HUMAN CAPITAL, ABILITY, AND THE DECISION TO BECOME AN ENTREPRENEUR

S. Ramakrishna Velamuri and S. Venkataraman
Darden Graduate School of Business Administration
University of Virginia
e-mail: velamuris00@darden.virginia.edu or rvelamuri@hotmail.com
Telephone: 1-804-9243212

April 16, 2002

We gratefully acknowledge the financial support of the Batten Institute of the Darden School of Business Administration, University of Virginia, for carrying out this research. We wish to thank Paul Allison, Srinath Baba, Greg Fairchild, Michael Moore, S. Sarasvathy, Scott Shane, Malathi Velamuri, and the participants of the 11th Annual Global Entrepreneurship Research Conference held in Lyon, France, for helpful comments on earlier drafts of this paper. We also want to thank Tecla Loup of the Panel Study of Income Dynamics for several clarifications on the dataset, and Prasanna Velamuri for assistance with downloading data.

HUMAN CAPITAL, ABILITY, AND THE DECISION TO BECOME AN ENTREPRENEUR

Abstract

We explore the relationship between the probability of a transition from paid work to self-employment and three explanatory variables: paid income, education, and ability. We use panel data for heads of households from the PSID SRC sample for eight pairs of years, from 1985-1986 to 1992-1993. Our results show that the transition to self-employment is a decreasing function of income and an increasing function of education, and that ability is a better predictor of self-employment than are income or education. We show that the relationship between ability and self-employment is U shaped.

JEL Classification: I2, J2, J3, J6.

I. Introduction

In this paper we address the question of who takes up self-employment and under what circumstances. There have been several studies that have found correlations between the probability of a transition from paid-employment to self-employment on the one hand, and income levels (Amit, Muller Cockburn, 1995; Evans and Leighton, 1989), human capital (Rees and Shah, 1986), race (Bates, 1997; Borjas and Bronars, 1989), wealth (Bates, 1997; Evans and Jovanovic, 1989), and gender (Devine, 1994) on the other. In this paper, we examine the theoretical arguments and empirical evidence for labor income and education as predictors of the transition to self-employment. We find, consistent with previous scholarship, that income is negatively related, and education positively related, to the likelihood that an individual will make the transition from paid employment to self-employment. These relationships hold when income and education are used separately, as well as when they are used together in the logistic regression models.

However, if labor income is determined by a worker's productivity, then the research on the ability bias suggests that innate abilities may explain both investment in education and productivity. According to Behrman and Rosenzweig (1999), "ability has been used as the rubric for all unmeasured earnings endowments, which may include genetic endowments of ability, pre-school human capital investment, or motivation." The focus of the ability bias literature (for example, Angrist and Krueger, 1991; Behrman, Hrubec, Taubman, and Wales, 1980; Blackburn and Neumark, 1992, 1995; Card 1999; Griliches and Mason, 1972) has been mainly on obtaining accurate parametric estimates of the returns to schooling. It is acknowledged that the decision to invest in schooling could very well be endogenous, in that high ability individuals choose to invest in schooling to signal their abilities to the market. Furthermore, there is a case for a direct

effect of ability on productivity, in addition to the indirect effect through education. We estimate an ability gap in labor income - the premiums and penalties that individuals receive for their innate abilities, over and above human capital, demographic, and spatial variables, and study its impact on the likelihood of becoming self-employed. If every individual had the same ability level, then human capital, race, sex, and spatial variables (industry, occupation, and state of residence) would correctly predict labor income. In other words, if the only distortions in the labor market are due to sex, race, and spatial variables, then the difference between actual income and predicted income can be taken as a rough measure of ability. We find that individuals of low and very high abilities are more likely to enter self-employment than those of moderate abilities. This result is consistent with prospect theory (Kahneman and Tversky, 1979), which suggests that individuals evaluate gains and losses with respect to a reference point. We submit that individuals evaluate their incomes in paid work based on either their human capital or their marginal productivity, whichever is higher. We argue that low ability individuals are penalized for their lack of ability and their incomes are lower than the incomes of individuals with similar human capital and average ability. Their incomes may seem lower to them relative to their reference point, which is their human capital. On the other hand, individuals of very high ability are likely to express dissatisfaction with their labor incomes for a different reason: their incomes, although very high in comparison to those of individuals with moderate and low abilities, are still lower than their own marginal productivity. This is because in paid employment, they are agents and are forced to share their marginal product with their principals.

A more nuanced understanding of who becomes an entrepreneur and under what circumstances can illuminate several issues of importance. First, there is the finding by Blau (1987) that the more developed the economy, the lower the percentage of individuals in non-

agricultural self-employment. This finding holds both for longitudinal US data over a hundred years until 1970, as well as for a cross-sectional comparison of countries at varying stages of economic development. This suggests that the quality of entrepreneurship matters more than sheer numbers. Specifically, it would appear that it is in the interest of society for high human capital individuals to take up self-employment, and to bid low human capital individuals away from subsistence self-employment. Kangasharju and Pekkala (2001) have found that firms founded by highly educated individuals have higher growth potential. Second, the reversal of the trend since the early 1970s of declining self-employment suggests that structural changes in the economies of the developed nations favor industries in which economies of scale are not as important as in others (Blau, 1987). Other possible explanations for this reversal are the increase in the percentage of women and minorities in the labor force; the self-employment rate for both these groups has been rising steadily (Leung and Robinson, 2001). Third, since there is no market for self-employment (in the way in which there is a market for managers and other workers), individuals who wish to become self-employed must necessarily hire themselves (Knight, 1921). This leads to the question of why individuals in paid employment make the transition to self-employment, and is all the more interesting because very few individuals join self-employment when they join the labor force for the first time; the percentage of individuals becoming self-employed increases as a function of age (Rees and Shah, 1986).

