

NBER WORKING PAPER SERIES

SELECTION, INVESTMENT, AND
WOMEN'S RELATIVE WAGES SINCE 1975

Casey B. Mulligan
Yona Rubinstein

Working Paper 11159
<http://www.nber.org/papers/w11159>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2005

We appreciate the comments of Josh Angrist, Gary Becker, Jim Heckman, Aitor Lacuesta, Amalia Miller, Kevin M. Murphy, June O'Neill, John Pepper, Chris Rohlfs, Ed Vytlačil, Yoram Weiss, seminar participants at Chicago, Houston, Rice, and Texas A&M, the research assistance of Ellerie Weber, and the financial support of the National Science Foundation (grant #0241148). We were encouraged by Larry Katz to explore in detail the links between wage inequality within and between genders, for which we thank him. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2005 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.

Selection, Investment, and Women's Relative Wages Since 1975
Casey B. Mulligan and Yona Rubinstein
NBER Working Paper No. 11159
February 2005
JEL No. J24, J31, J16, C34

ABSTRACT

In theory, growing wage inequality within gender should cause women to invest more in their market productivity and should differentially pull able women into the workforce, thereby closing the measured gender gap even though women's wages might have grown less than men's had their behavior been held constant. Using the CPS repeated cross-sections between 1975 and 2001, we use control function (Heckit) methods to correct married women's conditional mean wages for selectivity and investment biases. Our estimates suggest that selection of women into the labor market has changed sign, from negative to positive, or at least that positive selectivity bias has come to overwhelm investment bias. The estimates also explain why measured women's relative wage growth coincided with growth of wage inequality within-gender, and attribute the measured gender wage gap closure to changing selectivity and investment biases, rather than relative increases in women's earning potential. Using PSID waves 1975-93 to control for the changing female workforce with person-fixed effects, we also find little growth in women's mean log wages. Finally, we make a first attempt to gauge the relative importance of selection versus investment biases, by examining the family and cognitive backgrounds of members of the female workforce. PSID, NLS, and NLSY data sets show how the cross-section correlation between female employment and family/cognitive background has changed from "negative" to "positive" over the last thirty years, in amounts that might be large enough to attribute most of women's relative wage growth to changing selectivity bias.

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
Tel Aviv University
yonar@post.tau.ac.il

Table of Contents

| | |
|--|----|
| I. Introduction | 1 |
| II. Empirical Methods for Estimating Potential Wages in the Presence of Selection, Investment, and Growing Wage Inequality | 7 |
| Investment and Selection Interpretations of Control Functions | 7 |
| Observed and Unobserved Selection in an Extended Roy Model | 9 |
| III. Control Function Estimates of Unobserved Female “Selection” in the U.S. Data since 1975 .. | 14 |
| Heckit Estimates for Repeated Cross-Sections | 14 |
| Sources of Variation: Wage Growth and Labor Supply Across Various Groups | 19 |
| (In)Sensitivity of the Estimates to Alternative Specifications and Functional Forms | 22 |
| Estimates Based on Women who Leave and Join the Labor Force | 24 |
| IV. Changing Selection into the Female Labor Market: Evidence from Measures of Family and Cognitive Background | 26 |
| Consistency of the Heckit’s Reduced form Labor Supply Function with Background Indicators | 26 |
| Inferring the Amount of Wage Selection Bias From Observed Background Indicators | 31 |
| V. Conclusions | 32 |
| VI. Appendix I: Common Wage Implications of Selection and Investment | 34 |
| Wage Concepts | 35 |
| Selection and Investment Simultaneously Determined | 36 |
| VII. Appendix II: Converting Background Selection Bias into Wage Selection Bias | 37 |
| VIII. References | 39 |

I. Introduction

The U.S. labor market has experienced some dramatic changes over the past thirty years. First of all, within-gender earnings inequality has increased (see Levy and Murnane, 1992, and Katz and Autor, 1999, for comprehensive surveys). Inequality grew over this period not only from an increase in the Mincerian returns to education but also due to growing 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). At the same time, the measured market labor supply and earnings of women have substantially, although not fully, caught up with the earnings of men. The purpose of this paper is to further link these three trends, and show how doing so has important implications for understanding the structure of labor demand.

Various observers have noted that wage inequality within-gender and wage equality between genders have been curiously coincidental (Card and DiNardo, 2002, p. 742; Blau and Kahn, 1997, p. 2). Figure 1's 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*).¹ The dashed line is a measure of *inequality* within gender, namely, the ratio of the 90th percentile to the 10th percentile in the cross-section earnings distribution of men working *ftfy*. We see that both were flat until 1977 or so (see also O'Neill, 1985, on the apparent constancy of the gender wage gap prior to 1977). Both rose – most rapidly at first – from the late 1970's until about 2000. What can we conclude about the structure of labor demand over this period? Have there been changes in discrimination and other nonwage factors in the operation of the labor market

¹The solid line's trends would be quite similar if gender equality were measured in terms of hourly wages rather than earnings, at least for the years since 1979 when relatively good hours data is available from the CPS (see Gunderson, 1989). Also notice that the scope for a measured gender hours gap is limited, given that we limit our samples to those working *ftfy*.

(Becker, 1985; Katz and Murphy, 1992)? Can we conclude that earnings have multiple and largely independent determinants, such as brains versus brawn (Welch, 2000)? Maybe an outward shift in the demand for brains helps women and hurts men at the bottom of the male wage distribution? Is part of the coincidence due to a depressing effect of female labor supply on the wages of men (eg., Topel, 1994; Juhn and Kim, 1999; Fortin and Lemieux, 2000)? Our paper suggests that (a) apparent gender equality is a direct consequence of inequality within gender, which stimulates changes in female investment and labor supply behavior, 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.

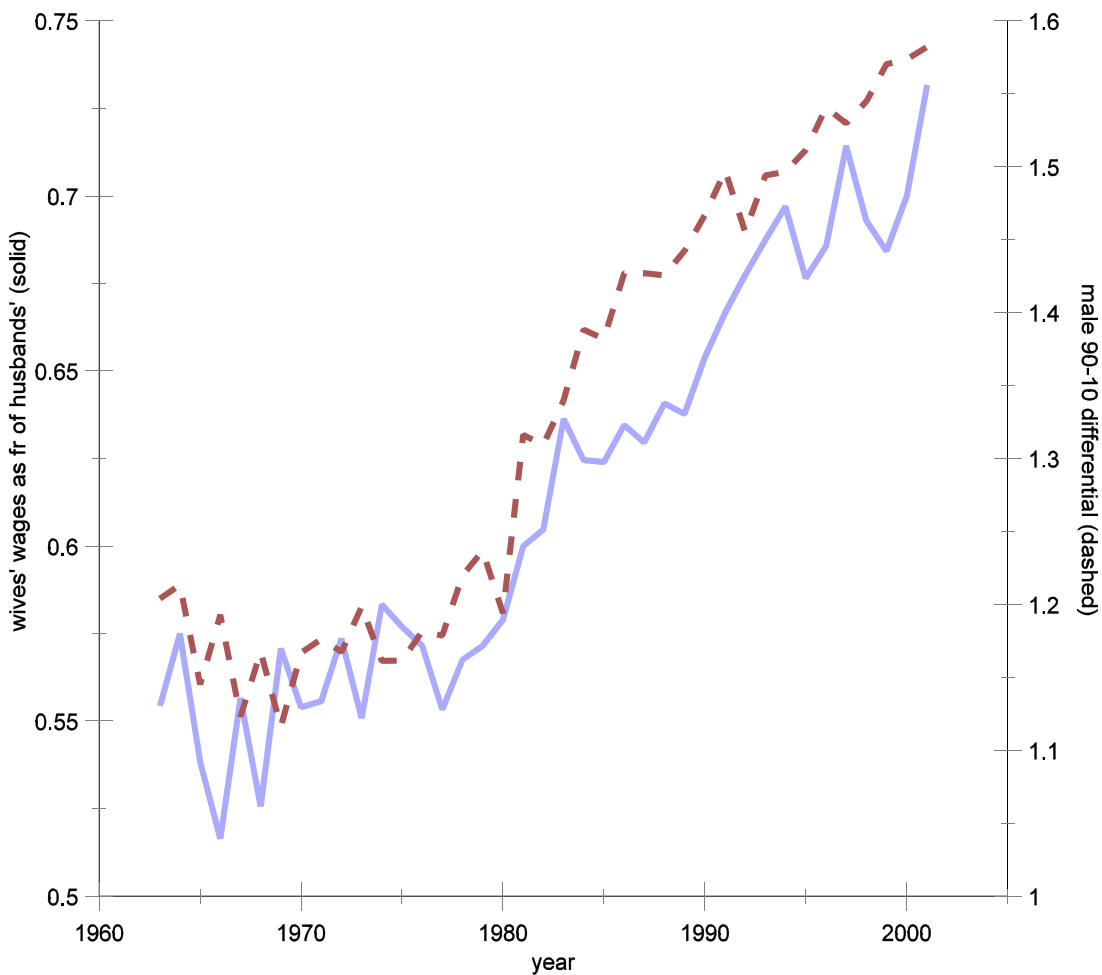


Figure 1 Wage Inequality between and within Genders

Suppose for the moment that wages were determined by some combination of gender and human

capital, and that growing wage inequality within-gender is an indicator of a shift in the demand for human capital in favor of those with relatively large amounts of it. One female response might be to increase their market human capital investment.² Some of this response might be readily observed, for example, as more women accumulating years of work experience, attaining college degrees, or studying subjects like business, law, and medicine with greater market returns. Some of the response, perhaps in the form of effort or on-the-job training, may be less readily observed by labor economists, and hence create an increase in women's measured wages conditional on their observed characteristics. A second female response might be for those with less human capital to drop out of the labor force, and for those with more human capital to enter the labor force. The second response would be observed as an increase in various skill proxies – such as schooling, IQ, etc. – of the female workforce relative to the female population as a whole. To the extent that human capital is unmeasured, the second response would also be observed as an increase in women's measured wages conditional on their observed characteristics. A third possibility is that women do not change their behavior at all, but that the working women have been selected from the right tail of the female potential wage distribution. In this case, the measured wages of working women increase, while the unmeasured potential wages of nonworking women decrease, so that measured wage growth for women overstates the potential wage growth for the average woman.

