Estimating the Veteran Effect with Endogenous Schooling when Instruments are Potentially Weak[1]

Saraswata Chaudhuri[2] and Elaina Rose[3]

Date: September 14, 2010

Abstract

Instrumental variables estimates of the effect of military service on subsequent civilian earnings – i.e., the veteran effect – either omit schooling or treat it as exogenous. Because military service is associated with schooling, and both military service and schooling are potentially correlated with unobservables in earnings equations, estimates of the veteran effect will depend upon the treatment of schooling. In a general setting that also allows for the treatment of schooling as endogenous, we estimate the veteran effect for men who were born between 1944 and 1952 and thus reached draft age during the Vietnam era. We illustrate how alternative treatments of schooling affect the estimates. We note that key instrumental variables used in our study tend to be only weakly correlated with the endogenous regressors. Hence we apply a variety of state-of-the-art econometric techniques to diagnose the extent of the weak instrument problem and to subsequently overcome it. We find that approximately 9 – 10 years after service, Vietnam-era veterans experienced a penalty in terms of civilian earnings.

JEL Classification: C2; J24

Keywords: Veteran effect; Draft lottery; Instrumental Variables; Weak instruments

---

[1] We are grateful to Colin Cameron, Scott Carrell, Jim Hosek and Eric Zivot for their helpful comments.

[2] Department of Economics, University of North Carolina, CB 3305, Chapel Hill, NC 27599, USA. Tel: (919) 966-3962. Fax: (919) 966-4986. Email: saraswata_chaudhuri@unc.edu.

[3] Department of Economics, University of Washington, Box 353330, Seattle, WA 98195, USA. Tel: (206) 543-5237. Fax: (206) 685-7477. Email: erose@u.washington.edu.

1. Introduction

The ongoing conflicts in Iraq and Afghanistan cause us to consider the economic consequences of war and, in particular, the effects of military service on veterans (Stiglitz and Bilmes, 2008). Yet it is unclear *a priori* whether military service generates a net premium or penalty in terms of the veterans' subsequent earnings. On one hand, the military provides opportunities for human capital acquisition, which, for many veterans, would not otherwise be available. On the other hand, time spent in the military comes at the expense of civilian human capital and labor market experience. Wartime combat poses risks to both physical and mental health that can adversely affect subsequent productivity. It is too early to know the effects of military service on all but the very short term outcomes for veterans of the current wars. However, knowing the effect of service in prior eras helps inform our understanding of the consequences of military service now. For this reason a substantial literature has evolved on the *veteran effect* – the effect of military service on subsequent civilian labor market earnings.

The central issue in the literature on the veteran effect is the endogeneity of military service. Young men choose whether or not to volunteer. Even during the Vietnam draft era that we study, many young men availed themselves of a variety of opportunities to avoid service. Instrumental variables (IV) can correct for the endogeneity bias in the estimates of veteran effect. The challenge is finding a suitable instrument for military service.

There is considerable variation in the treatment of schooling in empirical models of the veteran effect. Sometimes schooling is omitted. Sometimes it is included and treated as exogenous. But there are many reasons to think that military service and schooling are correlated, and when they are estimates of the veteran effect will depend on the treatment of schooling. Our first objective is to estimate the veteran effect holding schooling constant and gauge how the estimates are affected by alternative treatments of schooling. Our data are from the National Longitudinal Survey of Young

Men (NLS-YM). The NLS-YM provides instruments often used in the literature on the veteran effect and the returns to schooling. These include draft lottery parameters, as introduced by Angrist (1989, 1990, 1991) to instrument for military service, and measures of distance from nearest college, as introduced by Card (1995) to instrument for schooling.