We propose to study the self-employment decision, and more particularly, how it is influenced by income in paid work, education, and the innate abilities of individuals. Our study examines the likelihood of self-employment using three measures. First, we use income percentiles to study if individuals at low-income percentiles are more likely to take up self-employment. We use income percentiles rather than log dollar income because of the

longitudinal nature of our data in which dollar incomes in different years are not directly comparable. Through this first measure, we essentially replicate the studies done by Amit, Muller, and Cockburn (1995) and Evans and Leighton (1989). Second, we study the relationship between education and the likelihood of self-employment. We explore the negative relationship of the probability of self-employment to income, and its positive relationship to education. Finally, we measure an ability gap, which we define as the difference between the actual income percentile of an individual, and the predicted income percentile based on human capital (tenure, labor market experience, education), demographic (race, sex), and spatial measures (state of residence, industry, and occupation). Thus, the ability gap is a residual measured in income percentile units that remains after human capital, demographic, and spatial variables have been controlled for. The ability gap could be either positive or negative. A positive gap indicates a premium for ability; a negative gap indicates a discount for the lack of it. We then examine which of the three measures - income percentile, education, or ability gap - is the strongest predictor of the transition from paid work to self-employment.

Section II briefly reviews the literature on the relationship between self-employment, income, education, and ability, and develops the theoretical arguments. Section III gives an overview of the methodology and data source, defines the variables chosen and outlines the regression models. Section IV presents the results, and Section V summarizes the conclusions.

III. Literature review – income, education, ability, and the likelihood of self-employment

Income and education as predictors of the transition to self-employment

Many scholars have written on the likelihood of an individual's transition to self-employment using as explanatory variables income in paid work, current employment status,

human capital variables, age, race, and sex. We briefly review the findings on income and education as predictors of the transition to self-employment. Amit, Muller and Cockburn (1995), using data from the Canadian Labor Market Survey (LMAS), found support for the hypothesis that the lower a person's opportunity cost (defined as wage earnings in the period immediately prior to the self-employment decision), the more likely she is to take up self-employment. Consistent with the finding of Amit et. al. (1995), Evans and Leighton (1989) reported from their study based on the National Longitudinal Study of Young Men (NLS) and the Current Population Surveys that lower paid wage workers are more likely to enter self-employment at a point in time, all else equal. Borjas and Bronars (1989) found that the self-employed are more likely to be college educated, a finding consistent with that of Rees and Shah (1986), which was based on data from the UK. Knudsen and McTavish (1989) also concluded that more educated individuals are more likely to take up self-employment.

We attempt to delve deeper into the positive relationship between self-employment and education, and the negative relationship between self-employment and income. In particular, we argue that there are strong arguments in favor of a U-shaped relationship between income and self-employment: theoretical arguments can be made to support both a positive and a negative relationship between income and the likelihood of a transition to self-employment.

To support a negative relationship, one could say that the decision to take up self-employment on a full time basis is a decision of considerable magnitude for those who are in paid employment. It requires sacrificing current income based on expectations of future income in the new career. A rational actor model would postulate that a paid worker analyzes all possible career alternatives, the payoffs of each alternative over time, and chooses that alternative whose net present value, weighted by the probability of its occurrence, is the highest. However, a

complete list of career states, their probabilities, and their corresponding payoffs is not readily available to any paid worker. Even if this information were available, the costs for the worker of obtaining all the information and her limited computational ability would lead to her behaving in a boundedly rational way (Simon, 1955). Current income in paid employment is immediately available information that is a reasonable surrogate variable for the opportunity cost of becoming self-employed (see Amit, Muller, Cockburn, 1995: 97). Income in paid work is therefore the reservation wage against which the option of becoming self-employed is evaluated. Inertial forces favor status quo, and paid employment is no exception. The old adage of a bird in hand being worth two in a bush is likely to apply to individuals in paid work contemplating self-employment. Prospect theory (Kahneman and Tversky, 1979) also provides support for this position by proposing that individuals are risk averse over gains, which means that an individual in paid work is more likely to overweight the prospect of sacrificing current income relative to foregoing a higher future income in self-employment. This phenomenon has also been referred to as the "endowment effect" (Knetsch and Sinden, 1984).

