DISCUSSION PAPER SERIES

I Z A

Self-Esteem and Earnings

Francesco Drago

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Self-Esteem and Earnings

Francesco Drago
University of Napoli Parthenope
and IZA

ABSTRACT

Self-Esteem and Earnings[*]

Recent research in economics suggests a positive association between self-esteem and earnings. A major problem in this literature is that from simple cross-sectional wage regressions it is not possible to conclude that self-esteem has a causal impact on earnings. While classical measurement error leads to an attenuation bias, reverse causality and omitted variable are likely to drive the OLS coefficient on self-esteem upward. Using the National Longitudinal Survey of Youth (NLSY) that administered the Rosenberg Self-Esteem Scale during the 1980 and 1987 interviews, I provide further evidence for the existence of a self-esteem premium by exploiting variation in these measures in the two years. I show that the estimated impact of self-esteem in 1987 on earnings is about two times greater than previous OLS estimates would imply. The main explanation for this result is the large extent of measurement error in the reported self-esteem measure.

Corresponding author:

Francesco Drago
Departimento di Studi Economici
Università di Napoli, Parthenope
Via Medina 40
80133 Napoli
Italy
E-mail: fdrago@uniparthenope.it

1 Introduction

Recent research in economics provides evidence that non-cognitive skills are important determinants of earnings (Bowles, Gintis and Osborne, 2001). A relevant non-cognitive skill is self-esteem, which is usually conceived as the perception that individuals have about their own ability. From the theoretical point of view, the basic idea according to which self-esteem may increase earnings is very simple. Ability and effort are complements, so that, under the premise that individuals are uncertain about their own ability, higher self-esteem causes higher effort and earnings (Benabou and Tirole, 2002). The effect of self-esteem on earnings can be persistent. As recent theoretical works show, rational agents may find it convenient to hold incorrect beliefs about their own ability even in the long-run (Benabou and Tirole, 2002; Comte and Postlewaite, 2004; Santos Pinto and Sobel, 2005; see also Costa and McCrae, 1998 in the psychological literature.). In this perspective, self-esteem is a productive skill that has a direct impact on earnings.

Some empirical studies find a positive effect of self-esteem in OLS wage regressions, even after controlling for cognitive tests scores and demographic characteristics (Goldsmith, Veum and Darity, 1997; Murnane, Willett, Braatz and Duhaldeborde, 2001; Waddell, 2006).[1] However, it is difficult to conclude that self-esteem has a causal impact on earnings (Heckman, Stixrud and Urzua, 2006). A positive correlation between self-esteem and earnings is not surprising given that individual labour market outcomes (e.g. earnings, promotions) likely impact the self-image of a worker. Moreover, the association between self-esteem and earnings could reflect the fact that the former is systematically correlated to unobservable characteristics which determine earnings. Both these effects cast doubts on the existence of a causal impact of self-esteem on earnings since it is likely that they bias the OLS coefficient upward. In this paper, by using the National Longitudinal Survey of Youths (NLSY) I investigate if these concerns find empirical support in the data. I show that, despite the above

---

[1] Goldsmith et al. (1997) find a positive impact of psychological capital measured by "one's locus of control" on wages. Murname et al. (2001) document a positive impact of self-esteem on subsequent wages in the NLSY. They find that cognitive skills and self-esteem have different roles in explaining wage gaps between white black and hispanic males. Waddell (2006) uses the National Longitudinal Study of the High School Class of 1982 and finds that individuals with poor attitude and self-esteem attain fewer years of post-secondary education and lower earnings. Graham, Eggers and Sukhtankar (2004) find that happiness – determined by self-esteem and optimism - positively affects income especially for those at lower levels of income.

mentioned concerns, the estimated impact of self-esteem on earnings is about two times greater than previous OLS estimates imply. I argue that measurement error in self-esteem is the most plausible candidate that leads to a downward bias of the OLS estimates.

The question of whether the association between self-esteem and earnings is simply a correlation or a causal relationship is of practical and theoretical relevance. Assessing a significant impact of self-esteem on earnings might help to explain a part of the large residual wage inequality. In addition, the investigation of whether and how strongly higher self-esteem translates into higher earnings can inform parents and teachers about the importance of focusing on the development of the non-cognitive skills of their children. For example, if self-esteem is at least as important as verbal skill in explaining why individuals are paid differently, then educational institutions should be evaluated also on the grounds of the development of such non-cognitive skills.

