RETURNS TO COMPUTER SKILLS AND BLACK-WHITE WAGE DIFFERENTIALS

Barton Hughes Hamilton

John M. Olin School of Business
Washington University in St. Louis
Campus Box 1133
One Brookings Drive
St. Louis, MO 63130

e-mail: HAMILTONB@WUOLIN.WUSTL.EDU
Phone: (314) 935-8057. Fax: (314) 935-6359

March, 1997

I would like to thank Franque Grimard, Vivian Hamilton, Robert Pollack, and seminar participants at Washington University for helpful comments, and Raymond Yuen for expert research assistance.

RETURNS TO COMPUTER SKILLS AND BLACK-WHITE WAGE DIFFERENTIALS

ABSTRACT

We construct alternative measures of computer skill based on familiarity with software packages and programming languages using High School and Beyond data. Utilizing an endogenous switching model, we find a computer wage premium of 13% - 25%, depending on the skill measure. Conditional on quantitative test scores, young black males are significantly more likely than whites to acquire computer skills. The wage premium associated with computer skills is more than 50% larger for blacks than it is for whites. No significant black-white pay gap is found among the computer skilled, while blacks earn significantly less than whites among those without such skills.

JEL Classifications:

J31 Wage Level and Structure; Wage Differentials by Skill, Training, Occupation, etc.

J71 Discrimination.

RETURNS TO COMPUTER SKILLS AND BLACK-WHITE WAGE DIFFERENTIALS

I. INTRODUCTION

A number of recent studies have documented a substantial increase in wage inequality both between and within skill groups over the past 15 years.[1] One of the leading explanations for these trends is technological change, which is thought to have increased the productivity, and hence the wages, of more skilled workers.[2] Some argue that much of this technological change results from a rapid increase in the utilization of computers in the production process.[3] If computer skills increase productivity, individuals investing in such skills should be expected to earn a wage premium.

The increased diffusion of computer use in the 1980s coincided with a rise in the black-white wage gap among young workers (Bound and Freeman (1992)). Lack of investment in computer skills or lower returns to these skills for blacks may play a role in explaining the trend toward increased racial earnings inequality. For example, Boozer et al (1992) find that 28% of whites and 20% of blacks used a computer at work in 1984. By 1989, the fraction had increased to 42% for whites, but only 28.7% for blacks.

While technological change associated with the rapid spread of computers in the workplace has been pointed to as a leading cause for increasing earnings inequality, and perhaps a factor in the widening racial earnings gap, little direct evidence is available concerning the impact of computer skills on wages or on the black-white pay differential. Krueger (1993) investigates the relationship between computer use on the job and wages, and finds a premium of 10% - 15% for computer users. However, the estimated wage gap may in part arise from the unobserved characteristics of firms that are likely to employ computer technology in their workplaces. For instance, Reilly (1995) finds a

[1]See, for example, Bound and Johnson (1992), Juhn, Murphy, and Pierce (1993), Katz and Murphy (1992), and Bound and Freeman (1992).

[2]See, for example, Bound and Johnson (1992), Krueger (1993), and Mincer (1989). Levy and Murnane (1992) provide an excellent summary of both the trends in wage inequality and the explanations proposed in the literature for these phenomena.

[3]Krueger (1993) finds that the fraction of individuals using computers at work increased from .25 to .37 between 1984 and 1989.

positive correlation between computer use and establishment size. The estimated premium may also in part result from unobserved differences in worker abilities. Individuals acquiring computer skills may also have greater quantitative abilities, which will lead to wage differentials if these abilities are rewarded in the labor market. Thus, the premium estimated in the literature potentially reflects the returns to computer skills, as well as firm heterogeneity and worker self-selection.

This paper seeks to build upon the results in the literature in three ways. First, we focus on estimating the return to computer skills, which may differ to some degree from the premium associated with computer use on the job, since the latter may capture firm characteristics. In addition, some might argue that individuals with computer skills will earn a premium even if they do not use a computer at work. Computer skills may signal that a worker possesses some unobserved characteristic, such as logical reasoning ability, that is valued by employers (Spence (1973)). Second, we explicitly examine the role of computer skills in explaining earnings differences between young black and white men. Third, unlike previous studies, we estimate the return to computer skill accounting for worker self-selection. The analytical framework is based on Roy's (1951) model of occupational choice and self-selection in the labor market, which has been applied in a number of contexts, most notably (for this paper) in examining the relationships between the choice to acquire additional years of schooling and earnings (Willis and Rosen (1979)) and between training and wages (Heckman and Robb (1985)). In the Roy model, individuals choose whether to acquire computer skill based in part on their relative abilities at computer and non-computer related tasks. Consequently, the observed premium associated with computer skill may over or understate the premium for a randomly selected worker from the population, depending on whether positive or negative selection is observed among individuals with computer skills (denoted by C) and those without computer skills (NC).

Data from the High School and Beyond survey on young (1980 high school graduates) male workers is used to construct alternative measures of computer skill based on the individual's knowledge of computer software packages and programming languages. These measures may be more appropriate when evaluating the implications of the results for policy purposes; government training programs may be unable to place individuals in firms using computers, but they are able to provide instruction on the use of software packages and programming languages. We find an estimated OLS computer wage premium of 4% - 18% for these young workers, depending on the skill measure used, with individuals who are able to program in an advanced computer language earning the largest premium. For workers with knowledge of computer software packages, the premium increases to 25% after accounting for non-random sectoral selection using a robust two-step sample selection correction procedure. Surprisingly, the computer wage premium is substantially larger for blacks than for whites. In particular, blacks with computer skills earn *more* than whites, all else equal. On the other hand, among workers without computer skills, blacks earn significantly less. Acquisition of computer skill may thus play an important role in explaining the widening black-white wage gap.

The paper proceeds as follows: Section II outlines the empirical framework employed to measure the wage differential associated with computer skills. Section III describes the High School and Beyond data and the construction of the four alternative measures of computer skill. Section IV presents the earnings equations estimates and the predicted wage differentials. Concluding remarks are found in Section V.