On the other hand, several equally strong arguments can be put forward in favor of a positive relationship between paid income and the likelihood of self-employment. First, it can be argued that higher income individuals will see more valuable opportunities than lower income individuals, based on the finding that prior knowledge influences the quality of opportunities recognized. Shane (2000) studied eight sets of entrepreneurs who pursued eight different entrepreneurial opportunities arising from a single invention, a three dimensional printing process that was developed by the faculty and doctoral students at the MIT Labs. He found that prior knowledge is an important factor in determining the opportunities that different individuals will recognize and pursue. High-income individuals, who possess higher value knowledge, may

perceive the expected value of their self-employment opportunities to be higher than the expected value of continuing in paid employment. Second, higher income individuals are more likely to have accumulated higher levels of wealth during their careers. This higher wealth could provide protection against possible downside losses in self-employment. Bates (1997) found that one reason for the lower rates of self-employment entry for blacks relative to whites is their lower wealth holdings. Evans and Jovanovic (1989) found that capital market imperfections mean that the wealthy are more likely to take up self-employment. The relationship between wealth and the likelihood of self-employment may well be curvilinear; at very high levels of wealth, the individual may not participate in the labor force at all. This is especially true if the individual has received the wealth in the form of an inheritance (see Holtz-Eakin, Joulfaian, and Rosen, 1993). Third, based on the assortative matching theory of Becker (1973), married high income individuals are also likely to have spouses who are employed in high-income positions, based on the fact that marriages tend to take place between people of similar education and social and economic standing. These dual high-income households would have greater income stability than households in which there is either only one earner, or both earners are in jobs with modest incomes, thus making it easier for one of the spouses to venture into self-employment. Fourth, Blau (1987) suggested that higher marginal tax rates could be a key determinant for the self-employment decisions of high income individuals, because self-employment makes it easier to under-report income relative to paid work.

As already stated, the above discussion does nothing to disentangle the income effect on self-employment, which has been found to be negative (Amit, Cockburn, and Muller, 1995; Evans and Leighton, 1989) from the education effect, which has been found to be positive (Borjas and Bronars, 1989; Rees and Shah, 1986). We find it puzzling that the relationship

between paid income and the probability of self-employment is negative rather than U-shaped. One explanation is that it is not income per se that matters, but income relative to human capital and ability. We take up this discussion in the next sub-section.

Ability

Observation of the real world of employment leads to an unequivocal conclusion: there are large differences in the income levels of individuals with the same level of human capital. This is so even for individuals in the same occupation and industry, and in the same geographical location. We attribute these income differentials to individual differences in attributes such as the ability to get things done, energy levels, the ability to get along with people, motivation, commitment etc. All these different characteristics, which are not captured by human capital variables, can be grouped under one general heading: ability. The focus of the research on ability bias has been on how to obtain an accurate parametric estimate of the returns to education. The difficulty lies in separating the actual returns to education from the returns to innate abilities that are independent of education. It is recognized that education may be an endogenous variable with ability determining the amount of education related investment an individual will make. There are those who suggest that education may serve a function beyond that of productivity enhancement, as is suggested by human capital theory (Becker, 1993; Mincer, 1962; Schultz, 1961): it may serve as a signaling mechanism of an individual's ability (Spence, 1974). Thus, the ability variable may permit us to explore a lower level explanation for the probability of a transition to self-employment.

Blackburn and Neumark (1995) outline three broad approaches that have been used to control for ability in the estimation of returns to schooling: first, explicit measures, such as IQ test scores, have been used as proxies for ability. This approach tends to reduce the impact of

schooling on income relative to OLS estimates that do not control for ability. Second, data from identical (monozygotic) twins have been used to control for the ability factor. This approach reduces even more the estimates of returns to schooling than the first approach. Finally, "natural experiments" such as interactions between quarter of birth and compulsory schooling laws (Angrist and Krueger, 1991) are used to create instrumental variables that are uncorrelated with ability. This approach either leads to no change in the returns to schooling estimate relative to OLS estimates without any controls for ability, or to an *increase* in the estimate.

The conclusion one can draw from a review of the literature on ability bias is that the concept of ability is of great theoretical importance, but no fully satisfactory method has yet been found of measuring it. We offer a different approach to the ones that have been used so far; one that benefits from the fact that income estimation is not our primary research interest in this paper. We hypothesize that an individual's income in paid work can be explained by four broad constructs: first, human capital, which is made up of education, employment tenure, and labor market experience; second, discrimination, which based on the well-documented income penalties suffered by non-whites and females; third, occupation, industry, and state of residence; and fourth, innate abilities. Our model can therefore be written as

$$I = á + â_1H + â_2D + â_3S + â_4A + å,$$

where I is income in paid employment, H is a vector of human capital variables, D is a vector of demographic (discrimination) variables, S is a vector of spatial variables, and å is an error term uncorrelated with I. We assume that the market is efficient in rewarding the worker's productivity. Therefore, if all the variables except Ability are controlled for by means of an OLS regression, then the residual (Income-Income') should contain Ability and the error term. We call this residual the ability gap and take this residual to be a rough measure of ability. This

residual can take both positive and negative values, reflecting a premium for ability and a penalty for the lack of it. We recognize that there are problems with this measure. First, in predicting income, we have actually excluded ability, and thus all the parameter estimates could be biased. Since most studies report a drop in the effect size of the education parameter when ability is controlled for, our approach most likely understates the size of the residual. Second, we are unable to separate the error term from ability. Third, we make the assumption that the proportion of the true ability component in the residual is the same across individuals. Fourth, our ability measure is orthogonal to education, whereas the theory suggests that they should be substantially correlated.

We hypothesize that individuals with very low ability are more likely to take up self-employment. These individuals may simply lack the discipline to work under someone else's authority, or in teams. They may also compare their low positions in their organizations and their low remunerations with those of higher ability individuals with similar human capital, and feel frustrated at the difference. They would thus be tempted to strike it out on their own. It is an empirical question whether their self-employment income would be higher than their paid employment income. On the other hand, individuals with very high ability are those with high energy levels, who get things done, who have strong interpersonal skills, and who are creative problem solvers. These individuals may feel that, in spite of their above average remuneration, they can do better on their own. This is because they have to share with their principals a substantial percentage of the value they contribute to their organizations. Finally, individuals of average ability are likely to be compensated in line with their human capital – they are therefore less likely to search for self-employment opportunities. Even if they come across one, they are

likely to assign lower weights to the expected future earnings from it relative to their current incomes in paid work (Kahneman and Tversky, 1979).

