

NBER WORKING PAPER SERIES

THE CLOSING OF THE GENDER GAP AS A ROY MODEL ILLUSION

Casey B. Mulligan
Yona Rubinstein

Working Paper 10892
<http://www.nber.org/papers/w1089>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2004

We appreciate the comments of Gary Becker, Aitor Lacuesta, Amalia Miller, Kevin M. Murphy, John Pepper, Ed Vytlačil, the research assistance of Ellerie Weber, and the financial support of the National Science Foundation (grant #0241148). The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

© 2004 by Casey B. Mulligan and Yona Rubinstein. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Closing Gender Gap as a Roy Model Illusion
Casey B. Mulligan and Yona Rubinstein
NBER Working Paper No. 10892
November 2004
JEL No. J21, J31, J16, C31

ABSTRACT

Rising wage inequality within-gender since 1975 has created the illusion of rising wage equality between genders. In the 1970's, women were relatively equal (to each other) in terms of their earnings potential, so that nonwage factors may have dominated female labor supply decisions and nonworking women actually had more earnings potential than working women. By 1990, wages had become unequal enough that they dominated nonwage factors, so that nonworking women tended to be the ones with less earnings potential, and the wage gap between workers and nonworkers was large. Accounting for the growing selection bias using both parametric and semi-parametric versions of the Roy model, we show how the earning power of the median woman has not caught up to the earning power of a median man, even while the earning power of the median working woman has. As an illustration, we give some attention to wives with advanced degrees — they have high and stable labor force participation rates — and show how their measured wages have grown at about the same rate as those of men with advanced degrees.

Casey B. Mulligan
University of Chicago
Department of Economics
1126 East 59th Street, #506
Chicago, IL 60637
and NBER
c-mulligan@uchicago.edu

Yona Rubinstein
yonar@post.tau.ac.il

Table of Contents

I. Introduction	1
II. Selection and Composition Biases in the Calculation of Relative Wages	5
Labor Supply Constant	7
Labor Supply Varies	10
III. Recalculating Gender Gap Closure with Heckman and Related Selection Models	11
IV. Further Indicators of the Composition of Working Wives, Related to Identification at Infinity	17
The Composition of Advanced Degree Working Wives	17
Determinants of Marriage Rates	18
V. Conclusions	19
VI. Appendix: The Occupational Composition of Wives with Advanced Degrees	21
VII. References	23

I. Introduction

Perhaps one of the most dramatic changes in U.S. labor market outcomes over the past thirty years is the persistent growth, within-gender, in overall earnings inequality (see Levy and Murnane, 1992, and Katz and Autor, 1999, for comprehensive surveys). Inequality in earnings grew over this period not only from an increase in the Mincerian returns to education but also due to an increase in inequality within groups of workers of similar age and education (Katz and Murphy, 1992). As first pointed out by Juhn, Murphy, and Pierce (1993) the inequality growth during the 1970s, the 1980s, and the 1990s, appears to have occurred throughout the earnings distribution as well as over people's life cycle (Gottschalk and Moffitt, 1995). By the end of the second Millennium, US wage inequality was higher than it has ever been since WWII.

The measured earnings of women have substantially, although not fully, caught up with the earnings of men. This closing gender wage gap is often said to indicate the importance of wages for bringing women into the labor force (Mincer, 1962; Goldin, 1990), and the status of discrimination and other nonwage factors in the operation of the labor market (Becker, 1985; Katz and Murphy, 1992). At the same time, it has been observed that this increased gender equality has been coincident with growing earnings inequality within genders, as shown in Figure 1. The solid line is a familiar measure of gender equality (e.g., Murphy and Welch, 1999), namely, the median earnings of women working full-time full-year as a ratio of the median earnings of men working full-time full-year (hereafter *ftfy*). The dashed line is the ratio of the 90th percentile to the 10th percentile in the cross-section distribution of men working full-time full-year. We see that both were flat until 1977 or so (see also O'Neill, 1985, on the apparent constancy of the gender wage gap). Both rose – most rapidly at first – from the late 1970's until about 2000. This has been regarded in the literature as a curious coincidence (Card and DiNardo, 2002, p. 742; Blau and Kahn, 1997, p. 2),¹ and perhaps indicative of earnings' having multiple and largely independent determinants, but can we still conclude that wages help pull women into the labor force, and that the

¹Perhaps also in small part an negative effect of women's labor market entry on unskilled male wages(eg., Topel, 1994; Juhn and Kim, 1999).

labor market has mitigated gender discrimination over time? The purpose of our paper is to suggest that (a) apparent gender equality is a direct consequence of inequality within gender, and (b) the apparent gender equality is not real in the sense that the average woman's earnings potential has not caught up with that of the average man.

Although more work on this topic is needed, our two results may suggest a third, namely that much of earnings inequality growth can be understood with a single attribute model. In such a model, women would have less earnings potential than men for the same reason that some men have less earnings potential than others: differences in the one attribute (call it "skill"). According to our estimates below, the average married woman in 1975 had 36% less earnings potential than the average married man, and would fit in the 17th percentile of the 1975 male wage distribution. That percentile lost 18% relative to the male average over the period 1975-95, so in the single attribute model women should have lost the same percentage. Some of our estimates do suggest that the gender wage gap widened 1975-95, although it is unclear whether it widened that much. Nevertheless, the single attribute model performs much less badly than it appeared when it was thought that inequality moved in opposite directions within and across genders.

We argue that, as suggested by the Roy (1951) model as applied to the choice between market and nonmarket work, working women are selected based on their wages, and the gap between the earnings potential of working and non-working women grows with inequality among women. As a consequence, the wage distribution of working women is different than the wage distribution of women, to a degree which declines with the fraction of women who work. Hence, a less biased estimate of the gender gap may be most easily calculated, and shown to close more slowly if at all, using samples of women for which the propensity to work is high, and stable over time. Figure 2 displays such an estimate as the solid line, namely the gender wage gap calculated as the average log wage for married women with advanced degrees working full-time, net of the average log wage of full-time men with advanced degrees. Work is very common for this group of women – about 80% of them have some earnings during the year and about half work full-time – so we expect earnings inequality to affect less significantly the trend in this group's relative wages. The solid line shows little growth in the relative earnings of women. In contrast, other schooling groups have fewer women working (the 1970 percentage of women working full-time is shown in parentheses in the legend), and thereby a greater potential for selection and composition biases, and show growth in the relative earnings of women much like that shown in Figure 1.² Indeed, while

²Figure 2's calculations of the advanced degree gender gap are based on 400+ annual CPS observations of married women working full-time, full-year.

selection bias growth may be less important for the advanced degree women, it still may have been positive; the solid line actually suggests that the earnings potential of women may have fallen over time relative to men's.