II. THE MODEL

The standard approach in the literature for evaluating the earnings premium associated with computer use is to estimate an OLS regression of the form:

$$(1)\quad \ln W_i = X_i\beta + C_i\alpha + \varepsilon_i,$$

where W_i is individual i's hourly wage rate, X_i is a vector of individual characteristics with associated parameter vector β, C_i is a measure of computer use, and ε_i represents other unobserved (to the econometrician) factors influencing earnings. The coefficient α measures the computer premium in this framework, with estimates ranging from .10 to .15 (Krueger (1993)). The estimates of α will be biased if unobserved factors influencing earnings are correlated with the measure of computer use (i.e., cov(C_i, ε_i) does not equal zero). For example, individuals with high quantitative skills may earn more and also be more likely to use computers. Consequently, in this example, α may reflect the returns to both computer skills and unobserved (to the econometrician) quantitative ability. In addition, since the measure for C_i is typically computer use on the job, unobserved firm characteristics which influence both wages and the probability of using a computer may bias the estimate of α.

In order to account for the possible correlation between the unobservables affecting wages and those influencing the acquisition of computer skill, we implement a version of Roy's (1951) model of self-selection. The application is similar to other studies examining the impact of human capital on earnings in the presence of self-selection, such as Willis and Rosen (1979), who examine the college-high school earnings premium accounting for non-random sorting into educational status, and to studies examining the impact of training on wages using a control function approach (e.g., Heckman and Robb (1985)). We presume that there are jobs in the economy for which computer ability enhances an individual's productivity, and those for which it does not. In addition, employers may also choose to pay workers with computer skills more, even if they do not use a computer on the job, if computer skills are signals of other abilities which are valued by the firm, such as logical reasoning.[4] Individuals compare the income stream associated with the acquisition of computer skill, W_C, with that available without such skills, W_{NC}, accounting for the monetary and psychic costs associated with the acquisition of skills. We generalize equation (1) and let the log wage of individual i in skill category j be given by:

$$(2)\quad \ln W_{ij} = X_i \beta_j + \varepsilon_{ij}, \quad j = C, NC.$$

[4]The argument here is analogous to Spence's (1973) signaling model in which education does not enhance productivity, but rather reveals the inherent productivity of workers. It might be possible to distinguish between the human capital and signaling arguments for the effects of computer skills on wages if data were available on both the level of computer skill and whether the worker used a computer on the current job. Unfortunately, the latter piece of information is unavailable in the High School and Beyond data.

We assume that the acquisition of computer skills involves a cost M_i which varies across individuals. These costs may be monetary, such as buying or paying for access to a computer, paying for computer courses, and so forth. There may also be psychic costs in the sense that individuals who are "afraid" of computers may incur substantial disutility if they are forced to use one. If these costs are a linear function of an observed set of characteristics X_{iM} and an unobserved component ε_{iM}, then

$$(3)\quad M_i = X_{iM}\beta_M + \varepsilon_{iM}.$$

The decision of whether to invest in computer skills is given by the sign of the index function:

$$(4)\quad I_i^* = lnW_{iC} - lnW_{iNC} - M_i = X_i(\beta_C - \beta_{NC}) - X_{iM}\beta_M + \varepsilon_{iC} - \varepsilon_{iNC} - \varepsilon_{iM} = Z_i\gamma + \nu_i,$$

where the Z_i vector contains the observed characteristics presumed to influence earnings and costs, and $\nu_i = \varepsilon_{iC} - \varepsilon_{iNC} - \varepsilon_{iM}$ is a composite error term. If the latent index I_i^* is greater than 0, the individual chooses to invest in computer skills and the indicator variable $I_i = 1$. $I_i = 0$ otherwise.

As is clear from equations (2) and (4), OLS estimation of the wage equations results in biased parameter estimates if the unobserved factors influencing skill choice, ν_i, are correlated with the unobserved influences on earnings in sector j, ε_{ij}. The standard approach in the literature is to introduce a nonlinear selection correction term, $\lambda_j(Z_i\gamma)$, in equation (2) to account for the possible non-zero expected value of ε_{ij}:

$$(5)\quad lnW_{ij} = X_{ij}\beta_j + \lambda_j(Z_i\gamma) + \eta_{ij},\quad j = C, NC,$$

where the error term in equation (2) has been replaced by its conditional expectation plus a mean zero error term η_{ij}.

Studies in the literature typically assume joint normality of (ε_j, ν) and set $\lambda_j(Z_i\gamma)$ equal to the inverse Mills ratio. However, this approach has been criticized because of the potential sensitivity of the parameter estimates to this distributional assumption. Consequently, we adopt the semi-parametric approach of Lee (1982) and Newey (1988) and approximate $?_j(Z_i?)$ by the series:

$$(6)\quad ?_j(Z_i?) \approx \sum_{k=1}^{K} d_{jk} g_{jk}(Z_i?)$$

where the d_{jk} are unknown coefficients and that the $g_{jk}(.)$ are known basis functions. We follow Lee (1982) and assume that K = 3, and the $g_{jk}(.)$ are functions of the inverse mills ratio, $m_j(Z_i?)$:[5]

$$(7)\quad \begin{aligned} g_{j1} &= m_j(Z_i?) \\ g_{j2} &= (Z_i?) \cdot m_j(Z_i?) \\ g_{j3} &= (1 - (Z_i?)^2) \cdot m_j(Z_i?) \end{aligned}$$

Our approach is to estimate equation (4) via probit in the first step, use the estimated parameters to construct the basis functions in equations (6) and (7), and then estimate the sectoral wage equations (5) by OLS after including the selection correction terms. Evidence of non-random sorting may be found by examining the joint significance of the $g_{jk}(.)$ functions. This specification nests bivariate normality as a special case, which may be tested by examining the joint significance of $g_{j2}(.)$ and $g_{j3}(.)$.