III. Data, variables, and method

We use data from the PSID SRC sample, a large longitudinal dataset begun in 1968 at the Institute for Survey Research. We focus on heads of households because they are the family members about whom the greatest amount of information is available. We take eight pairs of years: 1985-86, 1986-87, 1987-88, 1988-89, 1989-90, 1990-91, 1991-92, and 1992-93. The first year of each pair is year 1, and the second is year 2. We take 1985 as our starting point because some of the variables of our interest are not available for earlier years. When the data preparation for this research was concluded, 1993 was the last year for which full data was available. We first use OLS regression with income percentile as the dependent variable, and education, tenure, labor market experience, race, sex, occupation, industry, and state of residence. We use seven dummy variables for the eight years. We take the residual (Income – Income') as a rough measure of ability, to be used in the logistic regression models. We call this residual the ability gap. We then run four logistic regression models, with a binary dependent variable for year 2 employment status (self-employment=1, paid work=0). In year 1 of each pair, all subjects are exclusively in paid work; in year 2 we examine whether they continue in paid work or in self-employment. We use year 1 variables as predictors of the self-employment choice in year 2. Since we have multiple observations for each subject – a maximum of 8 – we use income quintiles rather than dollar income as one of the predictors. Additionally, we use General Estimation Equations both in the OLS and the logistic regressions to control for clustered data, in which within subject correlations are higher than across subject correlations (see Allison, 1999).

In model 1 we use income quintiles and education. In model 2 we use four dummy coded variables to capture the ability gap in place of income quintiles, while retaining education. In model 3 we use all three variables. In model 4 we only consider subjects who were not unemployed at all in year 1 (unemployed hours = 0). We run four logistic regression models in total.

There was a problem of quasi-complete separation with some of the dummy variables for occupation and state of residence, in that these variables either did not have any case with a 1 on the dependent variable, or had only one such case. The variables affected for the first three models were occupation 9, and states 2, 3, 7, 10, 11, 14, 15, 16, 35, 38, and 43. In model 4, in additional to the variables affected for the first three models, states 30, 44, 46, and 50 were affected. These variables were dropped from the respective models.

Dependent variable (self-employment in year 2): This variable is coded as 0 if the individual continued in paid employment in year 2, and 1 if the individual made the transition to self-employment. All individuals in year 1 were in full time paid-employment.

Income quintiles: We have four dummy coded variables for the five income quintiles. Quintile 1 (1-20th percentile) is taken as the reference variable, and we thus report parametric estimates for quintiles 2, 3, 4, and 5.

Ability gap: As explained earlier, ability gap is the residual between the actual income percentile and the predicted income percentile that we get from the OLS regression. The ability gap has a mean of 1.015, and a standard deviation of 16.95 income percentile points. This indicates that at the mean, the abilities of individuals are roughly in line with their human capital. We divide this residual into 5 categories, and obtain four dummy coded variables. The first category (Ability Gap 0) comprises ability gaps within one standard deviation on either side of the mean. This

category serves as our reference point and these are individuals who are reasonably well compensated vis-à-vis their human capital. Ability gaps between one and two standard deviations below the mean are included under the variable Ability Gap Negative 1. Those that are more than two standard deviations below the mean are included under Ability Gap Negative 2. Those that are between one and two standard deviations above the mean are referred to as Ability Gap Positive 1. Those that are more than two standard deviations above the mean are referred to as Ability Gap Positive 2. Parametric estimates are reported for Ability Gaps Negative 1 and 2, and Positive 1 and 2.

Control variables: In addition to the above, we control for several other variables. *Unemployed hours* is the total number of unemployed hours reported for year 1. For model 4, only subjects with no unemployment hours were considered. *Children* is the number of children. *Jobunion* is dummy coded with a value of 1 if the paid job in year 1 was under union contract and 0 otherwise. *Tenure* is employment tenure in months. *Labmktex* is total full time labor market experience since the age of 18. *Tenuresq* and *Labmktexsq* are quadratic expressions of tenure and labmktex respectively to control for positive but declining returns to these two attributes (Ben-Porath, 1967). *Marital status* is coded 1 if the subject was married or cohabiting with a partner, and 0 otherwise. *Education* is years of education, with the values 0-16 representing the actual years of schooling, and 17 representing a maximum value for all graduate education. Thus the mean of the education variable is somewhat understated in the dataset. *Age* is given in years, and only observations in the range of 18-65 are considered. *Religious preference* is coded 1 for Protestants and 0 otherwise, to control for the greater propensity of Protestants to take up self-employment, as reported by Carroll and Mosakowski (1987). *Race* is coded 1 for whites and 0 otherwise. *Sex* is coded 1 for females and 0 for males. *Father's occupation* refers to the usual

occupation of the subject's father when the subject was growing up, and is coded 1 for self-employed and 0 otherwise. We then include four sets of dummy variables: 11 *occupation* variables for a total of 12 occupations at the 3-digit level; 7 *year* variables for a total of 8 years; 11 *industry* variables for a total of 12 industries at the 3-digit level; and 51 *state* variables for the 50 states, District of Columbia, and other territories like Virgin Islands, Puerto Rico, etc.