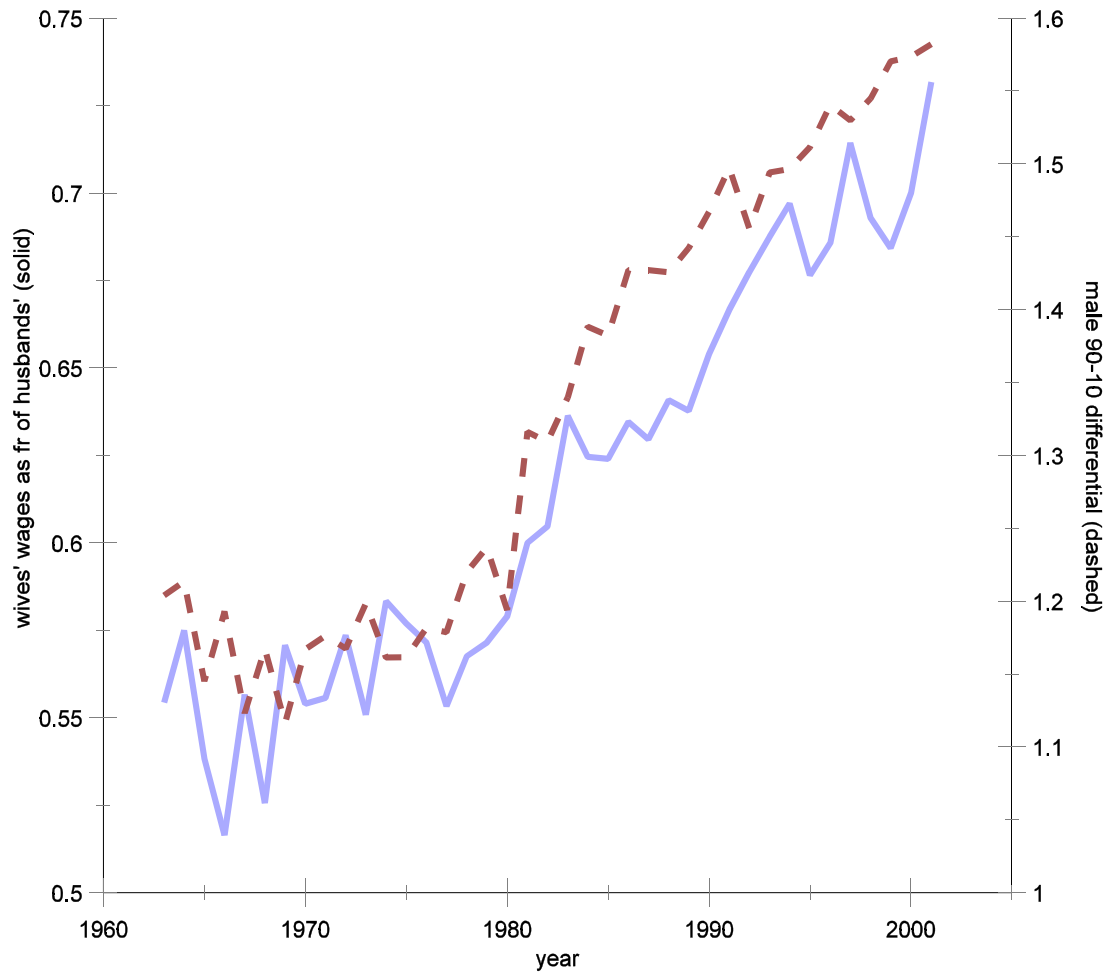


Figure 1 Wage Inequality between and within Genders

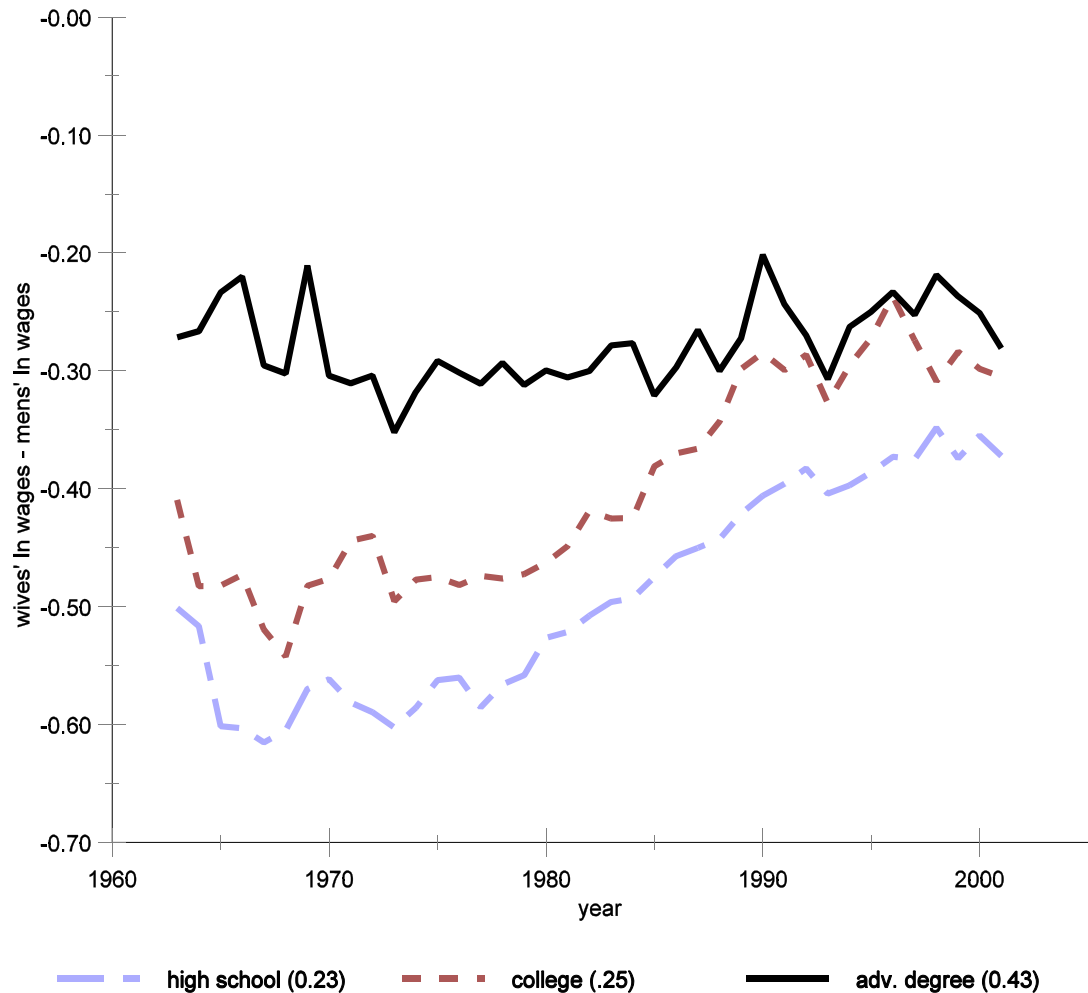


Figure 2 The Gender Gap by Schooling

Section II uses the Roy model to show how growing inequality within gender can give the illusion of a closing gender gap, due to the selection and composition biases involved with measuring the earnings potential of women. The Roy model predicts that a group's measured wage growth is an upward biased measure of its earnings potential growth, with the magnitude of the bias declining with the fraction of group members actually working. Section II also shows how a number of demographic groups, including married women, fit this pattern. Section III uses structural selection models to obtain numerical estimates of the amount of selection bias growth, and thereby the relative growth of male and female potential earnings. Section IV explores alternative interpretations of our Figure 2. Section V concludes.

Our data come from a series of 39 consecutive March Current Population Surveys and their Demographic Supplements (hereafter: March CPS) for survey years 1964 to 2002. The population sample (universe) consists of civilian non-institutionalized population of the US living in housing units and members of the Armed Forces living in civilian housing units on a military base or in housing units not on a military base. Each record contains information about an individual, the household in which the individual resides, and the family and the spouse of the individual. In addition to the standard monthly labor force data, these files contain supplemental data on work experience. This collection provides information on employment and wages in the preceding calendar year while demographic data refer to the time of the survey. Thus, the annual work experience data – from the CPS demographic supplement – cover the period of 1963 to 2001. We construct two data sets. The first file includes all individuals aged 24 to 54 (hereafter: individual file). The second file includes only husbands and wives. We restrict the second file to include only couples in which we observe both partners (1,248,117 couples in 1964 through 2002).

CPS observations are divided by school completion into five sub-groups: (i) high school dropouts – less than twelve grades, (ii) high school graduates (including those graduated by taking the GED exam), (iii) some college completed, (iv) college graduates with 16 (and 17) years of schooling (BA) and (v) college graduates with advanced/professional degree (MBA, Ph.D.) or, prior to 1993, persons with 18 or more years of completed schooling. We measure wages according to total annual earnings deflated by the US CPI, giving most of our attention to full-time workers (namely full-time workers who report working at least 50 weeks of the previous year).

II. Selection and Composition Biases in the Calculation of Relative Wages

Wages are often interpreted as measures of important economic concepts, such as human capital or discrimination. However, one nuisance in the measurement process is the difficulty of measuring wages for people who are not working. Labor economics suggests that the people who work are different from the people who do not work, and statistical theory shows that the wage difference between those working and those not working increases with the amount of wage inequality. The purpose of this paper is to show that, in the context of the rising inequality 1975 until the present, this effect is important and explains a lot of the cross-group pattern of wage gains over that period. Perhaps most important is the real possibility that the measured wage gains of married women relative to married men are consistent with the view that the wage distribution for wives was actually unchanged relative to the wage

distribution for husbands.

