}{p_{i,wm}}\right)^{-\frac{1}{\theta}} \left(\frac{\overline{\text{wage}}_{wm}}{\overline{\text{wage}}_{g}}\right)^{1-\eta}. \tag{26}$$

This equation has the following interpretation. If a group is either underrepresented in an occupation or if it faces a large average wage gap, the right-hand side of this equation will be high. The model can explain this in one of two ways (on the left side): either the group faces a large composite barrier, or it has a relatively low mean talent in that occupation (e.g. women in occupations where brute strength is important). We observe the right-hand side of this equation in the data and therefore use it to back out the average relative distortion or talent between groups, $\hat{\tau}_{ig}$.

To implement this calculation, we require estimates of θ and η. The parameter θ is a key parameter that governs the dispersion of wages. Given the occupational choice model developed above, one can show that the dispersion of wages across people within an occupation-group obeys a Fréchet distribution with the shape parameter $\theta(1-\eta)$: the lower is this shape parameter, the *more* wage dispersion there is within an occupation. Wage dispersion therefore depends on the dispersion of talent (governed by $1/\theta$) and amplification from accumulating human capital via spending (governed by $1/(1-\eta)$). In particular, the coefficient of variation of wages within an occupation-group in our model satisfies:

$$\frac{\text{Variance}}{\text{Mean}^2} = \frac{\Gamma(1-\frac{2}{\theta(1-\eta)})}{\left(\Gamma(1-\frac{1}{\theta(1-\eta)})\right)^2} - 1. \tag{27}$$

To estimate $\theta(1-\eta)$ in a given year, we need to know the dispersion of ability for a given occupation. As a starting point, we look at wage dispersion within occupation-groups. We take residuals from a cross-sectional regression of log worker wages on 66x4 occupation-group dummies. These span the 66 occupations (excluding Home) and the four groups of white men, white women, black men, and black women. The wage is the hourly wage, and the sample includes both full-time and part-time workers. The dummies should capture the impact of schooling requirements (ϕ_i levels) on average wages in an occupation, and the wage gaps created by frictions (the average τ_{ig} across occupations for each group). We calculate the mean and variance across workers of the exponent of these wage residuals. We then solve equation (27) for the value of $\theta(1-\eta)$. Sampling error is trivial here because there are 300-400k observations per year for 1960 and 1970 and 2-3 million per year for 1980 onward. The point estimates for $\theta(1-\eta)$ average 3.12. They drift down over time, from around 3.3 in 1960 to 2.9 in 2006-2008, as one would expect given rising wage inequality.

We are concerned that this way of estimating θ wrongly attributes all of the dispersion of wages within occupation-groups to comparative advantage. We thus make several adjustments, all of which serve to reduce residual wage dispersion. First, we compress the variance of the residuals by 14% and 4%, respectively, to reflect estimates of transitory wage movements from Guvenen and Kuruscu (2009) and how much wage variation can be explained by AFQT scores from Rodgers and Spriggs (1996). Temporary wage differences across workers are not a source of enduring comparative advantage, and AFQT score is arguably correlated with absolute ability across many occupations. Second, we controlled directly for individual education, hours worked, and potential experience in the Census data. Like AFQT score, a worker's education might be correlated with absolute advantage across many occupations. Though we examine the hourly wage, there could be compensating differentials associated with the workweek. And experience reflects the life cycle rather than lifelong comparative advantage.

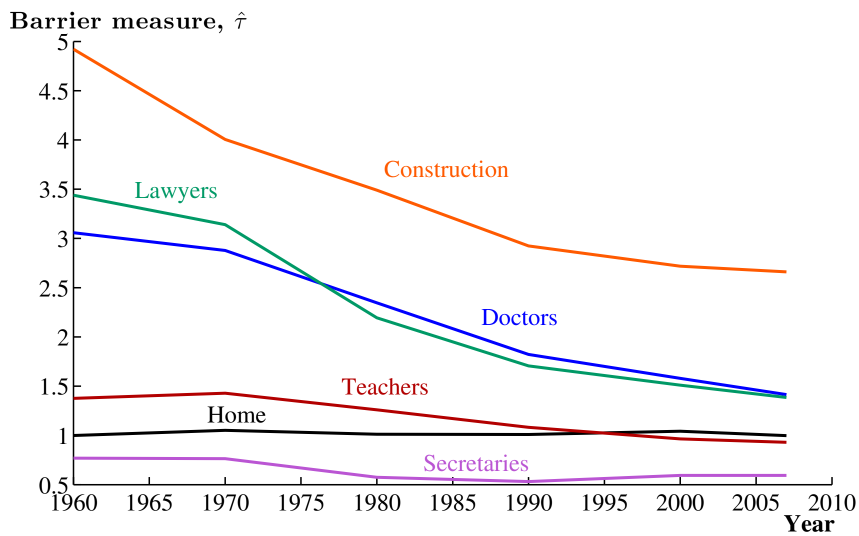
These adjustments cumulatively explain 25% of wage variation within occupation-groups. Attributing the remaining 75% of wage dispersion to comparative advantage, we arrive at a baseline value of $\theta(1-\eta) = 3.44$. Given that $\theta(1-\eta)$ is important when thinking about the productivity gains of labor market misallocation across groups, we explore the sensitivity of our results to alternate values. When computing our counterfactuals in Section 5, we also show results where we set $\theta(1-\eta)$ such that only 50%, 25% or 10% of the within occupation-group wage dispersion is attributable to comparative advantage.

The parameter η denotes the elasticity of human capital with respect to education spending. Related parameters have been discussed in the literature, for example by Manuelli and Seshadri (2005) and Erosa, Koreshkova and Restuccia (2010). In our model, η will equal the fraction of output spent on accumulating human capital in equilibrium, separate from time spent accumulating human capital. Absent any solid evidence on this parameter, we set $\eta = 1/4$ in our baseline and explore robustness to $\eta = 0$ and $\eta = 1/2$. In general, this parameter slightly affects the *level* of the τ_{ig} parameters, but not much else in the results.

Figure 3 presents our estimates of $\hat{\tau}_{ig}$ for white women for a select subset of our baseline occupations. First, we highlight the $\hat{\tau}_{ig}$ for the home sector. Second, we highlight high educated occupations for which white women were underrepresented in 1960. These occupations include doctors and lawyers. Third, we highlight high educated occupations like teachers for which white women were overrepresented relative to white men in 1960. Finally, we show the patterns for construction which is an occupation dominated by lower skilled male workers in 1960.

Many interesting patterns emerge from Figure 3. Consider the results for white women in the "home" occupation in 1960. Despite white women being 7 times as likely to work in the home sector as white men, we estimate $\hat{\tau}_{ig}$ for white women in the home sector to be just below 1 (0.99). This implies that white women in 1960 did *not* have an absolute advantage over white men in the home sector. Two factors underlie our estimate of $\hat{\tau}_{ig}$ being close to 1 in the home sector for white women. First, we are estimating that white women were choosing the home sector because they were facing disadvantages in *other* occupations. Those barriers show up in the average wage gap between white women and white men. Given that white women earned roughly 57 percent less when working than white men, our model predicts that women should be much more likely to work in the home sector relative to white

Figure 3: Estimated Barriers ($\hat{\tau}_{ig}$) for White Women



Note: Author's calculations based on equation (26) using Census data and imposing $\theta = 3.44$ and $\eta = 1/4$.

men all else equal. Second, how much more white women should be working in the home sector if the other sectors are less attractive for white women depends on θ. As the skill distribution becomes less dispersed (θ increasing), frictions in other sectors will push more women into the home sector. The reason for this is that the comparative advantage in a given occupation relative to another occupation gets stronger when θ is higher. Given our estimate of θ, the observed wage gap between white men and white women, and the relative propensity of each group to be in the home sector, we estimate that $\hat{\tau}_{ig}$ is roughly one for white women in the home sector.

