

CAHIER 9607

**RANK REGRESSIONS, WAGE DISTRIBUTIONS,  
AND THE GENDER GAP**

Nicole M. FORTIN<sup>1</sup> and Thomas LEMIEUX<sup>1</sup>

<sup>1</sup> Département de sciences économiques and Centre de recherche et  
développement en économique (C.R.D.E.), Université de Montréal

April 1996

We would like to thank Christian Gourtiéroux and Joel Horowitz for helpful comments and discussions. Fortin and Lemieux gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada and from the Fonds pour la formation de chercheurs et l'aide à la recherche of Québec.

C.P. 6126, succursale Centre-ville  
Montréal (Québec) H3C 3J7

Télécopieur (FAX) : (514) 343-5631  
Courrier électronique (E-Mail) : econo@ere.umontreal.ca

## RÉSUMÉ

Dans cet article, nous modélisons les interactions entre les distributions des salaires horaires des hommes et des femmes sous l'hypothèse que tout changement dans la distribution des salaires des femmes doit être compensé par un changement opposé dans la distribution des salaires des hommes. Cette hypothèse correspond à un modèle d'affectation pur dans lequel la position relative des femmes dans la distribution ne change pas la distribution des salaires de tous les travailleurs. Nous explorons les implications de cette hypothèse en estimant un modèle de détermination des salaires basé sur le rang. Ce modèle est estimé à l'aide d'un probit ordonné et de méthodes de lissage par noyaux.

Nous utilisons les données américaines du CPS de 1979 et 1991. Nos résultats indiquent que l'amélioration de la position relative des femmes explique pourquoi la croissance dans l'inégalité des salaires a été plus grande parmi les femmes que parmi les hommes. Nous expliquons aussi pourquoi cette inégalité s'est accrue plus rapidement dans le haut de la distribution que dans le bas chez les hommes, alors que le contraire est vrai chez les femmes. Enfin, nous expliquons l'évolution irrégulière de l'écart-type intra-groupe des salaires parmi les différents groupes de qualifications d'hommes et de femmes. Nous concluons que les changements dans la distribution des salaires féminins et masculins peuvent être décrits adéquatement par un accroissement (non linéaire) des prix des qualifications combiné avec le déclin du salaire minimum et l'amélioration dans la position relative des femmes dans le cadre d'un modèle d'affectation pur.

Mots clés : régression de rang, inégalité des salaires, écart de salaire hommes-femmes

## ABSTRACT

In this paper, we model the interactions between the distribution of male and female wages under the assumption that any change in the wage distribution of women must be offset by an opposite change in the wage distribution of men. This assumption is consistent with a pure assignment model in which the relative position of women in the wage distribution does not affect the overall distribution of wages. We explore the implications of this assumption by estimating a rank-based model of wage determination. The model is estimated using ordered probit and kernel smoothing methods.

Using CPS data for 1979 and 1991, we find that the improvement in the relative position of women explains why wage inequality increased faster among women than among men. It also explains why inequality increased faster in the upper tail than in the lower tail of the distribution for men, while the opposite happened for women. It finally explains the erratic pattern of changes in the within-group standard deviation of wages across different skill groups of men and women. We conclude that changes in the distribution of male and female wages are described well by a (non-linear) rise in skill prices combined with a declining minimum wage and an improvement in the relative position of women in the context of a pure assignment model.

Key words : rank regression, wage inequality, gender gap

Ce cahier a également été publié au Centre de recherche et développement en économique (C.R.D.E.) (publication no 1096).

Dépôt légal - 1996  
Bibliothèque nationale du Québec  
Bibliothèque nationale du Canada

ISSN 0709-9231

## 1. INTRODUCTION

During the 1980s, the U.S. labor market was characterized by two major trends: a pervasive increase in wage inequality and a striking decline in the male-female wage gap. There have been relatively few attempts to find a common explanation for these two trends. Explanations for the rise in inequality have focused on increasing returns to skills due to skill-biased technological shocks (Mincer (1992), Bound and Johnson (1992)) and international trade (Murphy and Welch 1991), and on institutional changes (DiNardo, Fortin, and Lemieux 1995). By contrast, the rise in women's relative experience levels and occupational status have been identified as the key factors in the closing of the gender gap over the 1980s (O'Neill and Polachek (1993), Blau and Kahn (1994a), Blau and Kahn (1994b)). Attempts to reconcile the two trends suggest that, if anything, wage inequality and the gender gap should move in opposite directions. A rise in returns to skills should increase the gender gap if part of the gap is due to skill differences between men and women (Juhn, Murphy, and Pierce (1991), Blau and Kahn (1994a), Blau and Kahn (1994b)). Topel (1994) has suggested a possible link between increased labor force participation of women and the rise in inequality. However, his model of factor demand implies a widening, as opposed to a closing, of the gender gap.

Using graphical techniques and a new rank-based procedure, this paper shows that the link between the changes in wage inequality and changes in the gender gap may be tighter than previously thought. We explain several important features of changes in the distribution of male and female wages by assuming that any change in the wage distribution of women must be offset by an opposite change in the wage distribution of men. This assumption is consistent with a pure assignment model in which the relative position of women in the wage distribu-

tion does not affect the overall distribution of wages. More generally, we argue that important features of the change in the distribution of wages are overlooked by considering men and women in isolation.

In Section 2, we use CPS data for 1979 and 1991 to illustrate the clear convergence in the shapes of the male and female wage distributions. We focus on the years 1979 and 1991 as the two years that encompass the 1980s and for which comparable data are available. We also show that the shape of the overall wage distribution changed much less than the respective shapes of the male and female wage distributions. This observation motivates our use of a pure assignment model for understanding the link between wage inequality and the relative position of women in the wage distribution.

In Section 3, we propose an econometric framework to estimate the relative position of women in the wage distribution of men, conditional on observed characteristics such as education and experience. This framework is based on a discretization of the wage distribution into approximately 150 wage intervals, similar to the centile-by-centile analysis of Juhn, Murphy, and Pierce (1993). The effect of individual characteristics on the position in the wage distribution is then estimated using an ordered probit model. The wage distribution *per se* is obtained using kernel smoothing methods.

This framework enables us to describe the distribution of male and female wages in terms of three components: 1) a "skill index" that captures both observed (experience, education, etc.) and unobserved measures of skills, 2) a returns-to-skill function that maps this skill index into wages, and 3) a "relative position locus" that indicates the relative position of women in the skill distribution of men. Changes in inequality over time can be decomposed in terms of changes in these three components using a procedure in the spirit of familiar Oaxaca de-

compositions. These decomposition are, however, partial equilibrium by nature. They cannot capture the interactions between the wage distributions of men and women. By contrast, the pure assignment model provides a simple way of modeling the interaction between the wage distributions of men and women in general equilibrium. We thus propose an alternative decomposition in which changes in the relative position of women do not affect the overall distribution of wages.

The estimation results and the decompositions are presented in Section 4. We conclude in Section 5 that changes in the distribution of wages for men and women are described well by a (non-linear) rise in returns to skill combined with a declining minimum wage and an improvement in the relative position of women in the context of a pure assignment model.

## 2. CONVERGENCE IN DISTRIBUTIONS AND WAGE INEQUALITY

### 2.1. Descriptive Evidence on the Convergence in Distributions

Studies of changes in the distribution of wages have focused traditionally on summary measures of wage inequality such as the variance of log wages and the Gini coefficient. More recently, however, graphical procedures have been used to illustrate what happens where in the distribution. These procedures have been successfully used to show that returns to skill increased everywhere in the distribution (Juhn, Murphy, and Pierce (1993)) and that the minimum wage played a major role in changes in wage inequality (DiNardo, Fortin, and Lemieux (1995)).

Following DiNardo, Fortin, and Lemieux (1995), we use weighted kernel methods to estimate the density of log hourly wages of men and women in 1979 and

1991.<sup>2</sup> These densities are estimated using the large samples available in the outgoing rotation group files of the Current Population Survey (CPS) for all workers aged 16 to 65.<sup>3</sup> As in DiNardo, Fortin, and Lemieux (1995), we use an hourly wage measure as opposed to the measure of average weekly earnings on all jobs that has been used in many studies of changes in the U.S. structure of wages. Each observation is weighted by the product of the CPS weight and usual hours of work per week. These "hours-weighted" estimates put more weight on workers who supply a large number of hours to the market. They reflect the distribution of wages per hour worked in the economy rather than the distribution of wages per worker. Summary statistics of the 1979 and 1991 CPS samples for men and women are reported in Table 1.

The estimated densities of real log wages are reported in Figure 1. The density for men and women (solid and broken lines) are weighted by their respective shares of the total workforce. These two weighted densities therefore add up to the overall density of log wages for all workers (dotted line). The figure illustrates the well-known fact that the male and female distributions, as well as the overall distribution of wages, widened between 1979 and 1991. It also indicates that the difference between mean wages (indicated by a vertical line) of men and women—the unadjusted gender gap—shrank during this period. As noted by DiNardo, Fortin, and Lemieux (1995), the minimum wage has a clear visual impact on the shape of the 1979 distribution for both men and women.

<sup>2</sup>These kernel density estimates use the optimal bandwidth calculated with (Sheather and Jones 1991)' selector and a Gaussian kernel function. This bandwidth selection method has been shown to be reliable in a number of real data sets and Monte-Carlo studies (Park and Turlach (1992), Sheather (1992)).

