Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper nº 94

Wage Inequality in a Frictional Labor Market

Joel Shapiro

November 2003

Wage Inequality in a Frictional Labor Market*

Joel Shapiro†
Universitat Pompeu Fabra
Barcelona Economics Working Paper #94

November 2003

Abstract

Wage inequality in the United States has grown substantially in the past two decades. Standard supply-demand analysis in the empirics of inequality (e.g. Katz and Murphy (1992)) indicates that we may attribute some of this trend to an outward shift in the demand for high skilled labor. In this paper we examine a simple static channel in which the wage premium for skill may grow - increased firm entry. We consider a model of wage dispersion where there are two types of workers and homogeneous firms must set wages and preferences for what type of worker they would like to hire. We find that both the wage differential and the demand for high skill workers can increase with the proportion of high skill workers - these high skill workers therefore "create" their own demand without exogenous factors. In addition, within group wage inequality can increase in step with the between group wage inequality. Simulations of the model are provided in order to compare the findings with empirical results.

JEL Classifications: D83, J31,J41

Keywords: wage posting, wage inequality, search

*I thank Patrick Bolton, Matthew Ellman, Maia Guell, Robert Shimer, Karen Conneely, and Kosuke Aoki for excellent comments and suggestions. I have also benefitted from comments from the Princeton Public Finance Working Group. Financial support from DGES under BEC 2000-1026, Barcelona Economics (CREA), and the Woodrow Wilson Foundation is gratefully acknowledged.

†Contact: Joel Shapiro, Universitat Pompeu Fabra, Departament D'Economia i Empresa, Ramon Trias Fargas, 25-27, 08005 Barcelona SPAIN. email: joel.shapiro@econ.upf.es phone: (34) 93 542 2718

1 Introduction

Wage inequality in the United States has grown substantially in the past two decades[1]. Most studies[2] trying to explain the wage premium have focused on demand factors; technology-skill complementarities and international trade's effect on skill composition are the most prominent explanations[3]. Nevertheless, a significant amount of the variation in wages is not explained by these factors.

This paper seeks to take one step back and ask whether a simple model of a frictional labor market can produce similar wage dynamics without such factors as technology-skill complementarities or international trade. There are two compelling reasons to pursue this avenue. The first is to understand more fully how imperfections in the labor market affect the wage structure. The second is to demonstrate that the current theoretical literature potentially overstates the contribution of heterogeneity in technology by combining it with labor market frictions.

We assume that there are two types of workers (high skill and low skill) and one type of firm. In a standard supply and demand framework this clearly implies that absent outside factors, a large increase in the supply of high skilled workers should result in a fall in the wage premium for skill. This lies in contrast to the data, and the empirical literature has overwhelmingly pointed to increases in the demand for skill. Where does this demand come from? In our model, high skill workers can create demand for themselves by making it more profitable for firms to enter.

The imperfection we consider comes from the matching between workers and firms. We assume that each firm has two tasks, a high skill and a low skill one. High skill workers may produce in either, while low skill workers may only produce in the low skill task. Firms open positions at a cost and set wages and preferences for the type of worker they would like to employ. Workers come into contact with a fixed number of jobs, apply, and then choose among the jobs that are offered to them. This model has similar characteristics and dynamics as other wage posting models (for example, Burdett and Judd (1983)), but abstracts away from worker search behavior in order to clearly study wage inequality.

In equilibrium, it is not necessarily the case that firms will prefer high

[1] For a quick view of the data on wage inequality, Murphy and Welch (1993), Figures 1-3 display the trends quite well.

[2] For surveys of the literature, see Levy and Murnane (1992) and Aghion et.al. (1999).

[3] Katz and Murphy (1992) also point compellingly to the rate of change in the supply of college graduates as one explanation of the data.

skill workers. If high skill workers' reservation wages are too large, firms will choose to employ low skill workers. The expected income of a worker can be represented in an extremely simple manner: the probability of having more than one job offer multiplied by the worker's output plus the probability of having only one job offer multiplied by the worker's reservation wage. This shows that the more likely it is for the worker to have an outside offer, the more rent can be extracted from the firms.

We prove that if wage inequality increases with the proportion of high skill workers, it must come from increased firm entry. Using a simulation, we are able to display cases where wage inequality is increasing in the proportion of high skill workers. It occurs in labor markets where firms prefer high skilled workers, where costs of entry are high and where the proportion of high skill workers is low. In such labor markets we may therefore find that high skill workers increase demand for themselves exclusive of any external factors. Given that we have wage distributions in equilibrium, we are also able to discuss wage dispersion and within group wage inequality. In regions where wage inequality between groups is increasing we find that there is also increasing wage inequality within each group.

Our work is most comparable with Acemoglu (1998, 1999). He assumes labor market frictions come from random matching and that there is bargaining between workers and firms. His main focus, however, is to show how endogenous investment decisions can change job composition and hence wage inequality. Shi (2001) and Shimer (2001) also discuss wage inequality in a directed search framework with heterogeneity in workers and capital. Lang, Manove, and Dickens (2000) focus on discrimination in a directed search framework, but as they assume two types of worker and one type of firm, their results are comparable with those here. They find a complete separation of markets based on type. Shimer (1998) also considers a two type worker population and explores how firms' preferences depend on how wages are determined (i.e. bargaining versus wage posting).

In section 2, we set up the model. In section 3 we provide our main results and in section 4 we conclude.

2 The Model

We construct a one period labor market in which:

- Firms borrow money to open a position at cost k, have two tasks, post wages for each one, and decide what task they prefer to be accomplished if possible.

- Two types of workers apply for positions and decide which task is appropriate for them to perform in each position.

While our results are quite clear and simple, we need some machinery to describe the equilibrium. We proceed by going through each part of the market.

2.1 Labor Market Frictions

The main assumption in this paper is that of an exogenous application/search process[4], which allows us to solve in a tractable manner. Workers meet randomly a fixed number of firms. This can be interpreted as workers finding a subset of the job opportunities available through advertisements (in the classifieds, on the internet) or through intermediaries. It is logical that firms within the same labor market meet workers in uniform ways.

