2005

Four essays on technology adoption and returns to skill in the US

Moohoun Song
Iowa State University

Four essays on technology adoption and returns to skill in the U. S.

by

Moohoun Song

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Peter F. Orazem, Major Professor
Steven B. Garasky
Joseph A. Herriges
Wallace E. Huffman
Brent Kreider

Iowa State University

Ames, Iowa

2005

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Moohoun Song

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

ABSTRACT

We investigate two plausible factors that may have affected changes in wage inequality: returns to higher education and returns to new technologies. The first two chapters examine the role of mathematics and verbal ability in estimating returns to advanced degrees. When average abilities within the major are treated as missing variables, we found that OLS estimates of returns to graduate education are underestimated. When average ability in the major is treated as part of an endogenous decision regarding whether to attend post-graduate degree programs, we find that students in majors with higher average mathematics skills are less likely to progress beyond the bachelor's degree while the opposite happens with average verbal skills.

The next two chapters examine the decision of whether to adopt Internet and computer technologies and the returns to adoption. We first identify demand-side and supply-side factors that affect technology adoption in urban and rural areas. Local access to high speed Internet plays an important role in the technology adoption decision. It increases the probability of using computers and the Internet for work from home and also increases the likelihood of using the Internet at work. That factor alone explains about half of the gap in Internet adoption at home or at work between urban and rural workers. Together, the demand and supply-side factors identified in the analysis completely explain the differences in technology adoption between urban and rural areas.

Using the previous model to identify the endogenous probability of adopting various information technologies, we estimate returns to adoption in the context of an earnings function. When treated as exogenous, adoption has an implausibly large positive and significant effect on earnings. When the endogeneity of the choice to adopt is controlled, the

estimated returns to adoption shrink in both sign and significance. Thus, while adoption is strongly tied to the availability of high-speed Internet in the home county, the higher income of adopters is due to factors that raise both the probability of adoption and earnings and not to the adoption *per se*.

CHAPTER 1. GENERAL INTRODUCTION

Wage inequality has increased since the 1980's in the United States. Using data from the March Current Population Survey (CPS), Juhn, Murphy, and Pierce (1993) showed that, for men, annual earnings inequality increased slowly in the 1970's and more rapidly in the 1980's. For women, annual earnings inequality declined modestly in the 1970's, but increased in the 1980's. For both men and women, increasing inequality in earnings was driven by increased wage variation rather than increased variation in hours worked. Inequality continued to increase in the 1990s among the youngest cohorts of workers (Card and Lemieux (2001)). Katz and Autor's (1999) summary of the literature on changes in the U.S. wage structure concluded that wage inequality increased substantially from the late 1970s to the mid-1990s. They also point out that wage inequality increased within demographic and skill groups so that there is greater variation in wages for workers with the same education level. These within group wage discrepancies were much larger in the mid-1990s than 1970s.

In this study we investigate two plausible factors that may have affected changes in wage inequality: returns to higher education and returns to new technologies, namely computer and the Internet skills. In chapter 2, we investigate returns to post college educations and the role of mathematical and verbal skills by major. Treating ability within the major as a missing variables problem, we estimate returns to graduate and professional degrees and direct returns to major level mathematical and verbal skills. Least squares estimates of returns to schooling that exclude major level ability are shown to be biased downward. Adding measures of major level ability shows much larger estimated returns to graduate education. Under plausible assumptions, we argue that these corrected estimates are lower bound measures of the true returns to graduate education.

In chapter 3, we treat the decision to attend graduate school as an endogenous choice. The graduate school entry decision is assumed to depend on average mathematical and verbal skills in the major as well as anticipated returns to graduate or professional education. Using parents' education level and graduate schooling cost measures as instruments in a first stage model explaining graduate and professional school entry, we are able to then use the predicted probability of completing a Master's, doctorate or professional degree in a second stage earnings function. Results show a statistically significant role for mathematics and verbal skills in the schooling choice. Consistent with the implication of chapter 2, students in majors with higher average mathematics skills are less likely to progress beyond the bachelor's degree while students in majors with higher average verbal skills are more likely to obtain post graduate education. Direct returns to mathematics and verbal skills were statistically insignificant after controlling for the endogenous choice of how long to go to school. As in chapter 2, consideration of the role of ability in the schooling decision greatly increases estimates of the return to post-graduate education.

Chapter 4 analyzes the computer and the Internet technology adoption decisions. Using a probit model, we identify demand-side and supply-side factors that affect various technology adoption decisions in urban and rural areas. Results show that high speed Internet access plays major a role in computer or Internet use. Differences in individual demand factors and Internet access between urban and rural areas completely explains observed regional differences in technology adoption often referred to as the urban-rural digital divide.

We evaluate the returns to information technology adoption in chapter 5. Using the estimation results from chapter 4, we identify the endogenous technology adoption decision in the context of an earnings function that is used to estimate returns to computer and the Internet use. We find positive, large and statistically significant returns to information technology adoption when the technology uses are treated as exogenous. However the earnings premium associated with technology adoption shrinks in magnitude and becomes statistically insignificant when we control for the endogeneity of technology use.

References

Card, David and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men?" *Quarterly Journal of Economics*, 2001, vol. 116, no. 2, 705-746.

Juhn, Chinhui, Kevin, M. Murphy, and Brooks Pierce, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 1993, vol. 101, no. 3, 411-442.

Katz, Lawrence and David Autor (1999), "Changes in the Wage Structure and Earnings Inequality," *Handbook of Labor Economics*, 1999, vol. 3A, 1463-1555

CHAPTER 2. THE ROLE OF MISSING MEASURES OF MATHEMATICAL AND VERBAL SKILLS ON RETURNS TO HIGHER EDUCATION

I. Background

Wage inequality has increased since the 1980's in the United States. Using data from the March Current Population Survey (CPS), Juhn, Murphy, and Pierce (1993) showed that, for men, annual earnings inequality moved from stability or gradual increases in the 70's to rapid increases in the 80's. For women, annual earnings inequality moved from a modest decline in the 70's to increases in the 80's. For both men and women, increased inequality in earnings was driven by increased wage variation rather than increased variation in hours worked.

The increasing wage inequality has been accompanied by increasing returns to college education during the 1980's (Balckburn, Bloom, and Freeman (1990), Levy and Murnane (1992), Katz and Revenga (1989) and Murphy and Welch (1992)). The college wage premium relative to high school graduates sharply increased in the late 80's (Katz and Murphy (1992) and Murphy and Welch (1989)) and increased further in 1990's especially for young college graduates (Card and Lemieux (2001)).[1] While steadily increasing returns to schooling seem to be able to explain some of the recent rise in wage inequality, much of the increase is due to skills other than years of schooling or experience (Juhn, Murphy and Pierce (1993)). Katz and Murphy (1992), using March Current Population Survey from 1964-1988, found that within group inequality that is due to differences in individual skills and ability increased 30 percent in 1987 relative to 1970.[2]

One line of research focuses on the role of college majors on within group wage inequality. Some majors, business and engineering for example, pay better than others (Bishop (1991), Berger (1988a, b), James, Alsalam, Conaty, and To (1989) and Rumberger (1984)). For example, Rumberger and Thomas (1993), using the 1987 Survey of Recent College Graduates, found that, for both men and women, graduates in engineering and health have the highest relative salaries. Women majoring in engineering and health have a higher relative wage advantage than men. Next to follow are science or mathematics and business majors. Education, social science and humanities majors receive the lowest earnings. Berger (1988b) found that, for male workers, engineering undergraduates command the highest starting salaries while liberal arts graduates are paid the least. Business graduate salaries lie in the middle as are all other majors, presumably. Moreover, the gap in salaries has been widening. Over the period from 1960 to 1981, engineering salaries have grown the fastest, followed by science, business and the liberal arts.

Many studies have examined the role of ability on estimated returns to schooling. Appropriate ability measures can be used to explain within group inequality controlling for levels of education as well as between groups wage inequalities. If latent abilities or cognitive skills are excluded, estimates of returns to schooling will be biased, presuming that these abilities and years of schooling are correlated.[3] However there has been no concrete evidence whether excluded ability measures could explain some of the recent increase in returns to schooling or the college wage premium.[4]

Some evidence has been advanced that returns to cognitive skills have risen (Juhn, Murphy and Pierce (1993)). Researchers have used measures of mathematics and verbal abilities to control for cognitive skills in estimating earnings functions. Murnane, Willett, and

Levy (1995), using data from two longitudinal surveys of American high school seniors, the National Longitudinal Study of the High School Class of 1972 (NLS72) for 1972 and High School and Beyond (HS&B) for the class of 1980, showed that the impact of basic cognitive skills on wages increased between 1972 and 1986.[5] They argue that this increase in returns to cognitive skills is due to recent technology favoring skilled workers. In addition to that, they showed that the mathematics score in the model completely absorbs the wage premium from college students for females and reduces its magnitude for males from 100 percent to 62 percent. Levine and Zimmerman (1995) tested whether taking more high school math and science classes had an impact on wages. Using data sets from the National Longitudinal Survey of Youth (NLSY) of 1979 and HS&B of 1980, they found that additional high school math classes increased the wages of women college graduates. For both men and women, high school science classes appear to have no significant impact on wages. Grogger and Eide (1995) tried to measure cognitive ability using standardized test scores and high school grades. Using data from NLS72 and HS&B of 1980, they found that skills attained prior to college had no impact on changes in the college wage premium for men. However they found that the math score had a significant impact on wages for women. They argue that recent estimates of the increasing college wage premium that fail to account for math ability may overstate the increase in the value of a college education for women. Moll (1998), using South Africa data, found that even mastery of a primary school level of cognitive skill had a significant impact on wages. Boissiere, Knight, and Sabot (1985) argue, in their work on East Africa that the main cause of the difference in earnings between primary and secondary educated workers was their cognitive skill differences.

This study explores several questions that have not been addressed in the previous literature. First, while the increase in average returns to schooling and the college wage premium have been well documented, few studies have focused on higher levels of education. Most studies on returns to schooling have focused on the earnings difference between high school graduates and college graduates. However, as shown in Figure 1, relative returns to post graduate degrees have risen steadily in the 1980's and have outpaced the gains to bachelor's degrees in the 1990's. This rise in returns to post graduate education has not been investigated.

Second, studies using college major variables to explain wages show a significant difference in returns to college majors, but the reasoning is somewhat vague. Berger (1988) found that the cohort size effects of business and engineering majors are smaller or insignificant where as the effect is negative for science and liberal arts graduates. This could imply that usual demand and supply analysis for different wage premiums of different college majors may not be appropriate. Grogger and Eide (1995) found that there are significant differences in wage premium among different college majors and they also found that there are substantial amount of wage premium due to college skills by different college majors. Studies have also found that the significant impact of cognitive skills on earnings functions, but they have concentrated on measures of cognitive skills at very basic levels of math or reading ability.[6] However, wage inequality is increasing more rapidly at higher levels of education. Grogger and Eide (1995) also reported that pre-college test scores and high school grades have no effect on changes in the college wage premium for men, but that choice of college major is highly significant. Assuming college major dummy variables indirectly control for college skills, their results may reveal that lower levels of cognitive skills are not as important as the skills acquired in specific college majors, and it is these latter skills that drive the changes in the college premium. Since college skills seem to reflect a substantial portion of the college wage premium, at least for men, it is interesting to examine the role of more educated cognitive skills on earnings.[7] In this paper we are going to estimate returns to college major level cognitive skills. Finding its significance will give us one reason of persistence and even increasing major specific college wage premium.

Lastly, this paper deals with the bias on measured returns to schooling coefficients. Studies argue that unmeasured skills are positively correlated with years of schooling, but the evidence is not strong. Using high school or lower levels of cognitive skills to proxy missing ability, several studies have found that missing ability causes a positive bias on estimated returns to schooling.[8] Although the measures of cognitive skills in this paper are aggregated at the college major level, it would be interesting to see whether this result holds at higher levels of schooling and cognitive skills.

Section II discussed theoretical background of the log earnings function and missing ability bias in least squares estimation. Section III discusses the data, section IV discusses the estimation strategy, section V reports the estimation results, and section VI includes conclusions.

II. Theoretical Background

Human Capital Earnings Function and Ability Bias in OLS Estimation

Almost all the recent empirical research on education and earnings is based on the human capital earnings function (HCEF) developed by Mincer (1974) and Becker (1967).[9] Their log-linear earnings function is statistical in nature, although it fits the data remarkably

(Heckman and Polacheck (1974), Card and Kruger (1992), and Park (1994)).[10] Followed by Mincer (1974), much of applied work in economics of education has used a version of the following equation in estimation:

(1) $y_i = \ln Y_i = \beta_0 + \beta_s S_i + X_i \beta_X + u_i$,

where Y_i is a measure of earnings, S_i is a measure of schooling, X_i is a vector of other variables assumed to affect earnings such as experience, experience squared, race, and so on, and u_i is a random disturbance term.

To derive unbiased estimates of the coefficient in earnings function, it must be that the disturbance term is not correlated with the independent variables. However, if individual ability, which is believed to be correlated with schooling, is omitted from the earnings function, then this condition will be violated. A simple example by Griliches (1977), suppressing subscripts, illustrates this point very well. Let

(2) $Y = w H e^{u}$,

(3) $H = e^{\delta_S S + \delta_\mu \mu}$,

(4) $y = \ln Y = \ln w + \delta_S S + \delta_\mu \mu + u$,

where w is market rental price of a unit of human capital that may vary over time and space, H is the unobservable quantity of human capital, and u is other, random influences on earnings that is normally distributed with the zero mean.[11] Equation (3) is an implicit human capital production function with time spent in schooling (S) and ability (μ) that affects quality of human capital and the efficiency of schooling, and δ_S and δ_μ are parameters.[12] As we mentioned, to have consistent estimates of coefficients, S and μ should not be correlated with u. However, when we estimate equation (4) without μ, we end up estimating

(5) $y = \ln Y = \ln w + \hat{\delta}_S S + \upsilon$,

where $\upsilon = \delta_\mu \mu + u$.

The least square estimate of schooling coefficient will be

(6) $$\hat{\delta}_S = \frac{\sum (S-\bar{S})(y-\bar{y})}{\sum (S-\bar{S})^2},$$

where $\bar{S}$ *and* $\bar{y}$ are sample means of schooling and income. Substitute υ for $(y-\bar{y})$ in (6) and rearrange to get

(7) $$\hat{\delta}_S = \delta_S + \frac{\sum (S-\bar{S})\upsilon}{\sum (S-\bar{S})^2}$$

$$= \delta_S + \delta_\mu \frac{\sum (S-\bar{S})\mu}{\sum (S-\bar{S})^2} + \frac{\sum (S-\bar{S})u}{\sum (S-\bar{S})^2}$$

$$= \delta_S + \delta_\mu \frac{\sum (S-\bar{S})(\mu-\bar{\mu})}{\sum (S-\bar{S})^2}.$$

Asymptotically, this converges to

(8) $$\hat{\delta}_S = \delta_S + \delta_\mu \frac{Cov(\mu, S)}{Var(S)}.$$

The direction of the bias will depend on the sign of δ_μ and the sign of the covariance between ability and schooling. Theoretically, the direction of the bias is ambiguous. The δ_μ, the market price of the individual cognitive skills, is assumed to be positive. The ambiguity of the bias comes in the correlation between μ and *S*. Years of schooling could increase the value of the cognitive skills, but the higher value of the cognitive skills would make the time out of school more valuable.[13] The sign of the correlation between schooling and cognitive skills is not clear.

Many empirical studies using less than college level of test scores on mathematics and language ability as a proxy of cognitive skills, report an upward bias in the schooling coefficient in the absence of measures of cognitive skills. This implies that the correlation between ability and years of schooling is positive. One of our objectives is to establish whether this relationship between cognitive skills and schooling variables holds for higher levels of education.

III. Data

The data set analyzed in this paper is collected from three sources. The main individual information is taken from the Scientist and Engineer Statistics Data System (SESTAT) of the National Science Foundation (NSF). To this, we merge in data on average cognitive skills by major from the Educational Testing Service (ETS). We also use information on the total number of doctorate recipients and the number of doctorate recipients who are not US citizens by major from the Survey of Earned Doctorates (SED) collected by the Division of Science Resources Statistics (SRS) of NSF. Brief explanations of each data set follows in order.

SESTAT Data

We use the 1993 Scientist and Engineer Statistics Data System (SESTAT) data set collected by National Science Foundation (NSF). SESTAT is a database of the employment, education, and demographic characteristics of the nation's scientists and engineers. The data are collected from the following three surveys: the National Survey of College Graduates, the

National Survey of Recent College Graduates, and the Survey of Doctorate Recipients. What is special about 1993 SESTAT is that it also includes the non-science and engineering (S&E) population. The 1993 National Survey of College Graduates was a once-a-decade baseline survey that also covered the non-S&E population with bachelor's or higher degree of about 29 million people. The 1993 SESTAT, therefore, represents the population of all bachelor's degree recipients in the United States.

The SESTAT data includes college graduates whose bachelor's degrees were earned between 1939 and 1992. The data includes information on individual demographics, parents' education level, year of college graduation, year of highest degree received, and salary. Descriptive statistics for the variables included in this study are reported in Table 1. The education categories of SESTAT are listed under 6 major groups. They are further subdivided into detailed majors. To correspond with the ETS and SED data sets, the detailed majors were allocated to 27 major groups and a catchall group of majors that did not fit the ETS fields. Engineering is the most populated major and public administration is the least. About 35 percent of the sample has a highest degree at the bachelor's level, 23 percent are master's degree holders, and 37 percent have the PhD. About 5 percent hold professional degrees in the sample. The average salary is slightly less than $55,000. Average salary, as we can see in Table 2, is higher for Whites and Asians than Blacks and Native Americans. Male salary is 25.6 percent higher than female. Chemistry and physics are among top paid majors when linguistics and social work are lowest paid majors. Average age is about forty years old.

GRE Data

Our measure of cognitive skills is based on average quantitative and verbal test scores by intended major on the Graduate Records Examination (GRE). The Educational Testing

Service (ETS) provided this data by selected years: 1963, 1974 to 1976, 1983 to 1986, and 1997 to 2000. The number of majors included in the report varied from 21 majors in 1963; to 92 majors in 1974 – 76; to 98 majors in 1983 – 86; and to 200 majors in 1997 - 2000. These were aggregated into 28 major groups. Of the 28 majors, 9 were not available in 1963. These were placed into the closest included major, so computer science was placed in mathematics, architecture was placed in music and so on. Once consistent data series were generated for the four reporting dates, the values were interpolated to generate continuous values for the intervening years. As most average scores change very slowly, this process is unlikely to generate wildly inaccurate estimates of average scores by major.

The trends of the quantitative, the verbal, and the Iowa Test of Basic Skills scores are reported in Figure 2. Scales of all series are normalized to 1 at 1963. The data show that both quantitative and verbal scores increased from 1963 to 1975, but decreased from 1975 to 1985. From 1985 to 1998, quantitative scores increased again but verbal scores kept decreasing at a slower rate. Bishop (1989) reported annual average scores on the Iowa Test of Basic Skills (ITBS) by Iowa Test of Educational Development (ITED) for Iowa students from 1940 to 1990. Bishop argues that, since the data set includes 95 percent of the public and private schools in the state of Iowa that are regularly participated in the testing program, ITED data for Iowa is free of selectivity bias and this feature makes ITED trends of test scores better representation of national trends prior to 1970 than the American College Test (ACT), the Scholastic Aptitude Test (SAT), and the American Council on Education Psychological Exam. Since these other tests were at first taken by a highly selected group and only more recently by more representative samples of college students, the trends of these tests' scores are biased by the decreasing selectivity of those who took the test. The scores on the 12th

grade test increased from 1940 to 1970, decreased from 1970 to 1980, and increased thereafter. This trend shows almost the same pattern as the GRE quantitative score. The time difference of the peak and the trough for the two series is about 6.5 years, which makes sense considering it takes four to five years to graduate from college.

Individuals are allocated GRE scores based on their major. Based on that, sample statistics on GRE scores by different categories of variables are reported in Tables 3 and 4. Female GRE scores are 1.6 percent higher on the verbal test and 6.8 percent lower on the quantitative test than males. Asians get the highest average quantitative score among different races and bachelor's degree recipients have highest quantitative scores than other degree recipients. Professional degree holders and Native American get the best average scores in the verbal test.

Survey of Earned Doctorates (SED) Data

The decrease in the GRE verbal score after 1985 is somewhat different from the ITED score trend. We suspect that the decline might be related to an increase in the number of foreign students taking the GRE. The data from Survey of Earned Doctorates (SED) by the Division of Science Resources Statistics (SRS) of the NSF includes the total number of doctorate recipients and the number of doctorate recipients who are not US citizens by major. We took the ratio of these foreign graduates to total graduates as an approximation of the ratio of foreign students taking the GRE. Since the SED reports the year when the student received the doctorate degree, we subtracted six years from the SED data to make it consistent with the year an individual would be expected to have received the bachelors' degree and taken the GRE. The time series of the ratio is shown in Figures 3 and 4. The ratio increases from 10 percent at 1963 to 27 percent in 1985 and stabilizes at about 25 percent after that. The trend increases sharply between the late 1960's and the mid 1980's. This is consistent with the pattern of GRE verbal scores during the same period. The correlation coefficient between the ratio and the GRE verbal score is –0.34 and the correlation coefficient between the ratio and the quantitative is 0.72. The percentage of foreign students is highest in chemistry and lowest in psychology.

IV. Estimation Model

Griliches (1977) dealt with missing observations on individual level ability. Our application includes a level of average ability at the major level, but we do not observe how well one individual performs relative to other majors. To demonstrate the implications for estimation, suppose an individual *i* in major *j* has ability

(9) $\mu_{ij} \equiv \mu_j^M + e_{ij}$,

where μ_j^M is average ability across all individuals in major *j*, and $e_{ij} = \mu_{ij} - \mu_j^M$, is individual i's ability relative to others in major j. We call the former "major component" and the latter "individual component." By definition $E\,[e_{ij}|j] = 0$ and $Cov(\mu^M, e) = 0$. The covariance between schooling and individual ability will have two terms: $Cov(S, \mu^M)$ and $Cov(S, e)$. Inserting (9) into (4), we have

$$(10)\ \ y_{ij} = \ln Y_{ij} = \ln w_i + \delta_S S_i + \delta_\mu (\mu_j^M + e_{ij}) + \varepsilon_i$$

$$= \ln w_i + \delta_S S_i + \delta_\mu \mu_j^M + \delta_\mu e_{ij} + \varepsilon_i .$$

If we do not have any measure of individual or major level ability, we estimate

(11) $y_{ij} = \ln Y_{ij} = \ln w_i + \hat{\delta}_S S_i + \upsilon_{ij}$,

where $\upsilon_{ij} = \delta_\mu (\mu_j^M + e_{ij}) + \varepsilon_i$.

Then the least squares estimate of the schooling coefficient will be

(12) $$\hat{\delta}_S = \delta_S + \frac{\sum_i (S_i - \bar{S})\upsilon_{ij}}{\sum_i (S_i - \bar{S})^2}$$

$$= \delta_S + \delta_\mu \frac{\sum_{j=1}^{J} \sum_{i=1}^{n_j} (S_i - \bar{S})\mu_j^M}{\sum_i (S_i - \bar{S})^2} + \delta_\mu \frac{\sum_i (S_i - \bar{S}) e_{ij}}{\sum_i (S_i - \bar{S})^2} + \delta_\mu \frac{\sum_i (S_i - \bar{S})\varepsilon_i}{\sum_i (S_i - \bar{S})^2}$$

$$= \delta_S + \delta_\mu \frac{\sum_{j=1}^{J} \sum_{i=1}^{n_j} (S_i - \bar{S})(\mu_j^M - \bar{\mu}^M)}{\sum_i (S_i - \bar{S})^2} + \delta_\mu \frac{\sum_i (S_i - \bar{S}) e_{ij}}{\sum_i (S_i - \bar{S})^2},$$ where $j = 1, 2, \ldots, J$, $i = 1, 2, \ldots, I$, and $I = n_1 + n_2 + .. + n_J$.

Asymptotically, this converges to

(13) $$\hat{\delta}_S = \delta_S + \delta_\mu \frac{Cov(\mu^M, S)}{Var(S)} + \delta_\mu \frac{Cov(e, S)}{Var(S)}.$$

Therefore, the bias due to missing ability measure is

(14) $$B = \hat{\delta}_S - \delta_S = \delta_\mu \frac{Cov(\mu^M, S)}{Var(S)} + \delta_\mu \frac{Cov(e, S)}{Var(S)},$$ which is equal to the bias term in equation (8). We call the first term B_1 and the second term B_2 in the right hand side of the equation (14). When major level ability, μ_j^M, is included in the estimation, the bias term is to B_2 $= \delta_\mu \frac{Cov(e, S)}{Var(S)}$.

There are three possibilities in terms of the bias. One is when $Cov(\mu^M, S) = Cov(e, S) = 0$. In this case, we would get an unbiased estimate of schooling from direct estimation of (11). Second is when $Cov(\mu^M, S) \neq 0$ but $Cov(e, S) = 0$

so that years of schooling is related to average ability in the major but not to the difference between the individual's ability and the average major. We can still get an unbiased estimate of the schooling coefficient by including a measure of major level ability in the earnings function as in (10).

The last case is when $Cov(\mu^M, S) \neq 0$ and $Cov(e, S) \neq 0$. In this case, there is no guarantee that inclusion of major level ability will yield a smaller bias. If the covariance terms have opposite signs , the size of B will depends on the relative sizes of B_1 and B_2. Only when $Sign(Cov(S, \mu^M)) = Sign(Cov(S, e))$ will $B > B_2$. It might seem likely that if average ability in the major is positively correlated with schooling, then individuals with above average ability within the major would be the most likely to obtain an advanced degree, but there is no guarantee.

In our empirical work below, we present evidence that the least squares estimate excluding major ability is smaller than the least squares estimate when major ability is included. If $Sign(Cov(S, \mu^M)) = Sign(Cov(S, e))$ and if $\delta_\mu > 0$, then it must be true that $Cov(\mu^M, S) < 0$ and taking the assumption that the signs of the covariance terms should be equal, that $Cov(e, S) < 0$. If $\delta_\mu < 0$, then the opposite covariance signs would hold. It turns out that we use two different measures of major-level ability whose inclusion raise the estimated δ_S. If our assumption that $Cov(e, S)$ will have the same sign holds, then any remaining bias due to unmeasured individual ability will be in the same direction, suggesting that we will have a lower bound estimate of the true returns to schooling.

To estimate the role of cognitive ability and the effect of college majors on earnings, we estimate an ordinary least squares (OLS) regression models as follows:

(15) $\ln Y_i = \beta_0 + Edu_i\,\beta_1 + X_i\beta_2 + u_i$,

(16) $\ln Y_i = \beta_0 + Edu_i\,\beta_1 + X_i\beta_2 + GRE_i\beta_4 + u_i$,

(17) $\ln Y_i = \beta_0 + Edu_i\,\beta_1 + X_i\beta_2 + Major_i\beta_5 + u_i$,

(18) $\ln Y_i = \beta_0 + Edu_i\,\beta_1 + X_i\,\beta_2 + GRE_i\beta_4 + Major_i\beta_5 + u_i$,

where Edu_i includes education level dummy variables such as bachelor's degree, master's degree, PhD and professional degree, GRE_i includes the average GRE quantitative and verbal scores by major, X_i includes control variables such as experience, experience squared, and dummy variables indicating whether the individual is male, a US citizen, or Hispanic, White, Black, Asian, or Native American. The equation (15) is the base equation to be compared with the other three equations: the reference is a white female non-US citizen who has bachelor's degree.

To see the effect of cognitive skills on earnings, equation (16) includes the variables in (15) plus a vector of measures of the average cognitive skills in major. This vector includes the average GRE quantitative and verbal scores by major.[14] Comparing equation (15) and (16), we would capture the role of cognitive skills on earnings function and how the cognitive skills are related to schooling variables. We expect coefficients of both verbal and quantitative score variables are positive.

Equation (18) includes the variables in (17) plus $Major_i$, a vector of 28 major dummies used as an alternative skills measure: the reference major is public administration. The comparison of the regression results of equation (16) with (17) will determine the extent to which differences in GRE scores by major explain the differences in returns to major.