The NLSY administered the Rosenberg Self-Esteem Scale during the 1980 and 1987 interviews. This scale, designed for adolescents and adults, is composed of ten questions that capture a degree of approval or disapproval toward oneself. This scale was extensively used in psychology and recently also in economics (Bowles et al., 2001; Heckman et al., 2006). The first measure, i.e. the one in 1980, reports the self-image of young individuals, the 50 percent of whom are adolescents (they had not yet entered the labour market). The second one (in 1987) measures the self-image of the same individuals who, for the most part, already entered the labour market. What is important for the analysis is that for most of the sample there is considerable variation in the two measures of self-esteem in 1980 and 1987. The positive correlation between these measures and education and achievement test scores confirms the behavioural validity of these variables reflecting individuals' self-image. However, the low partial correlation between the two measures of self-esteem suggests that these are only noisy measures of the perception that individuals have about their own ability and that maintain in the long-run.

In section 3 I present OLS estimates of the effect of self-esteem in 1987 on the log wage observations that come from 1988. The measure of self-esteem in 1987 has a positive and

statistically significant impact on earnings even when I control for measures of ability, parental background and education. To shed light on whether OLS estimates understate the true effect of self-esteem on earnings, the identification strategy exploits variation of the self-esteem measures in 1980 and 1987. In particular, the measure in 1980 is used as an instrument for self-esteem in 1987. To lend credibility to the IV estimates, the empirical analysis is performed on a sub-sample of workers born after 1961 for whom the schooling choices were constrained by compulsory schooling laws until at least 1978 (Neal and Johnson, 1996).[2] The idea to focus on individuals born after 1961 takes care to avoid that variation between self-esteem in 1980 (the IV) and 1987 is correlated with variation between wages in 1980 and 1987, which would render the IV not valid.[3] This approach takes care of another concern that typically plagues empirical estimates of the effects of non-cognitive skills on earnings. As Bowles and Gintis (1976) and more recently Heckman et al. (2006) show, schooling increases non-cognitive skills including self-esteem. This observation and the fact that schooling is a choice variable correlated with ability might cast doubts on the validity of the instrument, i.e. reported self-esteem in 1980. However, by using this sub-sample of workers born after 1961, the impact of education on self-esteem in 1980 should be equal for everyone once I partial out age effects from the self-esteem measure.

One of the main results is that IV estimates of the impact of self-esteem in 1987 on wages are about two times higher than the corresponding OLS. This result is robust to controlling for several variables potentially correlated to self-esteem and at the same time predictors of earnings. In a complementary analysis I use the sub-sample of siblings in the NLSY and I instrument individual self-esteem with siblings' self-esteem. This analysis supports the robustness of the results of the paper. The main explanation for the fact that the OLS estimator suffers from such a large downward bias is the presence of classical measurement error in the measure of self-esteem. Indeed, the other sources of unobserved heterogeneity, i.e. omitted variable and reverse causality, likely lead the OLS coefficient to be upward biased.

---

[2] As Neal and Johnson (1996) report, "most of these workers had neither entered the labour market full time nor started post-secondary schooling" at the date they were interviewed in 1980.
[3] To be clear, the IV should not have an effect on wages in 1980 and, given correlation in wages, a direct impact on wages in 1988, simply because most of the workers considered did not enter labour market by 1980.

4

Overall, the results indicate that measurement error that drives the OLS coefficient downward more than offsets the unobserved heterogeneity that drives the OLS estimator upward.

This paper contributes to the literature on the behavioural determinants of earnings (Bowles, et al. 2001; Cawley, Heckman and Vytlacil, 2001). While some recent econometric frameworks provided some evidence for the existence of a causal relationship between self-esteem and earnings (in particular see Heckman et al. 2006 and Borghans et al. 2008), the issue of how large is the bias in the OLS estimates has not been discussed so far. Understanding this issue is important to assess whether previous estimates in the literature should be interpreted as causal effects of self-esteem on earnings. Perhaps contrarily to what one could have expected, the evidence provided here suggests that the OLS estimator is downward biased. Hence, another contribution of this paper is to provide further evidence on the causal effect of self-esteem on earnings.

The paper is structured as follows. In the next section I describe the data. In sections 3 I discuss the potential biases of OLS estimates. Section 4 presents the IV estimates and the last section concludes

2 Data

Data come from the National Longitudinal Survey of Youth (NLSY). The NLSY is a representative sample of 12,686 young men and women residents in the US who were 14-22 years old when they were first interviewed in 1979. These individuals were interviewed annually through 1994 and are currently interviewed on a biennial basis. Individuals were interviewed on a large number of questions about their labor market history, family and social background, education and personality traits.