The Roy (1951) model, as applied to the choice between market and nonmarket work (see also Gronau, 1974, and Heckman, 1974), illustrates this. Each person is described by two variables: his or her (potential) market wage, and his or her reservation wage (a.k.a., his or her productivity in the nonmarket sector). A person works in the marketplace if and only if the market wage exceeds the reservation wage. Each person's market and reservation wages are drawn from a joint lognormal distribution, whose parameters may vary over time and across groups,

$$\begin{pmatrix} w \\ r \end{pmatrix} \sim N \left(\begin{pmatrix} \hat{w} \\ \hat{r} \end{pmatrix}, \begin{pmatrix} \sigma_w^2 & \rho_{wr} \sigma_w \sigma_r \\ \rho_{wr} \sigma_w \sigma_r & \sigma_r^2 \end{pmatrix} \right)$$

where w and r denote log market and reservation wages, respectively, and hats denote medians. The workers L are distinguished from the nonworkers by the condition $z \equiv w - r > 0$, where z is the net gain from working. Since wages are unmeasured for nonworkers, the average measured wage is $E(w|z>0)$:

$$E(w|z > 0) = \hat{w} + \lambda(L) \rho_{wz} \sigma_w$$

$$L = \Phi \left(\frac{\hat{w} - \hat{r}}{\sigma_z} \right) \quad (1)$$

$$\rho_{wz} \equiv \frac{\sigma_w - \rho_{wr} \sigma_r}{\sigma_z}, \quad \sigma_z^2 \equiv \sigma_w^2 + \sigma_r^2 - 2 \rho_{wr} \sigma_w \sigma_r, \quad \lambda(L) \equiv \frac{\varphi(\Phi^{-1}(L))}{L}$$

where σ_z is, roughly speaking, the inverse of the group labor supply elasticity. λ is the inverse Mill's ratio, and slopes down as a function of L .

ρ_{wz} is the correlation between log wages and the (log) net gain from working, which can either be positive or negative, according to whether workers have higher or lower wages than nonworkers, respectively. Just as important, growth in σ_w should *increase* ρ_{wz} and *could even change its sign*. Remember that σ_w was much lower in the 1970's, at a time when ρ_{wz} was found to be negative for married women (e.g., Heckman, 1974). $\rho_{wz} < 0$ is equivalent to $\sigma_w < \rho_{wr} \sigma_r$, which should be less likely to hold as σ_w gets larger. Indeed, we find that ρ_{wz} changes sign for married women in the early 1980's. Intuitively, nonwage factors dominated female labor supply decisions in the 1970's when σ_w was relatively small. By 1990, wages had become unequal enough that they dominated nonwage factors, so

that nonworking women tended to be the ones with less earnings potential.

Equation (1) decomposes the average measured log wage into four components, two of which have been emphasized in the gender wage gap literature. The first and obvious one is the median wage. For example declining gender discrimination is sometimes said to uniformly increase the potential market wage of all women, perhaps as modeled by shifting the median wage. Second is a form of composition bias emphasized by O’Neil (1985), Blau and Kahn (1997), and others: when $\rho_{wz} > 0$, labor supply shifts move relatively low wage people into (or, if the shift is in the direction of less labor supply, out of) the labor market.³ The magnitude of this composition bias depends on the Mill’s ratio, which is higher when a smaller fraction of the group is in the labor market. Third is another form of composition bias. In general, at least with $\rho_{wz} > 0$, workers are some combination of high market wage and low reservation wage. ρ_{wz} indicates the relative importance of these. Fourth, to the extent that workers are selected on wages, workers have higher wages. Gronau (1974, pp. 1127-8) and others recognize that the magnitude of the selection bias decreases with the amount of labor supply L , and increases with the amount of wage inequality σ_w . However, Gronau’s result has been ignored when considering wage trends since 1975, namely when σ_w was growing.

II.A Labor Supply Constant

Henceforth, we refer to equation (1)’s bias term without subscripts – namely as $\lambda\rho\sigma$ – except when needed for clarity. The change over time in a group’s average measured log wage has four components corresponding to the four biases mentioned above.

$$\Delta E(w|z > 0) = \Delta \hat{w} + \lambda(t)\rho(t)\Delta\sigma + \lambda(t-1)\Delta\rho\sigma(t-1) + \Delta\lambda\rho(t)\sigma(t-1) \quad (2)$$

where t denotes time and Δ denotes a change from time $t-1$ to time t . For the time period 1975-2000, $\Delta\sigma$ and $\Delta\rho$ are presumably positive, since within-group wage inequality grew during this period, and ρ increases with σ (see equation (1)). The sign of $\Delta\lambda$ depends on the particular group, namely whether the group increased or decreased its labor supply. Hence we begin by considering two groups whose labor supply was little changed during the period, namely high school educated men aged 40-49 and high

³As shown in the second formula (1), an increased labor supply might come from higher median market wages, lower median reservation wages, or a change in the labor supply elasticity. The labor supply elasticity is determined by the amount of inequality in the net gain z from working.

school educated men aged 66-70.⁴ For them, and any other group whose labor supply constant is constant because changes in the median reservation wage offset changes in the market wage distribution, $\lambda(t) = \lambda(t-1)$ and equation (2) becomes (2)'.

$$\Delta E(w|z > 0) = \Delta \hat{w} + \lambda(t-1)[\rho(t)\Delta\sigma + \Delta\rho\sigma(t-1)] \quad (2)'$$

In the case that workers earn more than nonworkers would ($\rho > 0$), the square bracket term is positive, and growing inequality causes measured average log wages to grow more than do median log wages. Furthermore, the magnitude of the bias is proportional to $\lambda(t-1)$, which declines with group labor supply $L(t-1)$, and should be close to zero for groups like men aged 40-49 for whom L is practically one.

We expect gender and schooling to be the more important determinants of the $\Delta\hat{w}$ term, with age less important because most work on the age structure has shown fairly little change over time in the returns to experience. In order to bring our attention to the bias term, we compare groups of the same gender and schooling, and for the moment treat the $\Delta\hat{w}$ as a constant. In terms of the bias term, the second term in square brackets could be as large the first. To see this, notice that inequality as measured by σ changes by about 0.13 1975-2000, and that $\Delta\sigma$ is multiplied by ρ which may be a lot less than 1, so that the product is less than 0.13. The initial level of σ is about 0.6, and is multiplied by $\Delta\rho$, which from the formula (1) could be as large or larger than 0.1. Hence 0.1 or 0.15 is a reasonable guess for the square bracket term. The group-difference in inverse Mills ratios is 1.19, so we expect, to an order of magnitude, the group difference in $\Delta E(w|z > 0)$ to be roughly 0.15.

Figure 3 shows that, empirically, the group difference is 0.15. The horizontal axis shows $L(t-1)$. The vertical axis shows $\Delta E(w|z > 0)$ for the two groups relative to $\Delta E(w|z > 0) = -0.255$ for all prime-aged high school educated men. The $L=1$ intercept in Figure 3 is important because $\lambda(L=1) = 0$ and equation (2)'s bias term disappears. Obviously, the selection and composition biases are zero for a group with 100% labor supply. Since the $L=1$ intercept is about zero, it appears that most of the measured wage gains for elderly high school educated men (relative to high school educated men overall) may just be an illusion, in the sense that they would be observed even if \hat{w} were constant. Perhaps this conclusion is not particularly surprising or interesting, but the technique of using the $L=1$ intercept leads to some novel conclusions for women.

⁴Over the period shown in the Figure below (1974-78 to 1994-98), the fraction of high school educated men aged 40-49 working full time fell from 0.77 to 0.73 (authors' calculation from the CPS). The fraction for high school educated men aged 66-70 fell from 0.14 to 0.11.