Equation (18) includes all the variables discussed above. The other equations are restricted forms of (18), allowing for joint significance tests for the $Major_i$ and GRE_i vectors.

V. Estimation Results

Table 5 presents parameter estimates and associated standard errors for each regression model discussed in section III. The omitted category for the schooling dummies is the bachelor's degree; for parents' education level it is the high school graduate; and for the college major it is public administration. The racial dummies use White as the reference group.

Schooling, Experiences, and Other control Variables

The estimates of the regressions are consistent with theoretical expectation and estimates of other studies (Grogger and Eide (1995), Moll (1998), Blackburn and Neumark (1995), Murnane, Willett, and Levy (1995)).[15] Schooling dummy coefficients are positive and significant in all four regression models. The PhD degree has twice the impact on earnings as a master's degree. Professional degrees have the largest impact on earnings. The magnitudes of the experience coefficients lie between 0.026 in the base model and 0.041 in the model 18. The coefficients of squared experience were negative and small, which implies that the effect of experience on wage is positive but of decreasing rate. Minorities earn less than Whites and the difference is biggest between White and Native Americans. Post-doctorate (postdoc) position is considered as on the job training. We can also see that workers in that position receive about 40 percent less than others.

College Major and GRE Scores

The sign of the coefficient of the quantitative score is positive when the sign of the verbal score is negative and both are significant. This result is consistent with studies conducted at lower levels of education.[16] Most studies found that, when both math and verbal scores are included in a model, the math score was positive and significant and the verbal score was either positive or insignificant. Adding test scores raises the R-square by 21 percent relative to the baseline equation. This may imply that test scores explain substantial amount of the wage variation.

Adding college major dummies to the baseline equation as a measure of skill content instead of GRE scores results in a regression that is remarkably similar to the one using GRE scores averaged by majors as the measure of skill content. Comparing coefficients on the same variable between equations (16) and (17), we find that most are almost identical qualitatively and quantitatively. As we can see from equation (18), we do not gain much explanatory power by using both GRE test scores and college major dummies together. The R-square is only 5 percent higher for equation 18 than for equation 16. Again, coefficients on the common variables in equations (16) and (18) except the test scores and major dummies are also identical to those in equation (16). It appears that the test scores explain almost as much wage variations as the college major dummy variables do. The estimates of coefficients of majors in natural science, social science, or arts are significant and negative.

The coefficient estimates of variables such as gender, nationality, experience, and schooling are affected by adding either GRE test scores or college major dummies. Postdoc, race, and regional variables do not seem to be correlated with these ability variables. The most interesting results will be the changes in coefficient estimates of schooling variables due to addition of GRE scores. They are 8 to 10 percent higher with test scores than without them. As we can see in equation (14), the sign of the bias of OLS estimates is determined by the sign of the coefficients of major average ability variable and the correlation between the variable of the interest and the major average ability variable. The sign of the coefficient on the quantitative score is positive and the coefficient on the verbal score is negative. Table 6 shows the shares of the people whose GRE scores are low, middle, or high for both verbal and quantitative scores by different degree level.[17] For Quantitative score, the ratio of the share of the high score group to the share of the low score group of BA holders is about 70 percent larger than those of MA or Ph.D. holders. In contrast, the ratio of the share of the high score group to the share of the low score group of BA holders for GRE verbal score is about 10 percent smaller than that of MA and about 30 percent smaller than that of Ph.D. holders. That is, the relative proportion of high GRE quantitative scorer is higher in BA holders than in MA or in Ph.D. holders, when the relative proportion of high GRE verbal scorer is lower in BA holders than in MA or in Ph.D. holders. This is evidence that the correlation between the verbal score and schooling is positive, but that the GRE quantitative score and schooling are negatively correlated. One possible explanation for this result is that students who have good mathematics skills do not seek advanced degrees because of their higher opportunity cost of staying school. However, verbal skills are not valued as highly in the market, so students who have good verbal skills decide to stay in school longer.

Recall in equation (14) the bias term due to missing major level component in the coefficient estimate of schooling enters in the product form of δ_{μ} and $Cov(\mu^{M},S)$. For GRE quantitative scores, $\delta_{\mu}>0$, but the covariance is negative. For GRE verbal scores, $\delta_{\mu}<0$, but the covariance is positive. Therefore in both cases, the bias is negative, suggesting that the OLS estimate understates the true return. Furthermore, if the bias attributable to unmeasured individual ability within the major goes in the same direction, then our corrected measure of the returns to graduate education will be a lower bound of the true return.

If the bias due to the unmeasured individual ability component, B_1, has an opposite sign with the bias term due to the major component, B_2, we cannot be certain about the size or sign of the combined total bias term, B. To get an idea about the sign of the correlation between unmeasured individual ability and the choice to pursue an advanced degree, we examine the residuals from the earnings function. The residuals represent unmeasured individual ability as well as any other unmeasured individual heterogeneity that affects earnings including luck. We run an auxiliary regression of education choices on the earnings residuals using multinomial logit. The sign of the coefficient will show the direction of combined effect of unobserved math and verbal ability on schooling choices. However, the estimated coefficients will be biased toward zero because the residual measures ability with error.

Table 7 reports the estimated marginal effect of the earnings residuals on the probability to pursue each degree. The sign pattern suggests that higher unmeasured ability lowers the probability of pursuing doctoral or professional degrees. However the coefficients are very small and not statistically significant. One interpretation is that the covariance between unmeasured ability and schooling is zero and so our major level ability correction fully controls for the bias, B.

It is hard to explain what this result represents if there exist correlations between unmeasured math and verbal ability and schooling. Recall that we need to know both the sign of correlation between ability and schooling and the sign of return to that ability to determine the sign of the ability bias in the returns to schooling.

The unmeasured ability in the residuals includes both unmeasured math and unmeasured verbal ability. We cannot measure separate correlations between the unmeasured math and verbal abilities with schooling, but only the sum of the two correlations. If the unobserved ability effect is the same as the observed ability effects, they will go in opposite directions and they may cancel each other, consistent with our finding that the combined correlations are insignificant. However, if the true summed effect is negative (so that the finding of insignificance is due to measurement error), then we cannot infer the sign of the bias term because we do not know the sign of the return to unmeasured math and verbal ability. It is possible to have a sizable bias, B_2, even though the correlations cancel each other, if the returns to unmeasured math ability have a different sign than the returns to unmeasured verbal ability. It is also possible to have B_2 close to zero depending on sign of returns to unmeasured ability when the sum of the two correlations of unmeasured math ability and schooling and of unmeasured verbal ability and schooling is negative. Therefore, we can not be certain if the corrected effect with measured ability makes the returns to schooling closer to the truth or not.

This result is opposite to the results of other studies. Grogger and Eide (1995), Murnane, Willett, and Levy (1995), and Blackburn and Neumark (1995) used high school level test scores for measures of cognitive skills in earnings function and they found that schooling coefficient is overestimated due to omission of ability variables. These seemingly different results can be explained. At lower levels of education, students with more math ability stay longer in school because returns to cognitive skills and returns to additional schooling are relatively larger than the opportunity cost of schooling. After college, opportunity cost for post college education increases more rapidly than returns to schooling. For those who have good cognitive skills, the opportunity cost is even higher since the skills are highly valued in labor market.

Returns to GRE Scores and Returns to Changes of College Major Distribution

Based on the estimation results in table 5 we simulate the impact of changes in GRE scores and in choice of college majors on earnings of the sample from 1963 to 1993. First, to see the impact of GRE scores, we hold all other variables except the mean GRE verbal and the mean quantitative scores at their sample means over the period.[18] To get pure variations of GRE scores net of the effect of changes in the ratio of the foreign graduates to total graduates, we regress both GRE scores on the ratio and collect residuals to get "pure" GRE scores. This GRE scores net of the ratio effect are in figure 5. There is not much dramatic change in the shape of GRE verbal score, but the dipping point in GRE quantitative score around 1985 is down much deeper compared to the GRE unadjusted quantitative score in figure 2. We use this purged GRE scores in the following earnings simulation.

Figure 5 demonstrates the impact of changes in average verbal and quantitative skills on earnings by bachelor's degree cohort. Over time, as shown in Figure 5, GRE verbal scores rise from 1963 through 1975 and decrease thereafter. Quantitative scores rise from 1963 to 1975, decrease, and then rise again after 1986. Holding all other variables at their sample means, the simulations in Figure 6 show the comparative static impact of changes in GRE scores by major on the earnings of successive cohorts of college graduates. Three paths: 1) only the GRE verbal score changes, holding the quantitative score constant at its sample mean; 2) only the GRE quantitative score changes, holding the verbal score constant at its sample mean; and 3) both scores are allowed to change.

When only the verbal score changes the path of the simulated earnings shows the opposite pattern of the trend of the verbal score since the coefficient of the verbal score is negative. The simulated earnings decrease by 5 percent from 1963 to 1975, and then increase by 4.2 percent thereafter. Overall, changes in GRE verbal scores account for a very slight decrease in real earnings over the 1963 – 1993 period. The simulated earnings path of the quantitative score shows almost the same pattern of GRE quantitative score in figure 5 since the coefficient is positive. The earnings increase by 5 percent from 45,164 dollars in 1963 to 47,378 dollars in 1975. It decreases by 4.2 percent, and increases again by 3 percent thereafter. For the whole period, the increase of projected earnings due to the quantitative score change is about 1,600 dollars or 3.6 percent. Lastly, the simulated earnings when the both scores change increases steadily for the period. It increases moderately from 1963 to 1983, dipped down for next couple of years, and increase quite strongly after then. The total increase of simulated earnings from 1963 to 1993 due to the changes of both scores is 1,205 dollars; about 2.6 percent.

Figure 6 shows the impact of changes in choice of college major on earnings by bachelor's degree cohort, holding all other variables constant at their sample means. We divide the sample period into six four-year sub periods; from 1963 to 1966, 1967 to 1970, 1971 to 1974, 1975 to 1978, 1979 to 1982, 1983 to 1986, and from 1987 to 1992. We calculate the mean distribution of college majors for each sub period. This grouping is inevitable to have noticeable variations in choice of college majors since there is not enough observations for some college majors in a year.

Over the whole sample period, 1.8 percent of earnings change is attributed to the changes in the distribution of college majors. The simulated earnings trend shows that it decreases in 1960's and early 1970's by about 1.1 percent from 46,776 to 46,253 dollars, stabilizes in early 1970's and mid 1970's, and increases after then by 3.1 percent from mid 1970's to mid 1980's. It dropped slightly after then, about 0.2 percent.

One interesting aspect of this trend is that it coincides with the Freeman's (1976) findings of low returns to college and graduate degrees in the late 1960's and the early 1970's. He attributed the earnings decline to increasing supply of college graduates associated with the maturation of the baby boom. Our result reinforces Freeman's findings. In addition to substantial increase in the supply of college graduates in the labor market, college graduates atypically selected majors with relatively low returns in the late 60's and the early 70's. Graduates moved to majors with higher returns after that and the earnings change due to the distribution change in college major stabilizes in the late 80's and the early 90's.

VI. Conclusions

This study investigated three research questions on the earnings function: returns to post college degrees, impact of cognitive skills by majors on earnings, and ability bias in OLS estimation.

We found that OLS estimates of returns to graduate education are underestimated when controls for average ability are missing. Comparing returns to graduate or professional degrees of the equation (15) that is not controlled for average ability and of the equation (18) that is controlled for average ability with GRE verbal and quantitative scores and college major dummies, coefficients of higher degree dummies in the equation (18) are larger than those in the equation (15). This result differs from findings in previous studies that looked at lower levels of education. Less educated students with more math ability may stay longer in school because returns to cognitive skills and returns to additional schooling are relatively larger than the opportunity cost of schooling. After college, the opportunity cost for post college education increases more rapidly than returns to schooling. For those who have good cognitive skills, the opportunity cost is even higher since the skills are highly valued in the labor market.

From the estimation results of the equation (18) in the Table 5, Master's degree holder are paid about 12 percent, Ph.D. holders are paid about 25 percent, and professional degree holders are paid about 55 percent higher than those who have only bachelor's degree. This could be an explanation of increasing earnings inequality among different education groups. This result also suggests that we might need to put more attention on higher education. For example, if you compared earnings of high school graduates with those of college graduates that in fact include the post college educated as well, you might end up overstating the value of a college degree.

Average GRE verbal and quantitative scores by college major play an important role in explaining earnings variation. Coefficient on the GRE quantitative score is positive and that on GRE verbal score is negative. This could be the evidence that quantitative skills are increasing in value in the labor market while the value of verbal skills is steady or falling. We confirm this conclusion in the simulation results from Figure 6. The simulated earnings differential attributed to changes in the average GRE quantitative and verbal scores by college major is about 2.6 percent. Since the impact of the verbal score change on earnings is trivial for the sample period as a whole, the 2.6 percent earnings differential is due to changes in the quantitative skills only, which is about 1,609 dollars. Considering it in real terms, the impact of the quantitative score is substantial.

Changes in the distribution of major on earnings by cohort also have had an impact on the earnings for college graduates. Researches have found that college majors play an important role in explaining wage variation. Some majors get paid more than others. The simulated earnings in Figure 6 decrease in 60's and early 70's, stabilize in 70's, and increase thereafter. This supports the findings of Freeman (1976). He found that low returns to college and graduate degrees in late 60's and early 70's and argued that the reason for this decline is supply changes in the labor market for the period. The simulation result shows that, as well as the increase in the supply of college graduates in labor market, relatively over populated college graduates whose returns to major are low could be the reason for his finding in late 60's and early 70's.

References

Becker, Gary S., *Human Capital* , 1975, second edition, Chicago, University of Chicago Press.

Berger, Mark C., "Predicted Future Earnings and Choice of College Major," *Industrial and Labor Relations Review*, 1988a, vol. 41, no. 3, 418-429.

Berger, Mark C., "Cohort Size Effects on Earnings: Differences by College Major," *Economics of Education Review*, 1988b, Vol. 7. no. 4, 375-383.

Bishop, John H., "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review*, 1989, vol. 79, no. 1, 178-197.

Bishop, John H., "Achievement, Test Scores, and Relative Wages," in *Workers and their Wages: Changing Patterns in the United States*, ed. Marvin H. Kosters, 1991, Washington, D.C.: The AEI Press, 146-190.

Boissiere, M., Knight, J. B., and R. H. Sabot, "Earnings, Schooling, Ability, and Cognitive Skills," *American Economic Review*, 1985, vol. 75, no. 5, 1016-1030.

Blackburn, McKinley L. and David Neumark, "Omitted-Ability Bias and the Increase in the Returns to Schooling," *Journal of Labor Economics*, 1993, vol. 11, no. 3, 521-544.

Blackburn, McKinley L., Bloom, David E., and Richard B. Freeman, "The Declining Economic Position of Less-Skilled American Men," in *A Future of Lousy Jobs? The Changing Structure of U.S. Wages*, edited by Gary Burtless. Washington, D.C., Brookings Institution, 1990.

Card, David and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men?" *Quarterly Journal of Economics*, 2001, vol. 116, no. 2, 705-746.

Card, David and Alan B. Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 1992, vol. 100, 1-40.

Card, David, "Earnings, Schooling, and Ability Revisited," *Research in Labor Economics*, 1995, vol. 14, page 23-48.

Card, David, "The Causal Effect of Education on Earnings," 1999, *Handbook of Labor Economics*, vol. 3, Edited by O. Ashenfelter and D. Card.

Freeman, Richard B., *The Over-Educated American*, 1976, New York, Academic Press.

Glewwe, Paul, “Schooling and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes,” *Journal of Economic Literature*, 2002, vol. 40, no. 2, 436-482.

Griliches, Zvi, “Estimating the Returns to Schooling: Some Econometric Problems,” *Econometrica*, 1977, vol. 45, no. 1, 1-22.

Grogger, Jeff and Eric Eide, “Changes in College Skills and the Rise in the College Wage Premium,” *Journal of Human Resources*, 1995, vol. 30, no. 2, 280-310.

Heckman, James J., and Solomon Polachek, “Empirical Evidence in the Functional Form of the Earnings-Schooling Relationship,” *Journal of the American Statistical Association*, 1974, vol. 69, 350-354.

James, Estelle, and Alsalam, Nabeel, Conaty, Joseph C., and Duc-Le To, “College Quality and Future Earnings: Where Should You Send Your Child to College?” *American Economic Review*, 1989, vol. 79, no. 2, 247-252.

Juhn, Chinhui, Kevin, M. Murphy, and Brooks Pierce, “Wage Inequality and the Rise in Returns to Skill,” *Journal of Political Economy*, 1993, vol. 101, no. 3, 411-442.

Katz, Lawrence W., and Kevin M. Murphy, “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *Quarterly Journal of Economics*, 1992, vol. 107, no. 1, 35-78.

Katz, Lawrence F., and Ana L. Revenga, “Changes in the Structure of Wages: The United States vs. Japan,” *Journal of the Japanese and International Economies*, 1989, vol. 3, 522-553.

Levy, Frank and Richard J. Murnane, “U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations,” *Journal of Economic Literature*, 1992, vol. 30, no. 3, 1333-1381.

Levine, Phillip B. and David J. Zimmerman, “The Benefit of Additional High-School Math and Science Classes for Young Men and Women,” *Journal of Business and Statistics*, 1995, vol. 13, no. 2, 137-149.

Mincer, *Jacob, Schooling, Experience and Earnings*, 1974, Columbia University Press, New York.

Mitra, Aparna, “Cognitive Skills and Black-White Wage in the United States Labor Market,” 2000, *Journal of Socio-economics* 29, 389-401.

Moll, Peter G., “Primary Schooling, Cognitive Skills and Wages in South Africa,” *Economica*, 1998, vol. 65, 263-284.

Murnane, Richard J., Willett, John B., and Frank Levy, "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics*, 1995, vol. 77, no. 2, 251-266.

Murnane, Richard J., Willett, John B., Duhaldebore, Yves, and John H. Tyler, "How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?" 2000, *Journal of Policy Analysis and Management*, vol. 19, no. 4, 547-568.

Murphy, Kevin M., and Finis Welch, "Wage Premiums for College Graduates: Recent Growth and Possible Explanations," *Educational Researcher*, 1989, vol. 18, no. 1, 17-26.

Murphy, Kevin M., and Finis Welch, "The Structure of Wage," *Quarterly Journal of Economics*, 1992, vol. 107, no. 1, 285-326.

Park, Jin Heum, "Returns to Schooling" A Peculiar Deviation from Linearity," 1994, Working paper no. 335 (Industrial Relation Section, Princeton University).

Rumberger, Russell W. and Scott L. Thomas, "The Economic Returns to College Major, Quality and Performance: A Multilevel Analysis of Recent Graduates," *Economics of Education Review*, 1993, vol. 12, no. 1, 1-19.

Rumberger, Russell W., "The Changing Economic Benefits of College Graduates," *Economics of Education Review*, 1984, vol. 3, no. 1, 3-11.

Table 1. Descriptive Statistic of the Data: 1963-1986, observation number: 67565

	Variables	Mean	Stand. Err.
Demographics	Age	41.2	(0.027)
	Experience	17.4	(0.025)
	Male	0.723	(0.002)
	Citizen	0.956	(0.001)
	Rural Background	0.319	(0.002)
	Postdoc	0.004	(>0.001)
Education	BA	0.549	(0.002)
	MA	0.287	(0.002)
	PHD	0.063	(0.001)
	Professional Degree	0.101	(0.001)
GRE Score	Verbal GRE Mean	503.3	(0.103)
	Quant. GRE Mean	574.9	(0.260)
Race	Hispanic	0.031	(0.001)
	White	0.849	(0.001)
	Black	0.052	(0.001)
	Asian	0.066	(0.001)
	Native American	0.002	(>0.001)
	Salary	53864	(113.3)
Selected Majors	Computer Sci.	0.049	(0.001)
	Agriculture	0.024	(0.001)
	Biology	0.127	(0.001)
	Chemistry	0.041	(0.001)
	Physics	0.026	(0.001)
	Economics	0.052	(0.001)
	Psychology	0.104	(0.001)
	Engineering	0.205	(0.002)

Table 2. Average Salary in Different Categories

	Mean	Stand. Err.
Male	57803	(135.7)
Female	43601	(182.6)
BA	47900	(161.5)
MA	53325	(208.7)
PhD	59657	(165.4)
Professional Degree	84155	(727.3)
White	54935	(133.0)
Black	44187	(364.5)
Asian	50386	(285.8)
Native Am.	45295	(1316.7)
Computer Sci.	52163	(374.0)
Agriculture	41691	(566.6)
Biology	57673	(365.8)
Chemistry	63337	(472.6)
Physics	58968	(527.6)
Economics	59886	(780.8)
Psychology	46937	(388.4)
Engineering	59753	(193.6)

Table 3. Average GRE Verbal Score in Different Categories

	Mean	Stand. Err.
Male	501.1	(0.12)
Female	509.1	(0.19)
BA	500.9	(0.17)
MA	502.5	(0.23)
PhD	508.4	(0.17)
Professional Degree	515.5	(0.32)
White	503.7	(0.12)
Black	505.1	(0.36)
Asian	497.2	(0.29)
Native American	507.0	(1.22)

Table 4. Average GRE Quantitative Score in Different Categories

	Mean	Stand. Err.
Male	585.3	(0.30)
Female	547.8	(0.48)
BA	581.9	(0.44)
MA	568.7	(0.57)
PhD	573.0	(0.41)
Professional Degree	555.7	(0.86)
White	573.9	(0.30)
Black	553.2	(1.04)
Asian	604.0	(0.67)
Native American	563.2	(3.35)

Table 5. OLS Regression coefficients on Log Annual Salary

	15	16	17	18
Variables	Coefficient	Coefficient	Coefficient	Coefficient
MA	0.108****	0.116***	0.117***	0.117***
PhD	0.226***	0.240***	0.253***	0.253***
Professional Degree	0.498***	0.551***	0.552***	0.552***
Experience/10	0.255***	0.361***	0.410***	0.406***
(Experience/10)^2	-0.039***	-0.054***	-0.075***	-0.074***
Male	0.248***	0.175***	0.161***	0.161***
Citizen	0.075***	0.102***	0.105***	0.104***
Postdoc	-0.403***	-0.386***	-0.375***	-0.373***
Verbal Score/100		-0.122***		-0.176***
Verbal Standard Deviation/100		0.031***		0.008
Quantitative Score/100		0.151***		0.105***
Quantitative Standard Deviation/100		-0.026***		-0.006
Ratio of Foreign Graduate Students		-0.033		-0.259***

Table 5. (cont'd)

Computer Sci.			0.087	-0.030
Mathematics			-0.060	-0.153
Agriculture			-0.319***	-0.346***
Biology			-0.156*	-0.206**
Chemistry			-0.030	0.005
Earth Sci.			-0.135	-0.191*
Physics			-0.046	-0.110
Economics			-0.028	-0.057
Political Sci.			-0.144	-0.101
Psychology			-0.205**	-0.205**
Sociology			-0.273***	-0.263***
Social Work			-0.308***	-0.315***
Linguistics			-0.340***	-0.232**
Philosophy			-0.348***	-0.309***
Anthropology			-0.202**	-0.114
History			-0.250***	-0.186*
Education			-0.245***	-0.292***
Business			-0.028	-0.092
Public Administration (dropped)				
Literature			-0.175*	-0.045
Foreign Literature			-0.171*	-0.140
Music			-0.320***	-0.328***
Arts			-0.260***	-0.218**
Communication			-0.218**	-0.196*
Architecture			-0.040	-0.1011
Health & Medical Sci.			-0.106	-0.132
Other Majors			-0.240***	-0.245***
Engineering			0.086	-0.065
Hispanic	-0.058***	-0.053***	-0.052***	-0.053***
Black	-0.121***	-0.097***	-0.095***	-0.095***
Asian	-0.036***	-0.080***	-0.087***	-0.088***
Native American	-0.174***	-0.153***	-0.150***	-0.151***
Constant	10.12***	9.752***	10.093***	10.476***
R^2	0.183	0.224	0.237	0.237

* : significant at 10 % significance level
** : significant at 5 % significance level
***: significant at 1 % significance level

Table 6. GRE Score Share of Degrees

Verbal Score

	BA	MA	PhD	Professional
Low	33.1%	31.3%	20.1%	5.7%
Middle	62.4%	61.9%	70.8%	90.7%
High	4.6%	6.8%	9.2%	3.6%

Quantitative Score

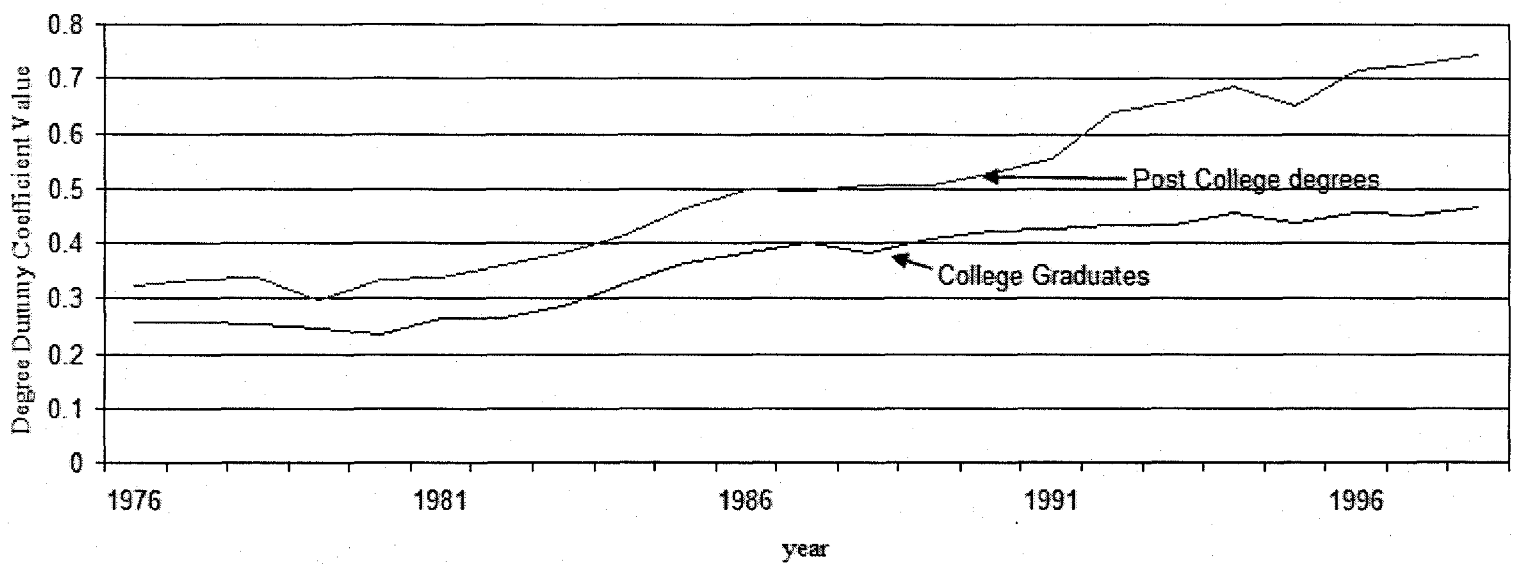
	BA	MA	PhD	Professional
Low	19.7%	28.6%	24.2%	25.4%
Middle	45.9%	42.6%	47.7%	63.2%
High	34.4%	28.8%	28.0%	11.4%

Table 7. Marginal Effect of Individual Ability on Probability to Pursue Advanced Degrees

Dependent Variable	Marginal Effect	Std. Err.
BA	0.011	(733020)
MA	0.035	(637883)
Ph.D.	-0.003	(173405)
Professional Degree	-0.043	(540997)

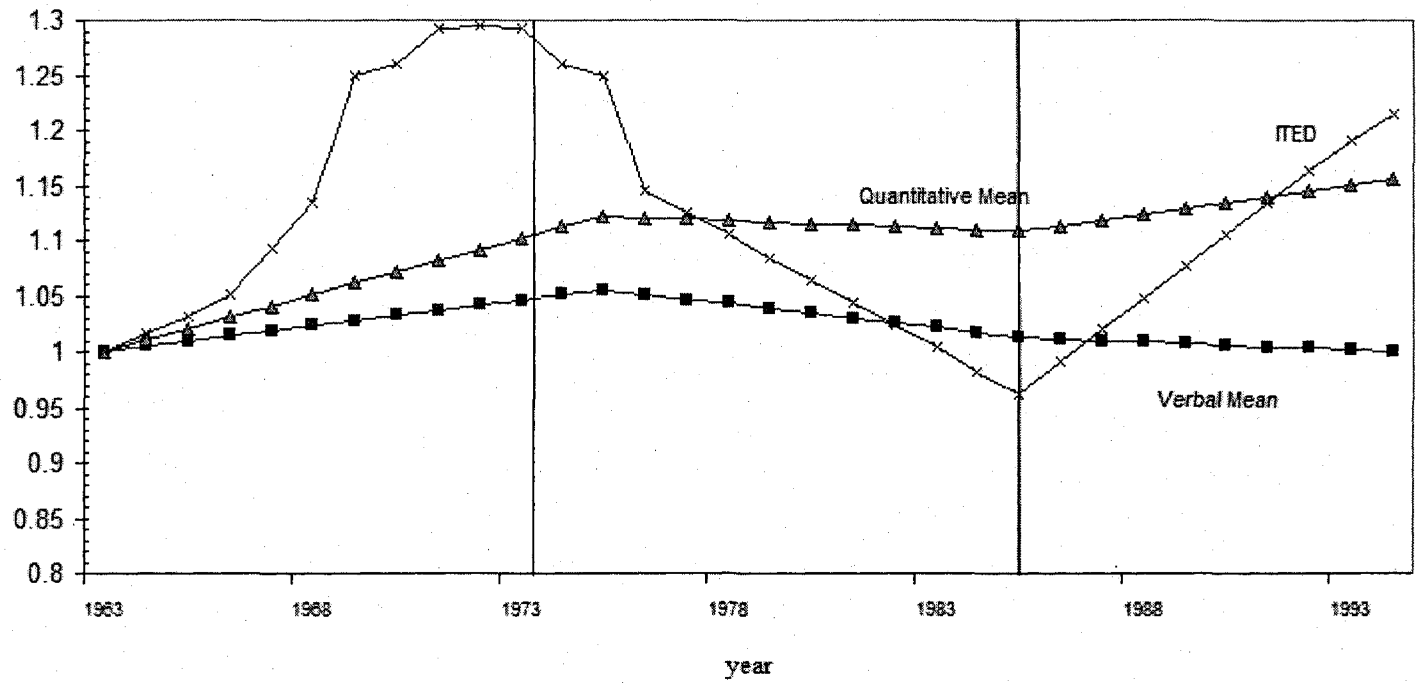
Estimates are scaled up by 100,000,000

Figure 1. Returns to Schooling Relative to High School Graduates: 1976-1998



Notes: Values based on coefficients from annual regressions of log weekly wage on a vector of dummy variables indicating educational attainment, age, age squared, and a vector of demographic dummy variables. Data taken from the March Current Population Survey (1976-1998).