Even if the amount and distribution of female employment were driven by something other than growing wage inequality within-gender (see Goldin and Katz, 2002, and Greenwood et. al., 2001, for some likely possibilities), it is still possible that measured women's wages have grown because women invest more or because women with high earnings potential have differentially entered the labor force. If we are to say something about the structure of labor demand, the empirical question becomes, for example, would a woman command a higher wage had she been born in 1970 rather than 1950, holding constant her behavior and other characteristics? How can we answer this question, given that measured wages have changed, at least in part, due to changing investment and labor supply behavior, and that the potential wages of nonworking women are not readily measured? Our data and empirical methods are discussed below in some detail, but the example of married women with advanced degrees illustrates the point. Work is very common for this group of women – about 80% of them have some earnings during the year and about half work fty – this has been true at least since 1970. Married women with high school and college degrees work a lot less, especially in the 1970's. Based on their higher and more

²This would also be a response for men; see below for a discussion of why women might respond more than do men.

stable labor force participation rates, we expect advanced degree women to be closer to their male counterparts in terms of their investment in market human capital, and that investment changes would do relatively little to close the gender wage gap among people with advanced degrees. The higher and more stable labor force participation rates also suggest that advanced degree women would be less selected into the labor force (as compared to women with just high school or college degrees) and whatever selection contributes to their measured wages is relatively stable over time. Figure 2 displays the gender wage gap calculated as the average log earnings for married women with advanced degrees working fty, net of the average log wage of fty men with advanced degrees. The solid line shows little growth in the relative earnings of women.³ In contrast, other schooling groups have fewer women working (the 1970 percentage of women working fty is shown in parentheses in the legend), and thereby a greater potential for selection, composition, and investment biases, and show growth in the relative earnings of women much like that shown in Figure 1.

³Figure 2's calculations of the advanced degree gender gap are based on 400+ annual CPS observations of married women working full-time, full-year.

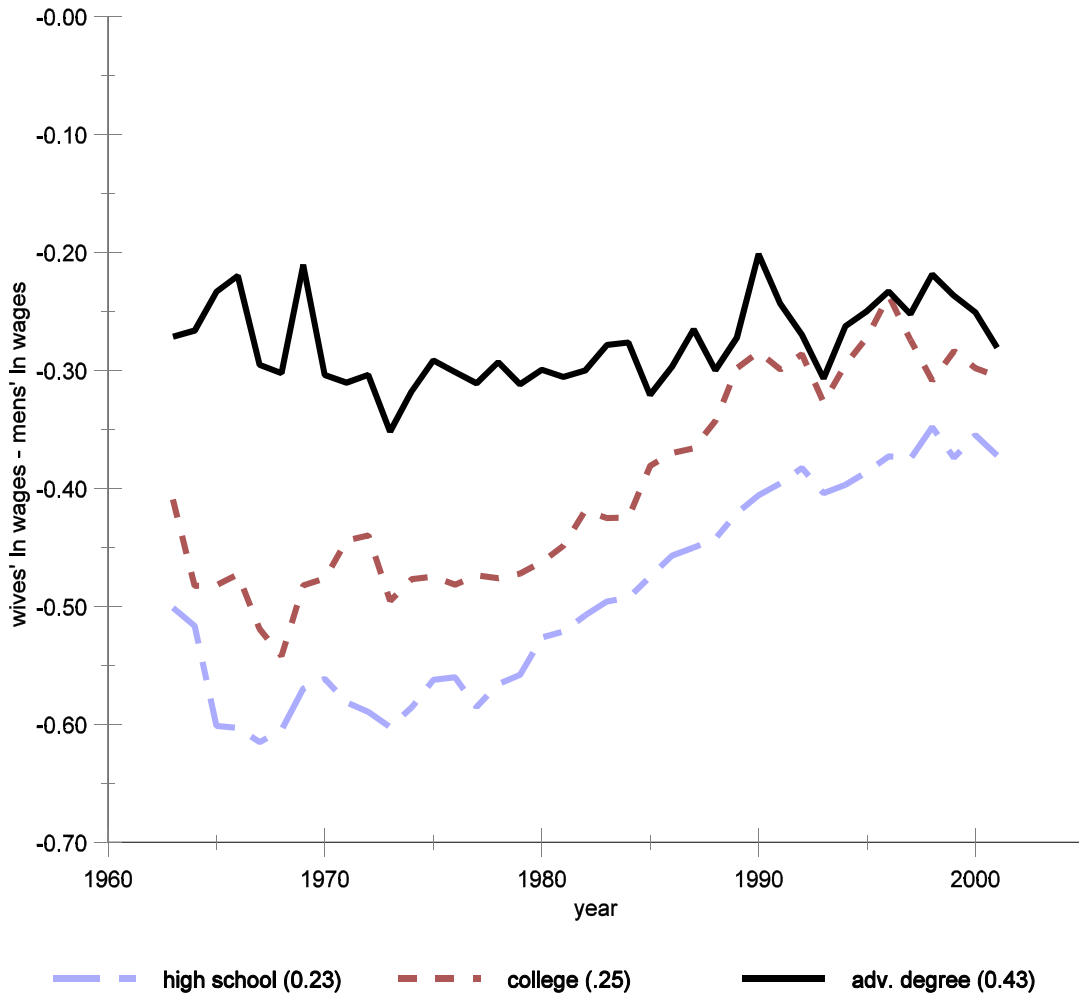


Figure 2 The Gender Gap by Schooling

Although we ultimately find that at least some measured relative wage growth for women must be attributed to growing selection bias, previous work such as Rosen (1992) and Goldin and Katz (2002) shows that women have also significantly increased their market-oriented human investments. Hence, empirical methods are needed to calculate – separately, or at least jointly – the contributions of investment and selection to measured wages. Section II explains how control function methods may be appropriate during an era of growing wage inequality, even though the methods were developed for models ostensibly of selection, but not investment. The economic theory behind control function methods, the Roy model, can, with a minor extension, be used to link observed and unobserved selection, and thereby suggest some alternatives to control function methods, and permit a first estimate of the

relative importance of selection versus investment biases. Section III uses control function methods to estimate the amount of selection bias, and to calculate a time series of women's wages that correct for the effects of their changing investment and labor supply behavior. The estimates suggest that selection of women into the labor market has changed sign, from negative to positive, or at least that positive selectivity bias has come to overwhelm investment bias. The estimates also explain why measured women's relative wage growth coincided with the growth of wage inequality within-gender, and attribute the measured gender wage gap closure to changing selectivity and investment biases, rather than relative increases in women's earning potential.

We acknowledge some of the criticisms of control function methods for estimating selection bias. Recent research has developed more robust control function methods (see Heckman and Vytlacil, 2004, for a survey), and many of our results suggest that control function methods are appropriate for the particular empirical question posed in this paper: how much would women's wages have grown if their labor force attachment had remained constant? Nevertheless, it would be nice to see observed selection be qualitatively and quantitatively consistent with control function estimates of unobserved selection. Section III uses the 1975-93 waves of the PSID to show how most married female relative wage growth can be attributed to work history and person fixed effects, rather than wage growth within person. Section IV looks at some measures of cognitive and family background, showing how women working in the 1980's and 1990's typically had better backgrounds than women not working, but women not working in the 1970's had better backgrounds than women working. Furthermore, the methods discussed in Section II are used in Section IV to calculate estimates of the amount of unobserved selection bias, and its change over time. These findings reinforce our control function estimates. They also suggest that the cross-section correlations between female labor supply and personal characteristics cannot be fully explained by investment behavior of women reacting to their labor market situation, but rather selection bias growth is an important contributor to measured women's wage growth.

Throughout the paper, we put most of our attention to married people. First of all, most of the growth in female labor supply and in female wages (in excess of male wages) has been among married women.⁴ For example, from 1975 to 1995 the gender earnings gap among ffy women aged 25-54 closes twice as much for married women: from -0.58 to -0.41 for married women, and from -0.25 to -0.16 for single women (-0.53 to -0.34 for all women combined). The fraction of married college women working ffy grew from 0.30 to 0.46 over that period, but hardly grew among single college women (from 0.67

⁴See also Mincer, 1962, Goldin, 1990, and much other important work on the female labor market that focuses on married women.

to 0.69). Just as important, our application of the Roy model emphasizes the systematic selection of significant numbers of people out of the workforce, which, aside from the very few independently wealthy, hardly makes sense for prime-aged single women.

II. Empirical Methods for Estimating Potential Wages in the Presence of Selection, Investment, and Growing Wage Inequality

II.A. Investment and Selection Interpretations of Control Functions

What would be the distribution of wages if all persons were forced to participate in the labor market? The answer to this question depends in part on the degree to which less-than-full participation suppresses investment, and creates selection bias. Consider first the effect of changing investment on measured wages. Decompose the investment change into a change conditional on labor force attachment, and a change in response to growing labor force attachment. Algebraically, we write the actual average log year t wage $w_t(X)$ into an average log potential wage $\hat{w}_t(X)$ and an investment component (more precisely, a lack of investment component) $f_t(p_t(X))$

$$w_t(X) = \hat{w}_t(X) + f_t(p_t(X)) \quad (1)$$

where X is the measured characteristics of the group being considered, and $p_t(X) \in [0,1]$ is the amount the women (of working age in year t) expect to work during their lifetime, as a fraction of lifetime full work. By definition of potential wage, $f_t(1) = 0$. f_t increases with p , and its shape depends on the nature of returns to investment. f might change its shape over time due to changes in investment and its returns conditional on p . Changes in p affect measured wages via movements along the f function. We expect the function f to shift for both men and women – both have an incentive to invest more and, even if they invest the same, both enjoy higher investment returns. For example, we expect f to have a greater slope since 1980, when the return to skill has been greater. Movements along f would be more important for women than for men, because women have increased their labor supply relative to men.