The NLSY administered the Rosenberg Self-Esteem Scale during the 1980 and 1987 interviews. Respondents were first asked to report their self-image in 1980 when they were aged 15-23 and most recently in 1987. This 10-item scale, designed for adolescents and adults, measures the self-evaluation that an individual makes and describes a degree of approval toward oneself. It contains 10 statements of self-approval and disapproval with

5

which respondents are asked to strongly agree, agree, disagree, or strongly disagree. The statements are the following: A) I am a person of worth, B) I have a number of good qualities, C) I am inclined to feel that I am a failure, D) I am as capable as others, E) I feel I do not have much to be proud of, F) I have a positive attitude, G) I am satisfied with myself, H) I wish I had more self-respect, I) I feel useless at times, J) I sometimes think I am "no good at all". Items A, B, D, F and G need to be reversed prior to scoring in order for a higher score to designate higher self-esteem. Hence, for each statement I assign 1 if the individual strongly agrees with the statement, 2 if she agrees, 3 and 4 if she disagrees and strongly disagrees, respectively. To create an aggregate measure of self-esteem in 1980 and in 1987, I sum all the values assigned to these statements. The results are robust when I use alternative measures that are derived using factor analysis (see sub-section 4.3).

To avoid confounding effects of race, gender and self-esteem, I focus my attention on white men. There are 3790 white male in the NLSY. I drop all individuals who were not interviewed in one of the two years (1980 and 1987). I also also drop observations for individuals that reported a wage per hour lower than 1 dollar or greater than 100 dollar in 1988 (the focus is on wages in 1988). I do not consider those enrolled in school in 1988. The number of individuals is 2250. In Table 1 I report summary statistics. In 1980 the average measure of self-esteem is equal to 29.55 (note that the lowest possible value is 10 and the highest possible value is 40) while for 1987 it is equal to 30.63. In Table 2 I report the means for the years of schooling completed by 1988, wages in 1988, AFQT scores, mother and father's education and height for individuals with self-esteem below and above the median.[4] (The AFQT score is an achievement test administered to the respondents of the NLSY in 1980 and it is conventionally interpreted as a strong correlate of individual productivity). As expected, in Table 2, individuals with higher level of self-esteem have higher wages, AFQT scores, years of schooling and height. The summary statistics indicates the behavioural validity of both measures of self-esteem. Education and achievement test score should predict the self-image of an individual. Indeed, by running regression of both measures of self-esteem

---

[4] Beauty and height are shown to have significant effects on earnings. To the extent that some of these effects can develop through self-esteem (Persico et al. 2004, Mobius and Rosenblat, 2006), it is reasonable to expect height and beauty to be correlated to self-esteem. The NLSY provides individual height in 1981 and 1985, the latter measure is used in the analysis.

on AFQT, education and height I find that all these variables have a strong and statistically significant effect in predicting self-esteem (results not reported). It is well known that AFQT increases with age and education (Neal and Johnson, 1996). In the rest of the analysis AFQT scores and measures of self-esteem are standardized to account for the differences in schooling levels across ages. Following the procedure of Altonji and Pierret (2001), I adjust the raw AFQT score by subtracting the mean score for a person of that age and dividing by the standard deviation for that age. The same procedure is adopted to standardize self-esteem in 1980 and 1987, so that AFQT and self-esteem have mean zero and standard deviation one in the sample.

2.1 Measures of self-esteem and measurement error

The raw correlation between the two self-esteem measures is 0.4215. A high degree of measurement error in the two measures is a potential explanation for such a relatively low correlation. Note also that given that the survey methods used to construct such measures are identical, a part of the raw correlation may reflect correlation in the errors. That the extent of measurement error for self-esteem can be large is intuitive for the very nature of the questions used to construct the self-esteem measures. Unlike respondents' answers to questions about objective measures (e.g., age, education), answers to questions about how one feels better than others can be affected by transitory and situational factors. However, we are interested in the effect of the "true" self-esteem, that is the perception that individuals have about their own ability and that maintain in the long-run. Overall, the low correlation between the two measures of self-esteem indicates that these should be interpreted as noisy measures of the "true" individuals' self-esteem.[5]

In OLS estimates of the effect of self-esteem on earnings, under a classical measurement framework, the measurement error yields an attenuation bias of the coefficient on self-esteem. This bias can be more severe when we include additional controls in the regression. One way to see this is to observe how the correlation between the two measures changes when we partial out the set of included controls in a wage regression. For example, the correlation

---

[5] Bertrand and Mullainathan (2001) discuss the critical issues in using subjective survey data.

between the two measures of self-esteem when I partial out AFQT score, age and education is 0.3479, a drop of about 20 percent relative to the raw correlation. This indicates that in estimating the effect of self-esteem, the attenuation bias induced by measurement error is very large, especially when we include additional controls. Overall, the extent of measurement error we expect is large relative to that affecting education or objective measures. The low correlation between the two measures of self-esteem provides evidence in this direction.