III. DATA

The choice of the data set for this study is motivated by the fact that we require measures of computer skill or computer human capital rather than indicators of whether the individual uses a computer at work, which may in part reflect the characteristics of the firm employing the worker. Information on computer skills might specify the quantity of these skills, such as the number of programming languages or software packages the individual is familiar with, and the quality of the individual's skill, i.e., how good a programmer the worker is. However, as in most studies of the

[5] Lee (1982) derives the specification for the $g_{jk}(.)$ functions from bivariate Edgeworth expansions of the joint error distributions.

impact of human capital on earnings, we focus on the quantity rather than quality dimension of computer skill due to data limitations.

The data employed in the empirical analysis is drawn from the 1980 senior cohort of the High School and Beyond Survey (HSB), which is a sample of high school seniors who are initially interviewed in 1980, and re-interviewed in 1982, 1984, and 1986.[6] The focus of the analysis is the effect of computer skills on the earnings of full-time young workers after leaving high school or university. Consequently, we use data from the 1986 survey to construct the hourly wage rate for the most recent full-time (greater than 20 hours) job held by the individual (after leaving school) as of the 1986 interview.[7] A complete description of the data set is provided in Appendix A. Females are excluded so that issues related to labor force participation are less important.

One of the main advantages of the HSB is that a number of measures associated with computer skill are reported in the survey. In particular, respondents are asked whether they have ever used each of the following types of software: word-processing; spreadsheet; database; statistical; educational; other. In addition, the survey asks whether the respondent has ever written a computer program in the following languages: BASIC; Fortran; Pascal; COBOL; PL/I; APL; SQL; assembly; other. The responses to these questions are used to construct two alternative measures of an individual’s computer human capital. The first measure is based on the individual's knowledge of various software packages. The variable SOFT is defined to equal one if the individual has ever used any type of computer software package, such as a word-processor, spreadsheet, or database package.

Computer programming ability is likely to be a good proxy for computer literacy. The second measure developed in this paper is thus based on the responses to the questions concerning whether the individual had ever written a computer program. The variable PROG equals one if the individual

---

[6]Black and Hispanic public and private schools were over-sampled in the survey. We use the HSB survey weights in the subsequent empirical analysis.

[7]Individuals working less than 20 hours per week on all jobs between 1984 and 1986 are dropped from the analysis since we wish to focus on the effect of computer skills on full time workers, and because it is difficult in some cases to determine whether individuals working less than 20 hours per week are still in school.

has ever programmed in an advanced computer language (a language other than BASIC), and zero otherwise.[8] Individuals are required to write a rudimentary program in BASIC as part of many high school (and college) computer courses, and knowledge of only this language may overstate the level of computer literacy. Consequently, PROG is likely to be more indicative of advanced computer skill, and thus individuals in this category may be expected to earn the largest returns.[9]

Panel A of Table 1 indicates the fraction of workers in the sample with computer skills as defined by the two measures described above, as well as the associated computer skill earnings premium, both overall and broken down by race. The first row shows that 31% of sample members have used a software package, while the third row indicates that a slightly lower fraction (28%) has written a computer program. The earnings premium associated with computer skill varies from 10% to 14%, depending on the measure, and is statistically significant. The racial breakdown in the remaining columns of Panel A indicates that whites are more likely to possess computer skills than other groups. However, the earnings premium is substantially higher for non-whites, particularly so for blacks. In addition, the computer wage premium is greater when the PROG definition is used, which we conjecture to be the measure reflecting the most advanced level of computer skill.

Panel B shows that at least some of the computer wage premium is likely to reflect differences in the characteristics of individuals with and without computer skills. For example, a substantially higher fraction of individuals with some knowledge of software programs or advanced programming languages are college graduates, and have higher scores on the standardized tests in mathematics, reading, and vocabulary administered to the individuals when they were high school seniors in 1980 (reading and vocabulary scores were averaged together to create the verbal score).[10] Individuals with

---

[8]Appendix A reports the actual questions asked of individuals regarding their knowledge of various software packages and programming languages. The knowledge of software package question does not refer to programming language, so there may be some individuals for whom SOFT = 0 and PROG = 1.

[9]The results are similar, though smaller in magnitude, when BASIC is included in the PROG measure.

[10]The results are unchanged when we include the reading and vocabulary scores separately in the regressions.

computer skills are also more likely to find math interesting or useful (sample members were queried while in high school). The final three rows of the panel indicate that individuals with greater exposure to technology when growing up, as measured by whether the family owned a pocket calculator, and when in high school, as measured by whether the student used a micro-computer or computer terminal, are much more likely to have computer skills than individuals without such exposure. The summary statistics presented in Panel A of Table 1 thus show a significant hourly wage differential between computer users and non-users, which Panel B suggests may in part reflect the higher ability of the former group.

Panel C of Table 1 breaks down the summary statistics by race, and shows four potential reasons why whites are more likely to possess computer skills than blacks or hispanics and asians. First, young whites have greater educational attainment in the sample, and have more work experience. Second, whites appear to have had greater exposure to technology at home, in the form of a pocket calculator, than did the other groups. Attitudes toward math and exposure to computers in school are fairly similar among blacks and whites. Finally, whites have quantitative and verbal test scores which are substantially higher than those of nonwhites. Neal and Johnson (1996) argue that the black-white wage gap among young workers reflects in large part differences in premarket skills, as measured by test scores. We investigate the extent to which these premarket differences in abilities explain racial differences in the diffusion of computer skills below.

IV. EMPIRICAL RESULTS

The first step in the empirical analysis is to estimate regressions similar to those found in the literature given by equation (1). Column (1) of Panels A and B of Table 2 presents estimates of the computer skill differentials by race using the SOFT and PROG definitions, respectively, without any other controls. Column (1) in each panel shows that computer users of each racial group earn significantly more than do individuals without computer skills. In addition, the estimates in column (1) also indicate substantial differences in the racial pay gap across computer skill categories. The first column of Panel A shows that blacks without software knowledge earn 13% less than whites without software knowledge. However, among those individuals with computer skill, the black-white pay gap is only 6.2% and is not statistically significant. Panel B shows similar variations in the racial pay gap when knowledge of advanced programming languages is used as the measure of computer skill.