Even if investment behavior were constant over time, a second difficulty is that wages are measured only for those people who work, and the types of people who work may change over time. As shown by Heckman (1979), selection bias also appears as a specification error like we see in equation (1). More precisely, when we average equation (1) for a cross-section of women working, we have

$$E[w_t(X)|X, \text{persons working}] = \hat{w}_t(X) + f_t(p_t(X)) + \beta_t \lambda(L_t(X)) \quad (2)$$

where λ is the inverse Mills ratio, $L_t(X)$ is the fraction of women in group X who work, and β_t is a coefficient whose sign and magnitude depends on the direction and degree to which wages are correlated with work status.⁵

Presumably L_t and p_t are positively correlated over time and across groups; among women who expected to be strongly attached to the labor force, we see a lot of them working, and vice versa. Furthermore, both of the last two terms of equation (2) disappear as women become strongly attached to the labor force, because $f_t(1) = \lambda(1) = 0$. This has three important implications for calculating the function of interest, $\hat{w}_t(X)$. First, both investment and selection models suggest including monotonic “control” functions of p or L in a cross-group regression of measured wages on characteristics X , as long as those functions disappear at p (or L) equal one. The exact control function to be used depends both on the shape of f and the distribution of unobservables determining work status, although in the latter case strategies have been devised to estimate $\hat{w}_t(X)$ even when distributional functional forms are unknown (see Newey, 1999, for one example, or Heckman and Vytlacil, 2004, for a survey). We suspect, but leave to future research to prove, that similar strategies would be robust to alternative specifications of f as long as f were monotonic and disappeared at $p = 1$.

Second, the gap between $\hat{w}_t(X)$ and measured wages can change sign over time. For example, f_t is an increasing function, while λ is a decreasing function, so the correlation between measured wages and p (or L) can change sign over time as β changes in magnitude. Even if $f_t = 0$, so there were no investment bias, β can change its sign over time (i.e., selection bias can be “positive” or “negative”) – a possibility that seems very real given the changing correlations between labor supply and background characteristics (see below). Mulligan and Rubinstein (2004) explain how control function methods of estimating selection bias are consistent with this possibility, while various imputation and matching methods are not. Also note how, even if β were always nonnegative, control function estimates might give the false impression of negative selectivity bias, although in such cases the estimates would still

⁵Our Appendix I shows in more detail how a Roy model amended with investment implies a wage equation like (2). As we know from Heckman (1979) and the subsequent literature, cross-section conditional wage distributions need to be lognormal in order for the last term in equation (2) to involve the inverse Mills ratio, rather than some other monotonic function disappearing at $L = 1$.

correctly imply that the average measured wage is less than the average potential wage, which is the more important implication for our purposes.

Third, both selection and investment suggest the same “identification at infinity” strategy to estimate $\hat{w}_i(X)$. The possibility of systematic selection of groups into the labor market motivates Chamberlin (1986) and Heckman (1990) to propose empirical strategies like this, but the alternative possibility of insufficient investment by weakly attached groups provides essentially the same motivation. Our calculation of a gender wage gap for advanced degree women is a heuristic version of these strategy. Based on their higher and more stable labor force participation rates, we expect advanced degree women to be closer to their male counterparts in terms of their investment in market human capital, and that selection bias should be smaller and more stable over time.

Investment and selection models may have so much in common that it is quite difficult to separately distinguish their effects on observed wages. Both of them may have significantly responded to the growing attachment of women to the labor market. Both models stimulate interest in the same counterfactual – what would be the distribution of wages if all persons were forced to participate in the labor market – and hence would be hard to distinguish with instrumental variable and other strategies for isolating special instances of high or low market participation. We leave it to future research to carefully and quantitatively contrast investment and selection.⁶ For now, we point out that the “potential wage function” $\hat{w}_i(X)$ is economically interesting, can be identified without separating the f and λ terms, and can be estimated with some of the tools that have been developed for models ostensibly of selection, but not investment. Henceforth, we combine equation (2)'s last two terms and refer to the combined specification error as a “selection bias.” As we explain below, there are good reasons to believe that selection bias, taken literally, is a quantitatively important contributor to the growth of measured women's wages. Nevertheless, the reader should remember that some of what sometimes appears to be selection bias is in fact a bias deriving from cross-sectional differences in investment behavior.

II.B. Observed and Unobserved Selection in an Extended Roy Model

The Roy (1951) model, as applied to the choice between market and nonmarket work (see also Gronau, 1974, and Heckman, 1974), describes each person 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. To these two

⁶Investment might cause the measured gender wage gap to widen at first, as women forgo market earnings in order to accumulate human capital more rapidly.

variables we add a “background indicator,” which has no direct impact on the labor supply decision, although it may be correlated with the reservation wage, the potential market wage, or both. As we explain below, adding a background indicator connects observed and unobserved selection, thereby accomplishing three things. First, we can calculate the sign of observed selection bias, and a rough estimate of the amount of unobserved selection bias, without using control function methods. Second, because selection and investment are convolved by control function estimates, but not by calculations of observed selection, we can make a first attempt at gauging the relative importance of selection versus investment biases. Third, the model reconciles with our results some previous (and small) estimates of the magnitude of composition and selection biases.

Each person's market wage, reservation wages, and background indicator are drawn from a joint lognormal distribution, whose parameters may vary over time and across groups,

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

where w and r denote log market and reservation wages, respectively, and hats denote medians. As above, X denotes the measured characteristics of the group being considered. b denotes the background indicator. Since it may be correlated with r and w , the model allows for possibilities like IQ or husband's wage affecting productivity at home and in the marketplace. b is measured for both workers and nonworkers, which is why we call it a “background indicator” and will refer to it as schooling, parents' schooling, husband's wage, or IQ in the empirical work.

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 $E(w|z>0)$:

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

$$L(X) = \Phi\left(\frac{\hat{w}(X) - \hat{r}(X)}{\sigma_z}\right) \quad (3)$$

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

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 (3) 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, 2004), 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.⁷ 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.

By analogy with equation (3), we can also calculate the average background indicator for the workforce, $E(b|z > 0)$:

⁷As shown in the second formula (3), 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.

$$E(b|z > 0) = \hat{b}(X) + \lambda(L)\rho_{bz}\sigma_b \tag{4}$$

$$\rho_{bz} \equiv \frac{\rho_{bw}\sigma_w - \rho_{br}\sigma_r}{\sigma_z}$$

We are not necessarily interested in the variable b *per se* and its relations with labor supply. Even if we were, the conditional and unconditional means of b are straightforward calculations, because b is observed for all persons, regardless of their work decision. Equation (4) is important because its selection bias has three relations with that shown in equation (3). First of all, equation (3)'s selection bias can also be interpreted as an investment selection bias. In contrast, the equation (4)'s selection bias would not be confused with an investment bias, at least with an appropriate background measure. For example, if we found that the female workforce had a higher average IQ than did the female population, we would interpret this as a selection bias, rather than women investing in their IQ in anticipation of strong labor force attachment.

Second, both (3) and (4) have the same inverse Mills ratio in their selection bias term. Hence, the form of composition bias for the wage equation (3) emphasized by O'Neil (1985) and Blau and Kahn (1997, 2004) should also be evident in the background equation (4). Third, the sign of the background selection bias is the sign of ρ_{bz} , which can be estimated by comparing the average background indicator for the workforce with the average background indicator for the population as a whole. If ρ_{br} were positive – we expect women with “good” backgrounds to be more productive outside the market – ρ_{bz} can be negative and become positive as wage inequality grows. In fact, we can bound the wage selection bias by converting the observed background bias from “background units” to “wage units.” One such conversion is to multiply the background selection bias by the background coefficient from the regression of log wage on characteristics X and the background indicator.⁸ This is related to the conversions made in the wage gap literature, and the resulting estimate tends to be small (see below for empirical examples). If the Roy model were right, this conversion is likely to be too small because selection into the labor force is on wages, and may only be incidentally correlated with background because background and wages happen to be (imperfectly) correlated. In the extreme case where background and wages were uncorrelated, this procedure would yield an estimate of zero wage selection bias even when wage selection bias were in fact quite large.

⁸If we thought the relation between background and potential market wage were the same for men and women, this regression could be estimated for men with less concern for selection bias.

Another conversion *divides* background selection bias by background's *reverse* regression coefficient, namely, the coefficient from a regression of the background indicator on log wage and the characteristics X . It is easy to show that this wage selection bias estimate is greater than the estimate cited above, for the same reason that the inverse of a reverse regression coefficient is larger than the forward regression coefficient. Nevertheless, the ratio of background selection bias to the reverse regression coefficient *underestimates* the wage selection bias as long as the partial correlation between background and reservation wage (holding constant potential market wage) were positive. Our Appendix II proves these results, but the intuition is clear. Using the reverse regression coefficient corrects for the fact that selection is on wages, and only incidentally on background. However, it does not correct for the possibility that somebody with high background and wages also has high reservation wage, and hence the magnitude of the covariance between b and w overstates the magnitude of the covariance between b and $z (= w-r)$.⁹ With some qualitative restrictions on the joint distribution of w , r , and b , which follow pretty easily from standard theories of labor supply, we can bound the amount of wage selection bias, although not estimate its precise magnitude.

In summary, the extended Roy model suggests two basic methods for estimating the amount of wage selection bias in a cross-section. The first method is familiar from Heckman (1979), namely to first consider the labor supply part of equation (3), ideally with a good instrumental variable, and second to enter an inverse Mills ratio as one of the regressions in a log wage regression. The ideal instrumental variable is correlated with the reservation wage and has a zero partial correlation with the potential market wage. By having these properties, the ideal instrumental variable would permit the isolation of groups of women who differ in terms of their labor supply L but are otherwise similar; comparing these groups on measured wages can then be interpreted as a selection effect. We pursue the first approach in Section III, with some attention to the instrumental variables, but with most of our attention on estimating selection bias in repeated cross-sections. The second method is to compare the female workforce to the female population in terms of a background indicator. A background indicator that is good for this purpose is likely a poor instrumental variable. The ideal background indicator is positively correlated with the market wage and uncorrelated (or at least positively correlated) with the reservation wage. Hence our empirical work uses variables like IQ, education, and parental education as background indicators for female labor supply models. Section IV shows how estimates based the two methods turn

⁹It follows that, if the partial correlation between b and r were negative, the magnitude of the covariance between b and w understates the magnitude of the covariance between b and z , and the proposed conversion overstates the amount of wage selection bias.

out to have similar repeated cross-section implications for female wages and labor supply.

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 fifty samples (namely full-time workers who report working at least 50 weeks of the previous year).