3 OLS estimates

In this section I first present OLS estimates of the regressions of log wages in 1988, i.e. soon after the second measure of self-esteem was taken, and then I discuss the potential biases. The results obtained are very robust to the choice of the year (1988). The approach to focus on cross-sectional variation in wages takes care to avoid controlling for variables such as experience that are endogenous. I restrict the attention to white men that in 1988 were not enrolled in school and for whom it was possible to derive all the relevant variables (e.g. self-esteem in 1980 and 1987, AFQT score). To facilitate subsequent comparisons, for AFQT and self-esteem I use the corresponding standardized measures as explained in section 2. The wage equation to estimate is the following:

$$w_{i,1988}=X'_i\beta+\alpha \ self\text{-}esteem_{i,1987} + u_i \qquad (1)$$

where X_i include standardized AFQT score, age and mother and father schooling. In Table 3 I report the results from this regression. The coefficient on self-esteem is precisely estimated. A two standard deviation increase in the measure of self-esteem leads to a 13 percent increase in the log wage.

A major problem with the interpretation of these results is that in equation (1) self-esteem in 1987 is very likely correlated to the error term u_i. For example, if expectations in 1987 about wages one year later influence self-esteem in 1987, it is the wage that causes self-esteem and not vice versa. If these expectations are fulfilled, the OLS coefficient α in (1) is upward biased. A second source of bias derives from omitted variables. If self-esteem is systematically associated to omitted ability, then the OLS coefficient in (1) is again biased

8

upward. Finally, under the assumption of classical measurement error, OLS estimates of α in (1) are subject to a standard attenuation bias. As discussed in the previous section, it is likely a large attenuation bias especially when we include in the regressions other controls. Overall, the only source of unobserved heterogeneity that leads the OLS coefficient to be biased downward is classical measurement error.

By exploiting variation in the two measures of self-esteem (1980 and 1987), the identification strategy uses an IV approach and addresses these three sources of unobserved heterogeneity. Self-esteem in 1980 is used as an instrumental variable for self-esteem in 1987. Note that when we want to obtain estimates purged of the attenuation bias and we have two noisy measures of the variable of interest, instrumenting one with the other is a standard approach to deal with this kind of measurement error problem (e.g, see Donohue and Levitt, 2006).[6] The identifying assumption is that self-esteem in 1980 can be excluded from equation 1 once we control for AFQT, parental background, educational attainment and other demographics characteristics. The second stage regression is equation (1). The first stage is:

$$selfesteem_{i,1987} = X'_i \pi + \gamma \, self\text{-}esteem_{i,1980} + \varepsilon_i \qquad (2)$$

Formally, the assumption in this approach is cov(selfesteem$_{i,1980}$, u$_i$|X$_i$) = 0. In the next section I argue that for the sub-sample I will use, it is reasonable to believe that this condition is satisfied.

3.1 Validity of the instrument and sub-sample used

In using self-esteem in 1980 as an IV for self-esteem in 1987 an immediate concern is that the first measure may be affected by past labor market experience. Labor market success in terms of satisfaction and earnings before 1980 could have driven a higher measure of self-esteem in 1980 and, given potential serial correlation in wages, this would lead to a violation of the exclusion restriction. The second problem is that self-esteem can be influenced by schooling

---

[6] The fact that errors in the two measures are not uncorrelated implies that IV estimates will understate the effect of self-esteem (see below).

(Bowles and Gintis, 1976 and Heckman et al. 2006). Given that educational attainment is endogenous and correlated to ability, self-esteem in 1980 might have a direct effect on wages in 1988. To address these two problems I focus on individuals born after 1961. The idea to focus on this sub-sample is that the schooling choices of workers born after 1961 were constrained by compulsory schooling laws until at least 1978 (see Neal and Johnson, 1996 who consider this sub-sample in the NLSY.). The majority of workers born after 1961 did not enter the labor market by 1980. Hence, past labor market outcomes should had not influenced the measures of self-esteem in 1980. Moreover, once I use the standardized measure of self-esteem, the effect of schooling on self-esteem should be equal for all individuals born after 1961. As Neal and Johnson (1996) point out: "No respondent in this sample had completed a year of schooling beyond high school by May 1980 and less than 1 percent had even enrolled in college by this date." In 1980, all individuals born after 1961 were interviewed by May 1980. In this way, in estimating the effect of self-esteem on wages, biases due to omitted variables and reverse causality are minimized.