Figure 2. Trends of GRE Verbal and Quantitative Scores, and ITED 12th Grade Score:1963-1994 (1963 normalized to 1)



Note: ITED: 12th grade Iowa Test of Educational Development, is taken from Bishop (1989). Because the test is taken in the senior year of high school, five years is added to match the high school data with the presumed year the same student would be taking the GRE.

Figure 3. Ratio of Foreign Students to Total Doctorate Degree Recipients in the U.S.: 1963-1994

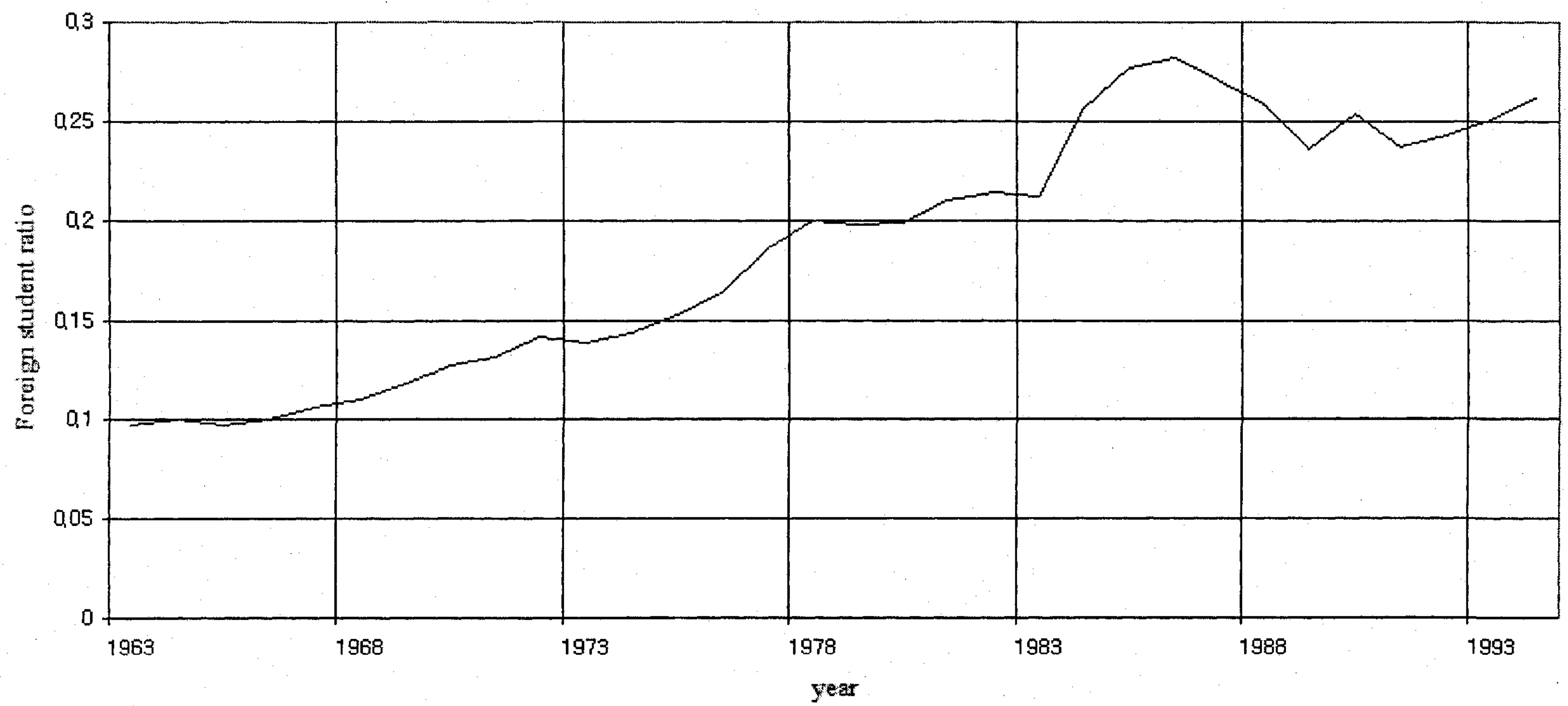


Figure 4. Ratio of Foreign to Total Doctorate Degree Recipients by Major

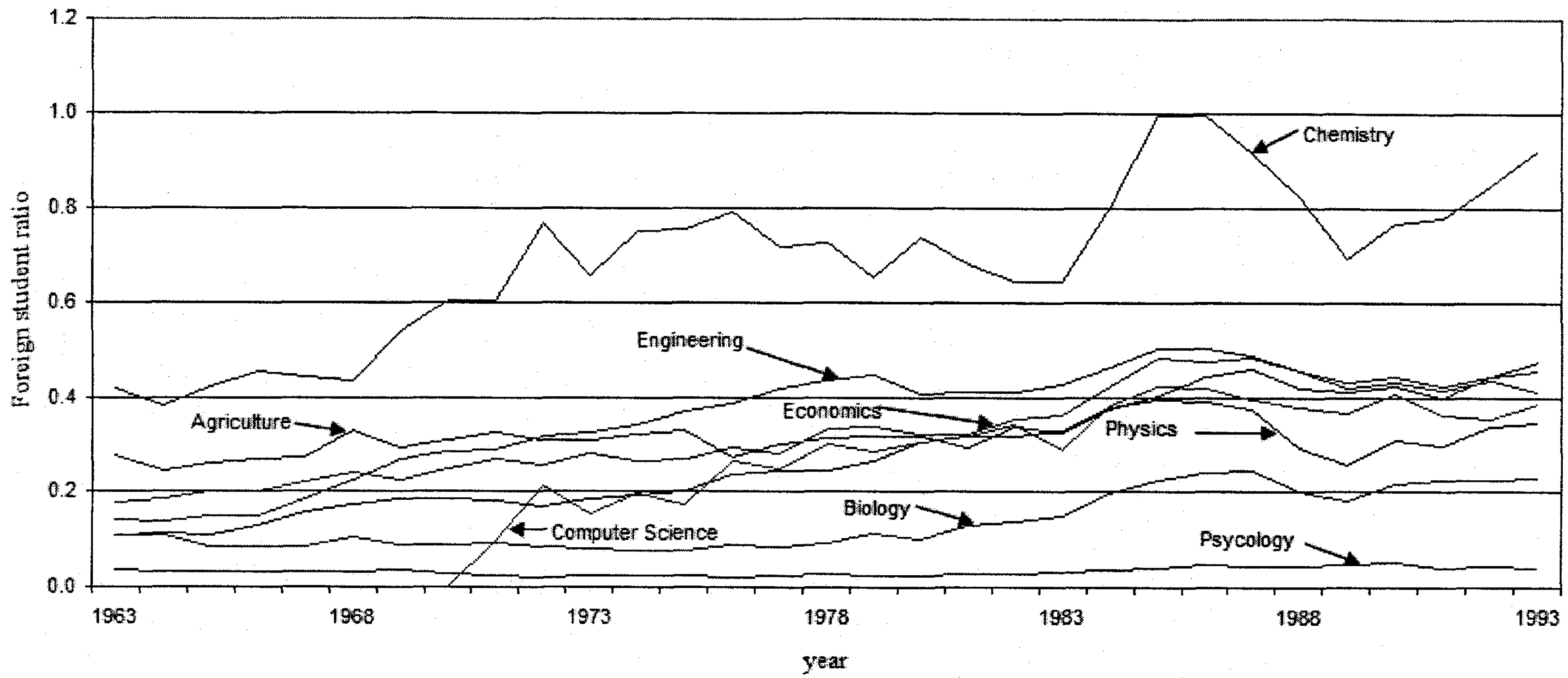


Figure 5. Residual Variation of GRE Scores

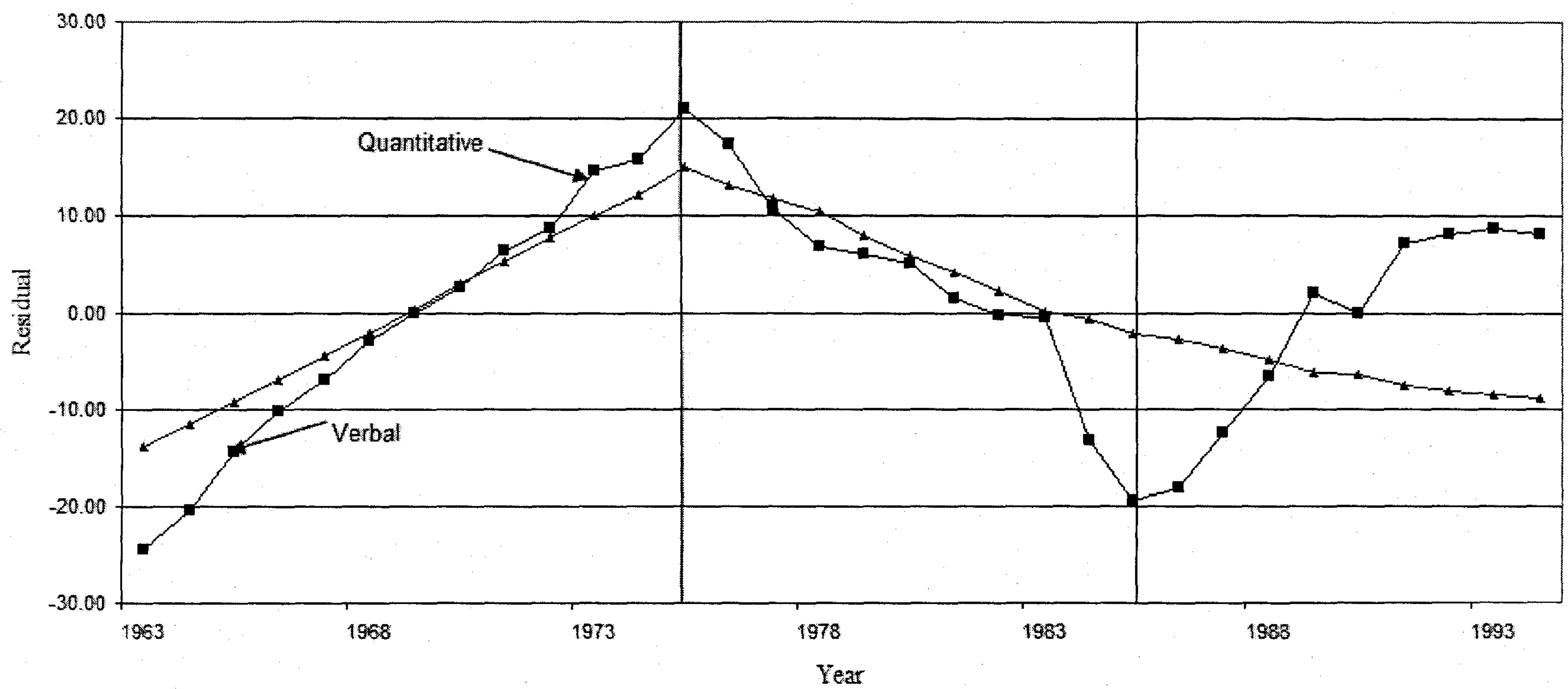


Figure 6. Simulated Impact of Average GRE Scores on Earnings by BA Cohort

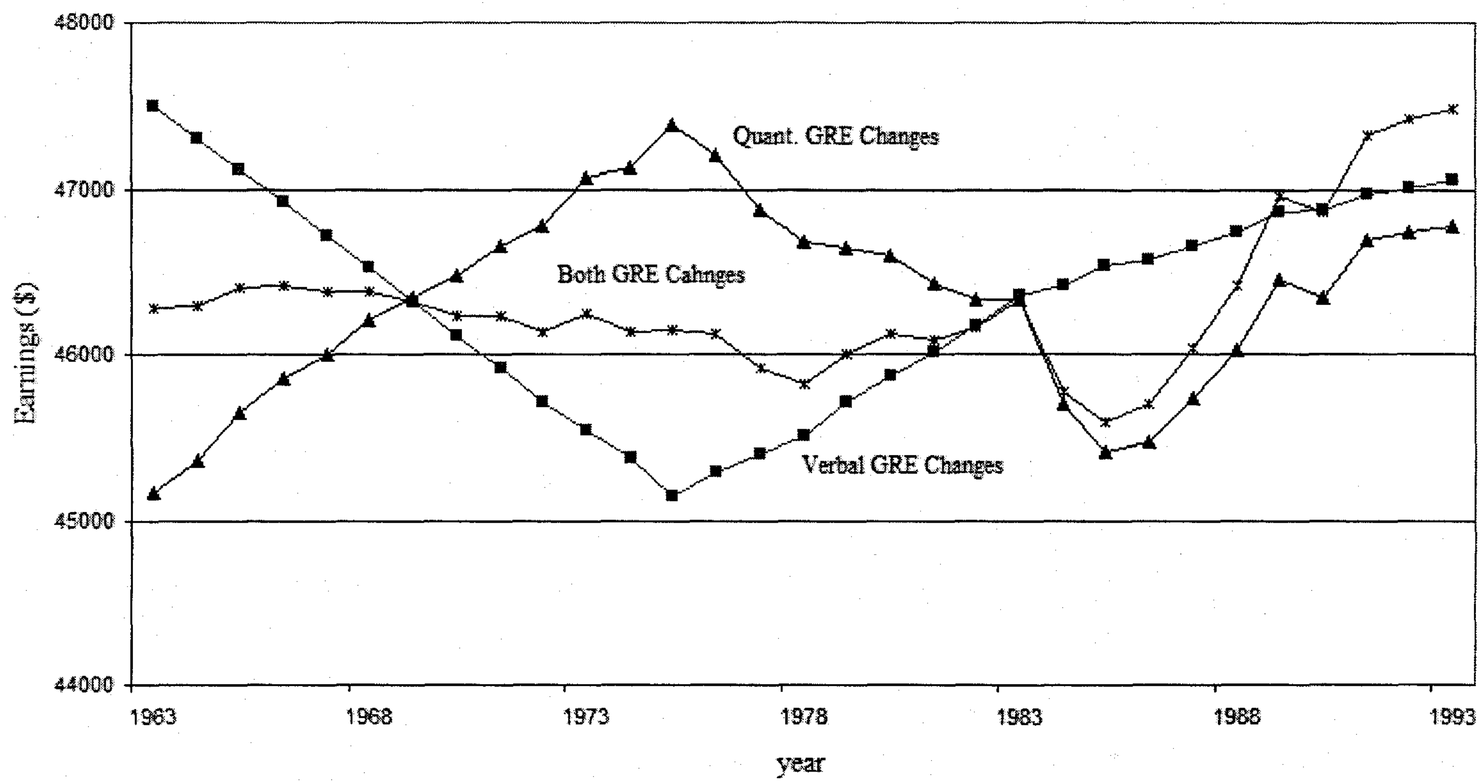
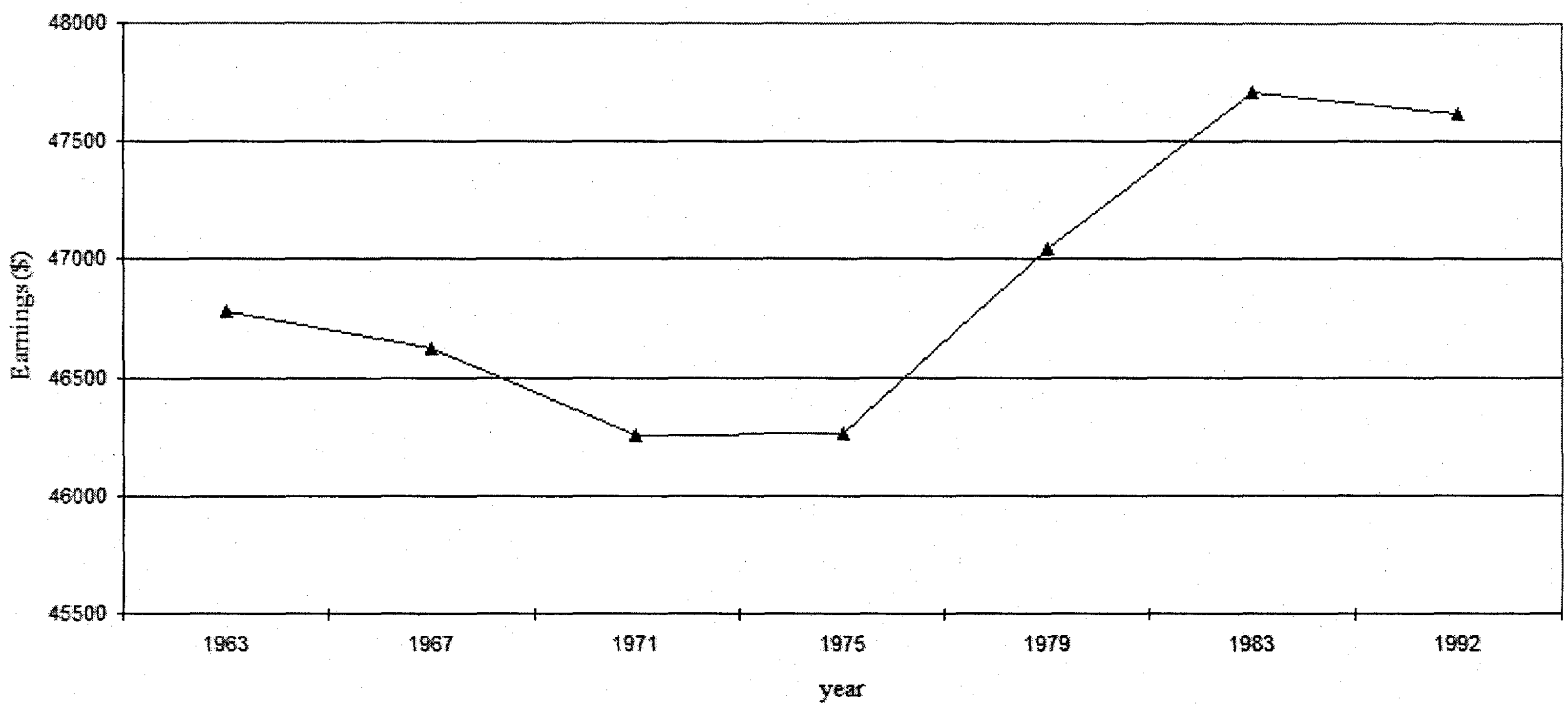


Figure 7. Simulated Impact of Changes in Distribution of College Majors on Earnings by BA Cohort



[1] This increasing trend of college wage premium gets stronger in 1990's. In addition, wage premium for higher degrees increases more steeply for last twenty years. See Figure 1, using March Current Population Survey, for estimated returns to schooling for college graduates and the post college degrees relative to high school graduates between 1976-1998.

[2] In this paper, they argue that, for supply shift, 1970's entry to the labor market of the large and relatively well educated baby boom cohorts. However, they mentioned that supply shift itself can not explain the trends in earned income inequality. A gradual increase in the demand for more skilled workers is also necessary.

[3] The definition or the notion of ability is still arguable. In this paper we use cognitive skills as an indicator of ability and use GRE verbal and quantitative test scores as a measure of cognitive skill. See Griliches (1977) for more discussion on the notion of ability.

[4] For in depth discussion of ability bias, see Blackburn and Neumark (1993).

[5] The magnitude of the impact of the cognitive skills on wage is about a third of that of years of schooling on wages in 1972 and about a half in 1986.

[6] Most studies focus on math and reading skills and use their test scores, some studies use more than math and reading tests. For example, Grogger and Eide (1995) used "mosaic" test that measures perceptual speed and accuracy.

[7] Paglin and Rufolo (1990) used GRE scores in their wage equation estimates. However their wage equation differ from typical earnings functions, so their results are not directly comparable to other studies'. Nevertheless, they also found that quantitative scores had a larger impact on wages than verbal scores.

[8] See Murnane, Willett, and Levy (1995) and Grogger and Eide (1995) for example.

[9] Their models are not different in nature. However, Becker's model allows heterogeneous individual ability to affect the decision on optimal schooling, so workers of higher ability receive higher returns to their human capital. Mincer assumes that workers have same opportunity for human capital investment, and so the returns to human capital are also identical for all workers.

[10] It also can be shown that Human Capital Earnings function can be derived from economic structure. Following Becker (1967), Card (1995) developed an analytically tractable model of the decision on optimal schooling and its implications for the earnings function.

[11] To estimate this equation correctly, we need to assume that schooling variable is exogenous. That is, $E(S,u) = 0$.

[12] We follow the model attributed to Griliches (1977). For more in depth discussion of the human capital function and ability bias, see Card (1999).

---

[13] See Glewwe (2002) for the in-depth discussion of the relation of the optimal schooling and cognitive skills.

[14] *GRE* also includes standard deviations of GRE scores by major and the ratio of foreign graduate students. These variables are prepared to attenuate possible measurement errors in GRE scores. This will be discussed in the data section.

[15] Moll (1998) had same results on experience in earnings function using the South Africa data.

[16] For example, Murnane et al. (1995) using only math test scores found its positive impact on earnings. Blackburn and Neumark (1995), Murnane et al. (2000), and Mitra (2000) using both measures of lower education levels of math and verbal abilities found that the sign of the estimates of the math coefficients are positive and significant, but the sign of the verbal coefficients are either negative or insignificant or both. Grogger and Eide (1995) using measures of the high school level of both math and verbal test scores and college major dummies found that the sign of the verbal test is negative for male, but positive for female. The Sign of math score was positive for both male and female, but it was only significant for female. Interesting result of Grogger and Eide is when they estimate earnings function with both major dummies and cognitive measures, performances of the test scores decreased.

[17] For this table, we subtracted the lowest score from the highest score and divided it by three for each verbal and quantitative score.

[18] The simulation has been done up to 1993 since the scores are available up to that point. However the sample means are calculated over 1963 to 1986, which is the sample period.

CHAPTER 3. THE ROLE OF MATHEMATICAL AND VERBAL SKILLS ON CHOICE OF AND RETURNS TO GRADUATE AND PROFESSIONAL EDUCATION

I. Background

A wealth of economic research has documented an increase in the returns to education in the 1980s. Most of this research has concentrated on the relative returns to a bachelor's degree relative to lower levels of education. Since the 1980s, there has been a well-documented increase in returns to a college education relative to lower levels of schooling. The trend in relative earnings for bachelor's degree holders relative to high school graduates between 1976 and 1998 is illustrated in Figure 1. The bachelor's degree premium over a high school degree rose from 25% in 1976 to 45% in 1998 with the gains beginning in the early 1980s. Not as commonly known is that returns for those who entered or completed some post graduate training rose in a parallel fashion through the 1980s, and then began to rise even more rapidly than did returns to bachelor's degrees in the 1990s. Over the period, the premium earned by those with graduate degrees relative to bachelor's degree recipients rose from 32% to 67%.

This study has two objectives. The first is to measure the returns to post graduate training, controlling for likely joint choices of years of schooling and their associated returns. The second objective is to determine if the rise in returns to post graduate training can be explained by changes in the quality of more recent cohorts of graduate students relative to their older colleagues or if we need to seek other explanations for the rising returns to graduate education.

Skill-biased technological change is believed to have progressively raised returns to college graduates since the 1970s. Given that graduate training is a heavy user of the information technologies believed to be a major source of technological innovations, one would expect that technological factors should have had a similar, if not a stronger, impact on post-graduate earnings as on bachelor's degree earnings. The rising graduate degree premium over the bachelor's degree premium in the 1990s might be a signal that graduate training has particularly benefited by skill-biased technical changes, although one might then have expected the premium to have risen earlier in the information technology adoption process.

To assess the role of technological change in explaining rising returns to graduate training, we focus on the role of quantitative skills on observed returns. Several studies have documented changes in the returns to quantitative skills in the 1980s. Murnane, Willett and Levy (1995) found that rising returns to mathematics skills can explain a substantial fraction of the observed increase in returns to college between 1978 and 1986. The effect was stronger for women than for men. Grogger and Eide (1995) and Levine and Zimmerman (1995) also reported that standardized mathematics scores or having taken more mathematics classes had a significant positive impact on women's wages but not men's wages.

The mechanism by which mathematical skills influence wages is not clear. It is likely that stronger quantitative skills are complementary with the use of information technologies that are widely suspected to have raised worker productivity and wages. However, quantitative skills may also affect the type of training individuals receive. Willis and Rosen (1979), Murnane, Willett and Levy (1995) and Taber (2001) all found that stronger mathematical skills in high school increased the likelihood of attending college. Paglin and Rufolo (1990) found that quantitative skills influenced choice of graduate major.

There is a presumption that quantitative and verbal skills increase in importance as the education level rises, and so changes in the value of these skills would be expected to affect the market for post-graduate training as well. Our review of the literature that concentrated on lower levels of education suggests that two effects are potentially at work:

1) Rising returns to cognitive skills may have increased the opportunity costs of attending graduate school, limiting incentives to pursue post-graduate education in the areas where the returns are rising the most rapidly. Consequently, the most able students opt not to pursue graduate education in favor of capturing returns to those skills in jobs they can acquire with a bachelor's degree.
2) The marginal product of cognitive skills may have risen atypically in post-graduate t raining, raising the returns to graduate training relative to lower education levels.