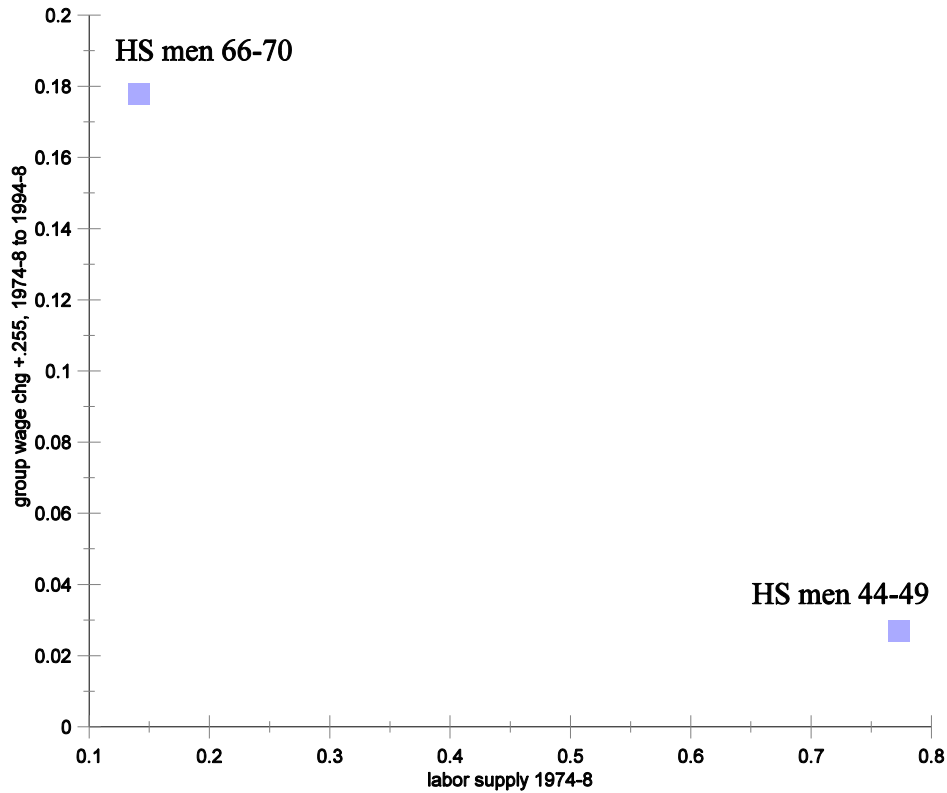


Figure 3 Measured Wage Growth Declines with Labor Supply (Men)

Most groups of married women significantly increased their labor supply in the 1970's, 1980's, and 1990's, so equation (2)' does not apply. However, this is much less true for married women with post-college education: 43% of them worked full-time full-year in the early 1970's compared with 55% in the late 1990's (the same percentage went from 27 to 47 for college grads and from 24 to 45 for high school grads). Where should they appear in Figure 3? ρ is probably smaller, and even negative early in the period, for women than for men because husbands and wives sort positively, and the husband's wage has a negative wealth effect on wife's labor supply. In the absence of gender-specific determinants of the $\Delta\hat{w}$ term, men and women's groups with constant labor supply should have the same $L=1$ intercept, but the slope of the wage growth-labor supply relation be smaller in absolute value for women. In fact, measuring wage growth relative to men with the same schooling, we find that post-college wives would be located at (0.43,0.06) in Figure 3. As predicted by the Roy model, their (schooling adjusted) wage growth is low like that for husbands aged 40-49.

II.B Labor Supply Varies

Labor supply trends for some groups, especially married women. For the purposes of using the formula (2), the various sources of the labor supply trend (which may include σ , see the formula (1) for L) all matter in the same way, namely as they contribute to changes in the inverse Mills term λ . Since married women's labor supply is increasing, their inverse Mills term λ is decreasing and equation (2)'s $\Delta\lambda$ term has the opposite sign as its $\Delta\sigma$ term. Equation (2)'' rewrites equation (2) by combining these terms:

$$\begin{aligned}\Delta E(w|z > 0) &= \Delta \hat{w} + \rho(t)[\lambda(t)\sigma(t) - \lambda(t-1)\sigma(t-1)] + \lambda(t-1)\Delta\rho\sigma(t-1) \\ &\approx \Delta \hat{w} + \lambda(t-1)\Delta\rho\sigma(t-1)\end{aligned}\tag{2}''$$

We measure σ as the standard deviation of log annual earnings among men, and λ by applying the inverse Mills formula to observed married female labor supply. The square bracket term turns out to be -0.14, and is multiplied by $\rho(t)$. Since the last term is $\Delta\rho$ times 0.71, it probably dominates the middle term for married women, which is our reason for using the approximation shown in the bottom line of equation (2)''. In words, measured wage growth for women is biased upward for men and, like the bias for men aged 40-49 or 66-70, the magnitude of the bias depends on the initial level of labor supply through the inverse Mills term $\lambda(t-1)$. Whether the bias is larger for male or female groups with similar $\lambda(t-1)$ depends on the relation between $\Delta\rho$ and gender. As we explain below, we expect $\Delta\rho$ to be positive and much larger for women.

Figure 4 is a female version of Figure 3. The horizontal axis measures group labor supply 1974-8. The vertical axis measures group wage growth relative to men with the same schooling. Again we see that the high initial labor supply groups (single women and/or college+ women) have lower wage growth. The relation is steeper within marital status than across marital status, which we expect if growing inequality has an additional wealth effect on wives (through the wages of their husbands). In any case, the high labor supply groups in Figure 4 tell us that the gender wage gap may not have closed at all, or closed at most 0.15, which is much less closing than suggested by the dashed and dash-dot line in Figure 2.

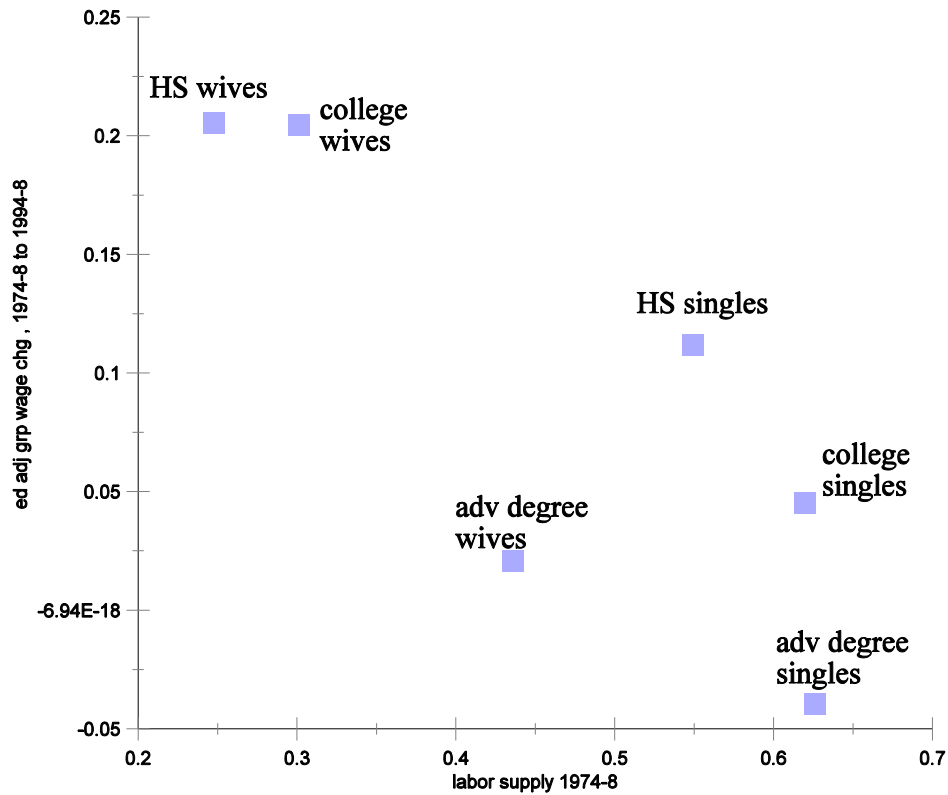


Figure 4 Measured Wage Growth Declines with Labor Supply (White Women)

III. Recalculating Gender Gap Closure with Heckman and Related Selection Models

If we modify equation (1) by allowing median reservation and market wages to be log-linear functions of demographic characteristics X , it becomes the Heckman (1979) selection model. Remember that the Heckman selection model can be interpreted as a least squares regression of log wages on X plus the inverse Mills ratio λ predicted for the worker based on her demographics; conversely that least squares regressions of log wages on X suffer from the bias resulting from the omission of the inverse Mill's ratio λ . Hence, if the relation between demographics and median wages were constant (our estimates below suggest that it is), then an increase over time in the $\lambda\sigma\rho$ term causes the constant term in the Heckman selection model to increase less (or decrease) more than the constant term in the least squares model. We explained above how the change over time in $\lambda\sigma\rho$ is qualitatively ambiguous because

λ falls and σ rises, but the Heckman selection model permits numerical estimates of $\lambda\sigma\rho$. We display some estimates in Table 1. The left part of the Table uses married women from the 1970's, and the right part uses married women from the late 1990's. On each side, a least squares and Heckman selection estimates are shown; of course the Heckman specifications have (or can be interpreted as having – see Heckman 1979) λ as an additional regressor.