<sup>3</sup>In addition, only individuals reporting an hourly wage between \$1 and \$100 (in 1979 dollars) were kept in the sample. Observations with allocated wages were not deleted from the sample because of a coding error in the CPS, it is impossible to identify most workers with allocated wages in 1991. The GDP deflator for personal consumption expenditures was used to convert nominal wages into 1979 dollars.

In addition to these well-known facts, three interesting patterns emerge from Figure 1. First, it is clear that the convergence between the male and female distribution of wages goes far beyond a simple convergence in means. Figure 1 shows that the shapes of the two distributions become more similar during this period. The distribution of wages of women is clearly skewed to the right in 1979 (coefficient of skewness of 0.511) whereas the men's distribution is skewed to the left (coefficient of skewness -0.129). By contrast, the 1991 distributions are much less skewed and have more similar shapes; the coefficients of skewness are 0.288 and -0.007 for women and men respectively.

It is important to note that this convergence in shapes was achieved by changes in both the shapes of male and female distributions. The shape of the female wage distribution looks increasingly like the male distribution and vice versa. Finally, the overall distribution did not change as much as the male and female wage distributions. For example, the standard deviation of log wages increased by 14.9 percent for men and 25.0 percent for women, but only by 11.9 percent for men and women combined. Furthermore, the coefficient of skewness of the overall distribution remained roughly constant from 0.173 in 1979 to 0.147 in 1991.

This suggests that the wage convergence between men and women occurred while leaving the overall distribution of wages relatively unchanged. In other words, changes in the distribution of female wages are offset by opposite changes in the distribution of male wages.

This corresponds to the case of a "dog bone" economy (Sattinger (1993)) in which the distribution of wages (size of bones) does not depend on who gets which wage. The assignment of wages to workers depends on an assignment rule that maps skills and other factors like gender into a particular position in the overall wage distribution. In a pure assignment model, the overall distribution

of wages in unaffected by changes in the assignment rule. The male and female wage distributions only depend on a fixed overall distribution of wages and on the relative position of women in the distribution of male wages. This relative position depends, in turn, on the role of gender in the assignment rule, conditional on skills.

This pure assignment model is based on a set of strong assumptions that we do not attempt to test here. Whether an improvement in the relative position of women tends to widen or compress the overall distribution of wages critically depends on why men and women earn different wages in the first place. A simple explanation for the unfavorable position of women in the distribution of wages is that they are not as "skilled" as men. In this case, women could only improve their position by acquiring more skills, which should in turn change the overall distribution of wages through supply and demand effects.<sup>4</sup> Alternatively, women may earn less than men because they are victims of taste discrimination (Becker 1957), expand lower effort (Becker 1985), have lower promotion probability (Lazear and Rosen 1990), or are crowded into low paying jobs (Bergman 1974). In this last case, it is not clear why an improvement in the position of women due to less occupational segregation should affect the overall distribution of wages.

The pure assignment model should thus be viewed as a "straw man" whose value depends on whether it can explain interesting features of the change in wage inequality during the 1980s. In the next section, we illustrate with a simple example some patterns of changes in wage inequality that can be explained by a pure assignment model. We look explicitly at the contribution of the model to observed changes in wage inequality in Section 4.

<sup>4</sup>See Teulings (1995) who estimates supply and demand effects in an assignment model of the distribution of wage in the Netherlands.

## 2.2. Gender Wage Differentials and Gender-Specific Inequality:

### An Illustrative Example

Consider an economy in which there are two types of jobs. A "skilled" job that pays 20 dollars an hour, and an "unskilled" job that pays 10 dollars an hour. An equal number of men and women are in the workforce. Ninety percent of women hold an unskilled job and ten percent of hold a skilled job. The proportions are just the reverse for men. The wage distributions of men and women are illustrated in Figure 2a. Both distributions have a variance equal to 9; the gender wage gap is equal to 8 dollars.

Now consider a change in the proportion of women holding unskilled jobs from 0.9 to 0.8 (0.1 to 0.2 for skilled jobs) and a symmetric change in proportions for men (from 0.9 to 0.8 in skilled jobs, from 0.1 to 0.2 in skilled jobs). The resulting distributions are illustrated in Figure 2b. This occupational upgrading of women has three main consequences. It reduces the gender gap from 8 to 6 dollars, increases the variance among men and among women from 9 to 16, and reduces the difference in the shapes of the two distribution. Note also that the change in proportions has no effect on the shape of the overall distribution (50 percent skilled, 50 percent unskilled) or on its variance which remains at 25.

This type of change in the distribution of jobs among men and women corresponds to some extent to the convergence in distributions described in section 2.1. In this simple example, when the proportion of the workforce in each job and the wage attached to these jobs remain fixed, the wage gap can only be closed by increasing wage dispersion among men and among women.

A generic increase in returns to skill could also increase wage inequality in this example. Figure 2c shows the case in which a rise in returns to skill increases the wage difference between skilled and unskilled jobs by 33 percent. Since women are



concentrated in unskilled jobs, this increase in the returns to skill has an adverse impact on the relative wages of women (Juhn, Murphy, and Pierce (1991), Blau and Kahn (1994a), Blau and Kahn (1994b)). As a result, the gender wage gap increases from 8 dollars to 10.67 dollars. The rise in returns to skill also increases the variance of wages among women (and among men) from 9 to 16. The overall variance increases to 44.44.

The occupational upgrading illustrated in Figure 2b and the rise of returns to skill in Figure 2c thus have a similar effect on the male and female variances. How can these two sources of change in wage inequality be distinguished in practice? We propose a distinction based on the relative position of women in the distribution of male wages. This relative position can be illustrated by simply plotting the location of women in the male wage distribution, that is, the percentiles of the male distribution as a function of the percentiles of the female distribution. For the sake of the exposition, it is useful to think of these percentiles as a normalized measure of skills uniformly distributed between 0 and 1. Jobs among men and among women are attributed on the basis of these skills. The least skilled workers hold the unskilled job while the most skilled hold the skilled job. However, perhaps because of discrimination (or tastes), women are crowded into low-paying jobs.

The relation between male and female skills is shown in Figure 2d. The solid line, that we call a "relative position locus", represents this relation in the base case. In the first ten percentiles of the distribution of skills, low skilled men and women hold the same unskilled job, women are in the same position as men with identical skills. The relative position locus, which indicates how much "extra skill" a woman needs to earn the same wage as a man with comparable skills, is thus a 45 degree line. For the next eighty percentiles, women are crowded in the low skilled jobs and do only as well as men with a skill level of 0.10. The relative

position locus is thus an horizontal line in this section of the graph. In the last ten percentiles of the distribution, high skilled men and women hold the same skilled job and the locus returns to the 45 degrees line.

The dotted line in Figure 2d is the relative position locus corresponding to the distribution of wages of Figure 2b. It shows that the convergence in the proportion of men and women in each job moves the relative position locus closer to the 45 degree line. By contrast, the relative position locus is unaffected by the increase in returns to skills depicted in Figure 2c. There is therefore a simple distinction between the increases in variance under the two above scenarios. The increase in Figure 2b is due to an improvement in the relative position of women, holding returns to skill constant. On the other hand, the increase in variance in Figure 2c is due to a rise in returns to skill, holding the relative position constant.

To distinguish between the two cases, one must thus decompose changes in wage inequality into a component due to changes in returns to skill for a given relative position locus, and a component due to changes in this locus. In the case illustrated in Figure 2d, there are no covariates and the relative position can be estimated simply by plotting percentiles of the wage distribution of men and women. We next develop an econometric framework to estimate this relative position in the presence of covariates.

### 3. ECONOMETRIC METHODS

#### 3.1. Wage Determination Model

We model the distribution of wages in terms of the distribution of a skill index, a returns-to-skill function, and a relative position locus. This framework allows us to account for changes over time in wage dispersion and other features of the

wage distribution such as the minimum wage.

Assume that there are two groups of workers in the population, males and females. Let  $\mathcal{P} = \mathcal{M} \cup \mathcal{F}$  represent the set of all workers, where  $\mathcal{M}$  and  $\mathcal{F}$  represent the sets of male and female workers, respectively. The cumulative distribution of log wages,  $w$ , is

$$F(w) = s_f F_f(w) + s_m F_m(w), \quad (1)$$

where  $s_f, g = f, m$ , represent the share of females and males in the workforce, and  $F_g$  represents the cumulative distribution.

We adopt a flexible formulation with the sole assumption that the rank of workers in the wage distribution is equal their rank in the distribution of a skill index. Let  $w_i$  denote the log wage of person  $i$ , and let  $r_i^g$  be a skill index such that

$$\begin{aligned} \text{rank}_{i \in \mathcal{M}}(w_i) &= \text{rank}_{i \in \mathcal{M}}(r_i^m) & \text{or equivalently} & & F_m(w_i) &= \bar{F}_m(r_i^m), & (2) \\ \text{rank}_{i \in \mathcal{F}}(w_i) &= \text{rank}_{i \in \mathcal{F}}(r_i^f) & \text{or equivalently} & & F_f(w_i) &= \bar{F}_f(r_i^f), \end{aligned}$$

where  $F_g(w)$  and  $\bar{F}_g(r_i^g)$ ,  $g = f, m$ , represent the cumulative distribution of the random variables  $w_i$  and  $r_i^g$ .

The skill index  $r_i^g$  can be interpreted as a latent variable that determines the rank, or position, of worker  $i$  in the distribution of wages. It is defined as the sum of a function of an observed skill vector  $X_i$  (education, experience, etc.) and an unobserved skill  $\epsilon_i$

$$\begin{aligned} r_i^m &= X_i \beta^m + \epsilon_i, & \text{for all } i \in \mathcal{M}, \\ r_i^f &= X_i \beta^f + \epsilon_i, & \text{for all } i \in \mathcal{F}, \end{aligned} \quad (3)$$

where the  $\epsilon_i$  are assumed to be i.i.d. normal with mean 0 and variance 1. Differences between  $\beta^m$  and  $\beta^f$  could, for example, be due to differences in the

relationship between potential and actual experience for men and women.