The application process that we specify utilizes the basic dynamics of nonsequential search established in Burdett and Judd (1983)[5]. Consider the labor market interactions of a mass of firms and a mass of workers. Each firm has one position available. Workers apply to n firms drawn at random from the available pool. If all workers only apply to just one firm, then firms will set wages at the workers' reservation level since workers will not have the option to refuse an offer. This is the 'Diamond-paradox'. However, if workers apply to more than one firm, some will have a positive probability of being able to refuse a low offer, putting pressure on firms to raise their wages in order to increase the probability of getting a worker. This creates a wage distribution in equilibrium[6]. Therefore, the existence of some probability with which workers can compare wages drives wage dispersion. In our model the probability of a worker's application succeeding (yielding a job offer) is between 0 and 1, implying that more than one application will allow some workers to compare wages. To obtain wage dispersion with closed form solutions, we limit the number of applications[7] to two.

[4] Many search models assume an exogenous search process, e.g. Varian (1980) and Stahl (1989), but obtain wage dispersion from the heterogeneity of agents. Here all workers searching for jobs are homogeneous and search in the exact same way. Ex-post heterogeneity creates the dispersion in wages.

[5] Burdett and Judd (1983) is actually a model of price dispersion. The description above translates it into a model of wage dispersion (price=wage, consumers=workers).

[6] Here we don't have to worry about the 'Bertrand result' where firms face so much pressure to raise wages that wage distribution is an atom at the highest possible value. This only occurs when workers sample more than one wage with probability 1, which can't occur in the labor market we describe.

[7] Albrecht, Gautier, and Vroman (2001) look at externalities caused by multiple appli-

2.2 Labor Market Participants

Workers

There are two types of workers, high skill and low skill. High skill workers represent a fraction μ of the total amount of workers, which is normalized to 1. A high skill worker can perform a high skill task, which yields output f_h, or a low skill task, which yields output f_l. The low skill worker may only perform the low skill task. Wages may be made conditional on output, a test, or on one's credentials; in any case, the high skill worker must be given incentives to perform the high skill task[8]. In addition, the reservation wages of the high and low skill workers are denoted by $\underline{w}_h$ and $\underline{w}_l$, respectively. We assume that $\underline{w}_h > \underline{w}_l$.

For a given probability λ of receiving a job offer at one firm, the expected wage of a worker can now be expressed as:

$$\lambda^2 \int_a^b 2wG(w)g(w)dw + 2\lambda(1-\lambda)\int_a^b wg(w)dw \tag{1}$$

where the first term represents the product of the probability of two successful applications and the expectation of the top wage given two samples, and the second term represents the product of the probability of one successful application and the expected wage given one sample.

Firms

The cost of opening a position in this labor market is $\tilde{k}$. Firms may borrow to open the position at an interest rate of r (we define $k = (1+r)\tilde{k}$). Should the firm not be able to pay k on time, they suffer some additional default cost d. We will call each position a firm, although it is possible to think of one firm recruiting separately for several positions. The assumption of a one worker - one firm match is also employed in Montgomery (1991) and Acemoglu and Shimer (1998) in order to discuss the impact of tightness in the labor market. There are a large number of potential firms, but in equilibrium a mass M of them enters, each receiving expected profits of zero. We assume that the equilibrium number of firms M is greater than 2 (there are more firms than applications).

Each position offered is flexible in the sense that the technology may accommodate either a low skill or a high skill worker, although a high skill

cations in matching functions.

[8] This can also be considered as a simple adverse selection problem for the high type. Since utility is not type dependent here, the high type will apply for the high skill task when the wage offered for it is greater than that offered for the low skill task.

worker may be more productive. We may think of the extremes, where a high school dropout has a huge productivity disadvantage in the field of nuclear physicist or where a diplomat has a reservation value so high that the compensation needed for her to work in McDonald's would be tremendous. However, many markets are somewhere in between. Berman, Bound and Griliches (1994) provide evidence that highly educated workers are working in manufacturing jobs that were previously the domain of those with lesser education.

The firm is free to offer distinct wages for a high skill or low skill task. Since high skill workers must be given incentives[9] to choose the high skill task, the wage for that task must be higher than the wage for the lower skill task. In equilibrium, therefore, we will see high skill workers at high skill tasks and low skill workers at low skill tasks.

As both types of workers may apply to a given firm, each firm needs some type of hiring policy. We allow this policy to be endogenous and take the form: "If some i types show up, we will choose randomly among them. If no i types show up and j types are around, we will choose randomly among j types." Essentially, the firm may have a preference, but if it is not able to exercise this preference, it will still try to fill its position to recover some surplus. This preference, and the wages offered for each type of task, are the choice variables of the firm.

We define β_i as the probability that at least one person who wants to perform task i shows up given that at least one worker shows up. Calculated, $\beta_h = \frac{1-e^{-\frac{2\mu}{M}}}{1-e^{-\frac{2}{M}}}$ and $\beta_l = \frac{1-e^{-\frac{2(1-\mu)}{M}}}{1-e^{-\frac{2}{M}}}$. We use these to discuss the preference of the firm over tasks. Writing the expected profits[10] for a firm posting wages w_h and w_l that has a preference for a type h worker (the expected profits for $E\pi_l$ can be written in an analogous manner):

$$\begin{aligned} E\pi_h(w_h, w_l) &= (1-e^{-\frac{2}{M}})\{\beta_h p_h(w_h)(f_h - w_h) + (1-\beta_h)p_l(w_l)(f_l - w_l)\} \\ &\quad -\{1-(1-e^{-\frac{2}{M}})(\beta_h p_h(w_h) + (1-\beta_h)p_l(w_l)\}d - k \end{aligned}$$

The first line represents the revenues. The first term represents the probability that at least one worker shows up. Within the brackets, there

[9] The firm may, on the other hand, set the high skill task wage lower than or equal the low skill task wage, thus encouraging high skill workers to perform the low skill task. This is not optimal in equilibrium no matter what task the firm prefers, essentially since the waste of the added productivity can't be balanced out by the shifts in probabilities. A proof of this is available from the author.