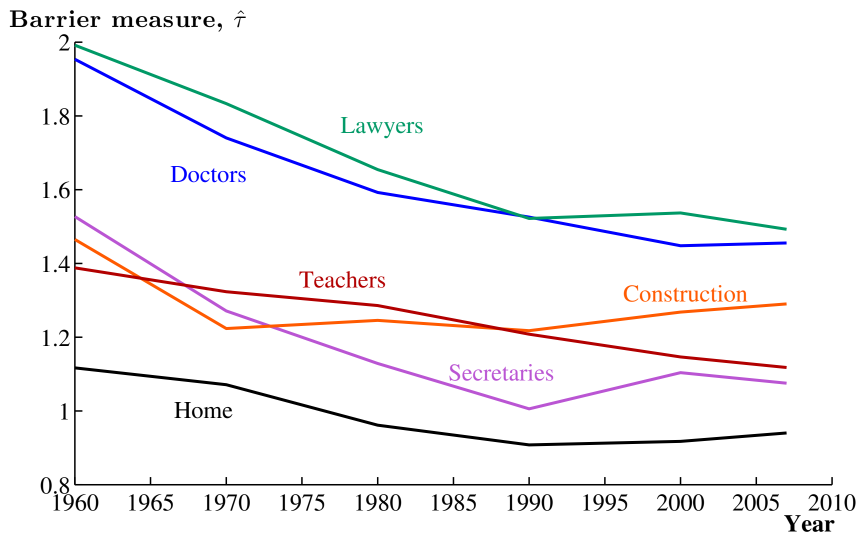
As seen from Figure 3, $\hat{\tau}_{ig}$ is also close to 1 for white women in the home sector in all years of our analysis. This suggests that women did not move out of the home sector because they lost any absolute advantage in the home sector. Instead, our results suggest that women moved into market occupations due to declining barriers in the market. Below we will show that changes in the productivity of the home sector relative to the market sector for all groups also contributed to women exiting the home sector over this time period. In order to quantify this effect, we need the general equilibrium analysis formulated in the next section. To preview our results, we find that the changing productivity of the home sector relative to the market sector explains roughly 70 percent of the movement of white women out of the home sector. The remaining 30 percent is due to changes in the $\hat{\tau}_{ig}$ in the market sector.

The remainder of the results from Figure 3 reinforce that the $\hat{\tau}_{ig}$'s for white women changed dramatically over time in certain occupations. For example, our estimates of $\hat{\tau}_{ig}$ for lawyers and doctors for white women in 1960 ranged from 3.0 to 3.5. In terms of our sorting model, the low relative participation of white women in these occupations in 1960 causes us to infer high values of $\hat{\tau}_{ig}$. If men and women draw from the same distribution of skill to be a lawyer or a doctor, the model will attribute the low propensity for white women to work in these occupations as reflecting barriers (either in the human capital market or the labor market directly). Interestingly, the $\hat{\tau}_{ig}$ for white women teachers is also greater than one in 1960. While white women were 1.7 times more likely than white men to work as teachers, this propensity is more than offset by the overall wage gap in 1960, where women earned about 0.57 times what men earned. If white women were not facing some friction or lower absolute advantage in the teacher occupation, our model predicts there should have been an even higher fraction of white women ending up as teachers in 1960.

Contrast this with secretaries in 1960. A white woman in 1960 was 24 times more likely to work as a secretary than was a white man. The model can only explain this enormous discrepancy by assigning a $\hat{\tau}_{ig}$ of 0.76 for white women secretaries. A τ below 1 is like a subsidy, so the model says either white women had an absolute advantage relative over white men as secretaries, or there was discrimination against white men secretaries. Also in 1960, white women had very high $\hat{\tau}_{ig}$ values in the construction, firefighting and vehicle mechanic professions.

For lawyers, and doctors, the $\hat{\tau}_{ig}$'s fell from around be around 3.0 to being around 1.4 between 1960 and 2008. School teachers also saw a substantial fall in their average $\hat{\tau}_{ig}$ from 1.37 to a value slightly below 1. While barriers facing white women fell in many skilled professional occupations, the $\hat{\tau}_{ig}$ values did not change much for low skilled occupations. This is particularly true after 1980. For example, the

Figure 4: Estimated Barriers ($\hat{\tau}_{ig}$) for Black Men



Note: Authors' calculations based on equation (26) using Census data and baseline parameter values.

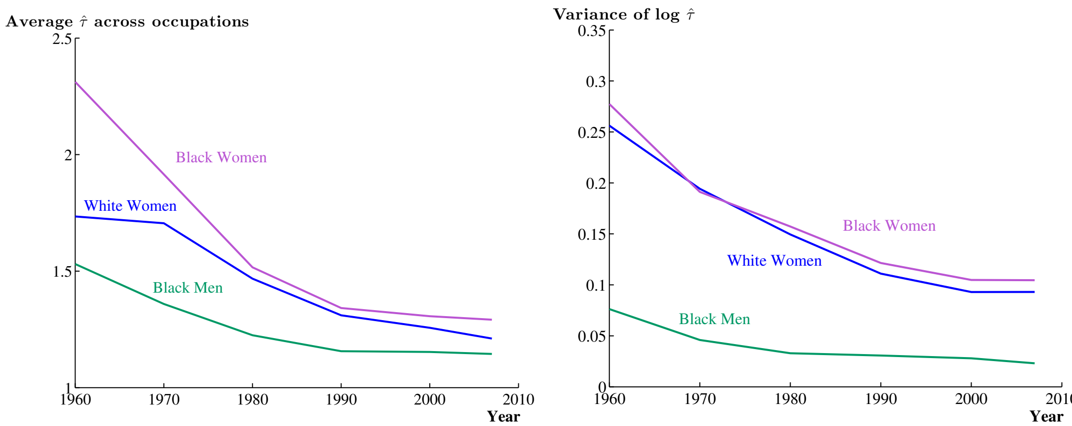
estimated $\hat{\tau}_{ig}$ for white women barely changed (or rose) for secretaries or construction workers between 1980 and 2008. Yet, the $\hat{\tau}_{ig}$'s for doctors, lawyers, and teachers continued to fall during this time period. These results are consistent with the results above showing that occupational convergence during the 1980 - 2000 period was primarily among high skilled individuals.

The $\hat{\tau}_{ig}$'s for black men — for these same select occupations — are shown in Figure 4. A similar overall pattern emerges, with the $\hat{\tau}_{ig}$'s being substantially above 1 in general in 1960, but falling through 2008. Still, they remained above 1 by 2008, especially for the high-skilled occupations, suggesting barriers remain. Unlike for white women, almost the entire change in the $\hat{\tau}_{ig}$ for black men occurred prior to 1980. The plots for black women look like a combination of those for white women and black men.

Figure 5 presents mean and variance of $\hat{\tau}_{ig}$ across occupations for each group over time.[13] The left panel shows the average $\hat{\tau}_{ig}$ falling over time for each of the

[13] When showing the mean and standard deviations of the $\hat{\tau}_{ig}$'s, we weight each occupation by their share of earnings in that occupation out of the total wage bill.

Figure 5: Means and Variances of $\hat{\tau}_{ig}$ Over Time



Note: The left panel shows the average level of $\hat{\tau}$ by group, weighted by total earnings in each occupation. The right panel shows the variance of $\log\hat{\tau}$, weighted in the same way.

groups. For women, the decline in average $\hat{\tau}_{ig}$ occurred continuously during the period. For black men, however, the decline was concentrated prior to 1980. The right panel of Figure 5 shows that, in 1960, the $\hat{\tau}_{ig}$'s were also dispersed across occupations for blacks and (especially) women. As we show below, it is this dispersion that can lead to a misallocation of talent across occupations. If there were no dispersion in the $\hat{\tau}_{ig}$'s across occupations for each group, there would be no misallocation of talent. All groups would have the same occupational distributions. The dispersion in the $\hat{\tau}_{ig}$'s leads to different occupational choices for each group – which indicates misallocation if the distribution of talent is the same in each group (same $T_{i,g}$'s) and if each group has the same efficiency in accumulating human capital (same $\bar{h}_{ig}$'s). We wish to note, however, that the decline in the mean $\hat{\tau}_{ig}$ for the groups can also explain some of U.S. productivity growth over the last half century. The extent of the productivity growth and the root cause of the growth depends on what was generating the changes in $\hat{\tau}_{ig}$. We turn to these issues next.

5. Estimating Productivity Gains

5.1. Parameter Values and Exogenous Variables

The key parameters of the model — assumed to be constant over time — are η, θ, ρ, and β. We discussed the estimation and assumptions for η (the elasticity of human capital with respect to goods invested) and θ (the parameter governing the dispersion of talent) above. The parameter ρ governs the elasticity of substitution among our 67 occupations in aggregating up to final output. We have little information on this parameter and choose $\rho = 2/3$ for our baseline value, which implies an elasticity of substitution of 3 between occupations. We explore robustness to a wide range of values for ρ.

The parameter β is the geometric weight on consumption relative to time in an individual's utility function (1). As schooling trades off time for consumption, wages must increase more steeply with schooling in equilibrium when people value time more (i.e. when β is lower). We choose $\beta = 0.693$ to match the Mincerian return to schooling across occupations, which averages 12.7% across the six decades.[14] Our results are essentially invariant to this parameter, as documented later.