Table 1: Women's wages over time, with and without selection corrections

independent variable	1975-79		1995-99		selection bias growth
	OLS	Heckit	OLS	Heckit	
(experience-15)	0.005 (0.002)	0.005 (0.002)	0.016 (0.002)	0.018 (0.002)	
(experience-15) ² /100	-0.011 (0.010)	-0.010 (0.010)	-0.065 (0.009)	-0.074 (0.010)	
high school dropout	9.787 (0.013)	9.888 (0.034)	9.686 (0.019)	9.460 (0.033)	0.327
high school graduate	9.999 (0.013)	10.085 (0.027)	9.960 (0.008)	9.788 (0.022)	0.258
some college	10.084 (0.011)	10.166 (0.027)	10.150 (0.008)	9.983 (0.021)	0.249
college graduate	10.282 (0.013)	10.353 (0.025)	10.492 (0.008)	10.334 (0.020)	0.228
advanced degree	10.500 (0.021)	10.556 (0.027)	10.774 (0.012)	10.637 (0.020)	0.193
teacher	0.033 (0.017)	0.035 (0.017)	-0.233 (0.013)	-0.231 (0.013)	0.000
observations	20,971	20,971	28,931	28,931	
$\sigma\rho$	0	-0.066 (0.020)	0	0.183 (0.022)	
adj-R ²	.07	.07	.18	.18	

Notes: (1) dependent variable is log weekly wage. sample is wives aged 25-54 from white households
(2) there is no constant term, but the schooling dummies sum to a constant
(3) selection bias growth is the growth over time of the OLS minus Heckit coefficient on the schooling dummy
(4) standard errors in parentheses
(5) experience measured as age - years of schooling - 6
(5) Heckit model estimated in two stages, with the first stage including wife's education and experience, husband's education and experience, and the number of children aged 0-6 in the family

The regressions shown in the Table have no constant term *per se*, although the schooling dummies sum to one. Hence the education coefficients estimate the mean (with the normal distribution, also the median) log wage for a nonteacher with 15 years of experience (experience measured as age minus schooling minus six). According to the least squares estimates, “some college” women’s median log wages increased by 0.066 log points. Since men’s wages were higher and declining over this period, this might be interpreted as a closing of the gender gap. However, the Heckman selection estimates say that mean log wages actually fell 0.183 log points; there was little or no gender gap closure. The reason for the different Heckman estimates is that the inverse Mill’s ratio coefficient was negative during the 1970’s and positive during the 1990’s. In words, the bias from not measuring the earning power of nonworking women has changed over time (for “some college,” by 0.249 log points), in large part because wage inequality has grown within gender.

Figure 5 displays time series for wife’s log wage selection bias. More specifically, Figure 5 is a graphical version of Table 1, with nine time periods rather than two: in each time period the Heckit constant term (for women with some college) is subtracted from the corresponding OLS constant term. During the 1970’s, the selection bias was negative (i.e., the selection correction was positive); women out of the labor force had more earnings potential than women in the labor force. Beginning in the early 1980’s, the selection bias became positive. Overall, women’s wage growth is 25-30% less when corrected for selection. Figure 5 suggests that all of the gender gap closure shown in Figure 1 is due to changing selection bias!

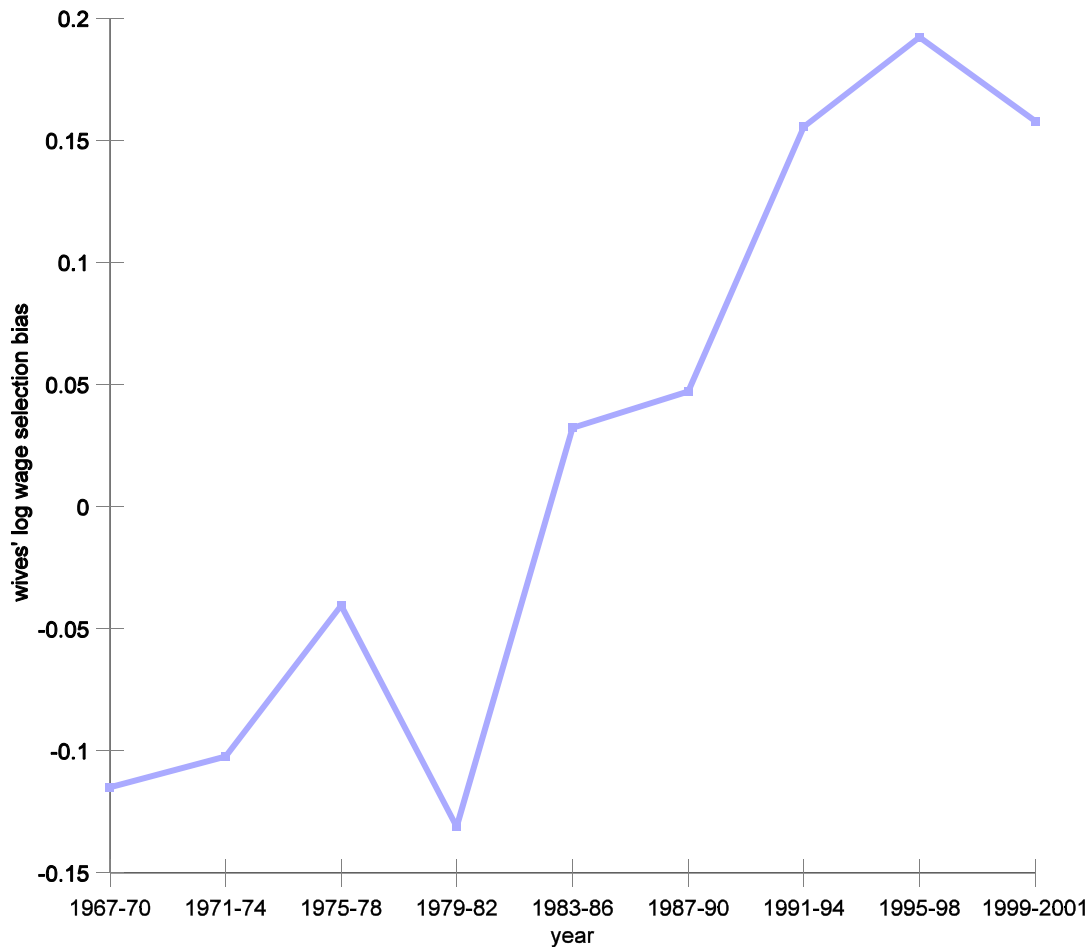


Figure 5 Wives' log wage selection bias over time

Although using different methods and concerned with wage gaps by race rather than gender, Neal (2004) has a result analogous to our Figure 5. More precisely, while we show that the selection bias for women is greater (and of the opposite sign) in recent decades than in the 1970's, Neal shows that the (1990) selection bias is greater (and perhaps of the opposite sign) for black women than for white women. Neal finds a (gender-) differential selection bias of 0.1, while we find a (time-) differential selection bias of as much as 0.3 (see Figure 5).

Although it may be surprising to see how much changing selection bias contributes to measured gender gap closure, it is not surprising that the sign of the selection bias might have changed over time as suggested by Figure 5. For example, married female employment rates have increased somewhat more among the more educated (Juhn and Kim, 1999, Table 2). In the early 1970's, the average education of

working wives' husbands was essentially the same as that for nonworking wives. By 1990, the average education of working wives' husbands exceeds that for nonworking wives by about 1 year. The more dramatic change in the relative quality of the married female work force is shown by Juhn and Murphy (1997) and Juhn and Kim (1999), who stratify married women by their husband's position in the married male wage distribution. In 1970, the employment rate of married women with husbands at the bottom of the male wage distribution was 0.44, as compared to 0.31 for married women with husbands at the top. By 1990, the wives' employment rate was essentially independent of husband's position, for example 0.60 at the bottom and 0.61 at the top. In other words, female labor force growth seems to have come disproportionately from skilled women (see also Topel, 1994).