These two possibilities would have opposite effects on incentives to attend graduate school and on observed wages. The former would suggest that the observed wage differentials between graduate and undergraduate degree holders would understate the true returns to graduate education because the earnings of those stopping at the bachelor's degree exceed the opportunity costs of those who attended graduate school. The latter would suggest the most able would attend graduate school, suggesting that the observed wage differential between graduate and undergraduate degree holders is an upward biased measure of the returns to graduate school. The latter argument would also potentially explain why we see rising graduate degree premia in the 1990s relative to earnings at the bachelor's degree level.

Even before we examine why returns to graduate training may have changed, we must document the returns to that training. There are many studies that examine incentives to enter individual majors and the returns to those decisions. However, more general studies of returns to graduate education are rare.[1] The main advantage to a general study of returns to graduate education for our purposes is that if one of the phenomena we wish to examine is how quantitative skills sort individuals across degrees, we need to have the sample cover the universe of students and not just a specific field or major. In addition, it is easier to compare estimated returns to an education level to the literature on returns to high school or college that do not distinguish by field than it is to compare returns to a specific graduate degree in, say, law or sociology.

A frequent challenge for studies measuring returns to schooling is that individuals are not randomly assigned to different schooling levels. If schooling choices are driven by individual comparative advantage, then the returns will reflect, at least in part, that nonrandom sorting of individuals across education levels. A large literature has developed assessing the impact on measured returns to schooling of various procedures aimed at controlling nonrandom sorting across school levels. A common tactic has been to use measures of parental education or other family background measures as instruments for education levels. Other instruments have included distance to school or other measures of school costs. As reviewed by Card (1999), these studies routinely obtain higher estimated returns to schooling when employing instrumental variables than they obtained using ordinary least squares.

However, these instruments are often challenged. For example, commonly used family background variables (Willis and Rosen (1979), Altonji and Dunn (1996), Deschenes

(2002)) may be correlated with unmeasured ability, rendering them invalid.[2] Another large body of research has utilized data on twins to better control for unmeasured ability. Interestingly, results based on twins are similar to the findings reported from instrumental variables using family background or school attributes as instruments. While this does not validate the use of instrumental variables, it at least suggests that instrumental variables can approximate the results obtained from presumably better controls for missing ability.

Sections II discusses the estimation strategy and III discusses the data. Empirical results are reported in Section IV, and Section V reviews the study's conclusions.

II. Estimation Model

Our analysis begins with the standard log-earnings framework:

1) $$\ln y_i = S_i\beta_S + X_i\,\beta_X + \mu_i\,\beta_\mu + u_i,$$

where $\ln y_i$ is the observed earnings of the i^{th} individual; S_i is the observed schooling level, taken as a vector of dummy variables with the value of one indicating the individual's highest degree earned; X_i is a vector of individual characteristics; μ_i is an individual-specific ability component that influences earnings; and u_i is a random error term that is uncorrelated with S_i, X_i and μ_i. The β_s and β_x represent the estimated returns to schooling levels and individual attributes, respectively.

If μ_i is not observable by the econometrician, then (1) becomes

1′) $$ln\, y_i = S_i\,\beta_s + X_i\,\beta_x + \varepsilon_i; \quad \varepsilon_i = \mu_i\beta_\mu + u_i,$$

where the error term ε_i will include both purely random components and unmeasured individual ability. If that ability is correlated with schooling success, then exclusion of μ_i

from the estimating equation will lead to $E(S_i \varepsilon_i) = E(S_i \mu_i \beta_\mu) \neq 0$, and so the estimates of β_s and β_x will be subject to missing variables bias.

In our application, individuals decide between stopping at the bachelor's degree or continuing on for additional schooling. The choice set at the time the individual finishes undergraduate training includes four schooling levels: Bachelor's, Master's, Doctorate and Professional degree (mainly law or medicine). These choices are denoted respectively by subscripts B, M, D, and P. For simplicity, we consider these choices mutually exclusive, and so we only consider the choice of the highest degree earned. This avoids complications related to sequential educational choices.

The schooling decision involves selecting the option that maximizes utility. This can be written as $S_i = max\ (S_{Bi}, S_{Mi}, S_{Di}, S_{Pi})$ where S_{li} is the utility from schooling choice l. Although the utility levels are not observable, we can observe how the elements of S_{li} affect the probability of selecting schooling choice l.

Suppose that the individual selects schooling level S_i at least in part on the basis of expected earnings at that education level. Then the individual will use knowledge of X_i and μ_i to forecast what he expects to earn from each of the four educational choices. Suppose also that there is a vector Z_i that contains factors that shift the individual's taste for or cost of schooling choice l. Then utility from each choice S_{li} can be approximated by

(2) $$S_{li} = X_{li}\,\theta_X + Z_{li}\,\theta_Z + \mu_{li}\,\theta_\mu + v_{li} \quad ; \quad l = B, M, D, P,$$

where v_{li} may include omitted variables, measurement errors, or specification errors of functional choice, and it is assumed to be independent of observed variables.

Now, even if $E(S_i \mu_{li} \beta_{\mu}) = 0$, direct estimation of (1) will yield biased estimates if $E(v_{li} u_i) \neq 0$. This endogeneity bias is caused by the joint selection of years of schooling with the expected returns from that schooling. A large literature on returns to schooling suggests that both sources of bias, missing measures of ability and endogeneity of the schooling choice, are likely to hold, although the biases are often small. However, we cannot infer from the past literature that there would be small biases in the context of estimated returns to post graduate education. Consequently, we need to derive a mechanism to address the two potential sources of bias.

To solve the problem, we follow two strategies commonly employed in the literature. First, we use graduate school tuition, medical school tuition, and the proportion of self-supporting graduate students in the year of receipt of the bachelor's degree as measures of the anticipated cost of attending graduate or professional education. These measures are included as elements of Z_i that are believed to alter the probability of continuing in school but do not affect what individuals expect to earn after completing school.

We also included measures of parental education as elements of Z_i. Card (1999) argued that parental education might not be a legitimate instrument for years of schooling because parental education is correlated with unobserved individual ability, even if parental education does not directly affect earnings. His argument suggested that when parental education is used as an instrument for years of schooling, the estimated returns would be biased upward. We also estimated equation systems that included parental education as elements of X_i that enter both the schooling and earnings equations. Those estimates showed that estimated returns were even larger when parental education was used as an instrument, although the differences were not large. In our application, use of parental education as an instrument for years of schooling does not appear to bias the coefficients upward. We report the estimates that exclude parental education from X_i. The other estimates are available on request.

One reason our measures of parental education appear not to cause problems may be that we are able to incorporate measures of verbal and quantitative ability into equations (1) and (2) that are typically missing in other studies. Let individual ability be given by

(3) $\mu_{li} = \mu_l^M + \eta_i$

where μ_l^M is the vector of average mathematical and verbal skills associated with the individual's undergraduate major and η_i is an individual-specific ability component that does not vary in productivity across schooling levels. The η_i would not affect choice of schooling level. However, verbal and mathematical skills can have different productivities at different schooling levels. Variation in μ_l^M across majors at one point in time or across cohorts can affect the graduate school entry decision. Elements of Z_i can still serve as legitimate instruments for years of schooling provided that $E(Z_i \eta_i) = 0$.

Inserting equation (3) into equation (2), we obtain

(4) $S_{li} = X_{li}\theta_X + Z_{li}\theta_Z + (\mu_l^M + \eta_i)\theta_\mu + v_{li},$

$$= V(X_{li}, Z_{li}, \mu_l^M) + \zeta_{li} \;\; ; V(X_{li}, Z_{li}, \mu_l^M) = X_{li}\theta_X + Z_{li}\theta_Z + \mu_l^M\theta_\mu, \;\; \zeta_{li} = v_{li} + \eta_i\,\theta_\mu,$$

$l = B, M, D, P.$

Therefore an individual chooses an alternative l over B if $I_{li}^* \geq 0$ where

(5) $I_{li}^* = g(X_{li}, Z_{li}, \mu_l^M) - \omega_{li}$; $g(X_{li}, Z_{li}, \mu_l^M) = V(X_{li}, Z_{li}, \mu_l^M) - V(X_{Bi}, Z_{Bi}, \mu_B^M)$, $\omega_{lu} = v_{Bi} - v_{li}$.

The probability of an individual to choose a schooling level l over B is

$$(6)\ \Pr[I_{li}^{*} \geq 0] = \Pr[g(X_{li}, Z_{li}, \mu_l^M) - \omega_{li} \geq 0]$$

$$= \Pr[\omega_{li} \leq g(X_{li}, Z_{li}, \mu_l^M)].$$

If the ω_{li} are drawn independently from an extreme value distribution, then (4) can be estimated using multinomial logit. The parameter estimates will generate predicted probabilities that individual i will select any of the four options S_{Bi}, S_{Mi}, S_{Di}, and S_{Pi}. Three of these are inserted into (1) in place of the endogenous S_i to generate unbiased estimates of β_s under the maintained hypothesis that $E(Z_i\, v_{li}) = E(Z_i \eta_i) = 0$.

This two-step procedure is inefficient because it does not incorporate the sampling errors in the parameter estimation of the multinomial logit estimates of (4) into the estimation of the log earnings equation (1). We correct the second-stage standard errors using a bootstrapping procedure in which the two-step estimation was replicated 100 times, sampling with replacement, and sampling variation in the resulting estimates used to compute the second-stage standard errors.

If major-specific skills at the bachelor's degree level are increasing in market value, then they will tend to lower incentives to pursue graduate work in that field. Conversely, majors whose skills are falling in value at the bachelor's level will have disproportionately high numbers of graduate students. If this sorting effect drives lower earning bachelor's degree recipients into graduate school and drives higher earning bachelor's degree recipients out of graduate school, it would tend to depress estimated returns to graduate work. If true, then least squares estimates of the returns to graduate school that ignored the role of major-specific ability measures would tend to understate the true returns. Our empirical work provides evidence consistent with this sorting story.

III. Data

The primary data source for this study is the Scientist and Engineer Statistics Data System (SESTAT) collected by the National Science Foundation (NSF). The 1993 wave of SESTAT also incorporated the 1993 National Survey of College Graduates, a once-per-decade survey that also covered fields outside of the sciences and engineering. The universe for the 1993 SESTAT was approximately 29 million individuals who received a bachelor's degree between 1939 and 1992. Our working sample included 67,565 individuals who received a bachelor's degree between 1963 and 1986. The 1963 limit was necessitated by the lack of information on Graduate Records Exam (GRE) scores by major before 1963. The 1986 limit was imposed because we needed to give bachelor's degree recipients sufficient time to enter and complete higher degrees. Through the use of sample weights, our subsample is representative of the population of all bachelor's degree recipients in the United States between 1963 and 1986.

Table 1 includes summary statistics on the variables included in the analysis. The dependent variables include the natural logarithm of annual salary in 1993 and a series of dummy variables indicating highest degree earned. Earnings of all college graduates in 1993 averaged just under $54,000. Bachelor's recipients averaged $48,000 while Master's recipients averaged $53,000, Ph.D.s averaged $60,000 and those with professional degrees averaged $84,000. Fifty-five percent of the college graduate population did not earn a degree beyond the bachelor's level. Twenty-nine percent had a Master's degree, 10 percent held professional degrees, and 6 percent had doctorates.

Variables included in the demographic vector X_i are potential work experience (1993 – graduation year of highest degree), gender, citizenship, and racial and ethnic dummy variables. The vector Z_i includes average real medical school and graduate school tuition, and the percentage of self-supporting graduate students for the year the individual received the first undergraduate diploma. Data on tuition and availability of graduate support were collected from the National Center for Education Statistics. Higher tuition levels should lower the probability of pursuing a graduate or professional degree. The percentage of graduate students who are self-supporting indicates a lower probability of obtaining a graduate assistantship or fellowship at the time the individual received the bachelor's degree. We also included information on whether the individual was raised in a rural area and the education levels of the individual's parents as reported in SESTAT. These measures are presumed to proxy tastes for graduate education: individuals from more educated households or from more cosmopolitan settings are expected to have stronger taste for graduate training.

Measures of X_{li} include a vector of dummy variables indicating bachelor's degree major. We also know the year of graduation. This allows us to append information on the average GRE mathematics and verbal score for the college major in the year of graduation.[3] The GRE scores are used to approximate the skill content of the major. These measures are not fixed over time, as can be seen in Figure 2. Average verbal scores rose until 1975 and then fell thereafter. Average quantitative scores rose about 12 percent until 1975, retreated slightly over the next ten years, and then resumed modest growth.

These changes may reflect changes in the composition of foreign graduate students taking the GRE. We computed the proportion of foreign doctorate recipients by major for each year in the sample period, using data from the Survey of Earned Doctorates. We then regressed the GRE scores by major on the proportion of foreign doctoral graduates in the major six year earlier.[4] The residual represents changes in the skill content of college graduates holding fixed the proportion of foreign test takers. These corrected GRE time paths are also shown in Figure 2. The corrected verbal GRE path is very similar to the uncorrected path. However, the corrected quantitative GRE path shows a much steeper decline in average scores after 1975 and a much steeper rebound after 1986.[5] The time series of average GRE scores does not demonstrate a systematic improvement in the quality of GRE test takers over time, suggesting that rising quality of graduate degree holders is not the explanation for the pattern of rising returns to graduate school in Figure 1.

The GRE scores also varied across majors, genders, races, and education levels. This variation provides cross-sectional variation in the skill content of bachelor's degree recipients. As shown in Table 2, students whose highest degrees were at the bachelor's level were in majors with the highest quantitative scores and the lowest verbal scores. This is consistent with the speculation that the sorting into graduate school may be based in part on cognitive skill content of majors as proxied by GRE scores. Undergraduate majors in the sciences and engineering had markedly higher average quantitative scores while Engineering and Business had markedly lower average verbal scores. If returns to these skills have changed over time, there will be asymmetric changes in the relative incentives to seek post-graduate training across majors. Because demographic groups concentrate in different majors, there is cross-sectional variation in major GRE scores by race, ethnicity and gender. Men tended to be in majors with higher average quantitative GREs and marginally lower verbal GREs. Asians also concentrate in majors with high quantitative and low verbal scores.

Together, the time series and cross-sectional variation in GRE scores should be sufficiently large to assess whether changes in cognitive skills developed in undergraduate programs have a role in explaining changes in the returns to post-graduate education in the United States. We proceed to that exercise in the next section.

IV. Estimation Results

A. Schooling Choices

Our primary interest is in deriving estimates of equation (1), but we also have an interest in assessing how bachelor degree recipients decide to continue on in school. Results from the weighted multinomial logit estimation of the schooling choice equation are reported in Table 3. The estimation uses stopping education at the bachelor's degree as the reference group, and so positive (negative) signs suggest an increased (decreased) probability of the educational choice relative to stopping at the B.A. level.

Family background variables are highly significant in influencing the choice of whether or not to pursue and advanced degree. As mother's and father's education levels rise, the probability of seeking an advanced degree increases. The effect is strongest at the PhD level. B.A. recipients who grew up in rural areas are less likely to pursue an advanced degree. U.S. citizens are less likely to seek a Master's or doctorate but are more likely to pursue a professional degree. Asians are more likely than whites to pursue a Master's or Ph.D., while Hispanics and Blacks are less likely to pursue the doctorate.

Measures of expected cost of pursuing a graduate degree performed as expected. Individuals who received the bachelor's degree in years with higher real graduate and medical school tuition levels were less likely to pursue an advanced degree. However, the negative effect is only statistically significant for the effect of graduate school tuition on PhD or Professional degrees. The percentage of self-supporting graduate students also significantly decreased the probability of pursuing an advanced degree. We also interacted the probability of self-support with a measure of parental education with the expectation that parents with higher education levels might moderate the adverse effects of a low probability of receiving graduate support.[6] That expectation was also realized in that all signs on the interacted terms were positive, although only significant in predicting the likelihood of obtaining a Master's degree.

GRE scores have an interesting impact on the probability of pursuing a higher degree. Undergraduates in majors with higher verbal scores and lower quantitative scores are more likely to pursue the doctorate or professional degrees. The standard deviation of GRE scores in the major tend to reinforce the effects of the mean scores: higher standard deviation of GRE verbal scores raises the likelihood of pursuing the doctorate, while increasing the standard deviation of the quantitative score lowers the likelihood of pursuing the doctorate. In separate regressions, we found that the impact of the quantitative score on schooling choice has not changed over time. If returns to quantitative skills have risen over time, the impacts must have been neutral across education levels. The GRE verbal score may have gained modestly in importance over time, but the effect is much smaller than the quantitative score.

Our main results concerning the impact of changing cognitive skills on graduate school choice are illustrated in Figures 3-5, using the results from Table 3. The simulations are carried through to 1993 because all necessary information was available, although the parameter estimates are based on data just through 1986. The most dramatic changes are due to changes in the GRE quantitative score. As shown in Figure 3, the proportion of students stopping at the bachelor's degree has risen since the mid 1980s while the likelihood of seeking doctoral or professional degrees has fallen due to rising average quantitative GRE scores. The finding that the marginal impact of the GRE quantitative score does not vary across graduation cohorts suggests that this is a result of rising quantitative skills and not rising returns to those skills.

Because verbal scores raise the likelihood of seeking advanced degrees, rising GRE verbal scores in the 1960s and 1970s tended to increase the likelihood of entering graduate school. However, the erosion in verbal skills indicated by the steady decline in average GRE verbal scores since 1975 have tended to reverse that effect. By 1993, most of the increase in predicted probability of seeking advanced degrees associated with verbal skills had disappeared.

Putting the two effects together, we show in Figure 5 that changes in quantitative skills increased the probability of seeking a doctorate until 1978 and then the probability began a slow, steady decline. The probability of stopping at the bachelor's degree level began to rise in the mid 1980s at the same time as the probability of seeking a professional degree began to fall. The net impact of changing verbal and quantitative skills of bachelor's degree cohorts has been to lower the supply of doctorates since the late 70s and to lower the supply of professionals since the mid 80s.

B. Estimated Returns to Post Graduate Education

Table 4 reports the results from Ordinary Least Squares and Two-stage estimation of the log earnings equation (1). Both sets of results correct for sample weights. Least squares estimates of returns to graduate education are positive and significant. However, the implied annual returns are small. Assuming a Master's program takes two years and a PhD program takes 6, implied annual returns are only 5.8% and 4.2% respectively.[7] Annualized returns to professional degrees are more reasonable at 14.1%, assuming a four year program. There is a significant positive return to GRE mathematics scores, but no measurable return to verbal skills. There is a significant premium for postgraduate degrees in business and a significant discount for postgraduate degrees in the sciences.

Controlling for the likely endogeneity of the schooling choices raises the measured returns to advanced degrees.[8] The implied annual return to a Master's degree rises to 14.5%, and the returns to a Ph.D. rises to 12.6%, very similar to instrumental variable estimates of the returns to a year of education obtained at lower levels of education. The annualized return to a professional degree rises to 20.9%.[9]

Returning to the two alternative possibilities discussed at the beginning of the paper, our findings are consistent with the hypothesis that students who would be atypically successful in graduate school are actually more likely to halt their education at the bachelor's level. Consequently, average earnings of bachelor's degree recipients overstate the opportunity cost faced by those opting to pursue advanced degrees.

Our assessment is that the sorting is most easily observed when examining the role of the average GRE quantitative score. As indicated before, higher average GRE quantitative scores actually lower the probability of pursuing graduate education, even though strong quantitative skills are presumed to increase the likelihood of success in graduate school. Consequently, atypically strong graduate school prospects are actually less likely to pursue graduate training.

We can illustrate the impact of changing GRE scores on observed returns to schooling. We simulate how GRE scores alter log earnings directly and indirectly through their implied impact on the probability of receiving an advanced degree illustrated earlier in

Figure 5. The results of the simulation are shown in Figure 6. The direct effect of increases in the GRE quantitative score is to raise earnings, although the coefficient is no longer precisely estimated. There was little direct effect of the verbal score on earnings. The rise in average GRE scores also lowers the likelihood of attending graduate school, which counteracts the positive direct returns to quantitative scores.

The GRE verbal score does have an impact on earnings through its influence on post graduate training. However, when GRE scores start to slide, the resulting earnings retreat to just 2% above their 1963 level. The summed effects of the changes in GRE scores is a modest increase in average earnings across all college graduates, suggesting that changing skill content of bachelor's degree cohorts can only explain about 2% of the 35% increase in relative earnings for graduate degree holders shown in Figure 1.

These are the average earnings effects, but they can be used to motivate the hypothesized sorting effect discussed above. Those who do not go on to graduate school are drawn atypically from the upper tail of the GRE quantitative distribution and the lower tail of the GRE verbal distribution, both of which are expected to raise their earnings. On the other hand, those who go on to graduate school are drawn disproportionately from the lower tail of the quantitative GRE distribution and from the upper tail of the GRE verbal distribution, both of which lower their opportunity costs of graduate school. Consequently, the observed premium of average earnings for post graduate degree holders over bachelor's degree recipients understates the true returns to graduate school. Correcting for the sorting raises the estimated returns, as found in Table 4.

C. Unobserved Ability

Unobserved individual abilities may also affect the likelihood of pursuing an advanced degree. To test that hypothesis, we follow Rosenweig and Schultz (1983) by collecting the residuals from the earnings equation. These residuals represent individual ability uncorrelated with education level, major level skills, parents' education level, or demographic variables included in the model. They will also include random noise in the earnings function, so they will measure the unobserved ability with error. An auxiliary multinomial logit estimation of education choices on the earnings residuals will illustrate the direction of the effect of unobserved ability to earn income on the probability of seeking graduate or professional education. Note that the measurement error inherent in this method will tend to bias the coefficients toward zero.

Table 5 reports the estimated marginal effect of the earnings residual on the probability of pursuing each degree. Those with higher unobserved ability to earn income were less likely to stop at the bachelor's degree level and were more likely to pursue advanced degrees of all types. Consequently, sorting on unobserved ability works in the opposite direction as sorting on observed quantitative skills.

V. Conclusions

Returns to advanced degrees are positive and significant. Least squares estimates are quite low, on the order of 5% per year. Estimates increase in magnitude after controlling for likely endogeneity of the choice of pursuing an advanced degree, with estimated returns of comparable size to those estimated for schooling more generally. The findings of downward bias in least squares estimates of returns to graduate education are similar to findings in other settings.

Our study points out an interesting role for cognitive skills in the market for advanced degrees. Students in majors with higher average quantitative GRE scores are less likely to attend graduate school, even though such students presumably are more likely to be successful in graduate education. The opposite happens for verbal skills—students in majors with higher average verbal GRE scores are more likely to attend graduate school. This leads to a sorting effect whereby students whose cognitive skills would suggest lower earnings at the bachelor's level are more likely to attend graduate school. This sorting effect appears to be part of the cause of the downward bias in estimated returns to graduate education—the average earnings of those who do not go to graduate school overstate the opportunity costs of graduate education for those who do pursue advanced degrees. Nevertheless, changes in verbal and quantitative skills over time do not explain the large increases in relative returns to graduate and professional education since 1980. Future work is needed to identify the source of those rising returns.

These conclusions are subject to the usual caveat that our instruments may not be valid, although our measures of the costs of graduate education perform as expected, and we do try to control for unmeasured ability to a greater extent than has been possible in most studies. Nevertheless, we must acknowledge that to the extent that remaining unmeasured ability is important and correlated with our measures of family background variables, our results may still be subject to biases that we cannot control with the data at hand.

References

Altonji, Joseph G. and Thomas A. Dunn, "The Effects of Family Characteristics on the Return to Education," *Review of Economics and Statistics* 78 (Nov. 1996), 692-704.

Ashenfelter, Orley and Joseph D. Mooney, "Graduate Education, Ability, and Earnings," *Review of Economics and Statistics* 50 (Feb. 1968), 78-86.

Ashenfelter, Orley and Cecilia Rouse, "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins," *Quarterly Journal Economics* 113 (Feb. 1998), 253-284.

Behrman, Jere R., Rosenzwig, Mark R. and Paul Taubman, "College Choice and Wages: Estimates Using Data on Female Twins," *Review of Economics and Statistics* 78 (Nov. 1966), 672-685.

Bishop, John H., "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review* 79 (Mar. 1989), 178-197.

Card, David, "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics* (Amsterdam: Elsevier Science Publishers, 1999).

Deschenes, Olivier, "Estimating the Effects of Family Background on the Return to Schooling," University of California at Santa Barbara, Economics Working Paper Series No. 1020 (Mar. 2002).

Ehrenberg, Ronald G. "The Flow of New Doctorates." *Journal of Economic Literature* 30 (June 1992), 800-875.

Freeman, Richard B. *The Overeducated American* (New York: Academic Press., 1976a).

Freeman, Richard B. "A Cobweb Model of the Supply and Starting Salary of New Engineers." *Industrial and Labor Relations Review* 29 (Jan. 1976a), 236-246.

Freeman, Richard B. "It's Better Being and Economist (But Don't Tell Anyone)." *Journal of Economic Perspectives* 13 (Summer 1999), 139-145.

Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica* 45 (Jan. 1977), 1-22.

Grogger, Jeff and Eric Eide, "Changes in College Skills and the Rise in the College Wage Premium," *Journal of Human Resources* 30 (Spring 1995), 280-310.

Jaeger, David A. and Marianne E. Page, "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education," *Review of Economics and Statistics* 78 (Nov. 1996), 733-740.

Levine, Phillip B. and David J. Zimmerman, "The Benefit of Additional High-School Math and Science Classes for Young Men and Women," *Journal of Business and Economic Statistics* 13 (April 1995), 137-149.

Murnane, Richard J., Willett, John B., and Frank Levy, "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics* 77 (May 1995), 251-266.

Paglin, Morton and Anthony M. Rufolo, "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences," *Journal of Labor Economics*, 8 (Jan. 1990, Part I), 123-144.

Rosenweig, Mark R. and Paul Schultz, "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight," *Journal of Political Economy* 91 (Oct. 1983), 723-746.

Taber, Christopher R. "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?" *Review of Economic Studies* 68 (2001), 665-691.