As our model is static, we infer exogenous variables separately for each decade. In each year, we have $6N$ variables to be determined. For each of the $i = 1, \ldots, N$ occupations these are A_i, ϕ_i, and τ_{ig}, where g stands for white men, white women, black women, or black men. We also allow population shares of each group q_i to vary by year to match the data. We normalize average ability to be the same in each occupation-group ($T_{ig} = 1$). Differences in ability across occupations (T_i) are isomorphic to differences in the production technology A_i. Across groups within an occupation, we think the natural starting point is *no* differences in mean ability; this assumption will be relaxed in our robustness checks. Additionally, we set $\tau_{i,wm} = 1$

[14] Workers must be more heavily compensated for sacrificing time to schooling the more they care about time relative to consumption. The average wage of group g in occupation i is proportional to $(1 - s_i)^{\frac{-1}{\beta}}$. If we take a log linear approximation around average schooling $\bar{s}$, then β is inversely related to the Mincerian return to schooling across occupations (call this return ψ): $\beta = (\psi(1 - \bar{s}))^{-1}$. We calculate s as average years of schooling divided by a pre-work time endowment of 25 years, and find the Mincerian return across occupations ψ from a regression of log average wages on average schooling across occupation-groups, with group dummies as controls. We then set $\beta = 0.693$, the simple average of the implied β values across years. This method allows the model to approximate the Mincerian return to schooling across occupations.

in all occupations. This restriction implies that white men face no occupational barriers. Again, we think this is a natural benchmark to consider. Finally, we normalize the $\bar{h}_{ig}$'s to be one for all groups in all occupations. As discussed above, the $\bar{h}_{ig}$'s are isomorphic to the τ^h_{ig}'s. Even though the magnitude of the productivity gains are the same, however, the underlying economic mechanisms differ under in the two cases. We discuss the $\bar{h}_{ig}$ interpretation in greater depth below. For now, we hold the $\bar{h}_{ig}$'s constant at 1.

To identify the values of the $6N$ forcing variables in each year, we match the following $6N$ moments in the data, decade by decade (numbers in parentheses denote the number of moments):

- $(4N-4)$ The fraction of people from each group working in each occupation, p_{ig}. (Fewer than $4N$ moments because the p_{ig} sum to one for each group.)
- (N) The average wage in each occupation.
- (N) The assumption that $\tau_{i,wm} = 1$ in each occupation.
- (3) Wage gaps between white men and each of our 3 other groups.
- (1) Average years of schooling in one occupation.

As shown above, the $3N$ $\hat{\tau}_{ig}$ variables are easy to identify from the data given our setup. Assuming that $\tau_{i,wm} = 1$, $T_{ig} = 1$, and $\bar{h}_{ig} = 1$, the τ_{ig}'s for the other groups are easy to infer using equation (26). But recall that $\tau_{ig} \equiv \frac{(1+\tau^h_{ig})^\eta}{1-\tau^w_{ig}}$. From the data we currently have, we cannot separately identify the τ^h and τ^w components of τ. That is, we cannot distinguish between human capital barriers and labor market barriers. We proceed by considering two polar cases. At one extreme, we assume all of the τ^w_{ig}'s are zero, so that τ_{ig} solely reflects τ^h_{ig}. At the other extreme, we set all of the $\tau^h_{ig} = 0$ and assume the τ^w_{ig}'s are responsible for the τ's. In short, we either assume only human capital barriers (the τ^h case) or only labor market barriers (the τ^w case).

The A_i levels and the relative ϕ_i's across occupations involve the general equilibrium solution of the model, but the intuition for what pins down their values is clear. We already noted that A_i is observationally equivalent to the mean talent parameter in each occupation T_i. The level of A_i helps determine the overall fraction of the population that works in each occupation. We also noted that ϕ_i is the key determinant of average wage differences across occupations. Thus, the data on employment shares and wages by occupation help us pin down the values of A_i and ϕ_i.

Recall from equation (10) that wages are increasing in schooling across occupations. From Proposition 1, we know that schooling increases with ϕ_i. From wages in each occupation, therefore, we can infer the *relative* values of ϕ_i across occupations. But we cannot pin down the ϕ_i levels, as absolute wage levels are also affected by the A_i productivity parameters. Thus we use a final moment – average years of schooling in one occupation – to determine the ϕ_i levels. We choose to match schooling in the lowest wage occupation, which is Farm Non-Managers in most years. Calling this the "min" occupation, we set ϕ_{min} in a given year to match the observed average schooling among Farm Non-Managers in the same year: $\phi_{min} = \frac{1-\eta}{\beta}\frac{s_{min}}{1-s_{min}}$.

5.2. Productivity Gains

Given our model, parameter values, and the forcing variables we infer from the data, we can now answer one of the key questions of the paper: how much of overall earnings growth from 1960 to 2008 can be explained by the changing labor market outcomes of blacks and women during this time period?

In answering this question, the first thing to note is that output growth in our model is a weighted average of earnings growth in the market sector and in the home sector. Earnings growth in the market sector can be measured as real earnings growth in the census data. Deflating by the NIPA Personal Consumption Deflator, real earnings in the census data grew by 1.32 percent per year between 1960 and 2008.[15] We impute wages in the home sector using the relationship between average earnings and average education across market occupations and from wage gaps by group in market occupations. (See the discussion in section 3.1. for additional details.) Taking a weighted average of the imputed wage in the home sector

[15]This might be lower than standard output growth measures because it is calculated solely from wages; for example, it omits employee benefits.

and the wage in the census data, we estimate that output (as defined by our model) grew by 1.47 percent per year between 1960 and 2008.

How much of this growth is accounted for by changing τ's, according to our model? We would like to answer this question by holding the A's (productivity parameters by occupation), ϕ's (schooling parameters by occupation), and q's (group shares of the working population) constant over time and letting the τ's change. But at which year's values should we hold the A's, ϕ's, and q's constant? We follow the standard approach in macroeconomics and use chaining. That is, we compute growth between 1960 and 1970 allowing the τ's to change but holding the other parameters at their 1960 values. Then we compute growth between 1960 and 1970 from changing τ's holding the other parameters at their 1970 values. We take the geometric average of these two estimates of growth from changing τ's. We do the same for other decadal comparisons (1970 to 1980 and so on) and cumulate the growth to arrive at an estimate for our entire sample from 1960–2008.

The results of this calculation are shown in Table 2. When the frictions are interpreted as occurring in human capital accumulation (the τ^h case), the change in occupational frictions explain 20.4 percent of overall earnings growth over the last half century. If we instead interpret the frictions as occurring in the labor market (the τ^w case), the changing τ^w's account for 15.9 percent of the cumulative earnings growth from 1960 to 2008. The gains are smaller in the τ^w case because some of the wage gaps are accounted for directly by labor market discrimination in this case, with no direct implications for productivity. There are still indirect effects operating through human capital accumulation and occupational choice, of course.

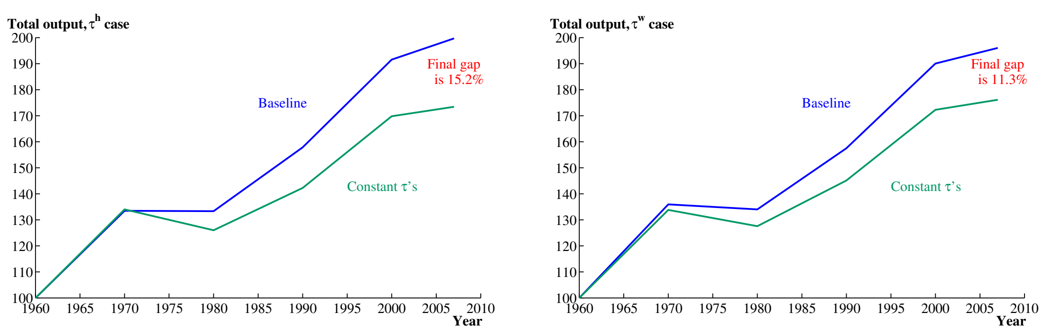
A related calculation, perhaps more transparent, is to hold the τ's constant and calculate the hypothetical growth due to the changes in the A's, ϕ's, and q's. Figure 6 plots the results of this calculation. The left panel considers the τ^h case. The large majority of growth is due to increases in A_i and ϕ_i over time, but an important part is attributable to reduced frictions. Allowing the τ^h's to change as they did historically raises output by 15.2 percent in the τ^h case. The right panel of Figure 6 presents the τ^w case. Here, reduced frictions raised overall output by 11.3% between 1960 and 2008.

Table 2: Productivity Gains: Share of Growth due to Changing Frictions

	τ^h case	τ^w case
Frictions in all occupations	20.4%	15.9%
Counterfactual: wage gaps halved	12.5%	13.7%
Counterfactual: zero wage gaps	2.9%	11.8%
No frictions in "brawny" occupations	18.9%	14.1%

Note: Average annual wage growth between 1960 and 2008 was 1.47%. Entries in the table show the share of labor productivity growth attributable to changing frictions according to our model under various assumptions. In the last line, we assume that there are no frictions for white women in occupations where physical strength is important. Instead, we allow $T_{i,ww}$ to change over time to match the occupational allocation for white women. For blacks in this case, we do allow for frictions, but also assume $T_{i,bw} = T_{i,ww}$.