Since the rankings (2) are monotonic, we can write equations (3) as the transformation models

$$\begin{aligned} \Lambda_m(w_i) &= X_i \beta^m + \epsilon_i, & \text{for all } i \in \mathcal{M}, \\ \Lambda_f(w_i) &= X_i \beta^f + \epsilon_i, & \text{for all } i \in \mathcal{F}, \end{aligned} \quad (4)$$

where  $\Lambda_g(\cdot) = \bar{F}_g^{-1}(F_g(\cdot))$  are unknown monotonic increasing functions, to be estimated nonparametrically.

The inverse functions  $\Lambda_f^{-1}(r_i^f)$ ,  $g = f, m$ , are returns-to-skill functions that indicate how the skill index  $r_i^g$  is valued in the labor market

$$w_i = \Lambda_m^{-1}(r_i^m), \quad \text{for all } i \in \mathcal{M}, \quad (5)$$

$$w_i = \Lambda_f^{-1}(r_i^f), \quad \text{for all } i \in \mathcal{F}. \quad (6)$$

Note that the returns-to-skill functions indicate how both observable and unobservable measures of skills are valued in the market. If skills of men and women were valued equally in the labor market, the returns-to-skill functions,  $\Lambda_g^{-1}(r_i^g)$ ,  $g = f, m$ , would be identical.

However, the returns-to-skill functions for male and female are typically not identical. A woman with given skills may earn less than a man with identical skills because of discrimination. Alternatively, men and women with the same skill index may not be equally productive because of systematic gender differences in omitted skills like physical strength.

We use a relative position locus to summarize how skills of women are under-valued relative to skills of men. This relative position locus is represented by a function  $\Psi(\cdot)$  defined as

$$\Psi(r^*) = \Lambda_m(\Lambda_f^{-1}(r^*)). \quad (7)$$



It represents the level of skills of a man who earns the same wage as a woman with skills  $r^*$ . Using the definition of the relative position locus, the wage equation (6) for women can be rewritten as

$$w_i = \Lambda_m^{-1}(\Psi(r_i^*)), \quad \text{for all } i \in \mathcal{F}. \quad (8)$$

The wage equations (5) and (8) and the skill index equation (3) represent our econometric model of the distribution of wages. The three basic elements of the model are the returns-to-skill function  $\Lambda_m^{-1}(\cdot)$ , the relative position locus  $\Psi(\cdot)$ , and the distribution of the skill index  $r^*$ .

### 3.2. Estimation

The estimation of transformation models, like equation (4), is a topic of continuing research. Very general (Horowitz (1993), Han (1987), Bickel and Doksum (1981)) or more specific transformation models (Heckman and Singer (1984), MacKinnon and Magee (1990)) have been considered in the literature. The Box-Cox transformation (Box and Cox (1964)) and the proportional hazard model of Cox (1972) are particularly well-known examples of such models. Here, we take advantage of our large samples to use a simple procedure. We divide the wage distribution into one to two hundred intervals and use an ordered probit model to approximate the transformation function with a step function. This procedure also yields predicted probabilities of belonging to each wage interval and facilitates the construction of predicted densities.<sup>5</sup>

To implement this procedure, we use  $J - 1$  thresholds,  $a_j, j = 1, \dots, J - 1$ , to

<sup>5</sup>This methodology is analogous to that of Donald, Green, and Paarech (1998) who use Cox's proportional hazard model.

divide the wage distribution into  $J$  intervals.<sup>6</sup> Equations (3), (5), and (8) can be used to compute the probability of being in each interval  $[a_{j-1}, a_j]$

$$\begin{aligned} \Pr(a_{j-1} \leq w_i < a_j) &= \Pr(\Lambda_j(a_{j-1}) \leq r_i^* < \Lambda_j(a_j)) \\ &= \Pr(\lambda_{j-1}^j \leq X_i \beta^j + \epsilon_i < \lambda_j^j) \\ &= \Pr(\epsilon_i < \lambda_j^j - X_i \beta^j) - \Pr(\epsilon_i \leq \lambda_{j-1}^j - X_i \beta^j) \\ &= \Phi(\lambda_j^j - X_i \beta^j) - \Phi(\lambda_{j-1}^j - X_i \beta^j), \end{aligned} \quad j = 1, \dots, J \quad (9)$$

where  $\Lambda_j(a_0) = \lambda_0^j = -\infty$ ,  $\Lambda_j(a_J) = \lambda_J^j = +\infty$ , and where  $\Phi$  denotes the cumulative normal distribution ( $\epsilon_i$  is i.i.d. normal). This formulation corresponds to an ordered probit model with a set of thresholds  $\lambda_j^j, j = 1, \dots, J$ . The parameters  $\lambda_j^j$  and  $\beta^j$  can be estimated with maximum likelihood techniques.

The estimates  $\hat{\lambda}_j^j$  can be used to approximate the transformation function  $\Lambda^j(\cdot)$  by the step function

$$\hat{\Lambda}_j(w_i) = \sum_{j=0}^J I_{\{\epsilon_{i-1} \leq w_i < \epsilon_i\}} \lambda_j^j, \quad (10)$$

where  $I_A$  is an indicator function for the set  $A$ .

The returns-to-skill functions are then obtained by inverting the step function  $\Lambda_j(\cdot)$

$$w = \hat{\Lambda}_j^{-1}(\tau) = \sum_{j=0}^J I_{\{\lambda_j^j \leq \tau < \lambda_{j+1}^j\}} a_{j+1}. \quad (11)$$

The relative position locus  $\Psi(\cdot)$  is similarly obtained as

$$\hat{\Psi}(\tau) = \hat{\Lambda}_m^{-1}(\tau) = \sum_{j=0}^J I_{\{\lambda_j^m \leq \tau < \lambda_{j+1}^m\}} \lambda_{j+1}^m. \quad (12)$$

To compute the predicted empirical densities, we first find the predicted prob-

<sup>6</sup>In practice, we choose the thresholds  $a_j$  to divide the sample into a series of wage intervals that contain approximately the same number of observations.

abilities of being in each wage interval

$$\begin{aligned} \hat{\pi}_j^g &= \pi_j(\hat{\Lambda}_g, \hat{\beta}^g, X_g) \\ &= \frac{1}{N_g} \sum_{i \in G} \left\{ \Phi(\lambda_i^g - X_i \hat{\beta}^g) - \Phi(\lambda_{i-1}^g - X_i \hat{\beta}^g) \right\} \end{aligned} \quad \text{for } j = 1, \dots, J, \quad (13)$$

where  $\Phi$  denotes the cumulative normal distribution,  $X_g$  is a matrix that contains the characteristics of individuals  $g \in G$ , and  $N_g$  is the number of workers in group  $g$ . Since equation (7) implies that  $\Lambda_j(w) = \Psi^{-1}(\Lambda_m(w))$ , the predicted probabilities for women can also be written as

$$\hat{\pi}_j^f = \pi_j \left[ \hat{\Psi}^{-1}(\hat{\Lambda}_m), \hat{\beta}^f, X_f \right] \quad (14)$$

These predicted probabilities can be used to estimate the density of wages using kernel smoothing methods. Let  $w_j$  be the mean log wage in the wage interval  $j$ . A relatively accurate kernel density estimate of  $f^j(w)$  is obtained by considering each wage interval as a bin

$$\hat{f}_h^j(w) = \sum_{i=0}^j \frac{\hat{\pi}_i^j}{h^*} K \left( \frac{w - w_i}{h^*} \right), \quad (15)$$

where  $h^*$  is the optimal bandwidth for the kernel density estimate over all observations and  $K(\cdot)$  is the kernel function.<sup>7</sup> Predicting the relative weights for each wage interval is thus sufficient to obtain a reliable density estimate.

<sup>7</sup>In this paper, we use the plug-in method of Sheather and Jones (1991) as bandwidth selector. This bandwidth selection method has been shown to be reliable in a number of real data sets and Monte-Carlo studies (Park and Turlach (1992), Sheather (1992)). The benefits of using discretization methods in kernel density estimation have been studied by Fan and Marron (1994), Hall and Wand (1993).

### 3.3. Decomposition Methods

The methodology proposed to obtain predicted densities of wages can also be used to obtain counterfactual densities and decompose changes in the density of wages over time. In general, the possible explanations for changes over time in the distribution of men and women wages can be classified in three categories: 1) "quantity" effects arising from changes in the distribution of labor market qualifications ( $X$ ) and changes in labor market participation ( $\theta_g$ ); 2) "price" effects arising from changes in the return-to-skill function ( $\Lambda_m^{-1}$ ) and changes in the relative rewards to particular skills ( $\beta^m$ ); 3) relative position effects due to changes in the relative position locus ( $\Psi$ ) and changes in the relative returns to skill ( $\beta^f/\beta^m$ ).<sup>8</sup> Note that a change in the minimum wage is a potential source of change in the returns-to-skill function.

We do not attempt to account for all these factors in this paper. In light of Figure 1, we simply focus on the role of the minimum wage and the relative position locus. The former has a large visual impact on the bottom of the distribution while the latter is a natural candidate to explain the convergence in the shapes of the distribution of male and female wages.