Taubman, Paul J. and Terence J. Wales, "Higher Education, Mental Ability, and Screening," *Journal of Political Economy* 81 (Jan.-Feb. 1973), 28-55

Willis, Robert J. and Sherwin Rosen, "Education and Self-Selection," *Journal of Political Economy* 87 (Oct. 1979, Part II), s7-s36

Table 1. Descriptive Statistics: 1963-1986 (N = 67565)

	Variable	Mean	Std. Err.
Demographics	Age	41.2	(0.027)
	Experience	17.4	(0.025)
	Male	0.723	(0.002)
	US Citizen	0.956	(0.001)
	Rural Background	0.319	(0.002)
Education	BA	0.549	(0.002)
	MA	0.287	(0.002)
	Ph. D.	0.063	(0.001)
	Prof. Degree	0.101	(0.001)
	Posdoc	0.004	(>0.001)
Race	Hispanic	0.031	(0.001)
	White	0.849	(0.001)
	Black	0.052	(0.001)
	Asian	0.066	(0.001)
	Native Am.	0.002	(>0.001)
BA Major Field	Science Majors	0.342	(0.002)
	Engineering Majors	0.205	(0.002)
	Social Sci. Majors	0.326	(0.002)
	Business Major	0.032	(0.001)
	Other Majors	0.095	(0.001)
Earnings (1993 dollar)	Overall	53,864	(113.319)
	BA	47,900	(161.490)
	MA	53,325	(208.694)
	Ph.D.	59,657	(165.362)
	Professional Degree	84,155	(727.269)

Table 1. (cont'd)

Parents Education	Mother Ed 11 -	0.154	(0.001)
	Mother Ed 12	0.398	(0.002)
	Mother Ed 12 – 15	0.211	(0.002)
	Mother Ed 16	0.151	(0.001)
	Mother Ed 17 +	0.086	(0.001)
	Father Ed 11 -	0.189	(0.002)
	Father Ed 12	0.268	(0.002)
	Father Ed 12 – 15	0.181	(0.001)
	Father Ed 16	0.176	(0.001)
	Father Ed 17 +	0.185	(0.001)
	Med. School Tuition (1993 dollar)	10,651	(10.324)
	Grad. School Tuition (1993 dollar)	3,501	(1.052)
	% Self-Supported	26.3%	(0.020)

Table 2: Average GRE Score for the major, by attributes of individuals in the major

Individual Attribute	Verbal GRE	Quantitative GRE
BA	500.8	581.9
MA	502.4	568.7
PhD	508.2	573.0
Professional Degree	515.4	555.7
Science Majors	512.0	606.0
Engineering Majors	469.2	649.5
Social Science Majors	518.6	518.5
Business Major	475.4	542.3
Other Majors	502.4	507.5
White	503.6	573.9
Black	504.9	553.2
Asian	497.1	604.0
Native American	506.9	563.2
Male	501.0	585.3
Female	509.0	547.8

Table 3. Multinomial Logit Estimation of Higher Education Choices

Variable	MA		PhD		Professional	
Mother Ed 11 -	0.113	(0.046)	0.196	(0.054)	-0.106	(0.095)
Mother Ed 12 – 15	0.045	(0.042)	0.214	(0.047)	0.178	(0.070)
Mother Ed 16	-0.228	(0.118)	0.152	(0.118)	0.123	(0.177)
Mother Ed 17 +	-0.041	(0.127)	0.641	(0.123)	0.352	(0.184)
Father Ed 11 -	-0.068	(0.047)	-0.149	(0.055)	-0.293	(0.092)
Father Ed 12 – 15	-0.012	(0.046)	0.013	(0.053)	0.092	(0.081)
Father Ed 16	-0.195	(0.117)	0.059	(0.115)	0.149	(0.179)
Father Ed 17 +	0.133	(0.119)	0.615	(0.118)	0.782	(0.178)
Experience/100	0.317	(5.316)	0.237	(5.442)	7.178	(8.429)
Experience squared/100	0.018	(0.120)	0.071	(0.125)	-0.331	(0.192)
Verbal mean/100	0.319	(0.099)	1.343	(0.104)	1.923	(0.145)
Quant. Mean/100	0.049	(0.056)	-0.494	(0.052)	-1.565	(0.092)
Verbal stdv/100	0.228	(0.038)	0.114	(0.040)	-0.307	(0.074)
Quant. Stdv/100	-0.202	(0.034)	-0.109	(0.037)	0.266	(0.065)
Foreign Student Ratio/100	-50.84	(15.86)	156.2	(12.21)	169.2	(17.76)
Science Majors	-0.890	(0.078)	-0.098	(0.085)	2.495	(0.187)
Engineering Majors	-0.620	(0.103)	-0.451	(0.117)	1.640	(0.267)
Social science Majors	-0.726	(0.065)	-0.952	(0.077)	1.160	(0.165)
Business Major	-0.628	(0.099)	-1.529	(0.179)	-1.085	(0.557)
Rural background	-0.180	(0.032)	-0.231	(0.035)	-0.396	(0.057)
Male	-0.230	(0.035)	0.343	(0.041)	0.664	(0.060)
Citizen	-0.389	(0.056)	-1.544	(0.060)	0.338	(0.117)
Hispanic	-0.042	(0.057)	-0.292	(0.084)	0.072	(0.088)
Black	-0.040	(0.050)	-0.330	(0.090)	-0.162	(0.087)
Asian	0.238	(0.043)	0.394	(0.050)	-0.069	(0.079)
Native Am.	0.111	(0.153)	0.256	(0.170)	-0.549	(0.276)
Medical School Tuition/100	-0.007	(0.005)	-0.003	(0.006)	-0.003	(0.009)
Graduate School Tuition/100	-0.010	(0.023)	-0.100	(0.025)	-0.065	(0.038)
% Self-Supported	-0.028	(0.010)	-0.035	(0.010)	-0.041	(0.016)
Parent Ed 16+*% Self-Supported	0.021	(0.001)	0.006	(0.008)	0.010	(0.012)
Constant	0.012	(0.763)	-1.902	(0.823)	-2.928	(1.190)

Pseudo $R^2 = 0.082$

Standard errors in parentheses. Tuition is in constant 1983-84 dollars.

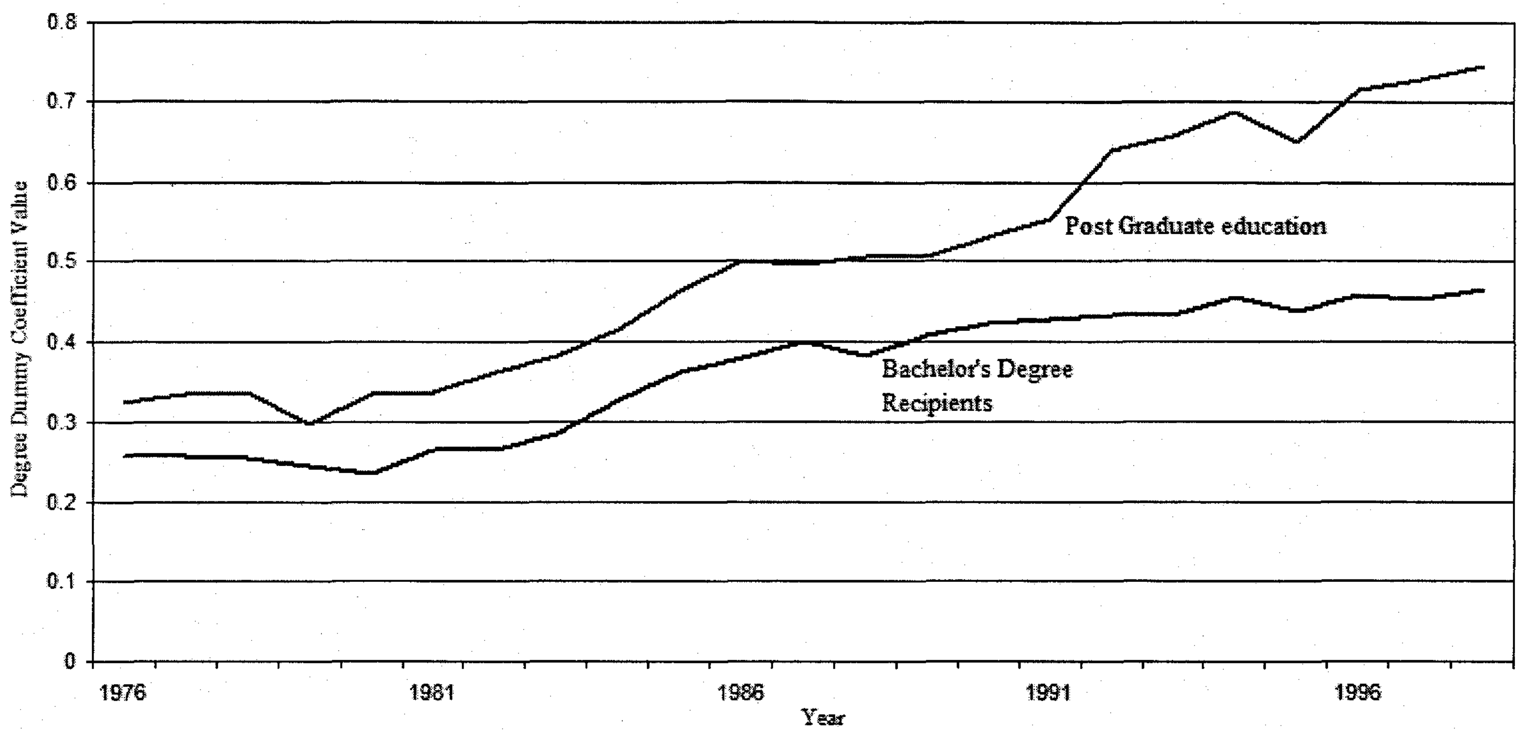
Table 4: Ordinary Least Squares and Two-Stage Estimation of the Log Earnings Function

Equation	OLS Estimates		IV Estimates	
Variables	Coefficient	Std. Err.	Coefficient	Std. Err.
MA	0.115	(0.007)	0.289	(0.158)
PhD	0.249	(0.008)	0.756	(0.109)
Professional Degree	0.565	(0.013)	0.837	(0.085)
Experience/100	3.189	(0.286)	3.058	(10.04)
Experience Squared/100	-0.039	(0.008)	-0.044	(0.027)
Male/100	16.695	(0.738)	15.592	(0.023)
Citizen/100	10.448	(1.256)	16.872	(0.011)
Posdoc/100	-36.757	(1.105)	-26.589	(0.010)
Verbal mean/100	-0.072	(0.020)	-0.154	(0.331)
Quant. mean/100	0.186	(0.012)	0.230	(1.515)
Verbal stdv/100	0.030	(0.007)	0.024	(0.009)
Quant. Stdv/100	-0.026	(0.007)	-0.020	(2.019)
Foreign Student Ratio/100	-4.692	(2.846)	-12.08	(1.095)
Science Majors	-0.066	(0.016)	-0.084	(0.011)
Engineering Majors	0.001	(0.021)	-0.012	(0.009)
Social Science Majors	0.015	(0.013)	0.042	(0.011)
Business Major	0.100	(0.019)	0.135	(0.034)
Hispanic	-0.053	(0.011)	-0.042	(0.051)
Black	-0.094	(0.009)	-0.078	(0.051)
Asian	-0.081	(0.009)	-0.098	(0.046)
Native Am.	-0.150	(0.034)	-0.141	(0.032)
Constant	9.342	(0.087)	9.394	(0.029)
R^2	0.228		0.139	

Table 5. Marginal Effect of Individual Heterogeneity on Probability to Pursue Advanced Degree

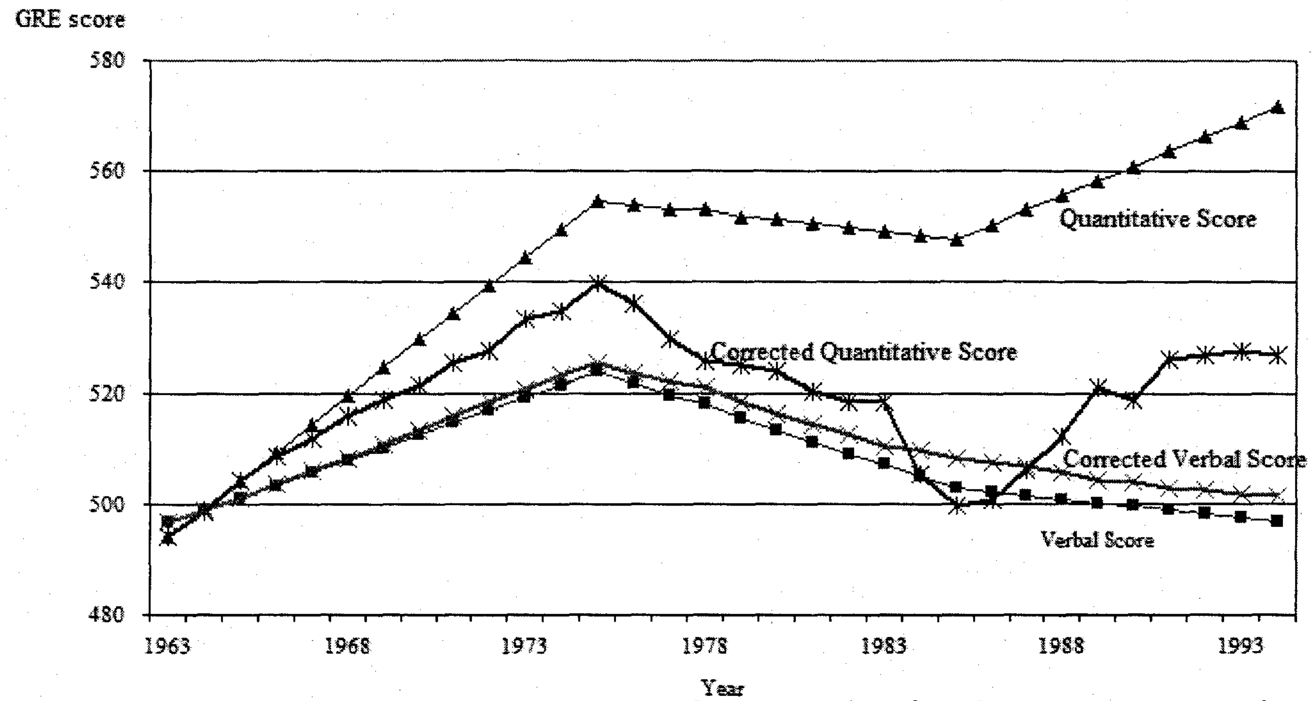
Dependent Variable	Marginal Effect	Std. Err.
BA	-0.224	(0.007)
MA	0.029	(0.007)
Ph.D.	0.038	(0.002)
Professional Degree	0.157	(0.004)

Figure 1. Estimated Returns to Schooling Relative to High School Graduates: 1976-1998



Notes: Values based on coefficients from annual regressions of log weekly wage on a vector of dummy variables indicating educational attainment, age, age squared, and a vector of demographic dummy variables. Data taken from the March Current Population Survey (1976-1998).

Figure 2: Trends of Observed and Corrected GRE Verbal and Quantitative Scores, 1963-1997



Note: Corrected Scores remove the estimated effect of foreign test takers from the mean score. Data taken from the Educational Testing Service and the Survey of Earned Doctorates, various years.

Figure 3: Simulated Probability of Schooling Choices from Changes in the Quantitative GRE Score, all else equal
(1963 normalized to 1)

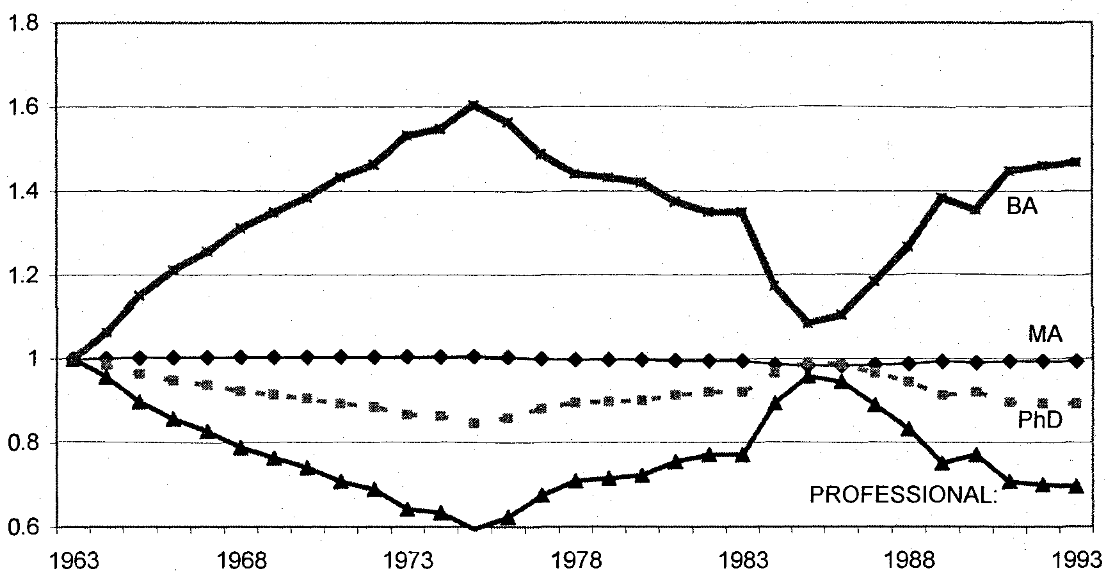


Figure 4: Simulated Probability of Schooling Choices from Changes in the Verbal GRE Score, all else equal (1963 normalized to 1)

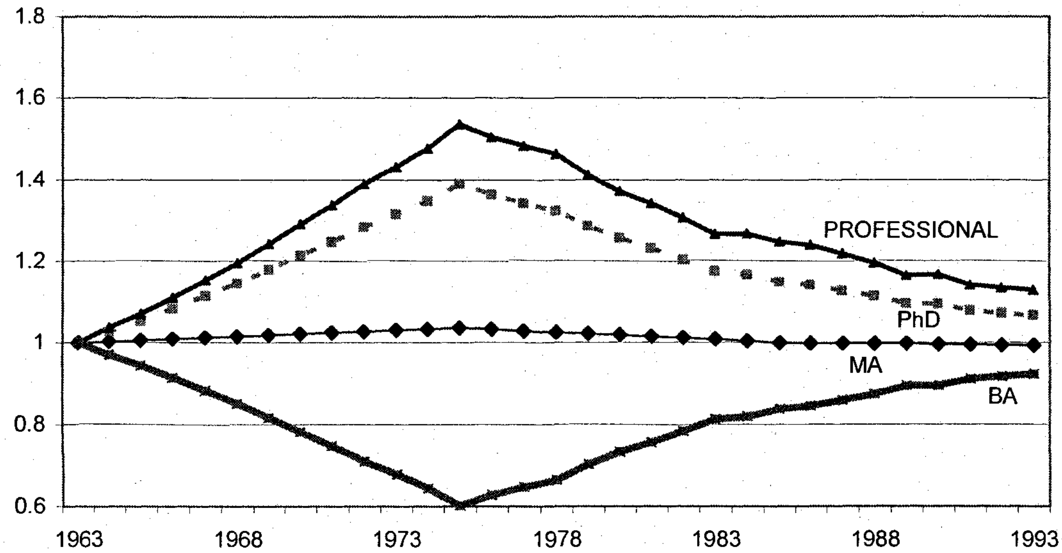


Figure 5: Simulated Probability of Schooling Choices from Changes in both the Quantitative and Verbal GRE Scores, all else equal (1963 normalized to 1)

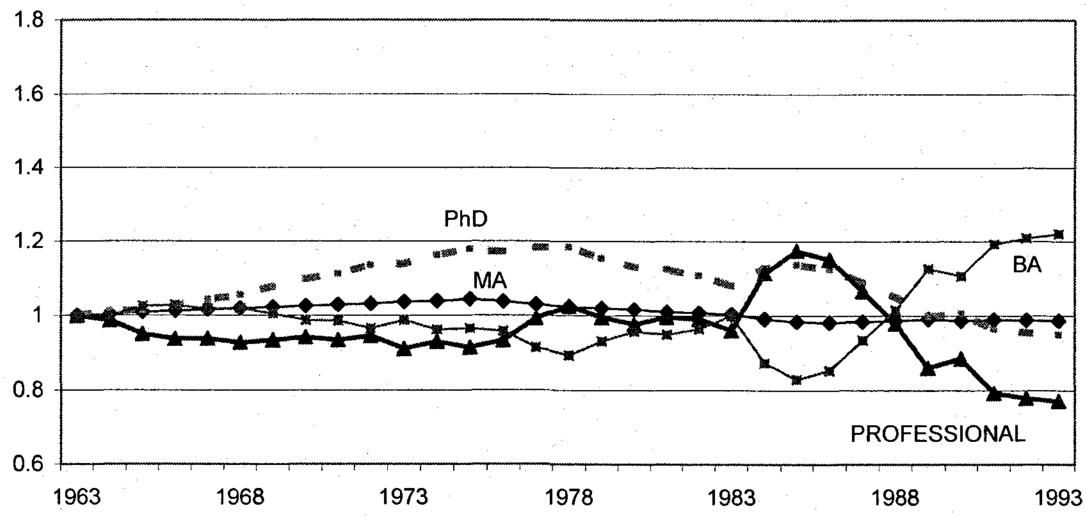
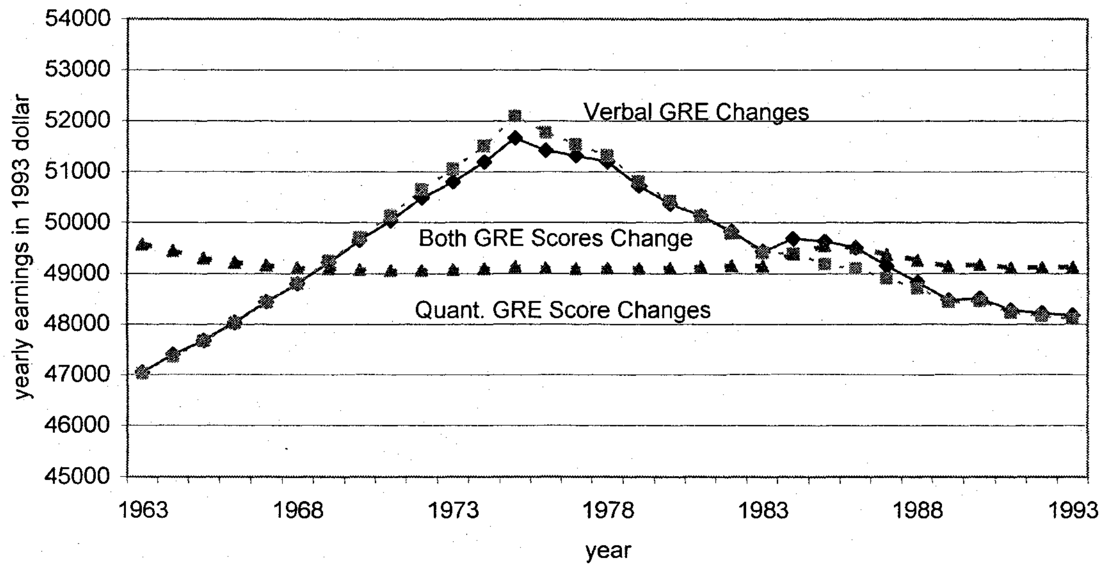


Figure 6: Simulated Direct and Indirect Impact of Changes in GRE Scores on the Average Earnings of Bachelor's Degree Recipients, 1963-1993 (in 1993 dollars)



[1] See Ehrenberg (1992) for a review. The most recent study of which we are aware is Jaeger and Page (1996). Earlier studies include Ashenfelter and Mooney (1968) and Taubman and Wales (1973). There is a vast literature on incentives to enter and returns to specific graduate degrees, pioneered by Richard Freeman (1976 a, b; 1999).

[2] In their study of twins data, Ashenfelter and Rouse (1998) found that family background variables strongly affected educational choices but did not affect earnings, exactly what one would want in an instrument. However, as Card (1999) argues, even that is not sufficient to validate family background measures as instruments if family background is correlated with unobservable ability.

[3] The Educational Testing Service provided this data for selected years:1963, 1974 to 1976, 1983 to 1986. The number of majors included in the report varied from 21 majors in 1963; 92 majors in 1974 – 76; and 98 majors in 1983 – 86. These were aggregated into 28 major groups to correspond with the majors reported in the SESTAT. The GRE did not report data on 9 of the majors 1963, and so the nearest included major was used: e.g. computer science was placed in mathematics; agricultural and food science was placed in biology; and so on. Once consistent data series were generated for the four reporting dates, the values were interpolated to generate continuous values for the intervening years. As most average scores change very slowly, this process is unlikely to generate wildly inaccurate estimates of average scores by major.

[4] We presume that the average doctoral program takes six years and that the percentage of foreign graduates completing the program is proportional to the percentage taking the GRE exam six years earlier.

[5] Bishop (1989) traced the time path of 12th grade high school cognitive skills. Our GRE scores would lag his measures by four years. The timing of the decline in verbal and quantitative scores is roughly consistent with the pattern of scores he reported for the Iowa Test of Basic Skills.

[6] Parents education level variable is 1 if both parents are more than college graduate, ½ if either one of them is more than college graduate, and 0 if both are less than college graduate.

[7] Jaeger and Page (1996) also estimate similarly small returns to Master's and PhD degrees under the assumption of exogenous education levels. Their estimation method includes both years of schooling as well as dummy variables indicating degree, so our annualized results are not directly comparable to theirs.

[8] Estimates that also included parental education in the second-stage earnings functions yielded comparable estimates of returns to graduate and professional education.

[9] These are likely to be overstated in that we do not incorporate tuition costs into the estimated return to professional degrees, and so these returns are gross of tuition costs. In contrast, tuition is often waived in doctoral programs, so those estimates are presumably closer to the true net return.

CHAPTER 4. THE COMPUTER AND THE INTERNET TECHNOLOGY ADOPTION AND THE URBAN-RURAL DIGITAL DIVIDE: DOES HIGH SPEED INTERNET ACCESS MATTER?

I. Background

There is widespread agreement that information technologies (IT) have contributed to U.S. economic growth since the mid 1990's. Jorgensen (2001) argued that productivity growth in IT-producing industries, especially in the semiconductor industry, led to substantial price declines in IT equipment. This IT cost decline has been a key element in the unprecedented economic expansion of the 1990s. Litan and Rivlin (2001) argued that IT technologies have effects beyond cost saving. IT can raise firm growth by improving customer service and by raising managerial efficiencies through enhanced supply-chain management. Blinder (2000) also concluded that IT accelerated the productivity growth of the U.S.

The rapid development of information technologies is also believed to be a factor explaining increasing U.S. wage inequality. Economists and commentators argue that the new information technologies are skill-biased, raising relative demand for highly educated workers relative to lower skilled workers (Acemoglu (2001), Autor, Katz and Kruger (1998)). To the extent that access to these technologies differs across races, genders, or regions, IT threatens to broaden earnings gaps. Differences in IT technology adoption and/or use between races, education levels, and regions has been labeled the "Digital Divide."

Previous studies have documented differences in adoption rates between the sexes, races, and education levels. However, most of these groups have similar access to technologies. In contrast, there can be large differences in access to technologies between urban and rural areas. The National Telecommunications and Information Administration (NTIA) of the U.S. Department of Commerce (2000) reported that, from 1998 to 2000, Black and Hispanic households have Internet adoption rates that are roughly half that of White and Asian households. There is only a modest gap in Internet use between men and women, Households in urban areas or in central cities have higher Internet access rates than households in rural area, although the rate of rural households is increasing faster than the other two areas. There are also significant differences in Internet and computer utilization rates by income and education level.

Prieger (2003) studied the supply-side of the Digital Divide. He investigated how various factors such as rural location, income, and race etc. affect broadband availability in a zipcode using the FCC, Census, and a telecommunications wire center database. He found no statistically significant differences in access to high speed Internet service by income. He also found no significant differences in access by ethnic or racial groups with the exception that Native Americans were significantly less likely to receive service. However, he did find rural location and demand side characteristics such as age, education or sex profiles of population in a zipcode are important determinants of broadband availability.

Fairlie (2004) analyzed the demand side of the Digital Divide. Using data from the Computer and Internet Use Supplement to the August 2000 Current Population Survey, he investigated the racial differences in rates of computer and Internet adoption. He explored the gap in utilization across races. He found that education, income and occupation are important determinants of computer ownership and having Internet access at home and that those factors also explain a substantial portion of racial gap, too. He speculated that there might be price or accessibility differences between rural and urban areas from the finding that living in a rural area had a negative effect on Internet access at home.

Gabe and Abel (2002) found that major telecommunications carriers deploy more ISDN (a high-speed, digital communications technology) infrastructure in metropolitan areas than in non-metropolitan areas in the U.S. and that the metro-nonmetro gap seems to be widening.

Past research suggests two possible reasons for an urban-rural digital divide. One is differences on the demand side such, as education, income or ethnicity that differ across regions that would lead to differences in technology adoption. The other is differences on the supply side, specifically differences in technology access across regions. Previous work that analyzes the demand side or the supply side separately may lead to misleading inferences if, for example, the same factor such as income is correlated with both use and availability of the technology. This study will investigate how both supply-side and demand-side factors affect adoption. Policy implications can be different depending on which factor is more crucial to the problem.

We explore two interrelated questions.

1) To what extent is there a difference in technology access and utilization between rural and urban areas?
2) How much of the difference in utilization is explained by differential technology access and how much is due to other socioeconomic factors?

We explore five measures of IT use on the job: use of the Internet at home for work; use of a computer at home for work: use of the Internet at work; use of a computer at work; Internet use for any purposes. The first two are viewed as measures of telecommuting, while the third and the fourth require physical commuting to access the technology. We investigate how local IT access, as measured by the number of high speed Internet providers in the county, affects the likelihood of adopting any of these five technology uses.

In section II, we discuss the empirical model of information technology adoption decision. Section III discusses data set used in this paper. The estimation model for the IT adoption decision and its results are included in section IV. Section V includes summary and conclusion.