[10] We assume that output is sold by the firm at a price normalized to 1.

are two terms: the expected payoff from task h and the expected payoff from task l. The expected payoffs depend on the preference of the firm for the task and the return from the task (output minus wage) multiplied by the probability that the worker who is offered the job accepts the offer. For now we refer to this probability as $p_i(w_i)$ for task i with wage w_i. The second line represents the revenues. If the firm does not match, it defaults and pays added cost d, and in any case is responsible for the cost of the machine k. This can be rewritten as:

$$E\pi_h(w_h, w_l) = (1 - e^{-\frac{2}{M}})\{\beta_h p_h(w_h)(f_h - w_h + d) + (1 - \beta_h)p_l(w_l)(f_l - w_l + d)\} - d - k \tag{2}$$

2.3 Labor Market Equilibrium

We begin the analysis with some definitions. The probability that a worker of type i will get selected for employment at a firm is equal to λ_i. We characterize the wage distribution for a type i worker by a cdf $G_i(w)$ and its associated pdf $g_i(w)$ defined over the range $[a_i, b_i]$. Using this, the probability that the randomly chosen worker will accept a firm's offer of w_i, $p_i(w_i)$, becomes $1 - \lambda_i + \lambda_i G_i(w_i)$, where the worker either does not have another offer (probability $1 - \lambda_i$) or has an offer which is smaller than w_i. This is a very important property; it says that expected payoffs of a firm X in task i are completely determined by the strategies of other firms (their preference over tasks) and the wage that firm X offers for task i, w_i, and *not* the wage from the other task. The only connection then between the two tasks is the restriction that $w_h \geq w_l$. We will ignore this constraint, solve for the equilibrium, and then verify that it holds. The next lemma characterizes the wage distributions.

Lemma 1 *Wages are distributed for task i along a continuous distribution* $[\underline{w}_i, b_i]$ *and expected payoffs for the firm along this distribution are equal.*

Proof. i) If all firms only offered one wage, then there is a profitable

deviation upward by ε, where the pay is larger, but the probability of the worker having another offer which is weakly preferred jumps discretely to 0. Should the one wage equal f_i, then there is a profitable deviation by lowering the wage - paying less and having a positive probability that a worker will accept.

ii) Suppose there is a gap in the wage distribution. Then a firm offering a wage at the top of the gap can lower its wage by ε, pay less, and have the same probability of attracting workers.

iii) Expected payoffs along the wage distribution must be equal. If not, a firms will deviate to the wages which offer higher payoffs.

iv) The firm which offers the lowest wage in the distribution will only get a worker if the worker has no other offer. Therefore, they have the incentive to lower the wage as much as possible, capturing profits from this worker. The lowest the firm can lower the wage is to the worker's reservation wage, $\underline{w}_i$. ■

Each task offers a fixed expected payoff to the firm of π_i, irrespective of the wage paid. Therefore for all w_i :

$$(1 - \lambda_i + \lambda_i G_i(w_i))(f_i - w_i) = \pi_i \tag{3}$$

The fact that the bottom wage equals the reservation wage $\underline{w}_i$ for type i workers allows us to solve explicitly for task i expected payoffs. Using the endpoint condition $G(\underline{w}_i) = 0$, the payoffs equal $(1-\lambda_i)(f_i-\underline{w}_i)$. This payoff makes sense intuitively; it increases with the output of the task, decreases with the reservation value of the workers, and decreases with the probability that the firm must give a specific worker the job. The bargaining power of the worker, based on her outside option and degree of competition with other workers, plays a strong role here. We will see later that despite the fact that the problem is based on wage posting, the results will often have clear bargaining interpretations. The equal profits condition also defines the cumulative distribution $G_i(w_i)$. We can easily determine that $G_i(w) = \frac{1-\lambda_i}{\lambda_i}(\frac{w_i - \underline{w}_i}{f_i - w_i})$ and the top wage $b = \lambda_i f_i + (1 - \lambda_i)\underline{w}_i$.

The final piece of the equilibrium lies in characterizing firms' preferences over tasks. Given that payoffs are fixed for each task, this consists in comparing the expected profits from prefering task h to the expected profits from prefering task l. This is straightforward, but in order to get a full characterization, we must elaborate and derive the formal matching probabilities. Workers face different probabilities of acceptance if all firms prefer task h than if all firms prefer task l. We calculate the probabilities using techniques similar to the ball-urn process of Butters (1977), and write them as λ_{ij}, where i is the task/type, and j is the task/type that all firms prefer (the calculations are in the appendix):

$$\lambda_{hh} = \frac{1-e^{-\frac{2\mu}{M}}}{\frac{2\mu}{M}} \qquad \lambda_{hl} = \frac{e^{-\frac{2(1-\mu)}{M}}-e^{-\frac{2}{M}}}{\frac{2\mu}{M}}$$
$$\lambda_{ll} = \frac{1-e^{-\frac{2(1-\mu)}{M}}}{\frac{2(1-\mu)}{M}} \qquad \lambda_{lh} = \frac{e^{-\frac{2\mu}{M}}-e^{-\frac{2}{M}}}{\frac{2(1-\mu)}{M}}$$

In the following lemma, we list some basic properties of these probabilities, as they will be useful in what follows. We focus on the probabilities λ_{hh} and λ_{lh}, which are used when firms prefer high skilled workers, but it is easy to see that these also describe the properties λ_{ll} and λ_{hl}, which are essentially the same probabilities with $1-\mu$ switched for μ.

Lemma 2 *i) Both λ_{hh} and λ_{lh} are decreasing with μ*
ii) The difference $\lambda_{hh} - \lambda_{lh}$ is decreasing with μ
iii) The difference $\lambda_{hh} - \lambda_{lh}$ is positive

The proof is in the appendix. Results (i) and (iii) are clearly intuitive. The probability of either worker getting a job offer when workers with high skills are preferred is hurt by a low skill worker being switched for a high skill worker. Moreover, since firms prefer high skill workers their probability of getting hired is larger. Result (ii) says that as low skilled workers get converted into high skill workers, the probability of a high skill worker getting a job is decreasing faster than that of a low skill worker.

We are now prepared to state equilibrium preferences of firms over tasks.