IV. Results and Discussion

Descriptive statistics are given in Table 1. The results of the OLS regression predicting income percentile through human capital, sex, race, occupation, industry, and state of residence are given in Table 2. As expected, the residuals (the ability gap) continue to be highly correlated with income percentiles, and their variance highly uniform, as can be seen from Figure 1. It is interesting to note that individuals above the 85th percentile hardly suffer any penalties for lack of ability.

Table 3 gives the parametric estimates for the four logistics regression models. Model 1 includes the income quintiles and the other explanatory variables. All four income quintile parameters are highly significant, showing that the likelihood of self-employment is a decreasing function of income. The other explanatory variables that are statistically significant predictors are Jobunion (negative), Tenure (negative), Labor market experience (positive), Education (positive), and Race (whites more likely to become self-employed).

Model 2 includes the Ability Gap variables in place of the income quintiles. Again, three of the four ability gap measures are predictors in a statistically significant way of the probability of taking up self-employment. It appears that those who receive incomes substantially below their expected incomes are more likely than the reference group to take up self-employment. On

the other hand, only those who are in income percentiles more than two standard deviations above the mean are more likely to take up self-employment. It is interesting to compare the education variable in models 1 and 2. While it is significant at the $p<0.01$ level in Model 1, it is only marginally significant at the $p<0.10$ level in model 2. This is particularly interesting given the orthogonality of education to our ability measure.

Model 3 includes both sets of variables: income quintiles and ability gaps. Interestingly, once we control for the ability gap, the explanatory power of the income quintile variables disappears, supporting the suggestion that income is an endogenous variable in the model with ability being exogenous. Again, Education is only marginally significant at the $p<0.10$ level. A plot of the predicted probabilities as a function of the ability gap demonstrates the U shaped effect - see Figure 2.

Model 4 only considers subjects who had no spell of unemployment in year 1. The three ability variables are again very strong predictors of self-employment, and the income percentile variables continue to be statistically insignificant. Education is no longer significant in predicting self-employment.

People of low abilities are more likely to take up self-employment than those of moderate ability. We hypothesize that these individuals are frustrated at their low incomes relative to their human capital. They may also be envious of individuals of similar human capital earning substantially more. People of very high abilities (the superstars) are also more likely to take up self-employment, in spite of earning substantially more than the mean income for individuals with similar human capital. One possible explanation for this is that these very high ability individuals, in spite of being highly rewarded, may feel that they unnecessarily hand over a substantial percentage of the value they create to the principals.

IV. CONCLUSIONS

Using panel data, we have shown that ability is a better predictor than labor income or education of the likelihood that a person will take up self-employment. We have used the residual derived from the difference between actual income and predicted income as a rough measure of ability. Individuals earning substantially below their expected incomes are considered low-ability individuals, whereas individuals earning substantially above their expected incomes are considered high ability individuals. Those who earn between one standard deviation below and one standard deviation above the predicted mean are considered roughly well compensated for their abilities – the difference could be attributable to market imperfections.

In contrast to the suggestion made by Rees and Shah (1986), we assume that success in self-employment does not require different skills from those required in paid work. Individuals who show high abilities in paid work can successfully transfer these abilities to self-employment. Although individuals with very high abilities receive considerable rewards in paid work, their compensation is by definition lower than their marginal product. Those individuals for whom the difference between their marginal product and their income is the greatest will be more likely to take up self-employment. Individuals with low abilities are penalized by the market, such that their earnings are considerably below what they should be given their human capital. We use prospect theory to explain why low and very high ability individuals are more likely to take up self-employment. Our results are consistent with non-monetary explanations also, such as the desire to take on the challenge of creating an economic institution, the desire for independence, in that high ability individuals are more likely to be motivated in this way.

References

Amit, R., E. Muller, and I. Cockburn, "Opportunity Costs and Entrepreneurial Activity," *Journal of Business Venturing* 10:1 (1995), 95-106.

Angrist, J. D., and A. Krueger, "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106:4 (1991), 979-1014.

Bates, T., *Race, Self-Employment, and Upward Mobility*, (Washington, D. C.: The Woodrow Wilson Center Press, 1997).

Becker, G. S., "A Theory of Marriages, Part (1)," *Journal of Political Economy* 81:4 (July-August 1973), 813-845.

Becker, G. S., *Human Capital*, 3rd edition, (Chicago: University of Chicago Press, 1993).

Behrman, J. R., Z. Hrubec, P. Taubman, and T Wales, *Socioeconomic Success: A Study of Genetic Endowments, Family Environment, and Schooling*, (Amsterdam: North Holland, 1980).

Behrman, J. R., and M. R. Rosenzweig., ""Ability" Biases in Schooling Returns and Twins: A Test and New Estimates," *Economics of Education Review* 18:2 (1999), 159-168.

Ben-Porath, Y., "The Production of Human Capital and the Life Cycle of Earnings," *The Journal of Political Economy* 75:4 (1967), 352-365.

Blackburn, M., and D. Newmark, "Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials," *Quarterly Journal of Economics* 107:4 (November 1992), 1421-1436.

Blackburn, M., and D. Newmark, "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look," this REVIEW 77:2 (May 1995), 217-230.