In our sample the instruments are only weakly correlated with the endogenous regressors and, therefore, the standard methods – the IV estimation and the t- and F-tests – tend to produce incorrect inference beyond the allowable margin of statistical error (see Staiger and Stock, 1997). This exercise, then, provides an ideal laboratory for applying the recently proposed weak-instrument-robust methods for inference in a model with two endogenous variables (see, for example, Kleibergen; 2004, 2008; Kleibergen and Mavroeidis, 2009a). Our second objective is to demonstrate the usefulness of recently developed methodological tools for inference with weak instruments.

Section II of this paper provides background on the literatures on the veteran effect and returns to schooling and considers the issues at the nexus of the two. Section III outlines the empirical model, focusing on issues associated with weak instruments. Section IV describes the data. Section V discusses the result and Section VI concludes. Our key result is that even with the generally conservative weak-instrument-robust methods of inference we still confirm Angrist's (1989, 1990) findings that Vietnam-era veterans experienced a penalty, net of the effect of schooling, in terms of earnings about 9-10 years after service.

2. **The Veteran Effect with Endogenous Schooling**

2.1. The Veteran Effect: Premium or Penalty?

Does military service increase or decrease earnings? The answer to this question is unclear *a priori*. On one hand, there are opportunities for human capital acquisition, which, for many, would not otherwise be available. The military provides on-the-job training; and college education is subsidized pre-service, in-service and post-service through GI Bills.[4]

Servicemen gain less measureable forms of human capital as well. For instance, the military serves as a "bridging environment" in which youths from disadvantaged backgrounds can learn skills such as an understanding of how to function in a structured environment (Teachman and Call, 1996). Successful completion of a term of service signals favorable pre-market ability and acquired unobservable skills (DeTray, 1982). Taken together, these factors suggest that veterans will receive a premium when they return to the civilian sector.

Yet military service entails costs as well as benefits. Draftees are drawn away from their otherwise optimal human capital investment paths. So are young men who enlist in order to preempt being drafted into the infantry and those enlisting for non-financial motives such as patriotism or family tradition. Soldiers exposed to combat experience adverse physical and mental consequences. Hence it is unclear whether, on net, there is a veteran premium or veteran penalty.

Estimates of the veteran effect vary considerably by era of service and approach to estimation. For instance, Rosen and Taubman (1982) suggest that World War II (WWII) veterans earned a premium of 10 percent, while Vietnam-era veterans suffered a penalty of 19 percent.[5]

However, as Rosen and Taubman note, their OLS estimates may be asymptotically biased because military service is endogenous. Moreover, the direction of the bias in the OLS estimates is ambiguous. On one hand, youths with stronger opportunities in the civilian sector have less incentive to join the military and a greater incentive to avoid the draft. On the other hand, those with

---

[4] Rostker (2006) outlines programs available from 1973 to 2004. New programs continue to be introduced. The best references on current offerings are the services' recruiting web pages.

[5] There appears to be no effect of military service on the earnings of Korean War veterans (Schwartz, 1986). OLS-type studies also find that the veteran effect varies by role in the military and by individual characteristics. For instance, Air Force veterans tend to earn more than veterans of other services (MacLean and Elder, 2007). Officers tend to fare better than enlisted personnel (MacLean, 2008). Technical skills transfer more readily to the civilian sector (Bryant and Wilhite, 1990; Goldberg and Warner, 1987). Costs of service are greatest for draftees and soldiers exposed to combat (Teachman, 2004; MacLean and Elder, 2007). Blacks achieve greater premia and suffer smaller penalties than whites (Bryant, Samaranayake and Wilhite, 1993; Teachman and Tedrow, 2004).

sufficiently low physical or cognitive ability cannot qualify. If WWII veterans were positively selected into military service and Vietnam-era veterans were negatively selected into military service, then the contrast between the OLS estimates of the WWII and Vietnam-era veteran effects could be explained entirely by this endogeneity.