Proposition 3 *If $\frac{f_h-w_h}{f_l-w_l} > \frac{1-\lambda_{lh}}{1-\lambda_{hh}}$ all firms prefer the high skill task. If $\frac{f_h-w_h}{f_l-w_l} < \frac{1-\lambda_{ll}}{1-\lambda_{hl}}$, all firms prefer the low task. For $\frac{1-\lambda_{ll}}{1-\lambda_{hl}} \leq \frac{f_h-w_h}{f_l-w_l} \leq \frac{1-\lambda_{lh}}{1-\lambda_{hh}}$ firms are indifferent and employ mixed strategies.*

The proposition follows directly from the equilibrium condition of no profitable deviations . The nature of the equilibrium depends on the ratio of the maximum returns of the firm from a type h worker versus that of a type l worker. This ratio of returns is exogenous and depends on the comparison between output and reservation wages of the types. The returns on the high skill workers must be sufficiently larger than the returns on the low skill worker in order for high skill workers to be preferred. This follows from the fact that $\lambda_{hh} > \lambda_{lh}$, making $\frac{1-\lambda_{lh}}{1-\lambda_{hh}} > 1$. If, on the other hand, the returns from low skill workers are actually larger due to their much lower reservation values, the firm will prefer the low skill workers (similarly it is

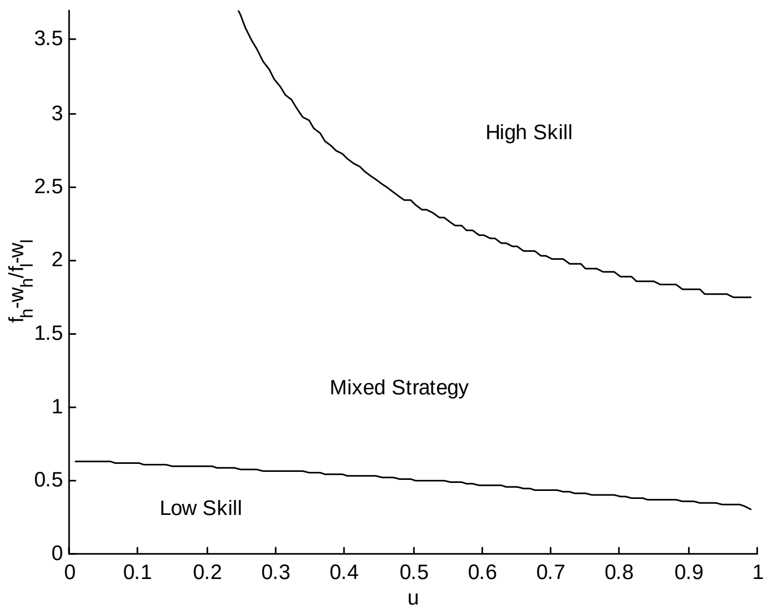

Figure 1: Equilibrium Firm Preferences over Tasks ($\frac{d}{f_l - \text{w}_l} = .2$, $\frac{k}{f_l - \text{w}_l} = .5$)

clear that $\frac{1-\lambda_{ll}}{1-\lambda_{hl}} < 1$). For intermediate values, firms are indifferent and mix among the types[11].

Figure 1 displays the equilibria for different values of the proportion of high skilled workers μ and the ratio of returns of the high skill to the low skill worker $\frac{f_h - \text{w}_h}{f_l - \text{w}_l}$. We can see that the high skill equilibrium requires a large proportion of high skill workers present in the market and a high return to these workers. The low skill equilibrium requires a higher return to low skill workers, which could mean that high skill workers have similar productivity to low skill workers in these jobs but a larger reservation wage. If we perform the comparative static experiment of an increase in the proportion of high

[11] While firms may use different mixed strategies, a quick way to observe that all intermediate values of parameters can be achieved is to suppose that all firms choose the same strategy: with probability ϕ they prefer H and with probability $(1-\phi)$ they prefer L. Then, for all $\phi\epsilon[0,1]$ all values between $\frac{1-\lambda_{ll}}{1-\lambda_{hl}} \leq \frac{1-(\phi\lambda_{lh}+(1-\phi)\lambda_{ll})}{1-(\phi\lambda_{hh}+(1-\phi)\lambda_{hl})} \leq \frac{1-\lambda_{lh}}{1-\lambda_{hh}}$ are spanned, and the ratio $\frac{f_h - \text{w}_h}{f_l - \text{w}_l}$ can be matched with a ϕ. Due to the infinite number of mixed strategy equilibria in this region, we restrict comparative static exercises to the regions where preferences are strict.

skill workers we find that it is possible for a market to switch from low skill to mixed or from mixed to high skill. So here high skill workers are preferred when there are more of them around. As we will see later, the equilibrium configuration depends on the costs as well. For high costs of entry, the low skill equilibrium will disappear, as the maximum return on it can't recover the costs.

In the wage posting environment in Shimer (1998), high skill workers are strictly preferred. In the model of Lang, Manove and Dickens (2001), the market separates completely. How realistic is the notion that firms may sometimes prefer lower skilled workers? In our model this preference comes about explicitly because these firms believe that the high skilled workers have many outside options and hence a high reservation wage. Another motivation can be seen from a recent court case in New London, Connecticut. A man sued the local police force for not hiring him on the basis that he was too intelligent (as determined by a standardized test). The deputy police chief was quoted[12] as saying that this man was, "exactly the type of guy [they] want to screen out...Police work is kind of mundane". This points to the fact that high skill workers may actually be relatively unproductive in low skill jobs. The issues of retention and job satisfaction of prospective applicants, while not covered explicitly by the model, provide additional motivation for the results.

The final piece of the equilibrium is the zero profit condition for firms. The number of firms M who enter in equilibrium (if task H is preferred) is determined by equation 2 which we can now rewrite:

$$(1-e^{-\frac{2}{M}})\{\beta_h(1-\lambda_{hh})(f_h-\underline{\mathrm{w}}_h+d)+(1-\beta_h)(1-\lambda_{lh})(f_l-\underline{\mathrm{w}}_l+d)\}=d+k \tag{4}$$

An analogous condition holds if task L is preferred.

All that remains is to verify whether the restriction that $w_h \geq w_l$ holds under the equilibrium we described. Since firms maximize the payoffs from each task separately, we can't specify a relationship $w_h(w_l)$ that assigns a specific high and low wage to each firm, but we should check if the restriction holds for some configuration of firms in the equilibrium. A good method for doing this is checking whether for each wage w, there are more firms still offering high skill tasks than there are offering low skill tasks, or, more concretely, $1-G_h(w) > 1-G_l(w)$. We do this in the appendix for the both the high task market and the low task market, and find that it holds.