It is well known that the slope coefficients in women's wage and labor supply equations are sensitive to alternative specifications (e.g., Mroz, 1987). But what about the growth over time in the selection bias terms? First, selection bias growth is similar for various schooling groups.⁵ The various entries in Table 1's last column are selection bias growth for various schooling groups. In all cases the selection bias growth is in the range 0.20-0.32, so that all of the measured gender gap closure for these groups appears to be selection bias growth. Second, married women selection bias growth is not sensitive to the reassignment of variables from the regression equation to the selection equation, or vice versa. Table 2 displays estimates of selection bias growth for Heckit specifications that differ according to the independent variables used in the selection and/or regression equations; six of the eight estimates are in the range 0.21-0.41, with the two extremes as 0.05 and 0.88.⁶ Figure 4 does suggest some specification sensitivity when we include single women because, for example, the selection bias for married high school graduates appears less when we compare them with single high school grads than when we compare them with married advanced degree women. Nevertheless, Figure 4 clearly shows that wage growth falls significantly with the labor force participation – the only question raised by that Figure is whether all, or just half, of the measured gender gap closure is selection bias.

⁵Selection bias growth varies little across schooling groups in all of the specifications we have tried, so henceforth we display selection bias for the “some college” group without reference to those for the other groups.

⁶The level of the selection bias does depend on specification. For example, the selection bias is significantly more negative (less positive) for Table 2's specification (1) than specification (2).

Table 2: Selection Bias Growth from
Various Specifications of the Regression (R) and Selection (S) Equations

independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
number of children 0-6	S	S	R, S	R, S	R, S	S	S	S
husband's age	S	R, S	R, S	R, S	S	S		
husband's education	S	R, S	R, S	S	R, S			S
selection bias growth	0.249	0.217	0.876	0.054	0.413	0.222	0.224	0.208

Notes: (1) each column is a different specification of the Heckit model

(2) wife's experience and schooling included in all regression and selection equations

(3) samples of married white women from the CPS

(4) selection bias growth is the growth from 1975-79 to 1995-99, for the "some college" schooling group, and 15 years of experience

Third, we expect quantitative, but not qualitative, results to be sensitive to wage distribution functional form. As discussed above in relation to equation (1), Figure 3, and Figure 4, the selection bias growth is negligible for groups of women with high and stable labor supply: a qualitative result which does not rely on the assumed lognormal distribution. Since we see log wage growth sloping down with the level of labor supply, we expect the high and stable labor supply groups to tell us the most about genuine gender gap closure, even if the wage distribution were not lognormal. Of course, in order to obtain a *quantitative* estimate of the selection bias growth, the Heckman two-step procedure uses the inverse Mills ratio, which derives from normal functional form. Moffitt (1999) and Newey (1999) have suggested using, as a robustness check on the distributional functional form, monotone transformations of the inverse Mills ratio (equivalently, monotone transformations of the predicted probability of working). Table 3 uses different transformations predicted probability of working and different instrumental variables. Using $1-L$ as the additional regressor (see the middle column), we find the selection bias growth to be essentially the same as with the lognormal model (see the first column). Using both the inverse Mills ratio and $1-L$, we find less selection bias growth (see the last column), although in the same direction.⁷

⁷Interestingly, Newey (1999) suggests that $1-L$ may be more robust to changes in distributional functional forms than the other specifications.

Table 3: Selection Bias Growth from
Various Specifications of Distributional Functional Forms

instrumental variables	$\lambda(L)$	$1-L$	$\lambda(L), 1-L$
number of children 0-6	0.265	0.350	0.145
number of children 0-6, husband's age, and husband's education	0.249	0.267	0.090

Notes: (1) selection bias growth from 1975-9 to 1995-9, for the “some college” schooling group. The 0.249 estimate is the same estimate as shown in the last column of Table 1
(2) L denotes the predict probability of working ffy, and λ denotes the inverse Mills ratio

IV. Further Indicators of the Composition of Working Wives, Related to Identification at Infinity

Growing inequality within gender likely contributes to the closing of the measured gender gap because of its effect on the nature of the selection of women into the labor force. As we point out above, the direction of the bias seems clear, as on various observable characteristics the female workforce has improved its composition. Structural selection models are one way of calculating a numerical gender gap that is free from selection bias. In the spirit of Chamberlin's (1986) and Heckman's (1990) “identification at infinity” argument, we have also proposed to focus on wage growth for groups of women with high and stable labor supply, namely the single women and advanced degree wives featured in our Figures 2 and 4. However, these estimates may still be biased to the extent that the composition of groups of women with high and stable labor supply have changed over time relative to their male comparison group. Subsection A makes two comparisons for advanced degree wives: working wives as compared with the general population of advanced degree wives (ie., regardless of labor force status), and advanced degree wives as compared to the general population of wives with a college diploma or higher. Subsection B displays comparisons of married and single women.

IV.A The Composition of Advanced Degree Working Wives

It is well known that husbands and wives sort positively on many characteristics: height, race, schooling, and (among dual-earner couples) even earnings. Our strategy here and in subsection B is to use the earnings of a woman's husband as a proxy for her own earning ability. Figure 6's solid line is the difference between the average log weekly wages of two groups of husbands: husbands of women who have advanced degrees and the husbands of women who have at least graduated college. The series has trends up since 197, which suggests that the group of advanced degree women has grown via the

addition of women with relatively high earnings potential.

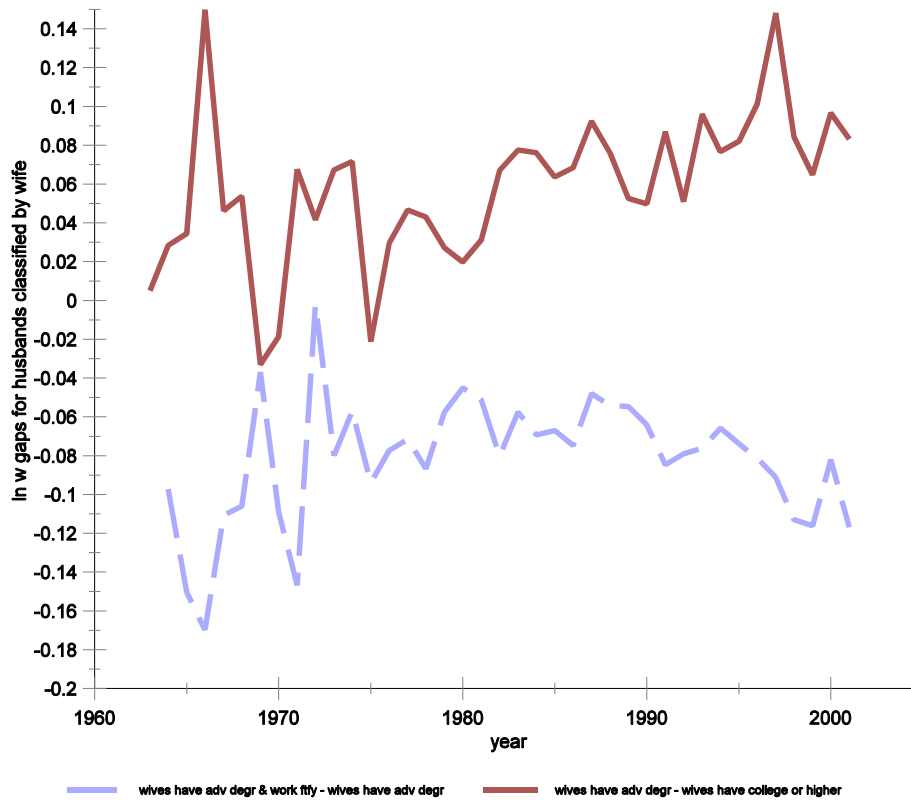


Figure 6 Composition of Advanced Degree Working Wives, using Husband's Earnings as a Proxy

IV.B Determinants of Marriage Rates

Figure 4 also shows that prime-aged single women as a group: (a) supply more labor, and (b) have enjoyed less wage growth than married women. Has single female wage growth been less because their human capital (as a group) has declined relative to that of married women? It is well known that prime-aged single women have become increasingly black, but that is not the explanation of Figure 4 because it uses only samples of white women. Figure 7 displays the fraction of white women, stratified by marital status, with a college or advanced degree. Here we see that human capital has grown for both groups in the same amount until the early 1990's. Since then, the fraction with degrees has grown only for married women, by about eight percentage points.