Instrumental variable techniques allow researchers to overcome the problem of endogeneity and obtain consistent estimates of the veteran effect. Several studies exploit the randomness of draft lotteries as a valid instrument. Angrist and Krueger (1994) find no evidence of a WWII veteran premium; in fact there may be a small penalty. Angrist (1989, 1990) finds a Vietnam-era penalty of 15 percent for whites (but no effect for blacks) about 10 to 20 years after service. Angrist and Chen (2008) look at longer-run outcomes of military service on earnings and find that the penalties found in Angrist's earlier work dissipated by about age 50. These IV studies indicate that both the WWII and Vietnam-era veteran effects are asymptotically biased: the former being biased upward and the latter being biased downward. Inferences about the effect of military service on earnings based on the OLS estimates would be very misleading.

2.2. Military Service and Schooling

As Rosenzweig and Wolpin (2000) note, the interpretation of all these estimates hinges on the treatment of schooling in the empirical model. There is an extensive literature on the returns to schooling emphasizing the fact that schooling is endogenous with respect to unobserved ability (see Card, 1999, for a survey). Those with more favorable labor market unobservables obtain more schooling, leading to an upward bias in the OLS estimates of the returns to schooling. On the other hand, measurement error in schooling will tend to bias estimates towards zero. A variety of IV approaches have been used to generate consistent and asymptotically unbiased estimates. The literature generally reports IV estimates exceeding comparable OLS estimates.

There are several reasons to think that military service is related to schooling. During the Vietnam-era, some young men deferred their obligations by going to college in order to avoid the draft (Card and Lemieux, 2001). Others attended college because their post-service education was funded by the GI Bill (Stanley, 2002; Bound and Turner, 2003; Angrist and Chen, 2008). Young men with service-related disabilities may have been less capable of returning to school.

Angrist and Chen (2008) is the first paper to estimate the effect of military service in the context of a model of human capital accumulation. They find no significant effect of Vietnam-era military service on labor market outcomes reported on the 2000 Census, that is, when the veterans were about 50 years old. They do find that military service is associated with higher levels of schooling and conclude that two to three decades after service, the initial veteran penalty due to lost work experience as found in Angrist (1989, 1990) was offset by gains from GI Bill-financed schooling.[6]

The joint endogeneity of schooling and military service has implications for the estimates of the veteran effect. When both are treated as exogenous, OLS estimates of the veteran effect will be asymptotically biased. IV estimates of the veteran effect will also be asymptotically biased if schooling and veteran status are correlated and schooling is treated as exogenous. In the most general case, when both schooling and veteran status are treated as endogenous, IV estimates with appropriate instruments will be consistent and asymptotically unbiased, and, therefore, reliable in sufficiently large samples.

But, as long as we instrument for veteran status, does it matter empirically how we treat schooling? We address this question by comparing results from a general model in which both veteran status and schooling are treated as endogenous with results from more restricted models. The instruments include lottery parameters, as used in Angrist (1989, 1990, 1991) and Angrist and

[6] GI Bills also increased schooling among veterans of WWII (Bound and Turner, 2002) and Korean War (Stanley, 2003) veterans.

Chen (2008) to instrument for military service, and measures of distance to nearest college, as used in Card (1995, 1999, 2001) and Kling (2001) to instrument for schooling. In the next section we specify our empirical models and outline our strategy for gauging the sensitivity of the veteran effect to alternative treatments of schooling. Because in our sample, instruments are only weakly correlated with the endogenous regressors, we undertake a variety of state-of-the-art approaches to make our results robust to the presence of weak instruments.

3. Estimation

3.1. Empirical Model:

We consider a limited information framework to estimate the net effect of a man's veteran status on his wage in the civilian market (later in his life cycle), after controlling for his years of schooling and other background characteristics. Accordingly, we model w, i.e., the logarithm of the real wage of a man in the civilian labor market as

$$w = V\gamma + S\beta + X\theta + \xi, \tag{1}$$

where V (=1) is a dummy variable indicating whether he ever served in the military, S is his years of schooling, and X contains an intercept term and a set of background variables including his demographic, household and locational characteristics. ξ represents the model error and includes components such as the unobservable human capital and the ability of the man.