III. Control Function Estimates of Unobserved Female “Selection” in the U.S. Data since 1975

III.A Heckit Estimates for Repeated Cross-Sections

If we modify equation (3) 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 $\lambda\sigma_w\rho_{wz}$ 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. The change over time in $\lambda\sigma_w\rho_{wz}$ is qualitatively ambiguous because λ falls and σ_w rises, but the Heckman selection model permits numerical estimates of $\lambda\sigma_w\rho_{wz}$. 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) | 0.003 (0.001) | 0.003 (0.001) | 0.010 (0.001) | 0.010 (0.001) | |
| (experience-15) ² /100 | -0.004 (0.005) | -0.007 (0.005) | -0.046 (0.005) | -0.043 (0.005) | |
| high school dropout | 9.799 (0.013) | 9.914 (0.034) | 9.723 (0.019) | 9.525 (0.032) | 0.313 |
| high school graduate | 10.012 (0.007) | 10.108 (0.027) | 9.999 (0.007) | 9.850 (0.021) | 0.245 |
| some college | 10.099 (0.011) | 10.193 (0.028) | 10.194 (0.007) | 10.050 (0.020) | 0.238 |
| college graduate | 10.303 (0.014) | 10.386 (0.026) | 10.548 (0.008) | 10.412 (0.020) | 0.219 |
| advanced degree | 10.519 (0.021) | 10.585 (0.028) | 10.827 (0.011) | 10.709 (0.019) | 0.184 |
| teacher | 0.032 (0.017) | 0.033 (0.017) | -0.235 (0.013) | -0.233 (0.013) | -0.001 |
| observations | 20,971 | 20,971 | 28,931 | 28,931 | |
| $\sigma_w \rho_{wz}$ | 0 | -0.075 (0.020) | 0 | 0.161 (0.021) | |
| adj-R ² | .08 | .08 | .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 0.095 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 0.143 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 0.238 log points), in large part because wage inequality has grown within gender.

Figure 3 displays time series for wives' log wage "selection" bias.¹⁰ More specifically, Figure 3 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 20-30% less when corrected for selection. Figure 3 suggests that all of the gender gap closure shown in Figure 1 is due to changing selection bias!

¹⁰As discussed above, our references to "selection bias" may, for the moment, also refer to "investment bias" or some combination.

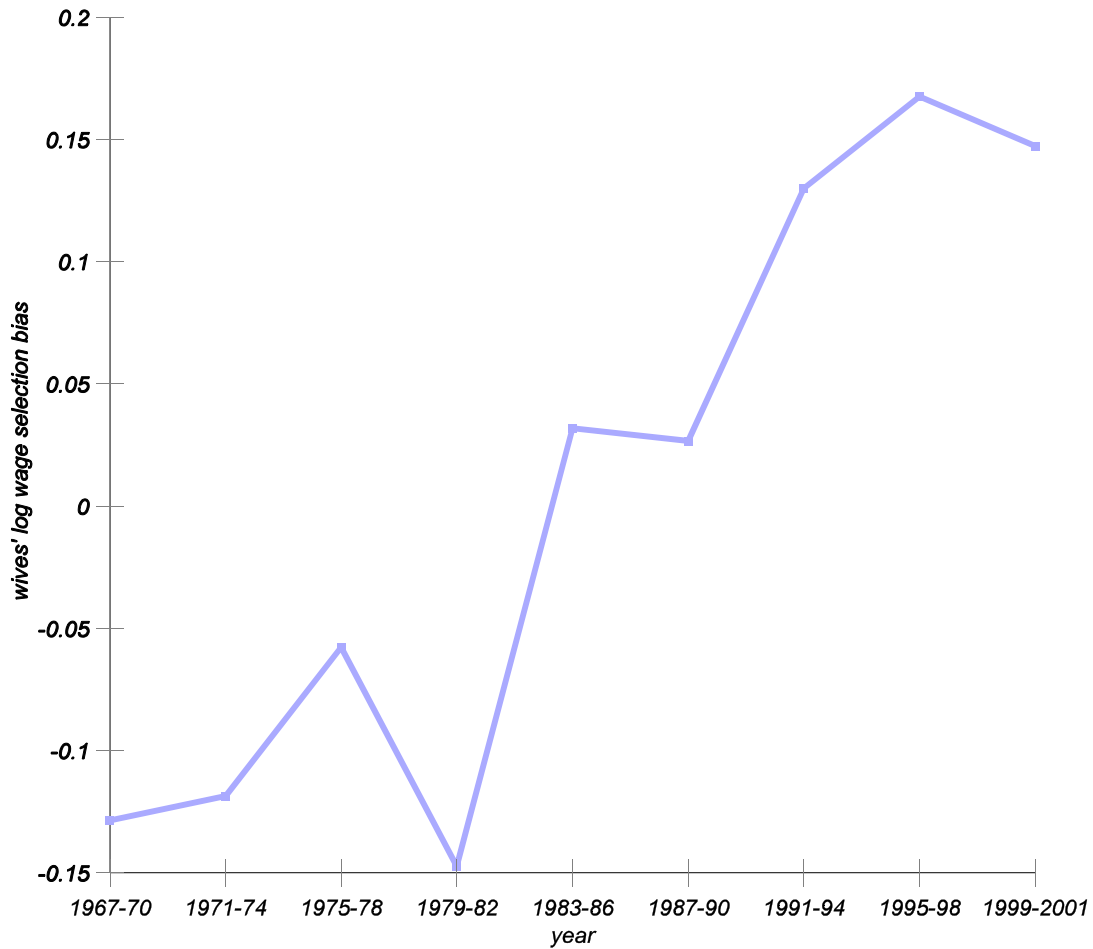


Figure 3 Wives' log wage selection bias over time

Selection bias and its growth is presumably much less important for married men, and essentially nil for married men with college degrees, because of their stronger attachment to the labor force. Nevertheless, growing inequality probably pulled some low potential wage men out of employment (Murphy and Topel, 1997, Welch, 1997, and Autor and Duggan, 2003, explain how) and a complete calculation of the gender gap in potential wages would include selection bias corrections for the husbands. Calculating the male corrections is beyond the scope of this paper because control function methods are less advantageous for male samples than for female samples, but Mulligan and Rubinstein (2004) report that the male corrections are small enough that the gender potential wage gap is well approximated, especially for college and high school graduates, by subtracting Figure 3's series from the

college and high school series in Figure 2.

Although using different methods and concerned with wage gaps by race rather than gender, Neal (2004) has a result analogous in some ways to our Figure 3. 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 3). Mulligan and Rubinstein (2004) also explain how different empirical methods are needed for estimating selection bias in the present context, namely when growing inequality is closing the gender gap over time.

III.B Sources of Variation: Wage Growth and Labor Supply Across Various Groups

The Heckit estimates of wage selection bias grow over time because groups of women with initially high levels of labor supply had less wage growth over time. In order to critically evaluate those estimates, it helps look directly at wage growth and labor supply across groups. Figure 4's horizontal axis measures group labor supply 1974-8. The vertical axis measures group wage growth relative to men with the same schooling. For example, the two data points nearest the bottom of the graph have vertical position measured as the wage growth of the married and single, respectively, advanced degree fty workforce, both measured relative to the wage growth of men (regardless of marital status) with advanced degrees. Their horizontal position by the fraction of the women-years in the schooling-marital subsample with fty work. For example, 25% of high school wives worked fty in a typical year 1974-78. The Figure shows how 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 a additional wealth effect on wives (through the wages of their husbands).

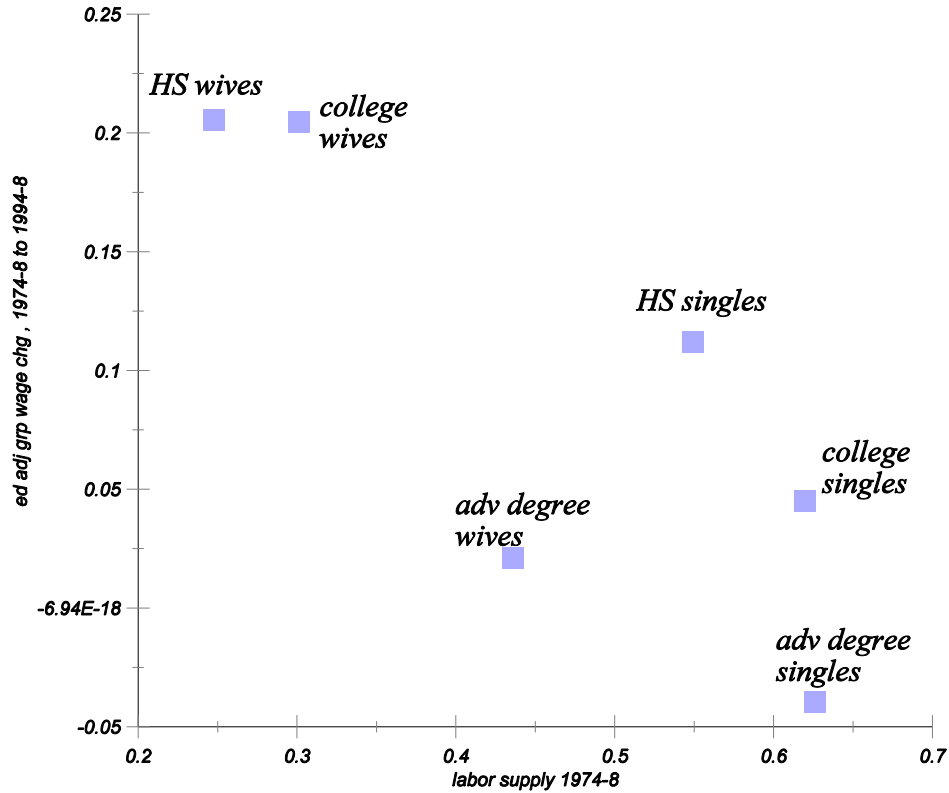


Figure 4 Measured Wage Growth Declines with Labor Supply (White Women)

According to the selection and investment bias interpretations, the wages of advanced degree women increase less over time because they have always been strongly attached to the labor force. Might this be better explained by a declining relative “quality” (i.e., earnings potential) of the advanced degree group? After all, the fraction prime-aged women with 18+ years of schooling tripled 1970-90, while the male fraction only doubled. This logic also implies that wages should have grown less for high school wives, and perhaps also college wives, which is not supported by the data. Elsewhere we have looked at various proxies for earnings potential such as husband’s wage, and have found no evidence for this effect. We interpret Figure 4 as saying that if all wives had been as attached to the labor force as single women or advanced degree wives, their wages would not have grown relative to men’s.

Figure 5 is an even closer analogue to the Heckit estimates, because the groups are defined according to the number of young children – one of the instrumental variables used in our Heckit models.