4 Empirical Results

4.1 Reduced form estimates

In Table 4 I report the results from the reduced form regression (i.e. OLS estimates of the the IV - self-esteem in 1980 - on log wages in 1988 for white males born after 1961). Column (1) reports the results including only the IV and age as regressors. A positive self-image in 1980 has a strong impact on wages. A two standard deviation increase in this measure leads to a 18 percent increase in log wage. In column (2) I report the OLS estimate for the log wage in 1988 including among the regressors father and mother education. For further analysis of the self-esteem premium that emerges from the reduced form regression, in column (3) I report the OLS estimates including the AFQT scores (standardized). In this case ability is allowed to operate both as a direct effect and as indirect effect through subsequent choices such as schooling and occupation. As expected the coefficient on AFQT is precisely estimated. The estimated self-esteem premium is about 50 percent of what it was in the absence of conditioning on AFQT, but still positive and statistically significant. Next, I control for the years of schooling and I find that the estimated self-esteem premium is about 65 percent of what it was in the absence of conditioning on schooling. By considering individuals born after

10

1961, I minimize the problem that education might have an effect on self-esteem in 1980. As mentioned above, the schooling choices for the majority of these workers were constrained by compulsory schooling laws at the date of the interview in 1980. There is, in fact, indication that a large part of the effect of schooling on the coefficient on self-esteem is mediated by differences in AFQT scores.[7] The last column of Table 4 reports OLS estimates including all the variables considered before. In this case the coefficient on self-esteem is equal to .0370. The empirical patterns in Table 4 show that family background explains a small part of the self-esteem resulting from the reduced form regression. Comparing columns 1-3 and 3-5 it appears that the effects that family background have on the coefficient on self-esteem are well accounted by the effect of AFQT score. An alternative explanation for why the self-esteem coefficient declines to the inclusion of AFQT score and other covariates is the measurement error in the self-esteem measure (e.g., if AFQT is affected by measurement error and errors in self-esteem and AFQT are correlated).

4.2 Alternative explanations

The NLSY allows to consider several other explanations for the results obtained in the reduced form regressions. In principle it is possible that the coefficient on self-esteem in 1980 is positive and significant even after controlling for many human capital variables because it captures some omitted variable. For example, the self-esteem premium could be explained by individual differences in outward characteristics, e.g. height and weight. In the followings I check the robustness of the results by analyzing some alternative explanations. Table 5 presents the results. In all specifications, age, standardized AFQT, education and educational parental background are included among the regressors so that the benchmark coefficient on the IV is equal to .0370 (Table 4, column (5)). As it is clear from column (1) and (2), differences in height and in weight do not to explain the association between self-esteem in 1980 and wages in 1988 (for height, I consider the height in 1985 that is the closest measure to 1988 in the NLSY). In column (3) I include dummy on social activities measured by the participation to non-vocational, non-academic school clubs in which the responded

---

[7] Although conditioning on schooling reduces the self-esteem premium, this effect is smaller than the effect of conditioning on AFQT. Moreover, including education in the wage regression does not substantially reduce the coefficient on self-esteem when AFQT and age are included.

participated when he was young.[8] These activities have been shown to have a positive impact on wages (e.g. Persico et al., 2004) and can have also an effect on adolescent self-esteem. Hence, participation in clubs could be in principle a channel through which self-confident workers earn more. However, from Table 5, the coefficient on self-esteem remains essentially unchanged to the inclusion of these variables. The results in Table 5 provide indirect evidence for the validity of the instrument. Variables such as height, weight and participation in clubs have been previously shown to have an impact on wages. The fact that the inclusion of these variables do not affect the coefficient on the IV suggests that the IV is orthogonal to these characteristics (once AFQT, parental background and education are included) and to unobservables that have an effect on earnings in 1988.

4.3 Instrumental Variable estimates

In the first stage regression, as reported in the first panel of Table 6, the measure of self-esteem in '80 has a strong and statistically significant effect on the measure in '87. An advantage of the instrument used is that we should not worry about the weakness of instrument. Given that the survey methods used to construct such measures are identical, it is likely that the measurement error in the two measures is positively correlated. Note that the correlation between the measurement error of the two measures of self-esteem leads to overestimate the coefficient γ in the first stage regression (2). This in turn translates into IV estimates that understate the impact of self-esteem on earnings. The IV coefficients on the measure in '87 are reported in the second panel of Table 6. Five different specifications estimated with IV and OLS are presented. Self-esteem in 1987 has a strong and statistically significant impact on wages across different specifications. A two standard deviation increase in the measure of self-esteem causes an increase in log earnings between 18 and 26 percent, depending on the specification. The main result is that the IV estimates indicate an impact of self-esteem that is two to three times higher than OLS estimates imply (see the last panel of the table). As the unobserved heterogeneity due to omitted variables and reverse causality probably leads the OLS coefficient upward, measurement error seems to be the main

---

[8] The clubs include performing arts clubs, hobby and athletics groups, newspapers and government groups, among others.