We are interested in the coefficient γ, which we would interpret as the net percent change in real wage attributable to veteran status, after controlling for years of schooling and other background characteristics. Halverson and Palmquist (1981) note that for not so small values of γ, the appropriate measure of percent change in real wage should be $\delta = (e^{\gamma} - 1)$. Our estimates for γ are not small and hence we also report estimates for δ. The coefficient β measures the net returns to schooling, after controlling for the veteran status and other background characteristics.

One challenge in conducting inference on γ is the endogeneity of V and S. As discussed in Section 2, it is quite likely that the characteristics (unobserved to the researcher) based on which the military accepted or rejected individuals, and the considerations that led individuals to volunteer for the service, also affected their wages in the civilian labor market. This leads to a correlation between V and ξ (probably even after partialing out S). In terms of schooling, the potential correlation between S and ξ is emphasized in the literature on the returns to schooling. The endogeneity, coupled with measurement error in either variable, prohibits the formulation of the error term in (1) as $\xi = w - E[w|V, S, X]$ and rules out the possibility of consistent estimation of γ (and β) using OLS.

Even if we instrument for veteran status, the veteran effect will be asymptotically biased if schooling is endogenous and schooling and veteran status are correlated. We cannot sign the bias *a priori*. As discussed in Section 2.2, veteran status may be either positively or negatively correlated with schooling. In our sample, veterans have, on average, slightly more schooling than the non-veterans. Of the 1080 veterans included in our sample, more than 61.5 percent went for additional schooling since their first (or only) term with the armed forces; and on average they got about 1.32 additional years of schooling. While it does not necessarily establish a causal effect of one's veteran status on schooling, this certainly calls for a thorough inspection of the treatment of schooling in the specification described by (1).

3.2. Estimation using Instrumental Variables

Likely endogeneity of veteran status and schooling and evidence of correlation between them suggest that simple OLS methods cannot consistently estimate γ. Therefore, the OLS-based estimate of $\delta = e^{\gamma} - 1$, i.e., the veteran effect net of schooling, is also inconsistent. IV is the most common method of inference in these situations. We use the following four instruments – (i) the lottery number assigned to the young man based on his date of birth, (ii) the lottery ceiling for the year when this young man attained draft age, (iii) a dummy variable indicating the presence of a four year accredited public college and (iv) a dummy variable indicating the presence of a four year accredited private college in the neighborhood of the young man's residence in 1966. The first two instruments, following Angrist (1989, 1990, 1991) and Angrist and Chen (2008), characterize the draft eligibility of the respondents. The other two instruments, following Card (1995, 1999, 2001) and Kling (2001), capture proximity of the respondents' residence to an accredited four-year college.

In our setup, IV estimation amounts to either two-stage least squares (TSLS) or limited information maximum likelihood (LIML). A necessary condition for the convergence of these estimators to the true value of the parameters is that the instruments are exogenous which, in turn, implies that the variation induced by the instruments in the endogenous regressors is uncorrelated with the unobserved error ξ in (1). Unfortunately, it is not possible to test exogeneity of the instruments without the prior assumption that the model is over-identified, i.e., there are at least three instruments, and at least two independent linear combinations of these instruments are exogenous. As discussed in Section 5, we try to overcome this limitation by considering various alternative specifications while testing exogeneity of the instruments and also by using joint tests of hypotheses on parameter values and the exogeneity restrictions.