It examines the partial relation¹¹ between wages and number of children, using two CPS samples. The horizontal axis measures the number of children 0-6, which we have ordered right-to-left so that, by analogy with Figure 4, the level of labor supply increases from left-to-right. The first sample uses married white women for the years 1975-9, and its sample averages are displayed as a dashed line. It slopes up, which is no surprise given that women with more children are less attached to the labor force. The solid line shows how the correlation between number of children and measured log wages became positive on the 1990's. In order measure wage growth rates, take the vertical distance between the solid and dashed lines in Figure 5. As in Figure 4, we see a negative relation between measured group wage growth and the level of labor supply.

¹¹The residuals from a regression of log wage on experience, experience squared, and education dummies are normalized to be zero for women with zero children aged 0-6. The average residual conditional on number of children is shown in the Figure.

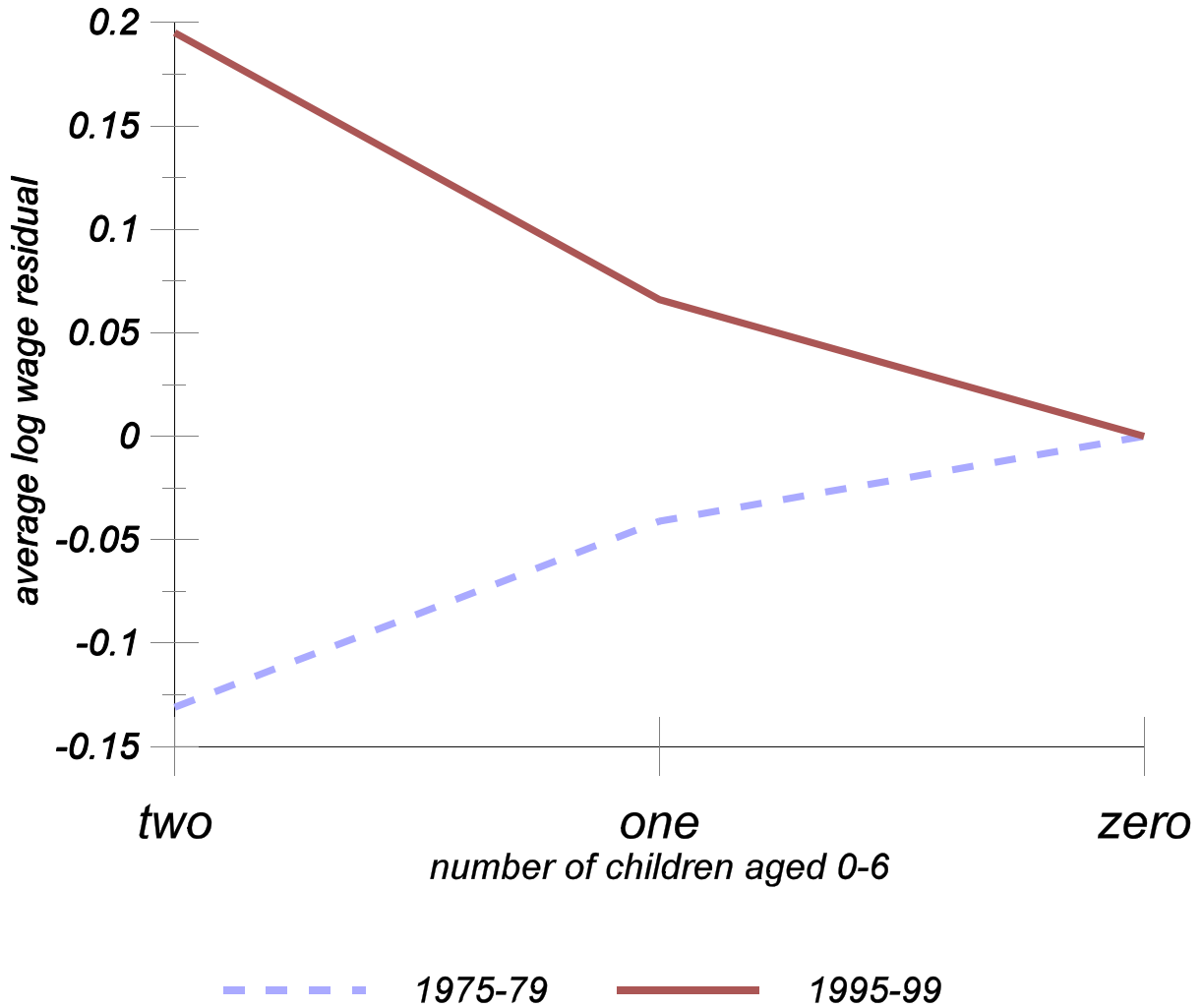


Figure 5 Children and Wives' Measured Wages

III.C (In)Sensitivity of the Estimates to Alternative Specifications and Functional Forms

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? We show elsewhere (Mulligan and Rubinstein, 2004, Table 2) how selection bias growth is positive and economically significant regardless of what combination of instruments we use. Moffitt (1999) and Newey (1999) raise the question of whether results are sensitive to the normal distribution assumption—a question which arises again when thinking about the possibility of investment bias (see the introduction and Appendix I of this paper), in which case the wage equation specification error depends on the shape of the human investment technology as well as the underlying distribution of unobservables. Mulligan and Rubinstein (2004, Table 3) find statistically and economically

significant selection/investment bias growth when the inverse Mills ratio is replaced by other control functions (i.e., monotone functions of L that disappear at $L = 1$).

The use of children as an instrumental variable in the female labor supply equation dates back to the early work applying the Roy model to the female labor market (e.g., Heckman, 1974). It may not be the ideal instrumental variable because high earnings potential may discourage a woman from having children, but how does this possibility affect our findings of selection bias growth over time? In order to derive an answer, we consider the Heckit wage equation, which includes the inverse Mills ratio as a regressor, and note that the error term may be negatively correlated with children (i.e., children are endogenous). The inverse Mills ratio is positively correlated with children, because wives with children work less. Hence the estimated regression coefficient on the inverse Mills ratio suffers from a downward omitted variable bias, and may even give the false impression of negative selectivity bias. Presumably the downward omitted variable bias would exist in both the 1970's and 1990's cross-sections, so would it affect our selection bias *growth* estimates? In what direction? Arguably, the endogeneity bias is worse in the 1990's, because the magnitude coefficient on kids (or determinants of kids) in the wage equation increased over time with inequality, like the magnitude of so many other coefficients in the wage equation. If so, the endogeneity of children causes us to underestimate the amount of selection bias growth.

Because measured wages for prime-aged men are less subject to selection bias than are measured wages for women, male wage regressions offer a test of the hypothesis that selection-corrected children coefficients have increased their magnitude over time. For example, the coefficient on number of children aged 0-18 in a married male wage regression is 0.005 (s.e. = 0.0015) for the (pooled) 1975-80 cross-sections, and 0.016 for the (pooled) 1995-2000 cross-sections.¹² The magnitude of the coefficient on kids in the male wage regression increases over time, as does overall wage inequality. Hence, if we had selection-corrected female wage regressions, we would expect the magnitude of the children coefficient to also increase over time, which, as argued above, means that the magnitude of the endogeneity bias induced by excluding children from the Heckit wage equation also increases its magnitude over time.

¹²Also included in the regression are schooling and experience variables for the husband. Results are similar if schooling and experience for the wife are included too.

III.D Estimates Based on Women who Leave and Join the Labor Force

If we are right to claim that labor demand did not shift in favor of women, then the measured relative wages of women should increase over time merely because the women who enter the workforce tend to have higher lifetime wages than the women who leave it. Table 2 uses panel data from the Panel Study of Income Dynamics to examine this point. Each column is a log wage regression, for which the dependent variable is measured only for person-years in which there was full-time employment. The regressors are gender, year dummies, gender interacted with 2-year dummies, schooling, and (when indicated) an experience dummy. Reported in the Table's first row is the 1992-75 difference in the coefficients on female-year dummy interactions. The first column is calculated from our repeated CPS cross-sections, so the estimate of 0.075 is analogous to the 0.100 increase in height of the dashed line in Figure 1 during the years 1975-92 (except for the fact that the regression controls for personal characteristics, while the Figure only displays cross-section means). The second column is the same specification estimated with PSID data, treating it as repeated cross-sections. The third and fourth column add a work experience dummy and person-fixed effects, respectively, to the previous specification. We see that female wage growth over the years 1975-92 is economically and statistically insignificant once we control for work experience and person fixed effects.

Table 2: Within-person Estimates of Women's Relative Wage Growth, 1975-92

| | dependent variable: log real hourly wage | | | |
|--|--|------------------|------------------|------------------|
| | (1) | (2) | (3) | (4) |
| coefficient on 1991-2 dummy*female [benchmark years are 1975-6] | 0.075 (0.011) | 0.087 (0.046) | 0.046 (0.045) | 0.000 (0.035) |
| work experience variable? | no | no | yes | yes |
| person fixed effects? | no | no | no | yes |
| R-squared | .03 | .12 | .14 | .77 |
| sample | CPS | | PSID | |
| observations | 173,894 | | 26,437 | |
| individuals | 173,894 | | 4,995 | |

Notes: (1) OLS standard errors in parentheses.

(2) years are 1975-92.

(3) sample is all white married persons aged 25-54, working full-time, and with real (year 2000 \$) hourly wage at least \$1.

(4) other regressors are gender, 2-year dummies (1975-76, 1977-78, etc.), and (when indicated) a dummy for having worked full-time two years prior.

Although Table 2 is consistent with the hypothesis that measured female relative wage growth derives mainly from some combination of selection and investment bias growth, it is also (by itself) consistent with the hypothesis that selection and investment bias growth are nil, and that measured female wages grew because the new cohorts of women entering the workforce enjoyed a relative demand shift. For example, discrimination might affect women mainly upon labor force entry, so that declining discrimination appears as a female cohort wage growth rather than wage growth within-person. Another possibility is that growing inequality by itself widened the gender gap within person, because women tend to have wages like men from the left tail from the male wage distribution, and that the reason that we see essentially no wage growth in Table 2's second column is that this effect was counterbalanced by labor demand's shifting in favor of women. Nevertheless, the panel estimates are worth seeing because they could have refuted the selection hypothesis if we had found a lot of female relative wage growth within-person.