Since individuals with computer skills appear to have greater levels of pre-market ability, as measured by quantitative and verbal test scores, column (2) of each panel includes the test score variables in the regressions to examine the extent to which these factors explain the computer pay premium. While the computer skill premium falls somewhat, it is still positive and strongly significant for all racial groups. The black-white pay gap also narrows for each skill group, which is consistent with Neal and Johnson's (1996) argument that much of the black-white wage differential reflects differences in pre-market abilities across races. However, it is still the case that the racial pay gap is much smaller (and not statistically significant) among those with computer skills than among those without them.

Columns (3) and (4) of the panels in Table 2 examines the extent to which the computer skill premium reflects demographic and job variables, and family background, as measured by the maximum educational attainment of the parents. In this case, Panel A shows that the return to software knowledge for whites declines to approximately 4% and is no longer statistically significant. However, the premium for blacks and hispanics and asians remains substantial at 9%-13%, and is strongly statistically significant for blacks. Blacks still earn significantly less than whites among NC workers, but the columns imply that young black males actually earn approximately 2% more than whites among those with computer skills, although the difference is not significant. Panel B shows that the computer skill premium remains large (8%-9%) and statistically significant for whites as well as non-whites for this measure of more advanced computer skill.

It may be argued that the wage premium associated with computer skill reflects in part the sorting of computer skilled workers into high wage industries. To examine this hypothesis, column (5) of Panels A and B includes industry dummies in the regressions. The results remain unchanged:

workers with computer skills earn a wage premium, particularly those with programming skills in advanced languages; the premium is especially large for blacks; and blacks with computer skills actually earn more than whites, all else equal, while the typical negative black-white pay gap persists for NC workers. With regard to the magnitude of the computer skill premium described above, our estimates are smaller than those found by Krueger (1993) in the Current Population Survey for individuals using computers at work, which may reflect the inclusion of older workers in that sample.[11] Krueger's estimates for the HSB survey (for ever having used a computer at work) are slightly higher than those presented here.

IV.1 ESTIMATES OF THE COMPUTER PREMIUM ACCOUNTING FOR SELECTION

We now estimate the computer skill premium accounting for unobserved factors which may be correlated with the choice of both skill and earnings. We first present the results from the estimation of the probit model described in equation (4). Included in the vector of independent variables Z_i are the exogenous variables from the wage regressions, X_i, as well as a set of variables hypothesized to be correlated with an individual's cost of acquiring computer skill (but which do not affect hourly wages). Exposure to technology while growing up, as measured by whether the family owned a pocket calculator, is likely to influence the decision. Such individuals may have less "computer fright" and hence lower psychic costs when acquiring computer skill. In addition, individuals who used a micro-computer or computer terminal while in high school may have found it less costly to learn to use a computer. Finally, we hypothesize that sample members who believed that math was interesting or useful when queried in high school are likely to have greater preferences toward acquiring computer skills and using computer technology.

The coefficient estimates from the probit regression are shown in Table 3. Columns (1) and (2) present the results for the SOFT skill measure, while the third and fourth columns present the PROG estimates. In order to determine the extent to which variations in the probabilities of acquiring

[11] Krueger's estimates are also not broken down by race.

computer skill across races reflects pre-market abilities, the odd numbered columns exclude the test score variables. Columns (1) and (3) indicate that, as might be expected, more educated workers are more likely to invest in computer skills. Individuals attending college may have had more exposure to computers through their course work and thus find the investment less costly. Individuals with more work experience are less likely to acquire programming skills, perhaps because they have had less time to obtain formal training in this area. For the variables assumed to affect investment in computer skill but not wages, preferences towards and interest in math, as well as exposure to technology in the home or use of computers in high school have a significant positive impact on the probability of acquiring computer skills. Finally, after controlling for these other factors, little difference is found in the relative propensities of blacks to acquire computer skills.

In columns (2) and (4) the probit model is re-estimated including the test score variables. Not surprisingly, quantitative ability has a strong positive impact on the probability of acquiring computer skills, while the verbal test score variable has a smaller effect and is only significant in the PROG regression. The most notable impact of the inclusion of the pre-market ability variables is found for the propensity of blacks to acquire computer skills. For both measures, blacks are significantly more likely than whites to possess such skills, all else equal. This is consistent with our finding in Table 2 that the computer wage premium is substantially higher for blacks than for whites, and that there is no racial pay gap among young workers with computer skills.

The probit estimates are used to construct the selection correction variables described by equations (6) and (7), which are then substituted into equation (5) to correct for the potential impact of non-random worker sorting on the parameter estimates of the wage equations. We estimate sectoral wage regressions for each measure of computer skill developed in this paper.

The first column of Table 4 provides some evidence of non-random sorting among the computer skilled, using the SOFT measure. In this case, the estimates imply that the SOFT wage premium found in Table 2 may understate the earnings gain associated with the acquisition of computer skills for a randomly selected worker. The estimates of $g_{C2}(.)$ And $g_{C3}(.)$ also imply that the assumption of bivariate normality in the SOFT sample selection model may be unduly restrictive. On the other hand, the PROG earnings regressions show little evidence of non-random sorting. In this case, the selection correction terms are individually and jointly insignificant.

For each of the skill measures, the hypothesis that the earnings equations differ only by an intercept can be rejected at the 1% level. Consequently, the specification given in equation (1) is unduly restrictive. Perhaps the most notable difference in the earnings structures implied by coefficient estimates from the computer and non-computer regressions is the effect for blacks. In both sets of regressions, young black males without computer skills earn approximately 7% less than whites, the difference being statistically significant at the .06 level. However, investment in computer skills virtually eliminates the black-white pay gap. In fact, blacks with computer skills earn 6% more than their white counterparts, although the difference is not statistically significant. In addition, there also appears to be an earnings premium for Hispanics and Asians relative to whites among the computer skilled.

Among the other coefficient estimates, at first glance it is somewhat surprising that the educational attainment variables often have an insignificant, and in some cases negative, effect on hourly earnings among both computer users and non-users. Recall that the individuals in the sample are 23 to 25 year-olds who graduated from high school in 1980. The college dropouts and graduates are thus generally in their first couple of years on the job and are likely to be making significant human capital investments, as opposed to the high school graduates who have been in the labor force for five to six years. Consequently, the insignificant or negative coefficients on the educational attainment dummies likely reflect the fact that these individuals are at the beginning of their lifetime earnings profiles.[12] In the case of the PROG measure, college graduates with computer skills earn a significant wage premium relative to high school graduates, which is not found among those in the

[12]Presumably, the months worked variable should pick up some of the difference in labor market experience among the educational groups. However, the variable includes months of part-time work while the worker was in school.