We evaluate the effect of a change in relative position by constructing a counterfactual density in year  $t$  using the relative position locus  $\hat{\Psi}_t$  that prevailed in year  $s$ , and by contrasting this counterfactual density with the actual density of wages in year  $t$ . The effect of a change in the relative position locus on summary measures of wage dispersion, such as the variance of log wages, is obtained similarly. Using the notation introduced in equation (14), the "counterfactual"

<sup>8</sup>The change in relative returns to skill may be due, however, to mismeasurement of experience.

probabilities  $\hat{\pi}_{jt}^d$  for women are defined as

$$\hat{\pi}_{jt}^d = \pi_j [\Phi_s^{-1}(\hat{\Lambda}_{mt}), \hat{\beta}_t^d, X_{jt}], \quad (16)$$

where the subscripts  $t$  and  $s$  indicate the time period. These counterfactual probabilities are used to compute a counterfactual density of wages using equation (15).

This decomposition exercise is in the spirit of familiar Oaxaca decomposition. It is a partial equilibrium exercise by nature. Our framework can also be used, however, to carry over a general equilibrium decomposition in the context of a pure assignment model.

The predicted cumulative distribution for group  $g$  in year  $t$  is obtained as the sum of the predicted probabilities

$$\bar{F}_{gt}(w_j) = \sum_{k=1}^J \hat{\pi}_{kt}^g, \quad (17)$$

and the overall cumulative distribution is given by

$$\bar{F}_t(w_j) = s_{ft}\bar{F}_{ft}(w_j) + s_{mt}\bar{F}_{mt}(w_j). \quad (18)$$

Since the overall cumulative distribution remains unchanged in the pure assignment model, the following relation must hold

$$s_{ft}\bar{F}_{ft}(w_j) + s_{mt}\bar{F}_{mt}(w_j) = s_{ft}\bar{F}_{ft}^*(w_j) + s_{mt}\bar{F}_{mt}^*(w_j), \quad (19)$$

where the superscript  $e$  refers to the counterfactual distribution in a pure assignment model. Since general equilibrium adjustments in a pure assignment model does not affect the relative position of men and women in the overall distribution,

it follows that

$$\frac{\bar{F}_{ft}^*(w_j)}{\bar{F}_{mt}^*(w_j)} = \frac{\bar{F}_{ft}(w_j)}{\bar{F}_{mt}(w_j)}, \quad (20)$$

where the superscript  $d$  refers to the corresponding counterfactual distribution obtained in partial equilibrium. The pure assignment counterfactual distributions  $\bar{F}_{ft}^*(w_j)$  and  $\bar{F}_{mt}^*(w_j)$  are obtained by solving equations (19) and (20) for  $j = 1, \dots, J$ . The predicted probabilities  $\hat{\pi}_{jt}^g$  can then be recovered from the cumulative distributions and used to estimate counterfactual densities using kernel methods.

Finally, we evaluate the effect of changing the minimum wage by estimating the returns-to-skill functions that would have prevailed in year  $t$  if the minimum wage had remained at its year  $s$  level. To estimate the counterfactual returns-to-skill function  $(\hat{\Lambda}_{gt}^w)^{-1}$ , we first fit a piecewise linear function to the returns-to-skill functions in year  $s$ . We then use the fitted function to predict the effect of a change in the minimum wage. The corresponding "counterfactual" probabilities are given by

$$\hat{\pi}_{jt}^w = \pi_j(\hat{\Lambda}_{gt}^w, \hat{\beta}_t^g, X_{gt}), \quad (21)$$

This procedure is discussed in more details in Section 4.2.

## 4. RESULTS

### 4.1. Estimates of the Ordered Probit Model

Table 2 reports the maximum likelihood estimates of the parameter vector  $\beta^g$  in the wage equation (5) for 1979 and 1991. As in a standard probit model, these parameters are estimated only up to scale. They are normalized by the standard deviation of the error terms  $\epsilon$  and represent the returns to observable

characteristics  $X_i$  relative to the unobservable characteristic  $\epsilon$ . The explanatory variables used in the model are years of education, a quartic in years of potential experience (age-education - 6), and a dummy variable for non-white.

The results indicate that, over the period 1979-91, the relative returns to education increased by a third for men and by a fourth for women. By contrast, the relative returns to experience remained unchanged for men.<sup>9</sup> For women, however, relative returns to (potential) experience increased significantly. One simple explanation for this difference between men and women is that, on average, the actual labor market experience of women increased for a given value of potential experience. It may thus be incorrect to interpret the changes in the parameters on potential experience as a rise in returns to skill.

The coefficient on non-white is large and significant for men but does not significantly change over the time. The corresponding coefficient for women is much smaller and decreases significantly between 1979 and 1991. In other words, the relative position of non-white women in the distribution of white women decreased between 1979 and 1991 while the relative position of non-white men was unaffected. Card and Lemieux (1995) find similar results for the 1979-89 period using a different estimation technique.

The estimated thresholds ( $\lambda_j^*$ ) for 1979 and 1991 are plotted as returns-to-skill functions in Figure 3a and 3b respectively. These functions are obtained by plotting the mean log wage of each wage interval  $j$  against the corresponding  $\lambda_j^*$ . Note that the skill index  $r^* = X\beta + \epsilon$  has been normalized to have a mean zero across all workers in 1979 and 1991. The value of the returns-to-skill function evaluated at  $r^* = 0$  thus represents the wage of a worker with average observed and unobserved characteristics.

<sup>9</sup>The experience profiles for men in 1979 and 1991 are almost indistinguishable from each other.

One noticeable feature of the returns-to-skill functions is the clear break at values around the minimum wage. There is a clear plateau at the 1979 minimum wage and steep returns to skill below this plateau. The situation is not as clear in 1991 because the minimum wage was raised from 3.85 to 4.25 dollar in April 1991. In addition, there is a substantial plateau at 5.00 dollars (2.73 in 1979 dollars) because of a spike in the wage distribution at this value. As in 1979, however, the returns to skill in 1991 are steeper below the minimum wage. The remainder of the 1991 returns-to-skill function is quite linear. In fact, the middle and upper parts of the returns-to-skill functions are almost indistinguishable from a linear function except for the small plateaus due to spikes in the wage distribution. By contrast, the corresponding section of the 1979 returns-to-skill function is slightly concave for men and slightly convex for women. Relative to women, the returns-to-skill function for men becomes "more convex" between 1979 and 1991.<sup>10</sup>

The minimum wage is the most important source of non-linearity in the returns-to-skill functions. This highlights the importance of approximating these functions with a flexible functional form such as a step function. More restricted functional forms like the Box-Cox are not flexible enough to account for these minimum wage effects and could yield misleading results.<sup>11</sup>

The estimates of the relative position locus for 1979 (dotted line) and 1991 (solid line) are plotted in Figure 3c. Between 1979 and 1991, the relative position locus moved closer to the 45 degrees line in everywhere except in the two tails of the distribution of skills. In other words, women do not need as much extra skills

<sup>10</sup>Card and Lemieux (1995) find the same relative pattern between 1979 and 1989 using mean wages for detailed skill categories based on experience and education.

<sup>11</sup>One simple function that would fit the data reasonably well is a step-wise linear function with three segments: one segment for the minimum wage plateau, one below it, and one above it.

as in 1991 to earn the wage of a similarly skilled man. The change is especially striking in the upper middle of the skill distribution where the "skill gap" declined by almost a half.

The minimum wage compresses male-female wage differences, or indeed all wage differences, in the lower part of the distribution. This could explain the absence of change in the relative position locus in that part of the distribution. Similarly, the topcoding of weekly earnings at 999 dollars in 1979 and at 1923 dollars in 1991 may explain the similar absence of change at the higher end of the distribution.

#### 4.2. Counterfactual Returns-to-Skill Functions

The effect of changes in the minimum wage and in the discrimination function on the distribution of wages can be evaluated by modeling the effect of these factors on the returns-to-skill functions. In Figure 4a and 4b, we show how the returns-to-skill function for men and women in 1991 would have changed if the minimum wage had been raised back to its (real) 1979 value. The solid line illustrate the original 1991 functions while the dotted line represent these functions with the 1979 minimum wage. The dotted line is obtained by extending the the 1991 returns-to-skill function below the 1991 minimum to the 1979 minimum wage and by constructing a plateau at the 1979 minimum wage.<sup>12</sup> The resulting counterfactual function can be used to compute counterfactual probabilities using equation (21). They result in the counterfactual densities pictured in Figure 4c,

<sup>12</sup>These "counterfactual" functions are obtained in two steps. We first fit a robust regression for values of the 1991 returns-to-skill below the 1991 minimum. The estimated regressions are then used to extrapolate this segment of the function up to the 1979 minimum wage. We then create a minimum wage "plateau" using the regression line obtained by fitting a robust regression through the 1979 plateau. This line is used to join the actual 1991 function to the segment extrapolated from below the minimum wage.

where the density for all workers is obtained as a weighted sum of the male and female densities.

Up to this point, we have implicitly assumed that partial equilibrium changes in the relative position of women only has an effect on the returns-to-skill function of women. An alternative assumption is that it both affects the male and female returns-to-skill function. This would be the case if women were underpaid because of discrimination while men were overpaid because of nepotism. As mentioned earlier, Figure 1 suggests that the shape of the male and female wage distributions converged toward each other during the 1980s. The latter assumption thus seems more realistic on empirical grounds.