[12] This story was related in "Help Wanted: The Not-Too-High-Q Standard" (New York Times, September 19, 1999).

3 Wage Inequality

Each equilibrium defines expected wages for workers as a function of the parameters. In this section we seek to identify exactly how an increase in the proportion of skilled workers may increase the wage gap between the skilled and unskilled. The wage formation process depends on the matching frictions, and we will explore how the labor markets therefore differ from standard supply and demand results.

We can simplify the problem considerably by substituting the wage distributions into the workers' expected wage, given in equation 1. This yields the expression $\lambda_{ij}^2 f_i + 2\lambda_{ij}(1-\lambda_{ij})\underline{w}_i$, where once again i is the type of worker and j is the type of worker that is preferred by firms. This expression is quite elegant, as it has two clear intuitive interpretations. First, if we take into account the 'Diamond paradox' (if all applicants have only one successful application, the wage will be the reservation wage of $\underline{w}_i$) and the 'Bertrand result' (if all applicants have two successful applications, the wage will be bid up to f_i) expected wage can be seen as the probability of two successful applications multiplied by its payoff (f_i) plus the probability of one successful application times its payoff ($\underline{w}_i$). Second, if we think about Nash Bargaining between a worker and a firm, the result[13] given our setting would be $\beta f_i + (1-\beta)\underline{w}_i$, where β is the relative bargaining power of the worker and is between 0 and 1. Here we find that the relative bargaining power arises *endogenously* from the ability of the worker to compare wages[14].

Although the expected wage is quite useful in itself, wage inequality is concerned with observed wages, that is, measures of wages among employed people. We can adjust the expected wage to find the average wage for a type i worker by dividing by the probability that the worker finds employment $\lambda_{ij}^2 + 2\lambda_{ij}(1-\lambda_{ij})$. Rewriting, the average wage for a type i is: $f_i - \frac{2-2\lambda_{ij}}{2-\lambda_{ij}}(f_i - \underline{w}_i)$. Wages of high skill workers are always greater than wages of low skill workers (given $\underline{w}_h > \underline{w}_l$).

[13] Maximizing $((f_i - w) - 0)^{1-\beta}(w - \underline{w}_i)^{\beta}$ yields this result.

[14] While the expected wage is $\lambda_{ij}^2 f_i + 2\lambda_{ij}(1-\lambda_{ij})\underline{w}_i$, the worker expects to earn $\lambda_{ij}^2 f_i + (1-\lambda_{ij}^2)\underline{w}_i$ since with probability $(1-\lambda_{ij})^2$ she does not receive an offer and gets her reservation value $\underline{w}_i$.

3.1 Analysis

We focus the analysis on labor markets that prefer high skilled workers[15]. An increase in the proportion of high skill workers μ changes the average wages of both high skill and low skill workers through the probability of employment. This change comes through two effects. The direct effect of an increase in μ is to reduce the probability either type is employed. The indirect effect comes from the change in the number of firms M in response to the change in μ. An increase in M increases the probability that each type will get a job. The potential opposition of these effects prompts an inspection into firms' entry decisions and hence expected profits. The following lemma establishes important properties of expected profits:

Lemma 4 *Expected profits decrease in the number of firms M. For some μ^*, expected profits increase with the proportion of high skill workers when $\mu > \mu^*$.*

A sketch of this proof is provided in the appendix. As the number of firms increases, the competition (in the form of getting workers to accept their positions) among them becomes tougher, lowering profits. An increase in the proportion of high skill workers has two effects. It increases the returns to high skill workers since the firms have more high skill workers to choose from and each high skill worker has a lower probability of having another offer. However, it decreases the returns to low skill workers for the opposite reason. The effect of an increase in the proportion high skill workers is strongest when there are more of them (i.e. large μ) because the second effect becomes small. In addition, an increase in profits due to more high skilled workers is amplified when the return to a high skill worker $f_h - \underline{w}_h$ grows relative to the return to a low skill worker $f_l - \underline{w}_l$ (μ^* decreases).

Given costly entry into the labor market, a change in the proportion of high skill workers will affect the number of firms through the zero profit condition. This relationship can be found by implicit differentiation of the zero profit condition: $\frac{dE\pi_h}{d\mu} + \frac{dE\pi_h}{dM}\frac{dM}{d\mu} = 0$. Rearranging, $\frac{dM}{d\mu} = \frac{-\frac{dE\pi_h}{d\mu}}{\frac{dE\pi_h}{dM}}$. Using the results from Lemma 4, we can sign this expression. For $\mu < \mu^*$, it is positive, meaning that more high skilled workers encourage firms to enter. For $\mu > \mu^*$, it is negative, implying that the returns from entering are decreasing with the proportion of high skill workers.

[15] The method of analysis for markets in which firms prefer low skill workers is similar, but, unsurprisingly, we do not find increasing wage inequality there.

Wage inequality is expressed as the difference between skill groups of observed average wages:

$$\bar{W}_h - \bar{W}_l = f_h - f_l + \frac{2-2\lambda_{lh}}{2-\lambda_{lh}}(f_l - \underline{\mathrm{w}}_l) - \frac{2-2\lambda_{hh}}{2-\lambda_{hh}}(f_h - \underline{\mathrm{w}}_h)$$

The change in wage inequality with respect to the proportion of high skilled workers is:

$$\frac{d(\bar{W}_h - \bar{W}_l)}{d\mu} = \frac{2(f_h - \underline{\mathrm{w}}_h)}{(2-\lambda_{hh})^2}(\frac{d\lambda_{hh}}{d\mu} + \frac{d\lambda_{hh}}{dM}\frac{dM}{d\mu}) - \frac{2(f_l - \underline{\mathrm{w}}_l)}{(2-\lambda_{lh})^2}(\frac{d\lambda_{lh}}{d\mu} + \frac{d\lambda_{lh}}{dM}\frac{dM}{d\mu})$$

Here we have made it explicit that the change in wage inequality with respect to a change in μ is the result of the direct effect (change in the probability of matching) and the indirect effect (change in the number of firms that enter). The following proposition states that only the indirect effect can cause an increase in wage inequality.