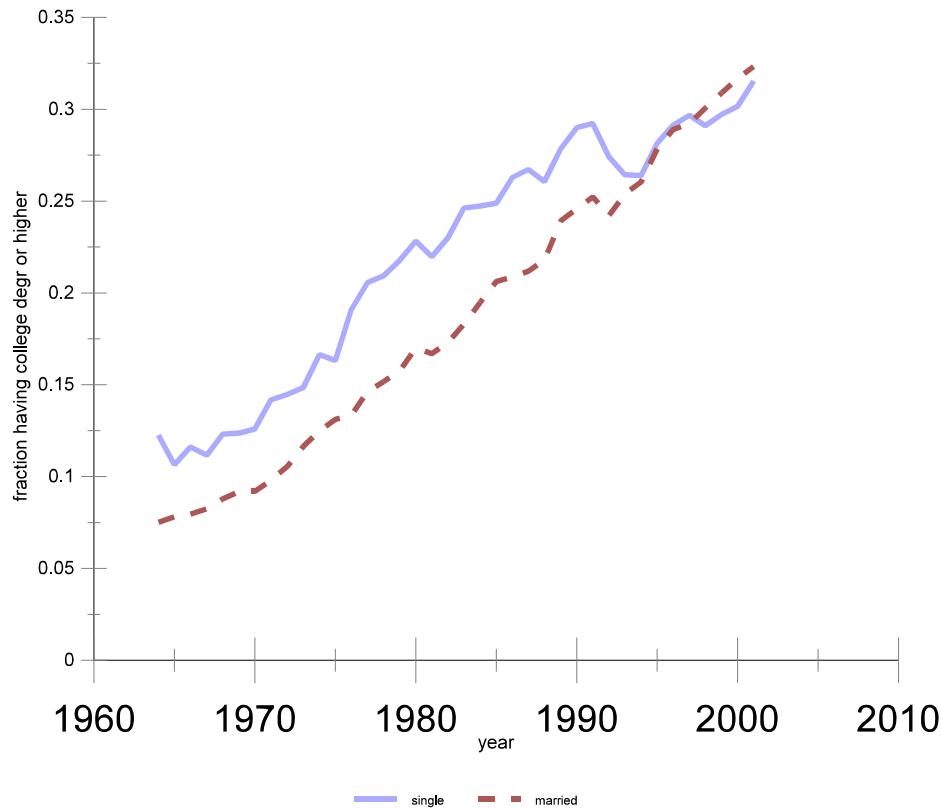


Figure 7 Fraction of White Women aged 25-54 with College Degree or Higher

If declining relative human capital of single were to explain the kinds of wage growth shown in Figure 4, then Figure 7 suggests that it would have occurred since 1990. More work comparing married and single women is needed, but for our calculation it appears that much of the relative decline in wages of single women occurred prior to 1990.

V. Conclusions

Growing earnings inequality has been associated with the loss in earnings potential for some groups of people, and the growth in earnings potential of others. Wives not working in the marketplace include a disproportionate share of people with lost earnings potential, but are not included in calculations of female earnings. Hence, growing earnings inequality within gender has created the illusion of a gender wage gap that closes more rapidly than the average woman has gained relative to the

average man, if she has gained at all. How large is this bias? We offer a parametric estimate and two simple nonparametric estimates, based on the principle that the effects of labor force selection are minimal among groups with high and stable labor supply. The first nonparametric estimate is the gender gap among married people with advanced degrees, which is about 0.25 or 0.30 log points throughout the years 1963-2000. This estimate, which implies that the average married women has gained only 0 or 0.05 log points in earnings relative to the average married man, may hide further selection bias growth and thereby hide an actual opening of the gender gap, because the employment rate of advanced degree women is still less than 100%. Our second nonparametric estimate is the gap between the earnings of single women and the earnings of men with similar education. The second gap closes, but only about half as much as the raw gender gap featured in the literature.

Finally, parametric estimates of the Heckit model (assuming lognormal distributions) suggest that the average married women did not gain, and may have even lost, relative to the average married man. All together, it appears that the relative earnings progress of women since 1963, if any, has been limited. Furthermore, we will see standard measures of the gender gap *widen* in the future, if and when earnings inequality within genders returns to 1970's levels.

We use the Roy model to decompose the “selection bias” into components related to changing labor supply, growing inequality within gender, and a changing cross-sectional correlation between wages and the net gain from working. The first two roughly cancel each other. However, the correlation change itself derives at least in part from growing inequality. Nonwage factors dominated female labor supply decisions in the 1970's when wage inequality was relatively small. By 1990, wages had become unequal enough that they dominated nonwage factors, so that nonworking women tended to be the ones with less earnings potential.

Many in the literature (see Moffitt, 1999, for a survey) have concluded that selection bias may be a relatively minor factor for understanding women's wages. Our Figure 5 shows that this may have been the case for U.S. cross-sections sampled in the late 1970's or early 1980's (remember that the sample made famous by Mroz, 1987, and subsequent work by Whitney et al, 1990, and others, was from 1975). However, our Figure 5 suggests that selection bias was significant in the early 1970's and, in the other direction, since the mid-1980's. At the very least, selection bias makes significant contributions to measured wage growth for women and other groups with relatively weak attachments to the labor force.

Our estimates suggest that women's earnings potential has grown, if at all, far less than previously estimated. Does this mean that female labor supply increases should not be attributed to wage changes, but rather to social forces or technological change in the nonmarket sector (e.g., Goldin and

Katz, 2002; Greenwood et al, 2001)? Answering this question is beyond the scope of this paper, but we argue elsewhere (Mulligan and Rubinstein, 2002) that even if the average woman's earnings potential had remained constant, growing wage *inequality* within gender might have pulled women into the labor force. Hence attributing female labor supply to wages versus other factors is still a topic for future research.

VI. Appendix: The Occupational Composition of Wives with Advanced Degrees

Figure 8's solid line shows how many of the working wives with advanced degrees are teachers. Might this bias our inference from Figure 2? On one hand, teachers have had less wage growth than other college graduates. On the other hand, teachers earn less than other advanced degree women, and the prevalence of teacher's among working wives with advanced degrees has been declining. Also, Figure 8's dashed line shows that the teachers are just as prevalent among working wives with college degrees, so teachers might not bias Figure 2's advanced degree wage growth estimate relative to that for college graduates.

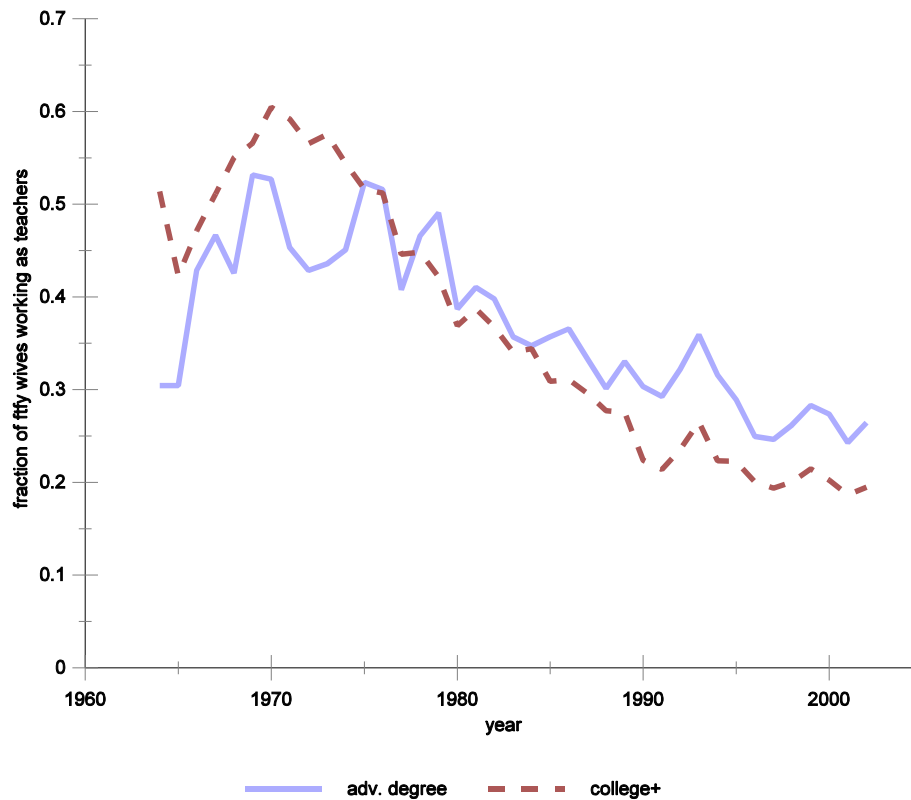


Figure 8 The Prevalence of Teachers by Degree

Figure 9's solid line is the same gender gap for advanced degree women as displayed in Figure 2. The dashed line is the gender gap for nonteachers, which has essentially the same (lack of) trend.