The counterfactual 1991 returns-to-skill functions with the 1979 relative position are obtained by increasing the horizontal distance between the returns-to-skill functions of men and women by an amount corresponding to the horizontal shift of the relative position locus. These counterfactual returns-to-skill functions are pictured in Figure 4c. Under the assumption that the shift is inversely proportional to the share of the group in the workforce, the counterfactual returns-to-skill functions  $(\Lambda_{mi}^d)^{-1}(r^*)$  and  $(\Lambda_{fi}^d)^{-1}(r^*)$  are

$$(\Lambda_{mi}^d)^{-1}(r^*) = \Lambda_{mi}^{-1}(r^*) - s_{fi} \{ [\Lambda_{mi}^{-1}(r^*) - \Lambda_{fi}^{-1}(r^*)] - [\Lambda_{ms}^{-1}(r^*) - \Lambda_{fs}^{-1}(r^*)] \}, \quad (22)$$

$$(\Lambda_{fi}^d)^{-1}(r^*) = \Lambda_{fi}^{-1}(r^*) + s_{mi} \{ [\Lambda_{mi}^{-1}(r^*) - \Lambda_{fi}^{-1}(r^*)] - [\Lambda_{ms}^{-1}(r^*) - \Lambda_{fs}^{-1}(r^*)] \}, \quad (23)$$

where  $t$  refers to year 1991 and  $s$  to year 1979. These counterfactual returns-to-skill functions are used to compute the corresponding counterfactual probabilities using the procedure of equation (16). Figure 5d illustrates the densities obtained by using these counterfactual returns-to-skill functions (in addition to the minimum wage modification) while allowing the densities to widen consequently. Figure 5e illustrates the effects of using these counterfactual densities in conjunction

with the restrictions implied by the pure assignment model (equations (19) and (20)).

#### 4.3. Predicted and Counterfactual Densities

The estimated models of Section 4.1 are used to compute the predicted 1979 and 1991 wage distributions. The corresponding kernel density estimates for men, women, and men and women together are reported in Figure 5a and 5b. Note that the individual densities for men and women (solid lines) are weighted by their respective shares of the workforce and sum up to the density for all workers. The densities based on the predicted distribution are indistinguishable from the raw densities of Figure 1, except in the tails.

The means and several standard measures of wage dispersion for these predicted distributions are reported in columns 1 and 2 of Table 3. All the statistics reported in Table 3 are based on trimmed distributions from which the lowest and highest one percent of observations have been deleted. The data was trimmed because small prediction errors in the extreme tails of the distribution have an unpredictable impact on the standard deviation and the coefficient of skewness. We find, like others, that all the measures of wage dispersion increased between 1979 and 1991. As mentioned in Section 2.1, we also find that although the coefficient of skewness for men and women have different signs, their absolute values decreased over this period.

Figure 5c reports the kernel estimates of the densities that would have prevailed if the minimum wage had remained at its 1979 level. Corresponding summary measures of wage dispersion are reported in column 3 of Table 3. The results obtained are very similar to those of DiNardo, Fortin, and Lemieux (1995) for the period 1979-1988 and will be discussed in more details in Section 4.4. In

the remainder of this section, we focus on the effect of changes in the relative position locus. We consider both the partial equilibrium case and the case of a pure assignment model. In both cases, we consider the distribution that would have prevailed in 1991 if both the minimum wage and the discrimination function had remained at their 1979 levels.

The kernel density estimates for the partial equilibrium case are reported in Figure 5d. The corresponding summary statistics of the distributions are reported in column 4 of Table 3. The shapes of the estimated densities for men and women are closer to the 1979 shapes after adjusting for changes in the relative position locus. More precisely, reverting the relative position locus back to its 1979 level makes the male density more skewed to the left and the female density more skewed to the right. Consistent with this visual change in shapes, the coefficient of skewness moves from 0.429 to 0.599 for women and from 0.079 to -0.076 for men (Table 3). Women become more concentrated in the lower part of the distribution and, as expected, measures of wage dispersion such as the standard deviation and the 10-90 differential decrease. By contrast, the same measures of wage dispersion indicate a small widening of the distribution of male wages.

Note that the spike at the minimum wage is larger in the counterfactual densities of Figures 5c and 5d than in the original 1979 density (Figure 5a). Because the lower end of the distribution expanded between 1979 and 1991, more workers are "at risk" of being affected by the minimum wage in 1991 than in 1979. This spike masks some of the similarities between the 1979 density and the counterfactual densities in Figures 5d and 5e.

Qualitatively similar results are obtained in the case of a pure assignment model where changes in the relative position locus have no effect on the overall distribution of wages. The kernel density estimates for this case are reported in



Figure 5e and summary measures of wage dispersion are reported in column 5 of Table 3. Reverting the relative position locus back to its 1979 level has a larger effect on inequality among women than in the partial equilibrium case. It also decreases slightly the standard deviation for men but has no effects on the 10-90 differential.

Although changes in the relative position locus may play an important role in the rise in inequality among women, it plays a relatively small role for men. A simple explanation for this difference can be developed by looking at the two tails of the densities in more details. In all the densities plotted in Figure 5, the lower tail is dominated by women but the fraction of men is substantial. By contrast, the upper tail is strongly dominated by men and the fraction of women is very small. In a pure assignment model, women at the top have some room to move up in a fixed overall distribution but men at the bottom are already at the bottom and cannot move down since they are already at the minimum wage. The top and the bottom of the male distribution are thus "pinned down" at the two extreme of the overall distribution and remain unaffected by changes in the relative position locus. By contrast, only the bottom of the women's distribution is "pinned down" at the lower end of the overall distribution. When the relative position of women improves, the spread between the two ends of the distribution is likely to expand for women but to remain constant for men.

Note, however, that the middle (median) of the distribution is always free to move for both men and women. The median man can always be reshuffled into the job of a woman above or below him and vice versa. The median male wage is thus likely to fall when the relative position of women improves while the median female wage is likely to increase. Although the effect of changes in the relative position of women has no effect on the 10-90 wage differential for men (comparing in columns 3 and 5 of Table 3, column 3 of Table 4), this masks offsetting changes

in the 10-50 and the 50-90 differentials. The improvement in the relative position of women between 1979 and 1991 has in fact increased the 50-90 differential for men but decreased the 10-50 differential. This can be explained by a fall in the median with the 10th and 90th percentiles remaining "pinned down" at the two extremes of the overall distribution. By contrast, the improvement in the relative position of women had a positive effect on their own 10-50 and 50-90 differentials. While the 10th percentile remained pinned down at the bottom, the median and especially the 90th percentile were free to move up.

#### 4.4. Reconciling the Pattern of Changes in the Distribution of Wages for Men and Women

In a pure assignment model, jobs are reshuffled between men and women for a given overall distribution of wages. As argued in Section 2.1, this model is a simple and effective way of thinking about the interaction between the wage distributions of men and women. The results of Section 4.3 indicate, however, one critical asymmetry between the wage distributions of men and women. While women at the top of their distribution can move up in the overall distribution, men at the bottom of their distribution are already at the minimum wage and cannot move down. Because of this asymmetry, we argue that our framework can be used to resolve several puzzling patterns about relative changes in the distribution of wages for men and for women. These puzzles are significant since they are inconsistent with a simple story based on a relatively uniform increase in the returns to skill (Juhn, Murphy, and Pierce 1993).

To illustrate these patterns, we decompose the 1979-1991 changes in wage dispersion into a minimum wage effect, a relative position locus effect, and a residual. These decompositions are reported in Table 4. We also decompose

changes in the between- and within-group standard deviation of log wages. We divide both men and women into 16 skill groups based on four categories of potential experience (0-9, 10-19, 20-29, 30+) and four categories of education (0-11, 12, 13-15, 16+). These between/within decompositions are reported in Table 4. We also regroup the 16 skill categories into three coarser skill groups to illustrate differential changes in the within-group standard deviation for less-skilled, medium-skilled, and highly-skilled men and women. These results are reported in Figure 6.

Table 4 and Figure 6 illustrate three key differences between changes in the distribution of wage for men and women.

1. Measures of wage dispersion increased about 50 percent faster for women than for men.
2. The 50-90 differential increased faster than the 10-50 differential for men, while the opposite happened for women (this is an alternative way of describing the relative changes in shapes documented in Section 2.1).
3. The within-group standard deviation increased substantially for medium- and highly-skilled men but did not increase much for less-skilled men. The within-group standard deviation increased even more among women with no clear pattern across skill groups.

The first difference can be fully accounted by changes in the minimum wage and in the relative position locus. For example, the standard deviation increased by 0.028 more for women than for men (Table 4, column 1). The minimum wage explains 0.006 of this differential increase (column 2) while the relative position locus explains 0.032 (column 4). We actually over-explain the difference by 0.01, which corresponds to the relative effect of de-unionization for men found by DiNardo, Fortin, and Lemieux (1995). In other words, the differential growth in inequality for men and women could be fully accounted for by changes in minimum wages, unionization, and especially in the relative position locus. Once

again, the latter effect can be traced back to the fact that women in the upper end of their distribution were able to move up in the overall distribution.

The same explanation applies to peculiar patterns of changes in the 10-50 and 50-90 differentials. Between 1979 and 1991, the 10-50 (50-90) differential increased by 0.067 (0.126) for men and by 0.196 (0.113) for women. Once we account for changes in the minimum wage and the relative position locus, changes in these differentials go in the same direction for men and women. The remaining changes in the 10-50 differential are 0.048 for men and 0.012 for women. The corresponding changes in the 50-90 differential are 0.082 and 0.064 respectively. For reasons discussed above, changes in the relative position locus had a negative effect on the 10-50 differential for men, but positive effect on the 10-50 for women and on the 50-90 for both men and women.