II. Empirical Model

In our application, individuals decide whether or not to use technologies, namely the Internet or computers. The decision involves selecting the option that maximizes expected utility. We define that the utility function of an individual i is

1) $U(I_{ij};X_{ij},Z_{ij},\varepsilon_{ij})$,

where X_{ij} denote the observed characteristics of the individual, I_{ij} is a dummy variable that indicates if an individual i chooses to use technology j, the vector Z_i is a vector of variables that shifts the individual's taste for or cost of computer or Internet uses, and ε_{ij} is an unobserved random component. Although the utility level is not observable, we can specify a functional form for a representative agent into two components

2) $U = V(I_{ij};X_{ij},Z_{ij}) + \varepsilon_{ij}$.

Individual i is assumed to choose to use the new technology, $I_{ij} = 1$, if

3) $V(I_{ij}=1;X_{ij}^{1},Z_{ij}^{1}) + \varepsilon_{ij}^{1} \geq V(I_{ij}=0;X_{ij}^{0},Z_{ij}^{0}) + \varepsilon_{ij}^{0}$, or

4) $I_{ij}^{*} = g(X_{ij},Z_{ij}) - \eta_{ij} \geq 0$,

where $g(X_i) = [V(I_{ij} = 1; X_{ij}, Z_{ij}) - V(I_{ij} = 0; X_{ij}, Z_{ij})]$, $\eta_{ij} = \varepsilon_{ij}^0 - \varepsilon_{ij}^1$, and I_{ij}^* is the latent variable of an individual *i* to choose to accept the new technology *j*. By approximating (4) into a linear form, we get

5) $I_{ij}^* = X_{ij}\,\beta_j + Z_{ij}\,\beta_j - \eta_i$, *where* $I_{ij} = 1$ *if* $I_i^* \geq 0$ and 0 otherwise.

Therefore, the probability for an individual *i* to choose to use technology *j* is

6) $\Pr[I_{ij} = 1] = \Pr[\,g(X_{ij}, Z_{ij}) - \eta_{ji} \geq 0\,]$

$$= \Pr[\,\eta_{ij} \leq g(X_{ij}, Z_{ij})\,].$$

If the η_{ij} is assumed to be randomly and normally distributed, equation (5) can be estimated using probit.

III. Description of the Data Set

The data set analyzed in this paper is collected from three sources. The main individual information is taken from a survey conducted by the UCLA Center for Communication Policy (CCP) from 1999 to 2001. The survey is representative of the United States population. Each year, the survey covered about 2,000 people. The survey also included information on the respondent's zip code, which allowed us to merge in Federal Communications Commission (FCC) data on the number of high speed Internet access providers by zip code. The zip code also allowed us to identify the county of residence. That enabled us to merge in USDA Economic Research Service data on Rural-Urban Continuum (Beale) codes that identify counties as rural urban or metropolitan. We used only the 1999 and 2001 waves of the UCLA CCP surveys. The 2000 CCP could not be used because it lacked the zip code information needed to merge zip code- or county-level information. The survey questions include the individual's employment status, education level, demographic characteristics, and attitudes regarding Internet or computer use. The FCC has reported the number of high speed Internet providers at the zip code level every six months since December 1999. The FCC reports the actual number of providers if there are four or more, but reports only the existence of at least one provider if there are one to three. For those zip codes, we set the number of providers at 2. We generated two zip code-level and two county-level high speed access measures: the number of providers in the zip code, a dummy variable indicating whether there is at least one provider in the zip code, the county-level averages corresponding to these measures. All the measures should increase as the probability of having local high speed service increases. In addition, a higher number of local high speed Internet providers should increase competition among the providers and lower the access price.

The sample statistics are reported in table 1. The working data set includes 2665 observations on individuals between ages 23 and 65 who are not retired. Younger and older respondents are excluded to avoid complications caused by computer adoption and commuting decisions that would interact with decisions regarding education and retirement.

The useable sample is divided between 1,406 respondents in 1999 and 1,259 in 2001. All statistics are calculated using the weights provided by CCP. The education level was divided into 7 groups: less than 12 years, high school graduate, some college, college graduate, master's degree, Ph. D., and professional degrees. We generated a continuous measure of education by multiplying the expected years of education for each group.[1] The average respondent in the working sample was 42 years old and had 13.5 years of education. About 60 percent were married. A dummy variable, Minority, was set equal to 1 if an individual is neither white nor Asian. About 80 percent are Whites or Asians. Techlove is a dummy variable indicating an individual has a taste for technology. Techlove equals 1 if the respondent answered positively to the question "Overall, do you think that new technologies such as the Internet, cell phones, and pagers have made the world better place, a worse place, or neither better nor worse?" Forty five percent thought new technology is good for the world.

Zip codes were matched to a corresponding county using a file prepared by the Bureau of Census from the U.S. Postal Service (USPS) City-State file (November, 1999). Counties are defined as metropolitan, urban or rural using Beale codes. In general, higher Beale codes imply a more rural county. Metropolitan counties have Beale codes 0 to 3, urban counties have codes 4 -7 and rural counties have codes 8 and 9. Non-metropolitan counties with codes 4, 6 and 8 are adjacent to a metropolitan area.

We also generated average travel time from home to work. The Bureau of Census data reports total workers in the county aged 16 and older and aggregate travel time to work for the county in minutes. We divided the travel time by the total number of workers to get average travel time per worker by county, TTIME. We expected that the distance to work might affect the reloative incentives to use IT at home versus driving tow work to access IT. We also expected that differences in travel time between urban and rural areas might affect relative incentives to telecommute in urban versus rural markets. However, there was only a modest difference in average commuting time between urban and rural markets. In fact, average commuting time is higher by 1.2 minutes in urban than in rural markets.

Figure 1 through 5 show the percent average of five types of information technology uses by Beale code. Five types of uses include use of the Internet for the job at home (IJH), use of a computer for the job at home (CJH), use of the Internet at work (IW) use of a computer at work (CW) and Internet use for any purposes (IA). Fourteen percent used the Internet from their home for work; 32 percent used the Internet at work; 47 percent used a computer at work; 19 percent used a computer at home for their job; and 73 percent used the Internet for any purposes. In all four job-related cases and the any Internet use, computer and the Internet use for work is greatest in the metropolitan areas. In the non-metropolitan counties, the use of the information technology is lowest in the rural counties. In urban counties, use of information technology at home is bigger in counties not adjacent to a metropolitan area. There is a less clear pattern of information technology use at home across non-metropolitan counties. The internet use for any purpose shows a similar patter to the use of the Internet at work; the use is higher for metropolitan counties than for non-metropolitan counties.

We have two measures of the number of high speed Internet access. One is the number of high speed Internet providers in a zipcode (NHZ) being 0, 2, 4 and 5 to 17. The other (NHC) is the county level average of NHZ ranging from 0 to 15.25. We also generated two measures of high speed Internet access availability based on NHZ and NHC. DHZ, generated from NHZ, is a dummy variable indicating whether there is at least one provider in a zip code. DHC is the average of DHZ by county, measuring a concentration rate of high speed Internet access in a county. The average number of high speed Internet providers in a zip code was 3.7, with 90% of respondents having at least one high speed Internet provider servicing their zip code. The differences between NHZ and NHC or between DHZ and DHC are very small.

Our focus is to investigate the difference in technology adoption between urban and rural areas, and so it is important to define those areas precisely. Past work (Gabe and Abel

(2002), Fairlie(2004)) has focused on a division of metropolitan versus nonmetropolitan areas, but this groups rural counties with suburban counties that may differ little from the metropolitan counties. Another option is to group metropolitan counties and counties adjacent to the metropolitan area, and to consider all other counties as the rural group (Goetz and Rupasingha (2002)). This groups relatively urbanized counties (e.g. Beale Code=5) with truly rural counties, that have appreciably lower measures of high-speed access.

Our working sample is divided into two subgroups by Beale code: the urban group and the rural group. The urban group includes counties whose Beale codes are 0, 1, 2, 3, 4 or 5 (Metropolitan counties plus Urbanized counties). The rural group includes counties whose Beale codes are 6, 7, 8 or 9 (Less Urban counties plus Rural counties). We decided to divide our sample in this way for two reasons. One is the population of counties. By the definition of Economic Research Service, Non-metropolitan counties (BC = 4-9) are subdivided into three groups by urban population in a county: Urbanized counties of 20,000 or more (BC = 4, 5), Less Urbanized counties of 2,500 to 19,999 (BC = 6,7), and Rural counties of less than 2,500 (BC = 8,9). The urban population of counties in the urban group is more than 20,000 and the urban population of counties in the rural group is less than 20,000. The other reason is the number of high speed Internet providers in a zipcode by Beale code. As we can see in Figure 5, average numbers of NHZ by Beale code are almost same ranging from 2.6 to 2.4 for Beale codes, 3, 4 and 5. The average numbers NHZ by Beale code are 1.9 for BC = 6 and 1.6 for BC = 7. For these reasons we found that BC =5 is a reasonable cutoff point for our rural-urban division.[2]

The second and the third columns of Table 1 report summary statistics for these groups. 1,925 observations resided in urban-metro counties and 740 resided in rural counties.

The urban group averaged about one year younger than the rural group. The rural group was 5 percent more likely to be married. Average education is marginally higher in the urban group. The urban group was 15 percentage points more likely to hold favorable opinions about new information technologies.

There are significant regional differences in information technology use. For all five measures of the Internet and/or a computer uses, the percentage of users is higher in the urban group. Urban residents are 6.7 percent more likely to use the Internet at home for work purposes; 8.7 percent more likely to use a computer for work at home, 11.3 percent more likely to use the Internet at work, 2.3 percent more likely to use a computer at work, and 11.3 percent more likely to use the Internet for any purposes. The average number of high speed Internet access providers is 4.1 in urban areas but only 1.7 in rural areas. Ninety three percent urban residents had at least one high speed Internet provider compared to only 72 percent of the rural residents. Similar patterns hold when the county average measures of Internet access are compared.

IV. Estimation Model and Results

First, we estimate the following base technology adoption equation using probit for the five binary dependent variables; use of the Internet for the job at home, use of a computer for the job at home, use of the Internet at work, use of a computer at work, and use of the Internet for any purposes:

(7) $IT_i = X_i\,\beta + \phi\, TTIME_i + \gamma\, techlove_i + \theta_1\, NHC_i + \theta_2\, NHC_i^2 + e_i$.

IT_i represents one of the five information technology uses; the vector X_i includes variables such as experience, experience squared, race, gender, marital status, and years of education; *TTIME* is the average travel time per worker by county; *NHC* is the average number of high speed Internet providers in a county. The *techlove* is the technology attitude dummy variable. To check the consistency of the results, we also estimated the same five equations using *DHC, NHZ* and *DHZ*. The estimation results were almost identical qualitatively as well as quantitatively.

The use of different measures of Internet access is in part a means of assessing the possible role of residential mobility invalidating the treatment of high-speed access as exogenous. If people move to areas with access, then local access measures may be endogenous. Our use of a 1999 measure mitigates this somewhat, as home access was first introduced in 1998, and so households would not have had much time to react. Nevertheless, as it is less costly to move across zip codes than across counties, any endogenous mobility is more likely to involve moves across zip codes within a county than across counties. In that case, the zip code-level data is more likely subject to endogenous moves than would be the county-level measures. The lack of large differences across the two sets of measures suggests that the problem of households moving to get home Internet access is not a serious problem. The other results are available in Table 5 in the Appendix to this chapter.

Table 2 reports estimated the coefficients and marginal effects from the five probit regressions. The estimates of the regressions are consistent with our expectations. Minorities (in our case, Blacks and Hispanics) are less likely to engage in all five forms of information technology use. The estimated marginal effects range from -0.045 for IJH to -0.138 for CW. Married respondents are more likely to use computers or internet at home but not at work.

Men are more likely to work from home using a computer and the Internet and are also more likely to use the Internet at work. More experienced workers are less likely to use the Internet at work and to use the Internet for any purposes but are more likely to use computers at work.

As Huffman and Mercier (1991) found that schooling is important in the farm computer service adoption decision, we also found that the years of education increases likelihood of adopting all five technology uses. Individuals who are favorably disposed toward information technologies generally are also more likely to use the Internet at home or at work as well as for any purposes and are also more likely to use computers at home for work purposes. Holding human capital, demographic data and attitudes constant, we still find evidence that access matters for adoption. The marginal effect of travel time is positive, but the size is small and is also statistically insignificant in all 5 equations. It seems that differences in travel time among counties may not have a substantial impact on IT adoption. The number of high speed Internet providers in a county increases the probability of all the technology utilization measures except the use of computers at work, the case least connected to high-speed Internet access. The marginal effect of NHC on the use of the Internet for any purposes is almost twice as large as the effect on the other four uses. The marginal effect of NHC is also larger for home use of the Internet or computers for work purposes but the effect is not significant in cases of IT use at work. The pattern of marginal effects is consistent with the presumption that high-speed Internet access atypically raises the productivity of Internet use and computer use from home, leading to increased telecommuting.

Regional Comparison: Urban-metro Area vs. Rural Area

As noted above, there is a gap in Internet and computer use between urban-metro and rural areas. Next, we will examine the extent to which these differences can be attributed to differences in high-speed Internet access across the two regions. Let D^r be a dummy variable indicating rural residence.

$$8)\quad \begin{aligned} IT_i &= X_i\,\beta^1 + \phi^1\,TTIME + \gamma^1\,techlove_i + \theta_1^1\,NHC_i + \theta_2^1\,NHC_i^2 \\ &+ D^r X_i\,\beta^2 + D^r\,\phi^2\,TTIME + \gamma^2\,D^r techlove_i + \theta_1^2\,D^r NHC_i + \theta_2^2\,D^r NHC_i^2 + \tilde{e}_i\,. \end{aligned}$$

The coefficients $\beta^2, \phi^2, \gamma^2, \theta_1^2$ *and* θ_2^2 measure the difference in coefficients between the urban and the rural samples, where ϕ^2 measures the rural-urban difference in IT use in response to the travel time; γ^2 measures the rural-urban difference in technology use in response to individual taste for technology; θ_1^2 *and* θ_2^2 measure the rural-urban difference in response to high-speed Internet access; and β^2 measures the difference in the response to the other variables in the equation. We define these differences in coefficients between the two groups as the structural difference. Individual t-statistics on these coefficients will tell us whether the variable has a different effect across urban and rural respondents. The joint test of the significance of these coefficients will be a global test of whether there is a significant difference in the model of technology use across regions.

The joint significance tests based on estimation of equation (8) are reported in Table 4 in the Appendix. The global likelihood ratio test of the null hypothesis that $\beta^2 = \phi^2 = \gamma^2 = \theta_1^2 = \theta_2^2 = 0$ could not be rejected at the 5 % significance level for all of the technology choices except the case of Internet use at work. For the rest of the four IT uses, the null hypothesis can not be rejected at the 10 % significance level. For the use of the

Internet at work choice, only two coefficients, those on minority and on experience, differed significantly across the urban and the rural samples at the 10 % significance level. There are no coefficients that significantly differ across urban and rural samples at the 5 % significance level. We conclude that there is no significant structural difference between rural and urban areas in how probability of work-related computer or Internet use is affected by human capital, demographics, attitude toward new information technology or high-speed Internet access.

Decomposition of Regional Gaps in Computer or the Internet Uses

The lack of significant differences in the parameters of the IT use models between urban and rural areas implies that the observed difference can be explained entirely by rural-urban differences in average human capital, demographics, attitude toward new information technology and Internet access. We estimate these impacts using a variation of Blinder-Oaxaca (Blinder (1973); Oaxaca (1973)) adapted to the probit regression model. The regional gaps in the five types of uses of information technology can be expressed as:

(9) $IT_i^u - IT_i^r = (\bar{Z}_i^u - \bar{Z}_i^r)\delta^u + \bar{Z}_i^r(\delta^u - \delta^r)$,

where $\bar{Z}_i^j$ is a vector of average values of the all independent variables and δ^j is a vector of coefficients for group *j*. The first term represents the explained portion of the regional difference. It is the portion of the gap attributable to group differences in average values of the independent variables. The second term represents the unexplained portion of the gap. It is the portion attributable to differences in the coefficients between the two groups. As we have shown, there is no significant difference in coefficients across rural and urban respondents, so the unexplained portion is insignificant. Because there is no significant difference in the coefficients, we use the estimated parameters for the pooled populations, δ^o.

The original Blinder-Oaxaca decompositions assumed a linear regression model. Since the probability function of a probit model is nonlinear, it cannot be decomposed exactly. Some studies use the coefficient estimates from a linear probability model (LPM) to approximate the decomposition (e.g. Fairlie (2004) and Kilkenny and Huffman (2003)). The potential problem is that the linear probability model is sensitive to outliers in that it is possible to have an estimated probability over 1 or under 0. Our strategy uses the estimated coefficients from the probit model as follows. The explained difference between urban and rural technology use is

$D_{IT} = F(\overline{Z}^u\delta^o) - F(\overline{Z}^r\delta^o)$ where F is a normal density function. We calculate the share of each variable i in explaining this gap by

$$\alpha_i = \frac{(\overline{z}_i^u - \overline{z}_i^r)\delta_i^o}{(\overline{Z}^u - \overline{Z}^r)\delta^o},$$

where $\overline{z}_i^j$ is an average value of independent variable i in area j and δ_i^o is the associated coefficient estimate. By multiplying D_{IT} by α_i, we can estimate the explained regional difference attributable to independent variable i.[3] This is the counterpart of $(\overline{z}_i^u - \overline{z}_i^r)\delta^{LP}$ that is the impact of regional difference in independent variable i calculated using coefficient estimates from a LPM.

Table 3 reports the results from this decomposition. Negative values mean the variable lowers the difference between urban and rural areas while positive numbers expand the gap. Marital status and gender explain little of the urban-rural gap. Minority status serves to shrink the gap. Lower education levels in rural areas explain a substantial fraction of the gap. The attitude toward new IT (Techlove) also explains 10 to 22 percent of the gap except for use of computers at work. However, differential access, measured by differences in the number of high speed Internet providers between rural and urban areas, plays the largest role in the differences in adoption rates. Differences in the number of high speed Internet providers by county explain 56 percent of the difference in the Internet use for job at home, 45 percent of the gap in computer use for the job at home, 40 percent of the gap in Internet use at work, and 71% of the gap in Internet use for any purposes. The gap in access actually lowers the gap in computer use at work by 88 percent. When its impact is positive, Internet access matters most in encouraging Internet use from home, either for work or non work purposes.

IV. Conclusions

We have examined the adoption of new information technologies defined as use of the Internet for the job at home, use of a computer for the job at home, use of the Internet at work, use of a computer at work and use of the Internet for any purposes using pooled data in 1999 and 2001. Adoption is positively affected by schooling, and is also correlated with gender, race, and marital status. Individuals who have a favorable impression toward new technologies are also more likely to use computers and the Internet. However, even when those factors are controlled, local access to high speed Internet plays an important role in the technology adoption decision. It increases the probability of using computers and the Internet for work from home and also increases the likelihood of using the Internet at work. The pattern of marginal effects is consistent with the presumption that access is more important for telecommuting than for the use of the technology at work. There is no statistically significant difference in the coefficients governing the technology adoption process between urban and rural areas. This implies that the observed difference in IT uses can be explained by rural-urban differences in average human capital, demographics and Internet access. Differences in the number of high speed Internet providers by county explain 56 percent of the difference in the Internet use for job at home, 45 percent of the gap in computer use for the job at home, 40 percent of the gap in Internet use at work, and 71% of the gap in Internet use for any purposes. The gap in access actually lowers the gap in computer use at work by 88 percent. When its impact is positive, Internet access matters most in encouraging Internet use from home, either for work or non work purposes. Results suggest that as high-speed Internet access increase, there will be increased substitution of telecommuting for commuting.

References

Acemoglu, Daron, "Technological Change, Inequality, and Labor Market," 2002, *Journal of Economic Literature*, vol. 40, no. 1, 7-72.

Autor, David H., Katz, Lawrence F. and Alan B. Krueger, "Computing Inequality: Have Computers Changed the Labor Market?" 1998, *Quarterly Journal of Economics*, vol. 113, no. 4, 1169-1213.

Blinder, Alan S., "Wage Discrimination: Reduced Form and Structural Variables," 1973, *Journal of Human Resources*, vol. 8, 436-455.

Blinder, Alan S., "The Internet and the New Economy," *The Internet Policy Institute*, policy brief #60, Available: http://www.brookings.edu/comm/policybriefs/pb60.htm, date retrieved: 26 May 2005.

Fairlie, Robert W., "Race and The Digital Divide," 2004, *Contributions to Economic Analysis & Policy*, vol. 3, no. 1, Article 15, 1-38.

Gabe, Todd M. and Jaison R. Abel, "Deployment of Advanced Telecommunications Infrastructure in Rural America: Measuring the Digital Divide," 2002, *American Journal of Agricultural Economics*, vol. 84, no. 5, 1246-1252.

Goetz, Stephan J. and Anil Rupasingha, "High-Tech Firm Clustering: Implications for Rural Areas," 2002, *American Journal of Agricultural Economics*, vol. 84, no. 5, 1229-1236.

Huffman, Wallace E. and Stephanie Mercier, "Joint Adoption of Microcomputer Technology: An Analysis of Farmers' Decisions," *Review of Economics and Statistics*, 1991, vol. 73, no. 3, 541-546.

Jorgensen, Dale W., "Information Technology and the U.S. Economy," *American Economic Review*, 2001, vol. 91, no. 1, 1-32.

Kilkenny, Maureen and Sonya K. Huffman, "Rural/Urban Welfare Program and Labor Force Participation," 2003, *American Journal of Agricultural Economics*, vol. 85, no. 4, 914-927.

Litan, Robert E. and Alice M. Rivlin, "Projecting the Economic Impact of the Internet," 2001, *American Economic Review*, vol. 91, no. 2, 313-317.

National Telecommunications and Information Administration, "Falling Through the Net: Toward Digital Inclusion: A Report on American's Access to Technology Tools," 2000, http://www.ntia.doc.gov/ntiahome/fttn00/Falling.htm#2, date retrieved: 26 May 2005.

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets" 1973, *International Economic Review*, vol. 14, 693-709.

Prieger, James E., "The Supply Side of the Digital Divide: Is There Equal Availability in the Broadband Internet Access Market?" 2003, *Economic Inquiry*, vol. 41, no. 2, 346-363.

Table 1. Descriptive statistics of the Working Sample: years 1999 and 2001

	Total Sample			Urban Sample			Rural Sample		
Variable	Obs	Mean	Std. Err.	Obs	Mean	Std. Err.	Obs	Mean	Std. Err.
Male	2665	0.462	(0.010)	1925	0.465	(0.011)	740	0.454	(0.018)
Married	2665	0.583	(0.010)	1925	0.570	(0.011)	740	0.615	(0.018)
Age	2665	41.9	(0.219)	1925	41.6	(0.256)	740	42.7	(0.419)
Experience	2665	22.5	(0.222)	1925	22.0	(0.260)	740	23.5	(0.426)
Education	2665	13.5	(0.046)	1925	13.6	(0.055)	740	13.2	(0.084)
Travel Time	2319	24.0	(0.101)	1925	24.166	(0.107)	394	23.0	(0.273)
Technology Interest	2665	0.455	(0.010)	1925	0.498	(0.011)	740	0.347	(0.018)
White	2665	0.771	(0.008)	1925	0.743	(0.010)	740	0.842	(0.013)
Black	2665	0.106	(0.006)	1925	0.118	(0.007)	740	0.078	(0.010)
Asian	2665	0.023	(0.003)	1925	0.028	(0.004)	740	0.009	(0.004)
Native Am.	2665	0.016	(0.002)	1925	0.017	(0.003)	740	0.011	(0.004)
Minority	2665	0.200	(0.008)	1925	0.225	(0.010)	740	0.136	(0.013)
Internet for Job Home	2665	0.137	(0.007)	1925	0.156	(0.008)	740	0.089	(0.010)
Computer for Job Home	2665	0.185	(0.008)	1925	0.210	(0.009)	740	0.123	(0.012)
Internet at Work	2665	0.317	(0.009)	1925	0.349	(0.011)	740	0.236	(0.016)
Computer at Work	1259	0.466	(0.014)	762	0.475	(0.018)	497	0.452	(0.022)
Any Internet Use	2665	0.727	(0.009)	1925	0.759	(0.010)	740	0.646	(0.018)
NHZ	2319	3.665	(0.060)	1925	4.091	(0.067)	394	1.657	(0.064)
DHZ	2319	0.895	(0.006)	1925	0.933	(0.006)	394	0.715	(0.023)
NHC	2319	3.698	(0.055)	1925	4.130	(0.061)	394	1.659	(0.063)
DHC	2319	0.899	(0.006)	1925	0.937	(0.004)	394	0.717	(0.022)

Computer at Work is available only in the 2001

Table 2. Probit Estimation Results on Each Dependent Variable

	Internet Job Home			Computer Job Home			Internet at Work			Computer at Work			Any Internet Use		
Variables	dF/dx	Coeff.	Std. Err.	dF/dx	Coeff.	Std. Err.	dF/dx	Coeff.	Std. Err.	dF/dx	Coeff.	Std. Err.	dF/dx	Coeff.	Std. Err.
Minority	-0.045	-0.271	(0.107)	-0.068	-0.310	(0.098)	-0.098	-0.292	(0.086)	-0.138	-0.356	(0.144)	-0.132	-0.411	(0.087)
Married	0.037	0.208	(0.074)	0.053	0.223	(0.068)	0.030	0.084	(0.062)	0.023	0.057	(0.103)	0.043	0.144	(0.067)
Male	0.051	0.273	(0.070)	0.047	0.195	(0.065)	0.045	0.128	(0.060)	-0.050	-0.125	(0.100)	0.038	0.127	(0.066)
Education	0.035	0.187	(0.016)	0.048	0.198	(0.015)	0.062	0.176	(0.013)	0.079	0.199	(0.022)	0.055	0.187	(0.019)
Experience/10	-0.010	-0.056	(0.062)	>0.001	>0.001	(0.056)	-0.038	-0.108	(0.048)	0.255	0.644	(0.179)	-0.096	-0.323	(0.048)
$(Experience/10)^2$	-0.002	-0.011	(0.013)	-0.005	-0.022	(0.012)	-0.002	-0.005	(0.009)	-0.070	-0.176	(0.036)	0.005	0.015	(0.008)
Rural Area	-0.033	-0.195	(0.112)	-0.047	-0.207	(0.100)	-0.070	-0.206	(0.091)	-0.064	-0.163	(0.150)	-0.046	-0.148	(0.092)
Technology Interest	0.059	0.318	(0.073)	0.068	0.281	(0.067)	0.065	0.185	(0.061)	0.006	0.015	(0.106)	0.073	0.246	(0.067)
Travel Time	0.001	0.007	(0.007)	0.001	0.005	(0.007)	0.002	0.006	(0.007)	0.001	0.003	(0.011)	0.001	0.003	(0.007)
NHC/10	0.196	1.059	(0.460)	0.184	0.762	(0.434)	0.159	0.451	(0.385)	-0.077	-0.195	(0.633)	0.531	1.792	(0.425)
$(NHC/10)^2$	-0.155	-0.839	(0.397)	-0.174	-0.719	(0.387)	-0.128	-0.362	(0.329)	0.032	0.082	(0.496)	-0.382	-1.288	(0.386)
Constant		-4.297	(0.313)		-4.085	(0.291)		-2.927	(0.262)		-2.944	(0.463)		-1.733	(0.323)
Pseudo R^2		0.154			0.152			0.114			0.142			0.157	

Obs.: 2319 except Computer at Work (924)

Table 3. Blinder-Oaxaca Decomposition:
Weighted Percentage of Explained Variation Using Probit Model

	IJH	CJH	IW	CW	IA
Minority	-9.65%	-14.62%	-17.49%	-76.55%	-10.04%
Married	0.89%	1.26%	0.60%	1.47%	0.42%
Male	-4.91%	-4.61%	-3.84%	13.57%	-1.57%
Education	28.3%	39.5%	44.4%	181.2%	19.3%
Experience	6.1%	7.6%	13.0%	55.5%	10.1%
TECHLOVE	19.2%	22.4%	18.7%	5.6%	10.2%
TTIME	3.5%	3.5%	4.6%	7.4%	0.8%
NHC	56.6%	45.0%	40.0%	-88.2%	70.7%
Sum	100.0%	100.0%	100.0%	100.0%	100.0%