NC category. Further investigation indicates that 18% of college graduates with advanced programming skills enter occupations listed as computer programming and electrical engineering. There appears to be some type of certification effect, since few high school graduates (6%) with programming skills are in those occupations. Consequently, the college graduate coefficient may partly reflect the wage premium earned by workers in computer programming and electrical engineering occupations.

The test score variables show that quantitative ability has a positive impact on earnings for all workers, although the magnitude is larger for those with computer skills. This may also reflect an occupational effect. On the other hand, verbal ability does not significantly affect earnings in any sector. When these variables are removed from the wage (and probit) regressions, the coefficient estimates on the black indicator variable drop to zero in the SOFT = 1 and PROG = 1 regressions, while the wage gap widens to 8.5% among those without computer skills.

IV.2 PREDICTED COMPUTER SKILL WAGE DIFFERENTIALS BY WORKER TYPE

In order to illustrate more fully the impact of the different parameter estimates, the selection-corrected coefficients reported in Tables 4A and 4B are used to construct the predicted computer skill premium for a variety of worker types. For a given set of worker characteristics X, the unconditional wage differential associated with computer skill, denoted by dW, may be calculated using the formula

$$(8)\ dW = \frac{E(W_C|X)}{E(W_{NC}|X)} - 1 = \frac{\exp(X\hat{\beta}_C + \frac{1}{2}\sigma_C^2)}{\exp(X\hat{\beta}_{NC} + \frac{1}{2}\sigma_{NC}^2)} - 1,$$

where σ_j^2 is the untruncated residual variance for the sector j wage equation. Calculation of equation (8) thus provides the expected wage differential for a randomly selected worker from the population with characteristics X.

Table 5 presents predicted percentage wage differentials (dW*100%) constructed using equation (8) for various demographic groups. The individual characteristics used to construct the differentials are the mean values of the independent variables for all workers in the sample. The first row of the table indicates that a randomly selected worker from the population earns a premium of roughly 25% if he has computer skills as given by the SOFT measure, and approximately 13% using the PROG definition of skill. The increase in the SOFT premium reported in Table 5 relative to that shown in Panel A of Table 2 appears to reflect selection effects.

Since blacks are unconditionally (on observed characteristics) less likely to acquire computer knowledge, the large computer pay premiums found in the first row suggest that a portion of the increase in the black-white wage gap observed during the 1980s (Bound and Freeman (1992)) reflects a lack of diffusion of computer skills among blacks. However, the second and third rows of the table indicate that those young black males who do acquire computer skills earn extremely large premiums. For each skill measure, the wage premium associated with computer knowledge for blacks is more than 50% higher than the premium for whites, due to the lack of a racial earnings penalty among computer users. Hispanics and Asians also earn a higher premium than whites. Consequently, for young nonwhite males with computer skills, racial pay gaps have virtually been eliminated.

The return to computer skills are greater for college graduates than for high school graduates with advanced programming skills, perhaps reflecting their entry into engineering occupations. High school graduates typically are unable to obtain these types of position, and hence advanced programming skills may be relatively less valuable for them. On the other hand, knowledge of word-processing, spreadsheet, and database management software may allow high school graduates to obtain higher paying clerical and administrative positions, which may explain the relatively large computer pay premium observed for these individuals.

The seventh and eighth rows of Table 5 indicate that the computer premium falls with labor market experience using each of the skill measures. The drop is fairly substantial, about one-third, for each measure. One interpretation of this finding is that workers without computer skills take jobs with greater training opportunities, and thus lower starting pay, in order to make up for their shortfall in computer literacy. Computer skills may give workers a head start in the labor market, allowing them to take jobs requiring less on the job training. Finally, the bottom row of the table indicates that the substantial computer wage premiums are being earned outside of the manufacturing sector. This is consistent with the finding of Berman, Bound, and Griliches (1994) that the extent of computerization has been greater in the services than in manufacturing, suggesting that the demand for computer skills may be higher outside of the manufacturing sector.

V. CONCLUSION

A number of studies have argued that the increasing wage inequality observed during the 1980s reflects in part the impact of skill-biased technological change and the consequent increase in demand for workers with these skills. This paper presents empirical estimates of the wage premium associated with computer skill, using two alternative measures of computer knowledge. The results are similar across computer skill measures. Simple OLS estimates imply that individuals investing in computer skills earn 4%-18% more than the unskilled among young black, white, hispanic, and asian workers. After accounting for non-random worker selection into skill categories using a robust sample selection correction procedure, we find that these estimates may understate the premium earned by workers acquiring computer software knowledge. After controlling for sample selection, the wage equation estimates imply a computer wage premium on the order of 25% for this measure. Little evidence of selection bias is found among those with knowledge of advanced programming languages.

These results provide useful insights into a variety of issues surrounding the debate on the explanations for increasing earnings inequality. In particular, college graduates are more likely to acquire computer skills than the less educated, and, in the case of the measure of advanced programming languages, experience greater returns to these skills. To the extent that demand for these skills is increasing in the economy, the findings provide some explanation for the increasing between-group inequality, although it is difficult to make such inferences from cross-sectional data. It should be emphasized that these findings refer to new entrants to the labor market during the mid-1980s, and it is difficult to generalize to older workers. Examination of older cohorts of workers is clearly necessary to determine whether the computer skill premium for more experienced individuals is similar to that for the new labor market entrants examined in this paper.