It is interesting to note that residual changes in the differentials that pertain to the lower part of the distribution are smaller than differentials that pertain to the upper part of the distribution. For example, the 10-50 differential for women exhibits the smallest change while the 50-90 differential for men exhibits the largest change. This pattern of change is consistent with skill prices rising faster at the high end than at the low end of the skill distribution.

This non-linear change in skill prices could also explain the change in the within-group standard deviations once the effect of changes in the minimum wage and in the relative position locus have been accounted for. Figure 6 shows that the erratic pattern of changes is mostly due to the differential effect of changes in the relative position locus on the different skill groups.

This effect of changes in relative position has a negative effect on the standard deviation of less-skilled men since these changes push their wages downwards and

cause more compression at the minimum wage. This compression does not affect high skilled men since the lower end of their distribution is above the minimum wage. The lower end of this group can thus move down while the high end remains at the top of the overall distribution. Changes in the relative position locus thus has a positive effect on the within-group standard deviation of skilled men.

The distribution of wages of less-skilled women is even more compressed at the minimum wage than the distribution of low skilled men. However, changes in the relative position locus are favorable to women and push their distribution upwards. This results in an increase, as opposed to a decrease for men, in their within-group standard deviation. The puzzling contrast between changes in the within-group standard deviation for less-skilled men and women is thus explained by the fact that the relative position locus is pushing their respective distributions in opposite direction. These effects would be missed by considering the distribution of male and female wages in isolation.

In summary, the main features of changes in wage inequality for men and women between 1979 and 1991 are well characterized by three factors. The first factor is the decline in the real value of the minimum wage. The second factor is the improvement in the relative position of women for a given overall distribution of wages. The last factor is a non-linear increase in skill prices. One possible reason why skill prices increased faster at the high end than at the low end of the skill distribution is that the minimum wage mitigates these effects at the low end. We conjecture that this non-linear change in skill prices is due to the interaction between the minimum wage and a relatively uniform change in skill prices.

## 5. CONCLUSION

In this paper, we develop an econometric framework to analyze the interaction between changes in the relative position of women in the wage distribution and changes in wage inequality among men and among women. We model the interaction between the wage distributions of men and women using a pure assignment model in which changes in the relative position of women have no effect on the overall distribution of wages. Jobs are simply reshuffled between men and women.

The pure assignment model explains why inequality grew faster among women than men between 1979 and 1991. Changes in the relative position of women are found to explain why, for men, the 50-90 differential grew faster than the 10-50 differential, while just the opposite happened for women. Our framework also helps reconcile the erratic pattern in the within-group standard deviation across skill groups for men and women. Unlike women, highly-skilled men were more affected by changes in relative positions than less-skilled men. Overall, changes in the wage distribution for both men and women are well characterized three factors: a change in relative position of women for a fixed overall distribution of wages, changes in the minimum wage, and a non-linear change in skill prices.

## REFERENCES

- BECKER, G. S. (1957): *The Economics of Discrimination*. University of Chicago Press, Chicago.
- (1985): "Human Capital, Effort and the Sexual Division of Labor," *Journal of Labor Economic*, 3, S33-S58.
- BERGMAN, B. R. (1974): "Occupational Segregation, Wages and Profits when Employers Discriminate by Race or Sex," *Eastern Economic Journal*, 82, 103-110.
- BICKEL, P., AND K. DOKSUM (1981): "An analysis of transformations revisited," *Journal of the American Statistical Association*, 76, 296-311.
- BLAU, F. D., AND L. M. KAHN (1994a): "The Impact of Wage Structure on Trends in U.S. Gender Wage Differentials: 1975-87," NBER Working Paper 4748, National Bureau of Economic Research, Cambridge, MA.
- (1994b): "Rising Wage Inequality and the U.S. Gender Gap," *American Economic Review - AEA Papers and Proceedings*, 84, 23-28.
- BOUND, J., AND G. JOHNSON (1992): "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations," *American Economic Review*, 82, 371-392.
- BOX, G., AND D. COX (1964): "An analysis of transformations," *Journal of the Royal Statistical Society B*, 26, 211-252.
- CARD, D., AND T. LEMIEUX (1995): "Wage Dispersion, Returns to Skill, and Black-White Differentials," *Journal of Econometrics*, forthcoming.
- COX, D. (1972): "Regression models and life tables," *Journal of the Royal Statistical Society B*, 34, 187-202.
- DINARDO, J., N. FONTIN, AND T. LEMIEUX (1995): "Labor Market Institutions and the Distribution of Wages: A Semiparametric Approach," NBER Working Paper 5093, National Bureau of Economic Research, Cambridge, MA.
- DONALD, S. G., D. A. GREEN, AND H. J. PAARSCH (1995): "Differences in Earnings and Wage Distributions between Canada and the United States: An Application of Generalized Lorenz Curve Estimation with Covariates," Discussion paper, Boston University, paper presented at the 5th Canadian Labor Conference.
- FAN, J., AND J. MARRON (1994): "Fast implementation of nonparametric curve estimators," *Journal of Computational and Graphical Statistics*, 3(1), 35-56.
- HALL, P., AND M. WAND (1993): "On the Accuracy of Binned Kernel Density Estimators," Discussion paper.
- HAN, A. K. (1987): "A Non-parametric analysis of transformations," *Journal of Econometrics*, 35, 191-209.
- HECKMAN, J., AND B. SINGER (1984): "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica*, 52, 271-319.
- HOROWITZ, J. L. (1993): "Semiparametric Estimation of a Regression Model with an Unknown Transformation of the Dependent Variable," mimeo 315, University of Iowa, Iowa City, IA. forthcoming in *Econometrica*.

- JUHN, C., K. MURPHY, AND B. PIERCE (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101, 410-442.
- JUHN, C., K. M. MURPHY, AND B. PIERCE (1991): "Accounting for the Slowdown in Black-White Wage Convergence," in *Workers and their wages: changing patterns in the United States*, ed. by M. H. Koster, pp. 107-143. American Enterprise Institute, Washington, D.C.
- LAZEAR, E. P., AND S. ROSEN (1990): "Male-Female Wage Differentials in Job Ladders," *Journal of Labor Economic*, 8, S106-S123.
- MACKINNON, J., AND L. MAGEE (1990): "Transforming the Dependent Variable in Regression Models," *International Economic Review*, 31(2), 315-339.
- MINCER, J. (1992): "Human Capital, Technology, and the Wage Structure: What do the Time Series Show," NBER Working Paper 3581, National Bureau of Economic Research, Cambridge, MA.
- MURPHY, K., AND F. WELCH (1991): "The Role of International Trade in Wage Differentials," in *Workers and Their Wages: Changing Patterns in the United States*, ed. by M. Koster. American Enterprise Institute, Washington, D.C.
- O'NEILL, J., AND S. POLACHEK (1993): "Why the Gender Gap in Wages Narrowed in the 1980s," *Journal of Labor Economic*, 11, 25-33.
- PARK, B. U., AND B. A. TURLACH (1992): "Practical Performance of Several Data Driven Bandwidth Selectors," *Computational Statistics*, 7, 251-270.
- SATTINGER, M. (1993): "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31, 831-880.
- SEATHER, S. (1992): "The Performance of Six Popular Bandwidth Selection Methods on Some Real Data Sets," *Computational Statistics*, 7, 225-250.
- SHEATER, S., AND M. JONES (1991): "A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation," *Journal of the Royal Statistical Society, B*, 53(3), 683-690.
- TEULINGS, C. N. (1995): "The Wage Distribution in a Model of Assignment of Skills to Jobs," *Journal of Political Economy*, 103, 280-315.
- TOPEL, R. H. (1994): "Regional Labor Market and the Determinants of Wage Inequality," *American Economic Review - Papers and Proceedings*, 84(2), 5-21.

TABLE 1  
Sample Means from the Current Population Survey

	Men		Women	
	1979	1991	1979	1991
Log wage	1.863 (0.512)	1.794 <sup>a</sup> (0.581)	1.486 (0.396)	1.552 (0.486)
Education	12.528 (3.122)	13.126 (2.991)	12.560 (2.399)	13.254 (2.393)
Experience	18.950 (13.951)	19.100 (12.146)	18.158 (12.825)	18.967 (11.284)
Non-White	0.111 (0.327)	0.132 (0.353)	0.142 (0.330)	0.157 (0.347)
No. of observations	93544	90099	73940	85192

Note: Figures in parentheses are standard deviations.  
<sup>a</sup> In 1979 dollars.

TABLE 2  
Ordered Probit Estimates of the Returns to Individual Characteristics

	Men		Women	
	(1) 1979	(2) 1991	(3) 1979	(4) 1991
Education	0.1528 (0.0012)	0.2093 (0.0013)	0.1892 (0.0017)	0.2362 (0.0016)
Experience	0.2201 (0.0038)	0.1991 (0.0044)	0.2018 (0.0045)	0.1871 (0.0048)
Experience <sup>a</sup> (+ 100)	-0.9656 (0.0308)	-0.7484 (0.0353)	-1.1823 (0.0387)	-0.8828 (0.0396)
Experiences (+ 1000)	0.1915 (0.0092)	0.1220 (0.0106)	0.2846 (0.0120)	0.1729 (0.0123)
Experiences <sup>a</sup> (+ 10000)	-0.0143 (0.0009)	-0.0069 (0.0011)	-0.0239 (0.0012)	-0.0118 (0.0013)
Non-White	-0.3380 (0.0101)	-0.3591 (0.0095)	-0.0844 (0.0112)	-0.1235 (0.0099)
No. of thresholds	153	153	159	159
-2 Loglikelihood	906808	878188	584922	697845
No. of observations	93544	90099	73940	85192

Note: Estimated standard errors are in parentheses.