Figure 1. Use of the Internet for Job at Home by Beale Code

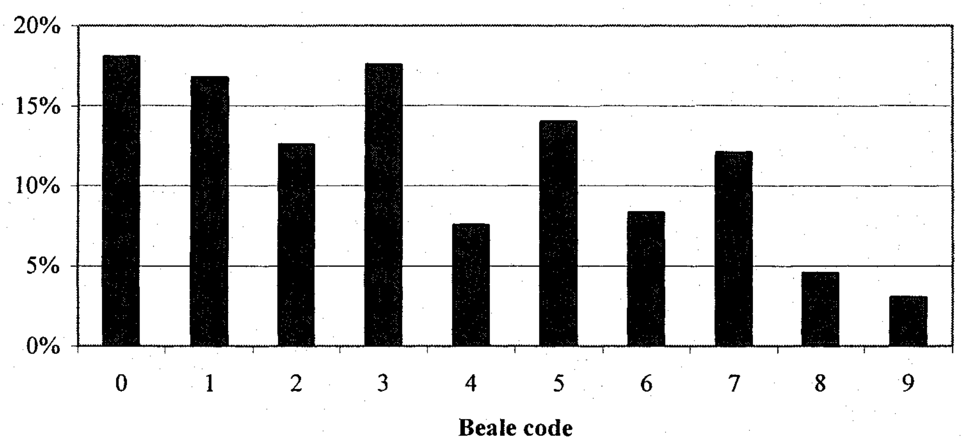


Figure 2. Use of Computers for Job at Home by Beale code

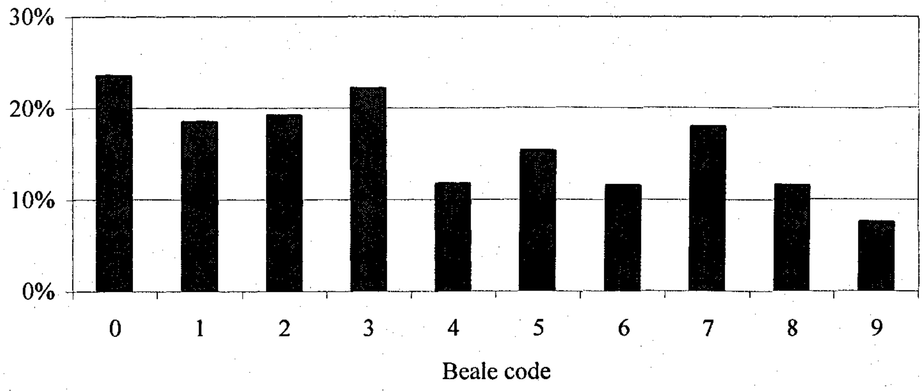


Figure 3. Use of the Internet at Work

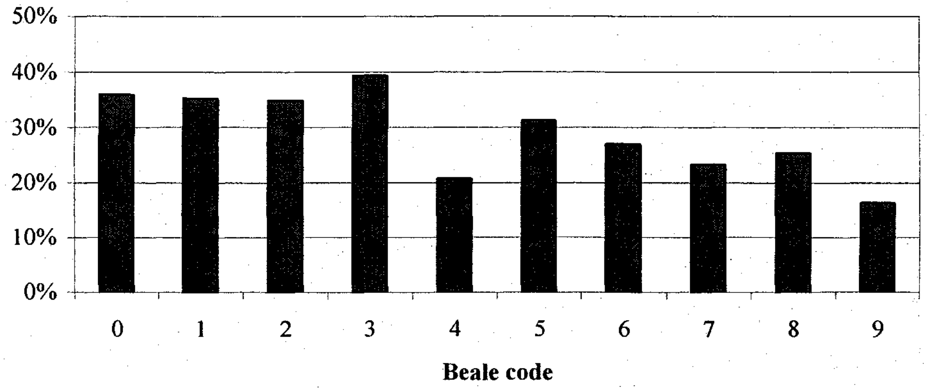


Figure 4. Use of Computers at Work

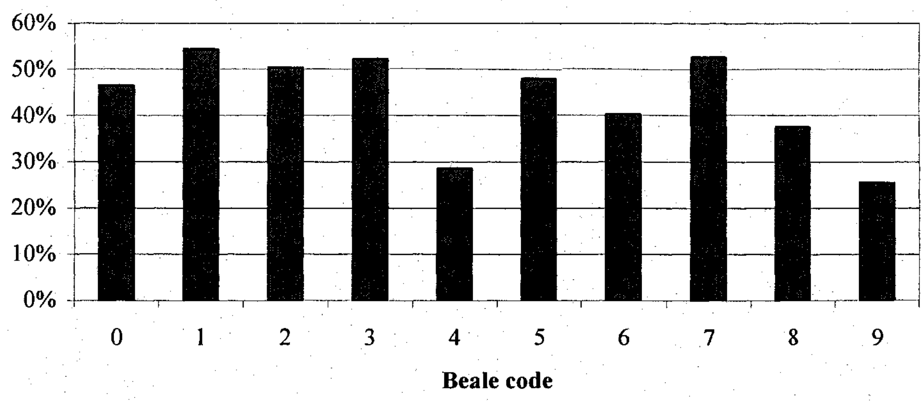


Figure 5. Internet Use for Any Purposes by Beale Code

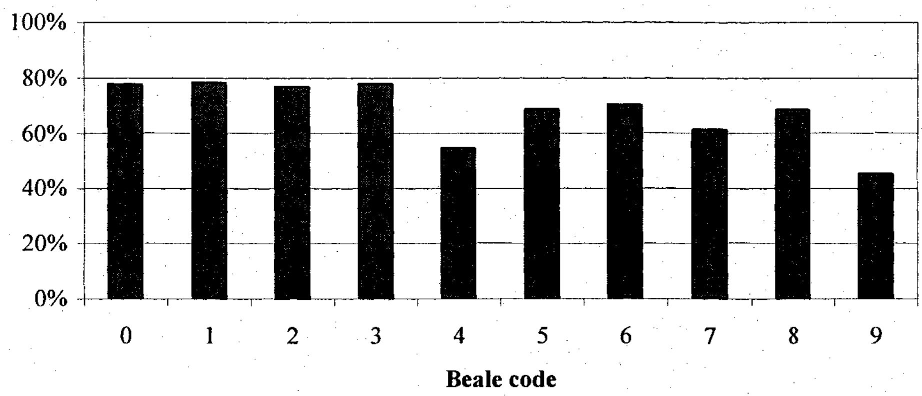


Figure 6. The Number of High Speed Internet Access Providers within a Zipcode by Beale Code

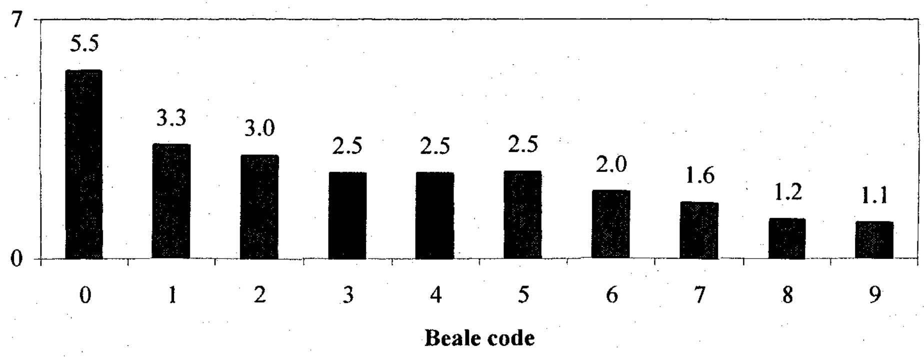


Table 4. Probit Estimation Results for Each Dependent Variable with 3 Regional Divisions

	IJH (Obs: 2319)		CJH (Obs: 2319)		IW (Obs: 2319)		CW (Obs: 924)		IA (Obs: 2319)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Minority	-0.266	(0.107)	-0.308	(0.098)	-0.290	(0.086)	-0.352	(0.145)	-0.409	(0.087)
Married	0.207	(0.074)	0.222	(0.068)	0.082	(0.062)	0.059	(0.103)	0.143	(0.067)
Male	0.271	(0.070)	0.195	(0.065)	0.129	(0.060)	-0.121	(0.100)	0.129	(0.066)
Education	0.187	(0.016)	0.198	(0.015)	0.176	(0.013)	0.198	(0.022)	0.187	(0.019)
Experience/10	-0.053	(0.062)	0.002	(0.056)	-0.108	(0.048)	0.648	(0.179)	-0.322	(0.048)
$(Experience/10)^2$	-0.011	(0.013)	-0.022	(0.012)	-0.005	(0.009)	-0.177	(0.036)	0.015	(0.008)
BC567*	-0.094	(0.115)	-0.186	(0.104)	-0.169	(0.092)	-0.081	(0.149)	-0.124	(0.093)
BC89**	-0.541	(0.255)	-0.338	(0.197)	-0.260	(0.172)	-0.430	(0.256)	-0.259	(0.169)
TECHLOVE	0.315	(0.073)	0.279	(0.067)	0.183	(0.060)	0.011	(0.106)	0.242	(0.067)
TTIME	0.009	(0.007)	0.005	(0.007)	0.005	(0.007)	0.006	(0.012)	0.003	(0.007)
NHC	0.101	(0.048)	0.070	(0.045)	0.044	(0.039)	-0.031	(0.064)	0.173	(0.043)
NHC^2	-0.008	(0.004)	-0.007	(0.004)	-0.004	(0.003)	0.001	(0.005)	-0.012	(0.004)
Constant	-4.321	(0.315)	-4.057	(0.294)	-2.912	(0.267)	-2.969	(0.472)	-1.722	(0.330)
R^2	0.155		0.152		0.113		0.144		0.158	

*: BC567: counties of BC = 5,6 or 7

**: BC89: counties of BC = 8 or 9

Reference region is the group of counties of BC = 0 - 4

Table 5. Probit Estimation Results of Equation (8) and Test Statistics on Regional Difference

Variable	IJH Coef.	IJH Std. Err.	CJH Coef.	CJH Std. Err.	IW Coef.	IW Std. Err.	CW Coef.	CW Std. Err.	IA Coef.	IA Std. Err.
Minority	-0.277	(0.112)	-0.315	(0.102)	-0.249	(0.090)	-0.342	(0.155)	-0.410	(0.092)
Married	0.174	(0.079)	0.187	(0.073)	0.039	(0.067)	-0.076	(0.114)	0.208	(0.074)
Male	0.255	(0.075)	0.183	(0.070)	0.165	(0.066)	-0.082	(0.112)	0.129	(0.074)
Education	0.188	(0.017)	0.197	(0.016)	0.180	(0.014)	0.205	(0.023)	0.188	(0.019)
Experience/10	-0.081	(0.067)	-0.034	(0.060)	-0.141	(0.051)	0.532	(0.193)	-0.339	(0.053)
$(\text{Experience}/10)^2$	-0.010	(0.015)	-0.016	(0.013)	0.000	(0.010)	-0.159	(0.040)	0.019	(0.009)
TECHLOVE	0.321	(0.078)	0.283	(0.072)	0.206	(0.066)	0.008	(0.119)	0.277	(0.075)
TTIME	0.012	(0.008)	0.012	(0.007)	0.007	(0.007)	0.011	(0.013)	0.006	(0.008)
NHC	0.105	(0.051)	0.074	(0.048)	0.053	(0.043)	-0.020	(0.073)	0.025	(0.017)
NHC2	-0.009	(0.004)	-0.007	(0.004)	-0.004	(0.004)	0.000	(0.006)	0.307	(0.163)
Dr*Minority	-0.047	(0.434)	0.027	(0.388)	-0.632	(0.341)	-0.071	(0.443)	-0.008	(0.292)
Dr*Married	0.396	(0.281)	0.380	(0.234)	0.294	(0.188)	0.742	(0.290)	-0.330	(0.186)
Dr*Male	0.231	(0.213)	0.156	(0.188)	-0.168	(0.163)	-0.109	(0.253)	-0.027	(0.169)
Dr*Education	-0.027	(0.030)	-0.006	(0.028)	-0.025	(0.025)	-0.040	(0.046)	0.015	(0.029)
Dr*(Experience/10)	0.184	(0.143)	0.206	(0.123)	0.204	(0.116)	0.370	(0.471)	0.067	(0.122)
$\text{Dr}^*(\text{Experience}/10)^2$	-0.006	(0.028)	-0.037	(0.025)	-0.035	(0.023)	-0.058	(0.096)	-0.015	(0.022)
Dr*TECHLOVE	-0.056	(0.218)	-0.071	(0.196)	-0.189	(0.162)	-0.041	(0.262)	-0.139	(0.168)
Dr*TTIME	-0.029	(0.016)	-0.034	(0.015)	-0.003	(0.013)	-0.028	(0.022)	-0.010	(0.014)
Dr*NHC	0.013	(0.069)	0.006	(0.067)	-0.013	(0.064)	0.047	(0.088)	0.051	(0.073)
Constant	-4.341	(0.312)	-4.122	(0.289)	-2.999	(0.260)	-2.993	(0.465)	-1.828	(0.337)
Pseudo R^2	0.159		0.156		0.118		0.154		0.158	
χ_9^2	14.04		14.44		17.24		12.520		9.18	

Critical Value of χ_9^2 = 14.68 (with 90% CI), 16.92 (with 95% CI), or 21.67 (with 99% CI)

Table 6. Probit Estimation Results on Each Dependent Variable Using Different Measures of IT Use

Panel 1: Results with NHZ

	IJH		CJH		IW		CW		IA	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Minority	-0.265	(0.107)	-0.309	(0.098)	-0.292	(0.086)	-0.354	(0.144)	-0.422	(0.088)
Married	0.208	(0.074)	0.222	(0.068)	0.087	(0.062)	0.066	(0.104)	0.140	(0.068)
Male	0.271	(0.070)	0.192	(0.065)	0.124	(0.060)	-0.130	(0.100)	0.123	(0.067)
Education	0.188	(0.016)	0.199	(0.015)	0.175	(0.013)	0.196	(0.022)	0.189	(0.019)
Experience/10	-0.055	(0.062)	0.001	(0.056)	-0.108	(0.048)	0.653	(0.179)	-0.322	(0.048)
$(\text{Experience}/10)^2$	-0.011	(0.013)	-0.022	(0.012)	-0.006	(0.010)	-0.178	(0.036)	0.014	(0.008)
Rural Area	-0.240	(0.111)	-0.238	(0.099)	-0.214	(0.089)	-0.111	(0.142)	-0.176	(0.089)
TECHLOVE	0.317	(0.072)	0.281	(0.066)	0.190	(0.060)	0.021	(0.106)	0.245	(0.067)
TTIME	0.008	(0.007)	0.006	(0.007)	0.004	(0.006)	-0.003	(0.011)	0.002	(0.007)
NHZ	0.037	(0.037)	0.027	(0.035)	0.010	(0.032)	-0.029	(0.049)	0.144	(0.034)
NHZ^2	-0.002	(0.003)	-0.002	(0.003)	0.000	(0.003)	0.003	(0.004)	-0.009	(0.003)
Constant	-4.185	(0.316)	-4.010	(0.293)	-2.846	(0.262)	-2.841	(0.461)	-1.668	(0.321)
Pseudo R^2	0.152		0.150		0.114		0.143		0.157	

Table 6. (cont'd)
Panel 2: Results with DHC

	IJH		CJH		IW		CW		IA	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Minority	-0.256	(0.108)	-0.305	(0.099)	-0.287	(0.087)	-0.361	(0.144)	-0.395	(0.088)
Married	0.208	(0.074)	0.227	(0.068)	0.085	(0.062)	0.059	(0.103)	0.149	(0.068)
Male	0.273	(0.070)	0.192	(0.065)	0.127	(0.060)	-0.124	(0.100)	0.129	(0.067)
Education	0.190	(0.016)	0.199	(0.015)	0.176	(0.013)	0.198	(0.022)	0.190	(0.019)
Experience/10	-0.056	(0.061)	0.000	(0.056)	-0.110	(0.048)	0.646	(0.179)	-0.334	(0.048)
$(Experience/10)^2$	-0.010	(0.013)	-0.022	(0.012)	-0.005	(0.009)	-0.177	(0.036)	0.018	(0.008)
Rural Area	-0.247	(0.110)	-0.223	(0.097)	-0.221	(0.087)	-0.155	(0.130)	-0.220	(0.089)
TECHLOVE	0.312	(0.071)	0.285	(0.066)	0.183	(0.060)	0.014	(0.106)	0.227	(0.066)
TTIME	0.009	(0.007)	0.005	(0.006)	0.007	(0.006)	0.002	(0.011)	0.007	(0.007)
DHC	0.184	(0.140)	0.145	(0.128)	0.103	(0.117)	-0.253	(0.220)	0.470	(0.120)
Constant	-4.316	(0.329)	-4.086	(0.306)	-2.954	(0.275)	-2.756	(0.481)	-1.886	(0.334)
Pseudo R^2	0.152		0.151		0.113		0.143		0.155	

Table 6. (cont'd)

Panel 3: Results with DHZ

	IJH		CJH		IW		CW		IA	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Minority	-0.257	(0.108)	-0.306	(0.098)	-0.288	(0.087)	-0.359	(0.144)	-0.397	(0.088)
Married	0.206	(0.074)	0.225	(0.068)	0.084	(0.062)	0.060	(0.103)	0.141	(0.067)
Male	0.273	(0.070)	0.192	(0.065)	0.127	(0.060)	-0.125	(0.100)	0.130	(0.067)
Education	0.190	(0.016)	0.199	(0.015)	0.177	(0.013)	0.198	(0.022)	0.191	(0.019)
Experience/10	-0.055	(0.061)	0.001	(0.056)	-0.109	(0.048)	0.643	(0.178)	-0.331	(0.048)
$(\text{Experience}/10)^2$	-0.010	(0.013)	-0.022	(0.012)	-0.005	(0.009)	-0.176	(0.036)	0.017	(0.008)
Rural Area	-0.251	(0.109)	-0.229	(0.097)	-0.226	(0.086)	-0.134	(0.129)	-0.233	(0.087)
TECHLOVE	0.311	(0.071)	0.284	(0.066)	0.182	(0.060)	0.017	(0.106)	0.225	(0.066)
TTIME	0.010	(0.007)	0.005	(0.006)	0.007	(0.006)	0.001	(0.011)	0.007	(0.007)
DHZ	0.168	(0.127)	0.119	(0.114)	0.081	(0.103)	-0.091	(0.192)	0.415	(0.105)
Constant	-4.304	(0.331)	-4.065	(0.304)	-2.935	(0.272)	-2.882	(0.475)	-1.843	(0.328)
Pseudo R^2	0.152		0.151		0.113		0.142		0.155	

Table 7. Blinder-Oaxaca Decomposition:
Weighted Percentage of Explained Variation Linear Probability Model

	IJH	CJH	IW	CW	IA
Minority	-6.84%	-10.95%	-15.80%	-86.57%	-10.67%
Married	0.95%	1.27%	0.57%	2.04%	0.46%
Male	-4.73%	-4.33%	-3.76%	14.87%	-1.55%
Education	33.46%	44.16%	46.38%	198.70%	16.90%
Experience	5.48%	6.56%	12.50%	56.56%	10.01%
TECHLOVE	17.26%	19.71%	17.58%	8.75%	10.29%
TTIME	4.65%	4.46%	5.57%	6.58%	0.55%
NHC	49.77%	39.12%	36.95%	-100.93%	74.01%
Sum	100.00%	100.00%	100.00%	100.00%	100.00%

[1] The equation we used to calculate years of education is Education = 10*(less than H.S.) + 12*(H.S. Grad.) + 14*(some college education) + 16*(college grad.) +18*(professional degree) + 20*(Ph. D.).

[2] We also tried three regional divisions: metro area, urban area and rural area, as the percentage of zipcodes in the county with high speed Internet access. As seen in Figure 7 of the Appendix, Beale codes can be grouped into "metro" (0-4), "urban" (5-7) and "rural" (8-9). We ran the same regressions using this three-way regional division and the results are reported in table 4 in the Appendix. The estimated coefficients are almost identical to the estimated coefficients with the two-way regional division.

[3] We also calculated the decomposition using the coefficients from the linear probability model. The conclusions are unchanged. We report this alternative decomposition in Table 6 of the Appendix.

CHAPTER 5. RETURNS TO COMPUTER AND INTERNET TECHNOLOGY USE

I. Background

Many studies have examined the relationships between increased computer or Internet use and productivity growth in labor or wages. Studies have documented substantial growth in information technology related industries, but it is still uncertain whether the Internet or computer uses have a causal effect on the pace of productivity or wage growth.

Autor, Katz, and Krueger (1998) found that the most computer-intensive sectors of the U.S. economy had the fastest growth of demand for skilled labor over the last twenty years. They argued that skill-biased technological and organizational changes generated faster growth in relative skill demand within these industries. Jorgensen (2001) argued that productivity growth in IT-producing industries reduced IT cost and led to a surge in economic growth. It is more difficult to establish whether or not the installation of information technologies has altered the pace of economic growth. Several studies have raised doubts regarding the magnitude of the direct impact of technologies on the productivity growth of the economy. Oliner and Sichel (2000), using data from the Bureau of Economic Analysis (BEA) and the Bureau of Labor Statistics (BLS), reported that the labor productivity growth in the second half of the 1990s could be tied to the information technology. However, Internet usage and e-commerce contributed little to multifactor productivity growth during the period. They also speculated that the productivity gains in the information technology sector should spill over to productivity growth in the economy as a whole.

Gordon (2000), using the same 1995-99 data as Oliner and Sichel (2000), reported that the Internet had little effect on multifactor productivity growth except in durable manufacturing. They argued that since the Internet substitutes for existing communication methods and services such as the telephone, books, and music CDs, its impact on productivity growth on sectors other than manufacturing is small. Litan and Rivlin (2001) were more positive on the role of information technology. They estimated that the total cost saving due to the Internet was about 1-2 percent over the five year period from 1995 to 2000. They argued that much of benefit of the Internet is through improved consumer convenience and expanded choices and improving management efficiencies, rather than in higher productivity and lower prices.

Goss (2001), using data from U.S. Bureau of Economic Analysis, U.S. Bureau of Labor Statistics, and the Current Population Survey, directly estimated the impact of on-the-job Internet usage on output and productivity growth. He found that it added 0.25 percent to annual productivity growth between 1997 and 1999 in all industries. The annual impact of Internet usage on productivity growth was larger in the non-IT intensive industries, 0.52 percent per year, compared a very small 0.03 percent impact on the IT intensive industries. Goss argued that, given current low Internet usage rates, there would be significant room for enhancing productivity growth by increasing job-related Internet usage, particularly in IT intensive industries.

The causal relationship between computer or Internet use and wage is also a matter of debate. Acemoglu (2002) argued that the new information technologies are skill-biased technologies so it has accelerated income inequality. Krueger (1993), using data from CPS and High School and Beyond (HSB), found that returns to on-the-job computer use was about 10 to 15 percent during the 1984-1989 period. Because more educated workers are more likely to use computers in the job, he argued that computer use at work contributed to between one-third and one-half of the increase in the rate of return to education during the period. DiNardo and Pischke (1997), using data from the West German Qualification and Career Survey, found that workers who used a computer at work earned about 17 percent more in 1985 and about 19 percent more in 1991, similar to Krueger's (1993) results. However they also found a 13 percent wage premium from using other tools at work such as pencils, calculators, or telephones in 1991. They argued that the link between wages and computer use at work is weak at best. Levy and Murnane (1996) examined the impact of computers on skill demand in the custodian unit of "Tammany Bank." They found that computerization increased the bank's demand for college graduates, although the increase was induced by scale effects due to the computers' impact on increasing the size of the financial industry rather than the substitution effect from increasing skill requirements within the bank.

Lee and Kim (2004), using the Current Population Survey (CPS) conducted in 1997, 1998, 2000, and 2001, estimated returns to computer use and Internet use at work. They, first, estimated returns to Internet use and returns to computer use in separate wage equations. The estimated returns to Internet use at work decreased from 13 percent in 1997 to 5 percent in 2000. The estimated returns to computer use at work decreased from 14 percent in 1997 to 11 percent in 2001. When returns to Internet use and returns to computer use were estimated together, the returns to Internet use was 8 percent and the returns to computer use was 12 percent in 1997. Only the 1997 CPS included information on both Internet use at work and computer use at work but they calculated an approximate premium to Internet use net of the premium to computer use in 1998 and 2000. This net premium to Internet use was negligible in 1998 and -3 percent in 2000.[1] Oosterbeek (1996), using longitudinal survey data in the Netherlands in 1993, found that workers who use a computer on-the-job earned a wage premium of 10 to 20 percent. The return was associated with computer adoption at work and not to the frequency of computer use. This result stands against the hypothesis that the return to computer use is caused by skill-biased technological change, which implies that more frequent computer use should generate higher returns. Therefore, he argued that the return from computer use could be attributed to unobserved heterogeneity.

A few studies tried to control for individual ability or unobserved heterogeneity. Entorf and Kramarz (1996), using data in France from 1985 to 1987, found that cross-sectional estimates of returns to computer use was about 10 percent. However, when individual-specific fixed effects estimation is used, the wage premium disappeared. They concluded that workers who used computers were already paid better than other workers before using the new technology. Krashinsky (2004), using the data from the CPS from 1990 to 1995, the National Longitudinal Survey of Youth in 1993 and 1994, and the data collected at the Twinsburg Twins Festival in Twinsburg, Ohio, from 1991 to 1995, found that a wage premium associated with computer use that was much smaller than Krueger's (1993) estimates after controlling for individual ability with the Armed Forces Qualifying Test (AFQT) or a person-specific fixed effect.[2] He concluded that workers who use computers at work were already more productive than those who do not. Conflicting evidence was found by Dolton and Makepeace (2004), using panel data from the UK National Child Development Study (NCDS) in 1991 and in 2000. They found that workers who use computers at work earned a wage premium averaging 11% after controlling for individual ability and heterogeneity.

These studies are not free from criticism. It is possible that the endogeneity of new technology use would still be problematic, even after controlling for individual heterogeneity or individual ability. The fixed-effect model may overcorrect for individual heterogeneity, underestimating the wage premium from use of the new technology.

Liu, Tsou and Hammitt (2004), using data from the 1999 Taiwan Social Change Survey, estimated returns to computer use at work. The wage and computer-adoption equations are estimated simultaneously to deal with the endogeneity problem of technology adoption. To identify the computer-adoption equation, they included measures of attitude variables toward computers.[3] When computer use was treated as exogenous in the wage equation, the estimated return was 15 percent, consistent with Krueger's findings. When the simultaneous-equation model was estimated, however, the estimated return to computer use fell to 7.5 percent, but was statistically significant. This results support the implication of Krueger (1993) that computer use at work increases productivity. Goss and Phillips (2002), using data from the CPS from 1997 to 2001, estimated the wage equation and the Internet use equation simultaneously, using a state-level school technology index as an identifying variable for Internet use. They found a statistically significant wage gain for Internet users of 13.5 percent, with a range between 4.9 percent and 16.4 percent depending on the intensity of technology use in the industry. They also found a smaller Internet use wage premium for high-tech industries and a larger Internet use wage premium for low-tech industries. They argued that this result might imply that the wage premium would diminish over time as technology intensity grows.

We investigate this link between use of computer or the Internet and individual wages. We have five measures of technology use: use of the Internet for the job at home, use of a computer for the job at home, use of the Internet at work, use of a computer at work, and use of the Internet for any purposes. We adopt a two step estimation method. First, we estimate a technology adoption decision equation using our five technology use measures as dependent variables and use its predicted value in the earnings function in the second step estimation, the earnings function. Key to the first step of estimation is to have a good instrumental variable to identify the technology adoption. We use a measure of high speed Internet access and a measure of individual attitude toward new information technology as the instruments for the first step.

Empirical model is discussed in section II. Section III discusses the data, section IV discusses the estimation strategy and reports the estimation results. Section V includes the conclusions.