Blau, David. M., "A Time-Series Analysis of Self-Employment in the United States," *The Journal of Political Economy* 95 (1987), 445-467.

Borjas, G. J., and S. G. Bronars, "Consumer Discrimination and Self-Employment," *The Journal of Political Economy* 97:3 (1989), 581-605.

Card, D., "The Causal Effect of Education on Earnings," *Handbook of Labor Economics*, O. Ashenfelter and David Card, eds., (New York: North Holland, 1999).

Carroll, G., and E. Mosakowski, "The Career Dynamics of Self-Employment," *Administrative Science Quarterly* 32 (1987), 570-589.

Devine, T. J., "Changes in Wage-and-Salary Returns to Skill and the Recent Rise in Female Self-Employment," *American Economic Review* 82:2 (May 1994), 108-113.

Evans, D. S., and B. Jovanovic, "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy* 97 (1989), 808-827.

Evans, D. S., and L. S. Leighton, "Some Empirical Aspects of Entrepreneurship," *The American Economic Review* 79:3 (1989), 519-535.

Griliches, Z., and W. M. Mason, "Education, Income, and Ability," *Journal of Political Economy* 80:3 (May-June 1972), S74-S103.

Holtz-Eakin, D., D. Joulfaian, and H. S. Rosen., "The Carnegie Conjecture: Some Empirical Evidence," *The Quarterly Journal of Economics* Vol. CVIII (May 1993), 413-436.

Kahneman, D., and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47:2 (1979), 263-291.

Kangasharju, A., and S. Pekkala, "Self-Selected Self-Employed: The Role of Education in Success," PTT Working Paper #38, Helsinki (2000).

Knetsch, J.L. and J.A. Sinden, “Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value,” *Quarterly Journal of Economics* 99:3 (1984), 507-521.

Knight, F., *Risk, Uncertainty, and Profit* (Chicago: University of Chicago Press, 1921).

Knudsen, K. R., and D. G. McTavish, “Modelling interest in entrepreneurship: implication for business and development,” In R. H. Brockhaus, N. C. Churchill, J. A. Katz, B. A. Kirchhoff, K. H. Vesper, W. E. Wetzel, eds., *Frontiers of Entrepreneurship Research* (1989), 67-86.

Leung, D., and C. Robinson, “Explaining the Recent Rise in Self- Employment: Lifecycle, Cohort, and Aggregate Economy Effects,” University of Western Ontario mimeo, (2001).

Mincer, J., “On the Job Training: Costs, Returns, and Some Implications,” Journal of Political Economy 70:5 (1962), 50-79.

Rees, H., and A. Shah, “An Empirical Analysis of Self-Employment in the U. K,” *Journal of Applied Econometrics* 1 (1986), 95-108.

Schultz, T., “Investment in Human Capital,” *American Economic Review* 51:1, (March 1961), 1-17.

Shane, S., “Prior knowledge and the discovery of entrepreneurial opportunities,” *Organization Science* 11:4 (2000), 448-469.

Simon, H. A., “A Behavioral Model of rational Choice,” *Quarterly Journal of Economics* 59, (1955), 99-118.

Spence, M., *Market Signalling*, (Cambridge: Harvard University Press, 1974).

Table 1
Descriptive Statistics

Number of valid observations	18,860
Number of clusters	3,908

Categorical Variables:

Frequency of self-employed in year 2	524 (Missing Values 370)
Frequency Ability Gap Negative 1	2,307 (Missing Values 842)
Frequency Ability Gap Negative 2	506 (Missing Values 842)
Frequency Ability Gap 0	12,354 (Missing Values 842)
Frequency Ability Gap Positive 1	2,559 (Missing Values 842)
Frequency Ability Gap Positive 2	292 (Missing Values 842)
Job under union contract	3,955 (Missing Values 517)
Marital Status (1=married or cohabiting)	13,061 (Missing Values 0)
Religious preference (1=Protestant)	11,789 (missing Values 206)
Race (1=white)	17,013 (Missing Values 43)
Sex (1=Female)	3,405 (Missing Values 0)
Father's occupation (1=self-employed)	673 (Missing Values 0)

Continuous Variables:	Mean	Std dev.	Min	Max
Unemployed Hours	45.21	187.10	0	2,080
Number of Children	0.98	1.14	0	7
Tenure in months	99.44	100.65	0	564
Labor mkt experience (yrs)	13.85	10.33	0	47
Education (yrs)	13.39	2.35	3	17
Age	37.56	10.82	18	65

Table 2

OLS Regression - GEE Parameter Estimates
Dependent Variable = Income Percentiles

Parameter	Estimate	S.E.
Intercept	-4.6359*	2.6001
Tenure	0.0936***	0.0055
Tenure squared	-0.0001***	0.0000
Labor Market Experience	0.9667***	0.0792
Labor Market Experience squared	-0.0235***	0.0022
Education	3.3026***	0.1301
Race	6.8515***	0.8556
Sex	-9.4561***	0.6639