TABLE 3  
Measures of Wage Dispersion

Statistic	1979		1991		
	(1) Predicted	(2) Predicted	(3) with 1979 Minimum Wage	(4) (3) with 1979 Discrimination and Nepotism	(5) (4) with All as in (3) <sup>b</sup>
<b>All:</b>					
Mean	1.718	1.693	1.708	1.714	—
Standard Deviation	0.462	0.511	0.489	0.504	—
Skewness	0.174	0.136	0.257	0.265	—
10-90 <sup>a</sup>	1.273	1.437	1.345	1.384	—
10-50	0.599	0.694	0.601	0.607	—
50-90	0.675	0.743	0.743	0.777	—
25-75	0.729	0.785	0.785	0.826	—
5-95	1.566	1.822	1.680	1.709	—
<b>Men:</b>					
Mean	1.857	1.788	1.801	1.872	1.859
Standard Deviation	0.457	0.520	0.500	0.507	0.494
Skewness	-0.129	-0.029	0.079	-0.076	-0.060
10-90	1.278	1.471	1.409	1.440	1.408
10-50	0.687	0.754	0.690	0.758	0.735
50-90	0.591	0.717	0.718	0.682	0.673
25-75	0.677	0.785	0.784	0.786	0.751
5-95	1.599	1.844	1.725	1.739	1.719
<b>Women:</b>					
Mean	1.506	1.571	1.591	1.513	1.516
Standard Deviation	0.382	0.473	0.448	0.421	0.409
Skewness	0.494	0.284	0.429	0.599	0.523
10-90	1.015	1.324	1.222	1.128	1.091
10-50	0.410	0.606	0.492	0.411	0.422
50-90	0.605	0.718	0.730	0.717	0.669
25-75	0.580	0.715	0.714	0.660	0.644
5-95	1.300	1.662	1.541	1.477	1.420

Note: Lower and upper 1% of wage distributions are trimmed. In 1979, the statistics were computed on wages  $\geq 1.71$  and  $\leq 18.73$ . In 1991, the statistics were computed on wages  $\geq 1.64$  and  $\leq 20.50$  in 1979 dollars.

<sup>a</sup> Difference between the 90th and the 10th percentiles of the log wage distribution. The 10-50, 50-90, 25-75 and 5-95 statistics are defined similarly.

<sup>b</sup> Statistics of the distribution of all workers is the same as in (3).

TABLE 4  
Changes in Measures of Wage Dispersion  
Between and Within Skill Groups: 1979-1991

Statistic	Effect of:			Residual Change
	(1) Total Change	(2) Minimum Wage	(3) Discrimination and Nepotism with All unchanged from (2)	
<b>Men:</b>				
Standard Deviation Within Groups	0.059	0.011 (0.19)	0.003 (0.05)	0.045 (0.76)
Between Groups	0.038	0.016 (0.40)	0.006 (0.16)	0.017 (0.44)
Total	0.063	0.019 (0.30)	0.007 (0.11)	0.037 (0.59)
10-90	0.193	0.062 (0.32)	0.001 (0.01)	0.130 (0.67)
10-50	0.067	0.064 (0.96)	-0.045 (-0.67)	0.048 (0.72)
50-90	0.126	-0.001 (-0.01)	0.045 (0.36)	0.082 (0.65)
<b>Women:</b>				
Standard Deviation Within Groups	0.073	0.013 (0.18)	0.019 (0.26)	0.041 (0.56)
Between Groups	0.065	0.022 (0.33)	0.034 (0.51)	0.010 (0.16)
Total	0.091	0.025 (0.28)	0.039 (0.42)	0.027 (0.30)
10-90	0.309	0.102 (0.33)	0.131 (0.42)	0.076 (0.25)
10-50	0.196	0.114 (0.58)	0.070 (0.36)	0.012 (0.06)
50-90	0.113	-0.012 (-0.10)	0.061 (0.54)	0.064 (0.57)

Note: Percent of total variation explained in parenthesis. Lower and upper 1% of wage distributions are trimmed. In 1979, the statistics were computed on wages  $\geq 1.71$  and  $\leq 18.73$ . In 1991, the statistics were computed on wages  $\geq 1.64$  and  $\leq 20.50$  in 1979 dollars.