The attention to exogeneity, on the other hand, also leads us to be conservative in the choice of instruments and restricts us from capturing some variations in the endogenous explanatory variables veteran status and schooling. Ideally, asymptotic efficiency of the inference should be the only virtue at stake here. However, as Bound, Jaeger and Baker (1995) emphasize, this could also give rise to the so-called "weak instrument problems". As a consequence, the TSLS and LIML estimates can be inconsistent and asymptotically biased, and the t-test and F-test can (and tend to) over-reject the true value of the parameters (see Staiger and Stock, 1997).[7]

[7] It is important to distinguish between the two types of problems that can arise due to weak instruments. The first problem is a reduction in precision; this is natural because the data do not contain enough information to precisely

To overcome such problems we also consider the recently proposed "weak-instrument-robust" methods of inference. The broader aspects of our conclusion remain unchanged. The weak-instrument-robust methods provide a way to test the parameters of interest. These tests can subsequently be inverted to obtain confidence regions for the parameters. Unlike the t-test and F-test, these tests do not over-reject the true value of the parameters (as long as the instruments are exogenous). Hence the corresponding confidence regions for the parameters have at least correct asymptotic coverage probability. Extensive simulations in Kleibergen (2008), Kleibergen and Mavroeidis (2009a) and Chaudhuri and Zivot (2010) also show that the asymptotic results for the weak-instrument-robust methods provide good approximations to their behavior in finite samples.

In particular, we report results based on the weak-instrument-robust plug-in based subset K, subset KJ and subset conditional likelihood ratio (CLR) tests proposed by Kleibergen (2004, 2008) (also see Moreira, 2003). Results in Kleibergen and Mavroeidis (2009a) and Chaudhuri and Zivot (2010) indicate that these are the best plug-in based methods and should be generally preferred over their projection-based counterparts. Unlike the conventional methods such as TSLS or LIML estimation and the t and F tests based on them, these weak-instrument-robust methods for subset inference do not report incorrect parameter values with spuriously high precision or over-reject the true parameter value in the presence of weak instruments. Rather, they correctly reflect the lack of information due to weak instruments and produce large confidence regions with conservative (i.e., at least correct) coverage probability. However, despite their conservativeness, the use of the weak-instrument-robust methods allows us to strongly reject a zero or positive veteran effect.

4. **Data**

---

estimate the parameters in the model. The second problem is the so-called "weak instrument problem" and this refers to the case where the conventional first-order asymptotic results provide poor approximation to the finite sample behavior of the estimators and tests; namely, the usual estimates tend to precisely report wrong values of the parameters and the usual tests tend to over-reject the true value of the parameters. Weak-instrument-robust methods were developed to address the second problem and overcome such misleadingly spurious precision in the usual methods of inference.

We use data from the National Longitudinal Survey of Young Men (NLS-YM) to estimate the unknown parameters in (1). The NLS-YM is a nationally representative data set of young men aged 14–24 in 1966. Respondents were followed annually until 1971, and then annually or biennially until 1981.

Men born between 1944 and 1952, who constitute about 82 percent of the survey respondents, were subject to the annual lotteries from 1969 through 1972. This is the cohort we study. Veteran status is captured in two ways. First, there are a number of specific questions about military service. Second, the data indicate whether a respondent was unavailable for the survey in a given year because he was serving in the military. Schooling is measured as the highest grade completed that is reported on the survey.

The dependent variable, (log) real hourly earnings, is measured in 1981 dollars at the oldest age at which the respondent appeared on the survey. In order to capture the effect of veteran status (and schooling) as late as possible in the man's life-cycle, we further restrict our attention to men whose last recorded wage was earned at the age of 29 or more.[8] Ignoring the 1.69 percent of respondents with missing wage figures, 65.75 percent of men interviewed in the survey reported their last wage at age 29 or more.[9] Lastly we ignore one respondent with an implausible birthday and three respondents with missing information on the type of area (urbanized, urban place or rural) of the respondent's residence in 1966. Our final sample consists of 2754 respondents.

In all, 1080, or 39 percent of the final sample, were veterans and 1674 were non-veterans. Sample size and reporting issues preclude us from disaggregating by rank, service or military occupation. Highest year of schooling completed was, on average, slightly higher for veterans (13.6) than non-veterans (13.4). However, the partial correlations of veteran status and schooling,

---

[8] Among all the respondents born between 1944 and 1952, the wage figures are missing for 66 men and 6 reported 0 wage.