IV. Changing Selection into the Female Labor Market: Evidence from Measures of Family and Cognitive Background

Taken literally, the Heckit estimates suggest two important conclusions. First, selection bias has grown in part due to the effect growing inequality would have with fixed labor supply behavior. Second, selection bias was negative in the 1970's, and changed sign over time because labor supply has changed – women with high unobserved earnings potential disproportionately entered the labor market. The first conclusion also an investment interpretation, and for our purposes it is unimportant to distinguish the interpretations. The second conclusion would be more credible if women with high observed proxies for earnings potential had also entered the labor market, and did so in proportions consistent with the Heckit estimates. The purpose of this section is to present some evidence of negative selection on observables in the 1970's, and to show how the direction and magnitude of labor force changes of women grouped by their observables are consistent with the Heckit estimates for the unobservables.

IV.A Consistency of the Heckit's Reduced form Labor Supply Function with Background Indicators

Figure 6 displays the labor supply functions implied by our Heckit estimates for the 1970's (dashed) and 1990's (solid). The horizontal axis measures the percentile of the residual wage distribution and the vertical axis the fraction of women working ftfy.¹³ The Heckit estimates imply a downward slope for the 1970's, because the inverse Mill's ratio has a negative coefficient in the wage equation for the 1970's. In the language of the Roy model, high wage women were less likely to work because they have especially high reservation wages.

¹³These functions have a vertical position that depends on the values of the observables. For the purposes of making Figure 6, we positioned the functions so that the median labor supply was 0.245 in the 1970's and 0.448 in the 1990's, thereby matching the level of labor supply for entire sample of married women. More specifically, Figure 6 graphs the function

$$\Phi\left(\Phi^{-1}(\hat{L}(t)) + \frac{\rho_{wz}(t)\sigma_z(t)}{\sigma_r(t)}\Phi^{-1}(\theta)\right),$$

where θ is the wage percentile and t is the time period (1975-79 or

1995-99). $\hat{L}(t)$ is median labor supply (0.245 or 0.448). ρ_{wz} and σ_w^2 are the estimated correlation and variance, respectively, from the first stage of the Heckit model. σ_z is estimated as the ratio of an assumed value of the college coefficient (0.25 for the 1970's, 0.39 for the 1990's) to college's estimated coefficient in the probit equation. σ_r is calculated as $\sigma_r^2 = \sigma_z^2 + \sigma_w^2 - 2\sigma_z\rho_{wz}\sigma_w$.

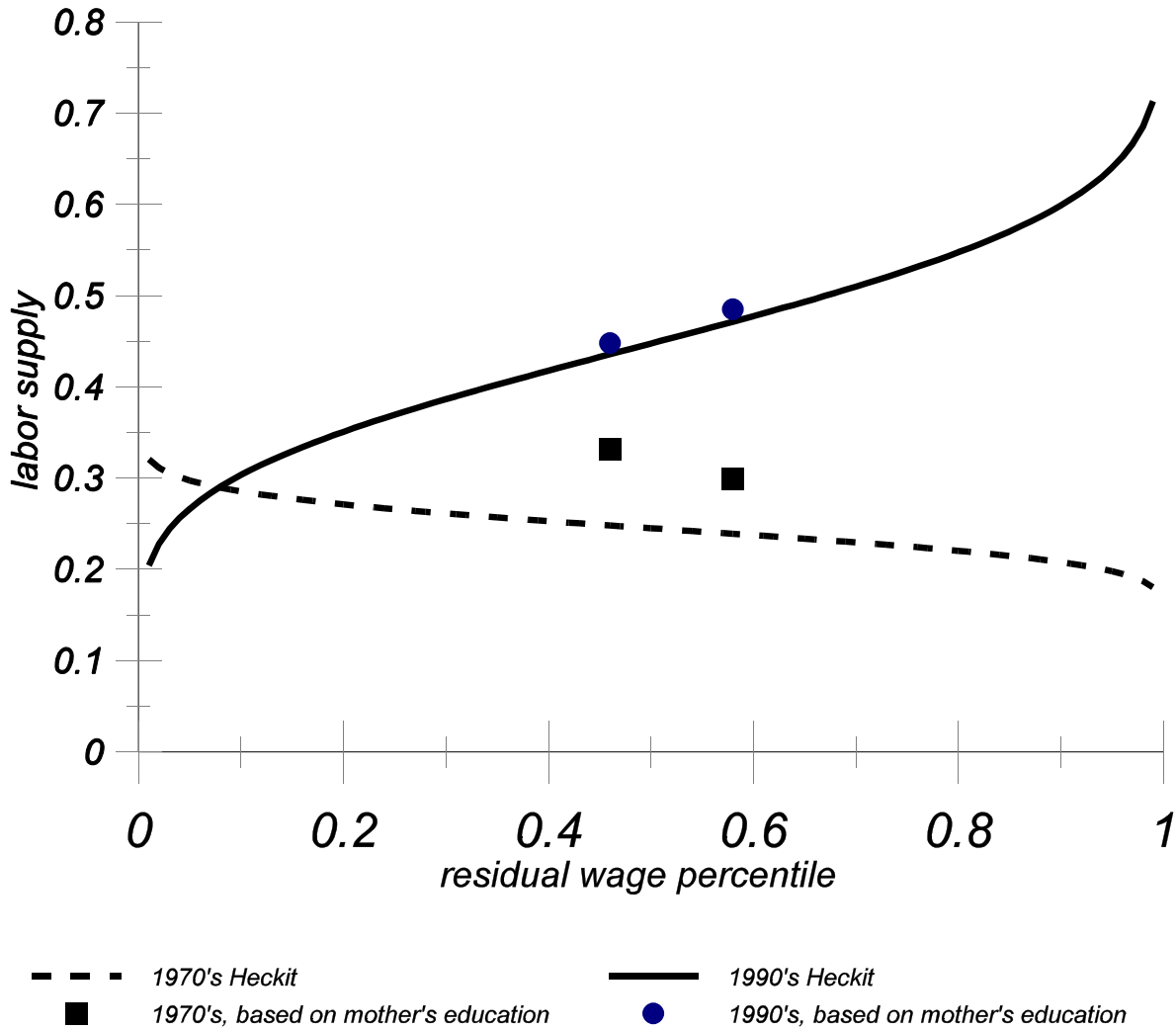


Figure 6 Labor Supply Functions Implied by Heckit Estimates

The labor supply functions can also be interpreted as the weights that should be convolved with the potential wage distribution in order to calculate the average measured wage. A perfectly horizontal supply curve corresponds to random selection, in which case the average measure wage is equal to the average potential wage. Hence, the more the downward (upward) sloped is the labor supply function, the more downward (upward) biased is the measured average wage as an estimate of the average potential wage. Furthermore, our finding of selection bias growth on the order of 0.20-0.25 implies that labor supply participation rates should have increased little for groups of women sampled from the left tail of the potential wage distribution, and more than double for groups of women sampled from the right tail. For example, our Heckit estimates imply that labor supply increased 30 percentage points at the 75th

percentile, and only ten percentage points at the 25th percentile. Indeed, if we could sample from, say, the 5th percentile of the potential wage distribution, our Heckit estimates predict that labor supply would have *fallen* over that time.

We used family background from the PSID to place the circles and squares in the Figure, as checks on the Heckit estimates. First, we regress log wage on age dummies, education dummies, a time dummy, and a dummy for whether mother had a high school degree or higher in a PSID sample of married *ffy* men, using the years 1974-9 and 1989-93. By analogy with our CPS specifications, we define the residual to be the log wage not explained by the age, education, and time dummy variables. Our second step is to calculate the position in the male log wage residual distribution of the median man with mother not high school educated, and of the median man with more high school educated (or better). These positions turn out to be the 46th and 58th percentile (among men with some college education), respectively. Under the assumption that the joint distribution of mother's schooling and log potential wages is the same for married women as it does for married men, our third step is to put the median married woman with mother not high school educated (at least high school educated) at the 46th (58th) percentile of the residual married female potential wage distribution.¹⁴ Fourth, we calculate the average labor supply for married women with (without) high school educated mother, holding own age and education dummies constant, for the periods 1974-9 and 1989-93. The third and fourth steps create four data points in Figure 6, which are marked as squares (circles) for 1974-9 (1989-93), respectively. We notice that 1970's labor supply is higher in the PSID than in the CPS (i.e., the squares are above the dashed line). Otherwise, the squares match the dashed line well in terms of slope, and the circles match the solid line well both in terms of levels and in terms of slope. If anything, the four data points suggest that labor supply growth increases more with residual wage percentile than indicated the the CPS estimates of the Heckit model, although the maternal schooling data is not giving us much information about very high and very low residual wage percentiles.

It is easy to see how married women with high potential wages tended to be less likely than other married women to be out of the labor market in the 1970's, and more likely to be in the labor market in the 1990's. 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

¹⁴Under the alternative assumption that mother's education puts a women in the same percentile of the married *ffy* female residual distribution as in the married female residual distribution, we find the percentiles to be 46 and 65, respectively, using a log wage regression for *ffy* married women.

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.

Recall how equations (3) and (4) link two correlations: the correlation between background b and labor supply, and the correlation between wages w and labor supply. In particular, if both were negative in the beginning, growing wage inequality could change their sign. Table 3 reports coefficients from labor supply regressions, which might inform us about the sign of the first correlation, ρ_{bz} . Columns (1) and (2) show how test score had a negative sign in a female labor supply equation in the 1970's, but a positive sign in the 1990's.¹⁵ Columns (3) and (4) show how women from more educated families were once less likely to work full-time, and now are more likely to work.¹⁶

¹⁵Test score is measured as IQ in the 1970's, AFQT in the 1990's, both of which are normalized to have standard deviation equal to one. The 1970's data are from the original 1968 NLS cohort (wages measured in 1973), and the 1990's data from the 1979 youth cohort (NLSY) with wages measured in 1993.

¹⁶We can pool the two PSID samples and reject the hypothesis that the mother's schooling coefficient is constant over time; the t-statistic is 3.7.