Figure 6: Counterfactuals: Output Growth due to A, ϕ versus τ



Note: These graphs show the counterfactual path of output in the model if the τ's were kept constant over time (in a chained calculation). That is, how much of cumulative growth is due to changing A's and ϕ's versus changing τ's. The left panel is for the τ^h case, and the right is for the τ^w case.

It is worth elaborating on the gains from the changing τ's. Recall that Figure 5 above showed that both the mean and variance of τ has fallen for women and blacks over time. Both are sources of output gains, but the decline in the dispersion of the frictions is directly the source of the gains from the improved allocation of talent.[16]

Could the productivity gains we estimate be inferred from a simple back-of-the-envelope calculation involving the wage gaps alone? In particular, suppose one takes white male wage growth as fixed, and calculates how much of overall wage growth comes from the faster growth of wages for the other groups. The answer is that faster wage growth for blacks and white women accounts for 12 percent of overall wage growth from 1960 to 2008. This is compared to our estimate of the productivity gains from changing τ's of 20 percent in the τ^h case and 15 percent in the τ^w case.

Our counterfactuals differ from the back-of-the-envelope calculation in two fundamental ways. First, we are isolating the contribution of changing τ's on labor productivity growth, whereas the back-of-the-envelope also reflects any impact of changing A's, ϕ's, and q's on the wage growth of women and blacks relative to white men. Second, our counterfactuals take into account the impact of changing τ's on white men. In our counterfactuals, we will show shortly, the wage gains to women and blacks come partly at the expense of white males. As women and blacks move into high-skill occupations, this crowds out white men by lowering the wage per unit of human capital in those occupations so long as $\rho < 1$, i.e. occupations are not imperfectly substitutable.

The middle rows of Table 2 illustrate how the average within-occupation wage gaps between groups relate to the productivity gains that we estimate. In particular, we consider counterfactuals in which we substantially reduce the average wage gaps fed into the model. Cutting the wage gaps in half in all years reduces the share of growth explained from 20.4% to 12.5% in the τ^h case. Setting the average wage gaps in the data to zero leaves only 2.9% of growth explained by changes in the hu-

[16]Even for women, most of the decline in the variance can be seen between market occupations. But a notable portion of the overall decline for women comes from falling barriers in market occupations relative to the home sector.

man capital frictions. This is not surprising given the theoretical results shown in Section 2. The wage gap is affected by frictions faced by a group. As the wage gap goes to zero, so do the frictions. In the τ^w case, in contrast, the gains are relatively insensitive to the wage gaps, falling from 15.9% to 13.7% to 11.8% as wage gaps are halved and then eliminated. In the τ^w case, misallocation of talent by race and gender can occur even if average wages are similar, as the misallocation of talent is tied to the *dispersion* in the τ's, whereas the wage gaps are more related to average τ's. In the τ^h case the distortions operate by affecting average human capital investments that do show up in the wage gaps.[17] As these contrasting cases show, model productivity gains cannot be gleaned from the wage gaps alone.

The final row in Table 2 considers the robustness of our productivity gains to relaxing the assumption that men and women draw from the same distribution of talent in all occupations. In particular, we consider the possibility that some occupations rely more on physical strength than others, and that this reliance might have changed because of technological progress. For this check, we go to the extreme of assuming no frictions face white women in the occupations where physical strength is arguably important, including construction, firefighters, police officers, and most of manufacturing.[18] That is, we estimate values for T_{ig} for white women that fully explain their observed allocation to these occupations for 1960, 1970, . . ., 2008. Our hypothesis going into this check was that most of the productivity gains were coming from the rising propensity of women to be like lawyers, doctors, scientists, professors, and managers, occupations where physical strength is not important. The results in Table 2 support this hypothesis. The amount of growth explained by changing frictions falls only slightly — for example, from 20.4% to 18.9% in the τ^h case — if we assume that all the movement of women into manufacturing, construction, police, firefighting and other brawny occupations were due to changes in relative comparative advantages in these occupations.

How much additional growth could be had from reducing the frictions all the

[17]The τ^h results are identical to the case where we assume human capital efficiency parameters $\bar{h}_{ig}$ are driving the differential labor market outcomes between groups. As a result, the wage gap is closely linked to the productivity gains in this case.

[18]These occupations are assigned based on Rendall (2010).

Table 3: Potential Remaining Output Gains from Zero Barriers

	τ^h case	τ^w case
Frictions in all occupations		
Cumulative gain, 1960–2008	15.2%	11.3%
Remaining gain from zero barriers	14.3%	10.0%
No frictions in "brawny" occupations		
Cumulative gain, 1960–2008	14.0%	10.0%
Remaining gain from zero barriers	11.7%	9.1%

way to zero? The answer is in Table 3. Consider first the τ^h case. Between 1960 and 2008, changing frictions raised output by 15.2% with our baseline parameters. If the remaining frictions as of 2008 were removed entirely, output would be higher by an additional 14.3%. For the τ^w case the gain from eliminating all frictions in 2008 would be 10.0%. These results suggest that there are still some gains to be had by removing frictions further.

5.3. Robustness

How sensitive is the growth contribution of changing τ's to our parameter choices? Tables 4 and 5 explore robustness to different parameter values. For each set of parameter values considered, we recalculate the τ_{ig}, A_i, and ϕ_i values so that the model continues to fit the occupation shares and wage gaps.

The first row checks sensitivity to the elasticity of substitution (ρ) between occupations in production. In the τ^h case, the share of growth explained ranges from 19.7 percent when the occupations are almost Leontief ($\rho = -90$) to 21.0 percent when they are almost perfect substitutes ($\rho = 0.95$). This compares to 20.4 percent with our baseline value of $\rho = 2/3$. Outcomes are more sensitive in the τ^w case, with the share of growth explained by changing τ^w's going from 12.3 to 18.4 percent (vs. 15.9 percent baseline). The gains are increasing in substitutability. Our intuition is

Table 4: Robustness Results: Percent of Growth Explained in the τ^h case

	Baseline				
	$\rho = 2/3$	$\rho = -90$	$\rho = -1$	$\rho = 1/3$	$\rho = .95$
Changing ρ	20.4%	19.7%	19.9%	20.2%	21.0%
	3.44	4.16	5.61	8.41	
Changing θ	20.4%	20.7%	21.0%	21.3%	
	$\eta = 1/4$	$\eta = 0.01$	$\eta = .05$	$\eta = .1$	$\eta = .5$
Changing η	20.4%	20.5%	20.5%	20.5%	20.3%

Note: Entries in the table represent the share of labor productivity growth that is explained by the changing τ^h's using the chaining approach. Each entry changes one of the parameter values relative to our baseline case. In the "Changing θ" row, the parameter values reported are $\theta(1-\eta)$.

that distortions to the total amount of human capital in one occupation versus another are greater with higher substitutability across occupations (higher ρ). We not only have too few women doctors, for example, but too little total human capital of doctors when women face barriers to the medical profession. This is particularly true when the allocation of talent is being directly distorted as in the τ^w case.

The second row indicates that the gains from changing τ's rise modestly as $\theta(1-\eta)$ rises above our baseline value (holding η fixed at 0.25). Recall that our baseline $\theta(1-\eta)$ of 3.44 was estimated from wage dispersion within occupation-groups controlling for hours worked, potential experience, and education – and making adjustments for AFQT scores and transitory wage movements. This baseline value attributes 75% of wage dispersion within occupation-groups to comparative advantage. Our baseline θ may overstate the degree of comparative advantage, as it imperfectly controls for absolute advantage. We thus entertain higher values of θ that attribute 50%, 75% and 90% of wage dispersion within occupation-groups to absolute advantage. These higher $\theta(1-\eta)$ values of 4.16, 5.61, and 8.41 attribute the remainder – 50%, 25% and 10% of wage dispersion – to comparative advantage. As shown in the second row of the robustness table, counterfactual gains rise from 20.4

Table 5: Robustness Results: Percent of Growth Explained in the τ^w case

	Baseline				
	$\rho = 2/3$	$\rho = -90$	$\rho = -1$	$\rho = 1/3$	$\rho = .95$
Changing ρ	15.9%	12.3%	13.3%	14.7%	18.4%
	3.44	4.16	5.61	8.41	
Changing θ	15.9%	14.6%	12.9%	11.2%	
	$\eta = 1/4$	$\eta = 0$	$\eta = .05$	$\eta = .1$	$\eta = .5$
Changing η	15.9%	13.9%	14.4%	14.8%	17.5%

Note: Entries in the table represent the share of labor productivity growth that is explained by the changing τ^w's using the chaining approach. Each entry changes one of the parameter values relative to our baseline case. In the "Changing θ" row, the parameter values reported are $\theta(1-\eta)$.

percent to 21.3 percent of growth as θ rises in the τ^h case. In the τ^w case, the percent of growth explained *falls* from 15.9 percent to 11.2 percent across the θ range.