[9] 60.39 percent satisfies the stricter criterion of last recorded wage being earned at age 30 or more.

controlling for the control variables X in equation (1) is negative. These controls include race, region,[10] urbanicity[11] and the age and year at which the wage was earned.

The NLS-YM provides suitable instruments for both veteran status and schooling. As discussed in Section (3), we use the following four variables as instruments: (i) the lottery number assigned to the young man based on his date of birth, (ii) the lottery ceiling for the year when this young man attained draft age, (iii) a dummy variable indicating the presence of a four year accredited public college and (iv) a dummy variable indicating the presence of a four year accredited private college in the neighborhood of the young man's residence in 1966.

The full set of sample statistics is reported in Table 1.

5. Results

5.1. Specifications

In order to illustrate the sensitivity of the estimates of veteran status to different treatments of schooling and treatment of potentially endogenous variables we compare results of five specifications which are variations of equation (1). Results are reported in Table 2(a). Corresponding estimates and standard errors for the control variables are reported in Appendix Table 2(b).

Our most general IV specification, (A), treats both veteran status and schooling as endogenous. The point estimate suggests a large veteran penalty: The coefficient γ is -.374 and corresponds to a veteran effect of -31.2 percent ($\delta = e^{\gamma} - 1$). The standard error of this estimate is rather large, around 15 percent (obtained by the Delta method); and a 95 percent confidence region suggests that the wage reduction for veterans can vary from 1 percent to 61 percent. Nevertheless, we can safely reject that the net veteran premium is zero or positive.

[10] The regions are northeast, mid-atlantic, east north central, west north central/mountain, east south central, west south central, pacific; and south atlantic as a default.
[11] The area is categorized as urbanized if its population is more than 125,000, as an urban place if the population is between 12,000 and 125,000, and rural otherwise.

The second panel reports the p-value from Sargan's test of over-identification. This is a specification test that, among other things, also tests for the exogeneity of the instruments. The test fails to reject the specification, and the extremely large p-value (.978) suggests that subsequent statistical inference based on specification (A) is unlikely to be adversely affected by the assumption of instrument exogeneity.

The third panel reports the results of test for endogeneity of the regressors – veteran status and schooling. The hypothesis that veteran status is exogenous is rejected at less than the 5 percent level. The hypothesis that schooling is exogenous is rejected at the 8.2 percent level. The hypothesis that schooling and veteran status are jointly exogenous is rejected at the 5.5 percent level. None of these results is surprising, given the attention paid to these points in the empirical literature.

In principle, failure to account for the endogeneity of schooling, given its correlation with veteran status, can adversely affect the estimate of γ, and consequently, δ, and alter their interpretation. One key question in this paper is whether incorrectly assuming that schooling is exogenous affects the estimate of the veteran effect. To answer this question we turn to specification (B) where schooling is treated as exogenous.

In specification (B), the point estimate of γ is -.234 and, consequently, the estimate of veteran effect is -20.9 percent, or about one-third the magnitude of the estimate from specification (A). The standard error is rather large, and a generalized Hausman test based on the difference in the estimated γ from specification (A) and specification (B) rejects the equivalence of these two specifications (in terms of the coefficient of veteran status) at the 12 percent level. The only difference between specifications (A) and (B) is that schooling is treated as endogenous in the former and exogenous in the latter. If specification (B) is correct, then using specification (A) results in loss of precision. On the other hand, if specification (B) is incorrect and the Hausman test, because of its low power, fails to detect that, then inference based on specification (B) is prone to significant statistical errors (see Guggenberger, 2010).[12] Given the potential adverse consequences, we view a p-value of .12 too small to prefer specification (B) over specification (A). The same issue arises when comparing specifications (A) and (C), and also in the other applications of the Hausman test (see the first three rows in the fourth panel of Table 2(a)).