Table 3: Selectivity Bias has Changed Sign

| | works full-time full-year | | | |
|--|---------------------------|----------------|-------------------|------------------|
| | (1) | (2) | (3) | (4) |
| test score | -0.01 (0.02) | 0.02 (0.01) | | |
| mother has at least high school degree | | | -0.033 (0.015) | 0.038 (0.013) |
| years | 1973 | 1993 | 1974-9 | 1989-93 |
| pseudo R-squared | .02 | .02 | .02 | .02 |
| sample | NLS68 | NLSY | PSID | PSID |
| observations | 991 | 2046 | 16,658 | 16,658 |

Notes: (1) samples of married white women
 (2) coefficient estimates are marginal probabilities from a probit
 (3) other regressors include age dummies, education dummies, and a constant
 (4) NLS79 test score is an IQ test; NLSY test score is AFQT. Both test scores scaled so that the sample mean is zero and the sample standard deviation is one.
 (5) PSID standard errors are clustered by person

Recall Figure 5's partial relations between wages and number of children. For the years 1975-9, women with more children have lower measured wages, which is no surprise given that women with more children are less attached to the labor force. For example, investment models would predict that women with more children would invest less in their human capital, due to their lesser attachment. If selection were labor market "negative," selection models predict that working women with more children would have lower wages because they are a particularly select sample of all women with the same number of children. Both models predict that the slope of the dashed line depends on the returns to skill. In the investment model, a higher return to skill should increase the magnitude of the negative slope. In the selection model, the slope is likely to turn positive. The solid line shows how, in fact, the correlation between number of children and measured log wages became positive on the 1990's. Hence, Figure 5 suggests that selection bias has changed sign over time, and that at least some of the changes over time are selection bias rather than investment bias.¹⁷

¹⁷The upward slope for the 1970's can also be interpreted in the reverse direction, namely that women's wages decrease the number of children due to the substitution effect. Obviously, this cannot explain the downward slope for the 1990's, or why the slope might have changed over time.

IV.B Inferring the Amount of Wage Selection Bias From Observed Background Indicators

Table 4 adds to this evidence by exploring the composition of the married male and female fifty workforces according to the worker's cognitive and family background. The first three rows of the first column shows one instance of changing behavior by married women during the years 1975-99: the married female workforce has gained 1.2 years of schooling, while the married male workforce gained only 0.93. The married female workforce also got "better" in terms family background. Over time, married female workers were increasingly likely to have mothers with a college degree, more so than were married male workers.

Table 4: Selection Bias Growth as Indicated Family and Cognitive Background

| | own years of schooling | mother is high school grad |
|--|------------------------|----------------------------|
| (1) change for married white male workforce | 0.93 | 0.19 |
| (2) change for married white female workforce | 1.20 | 0.21 |
| (3) female-male difference | 0.27 | 0.02 |
| (4) log wage partial regression coefficient | 0.09 | 0.26 |
| (5) selection bias growth, assuming selection on background. (5) = (3)*(4) | 0.02 | 0.006 |
| (6) log wage partial reverse regression coefficient | 1.60 | 0.18 |
| (7) selection bias growth, assuming selection on wages. (7) = (3)/(6) | 0.17 | 0.13 |
| time period | 1975-9 to 1995-9 | 1974-9 to 1989-93 |
| addendum: measured gender wage gap closure for the period | 0.12 | 0.09 |

Notes: (1) Own schooling sample is from the CPS. Mother's schooling sample is from the PSID.
 (2) "workforce" is the sample of people working fifty
 (3) addendum is from Figure 1

In summary, the first three rows of Table 4 show us how selection into the female workforce has changed over time in the direction of women with apparently greater earnings potential joining the

Also recall our earlier calculations for men – namely that male analogue to Figure 5 would have a downward slope for the 1970's and the 1990's, but to a much lesser degree.

workforce. However, based on this information, how much might changing selection have closed the gender wage gap? After all, the (married) female workforce has gained only 0.27 years of schooling relative to male workers! We know from the male wage distribution that each year of schooling adds about nine percent to earnings (see the Table's fourth row); a regression of log wage on years of schooling and other demographic variables yields a coefficient of 0.09 in a cross-section of married men. So, if the effect of schooling on wages were about the same for men and women, we might guess that the increased relative schooling of the female workforce added only two percent ($0.02 = 0.027 \cdot 0.09$) to relative female wages. This is the kind of calculation made by Blau and Kahn (1997, 2004) and others, and would be appropriate if the changing selectivity into the female workforce were on schooling. However, the Roy model says that selectivity is on *wages* (and reservation wages), so that any changes in the female workforce in terms of schooling, family background, etc., is *only incidental to the selectivity on wages*. In other words, the 0.02 calculation would be correct only if the selection were 100% on schooling and 0% on the wage. As we showed above by adding the background indicator to the Roy model, an alternative calculation *divides* the 0.27 by the log wage coefficient in a schooling regression. The sixth and seventh rows of Table 4 report the reverse regression coefficients and the resulting estimate of selectivity bias growth: 0.17 and 0.13. According to these estimates, selectivity bias growth *exceeded* the measured gender wage gap closure (see the last row).

There are several good reasons to prefer the control functions estimates over the estimates shown in row (7). For example, row (7) could be biased down because own schooling and parental schooling is positively correlated with the reservation wage (see Appendix II). On the other hand, measured wages are imperfectly correlated with the theoretical concept of the wage,¹⁸ so that the reverse regression coefficient is biased down. We don't intend to present the seventh row as the best estimators of selection bias growth, but merely to illustrate how apparently small changes in the relative composition of the female workforce in terms of their educational and family background might be generated by large changes in the selectivity on market wages.

V. Conclusions

We present four arguments in favor of the hypothesis that recent apparent gender wage equality is not real in the sense that the average woman's earnings potential has not caught up with that of the

¹⁸For example, the model is about lifetime wages, whereas we measure the wage in a single year. Also, hours, and therefore hourly wages, are probably mismeasured.

average man. The first is theoretical: growing wage inequality within gender should cause women to invest more in their market productivity and should differentially pull into the workforce women with high earnings potential. Second, the empirical gender gap has not closed much among advanced degree women, perhaps because they have been strongly attached to the labor force throughout the sample period (the 1970's, 1980's, and 1990's).

Third, Heckit estimates in repeated cross-sections imply that selection bias, investment bias, or some combination thereof, has grown significantly over time. Namely, in the 1970's, women invested little in their market human capital, and/or the women who worked tended to be from the left tail of the female potential wage distribution. By the 1990's, women invested more in their market human capital, and/or the women who worked tended to be from the right tail of the female potential wage distribution. The Heckit estimates of combined selection and investment bias growth seem to be robust to alternative specifications of the instruments, and of the shape of the control function.

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. Ignoring for the moment the difficulties with distinguishing selection bias from investment bias, our Figure 3 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 Newey et al, 1990, and others, was from 1975). However, our Figure 3 suggests that selection bias was significant in the early 1970's and, in the other direction, since the mid-1980's. Selection bias may also be a major part of the answer to the question posed by Blau and Kahn, 2004, namely that the measured gender wage gap closed less rapidly in the 1990's because selection bias *grew* much less in the 1990's than it did in the 1980's.

Negative selection bias is hard to distinguish from investment bias, because both imply that measured women's wages understate their potential wages by an amount that declines with the degree to which the women are attached to the labor force. Our fourth finding is that selection on some observed background characteristics was indeed negative in the 1970's and became positive by the 1990's. Women working in the 1980's and 1990's typically had better backgrounds (in terms of own schooling, cognitive test scores, and parental schooling) than did non-working women, but women not working in the 1970's had better backgrounds than did working women. Hence, at least part of women's relative wage growth is due to growing selection bias, narrowly defined. Preliminary calculations suggest that selection bias growth may even account for all of the measured gender wage gap closure. However, the reader should note that much more work is needed to carefully separate selection bias from investment bias, and that the sum of selection and investment bias could easily explain more than 100%

of measured gender wage gap closure, because rising skill prices may have widened the true gender potential wage gap.

Our results are important for understanding the structure of labor demand, which created inequality in the first place, and caused inequality to grow over time. For instance, our results suggest that much 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 for married people aged 25-54 with some college education (results are similar for other education groups), the average woman in 1975 had 25% less earnings potential than the average man, and would fit in the 23rd percentile of the 1975 male wage distribution (see also Blau and Kahn, 2004, Table 1). That percentile lost 10% relative to the male average over the period 1975-95, so according to the single attribute model women should have lost the same percentage relative to men. 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 better than it appeared when it was thought (e.g., by Welch, 2000) that inequality moved in opposite directions within and across genders.

Discrimination is another factor determining labor demand (see Becker, 1957). If discrimination manifests itself in part as a wage premium for men, and had declined significantly over time due to government policy and other factors, then all else equal we expect women's wages to grow faster than men's. In fact, women's wages may have grown less than men's did. Perhaps, as mentioned above, the growing skill premium itself widened the gender wage gap, and that declining discrimination was real but not enough to counteract this trend. Nevertheless, our results show that some combination of the growing skill premium and its interactions with the actions women take to attach themselves to the labor force is primarily responsible for the closure in the measured gender wage gap over the last 30 years.

VI. Appendix I: Common Wage Implications of Selection and Investment

Consider a simple lifetime model in which market-oriented human investments h and the probability of working in the marketplace are simultaneously determined. A person with characteristics X and investment h earns $A(X)h^\alpha e^{\varepsilon_w}$ if she works, and 0 if she does not work. $\alpha \in (0,1)$ is a scalar. ε_w is a normally distributed stochastic wage component, with mean zero and variance σ_w^2 . Her (endogenously determined) probability of working is $p(X)$, so her expected net investment return is $p(X)A(X)h^\alpha e^{\sigma_w^2/2} - h$. The optimal human investment $h^*(X)$ solves

$$h^*(X) \equiv \underset{h}{\operatorname{argmax}} [p(X)A(X)h^\alpha e^{\sigma_w^2/2} - h]$$

VI.A. Wage Concepts

Several wage concepts are relevant. The potential market wage is the wage that a person would have earned *if* (a) she worked in the market place, and (b) she invested with the expectation $p(X) = 1$. The market wage option is the wage she would earn by working in the market place, using her actual human capital $h^*(X)$. The actual wage is zero if she does not work, and equal to the market wage option if she works. Another distinction is an individual's log potential (or market option) wage versus the average log potential (or market option) wage for persons with the same characteristics X . These two differ by the person-specific shock ε_w .

A person with characteristics X has log market wage option w , which is calculated by plugging $h^*(X)$ into the wage function:

$$w = \frac{\ln A(X) + \alpha \ln \alpha + \sigma_w^2/2 + \alpha \ln p(X)}{1 - \alpha} + \varepsilon_w$$

The group average ε_w is zero, so the average log market wage option for the group of persons with characteristics X is just the ratio term above, without the ε_w term.