The most surprising result in the paper is that blacks with computer skills suffer no significant wage penalty relative to whites, primarily because the pay premium associated with such skills is more than 50% higher for blacks than for whites. However, the typical racial wage gap found in other studies (e.g., Bound and Freeman (1992)) is present among those without skills. Reflecting the substantial earnings premium available to blacks with computer skills, we find that blacks are more likely than whites to acquire computer skills, all else equal. However, like Neal and Johnson (1996), who find that differences in pre-market skills explain much of the black-white wage gap, blacks have lower educational attainment, lower math scores, and less exposure to technology at home, characteristics which increase the propensity to acquire computer skill. For example, when the test score variables are removed from the probit regression, there is no significant racial difference in the probability of acquiring computer skills. The HSB data also show that, unconditionally, blacks are less likely to possess such skills. Boozer et al's (1992) finding that the rate of increase in computer use at work between 1984 and 1989 was smaller for blacks than whites implies that the widening black/white wage gap in the 1980s may in part reflect the relative lack of diffusion of computer skills among blacks. More research utilizing data over a longer period is needed to accurately assess this explanation.

A potential implication of the findings in this paper is that one way of improving the economic status of young blacks is for government sponsored training programs to emphasize the teaching of computer skills, since these appear to be particularly valuable to black males. Of course, the diffusion of computer skills across the population may erode the fairly large returns estimated here. In addition, as computer use becomes more widespread, any potential signaling value of computer skills to employers is likely to be less valuable. However, Krueger (1993) finds little decline in the wage premium associated with computer use at work between 1984 and 1989, a period which also saw a rapid expansion in computer use. Consequently, government programs emphasizing training in computer skills are likely to remain a good investment in the near future.

TABLE 1
ESTIMATED COMPUTER WAGE PREMIUMS AND SUMMARY STATISTICS

PANEL A: COMPUTER WAGE PREMIUM, BY RACIAL GROUP AND SKILL MEASURE

Measure	Group: All	White	Black	Hispanic & Asian
Fraction SOFT = 1	.31	.33	.28	.24
SOFT % Wage premium (t-stat)	10.3% (5.920)	9.1% (3.483)	16.8% (2.738)	13.9% (3.644)
Fraction PROG = 1	.28	.30	.24	.21
PROG % Wage premium (t-stat)	14.0% (8.104)	12.7% (5.281)	19.5% (3.269)	19.0% (4.890)

PANEL B: SUMMARY STATISTICS, BY SKILL MEASURE

Variable	SOFT = 1	SOFT = 0	PROG = 1	PROG = 0
College Dropout	.18	.18	.14	.19
College Graduate	.42	.13	.46	.13
Black	.07	.09	.07	.09
Hispanic or Asian	.09	.13	.08	.13
Months Worked Since 1980	48.2	52.2	47.0	52.5
Math Score	58.2	51.2	60.0	50.8
Verbal Score	56.1	51.3	57.8	50.8
Finds Math Interesting	.54	.41	.57	.40
Thinks Math is Useful	.77	.63	.79	.63
Family Owned Pocket Calculator	.95	.84	.96	.84
Used Micro-Computer in High School	.19	.04	.18	.05
Used Computer Terminal in High School	.24	.10	.24	.10
Number of Observations	733	1655	668	1720

TABLE 1 (continued)
ESTIMATED COMPUTER WAGE PREMIUMS AND SUMMARY STATISTICS

PANEL C: SUMMARY STATISTICS BY RACE

Variable	Whites	Blacks	Hispanics and Asians
College Dropout	.17	.16	.25
College Graduate	.25	.12	.12
Months Worked Since 1980	51.7	46.4	48.7
Math Score	54.7	46.4	48.7
Verbal Score	54.0	47.3	48.5
Finds Math Interesting	.44	.45	.47
Finds Math Useful	.68	.64	.62
Family Owned Pocket Calculator	.91	.65	.81
Used Micro-Computer in High School	.09	.08	.05
Used Computer Terminal in High School	.15	.13	.09
log(Hourly Wage)	1.995	1.880	1.981

Note: Summary statistics calculated using HSB sampling weights.

TABLE 2
OLS ESTIMATES OF RACIAL WAGE DIFFERENTIALS AND COMPUTER WAGE PREMIUM

Panel A: SOFT Measure of Computer Skill

Variable	Specification				
	(1)	(2)	(3)	(4)	(5)
Computer Skill	0.087 (3.019)	0.061 (1.993)	0.042 (1.400)	0.037 (1.244)	0.035 (1.198)
Black	-0.130 (-4.468)	-0.106 (-3.413)	-0.071 (-2.213)	-0.064 (-1.983)	-0.072 (-2.262)
Hispanic or Asian	-0.017 (-0.484)	-0.004 (-0.078)	0.010 (0.279)	0.019 (0.534)	0.019 (0.540)
Black*Computer Skill	0.068 (1.236)	0.062 (1.505)	0.088 (1.646)	0.088 (1.631)	0.100 (1.835)
Hisp or Asian*Computer Skill	0.043 (0.613)	0.047 (0.674)	0.059 (0.820)	0.055 (0.784)	0.055 (0.779)
Test Scores	No	Yes	Yes	Yes	Yes
Demog. and Job Vars	No	No	Yes	Yes	Yes
Parental Education	No	No	No	Yes	Yes
5 Industry Dummies	No	No	No	No	Yes
P-Values from Test of:					
Comp. + Black*Comp. = 0	< .001	.002	.004	.007	.002
Comp. + Hisp.*Comp. = 0	.042	.093	.132	.159	.145
Black + Black*Comp. = 0	.184	.600	.726	.624	.652
Hisp. + Hisp.*Comp. = 0	.667	.481	.288	.240	.288
Adjusted R^2	0.016	0.025	0.052	0.065	0.133

Note: t-statistics in parantheses. Standard errors adjusted for arbitrary forms of heteroscedasticity. All regressions based on 2388 observations using HSB sample weights. Demographic and job variables include age, months worked since 1980, and indicators for marital status, educational attainment, government worker, and year. Test scores include math and verbal test scores from the HSB survey. Parental education includes indicators for maximum educational attainment of parents.