interpretation for such differences between OLS and IV estimates.[9] The analysis so far was based on individuals born after 1961 because most of these individuals did not entered labor market at the date of their interview in 1980 (Neal and Johnson, 1966). As a robustness check I perform the overall analysis for individuals born after 1962 and 1963. For these workers there are stronger reasons to believe that self-esteem in 1980 has no direct effect on wages in 1988. Indeed, it is more likely that no one of them had any experience in the labor market in 1980 at the date of interview. As reported in Table 7, the IV estimates for these sub-samples are precisely estimated and much larger than the previous ones. To provide further evidence that OLS underestimate the impact of self-esteem on earnings, in the NLSY I use the sub-sample of siblings. For the individuals born after 1961, I use as instrumental variable the sibling's self-esteem in 1980. There is substantial variation between the measures of self-esteem in 1980 among siblings. Table 8 reports the main results of this analysis. The sub-sample in this case is composed of about 250 individuals. Although imprecise, it is remarkable that IV estimates are similar and in some cases higher than those obtained in Table 6. Finally, in Table 9 I present the results by using the measures of self-esteem constructed in a different way from that described in section 2. Table 9 reports results for the two measures in 1980 and 1987 constructed by using factor analysis on the 10 questions of the Rosenmberg Self-esteem Scale. Among the four factors retained, I use the scoring coefficients of the factor that delivers all factor loadings positive. From Table 9, self-esteem has a stronger and statistically significant impact than that obtained before. Again, the IV estimates are still two to three times higher than OLS estimates.

5 Conclusion

This paper provides further evidence on the causal effect of self-esteem on earnings. The identification strategy exploits variation of the self-esteem measures in 1980 and 1987 and addresses the three sources of unobserved heterogeneity (measurement error, omitted variable and reverse causality). I find a large and positive effect of self-esteem on earnings. In particular, the IV estimates indicate an impact of self-esteem that is two to three times higher

---

[9] One should also note that if there is also measurement error in the dependent variable, depending on the sign of the correlation between the errors in self-esteem and earnings, under the classical error framework, the attenuation bias is even larger (Hyslop and Imbens, 2002).

than OLS estimates imply. The main explanation for this result is the large extent of measurement error in the reported self-esteem measures. This is relevant because in the context of this application reverse causality and omitted variables would likely bias the OLS coefficient upward, casting doubts on the existence of a causal relationship between self-esteem and earnings. This paper contributes to the literature on the behavioral determinants of earnings. In the analysis of other research questions about the effects of behavioral traits on earnings (e.g., the impact of happiness on earnings) the attenuation bias induced by measurement error might prevail on the potential sources of unobserved heterogeneity which would probably lead to positive spurious correlations. While this paper shares with others all the criticisms about IV estimates, a caveat of this study is that it is not possible to completely rule out the alternative explanation that the instrument used (self-esteem in 1980) has a direct, large effect on wages eight years later. However, the fact that the results are obtained on individuals for whom the schooling choices were constrained by compulsory schooling laws until 1980, the various robustness checks (e.g., saturation of the model) and the results on siblings confirming the baseline empirical patterns, suggest that OLS estimates, in fact, understate the effect of self-esteem on earnings.

References

Benabou, R. and Tirole, J., 2002. Self-confidence and Personal Motivation. Quarterly Journal of Economics 117, 871-915.

Bertrand, M., Mullainathan, S., 2001. Do people mean what they say? Implications for subjective survey data. American Economic Review 91, 67–72.

Bowles, S. and Gintis, H., 1976. Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. New York: Basic Books.

Bowles, S., Gintis, H. and Osborne, M., 2001. The Determinants of Earnings: a Behavioral Approach. Journal of Economic Literature 39, 1137-1176.

Borghans, L., Duckworth, A.L., Heckman, J.J. and ter Weel, B., 2008. The economics and psychology of personality traits. Journal of Human Resources, forthcoming.

Comte, O. and Postlewaite, A., 2004. Confidence-Enhanced Performance. American Economic Review 94, 1536-1557.

Cawley, J., J. Heckman and Vytlacil, E., 2001. Three Observations on Wages and Measured Cognitive Ability. Labour Economics 8, 419-442.

Costa, P. T. and McCrae, R. R., 1988. From Catalog to Classification: Murray's Needs and the Five-Factor Model. Journal of Personality and Social Psychology 55, 258-65.

Donohue J. J. and Levitt, S. D., 2006. Measurement Error, Legalized Abortion, the Decline in Crime: A Response to Foote and Goetz. Mimeo, University of Chicago.

Heckman, J. Stixrud, J and Urzua, S., 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. Journal of Labor Economics 24, 411-482.

Hyslop, D. and Imbens, G., 2002. Bias from Classical and Other Forms of Measurement Error. Journal of Business and Economic Statistics 19 475-481.

Goldsmith, A. H., Veum, J. R. and Darity, Jr, W., 1997. The Impact of Psychological and Human Capital on Wages. Economic Inquiry 25, 815- 829.