When people are more similar in ability (θ is higher), smaller barriers are required to explain why women and blacks were underrepresented in high skill occupations. The more modest the barriers, the the smaller the gains from reducing them. This is why the gains shrink significantly with higher θ in the τ^w case. In the τ^h case, the gains are more tied to average human capital investments, which are not so sensitive to θ.

The third row considers different values of the elasticity of human capital with respect to goods invested in human capital (η). In the τ^h case, the gains hardly change as η rises from 0.01 to 0.25.[19] The τ^h values adjust as we change η to explain the observed wage gaps, and the productivity gains are tied to the wage gaps in the τ^h case. Gains are much more sensitive to η in the τ^w case, rising from 13.9 percent to 17.5 percent as η goes from 0 to 0.5. When η is high, wage gaps reflect human capital gaps more and the labor market discrimination less in the τ^w case.

[19] We must have $\eta > 0$ in the τ^h case as the only source of wage and occupation differences across groups is different human capital investments in this case.

Although not shown in the robustness tables, the gains are not at all sensitive to the value of β, the weight placed on time vs. goods in utility (β). The gains do not change to one decimal point as we move its value from 0.5 to 0.8 around the baseline value of $\beta = 0.693$.

As we imputed wages in the home sector, we now check whether the results are sensitive to higher or lower wages there. Specifically, we raise or lower home sector wages in every year by one standard error from the regression of market wages on education across occupation-groups controlling for group dummies – between 15% and 23% depending on the year. We find the gains fluctuate modestly around the baseline values. For example, in the τ^h case the fraction of growth explained rises to 21.9% with low home sector wages and falls to 18.6% with higher home sector wages, versus 20.4% in the baseline. Thus our imputation for the home sector does not have a big effect on our conclusions.

We can also exclude the home sector altogether, and analyze the gains conditional on working in the market. In this calculation, we allow the "population" share of women to rise over time to explain their increasing representation in the workforce. The τ changes explain an even higher fraction of growth in *market* wages: 24.9% in the τ^h case and 21.8% in the τ^w case (vs. 20.4% and 15.9% when combining the home sector with market occupations).[20]

Another robustness check we carry out is to look only at young workers, those between 25 and 35 years old (inclusive). The reason for this counterfactual is that our model is static and the data are inherently dynamic. For example, 25 year olds in 1960 will be 35 year olds in 1970. By focusing on a young cohort, we will include each cohort only once in our analysis, when they first enter the labor market. Focusing on this younger sample, the changing τ's account for an even higher fraction of growth, 28.7% in the τ^h case and 23.6% in the τ^w case.[21] The younger cohorts appear to be more responsive to the changing τ's than the entire population.

Finally, we test the sensitivity of our findings to the number of occupations con-

[20]Market wage growth is modestly lower, at 1.32% per year, than the 1.47% growth we infer for the market and home combined.

[21]Wage growth was notably slower for the young at 1.04% per year, so there was less to explain.

sidered. The gains are surprisingly robust. When we look at a broader set of 20 occupations (vs. 67 in our baseline), the gains are 20.1% vs. 20.4% baseline in the τ^h case, and 14.4% vs. 15.4% baseline in the τ^w case. For 1980 onward we can construct a consistent set of 331 more detailed occupations. Looking at the 1980–2008 subperiod, the gains are similar with more finely divided occupations: 21.1% vs. 20.9% baseline in the τ^h case, and 16.8% vs. 15.2% in the τ^w case.

5.4. Further Results

In this subsection, we describe a number of additional insights from the model.

In the Census data, the share of women working in the market rose from 32.9 percent in 1960 to 69.2 percent in 2008. One explanation is that women's market opportunities rose, say due to declining discrimination or better information. See Jones, Manuelli, and McGrattan (2003), Albanesi and Olivetti (2009), and Fogli and Veldkamp (2011) for empirical analysis of these hypotheses. As illustrated in Figure 3, the τ's fell in market occupations relative to the home sector for women. How much of the rising female labor-force participation rate can be traced to changing τ's? Table 6 provides the answer. Of the 36.4 percentage point increase, the changing τ's contributed 23.5 or 26.2 percentage points, or around 75 percent of the total increase. According to our model, the remaining 25 percent can be attributed to changes in technology such as the A's. The latter is in the spirit of the work by Greenwood, Seshadri and Yorukoglu (2005) on "engines of liberation". It is also consistent with studies attributing women's rising work to changes in the wage structure, such as Jones, Manuelli, and McGrattan (2003) and Fernandez and Wong (2011).

As we report in Table 7, gaps in average years of schooling narrowed from 1960 to 2008 for all three groups vs. white males: by 0.4 years for white women, 1.8 years for black men, and 1.55 years for black women. If the τ's for blacks and women fell faster in high schooling occupations, then the changing τ's contributed to this educational convergence. The table indicates how much. For white women, the changing τ's account for the trend and then some (0.6 years, vs. 0.4 in the data). For black men, falling frictions might have narrowed the schooling gap by 0.65 years,

Table 6: Female Participation Rates

	τ^h case		τ^w case
Women's LF participation	1960 = 0.329		2008 = 0.692
Change, 1960 – 2008		0.364	
Due to changing τ's	0.235		0.262
(Percent of total)	(72.3%)		(78.7%)

Note: Results are for white women and black women combined. Participation is defined as working in market occupations. The sampling weight of part-time workers is split evenly between the market sector and the home sector. Italicized entries in the table are data; non-italicized entries are results from the model.

Table 7: Education Predictions, All Households (Age 25–55)

	Actual 1960	Actual 2008	Actual Change	Change vs. WM	Due to τ's
White men	11.11	13.47	2.35		
White women	10.98	13.75	2.77	0.41	0.63
Black men	8.56	12.73	4.17	1.81	0.65
Black women	9.24	13.15	3.90	1.55	1.17

Note: Authors' calculations using Census data.

about one-third of the convergence in the data. For black women, declining distortions might explain three-quarters (1.17/1.55) of their catch-up in schooling.

How much of the productivity gains reflect changes in the occupational frictions facing women vs. those facing blacks? Tables 8 provides the answer. The columns presents the productivity gain from setting the τ's to their levels at the end of each period (1960–1980, 1980–2008, and 1960–2008). The first row does this for all groups, and the next rows do this for white women, black men, and black women, respectively. Take the τ^h case. Three-quarters (15.3 out of 20.4) of the total gains from reduced occupational frictions over the last fifty years can be explained by the

Table 8: Contribution of Each Group to Total Earnings Growth

	1960–1980	1980–2008	1960–2008
τ^h *case:*			
All groups	19.7%	20.9%	20.4%
White women	11.3%	18.2%	15.3%
Black men	3.3%	0.9%	1.9%
Black women	5.1%	1.9%	3.2%
τ^w *case:*			
All groups	21.1%	19.2%	20.0%
White women	8.7%	15.4%	12.6%
Black men	3.4%	0.8%	1.9%
Black women	4.7%	1.4%	2.8%

Note: Author's calculations using Census data and baseline parameter values.

changes facing white women. Falling frictions faced by blacks accounted for the remaining one-quarter of the gains.

The share of gains associated with falling frictions for white women vs. blacks differs across the time periods. Again, consider the τ^h case. Blacks accounted for a larger share of the gains in the 1960s and 1970s than in later decades. From 1960 to 1980, reduced frictions for blacks account for over 40% of the overall gains. From 1980 to 2008, reduced frictions for blacks account for less than 15% of the overall gains. This timing might link the gains for blacks to the Civil Rights movement of the 1960s.

What was the consequence of shifting occupational frictions for the wage growth of different groups? Table 9 tries to answer this question. The first column presents the actual growth of real wages for the different groups from 1960 to 2008. Real wages increased by 77 percent for white men, 126 percent for white women, 143 percent for black men, and 198 percent for black women. For brevity, consider the τ^h case. In the absence of the change in occupational frictions, the model says real wages for white men would have been almost 6 percent higher. Put differently, real

Table 9: Group Changes in Wages

	Actual Growth	Due to τ^h's	Due to τ^w's
White men	77.0 percent	-5.8%	-7.1%
White women	126.3 percent	41.9%	43.0%
Black men	143.0 percent	44.6%	44.3%
Black women	198.1 percent	58.8%	59.5%

Note: Authors' calculations using Census data and baseline parameter values.

income of white men declined due to the changing opportunities for blacks and women. At the aggregate level, this loss was swamped by the wage gains for blacks and women. Over 40 percent of the wage growth for white women was due to the change in occupational frictions – similarly for black men. For black women, over half of their earnings growth might reflect the increased opportunities they faced. The model explains the remainder of growth as resulting from changes in technology (A's) and skill requirements (ϕ's).