Some studies of the veteran effect have access to suitable instruments for veteran status, but because the researchers do not have access to data on schooling it is omitted. This is specification (C). The point estimate of γ is -.172 and, consequently, the veteran effect from (C) is -15.8 percent. This is smaller in absolute value than the (A) and (B) estimates. The specification does not allow the estimate of δ to be interpreted as the veteran effect holding schooling constant. It is likely that the estimate captures some combination of the effects of veteran status and schooling.[13]

It should also be noted that the relatively low p-value of the Sargan test in specifications (B) and (C) as compared to specification (A) provides less evidence for instrument exogeneity. This makes inference based on specifications (B) and (C) less convincing than inference based on specification (A).

The OLS specifications corresponding to specifications (A) and (C) are reported in columns (D) and (E), respectively**.** OLS is clearly inappropriate. Recall that the endogeneity tests from column (A) rejected the exogeneity of veteran status individually (p-value is .03) and jointly with schooling (p-value is .055). Moreover, despite its low power as pointed out by Guggenberger (2010), even the Hausman test does not fail to indicate that relative to specification (A), the (asymptotic) bias induced by specification (D) and (E) is substantial (p-values around .067). In each case the OLS estimates of the veteran effect are greater than in the three IV specifications (A, B and

[12] To see if the conclusion from the Hausman test is affected by the presence of weak instruments, we use all three forms of the statistic described in equation 3.9 (page 568) of Staiger and Stock (1997). The conclusion does not change with the other forms of the Durbin-Wu-Hausman statistic. These tests, under weak instruments, lack power.

[13] While it may be tempting to interpret the estimate as the overall veteran effect, i.e., the sum of the direct effect (as in specification (A)) and also the indirect effect operating through post-service schooling paid for by the GI Bill, this interpretation is not appropriate (see Joeffe et. al., 2008).

C). This indicates that OLS estimates are biased upward and suggests that servicemen were positively selected into military service – in terms of unobservables realized at about age 30 – during the Vietnam era.

Although not the focus of the study, it may be useful to consider the estimates of the return to schooling. Our OLS estimate in column (D) is about 5 percent. When schooling is treated as endogenous (as in column (A)) the estimate increases to about 16 percent. The latter is on the high side in light of the literature, but estimates based on this data using similar instruments are of comparable magnitude. Kling (2001) reports OLS estimates of 7 percent and IV estimates of about 16 percent. Card's (1995) estimate is about 7 percent under OLS and 13 percent under IV, and our point estimates are within one standard deviation of Card's.

5.2. Are these results affected by weak instruments?

Yes. The first stage F-statistics for testing the relevance of the (excluded) instruments are low: 8.46 for veteran status and 2.53 for schooling. The partial R^2 statistics are .012 and .004 respectively for the two endogenous regressors (see Shea, 1997). Hence there is evidence that the instruments are weak and do not explain much variation of the endogenous regressors. A more systematic test for (the effect of) weak instruments in the presence of multiple endogenous regressors is the test proposed by Stock and Yogo (2005). This test concludes the presence of weak instruments if the minimum eigen value of the "so-called" concentration matrix is less than the appropriate critical value (depending on the number of exogenous instruments and endogenous regressors). In addition, the test also provides an estimate of the (adverse) effect of weak instruments on inference. The Stock and Yogo test suggests that given our empirical model and the exogenous instruments, the maximum (asymptotic) bias of a standardized linear combination of the TSLS estimators of γ and β, relative to that of their OLS estimators is more than 30 percent. If the instruments were strongly correlated with the endogenous regressors, this would be close to 0. The test also suggests that the nominal size of a 5 percent Wald test for jointly testing the significance of veteran status and schooling is likely to be more than 25 percent. Again, if the instruments were strongly correlated with the endogenous regressors, this would be close to 5 percent.

To gauge how seriously such problems affect our main conclusion on the veteran effect (i.e. specification (A)), we construct confidence regions for the effect using the recently proposed weak-instrument-robust methods of inference. These methods are valid as long as the instruments are exogenous.