Graham, C., Eggers, A. and Sukhtankar, S., 2004. Does happiness pay? An exploration based on panel data from Russia. Journal of Economic Behavior & Organization 55, 319-342.

Mobius, M. and Rosemblat, T. S., 2006. Why Beauty Matters. American Economic Review 96, 222-235.

Murnane R. J., Willett J. B., Braatz M. J. and Duhaldeborde, Y., 2001. Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. Economics of Education Review 20, 311-320.

Neal, A., D. and Johnson, W. R., 1996. The Role of Premarket Factors in Black-White Wage Differences. Journal of Political Economy 104, 869-895.

Persico, N., A. Postlewaite and Silverman, D., 2004. The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height. Journal of Political Economy 112, 1019-1053.

Santos Pinto, L. and Sobel, J., 2005. A Model of Positive Self-Image in Subjective Assessment. American Economic Review 95, 1368-141

Waddell, G. R., 2006. Labor-Market Consequences of Poor Attitude and Low Self-Esteem in Youth. Economic Enquiry 44, 69-97.

Table 1. Summary statistics

	Mean	Std	Median	25th percentile	75th percentile
self-esteem in '80	29,55	3.37	29	27	32
self-esteem in '87	30.63	3.52	30	27	34

Table 1 reports summary statistics for white males (number of individuals 2250). I include only white males interviewed both in 1980 and 1987 and that reported a wage in their current job in 1988. I do not include individuals enrolled in school in 1988. I drop also the individuals that reported a wage per hour lower than 1 dollar or greater than 100 dollar in the current job in 1988. See section 2 for the construction of the sample.

Table 2. Summary statistics. Self-esteem and individual characteristics

	self-esteem in '80 Median or below	self-esteem in '80 Above median	self-esteem in '87 Median or below	self-esteem in '87 Above median
hourly wage in 1988 (in cents)	823	1022	827	1008
	(497)	(669)	(563)	(608)
AFQT score	41.66	60,59	41.93	59,31
	(28.78)	(26.50)	(29.29)	(26.61)
Years of education	10.39	11.75	12.08	13.67
	(1.85)	(1.80)	(2.30)	(2.37)
Mother's years of schooling	11.33	12.14	11.25	12.16
	(2.51)	(2.43)	(2.49)	(2.44)
Father's years of schooling	11.26	12.57	11.24	12.50
	(3.37)	(3.38)	(3.41)	(3.35)
Height in 1985	70,34	70.71	70.76	70.68
	(3.01)	(2.61)	(2.96)	(2.69)

Entries are means and in parenthesis standard deviation on white males.

Table 3. OLS estimates

	Self-esteem in '87	Age	AFQT	mother edu	father edu
N=1963	0.613**	.0257**	.1370**	−.0029	0.153**
R- squared=0.181	(.0112)	(.0049)	(.0134)	(.0056)	(.0040)

Dependent variable log wage in 1988. Robust standard errors in parenthesis. The AFQT scores and self-esteem are standardized with average zero and standard deviation equal to 1 as explained in section 2.
**: Significantly different from 0 at the 0.05 level.

Table 4. Reduced form regressions

	1	2	3	4	5
	N=887	N=787	N=877	N=887	N=777
self-esteem in '80	.0909**	.0695**	.0425**	.0613**	.0370**
(standardized)	(.0149)	(.0165)	(.0153)	(.0154)	(.0166)
Age	.0619**	.0583**	.0550*	.0564**	.0577**
	(.0185)	(.0196)	(.0177)	(.0154)	(.0189)
Mother's years of schooling	-	.0059	-	-	−.0064
		(.0085)			(.0082)
Father's years of schooling	-	.0224**	-	-	.0050
		(.0060)			(.0063)
AFQT score (standardized)	-		.1480**	-	.1082**
			(.0158)		(.0232)
Years of schooling	-		-	.0462**	.0278**
				(.0075)	(.0094)
R-squared	.0494	.0759	.1385	.1141	.1533

Dependent variable: log wage in 1988. Regressions on white males born after 1961. Robust standard errors in parenthesis. Self-esteem and AFQT score are standardized with average zero and standard deviation equal to 1 as explained in section 2.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

Table 5. Reduced form regressions, robustness checks

	1 N=770	2 N=775	3 N=770	4 N=764
Self-esteem in '80 (standardized)	.0375** (.0169)	.0372** (.0169)	.0366** (.0171)	.0367** (.0172)
Height in '85	.0076 (.0055)	-	-	.0098* (.0154)
Dummy on social activities	NO	NO	YES	YES
Dummy on weight in '88	NO	YES	NO	YES
R-squared	.1577	.1592	.1567	.1692