Tables 10 looks at the regional dimension of the decline in frictions confronting blacks and women. Here, we assume that workers are immobile across regions. With this assumption, a decline in occupational frictions in the South relative to the North will increase average wages in the South relative to the North. From 1960 to 2008, wages in the South increased by 10 percent relative to wages in the Northeast. In the τ^h case, about 7 percentage points of this convergence was due to reduced occupational frictions facing blacks and women in the South relative to the Northeast — with the bulk of the effect due to falling τ's for blacks.

From 1980 to 2008, we see a reversal of the North-South convergence, perhaps driven by the reverse migration of blacks to the U.S. South. Reverse migration is what one would expect to see if workers are responding to the improved labor market outcomes in the South by relocating to the South. In a long run with higher labor mobility, the main effect of declining occupational frictions for blacks in the South relative to the North might be to increase the number of blacks living in the South

Table 10: Contributions to Northeast - South Convergence

	1960–1980	1980–2008	1960–2008
τ^h *case:*			
Actual wage convergence	20.7%	-16.5%	10.0%
Due to all τ's changing	4.9%	1.5%	6.9%
Due to black τ's changing	3.6%	1.9%	5.6%
τ^w *case:*			
Actual wage convergence	20.7%	-16.5%	10.0%
Due to all τ's changing	2.2%	0.4%	2.8%
Due to black τ's changing	2.3%	1.2%	3.6%

Note: Author's calculations using Census data and baseline parameter values.

relative to the North. Persistent wage gaps might reflect skill differences between regions. Of course, to the extent mobility is costly even in the long run, frictions can contribute to wage gap differences across regions even in the long run.

5.5. Average Quality of Workers by Occupation

Using equations (9) and (10), the average quality of workers — including both innate ability and human capital — for group g in occupation i is given by

$$\frac{H_{ig}}{q_g p_{ig}} = \gamma\bar{\eta} \cdot \frac{1}{(1-\tau_{ig}^w)w_i} \cdot (1-s_i)^{-1/\beta} \cdot \left(\sum_{s=1}^{N} \tilde{w}_{sg}^{\theta}\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}}. \tag{28}$$

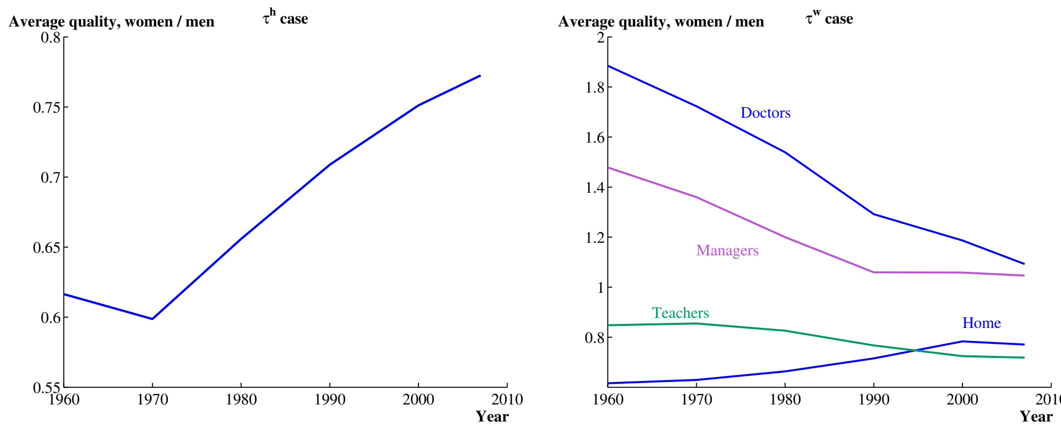
Average quality for a group *relative* to white men is therefore

$$\frac{H_{ig}/q_g p_{ig}}{H_{i,wm}/q_{wm} p_{i,wm}} = \frac{1-\tau_{i,wm}^w}{1-\tau_{ig}^w} \cdot \frac{\overline{\text{wage}}_g}{\overline{\text{wage}}_{wm}}. \tag{29}$$

That is, relative quality in an occupation is simply the wage gap adjusted by the τ^w frictions.

In the τ^h case (where the τ^w variables are set to zero), equation (29) implies that

Figure 7: Relative Average Quality, White Women vs. White Men



Note: The panels show relative average quality (human capital and innate ability) in various occupations for white women versus white men, in the τ^h and τ^w cases. Computed using equation (29).

average quality for a group relative to white men is *the same across all occupations*. In particular, relative quality is precisely equal to the wage gap. When the labor market friction are introduced, this changes. In this case, wages are not equal to marginal products, so that average quality differs from wages. More specifically, wages are less than marginal products, so average quality is higher when the frictions are larger.

One way to think about these quality differences is to consider the following question: if you were to see a doctor chosen at random in 1960 for a fixed fee, would you rather see a male doctor or a female doctor?

Figure 7 shows the ratio of average quality between white women and white men for several occupations, as in equation (29). In the τ^h case, as noted, relative qualities are equated for all occupations. Because the wage gap reflects quality, the average female doctor is less qualified than the average male doctor — the distortions lead women doctors to accumulate less human capital than their male counterparts. Over time, as the wage gaps have declined, the relative quality of women to men in each occupation rose substantially between 1960 and 2008, from 0.56 to 0.77.

The τ^w case presents a very different view of the data, as shown in the right panel. The relative quality of women is higher than their wage gaps suggest because they are paid less than their marginal products. In 1960, average quality was substantially higher for women vs. men doctors and managers. Only the most talented women overcame frictions to become doctors and managers in 1960, and some lesser talented white men entered these professions instead. According to this case, the difference in quality has faded substantially over time due to declining frictions, but remains present even in 2008.

The real world presumably has elements of both the τ^h and the τ^w cases. To this end, independent information on quality trends for occupation-groups could help us separate and identify human capital and labor market frictions.

6. Conclusion

How does discrimination in labor markets and in the acquisition of human capital affect occupational choice? And what are the consequences of the resulting misallocation of talent on aggregate productivity? We develop a framework to tackle these questions empirically. This framework has three building blocks. First, we use a standard Roy model of occupational choice, augmented to allow for labor market discrimination and discrimination in the acquisition of human capital. Second, we impose the assumption that the distribution of an individual's ability over all possible occupations follows an iid extreme value distribution. Third, we embed the Roy model in general equilibrium to account for the effect of occupational choice on the price of skills in each occupation, and to allow for the effect of technological change on occupational choice.

We apply this framework to measure the changes in barriers to occupational choice facing women and blacks in the U.S. from 1960 to 2008. We find large reductions in these barriers, largely concentrated in high-skilled occupations. We then use our general equilibrium setup to measure the aggregate effects of the reduction in occupational barriers facing these groups. We estimate that the reduction in these barriers can explain 15 to 20 percent of aggregate wage growth, 90 to 95 percent of the wage convergence between women and blacks and white men, and 75 percent of the rise in women's labor force participation from 1960 to 2008.

It should be clear that this paper only provides a preliminary answer to these important questions. It would be useful to develop a framework that does not rely on the assumption that the distribution of talent is iid across all occupations (we are less troubled by the extreme value distribution assumption).[22] It would also be useful to quantify the extent to which the barriers are due to labor market discrimination versus discrimination in the acquisition of human capital. As we've discussed, independent data on quality trends would be useful to distinguish between these two forces. Finally, we have focused on the gains from the reduction in barriers in occupational choice facing women and blacks over the last fifty years. However, we suspect that similar barriers facing children from less affluent families and from regions of the country hit by adverse economic shocks have *worsened* in the last few decades. If so, this could explain both the adverse trends in aggregate productivity and the fortunes of less-skilled Americans over the last decades. We hope to tackle some of these questions in future work.[23]

A Derivations and Proofs

The propositions in the paper summarize the key results from the model. This appendix shows how to derive the results.

Proof of Proposition 1. Occupational Choice

As given in equation (5), the individual's utility from choosing a particular occupation, $U(\tau_i, w_i, \epsilon_i)$, is proportional to $(\bar{w}_{ig}\epsilon_i)^{\frac{\beta}{1-\eta}}$, where $\bar{w}_{ig} \equiv w_i \bar{h}_{ig} s_i^{\phi_i}(1-s_i)^{\frac{1-\eta}{\beta}}/\tau_{ig}$. The solution to the individual's problem, then, involves picking the occupation with the largest value of $\bar{w}_{ig}\epsilon_i$. To keep the notation simple, we will suppress the g sub-

[22] For example, Lagakos and Waugh (2011) allow skill to be correlated for working inside and outside of agriculture. Mulligan and Rubinstein (2008) do the same for working in the home sector and in the market, and argue for changing selection of high ability U.S. women into the market over time.