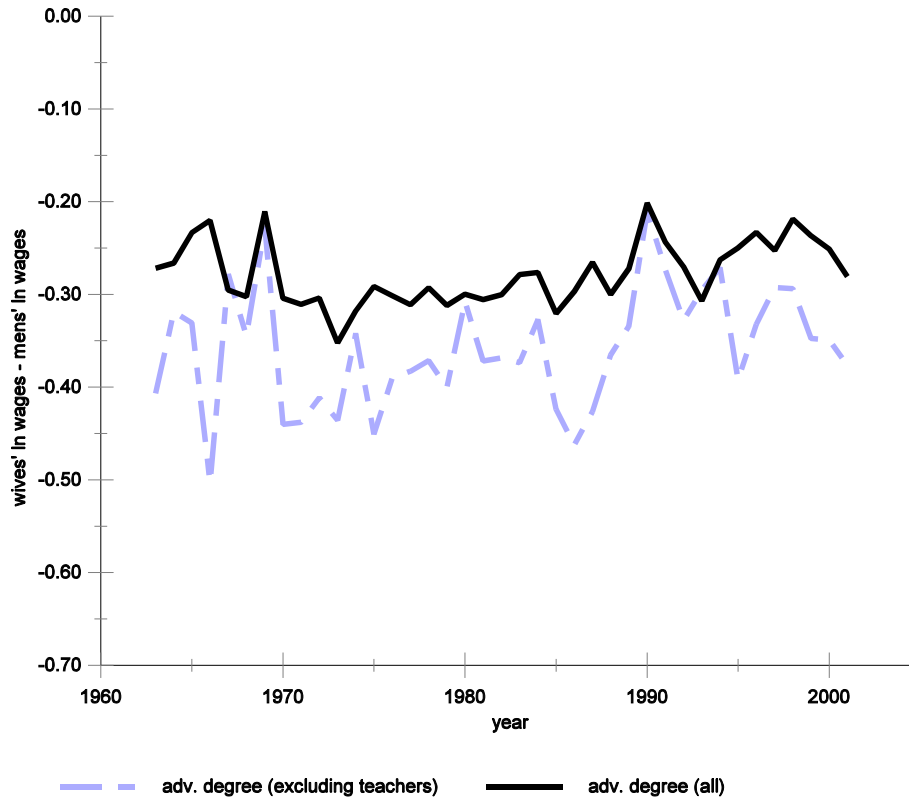


Figure 9 The Advanced Degree Gender Gap by Occupation

Aside from teaching, the main occupations for advanced degree women have been (in order of their prevalence) managers, nurses, physicians, professors, scientists, and lawyers.⁸ Within these six occupations, managers gained the most advanced degree married women, but relatively few married men. The importance and growth of the manager category by itself tends to close the advanced degree gender gap, because managers earn somewhat more than the average advanced degree wife, and because the gender gap among managers closed 0.08 log points. However, closure of 0.08 is much less than we see in Figure 2 for the other education groups. Furthermore, the gender gap seems to have widened for some occupations, such as physicians. In summary, the gender gaps by occupation have closed too little, and

⁸According to the CPS, these six occupations account for about 60% of married prime-aged nonteachers working full-time full-year with advanced degrees, among both women and men, throughout the period 1975-present.

shifts of women toward high wage occupations have been too little, for the overall gender gap among persons with advanced degrees to close significantly.

VII. References

- Ahn, Hyungtaik, and James L. Powell. “Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism.” *Journal of Econometrics*. 58(1), July 1993: 3-29.
- Becker, Gary S. “How the Market Acted Affirmatively for Women.” *Business Week*. May 13, 1985.
- Blau, Francine D. and Lawrence M. Kahn. “Swimming Upstream: Trends in the Gender Wage Differential in the 1980's.” *Journal of Labor Economics*. 15(1), Part 1, January 1997: 1-42.
- Chamberlain, Gary. “Asymptotic Efficiency in Semiparametric Models with Censoring.” *Journal of Econometrics*. 32(2), July 1986: 189-218.
- Das, Mitali, Whitney K. Newey, and Francis Vella. “Nonparametric Estimation of Sample Selection Models.” *Review of Economic Studies*. 70(1), January 2003: 33-58.
- Goldin, Claudia. *Understanding the Gender Gap*. New York: Oxford University Press, 1990.
- Goldin, Claudia, and Lawrence F. Katz. “The Power of the Pill: Oral Contraceptives and Women’s Career and Marriage Decisions.” *Journal of Political Economy*. 110(4), August 2002: 730-70.
- Gottschalk, Peter and Robert Moffitt. “The Growth of Earnings Instability in the U.S. Labor Market.” *Brookings Papers on Economic Activity*. 2, 1994: 217-54.
- Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. “Engines of Liberation.” mimeo, University of Rochester, 2001.
- Gronau, Reuben. “Wage Comparisons-A Selectivity Bias.” *Journal of Political Economy*. 82(6), December 1974: 1119-43.
- Heckman, James J. “Shadow Prices, Market Wages, and Labor Supply.” *Econometrica*. 42(4), July 1974: 679-94.
- Heckman, James J. “Sample Selection Bias as a Specification Error.” *Econometrica*. 47(1), January 1979: 153-61.
- Heckman, James J. “Varieties of Selection Bias.” *American Economic Review*. 80(2), May 1990: 313-318.
- Heckman, James J. and Bo E. Honore. “The Empirical Content of the Roy Model.” *Econometrica*. 58(5), September 1990: 1121-49.

-
- Heckman, James J. and Jose Scheinkman. “The Importance of Bundling in a Gorman-Lancaster Model of Earnings.” *Review of Economic Studies*. 54(2), April 1987: 243-55.
- Juhn, Chinhui and Dae Il Kim. “The Effects of Rising Female Labor Supply on Male Wages.” *Journal of Labor Economics*. 17(1), January 1999: 23-48.
- Juhn, Chinhui, and Kevin M. Murphy. “Wage Inequality and Family Labor Supply.” *Journal of Labor Economics*. 15(1), Part 1, January 1997: 72-97.
- Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. “Wage Inequality and the Rise in Returns to Skill.” *Journal of Political Economy*. 101(3), June 1993: 410-42.
- Katz, Lawrence F. and David H. Autor. “Changes in the Wage Structure and Earnings Inequality.” *Handbook of Labor Economics*. Volume 3A. New York: Elsevier Science, 1999: 1463-1555.
- Katz, Lawrence F. and Kevin M. Murphy. “Changes in Relative Wages, 1963-1987: Supply and Demand Factors.” *Quarterly Journal of Economics*. 107(1), February 1992: 35-78.
- Mincer, Jacob. “Labor Force Participation of Married Women: A Study of Labor Supply.” in H.G. Lewis, ed. *Aspects of Labor Economics*. Princeton: Princeton University Press, 1962.
- Moffitt, Robert A. “New Developments in Econometric Methods for Labor Market Analysis.” in Orley Ashenfelter and David Card, eds. *Handbook of Labor Economics*. Volume 3A. Amsterdam: Elsevier Science, North-Holland, 1999: 1367-97.
- Mroz, Thomas A. “The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions.” *Econometrica*. 55(4), July 1987: 765-99.
- Mulligan, Casey B. and Yona Rubinstein. “Specialization, Inequality, and the Labor Market for Married Women.” manuscript, University of Chicago, July 2002.
- Murphy, Kevin M. and Finis Welch. “Wage Inequality in the 1990's.” paper presented at the 1999 American Economic Association Meetings, January 1999.
- Neal, Derek. “The Measured Black-White Wage Gap among Women Is Too Small.” *Journal of Political Economy*. 112(1), Part 2, February 2004: S1-28.
- Newey, Whitney K. “Consistency of Two-Step Sample Selection Estimators Despite Misspecification of Distribution.” *Economics Letters*. 63(2), May 1999: 129-32.
- Roy, A.D. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers*. 3(2), June 1951: 135-46.
- Topel, Robert H. “Regional Labor Markets and the Determinants of Wage Inequality.” *American Economic Review*. 84(2), May 1994: 17-22.