Proposition 5 *An increase in the proportion of high skilled workers can increase wage inequality by attracting the entry of firms.*

The proof is in the appendix, and shows that the direct effect decreases wage inequality. Consider the following experiment: take one worker who is of low skill and make that person high skill. All low skill workers who were previously contending with this worker now face a little tougher competition from her. All high skill workers, who had no competition from this worker at all, now have a potential rival. Therefore, by holding the firm effect constant, an increase in the proportion of high skilled workers decreases the probabilities of matching of both types of workers, but affects the high type more. This then lowers her wage more relative to the low type. This direct effect is essentially the supply effect from a supply and demand framework.

What remains is to see whether an increase in the proportion of high skill workers can attract enough firms to enter to raise their wages. This is an endogenous demand[16] effect, driven solely by the fundamentals of the problem: relative returns on workers, entry costs, and matching frictions. In figure 2, we show that it is possible to have increasing wage inequality in our setting. Based on simulations, the requirements are high capital

[16] Indeed, when we consider firms as the number of positions opened, the interpretation of the increase as a demand effect is very clear.

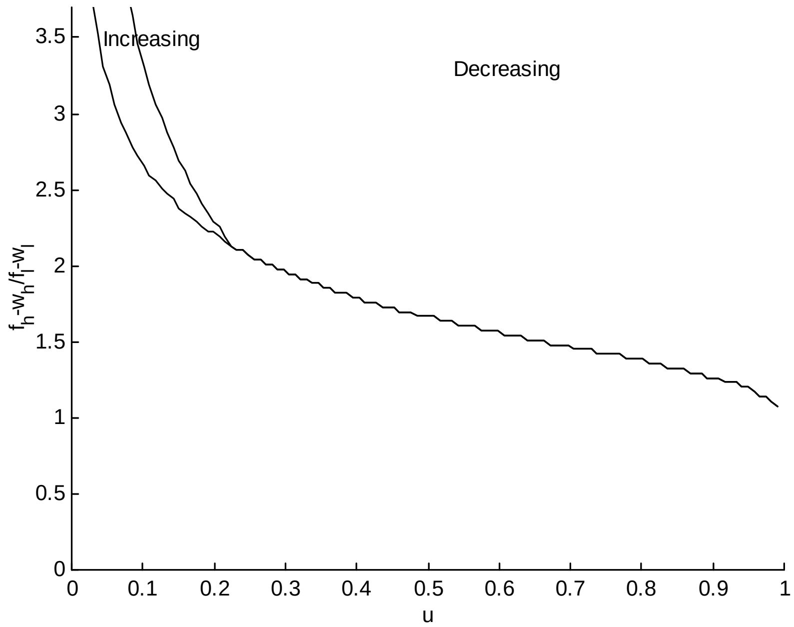

Figure 2: Changes in Inequality in a High Skill Labor Market ($\frac{d}{f_l - \underline{w}_l} = .2$, $\frac{k}{f_l - \underline{w}_l} = 1$)

costs and positive default penalties (for the simulations we measure them as percentages of the return on low skill workers $f_l - \underline{w}_l$). The high capital costs eliminate the presence of a low skill equilibrium as the maximum returns on a low skill individual are too low.

The increasing wage inequality takes place in the region of high skill equilibria where the proportion of high skill individuals is lowest, i.e. is driven by the increasing profits to firms of added high skill workers when there are few around. Wages for high skill workers are increasing strongly in this region. The wages of low skill workers are increasing by a small amount in both regions.

Another measure that we are interested is the range of wages, the top wage minus the bottom wage. This could give some insight into within group wage inequality. The range for type i in the high skill labor market is expressed simply as the top wage $\lambda_{ih} f_i + (1 - \lambda)\underline{w}_i$ minus the bottom wage

$\underline{w}_i$, or $\lambda_{ih}(f_i - \underline{w}_i)$. Evaluating the change in the ranges for the parameters above, we find that in the region where wage inequality is increasing, the ranges of wages for high skill workers is increasing sharply and the range of wages for low skill workers is also increasing but at a smaller rate. Note that as the bottom wage is fixed, all of the movement comes from the top wage. There is strong evidence on increasing inequality within groups (for example, see Juhn, Murphy and Pierce (1992)). In the region where wage inequality is decreasing, the range for high skill workers decreases, but the range for the low skill workers increases.

3.2 Relationship with the Literature

The issue of which channel increases the demand for skilled labor is quite important, as it has strong implications for inequality policy. Acemoglu (1998 and 1999) discusses how the increase in the size of the skilled population encourages skill complementary technological innovation, spurring demand for high skill workers. In contrast, we offer an explanation for the rising skill premium solely based on firm entry. High skill workers therefore "create" demand for themselves. Had the model included complementarities with new technologies, the results could be amplified substantially, making it important to differentiate between the two effects.

The search literature splits substantially on how wages are decided upon. Dynamic search models often assume Nash Bargaining (for example, Acemoglu (1999) and Shimer (1998)) while static models usually assume wage posting (for example, Montgomery (1991) and Shi (2001)). Our work offers a wage posting model that offers a micro-foundation for how bargaining power is determined. Among wage posting models, the most commonly used is a directed search model, which describes how workers may direct one application towards a firm of their choice having observed the wage distribution. The directed search approach involves equilibria selection and mixed strategy equilibria. The closest paper to ours in the directed search literature is Lang, Manove, and Dickens (2001), which discusses discrimination. While they are interested in wage differentials, they do not explore changes in wage differentials with the proportion of one type (reasonably so, given their focus).

4 Conclusion

Data shows that the proportion of the U.S. population with college degrees has been increasing constantly since the late 1970s[17]. Supply-demand analysis, such as that in Katz and Murphy (1992), indicates that the demand for highly educated workers far outstripped supply since 1980, pushing wages for these workers constantly upward. The huge demand has been explained by numerous factors, including technology and international trade. Here we present an explanation that does not depend on exogenous demand factors or technological change: increased firm entry. The increased entry is a result of increasing returns to matching with high skill workers, a result of the frictions in the labor market. Further theoretical and empirical work is needed to isolate the determinants of the wage premium. Separating the factors leading to the creation of demand for high skill workers would yield important insight into inequality and policy options.