The average potential log wage $\hat{w}(X)$ in this model is the average log wage that would be earned if all persons with characteristics X invested with the expectation of $p(X) = 1$:

$$\hat{w}(X) = \frac{\ln A(X) + \alpha \ln \alpha + \sigma_w^2/2}{1 - \alpha}$$

which we calculate by maximizing the net investment return with $p(X)$ set to one. Notice the relation between a person's log market wage option w and the average potential log wage $\hat{w}(X)$:

$$w = \hat{w}(X) + \frac{\alpha}{1 - \alpha} \ln p(X) + \varepsilon_w$$

The function $\hat{w}(X)$ tells us about the structure of labor demand because, for example, it tells us how the market wage option of a person with *given* human capital depends on her personal characteristics X .

VI.B. Selection and Investment Simultaneously Determined

Once a person has invested, her stochastic wage component ε_w is realized, and she works if and only if her log market wage option w exceeds her log reservation wage $r = r(X) + \varepsilon_r$, which is also normally distributed. The fraction $p(X)$ of persons with characteristics X who end up working is implicitly defined by:

$$p(X) \equiv \Phi \left(\frac{\ln A(X) + \alpha \ln \alpha + \sigma_w^2/2 + \alpha \ln p(X) - (1 - \alpha)r(X)}{(1 - \alpha)\sigma_z} \right)$$

where σ_z is the standard deviation of $\varepsilon_w - \varepsilon_r$.¹⁹

We can average w among working persons with characteristics X , using the formulas for conditional expectations of normal distributions:

$$E[w|X, \text{ persons working}] = \hat{w}(X) + \frac{\alpha}{1 - \alpha} \ln p(X) + \rho_{wz} \sigma_z \lambda(p(X))$$

where λ is the inverse Mills ratio, and ρ_{wz} is the correlation between $\varepsilon_w - \varepsilon_r$ and ε_w . Notice how this is a

¹⁹The zero function $p(X) = 0$ is always a stable rational expectations equilibrium, because persons expecting not to work invest zero, which causes them to have zero market wage option, which rationalizes their not working. This equilibrium has trivial comparative statics in the sense that marginal parameter shifts leave p unchanged. However, this equilibrium is not very interesting, because a more realistic model would allow for some small but positive earnings with zero investment, or nonpecuniary reasons for investment, etc. The more realistic model would have a stable rational expectations equilibrium near $p = 0$, with nontrivial comparative statics.

For a given X , there may also exist two rational expectations equilibria, one stable and one unstable. The stable one is at higher p , and probably more interesting for our purposes because it has the intuitive comparative statics (e.g., higher r lowers p). One topic for future research is the possibility that marginal parameter changes not only marginally shift the equilibria, but cause some groups of women to shift from one stable equilibrium to another. This possibility might be interesting not only to help explain some of the very rapid occupational shifts in the 1970's (see Rosen, 1992), but also help to separate investment models from selection models.

special case of the formula (2) shown in the text. Measured wages differ from potential wages due to a combination of selection and investment biases, both of which disappear as $p \rightarrow 1$. Also notice that growing wage inequality, which might be modeled as growing α (a.k.a., a growing skill premium) and/or growing σ_w , increase the magnitude of the p gradient in both the investment and selection terms.

VII. Appendix II: Converting Background Selection Bias into Wage Selection Bias

Using the extended Roy model from the text and the formulas for conditional expectation based on normal distributions, we can calculate the average log market wage w and average background indicator b :

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

$$E(b|z > 0) = \hat{b}(X) + \lambda(L(X))\rho_{bz}\sigma_b$$

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

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

The background indicator is measured regardless of whether the person works or not, so \hat{b} can be consistently estimated by averaging b over the entire sample, and $E[b|z>0]$ consistently estimated by averaging b over the subsample of workers. $\rho_{bz}\sigma_b\lambda$ is the selection bias from estimating \hat{b} from the workforce average b , and can be estimated as the difference between the workforce and population averages.

But how is $\rho_{bz}\sigma_b\lambda$ related to the wage selection bias $\rho_{wz}\sigma_w\lambda$? Propositions 1 and 2 give some insight. Here we normalize b so that $\rho_{wb} \geq 0$.

Proposition 1 If the partial correlation (within a randomly selected group of persons with the same X , holding w constant) between b and r were zero, then $\rho_{wz}\sigma_w\lambda$ equals $\rho_{bz}\sigma_b\lambda$ divided by the w coefficient from a regression of b on w . If the partial correlation were positive, then $\rho_{bz}\sigma_b\lambda$ divided by the w coefficient is less than $\rho_{wz}\sigma_w\lambda$.

Proof The partial correlation is nonnegative iff

$$\rho_{br} \geq \rho_{wr}\rho_{wb}$$

Multiplying by σ_r and then subtracting from $\rho_{wb}\sigma_w$ yields

$$\begin{aligned} \rho_{wb}\sigma_w - \sigma_r\rho_{br} &\leq \rho_{wb}\sigma_w - \sigma_r\rho_{wr}\rho_{wb} \\ \rho_{bz}\sigma_z &\leq \rho_{wb}\rho_{wz}\sigma_z \end{aligned}$$

where the second line follows from the definition of z and the additivity of covariance. Collecting terms and multiplying by $\lambda\sigma_b/\sigma_z$ yields:

$$\rho_{bz}\sigma_b\lambda \leq \frac{\rho_{wb}\sigma_b}{\sigma_w}(\rho_{wz}\sigma_w\lambda)$$

The term on the LHS is the b selection bias. The term in parentheses on the RHS is the w selection bias, and the ratio term is the regression coefficient referenced in the proposition. The intuition for the zero partial correlation is very simple: the background characteristic is correlated with labor force status only incidentally through its imperfect correlation with w . Hence, dividing by the reverse regression coefficient both converts from background units to wage units, but also corrects for the imperfect correlation between background and wages. When the partial correlation is positive, a high value of b not only indicates a high w , but also a high r , and hence the zero partial correlation assumption overstates the magnitude of the “effect” of b on labor force status. In this case, dividing by the reverse regression coefficient is not enough.

Proposition 2 further develops this intuition, and relates our calculations to those of Blau and Kahn (1997, 2004).

Proposition 2 If the partial correlation (within a randomly selected group of persons with the same X , holding b constant) between z and w were zero, then $\rho_{wz}\sigma_w\lambda$ equals $\rho_{bz}\sigma_b\lambda$ multiplied by the b coefficient from a regression of w on b . If the partial correlation were positive, then $\rho_{wz}\sigma_w\lambda$ exceeds $\rho_{bz}\sigma_b\lambda$ multiplied by the aforementioned regression coefficient.

Proof The partial correlation is nonnegative iff

$$\rho_{wz} \geq \rho_{bz} \rho_{wb}$$

Multiplying by $\lambda \sigma_w$ yields

$$\rho_{wz} \sigma_w \lambda \geq \frac{\rho_{wb} \sigma_w}{\sigma_b} (\rho_{bz} \sigma_b \lambda)$$

The term on the LHS is the w selection bias. The term in parentheses on the RHS is the b selection bias, and the ratio term is the regression coefficient referenced in Proposition 2. The intuition for the zero partial correlation is very simple: conditional on the background characteristic, wages are uncorrelated with selection. Hence we only need to correct for the fact that samples with different backgrounds may incidentally have different wages – a correction which is achieved with the “forward” regression coefficient.

This leaves two practical questions. First of all, how can the regression coefficient be estimated, given that it is supposed to come from a randomly selected sample of women with characteristics X ? Our approach, where “background” is measured by own schooling, parental schooling, or own IQ, is to estimate the regression in a sample of men with otherwise similar characteristics. Second, what is the sign of the partial correlation between background and reservation wages? For our measures, it seems that the correlation must be positive, because schooling, IQ, etc., not only enhance market productivity, but enhance nonmarket productivity too. Furthermore, since we believe that our background measures enhance market productivity, a positive partial correlation is needed to explain why, in some cases, the raw correlation between background and employment is weak or even negative.

VIII. References

- Autor, David H. and Mark G. Duggan. “The Rise in the Disability Rolls and the Decline in Unemployment.” *Quarterly Journal of Economics*. 118(1), February 2003: 157-205.
- Becker, Gary S. *The Economics of Discrimination*. Chicago: University of Chicago Press, 1957.
- 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.
- Blau, Francine D. and Lawrence M. Kahn. "The US Gender Pay Gap in the 1990s: Slowing Convergence." NBER Working Paper no. 10853, October 2004.
- Card, David and John E. DiNardo. "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics*. 20(4), October 2002: 733-83.
- Chamberlain, Gary. "Asymptotic Efficiency in Semiparametric Models with Censoring." *Journal of Econometrics*. 32(2), July 1986: 189-218.
- Fortin, Nicole M. and Thomas Lemieux. "Are Women's Wage Gains Men's Losses? A Distributional Test." *American Economic Review*. 90(2), May 2000: 456-60.
- 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.
- Gunderson, Morley. "Male-Female Wage Differentials and Policy Responses." *Journal of Economic Literature*. 27(1), March 1989: 46-72.
- 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 Edward Vytlacil. "Econometric Evaluation of Social Programs." in James J. Heckman and Edward E. Leamer, eds. *Handbook of Econometrics*. Volume 6. Amsterdam: Elsevier, 2005.
- 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.
- Mulligan, Casey B. and Yona Rubinstein. "The Closing Gender Gap as a Roy Model Illusion." NBER working paper #10892, November 2004.
- Murphy, Kevin M. and Robert H. Topel. "Unemployment and Nonemployment." *American Economic Review*. 87(2), May 1997: 295-300.
- Murphy, Kevin M. and Finis Welch. "Relative Wages 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.
- Newey, Whitney K., James L. Powell, and James R. Walker. "Semiparametric Estimation of Selection Models: Some Empirical Results." *American Economic Review*. 80(2), May 1990: 324-328.
- O'Neill, June. "The Trend in the Male-Female Wage Gap in the United States." *Journal of Labor Economics*. 3(1), Part 2, January 1985: S91-S116.
- Rosen, Sherwin. "The Market for Lawyers." *Journal of Law and Economics*. 35(2), October 1992:

215-246.

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.

Welch, Finis. "Growth in Women's Relative Wages and in Inequality among Men: One Phenomenon or Two?" *American Economic Review*. 90(2), May 2000: 444-49.

Welch, Finis. "Wages and Participation." *Journal of Labor Economics*. 15(1), January 1997: S77-103.