II. Empirical Model

Our analysis begins with the standard log-earnings framework:

1) $\ln y_i = X_i \beta_X + I_i \beta_I + u_i$,

where $ln\, y_i$ is the observed log earnings of the individual i; X_i is a vector of individual demographic and skill characteristics, and I_i is a measure of individual Internet or computer use. The u_i is unobserved individual heterogeneity plus an error component assumed to be randomly distributed. A major concern in estimating equation (1) is that the coefficient on I_i will be biased if $E(I_i\, u_i)$ is not equal to zero. The coefficient of I_i would be upward biased if $\beta_I > 0$ and $Corr(I_i, u_i) > 0$.

In our application, individuals decide whether or not to use the Internet or a computer. This decision involves selecting the option that maximizes expected utility. In functional form it can be written as:

2) $I_i = 1 \; if \; U(X_i^1, Z_i^1, I_i = 1) > U(X_i^0, Z_i^0, I_i = 0)$,

where $U(\bullet)$ is the expected utility of an individual. Although the utility level is not observable, we can observe how the elements of $U(\bullet)$ affect the likelihood of Internet use decision.

If the individual decides whether to use the Internet or a computer in part on the basis of expected returns from the usage, then $U(\bullet)$ will also include u_i. Then there will be a correlation between I_i and u_i, and so direct estimation of equation (1) will be biased. If the vector *Zi* shifts the individual's taste for or cost of Internet use but is not directly related with individual earnings, then Z_i can be used to identify I_i in equation (1). Expected utility from choice I_i can be approximated by

3) $I_i = X_i \alpha_X + Z_i \alpha_Z + v_i$,

where v_i is a random disturbance.

To deal with this endogeneity problem, we use the number of high-speed Internet providers in a zip code area as an instrument for Internet or computer use. We assume that $E(Z_i v_i) = 0$. Most types of high speed Internet, such as cable modem lines or DSL, are at least five times faster than typical telephone modem lines. Availability of high-speed Internet will alter the productivity of Internet use. In addition, as the number of high-speed Internet providers increases, the competition among them will be intensified and it will decrease the cost of high-speed Internet access or will increase the quality of services.

If the v_i is drawn independently from an extreme value distribution, then equation (3) can be estimated using probit model. By inserting the predicted probability of Internet use of individuals from the estimate into equation (1) in place of the endogenous I_i, we can get unbiased estimate of β_I.

III. Description of the Data Set

The data set analyzed in this paper is a subset of that used in the previous chapter on Internet technology adoption. We use only 2001 data since the Center for Communication Policy at UCLA included questions about individual income level only in the 2001 survey. The descriptive statistics for the data set are reported in table 1. Average years of education was 13.5 years, roughly consistent with the U.S. average (Baier et al. (2004)).

As before, we define four types of new information technology uses: use of the Internet for the job at home (IJH), use of a computer for the job at home (CJH), use of the Internet at work (IW), use of a computer at work (CW), and use of the Internet for any purposes (IA). About 47 percent used a computer at work and 13 percent used a computer at home for work purposes. Twenty-seven percent reported using the Internet at work and about 11 percent used the Internet at home for work. Individuals averaged just over $33,000 in annual income. The income distribution is skewed right: the majority earned less than $50,000 but 0.2 % earned more than $200,000.

As in previous chapter, we define the high speed Internet access as the number of high speed Internet providers in a zipcode. We have four measures of high speed Internet access, the number of high speed Internet providers in a zipcode (NHZ), the average of NHZ by a county (NHC), availability of high speed Internet access in a zipcode (DHZ), and the average of NHZ by county (DHC). We only use NHC in our analysis for the same reason in the previous chapter; results using the other three measures are similar. The average number of high speed Internet providers was 5.2 and about 93 percent of all zipcode areas had at least one or more providers.

IV Estimation and Results

We analyze returns to technology use under two different assumptions: that the technology adoption is exogenous and that the technology use is an endogenous choice based on anticipated returns. When exogenous technology use is assumed, we directly estimate the earnings function using a dummy variable indicator for each technology use in place of I_i in equation (1). That is,

(4) $\ln y_i = X_i\, \beta_X + \beta_{IT}\, IT_i + u_i$

where IT_i represents one of the four job related information technology uses; the vector X_i includes variables such as experience, experience squared, race, gender, marital status, rural region indicator, and years of education.[4]

Under the endogenous technology use assumption, first, we estimate the following base technology adoption equation (3) using probit for the four binary dependent variables; use of the Internet for the job at home, use of a computer for the job at home, use of the Internet at work, and use of a computer at work:

(5) $IT_i = X_i\, \beta + \gamma\, techlove_i + \theta_1\, NHC_i + \theta_2\, NHC_i^2 + v_i$.

IT_i and the vector X_i are same as in equation 5. NHC_i and $techlove_i$ are two instrumental variables: the average number of high speed Internet providers in a county and the technology attitude dummy variable, respectively. The predicted value of IT_i, $\widehat{IT}_i$, is inserted in place of IT_i in equation (4) to get

(6) $\ln y_i = X_i\,\beta_X + \beta_I\,\widehat{IT}_i + e_i$

Earnings and Exogenous Computer or Internet Use

Tables 2 reports the estimation results of the log-earnings function with exogenous technology using various specifications. First five columns of Table 2 report estimates of earnings functions with only one of the four technology use variables. We also extend equation (4) to allow multiple use of information technology. The rest columns report the estimation results of specifications with two of the five technology use variables.

Estimates are consistent with previous studies. Earnings are lower for minorities and for those living in rural areas. Males and married persons get higher pay. Turning to the main focus, estimated returns to technology adoption are substantial: 53.5 percent for Internet use at home; 44.8 percent for computer use at home; 45 percent for Internet use at work; and 47.5 percent for computer use at work. The return to any Internet use is substantially smaller and is only statistically significant at the 10 % level. All four estimates of returns to work related technology uses are highly statistically significant. It seems clear that these estimates must overstate the true returns as they are even larger than the estimates uncorrected for endogeneity that were reported by Krueger (1993) and Dolton and Makepeace (2004).

Estimated log earnings functions incorporating alternative combinations of exogenous technology adoption measures are reported in the last 10 columns of Table 2. Returns to technology use are larger at work than at home. Estimated returns to Internet use is larger than computer use when they are used at home, estimated returns to computer use is larger then Internet use when they are used at work. In general, when two technology uses are included, the returns to individual uses falls in comparison to the estimates when the technologies are treated separately. That suggests that one reason for the unreasonably large estimated returns to individual technology adoption in the top section of Table 2 is that these technologies are correlated with a menu of adoptions and not just an isolated technology. Nevertheless, the returns to Internet and computer use are still quite large when multiple adoptions are allowed.

Earnings and Endogenous Computer or Internet Use

Table 3 reports the estimation results of log-earnings function controlling for the endogenous use of a computer or the Internet. The probit estimation results of technology adoption decision is also reported in Table 4 in Appendix.

We tested the overidentification restrictions of the instrumental variables used to identify the technology adoption equation. We regressed the residuals from the log-earnings function (6) on the instruments NHC, HNC^2, and *techlove*. That is,

(7) $\hat{e}_i = \alpha_0 + \alpha_1 NHC_i + \alpha_2 NHC_i^2 + \alpha_3\, techlove_i + \eta_i$,

where $\hat{e}_i$ is the estimated residual from eq. (6) and η_i is a random disturbance. If these NHC, NHC^2 or *techlove* are correlated with earnings, then $\alpha_1 \neq 0, \alpha_2 \neq 0, or\, \alpha_3 \neq 0$. The results are reported in Table 5 in Appendix. We cannot reject the joint hypothesis that NHC, NHC^2, or *techlove* are endogenous.

The coefficients of the variables other than technology use are qualitatively and quantitatively similar between tables 2 and 3. However the estimated coefficients on computer or Internet use are very different. The returns to IT uses for job at home are positive but the coefficient is substantially smaller when endogeneity is controlled. The other two returns to technology use at work change sign. The estimated return to any Internet use is larger, but the coefficients of all five technology uses are statistically insignificant. After controlling for endogenous decisions regarding technology adoption, the wage premium from computer or Internet use disappears.

These results suggest that the large estimated returns to technology adoption that are found when the adoption is viewed as exogenous are due to unmeasured factors that are correlated with the adoption. When the choice to adopt is controlled, the estimated returns to adoption shrink in both sign and significance, and we cannot reject the null hypothesis that there is a zero return to adoption. Thus, while adoption is strongly tied to the availability of high-speed Internet in the home county, the higher income of adopters is due to factors that raise both the probability of adoption and earnings and not to the adoption *per se*. As to the issue of an urban-rural Internet divide, differential access between urban and rural markets does explain differences in adoption rates, but rural residents are not getting lower earnings as a result of their lower adoption rates.

V. Conclusions

Since Krueger (1993) and DiNardo and Pischke (1997), the existence of wage premium of a computer or the Internet use has been in debate. Most studies that treat information technology use as exogenous found positive and statistically significant returns. On the other hand, studies that controlled for unobserved individual factors or the endogeneity between earnings and information technology use found much smaller and sometimes insignificant returns.

In this study, we dealt with the endogeneity problem by using high-speed Internet access as an instrument for technology adoption. The number of high speed Internet providers by county increased the probability of using a computer or the Internet for work while at home and it also increased the probability of using the Internet at work. After controlling for the endogeneity of technology adoption, we conclude that the wage premium disappears.

Card and DiNardo (2002) criticized the view that rising wage inequality is due to skill-biased technological change as being inconsistent with stabilized wage inequality in 1990s when computer related technologies were still advancing. Our findings are consistent with their conclusion that the premium associated with using computers or the Internet is due to higher wage workers using computers and not because computers raise wages for those using them. The results suggest that the large estimated returns to technology adoption that are found when the adoption is viewed as exogenous are due to unmeasured factors that are correlated with the adoption. When the choice to adopt is controlled, the estimated returns to adoption shrink in both sign and significance. Thus, while adoption is strongly tied to the availability of high-speed Internet in the home county, the higher income of adopters is due to factors that raise both the probability of adoption and earnings and not to the adoption *per se*. We also found that differential access between urban and rural markets does explain differences in adoption rates, but rural residents are not getting lower earnings as a result of their lower adoption rates.

References

Autor, David H., Kats, Lawrence F., and Alan B. Krueger, "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, vol. 113, no. 4 (1998), 1169-1213.

Baier, Scott, Mulholland, Sean, Tamura, Robert, and Chad Turner, "Income and Education of States of the United States: 1840-2000," Working paper, Clemson University, 2004.

Card, David and John E. DiNardo, "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," 2002, *Journal of Labor Economics*, vol. 20, no. 4, 733-783.

DiNardo, John E., and Jorn-Steffen Pischke, "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?" *Quarterly Journal of Economics*, vol. 112, no. 1 (1997), 291-303.

Dolton, Peter, and Gerry Makepeace, "Computer Use and Earnings in Britain," *Economic Journal*, vol. 114 (2004), 117-129.

Gordon, Robert J., "Does the "New Economy" Measure up to the Great Inventions of the Past?" *Journal of Economic Perspective*, vol. 14, no. 4 (2000), 49-74.

Goss, Ernest, "The Internet's Contribution to U.S. Productivity Growth," *Business Economics*, vol. 36, no. 4 (2001), 32-42.

Krashinshy, Harry A., "Do Marital Status and Computer Usage Really Change the Wage Structure?" *Journal of Human Resources*, vol. 39, no. 3 (2004), 774-791.

Krueger, Alan B., "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989," *Quarterly Journal of Economics*, vol. 108, no. 1 (1993), 33-60.

Lee, Sang-Hyup, and Jonghyuk Kim, "Has the Internet Changed the Wage Structure too?" *Labour Economics*, vol. 11 (2004), 119-127.

Levy, Frank, and Richard J. Murnane, "With What Skills Are Computers a Complement?" *American Economic Review*, vol. 86, no. 2 (1996), 258-262.

Litan, Robert E., and Alice M. Rivlin, "Projecting the Economic Impact of the Internet," *American Economic Review*, vol. 91, no. 2 (2001), 313-317.

Liu, Jin-Tan, Tsou, Meng-Wen, and James K. Hammitt, "Computer Use and Wages: Evidence from Taiwan," *Economics Letters*, vol. 82 (2004), 43-51.

Oliner, Stephen D., and Daniel E. Sichel, "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives*, vol. 14, no. 4 (2000), 3-22.

Oosterbeek, Hessel, "Returns from Computer Use: A Simple Test on the Productivity Interpretation," *Economics Letters*, vol. 55 (1997), 273-277.

Table 1. Descriptive statistics of the Working Sample

Variable	Obs	Mean	Std. Err.
Male	1259	0.472	(0.014)
Married	1259	0.583	(0.014)
Age	1259	43.4	(0.329)
Experience	1259	23.9	(0.335)
Education	1259	13.5	(0.066)
Technology Interest	1259	0.297	(0.013)
White	1259	0.747	(0.012)
Black	1259	0.122	(0.009)
Asian	1259	0.028	(0.005)
Native Am.	1259	0.023	(0.004)
Minority	1259	0.211	(0.012)
Rural Area	1259	0.400	(0.014)
Internet Job Home	1259	0.111	(0.009)
Computer Job Home	1259	0.129	(0.009)
Internet at Work	1259	0.269	(0.013)
Computer at Work	1259	0.466	(0.014)
Internet For Any Purposed	1259	0.734	(0.012)
NHZ	924	5.224	(0.112)
DHZ	924	0.930	(0.008)
NHC	924	5.286	(0.102)
DHC	924	0.932	(0.007)
Income (2001 dollar)	845	33224	(879)

Table 2. Log-earnings Function Estimation Results with Exogenous Technology Uses

Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
IJH	0.533	(0.064)								
CJH			0.448	(0.070)						
IW					0.450	(0.056)				
CW							0.475	(0.051)		
IA									0.129	(0.073)
Minority	-0.070	(0.065)	-0.071	(0.066)	-0.041	(0.068)	-0.037	(0.066)	-0.071	(0.069)
Married	0.456	(0.050)	0.113	(0.065)	0.106	(0.065)	0.092	(0.061)	0.127	(0.066)
Male	0.120	(0.065)	0.461	(0.050)	0.442	(0.051)	0.475	(0.050)	0.457	(0.052)
Rural Background	-0.138	(0.052)	-0.139	(0.052)	-0.143	(0.049)	-0.162	(0.049)	-0.155	(0.052)
Education	0.083	(0.011)	0.084	(0.011)	0.077	(0.011)	0.079	(0.010)	0.104	(0.011)
Experience/10	0.226	(0.083)	0.227	(0.082)	0.265	(0.078)	0.182	(0.079)	0.287	(0.081)
$(Experience/10)^2$	-0.042	(0.017)	-0.042	(0.017)	-0.047	(0.016)	-0.027	(0.016)	-0.054	(0.017)
Constant	8.496	(0.163)	8.483	(0.163)	8.458	(0.164)	8.400	(0.157)	8.115	(0.158)
R^2	0.361		0.349		0.379		0.406		0.311	

Obs: 647

Table 2. (cont'd)

Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
IJH	0.527	(0.203)	0.298	(0.074)	0.357	(0.064)	0.524	(0.064)		
CJH	0.006	(0.198)							0.258	(0.075)
IW			0.340	(0.061)					0.370	(0.060)
CW					0.408	(0.052)				
IA							0.101	(0.072)		
Minority	-0.070	(0.065)	-0.041	(0.066)	-0.030	(0.063)	-0.054	(0.065)	-0.038	(0.067)
Married	0.120	(0.065)	0.105	(0.064)	0.088	(0.060)	0.115	(0.064)	0.099	(0.064)
Male	0.456	(0.050)	0.443	(0.050)	0.469	(0.048)	0.452	(0.050)	0.444	(0.050)
Rural Background	-0.138	(0.052)	-0.133	(0.050)	-0.143	(0.049)	-0.130	(0.052)	-0.131	(0.050)
Education	0.082	(0.011)	0.070	(0.011)	0.065	(0.010)	0.079	(0.011)	0.068	(0.011)
Experience/10	0.226	(0.082)	0.238	(0.079)	0.160	(0.079)	0.232	(0.082)	0.237	(0.079)
$(\text{Experience}/10)^2$	-0.042	(0.017)	-0.042	(0.017)	-0.023	(0.017)	-0.042	(0.017)	-0.041	(0.017)
Constant	8.497	(0.162)	8.573	(0.166)	8.591	(0.160)	8.455	(0.165)	8.591	(0.165)
R^2	0.361		0.392		0.429		0.364		0.391	

Table 2. (cont'd)

Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
IJH										
CJH	0.257	(0.070)	0.440	(0.070)						
IW					0.206	(0.066)	0.445	(0.057)		
CW	0.417	(0.052)			0.362	(0.063)			0.471	(0.051)
IA			0.102	(0.072)			0.032	(0.074)	0.025	(0.071)
Minority	-0.031	(0.064)	-0.054	(0.066)	-0.026	(0.066)	-0.036	(0.068)	-0.033	(0.066)
Married	0.085	(0.061)	0.109	(0.065)	0.089	(0.062)	0.104	(0.064)	0.091	(0.061)
Male	0.472	(0.049)	0.456	(0.050)	0.463	(0.050)	0.441	(0.051)	0.474	(0.050)
Rural Background	-0.146	(0.049)	-0.130	(0.053)	-0.152	(0.048)	-0.141	(0.050)	-0.160	(0.049)
Education	0.068	(0.011)	0.080	(0.011)	0.071	(0.011)	0.076	(0.011)	0.078	(0.011)
Experience/10	0.163	(0.078)	0.233	(0.081)	0.198	(0.078)	0.267	(0.078)	0.184	(0.078)
$(\text{Experience}/10)^2$	-0.023	(0.016)	-0.042	(0.017)	-0.030	(0.016)	-0.047	(0.016)	-0.027	(0.016)
Constant	8.555	(0.162)	8.441	(0.165)	8.479	(0.160)	8.443	(0.167)	8.389	(0.159)
R^2	0.419		0.352		0.416		0.379		0.406	

Table 3. Log-earnings Function Estimation Results with Endogenous Technology Uses

Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
IJH	0.177	(0.405)								
CJH			0.121	(0.387)						
IW					-0.081	(0.562)				
CW							-1.183	(0.856)		
IA									0.291	(0.432)
Minority	-0.084	(0.070)	-0.085	(0.073)	-0.101	(0.095)	-0.243	(0.135)	-0.068	(0.074)
Married	0.125	(0.070)	0.127	(0.071)	0.137	(0.071)	0.164	(0.080)	0.111	(0.071)
Male	0.457	(0.051)	0.459	(0.055)	0.466	(0.062)	0.410	(0.070)	0.450	(0.058)
Rural Background	-0.158	(0.055)	-0.160	(0.057)	-0.173	(0.072)	-0.228	(0.083)	-0.150	(0.063)
Education	0.100	(0.022)	0.102	(0.024)	0.115	(0.042)	0.190	(0.059)	0.095	(0.024)
Experience/10	0.264	(0.091)	0.267	(0.095)	0.285	(0.092)	0.509	(0.193)	0.301	(0.087)
$(\text{Experience}/10)^2$	-0.051	(0.019)	-0.052	(0.021)	-0.057	(0.021)	-0.118	(0.049)	-0.055	(0.018)
Constant	8.273	(0.299)	8.246	(0.320)	8.105	(0.417)	7.578	(0.459)	8.096	(0.178)
R^2	0.306		0.306		0.306		0.308		0.307	

Obs: 647

Table 4. Probit Estimation Results of Each Information Technology Adoption

	IJH		CJH		IW		CW		IA	
Variables	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Minority	-0.305	(0.217)	-0.354	(0.208)	-0.415	(0.162)	-0.353	(0.144)	-0.299	(0.155)
Married	0.316	(0.130)	0.368	(0.130)	0.156	(0.108)	0.057	(0.103)	0.267	(0.113)
Male	0.184	(0.128)	0.128	(0.123)	0.173	(0.104)	-0.126	(0.100)	0.204	(0.111)
Education	0.233	(0.025)	0.235	(0.025)	0.230	(0.021)	0.199	(0.022)	0.246	(0.035)
Experience/10	0.820	(0.248)	0.752	(0.256)	0.277	(0.192)	0.645	(0.179)	-0.505	(0.208)
$(\text{Experience}/10)^2$	-0.224	(0.053)	-0.201	(0.057)	-0.098	(0.039)	-0.176	(0.036)	0.045	(0.039)
Rural Area	-0.093	(0.218)	-0.101	(0.202)	-0.103	(0.166)	-0.160	(0.150)	-0.165	(0.168)
Technology Interest	0.158	(0.139)	0.091	(0.134)	-0.099	(0.112)	0.017	(0.106)	0.062	(0.118)
NHC/10	1.405	(0.788)	1.356	(0.763)	0.608	(0.653)	-0.194	(0.633)	1.316	(0.688)
$(\text{NHC}/10)^2$	-0.862	(0.602)	-0.770	(0.608)	-0.290	(0.499)	0.095	(0.497)	-1.103	(0.544)
Constant	-5.701	(0.508)	-5.567	(0.499)	-3.997	(0.408)	-2.890	(0.397)	-1.992	(0.557)
Pseudo R^2	0.227		0.225		0.175		0.142		0.185	

Obs: 924

Table 5. Overidentification Test for Instrumental Variables Results

variable	IJH residual		CJH residual		IW residual		CW residual		IA residual	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
NHC	-0.058	(0.052)	-0.055	(0.052)	-0.056	(0.052)	-0.047	(0.052)	-0.058	(0.052)
NHC^2	0.039	(0.032)	0.040	(0.032)	0.042	(0.032)	0.035	(0.032)	0.031	(0.032)
Techlove	-0.003	(0.003)	-0.003	(0.003)	-0.003	(0.003)	-0.003	(0.003)	-0.003	(0.003)
Constant	-0.064	(0.074)	-0.067	(0.075)	-0.075	(0.075)	-0.056	(0.074)	-0.051	(0.074)
R^2	0.0059		0.0058		0.0062		0.0046		0.0045	
χ^2_3	3.82		3.75		4.01		2.98		2.91	

χ^2_3 = Obs.*R^2; χ^2_3 at 10% significance level = 6.25

[1] Their approximation appears somewhat unreliable. Applying their method to calculate the net premium to Internet use in 1997 results in a 2.4 percent return compared to the 7.7 percent return from the direct estimation. That suggests that their approximation may be severely biased downward.

[2] Using the twins data, he found about 7 percent of the wage premium of computer use at work after controlling for person-specific fixed-effect but it was not statistically significant.

[3] One variable is about the attitude of an individual on the computer in general and the other variable is about the attitude of an individual on how to use the computer at work or for leisure. For detail see Liu et al. (1997).

[4] We also estimate specifications with more than one technology use. IT_i, in those specifications, will be a vector.

CHAPTER 6. GENERAL CONCLUSIONS

We investigated two plausible factors that may have contributed to rising wage inequality over the past 30 years: returns to higher education and returns to adoption of information technologies.

In chapter 2, we estimated returns to advanced degrees and direct returns to mathematics and verbal skills by college major. We found that OLS estimates of returns to graduate education are underestimated when controls for average ability by major are missing. When controls for average ability are included, returns to post-graduate education rise by about 11%. As argued in the paper, these are lower bound estimates of the true return if the covariance between years of schooling and individual ability have the same sign as the associated covariance between major average ability and years of schooling. Our result differs from findings in previous studies that focused on lower levels of education. At lower levels of schooling, students with more mathematics ability are more likely to attend college because returns to cognitive skills and returns to additional schooling are relatively larger than the opportunity cost of schooling. After attaining the bachelor's degree, my results suggest that the opportunity cost of post-graduate education dominate the returns to schooling for those with the most mathematical ability, but not for those with the greatest verbal ability

Average GRE verbal and quantitative scores by college major play an important role in explaining earnings variation. Returns to the major average GRE quantitative score are positive while returns to the major average GRE verbal score are negative. Rising GRE quantitative scores in the 1980s and falling GRE verbal scores since the mid 1970s have both contributed to the rising earnings for college graduates. The simulated earnings differential attributed to changes in the average GRE quantitative and verbal scores by college major is about 2.6 percent or $1,609 in constant 1993 dollars.

Changes in the distribution of major on earnings by cohort also have had an impact on the earnings for college graduates. Researches have found that college majors play important role in explaining wage variation. Some majors get paid more than others. The simulated earnings paths in Figure 7 show that choice of college major caused average returns to college to fall in the 1960's and early 70's, and increase since the mid 1970s. Freeman (1976) argued that low returns to college degrees in late 1960's and early 1970's were due to excess supply of college graduates, but the low returns were also apparently due to a disproportionate concentration of college students in lower paying majors.

In chapter 3, we investigated the same topic as in chapter 2 but focusing on the role of ability in sorting college graduates into different degree programs. We identified the schooling decision equation using parents' education level and schooling cost measures as instruments. This two step approach allowed us to investigate how the major level skills affect schooling decision as well as its direct returns.

Estimated returns to advanced degrees are positive and significant and are even larger than those found in chapter 2. Because of the inability to control for individual ability within the major, the chapter 2 results were lower bound estimates. Controlling for this unobserved source of heterogeneity in chapter 3 does lead to larger returns as hypothesized in chapter 2.

The estimation results pointed out an interesting role for cognitive skills in the market for advanced degrees. Students in majors with higher average quantitative GRE scores are less likely to attend graduate school, even though such students presumably are more likely to be successful in graduate education. The opposite happens for verbal skills. This leads to a sorting effect whereby students whose cognitive skills would suggest lower earnings at the bachelor's level are more likely to attend graduate school. This sorting effect appears to be part of the cause of the downward bias in estimated returns to graduate education—the average earnings of those who do not go to graduate school overstate the opportunity costs of graduate education for those who do pursue advanced degrees.

In chapter 4 we examined factors that might affect the adoption of new information technologies and urban-rural difference in the adoption. Adoption is positively affected by schooling, and is also correlated with gender, race, and marital status. Individuals who have a favorable impression toward new technologies are also more likely to use computers and the Internet. However, even when those factors are controlled, local access to high speed Internet plays an important role in the technology adoption decision. It increases the probability of using computers and the Internet for work from home and also increases the likelihood of using the Internet at work. The pattern of marginal effects is consistent with the presumption that access is more important for telecommuting than for the use of the technology at work. There is no statistically significant difference in the coefficients governing the technology adoption process between urban and rural areas. This implies that the observed difference in IT uses can be explained by rural-urban differences in average human capital, demographics and Internet access. Difference in the number of high speed Internet access by county explains 44 percent of the difference in the Internet use for job at home, 40 percent of the gap in computer use for the job at home, 32 percent of the gap in Internet use at work and 50% of the gap in Internet use for any purposes, but only -7.6 percent of the gap in computer use at work. Internet access matters most in encouraging any Internet use and work from home. Results suggest that as high-speed Internet access increase, there will be increased substitution of telecommuting for commuting.

In chapter 5 we investigated returns to computer and Internet uses. Previous studies have found positive and statistically significant returns to computer or Internet use when adoption is assumed to be exogenous. On the other hand, studies that controlled for unobserved individual factors or the endogeneity between earnings and information technology use found much smaller and sometimes insignificant returns. In this chapter, we dealt with the endogeneity problem by using high-speed Internet access as an instrument for technology adoption. The number of high speed Internet providers in the home county increased the probability of home use of a computer or the Internet for work purposes and it also increased the probability of using the Internet at work. After controlling for the endogeneity of technology adoption, we conclude that the wage premium disappears.

Our findings are consistent with their conclusion that the premium associated with using computers or the Internet is due to higher wage workers using computers and not because computers raise wages for those using them. The results suggest that the large estimated returns to technology adoption that are found when the adoption is viewed as exogenous are due to unmeasured factors that are correlated with the adoption. When the choice to adopt is controlled, the estimated returns to adoption shrink in both sign and significance. Thus, while adoption is strongly tied to the availability of high-speed Internet in the home county, the higher income of adopters is due to factors that raise both the probability of adoption and earnings and not to the adoption *per se*. We also found that differential access between urban and rural markets does explain differences in adoption rates, but rural residents are not getting lower earnings as a result of their lower adoption rates.

Reference

Freeman, Richard B. *The Overeducated American* (New York: Academic Press., 1976a).