Appendix

A Proof of Lemma 1

Suppose there are n firms and q workers. The general probability of a worker matching equals:

$$\sum_{i=0}^{q} \Pr(i \ others \ show \ up) \Pr(hired \mid i \ show \ up) \tag{5}$$

The probability that a type h has a successful application given that firms prefer type h (which is labelled λ_{hh}) sets the $\Pr(hired \mid i \ show \ up)$ equal to (given a proportion μ of high skilled workers):

$$\sum_{j=0}^{i} \binom{i}{j} \mu^j (1-\mu)^{i-j} \frac{1}{j+1}$$

Using the binomial theorem, this reduces to $\frac{1}{\mu(i+1)}(1-(1-\mu)^{i+1})$. We plug this into equation 5:

[17] See Murphy and Welch (1993) Figure 3(a).

$$\sum_{i=0}^{q} \binom{q}{i} (\frac{1}{n})^i (1 - \frac{1}{n})^{q-i} \frac{1}{\mu(i+1)} (1 - (1 - \mu)^{i+1})$$

Simplifying (and using the binomial theorem again), this yields:

$$\frac{n}{\mu(q+1)} (1 - (\frac{\mu}{n})^{q+1})$$

Since there is a positive mass of workers and firms, we take the limit as n and q approach ∞ (noting that the worker application-firm ratio equals $\frac{2}{M}$) and achieve our result.

The proof for the matching probability for a low skill worker when high skill workers are preferred (λ_{lh}) is calculated in a similar way, but the $\Pr(\textit{hired} \mid i \textit{ show up}) = \frac{(1-\mu)^i}{i+1}$. The probabilities for the labor market in which low skilled workers are preferred can be written directly using these results.

B Proof of Lemma ??

To simplify notation, let $v = \frac{2}{M}$.

(i) We prove that λ_{hh} is decreasing in μ in Shapiro (2003). To prove that λ_{lh} is decreasing in μ, consider the numerator of the derivative $\frac{d\lambda_{lh}}{d\mu}$, $v\{(1-v(1-\mu))e^{-v\mu} - e^{-v}\}$ It is easy to show this expression is increasing in μ and is equal to zero when evaluated at $\mu = 1$, implying that the expression and $\frac{d\lambda_{lh}}{d\mu}$ are negative.

(ii) The derivative $\frac{d}{d\mu}(\frac{1-e^{-v\mu}}{v\mu} - \frac{1-e^{-v\mu}}{v(1-\mu)}) = \frac{(v\mu(1-\mu)+1-2\mu)e^{-vu} - (1-\mu)^2 + \mu^2 e^{-v}}{v\mu^2(1-\mu)^2}$. As the denominator is positive, the sign of the expression is equal to the sign of the numerator. We call the numerator $F(u)$. Then:

$$F'(\mu) = (v^2\mu(1-\mu) - 2)e^{-v\mu} + 2(1-\mu) + 2\mu e^{-v}$$

$$F'''(\mu) = (-2v^2(1 + v(1-2\mu)) + v^2(v^2\mu(1-\mu) - 2))e^{-v\mu}$$

Evaluating at the endpoints, $F'(0) = 0$ and $F'(1) = 0$. Since $v\epsilon(0, 1]$ and $\mu\epsilon[0, 1]$, $F'''(\mu) < 0$, which then means that $F'(\mu)$ is concave and greater than or equal to zero for all μ. Using this fact, all $F(\mu)$ must be less than or equal to $F(1)$. Evaluating, $F(1) = 0$, which proves our conjecture.

(iii) From (ii) we know that $\frac{d}{d\mu}(\ \lambda_{hh} - \lambda_{lh}) < 0$. Evaluating $\lambda_{hh} - \lambda_{lh}$ at $\mu = 1$ (and using L'Hospital's Rule) give us $\frac{1-(1+v)e^{-v}}{v}$. The numerator is positive (it is increasing and equal to zero at $v = 0$) which gives us the result.

C Matching Firms with Wages

We are trying to prove that for each wage w, there are more firms still offering high skill tasks than there are offering low skill tasks, i.e. $1 - G_h(w) > 1 - G_l(w)$. When the high skill task is preferred, we can rewrite the inequality as $G_l(w) > G_h(w)$ or $\frac{1-\lambda_{lh}}{\lambda_{lh}}(\frac{w-\underline{w}_l}{f_l-w}) > \frac{1-\lambda_{hh}}{\lambda_{hh}}(\frac{w-\underline{w}_h}{f_h-w})$. Re-arranging yields $(\frac{1-\lambda_{lh}}{1-\lambda_{hh}})(\frac{\lambda_{hh}}{\lambda_{lh}}) > (\frac{f_l-w}{f_h-w})(\frac{w-\underline{w}_h}{w-\underline{w}_l})$, which is true since $f_h > f_l$, $\lambda_{hh} > \lambda_{lh}$, and $\underline{w}_h > \underline{w}_l$.

When the low skill task is preferred $\pi_l > \pi_h$, or $(1 - \lambda_{ll} + \lambda_{ll}G_l(w))(f_l - w) > (1 - \lambda_{hl} + \lambda_{hl}G_h(w))(f_h - w)$ for any w. Since $f_h > f_l$ we can then write $1 - \lambda_{ll}(1 - G_l(w)) > 1 - \lambda_{hl}(1 - G_h(w))$. Using the fact that $\lambda_{ll} > \lambda_{hl}$ gives us our result that $1 - G_h(w) > 1 - G_l(w)$.

D Proof of Lemma 4

We only sketch the proof here, due to the fact that it involves a large amount of algebra. The details are available upon request from the author.