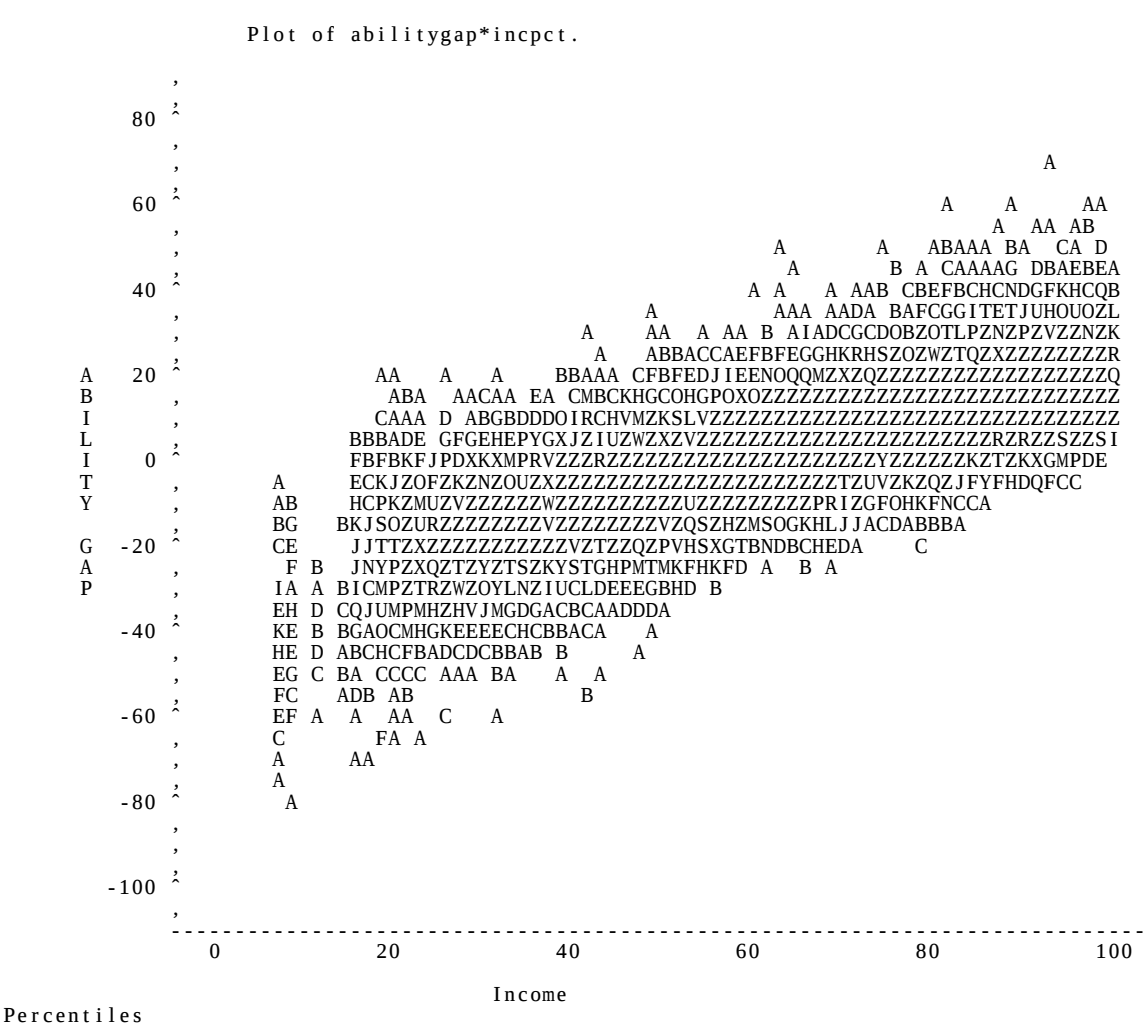
In addition to the variables listed above, dummy variables were used to control for occupation, industry, year, and state of residence. The estimates of these parameters have not been reported.

* p<.10
** p<.05
***p<.01

Note: The GEE tool used (PROC GENMOD in SAS with REPEATED statement for subject id) does not report a model R^2. The simple correlation between the observed and the predicted values of the dependent variable is 0.69.

Figure 1

PLOT OF ABILITY GAP AS A FUNCTION OF INCOME PERCENTILE



Legend: A = 1 obs, B = 2 obs, etc.

NOTE: 842 obs had missing values. 6609 obs hidden.

Table 3

Logistic Regression - GEE Parameter Estimates
Dependent variable = Self-employed in year 2 (=1); Paid work in year 2 (=0)

Parameter	Model 1	Model 2	Model 3	Model 4
Intercept	-5.3848***	-5.5367***	-5.6103***	-5.4419***
Income Quintile 2	-0.4946**		0.0504	0.0411
Income Quintile 3	-1.1142***		-0.1648	-0.1718
Income Quintile 4	-1.2935***		-0.2338	-0.1731
Income Quintile 5	-1.0362***		-0.1223	-0.0422
Ability Gap Negative 1		0.6469***	0.5286***	0.6332***
Ability Gap Negative 2		1.7269***	1.5627***	1.8621***
Ability Gap Positive 1		0.1777	0.1843	0.2398
Ability Gap Positive 2		1.4878***	1.5047***	1.5565***
Hours unemployed	0.0001	0.0001	0.0001	
Children	0.0059	0.0304	0.0277	0.0369
Job under Union Contract (Yes=1)	-0.6427***	-0.7407***	-0.7211***	-0.8512***
Tenure	-0.0049***	-0.0050***	-0.0047***	-0.0051***
Tenure squared	0.0000	0.0000	0.0000	0.0000
Labor Market Experience	0.0518**	0.0442*	0.0471*	0.0523**
Labor Market Experience squared	-0.0006	-0.0004	-0.0005	-0.0004
Marital Status (Married=1)	-0.0957	-0.0892	-0.0874	0.5243
Education in years	0.0956***	0.0565*	0.0641*	0.0636
Age	-0.0084	-0.0239	-0.0217	-0.0365
Age squared	-0.0001	0.0000	0.0000	0.0001
Religious Preference (Protestant=1)	0.0672	0.0877	0.0880	0.0769
Race (White=1)	0.6082***	0.4976**	0.5221**	0.4727**
Sex (Female=1)	-0.2253	-0.1050	-0.1254	-0.0531
Father's occupation (Self-employed=1)	0.3345	0.2919	0.2924	0.2674