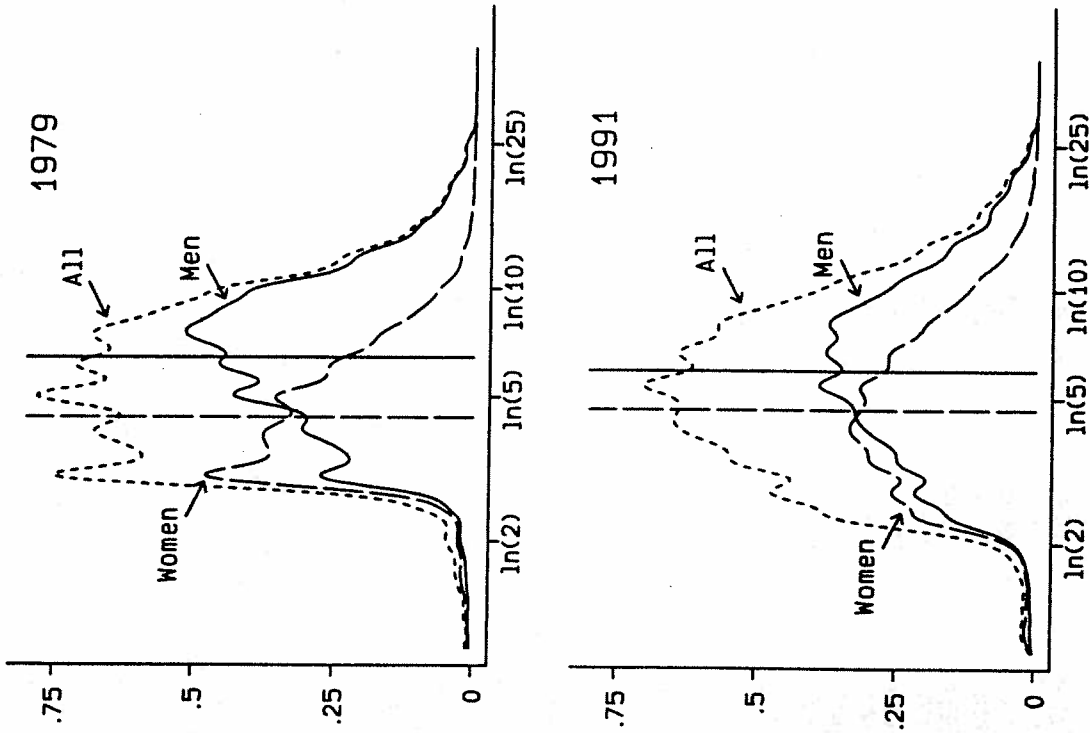


Figure 1.- Kernel Density Estimates of Real Log Wages (\$1979) - Hours Weighted

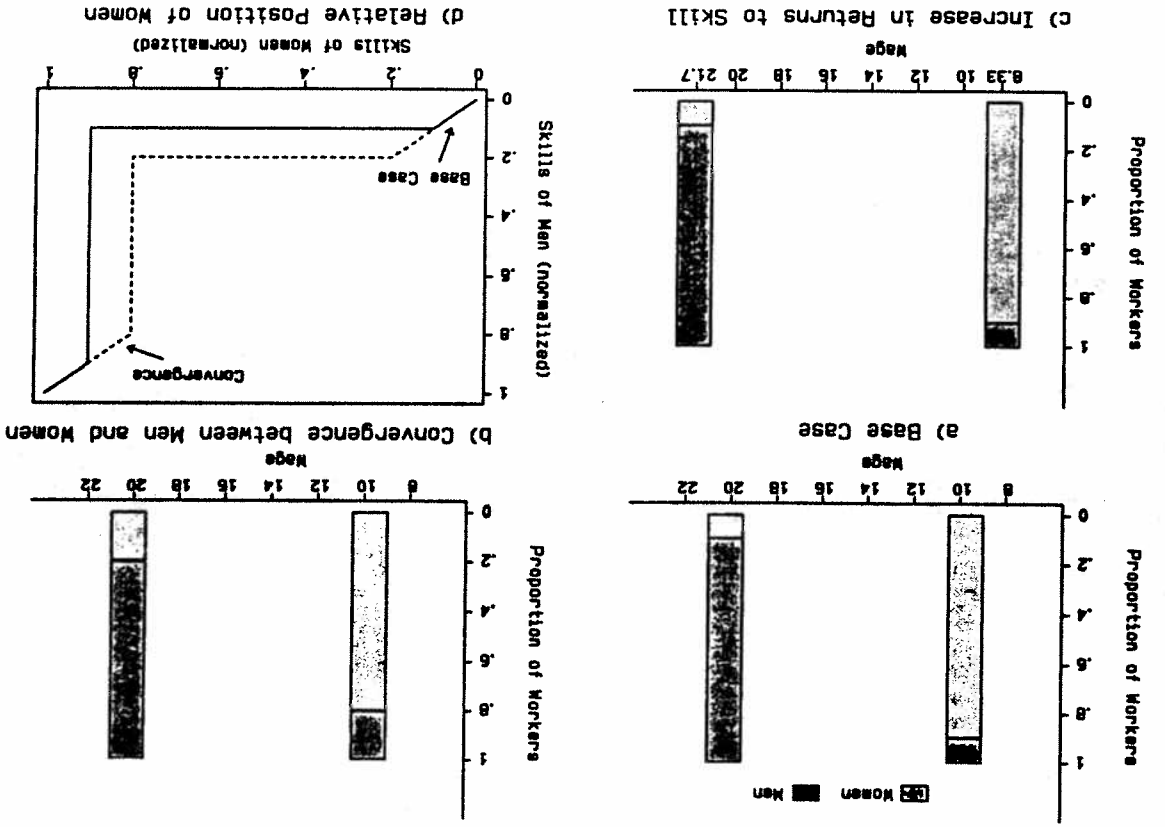


Figure 2.- Illustrative Example

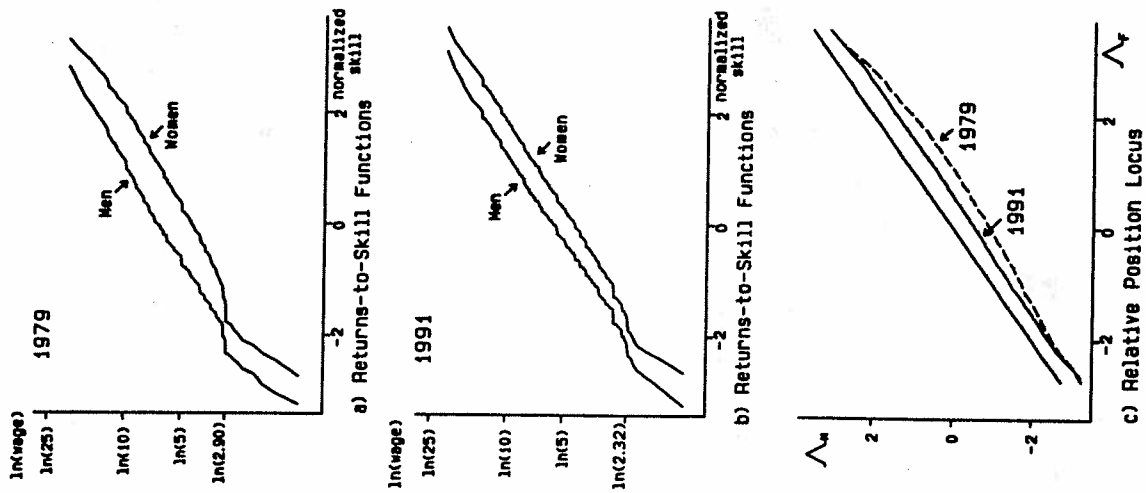


Figure 3.- Estimates of Returns-to-Skill and Relative Position Locus

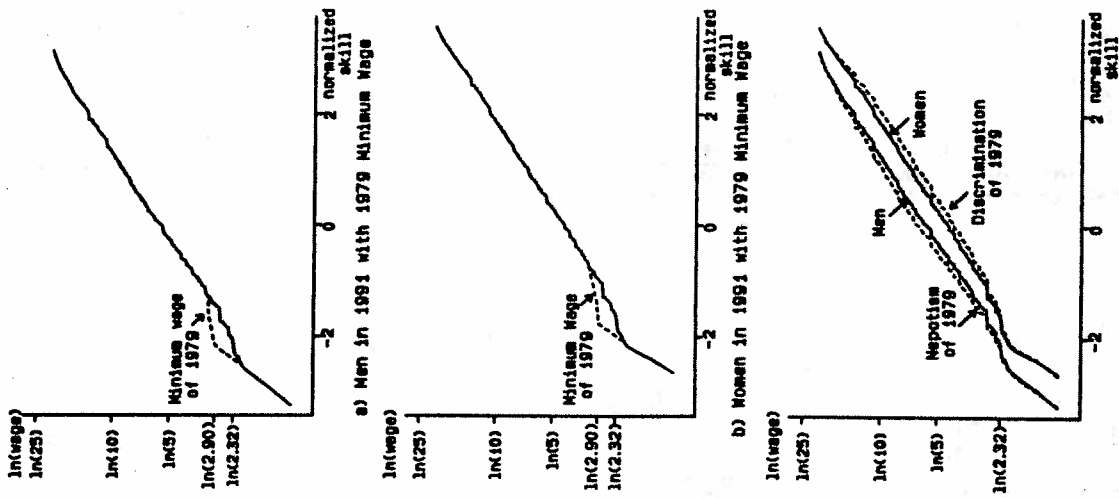


Figure 4.- Counterfactual Returns-to-Skill Functions

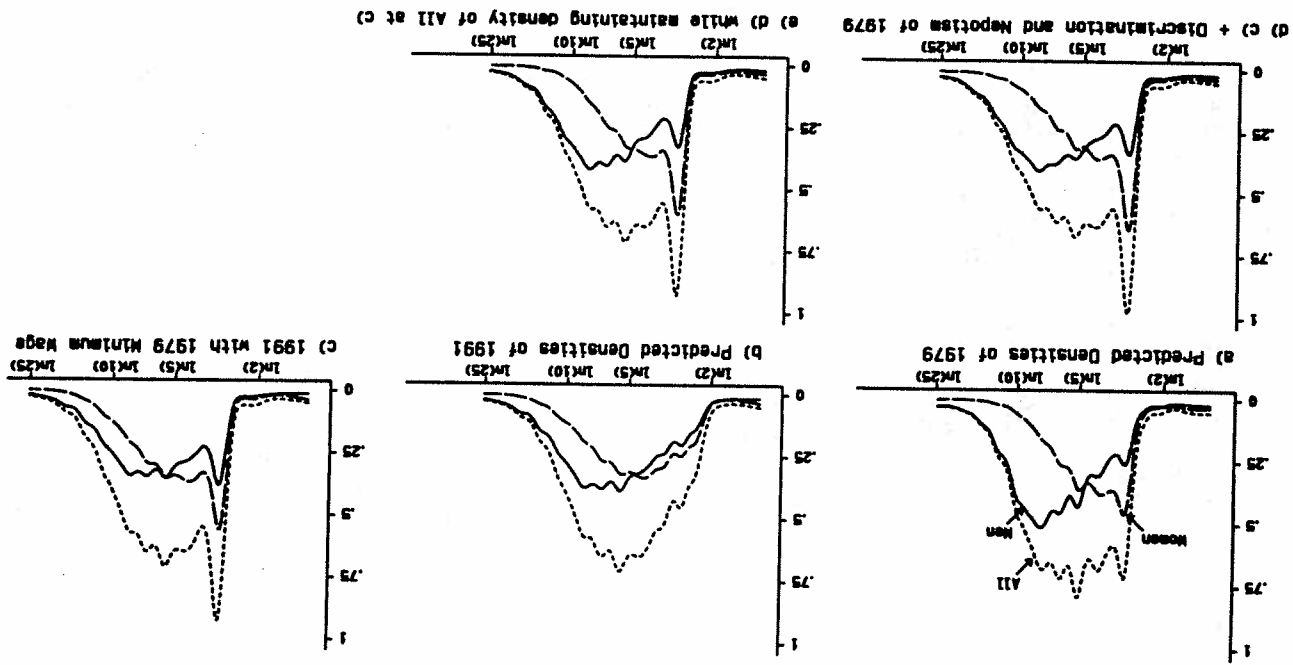


Figure 5.- Counterfactual Densities

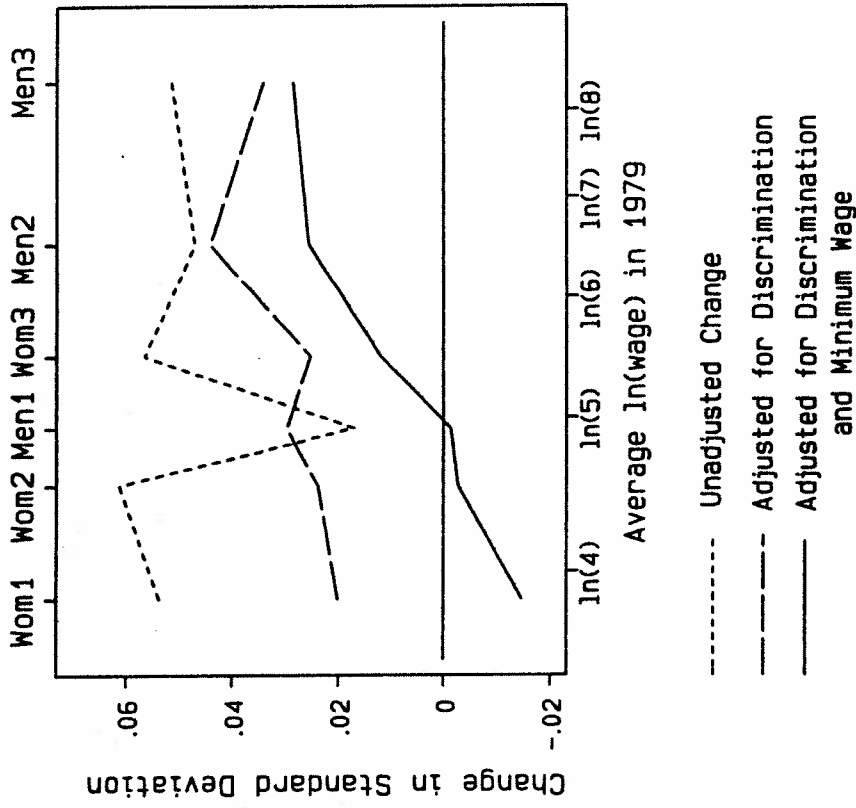


Figure 6.- Changes in Within Standard Deviation