TABLE 2 (continued)
OLS ESTIMATES OF RACIAL WAGE DIFFERENTIALS AND COMPUTER WAGE PREMIUM

Panel B: PROG Measure of Computer Skill

Variable	Specification (1)	(2)	(3)	(4)	(5)
Computer Skill	0.119 (4.160)	0.097 (3.197)	0.092 (3.034)	0.083 (2.739)	0.069 (2.365)
Black	-0.122 (-4.345)	-0.106 (-3.495)	-0.064 (-1.996)	-0.058 (-1.794)	-0.070 (-2.198)
Hispanic or Asian	-0.014 (-0.428)	-0.006 (-0.178)	0.005 (0.129)	0.014 (0.387)	0.011 (0.322)
Black*Computer Skill	0.059 (0.995)	0.067 (1.135)	0.049 (0.854)	0.055 (0.953)	0.091 (1.605)
Hisp or Asian*Computer Skill	0.054 (0.729)	0.053 (0.710)	0.079 (1.042)	0.078 (1.041)	0.089 (1.201)
Test Scores	No	Yes	Yes	Yes	Yes
Demog. and Job Vars.	No	No	Yes	Yes	Yes
Parental Education	No	No	No	Yes	Yes
5 Industry Dummies	No	No	No	No	Yes
P-Values from Test of:					
Comp. + Black*Comp. = 0	< .001	.002	.006	.008	.001
Comp. + Hisp.*Comp. = 0	.012	.034	.018	.024	.020
Black + Black*Comp. = 0	.227	.453	.763	.955	.843
Hisp. + Hisp.*Comp. = 0	.548	.485	.224	.180	.177
Adjusted R^2	0.024	0.030	0.059	0.071	0.138

Note: t-statistics in parantheses. Standard errors adjusted for arbitrary forms of heteroscedasticity. All regressions based on 2388 observations using HSB sample weights. Demographic and job variables include age, months worked since 1980, and indicators for marital status, educational attainment, government worker, and year. Test scores include math and verbal test scores from the HSB survey. Parental education includes indicators for maximum educational attainment of parents.

TABLE 3
PROBIT ESTIMATES: Dependent Variable is Computer Skill Measure

	Definition of Computer Skill			
	SOFT		PROG	
Variable	(1)	(2)	(3)	(4)
Intercept	-1.972	-5.280	5.963	1.092
	(-1.325)	(-3.349)	(3.678)	(0.620)
College Dropout	0.292	0.302	0.049	0.066
	(3.682)	(3.732)	(0.575)	(0.747)
College Graduate	0.804	0.627	0.798	0.519
	(10.558)	(7.845)	(10.369)	(6.337)
Married	-0.003	0.041	-0.063	0.027
	(-0.040)	(0.611)	(-0.911)	(0.369)
Black	0.060	0.276	0.021	0.416
	(0.546)	(2.398)	(0.182)	(3.329)
Hispanic or Asian	-0.115	-0.017	-0.087	0.109
	(-1.181)	(-0.170)	(-0.845)	(1.004)
Months Worked since 1980	-0.0044	-0.0041	-0.009	-0.0099
	(-2.210)	(-2.033)	(-4.496)	(-4.554)
Government Worker	-0.159	-0.205	-0.099	-0.213
	(-1.841)	(-2.316)	(-1.111)	(-2.282)
Math Score	-	0.0321	-	0.0429
		(7.089)		(8.792)
Verbal Score	-	0.0002	-	0.0183
		(0.032)		(3.416)
Finds Math Interesting	0.227	0.173	0.318	0.250
	(3.832)	(2.860)	(5.510)	(3.870)
Thinks Math is Useful	0.297	0.210	0.323	0.185
	(4.636)	(3.187)	(4.771)	(2.591)
Family Owned Pocket Calculator	0.454	0.387	0.590	0.478
	(4.307)	(3.579)	(4.880)	(3.723)
Used Micro-computer in High School	0.870	0.825	0.675	0.625
	(7.679)	(7.143)	(6.011)	(5.321)
Used Computer Terminal in High School	0.135	0.105	0.186	0.116
	(1.495)	(1.148)	(2.043)	(1.224)
Log-Likelihood	-1231.32	-1194.24	-1121.43	-1032.28

Note: t-statistics in parentheses. Each regression based on 2388 observations using HSB sampling weights and includes age, indicators for year, parental educational attainment, and 5 industry dummies.

TABLE 4
SAMPLE SELECTION-CORRECTED (SSC) WAGE EQUATION ESTIMATES
Dependent Variable is ln(Hourly Wage) for Appropriate Sector

	Definition of Computer Skill			
VARIABLE	SOFT		PROG	
	= 1	= 0	= 1	= 0
Intercept	2.145 (2.640)	1.218 (2.411)	0.801 (0.979)	1.304 (2.652)
College Dropout	-0.067 (-1.532)	0.123 (4.249)	0.013 (0.299)	0.082 (3.263)
College Graduate	0.048 (1.162)	0.087 (2.098)	0.166 (4.302)	-0.013 (-0.337)
Married	0.068 (2.084)	0.070 (3.361)	0.043 (1.280)	0.084 (4.128)
Black	0.062 (1.073)	-0.070 (-1.905)	0.066 (1.077)	-0.079 (-2.141)
Hispanic or Asian	0.076 (1.469)	0.009 (0.293)	0.104 (1.968)	0.003 (0.103)
Months Worked since 1980	0.0014 (1.520)	0.0037 (5.259)	0.0020 (1.988)	0.0032 (5.137)
Government Worker	-0.209 (-5.141)	-0.118 (-3.905)	-0.274 (-6.711)	-0.093 (-3.175)
Math Score	0.0085 (2.945)	0.0029 (1.510)	0.0082 (2.396)	0.0022 (1.039)
Verbal Score	-0.0016 (-0.655)	-0.0012 (-0.789)	-0.0021 (-0.796)	-0.0016 (-0.943)
$g_{j1}(.)$	0.431 (1.963)	-0.164 (-0.906)	0.097 (0.509)	0.006 (0.042)
$g_{j2}(.)$	0.524 (1.752)	0.353 (1.629)	0.116 (0.480)	0.142 (0.874)
$g_{j3}(.)$	0.471 (1.243)	-0.389 (-0.779)	-0.079 (-0.314)	-0.202 (-0.540)
R^2	.153	.138	.206	.123
N	733	1655	668	1720

Note: t-statistics in parentheses. All regressions weighted by HSB sampling weights. Each regression also includes age, indicators for year, education of parents, and 5 industry dummies.