First we examine how expected profits change with M. We break equation 4 down into two parts, the part associated with $f_h - \underline{w}_h$ and the part associated with $f_l - \underline{w}_l$. From Shapiro (2003) Lemma 2, we know that the part associated with $f_h - \underline{w}_h$ is decreasing in M. Define $a = \frac{1}{M}$ and $Y = e^{-2a\mu} - e^{-2a}$. The part associated with $f_l - \underline{w}_l$ can be written as: $Y(1 - \frac{Y}{2(1-\mu)a})$. Taking the derivative with respect to a and re-writing we get $\frac{2(1-\mu)a^2\frac{dY}{da} - 2a\frac{dY}{da}Y + Y^2}{2(1-\mu)a^2}$. After completing the square, the numerator becomes $2(1-\mu)a^2\frac{dY}{da}(1 - \frac{\frac{dY}{da}}{2(1-\mu)}) + (a\frac{dY}{da} - Y)^2$. If we show $\frac{dY}{da}$ is negative and that $\frac{\frac{dY}{da}}{2(1-\mu)} < 1$, the whole expression will be positive and we will have our result (by the chain rule, since $\frac{da}{dM}$ is negative). The derivative $\frac{dY}{da} = 2(e^{-2a} - \mu e^{-2\mu a})$. It is easy to show this is decreasing in μ and equal to zero when evaluated at $\mu = 1$, and hence is positive. The expression

$\frac{\frac{dY}{da}}{2(1-\mu)} = \frac{2(e^{-2a}-\mu e^{-2\mu a})}{2(1-\mu)} < \frac{e^{-2\mu a}-\mu e^{-2\mu a}}{1-\mu} = e^{-2\mu a} < 1.$

Now we look at how expected profits vary with μ. After taking the derivative of equation 4 and simplifying, we can write:

$$\begin{aligned}\frac{dE\pi}{d\mu} &= [e^{-\frac{2\mu}{M}}(1-\lambda_{hh}) + \lambda_{hh}(\lambda_{hh} - e^{-\frac{2\mu}{M}})]((f_h - \underline{w}_h) - (f_l - \underline{w}_l)) + \qquad (6)\\ &\quad [(\lambda_{hh} - \lambda_{lh})(\lambda_{hh} + \lambda_{lh} - 2e^{-\frac{2\mu}{M}})](f_l - \underline{w}_l)\end{aligned}$$

The first step is to analyze this expression when $f_h - \underline{w}_h = f_l - \underline{w}_l$, which eliminates the first term. From Lemma ??? we have shown that $\lambda_{hh} > \lambda_{lh}$. We can show that both $\lambda_{hh} - e^{-\frac{2\mu}{M}}$ and $\lambda_{lh} - e^{-\frac{2\mu}{M}}$ are increasing in μ. Moreover, when $\mu = 0$, the expression $\lambda_{hh} + \lambda_{lh} - 2e^{-\frac{2\mu}{M}}$ is negative, and when $\mu = 1$ the expression is positive. Hence the claim in the lemma follows when $f_h - \underline{w}_h = f_l - \underline{w}_l$. For $f_h - \underline{w}_h > f_l - \underline{w}_l$, we observe that the first term of equation 6 is positive since $\lambda_{hh} - e^{-\frac{2\mu}{M}}$ is positive (we noted above that this expression is increasing in μ; moreover it is equal to zero when μ equals zero (using L'Hopital's rule)). By adding this positive expression the cutoff for above which profits are increasing with μ remains but is lower than the cutoff when $f_h - \underline{w}_h = f_l - \underline{w}_l$.

E Proof of Proposition 5

The first step is to show that $\frac{2(f_h - \underline{w}_h)}{(2-\lambda_{hh})^2} > \frac{2(f_l - \underline{w}_l)}{(2-\lambda_{lh})^2}$. Since we are in a labor market where high skilled workers are preferred, we know $(f_h - \underline{w}_h) > (f_l - \underline{w}_l)$ and $\lambda_{hh} > \lambda_{lh}$, which together prove that the inequality holds.

Since we have proved $0 > \frac{d\lambda_{lh}}{d\mu} \geq \frac{d\lambda_{hh}}{d\mu}$ in Lemma ??, the result follows.

References

[1] Acemoglu, Daron. (1998), "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics*, CXII, 1055-1090.

[2] Acemoglu, Daron. (1999), "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence," *American Economic Review*, 89, 1259-1279.

[3] Acemoglu, Daron, and Shimer, Robert. (2000), "Wage and Technology Dispersion," *Review of Economic Studies*, 67, 585-608.

[4] Albrecht, James W., Gautier, Pieter A., and Vroman, Susan B. (2001), "Matching with Multiple Applications," mimeo, Georgetown University.

[5] Allen, Mike. (September 19, 1999), "Help Wanted: The Not-Too-High-Q Standard" *The New York Times*, Week In Review p.3.

[6] Berman, Eli, Bound, John and Griliches, Zvi. (1994), "Changes in the Demand for Skilled Labor within U.S. Manufacturing Industries: Evidence from the Annual Survey of Manufacturing," *Quarterly Journal of Economics*, CIX, 367-398.

[7] Burdett, Kenneth and Judd, Kenneth L. (1983), "Equilibrium Price Dispersion," *Econometrica*, 51, 955-969.

[8] Butters, Gerald R. (1977), "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies*, 44, 465-491.

[9] Juhn, Chinhui, Murphy, Kevin, and Pierce, Brooks. (1993), "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101, 410-442.

[10] Katz, Lawrence F. and Murphy, Kevin M. (1992), "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107, 35-78.

[11] Lang, Kevin, Manove, Michael, and Dickens, William. (2001), "Racial Discrimination in Labor Markets with Posted Wage Offers," mimeo, Boston University.

[12] Montgomery, James D. (1991), "Equilibrium Wage Dispersion and Interindustry Wage Differentials," *Quarterly Journal of Economics*, 109, 163-179.

[13] Shapiro, Joel. (1999), "Income Taxation and Wage Inequality in a Frictional Labor Market", mimeo, Princeton University.

[14] Shi, Shouyong. (2001), "Frictional Assignment I. Efficiency," *Journal of Economic Theory*, 98, 232-260.

[15] Shimer, Robert. (1998), "Do Good Guys Come in First? How Wage Determination Affects the Ranking of Job Applicants," mimeo, Princeton University.

[16] Shimer, Robert. (2001), "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," mimeo, Princeton University.

[17] Stahl, Dale O. (1989), "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review*, 79, 700-712.

[18] Varian, Hal R. (1980), "A Model of Sales," *American Economic Review*, 70, 651-659.