Dependent variable: log wage in 1988. Regressions on white males born after 1961. Robust standard errors in parenthesis. Self-esteem is standardized with average zero and standard deviation equal to 1 as explained in section 2.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

Table 6. IV results. Instrumental variable: self-esteem in '80

First stage	1 N=877	2 N=777	3 N=777	4 N=765	5 N=761
self-esteem in '80 (standardized)	.3251** (.0320)	.3457** (.0335)	.3374** (.0334)	.3276** (.0339)	.3225** (.0340)
R-squared	.1983	.2267	.2363	.2511	.2588
Second stage					
self-esteem in '87 (standardized)	.1308** (.0475)	.1176** (.0488)	.1092** (.0492)	.0934** (.0496)	.0847* (.0502)
OLS estimates					
self-esteem in '87 (standardized)	.0593** (.0161)	.0533** (.0175)	.0466** (.0174)	.0399** (.0170)	.0395** (.0171)
R-squared	.1445	.1446	.1562	.2418	.2410

Dependent variable: log wage in 1988. Robust standard errors in parenthesis. Regressions on white males born after 1961. In the first column age and standardized AFQT are the additional covariates. In the second one I add to AFQT and age, mother education and father education. The third column includes also educational attainment. The fourth, occupation, height and weight. In the last column dummy on social activities are included. Self-esteem is standardized with average zero and standard deviation equal to 1 as explained in section 2.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

Table 7. IV results for individuals born after 1962 and 1963

	1	2	3	4
Second stage	N=483	N=472	N=225	N=221
self-esteem in '87	.1579**	.1514*	.2291**	.2305**
(standardized)	(.0731)	(.0791)	(.1006)	(.1056)
OLS estimates				
self-esteem in '87	.0528**	.0411**	.0864**	.0790**
(standardized)	(.0215)	(.0220)	(.0283)	(.0286)
R-squared	.1852	.2734	.2099	.3281

Dependent variable: log wage in 1988. Instrumental variable: self-esteem in '80. Robust standard errors in parenthesis. Regressions on white males born after 1962 in column 1 and 2. Regressions on white males born after 1963 in column 3 and 4. In the first and third columns age, standardized AFQT are, years of education and parental background are included in the regression. In the second and fourth column I add to these variables dummy on occupation, height, weight and participation in social activities. Self-esteem is standardized with average zero and standard deviation equal to 1 as explained in section 2.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

Table 8. IV results. Instrument: sibling's self-esteem in 1980

	1	2	3	4	5
First stage	N=263	N=232	231	230	230
sibling's self-esteem in '80	.0552**	.0584**	.0750**	.0551**	.0480**
(standardized)	(.0169)	(.0179)	(.0175)	(.0171)	(.0174)
R-squared	.1120	.1237	15.09	17.52	.2235
Second stage					
self-esteem in '87	.1685	.2166	.1989	.1566	.2107
(standardized)	(.1396)	(.1467)	(.1494)	(.1569)	(.1820)

Dependent variable: log wage in 1988. Robust standard errors in parenthesis. Regressions on white males born after 1961. In the first column age and standardized AFQT are the additional covariates. In the second I add to AFQT and age, mother education and father education. In addition, the third column includes educational attainment. In the fourth, occupation, height and weight are included. In the last dummy on social activities are included. Self-esteem is standardized with average zero and standard deviation equal to 1 as explained in section 2.

*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

Table 9. IV results. Self-esteem constructed using factor analysis

	1	2	3	4	5
First stage	N=877	N=777	N=777	N=765	N=761
self-esteem in '80	.3410**	.3619**	.3437**	.3466**	.3398**
	(.0351)	(.0367)	(.0364)	(.0367)	(.0373)
R-squared	.1929	.2186	.2385	.2509	.2588
Second stage					
self-esteem in '87	.2753**	.2307**	.2065**	.1868*	.1789*
	(.1002)	(.1012)	(.1060)	(.1062)	(.1071)
OLS estimates					
self-esteem in '87	.1243**	.1046**	.0779**	.0668**	.0663**
	(.0332)	(.0343)	(.0345)	(.0336)	(.0338)
R-squared	.1369	.1370	.1528	.2370	.2403

Dependent variable: log wage in 1988. Robust standard errors in parenthesis. Regressions on white males born after 1961. In the first column age and standardized AFQT are the additional covariates. In the second one I add to AFQT and age, mother education and father education. The third column includes also educational attainment. The fourth, occupation, height and weight. In the last column dummy on social activities are included. The measures of self-esteem in 1980 and 1987 are constructed using factor analysis on the 10 survey questions of the Rosenmberg Self-esteem Scale. For each measure only four factors are retained. The measure are built by using the scoring coefficients of the factor for which the factor loadings are all positive.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.