TABLE 5
COMPUTER WAGE PREMIUM, BY WORKER TYPE AND SKILL DEFINITION

GROUP	COMPUTER SKILL DEFINITION	
	SOFT	PROG
Overall	24.6%	12.7%
Whites	22.3%	10.1%
Blacks	39.6%	27.3%
Hispanics and Asians	30.7%	21.7%
High School Grads	29.8%	9.0%
College Grads	24.8%	30.5%
24 Months Worked Since 1980	32.5%	18.5%
60 Months Worked Since 1980	21.9%	10.6%
Manufacturing Workers	6.7%	-7.0%

Note: All wage differentials constructed using the average characteristics of workers in the sample.

APPENDIX A: The HSB Data Set

This study uses the senior cohort of the High School and Beyond data set, which is a sample of high school seniors from randomly selected high schools in 1980. A subsample of individuals was re-interviewed in 1982, 1984, and 1986. Individuals were included in the sample if they responded to both the base questionnaire and each of the three re-interviews. Predominantly black and Hispanic public and private high schools were over-sampled. Consequently, in each regression we use the sampling weights provided in the HSB. Individuals were dropped from the sample if they did not hold a full-time (> 20 hours per week) job between 1984 and 1986, as were workers who reported earning less than $1.25 or more than $75 dollars per hour. Wage responses of this type appear to result from inaccurate reporting of the schedule by which the worker was paid. In addition, individuals were dropped if they had missing values for one or more of the independent variables. Approximately 20% of the observations were missing data on the math, reading, and vocabulary test scores or the computer skill questions. The math, reading, and vocabulary tests were administered to high school seniors in 1980.

The computer knowledge data used here comes from the 1984 interview wave. The computer skill variables are derived from responses to the following questions: First, for use of computer software packages, individuals were asked:

"Have you ever used any of the following types of computer software packages? [Mark all that apply]

(a) Statistical Packages (e.g., SAS, SPSS, BMD);

(b) Business application packages (e.g., Visicalc for financial reporting, inventory control, and billing);

(c) Word processing software (e.g., WORDSTAR);

(d) Data-base management systems (e.g., System 2000, TOTAL);

(e) Instructional/educational courseware (e.g., PLATO);

(f) Other (Write In);"

The SOFT variable was constructed using the responses to this question as described in the text.

For programming language use individuals were asked:

"Have you ever written a computer program in any of the following computer languages? [Mark all that apply]

(a) Fortran;

(b) BASIC;

(c) COBOL;

(d) PASCAL;

(e) PL/I;

(f) APL;

(g) ASSEMBLY;

(h) Other (Please Specify);"

The responses to this question were used to construct the PROG measure of computer skill.

The earnings information refers to the hourly wage on the individuals current or most recent full-time job (i.e., the individual works 20 hours per week or more) after the completion of schooling, as of the 1986 interview. We use the wage data from the 1986 interview rather than from the 1984 wave since a large fraction of the sample was in college in 1984. The hourly wage may refer to a job held in 1984, 1985, or 1986. The vast majority of jobs were held in 1985 or 1986. There may be some concern that because the computer skill questions refer to the individual's computer skills at the time of the 1984 interview, some workers who acquired computer skills between 1984 and 1986 will be incorrectly classified as without skill. Although the number of such individuals is likely to be small, to investigate this issue we first estimated the model using only the 1984 wave wage data for the subsample of individuals at work as of the 1984 interview. We then estimated the model using the 1986 wave wage data for this same sub-sample of individuals. We found no qualitative difference in the results using the 1984 versus the 1986 wage data. To provide another check on the results, we also estimated the model using the wage on the first (rather than most recent) full-time job held by the worker between 1984 and 1986. Most of these jobs were held in 1984 or 1985. Again, we found no significant difference in the results using this data versus the estimates reported in the paper. Finally, we note that if some workers who acquire skills after 1984 are inadvertently included in the no computer skill category, this should narrow the estimated computer wage premium reported in the paper.

REFERENCES

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

Bound, John, and Richard Freeman. 1992. "What Went Wrong? The Erosion of Relative Earnings and Employment Among Young Black Men in the 1980s." *Quarterly Journal of Economics* CVII (February): 201-232.

Bound, John, and George Johnson. 1992. "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations." *American Economic Review* 82 (June): 371-392.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (January): 153-162.

_______. 1990. "Varieties of Selection Bias." *American Economic Review* 80 (May): 313-318.

_______, and Richard Robb. 1985. "Alternative Methods for Evaluating the Impact of Training on Earnings." in J. Heckman and B. Singer, eds., *Longitudinal Analysis of Labor Market Data*. Cambridge: Cambridge University Press.

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

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

Krueger, Alan. 1993. "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989." *Quarterly Journal of Economics* CVIII (February): 33-60.

Lee, Lung-Fei. 1978. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative And Limited Dependant Variables." *International Economic Review* 19, (June): 415-433.

_______. 1982. "Some Approaches to the Correction of Selectivity Bias." *Review of Economic Studies* XLIX: 355-372.

Levy, Frank and Richard Murnane. 1992. "US Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations." *Journal of Economic Literature* XXX, (September): 1333-1381.

Mincer, Jacob. 1989. "Human Capital Responses to Technological Change in the Labor Market." National Bureau of Economic Research working paper 3207.

Neal, Derek, and William Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (October): 869-895.

Newey, Whitney. 1988. "Two Step Series Estimation of Sample Selection Models." Mimeo, Princeton University.

_______, Powell, James, and James Walker. 1990. "Semiparametric Estimation of Selection Models: Some Empirical Results." *American Economic Review* 80 (May): 324-328.

Reilly, Kevin. 1995. "Human Capital and Information: The Employer Size-Wage Effect." *Journal of Human Resources* 30, (Winter): 1-22.

Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3: 134-146.

Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87, (August): 355-374.

Willis, Robert J. And Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87 (Supplement): S7-S36.