[23] One could also investigate groups in other countries. For example, Hnatkovska, Lahiri and Paul (2011) look at castes in India and find narrowing differences in education, occupations, and wages in recent decades.

script in what follows.

Let p_i denote the probability that the individual chooses occupation i. Then

$$\begin{aligned} p_i &= \Pr\left[\bar{w}_i\epsilon_i > \bar{w}_s\epsilon_s\right] \ \forall i \neq s \\ &= \Pr\left[\epsilon_s < \bar{w}_i\epsilon_i/\bar{w}_s\right] \ \forall s \neq i \\ &= \Pi_{s\neq i} F_s(\bar{w}_i\epsilon_i/\bar{w}_s) \end{aligned} \tag{30}$$

if ϵ_i is known for certain. Since it is not, we must also integrate over the probability distribution for ϵ_i:

$$p_i = \int \Pi_{s\neq i} F_s(\bar{w}_i\epsilon_i/\bar{w}_s) f_i(\epsilon_i) d\epsilon_i, \tag{31}$$

where $f_i(\epsilon) = \theta T_i \epsilon^{-(1+\theta)} \exp\{-T_i\epsilon^{-\theta}\}$ is the pdf of the Fréchet distribution. Substituting in for the distribution and pdf, additional algebra leads to

$$\begin{aligned} p_i &= \int \theta T_i \left(\Pi_{s\neq i} \exp\{-T_s(\bar{w}_i\epsilon_i/\bar{w}_s)^{-\theta}\}\right) \epsilon_i^{-(1+\theta)} \exp\{-T_i\epsilon^{-\theta}\} d\epsilon_i \\ &= \int \theta T_i \epsilon_i^{-(1+\theta)} \exp\left\{-\sum_{s=1}^{N} T_s \left(\frac{\bar{w}_i}{\bar{w}_s}\right)^{-\theta} \epsilon_i^{-\theta}\right\} d\epsilon_i. \end{aligned} \tag{32}$$

Now, define $\bar{T}_i \equiv \sum_{s=1}^{N} T_s \left(\frac{\bar{w}_i}{\bar{w}_s}\right)^{-\theta}$. Then the probability simplifies considerably:

$$\begin{aligned} p_i &= \frac{T_i}{\bar{T}_i} \int \theta \bar{T}_i \epsilon_i^{-(1+\theta)} \exp\{-\bar{T}_i\epsilon_i^{-\theta}\} d\epsilon_i \\ &= \frac{T_i}{\bar{T}_i} \int d\bar{F}_i(\epsilon_i) \\ &= \frac{T_i}{\bar{T}_i} \\ &= \frac{T_i \bar{w}_i^{\theta}}{\sum_s T_s \bar{w}_s^{\theta}} \\ &= \frac{\tilde{w}_i^{\theta}}{\sum_s \tilde{w}_s^{\theta}} \end{aligned} \tag{33}$$

where $\tilde{w}_i \equiv T_i^{1/\theta} \bar{w}_i$.

Proof of Proposition 2. Average Quality of Workers

Total efficiency units of labor supplied to occupation i by group g are

$$H_{ig} = q_g p_{ig} \cdot \mathbb{E}\left[h_i \epsilon_i \,|\, \text{Person chooses } i\right].$$

Recall that $h(e,s) = \bar{h}_i s^{\phi_i} e^{\eta}$. Using the results from the individual's optimization problem, it is straightforward to show that

$$h_i \epsilon_i = \tilde{h}_i \left(\frac{w_i(1-\tau_i^w)}{1+\tau_i^h}\right)^{\frac{\eta}{1-\eta}} \epsilon_i^{\frac{1}{1-\eta}},$$

where $\tilde{h}_i \equiv (\eta^{\eta} \bar{h}_i s_i^{\phi_i})^{\frac{1}{1-\eta}}$. Therefore,

$$H_{ig} = q_g p_{ig} \tilde{h}_i \left(\frac{w_i(1-\tau_i^w)}{1+\tau_i^h}\right)^{\frac{\eta}{1-\eta}} \cdot \mathbb{E}\left[\epsilon_i^{\frac{1}{1-\eta}} \,|\, \text{Person chooses } i\right]. \tag{34}$$

To calculate this last conditional expectation, we use the extreme value magic of the Fréchet distribution. Let $y_i \equiv \bar{w}_i \epsilon_i$ denote the key occupational choice term. Then

$$y^* \equiv \max_i \{y_i\} = \max_i \{\bar{w}_i \epsilon_i\} = \bar{w}^* \epsilon^*.$$

Since y_i is the thing we are maximizing, it inherits the extreme value distribution:

$$\begin{aligned}
\text{Pr}\left[y^* < z\right] &= \Pi_{i=1}^N \text{Pr}\left[y_i < z\right] \\
&= \Pi_{i=1}^N \text{Pr}\left[\bar{w}_i \epsilon_i < z\right] \\
&= \Pi_{i=1}^N \text{Pr}\left[\epsilon_i < z/\bar{w}_i\right] \\
&= \Pi_{i=1}^N \exp\left\{-T_i \left(\frac{z}{\bar{w}_i}\right)^{-\theta}\right\} \\
&= \exp\{-\sum_{i=1}^N T_i \bar{w}_i^{\theta} \cdot z^{-\theta}\} \\
&= \exp\{-\bar{T} z^{-\theta}\}.
\end{aligned} \tag{35}$$

That is, the extreme value also has a Fréchet distribution, with a mean-shift parameter given by $\bar{T} \equiv \sum_s T_i \bar{w}_i^{\theta}$.

Straightforward algebra then reveals that the distribution of ϵ^*, the ability of people in their chosen occupation, is also Fréchet:

$$G(x) \equiv \Pr\left[\epsilon^* < x\right] = \exp\{-T^* x^{-\theta}\} \tag{36}$$

where $T^* \equiv \sum_{i=1}^N T_i \left(\bar{w}_i/\bar{w}^*\right)^\theta$.

Finally, one can then calculate the statistic we needed above back in equation (34): the expected value of the chosen occupation's ability raised to some power. In particular, let i denote the occupation that the individual chooses, and let λ be some positive exponent. Then,

$$\begin{aligned}\mathbb{E}[\epsilon_i^\lambda] &= \int_0^\infty \epsilon^\lambda dG(\epsilon)\\ &= \int_0^\infty \theta T^* \epsilon^{-(1+\theta)+\lambda}\, e^{-T^*\epsilon^{-\theta}} d\epsilon \end{aligned}\tag{37}$$

Recall that the "Gamma function" is $\Gamma(\alpha) \equiv \int_0^\infty x^{\alpha-1}e^{-x}dx$. Using the change-of-variable $x = T^*\epsilon^{-\theta}$, one can show that

$$\begin{aligned}\mathbb{E}[\epsilon_i^\lambda] &= T^{*\lambda/\theta}\int_0^\infty x^{-\lambda/\theta}e^{-x}dx\\ &= T^{*\lambda/\theta}\Gamma(1-\lambda/\theta). \end{aligned}\tag{38}$$

Applying this result to our model, we have

$$\begin{aligned}\mathbb{E}\left[\epsilon_i^{\frac{1}{1-\eta}} \mid \text{Person chooses } i\right] &= T^{*\frac{1}{\theta}\cdot\frac{1}{1-\eta}}\,\Gamma\left(1-\frac{1}{\theta}\cdot\frac{1}{1-\eta}\right)\\ &= \left(\frac{T_i}{p_{ig}}\right)^{\frac{1}{\theta}\cdot\frac{1}{1-\eta}}\Gamma\left(1-\frac{1}{\theta}\cdot\frac{1}{1-\eta}\right). \end{aligned}\tag{39}$$

Substituting this expression into (34) and rearranging leads to the last result of the proposition.

Proof of Proposition 3. Occupational Wage Gaps

The proof of this proposition is straightforward given the results of Proposition 1. Note that $\bar{\eta} \equiv \eta^{\eta/(1-\eta)}$.

References

Albanesi, Stefania and Claudia Olivetti, "Gender Roles and Medical Progress," NBER Working Papers 14873, National Bureau of Economic Research, Inc April 2009.

Bailey, Martha J., "More Power to the Poll: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply," *Quarterly Journal of Economics*, February 2006, *121* (1), 289–320.

__, Brad Hershbein, and Amalia R. Milleri, "The Opt-In Revolution? Contraception and the Gender Gap in Wages," NBER Working Papers 17922, National Bureau of Economic Research, Inc March 2012.

Becker, Gary, *The Economics of Discrimination*, Chicago, IL: The University of Chicago Press, 1957.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, July 2010, *2* (3), 228–55.