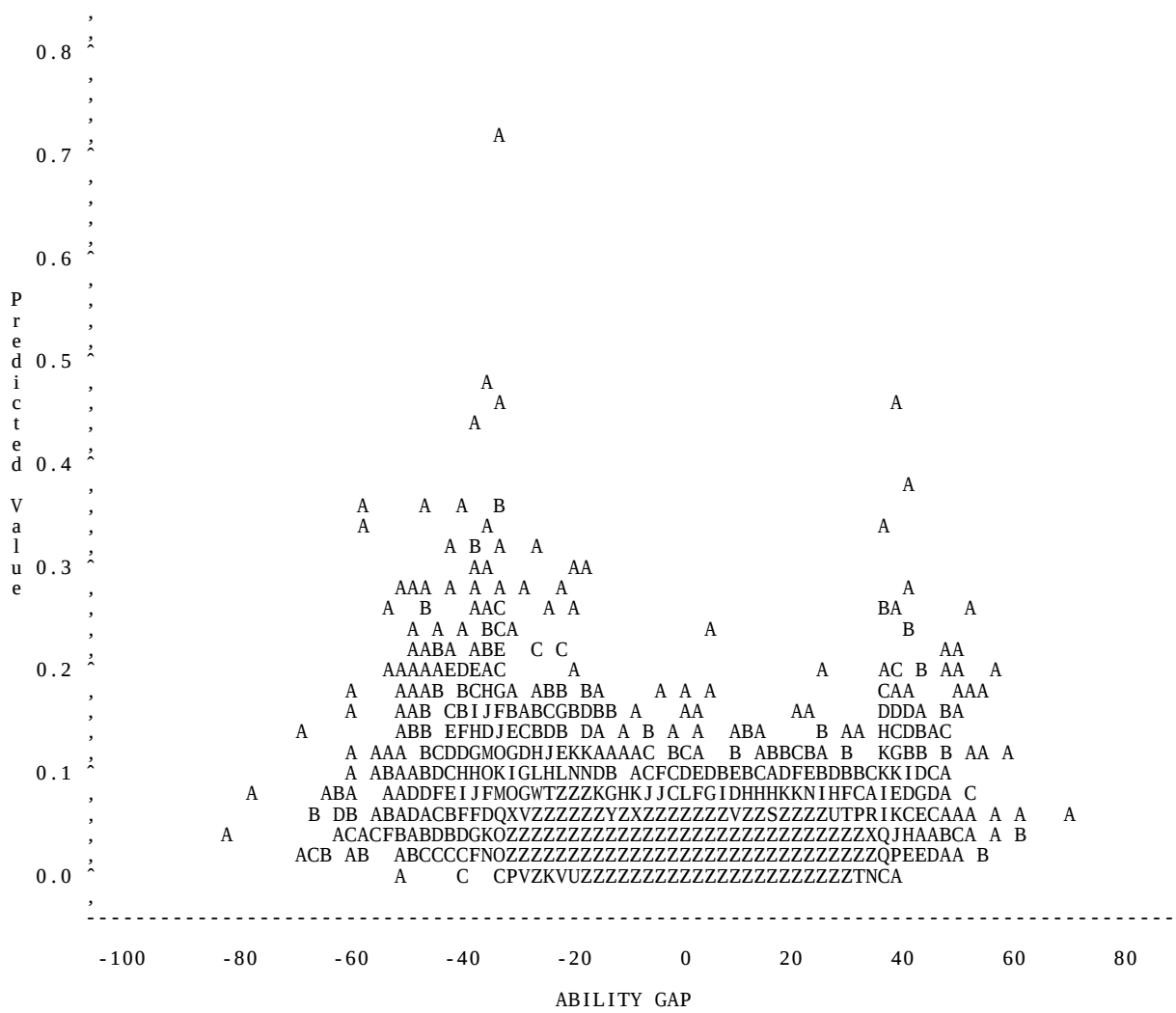
The parameter estimates for occupation, year, industry, and state of residence have not been reported. Model 4 holds hours unemployed constant at 0. Valid observations = 15454. Self-employed in year 2 = 402.

* p<.10, ** p<.05, ***p<.01

Model Goodness of Fit Criteria				
Deviance	4060.432	3987.992	3985.07	3323.99
Pearson Chi-Square	18356.6	18190.86	18249.79	16173.33
Log Likelihood	-2030.22	-1994	-1992.53	-1662.00

Figure 2 (based on Model 3)

Plot of Predicted Probability of Self-Employment as a function of the Ability Gap



Plot of probslfemp*incgap. Legend: A = 1 obs, B = 2 obs, etc.

NOTE: 1475 obs had missing values. 12883 obs hidden.