Biskupic, Joan, *Sandra Day O'Connor: How the First Woman on the Supreme Court Became Its Most Influential Justice*, New York, NY: HarperCollins, 2006.

Blau, Francine D., "Trends in the Well-Being of American Women, 1970-1995," *Journal of Economic Literature*, March 1998, *36* (1), 112–165.

Card, David and Alan B. Krueger, "School Quality and Black-White Relative Earnings: A Direct Assessment," *The Quarterly Journal of Economics*, February 1992, *107* (1), 151–200.

Cavalcanti, Tiago and Jos Tavares, "The Output Cost of Gender Discrimination: A Model-Based Macroeconomic Estimate," CEPR Discussion Papers 6477, C.E.P.R. Discussion Papers September 2007.

Charles, Kerwin Kofi and Jonathan Guryan, "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination," *Journal of Political Economy*, October 2008, *116* (5), 773–809.

Chay, Kenneth Y., Jonathan Guryan, and Bhashkar Mazumder, "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth," NBER Working Papers 15078, National Bureau of Economic Research, Inc June 2009.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

Donohue, John J. and James Heckman, "Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks," *Journal of Economic Literature*, December 1991, *29* (4), 1603–43.

Dupuy, Arnaud, "An Economic Model of the Evolution of the Gender Performance Ratio in Individual Sports," *The International Journal of Performance Analysis in Sport*, 2012, *12* (1), 224–245.

Eaton, Jonathan and Samuel S. Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Ellison, Glenn and Ashley Swanson, "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions," *Journal of Economic Perspectives*, 2010, *24* (2), 109–128.

Erosa, Andres, Tatyana Koreshkova, and Diego Restuccia, "How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *Review of Economic Studies*, October 2010, *77* (4), 1421–1449.

Fernandez, Raquel, "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century," *American Economic Review*, forthcoming 2012.

_ , Alessandra Fogli, and Claudia Olivetti, "Preference Formation and the Rise of Women's Labor Force Participation: Evidence from WWII," NBER Working Papers 10589, National Bureau of Economic Research, Inc June 2004.

_ and Joyce Chen Wong, "The Disappearing Gender Gap: The Impact of Divorce, Wages, and Preferences on Education Choices and Women's Work," 2011. unpublished paper.

Fogli, Alessandra and Laura Veldkamp, "Nature or Nurture? Learning and the Geography of Female Labor Force Participation," *Econometrica*, July 2011, *79* (4), 1103–1138.

Goldin, Claudia, *Understanding the Gender Gap: An Economic History of American Women*, New York, NY: Oxford University Press, 1990.

— and Lawrence F. Katz, "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy*, 2002, *110* (4), 730–770.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu, "Engines of Liberation," *Review of Economic Studies*, January 2005, *72* (1), 109–133.

Guvenen, Fatih and Burhanettin Kuruscu, "A Quantitative Analysis of the Evolution of the U.S. Wage Distribution, 1970-2000," in Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., *NBER Macroeconomics Annual*, Cambridge, MA: MIT Press, 2009, pp. 227–276.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh B. Paul, "Breaking the Caste Barrier: Intergenerational Mobility in India," 2011. U.B.C. manuscript.

Jones, Larry E., Rodolfo E. Manuelli, , and Ellen R. McGrattan, "Why Are Married Women Working So Much?," Research Department Staff Report 317, Federal Reserve Bank of Mineapois June 2003.

Karabel, Jerome, *The Chosen: the Secret History of Admission and Exclusion at Harvard, Yale, and Princeton*, New York, NY: Houghton Mifflin Harcourt, 2005.

Lagakos, David and Michael E. Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," 2011. Arizona State University and NYU manuscript.

Manuelli, Rodolfo and Ananth Seshadri, "Human Capital and the Wealth of Nations," March 2005. University of Wisconsin working paper.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers of Econometrics*, New York, NY: Academic Press, 1974, pp. 105–142.

Mulligan, Casey B. and Yona Rubinstein, "Selection, Investment, and Women's Relative Wages Over Time," *Quarterly Journal of Economics*, August 2008, pp. 1061–1110.

Rendall, Michelle, "Brain versus brawn: the realization of women's comparative advantage," IEW - Working Papers iewwp491, Institute for Empirical Research in Economics - University of Zurich September 2010.

Rodgers, William M. and William E. Spriggs, "What Does the AFQT Really Measure: Race, Wages, Schooling and the AFQT Score," *The Review of Black Political Economy*, Spring 1996, pp. 13–46.

Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, June 1951, *3* (2), 135–146.

Smith, James P. and Finis R. Welch, "Black Economic Progress after Myrdal," *Journal of Economic Literature*, June 1989, *27* (2), 519–64.

B Data Appendix

Table B.1: Sample Statistics By Census Year

	1960	1970	1980	1990	2000	2006-8
Sample Size	641,686	694,419	4,057,685	4,711,405	5,216,431	3,147,547
Share of White Men in Sample	0.432	0.433	0.435	0.435	0.431	0.431
Share of White Women in Sample	0.475	0.468	0.459	0.447	0.437	0.431
Share of Black Men in Sample	0.042	0.044	0.047	0.054	0.061	0.065
Share of Black Women in Sample	0.052	0.055	0.059	0.065	0.071	0.074
Relative Wage Gap: White Women	-0.578	-0.590	-0.476	-0.372	-0.299	-0.259
Relative Wage Gap: Black Men	-0.379	-0.289	-0.215	-0.158	-0.142	-0.150
Relative Wage Gap: Black Women	-0.875	-0.705	-0.479	-0.363	-0.317	-0.313

Table B.2: Occupation Categories for our Base Occupational Specification

Home Sector
Executives, Administrative, and Managerial
Management Related
Architects
Engineers
Math and Computer Science
Natural Science
Health Diagnosing
Health Assessment
Therapists
Teachers, Postsecondary
Teachers, Non-Postsecondary
Librarians and Curators
Social Scientists and Urban Planners
Social, Recreation, and Religious Workers
Lawyers and Judges
Arts and Athletes
Health Technicians
Engineering Technicians
Science Technicians
Technicians, Other
Sales, All
Secretaries
Information Clerks
Records Processing, Non-Financial
Records Processing, Financial
Office Machine Operator
Computer and Communication Equipment Operator
Mail Distribution
Scheduling and Distributing Clerks
Adjusters and Investigators
Misc. Administrative Support
Private Household Occupations
Firefighting
Police
Guards
Food Preparation and Service
Health Service
Cleaning and Building Service
Personal Service
Farm Managers
Farm Non-Managers
Related Agriculture
Forest, Logging, Fishers, and Hunters
Vehicle Mechanic
Electronic Repairer
Misc. Repairer
Construction Trade
Extractive
Precision Production, Supervisor
Precision Metal
Precision Wood
Precision Textile
Precision Other
Precision Food
Plant and System Operator
Metal and Plastic Machine Operator
Metal and Plastic Processing Operator
Woodworking Machine Operator
Textile Machine Operator
Printing Machine Operator
Machine Operator, Other
Fabricators
Production Inspectors
Motor Vehicle Operator
Non Motor Vehicle Operator
Freight, Stock, and Material Handlers

Table B.3: Examples of Occupations within Our Base Occupational Categories

Management Related Occupations

Accountants and Auditors
Underwriters
Other Financial Officers
Management Analysts
Personnel, Training, and Labor Relations Specialists
Purchasing Agents and Buyers
Construction Inspectors
Management Related Occupations, N.E.C.

Health Diagnosing Occupations

Physicians
Dentists
Veterinarians
Optometrists
Podiatrists
Health Diagnosing Practitioners, N.E.C.

Personal Service Occupations

Supervisors, personal service occupations
Barbers
Hairdressers and Cosmetologists
Attendants, amusement and recreation facilities
Guides
Ushers
Public Transportation Attendants
Baggage Porters
Welfare Service Aides
Family Child Care Providers
Early Childhood Teacher Assistants
Child Care Workers, N.E.C.

Table B.4: Occupation Categories for our Broad Occupation Classification

Home Sector	Sales, All
Executives, Administrative, and Managerial	Administrative Support, Clerks, and Record Keepers
Management Related	Fire, Police, and Guards
Architects, Engineers, Math, and Computer Science	Private Household and Food, Cleaning, and Personal Services
Natural and Social Scientists, Recreation, Religious, Arts, and Athletes	Farm, Related Agriculture, Logging, Forest, Fishing, Hunters and Extraction
Doctors and Lawyers	Mechanics and Construction
Nurses, Therapists, and Other Health Services	Precision Manufacturing
Teachers, Postsecondary	Manufacturing Operators
Teachers, Non-Postsecondary and Librarians	Fabricators, Inspectors, and Material Handlers
Health and Science Technicians	Vehicle Operators