5.3. Are the instruments exogenous?

Recall that the over-identification test cannot reject the exogeneity of the instruments even at 97 percent level. As noted in Section 5.1, the extremely large p-value (.978) for the over-identification test suggests that statistical properties of subsequent inference based on specification (A) are unlikely to be adversely affected by the assumption of instrument exogeneity.[14]

Intuitive justification of the exogeneity of these instruments is provided in the original papers by Angrist (1989, 1990, 1991) and Card (1995). The lottery parameters faced by each man was based on his date of birth and the demand for conscripts, and are arguably uncorrelated with his unobserved characteristics. It may, however, be less straightforward to intuitively justify the exogeneity of the other instruments – the presence of four year accredited public and private colleges in the vicinity of the respondent's residence in 1966. Nevertheless, given that the exogeneity of lottery ceiling and lottery number is more convincing, in Table 3 we report the results from the exogeneity test of the variables indicating the presence of colleges (individually and jointly) under the assumption that the variables involving the lottery are exogenous. The minimum p-value for the over-identification test is .836; and for specification (A), that we actually use, the p-

[14] Unlike some other studies (and also specifications (B) and (C)), the p-value in this case is possibly large enough to buffer for the fact that the over-identification test may lack power in certain directions.

value is .978.[15,16] This further supports the claim that treating the presence of colleges as exogenous instruments is also unlikely to affect the statistical properties of our results. Nevertheless, we have taken further measures (discussed) in the next subsection to avoid the adverse consequences of the instrument exogeneity assumption on our key result based on the weak-instrument-methods of inference.

5.4. <u>Results on the veteran effect from weak-instrument-robust methods of inference:</u>

Comprehensive p-value plots for all the weak-instrument-robust tests for γ are presented in Figure 1. The horizontal grid represents the levels of the test and, therefore, for example, the values of γ for which the p-values are above the 5 percent grid (line) are included in the 95 percent confidence region for γ. As stated in equation (5.13) of Dufour (1997), given a confidence region $C(\gamma)$ for γ, one can always use the projection technique to obtain a projected confidence region for $\delta = e^{\gamma} - 1$ as

$$C(\delta) = \{ \delta_0 = e^{\gamma_0} - 1 | \gamma_0 \in C(\gamma)\}. \tag{2}$$

The coverage probability of the projected confidence region cannot be less than that of the original one. In this case, the strict monotonicity of the functional form implies that the lower bound is sharp.

In the text we only report the results based on the subset KJ test. The other weak-instrument-robust tests, i.e., subset K test and the subset CLR test, for which the p-values are also reported in Figure-1, give very similar confidence regions. However, since the subset KJ test also explicitly accommodates for the simultaneous testing of instrument exogeneity (i.e., the test for moment/exclusion restrictions), we feel more comfortable with the confidence regions obtained by

---

[15] We also use the H_3 (for columns 1 and 2 of Table 3) and the H_4 statistics (for all the columns) described in Hahn, Ham and Moon (2008) to test for exogeneity of the instruments. These are variations of the Hausman statistic extended to test the instrument exogeneity when the instruments are actually weak. The p-value for all these tests is very large and does not provide evidence against the exogeneity hypothesis. The results are not reported here because the weighting matrix of the quadratic form is near-singular in all cases and there may be some concern with the ill-conditioned computations.

[16] We speculate that the results of the over-identification tests using alternative sets of instruments indicate that the usual interpretation of the TSLS estimator as the local average treatment effect (on the compliers) could possibly be extended to an interpretation as the average treatment effect on the entire population under reasonable assumptions (see Angrist, 2004). However, this is beyond the scope of the present paper and is not pursued any further.

inverting this test. Not only do these regions have at least correct asymptotic coverage probability (because they are weak-instrument-robust), but also the exogeneity of instruments are satisfied (within user-specified margins of statistical error) at each point belonging to these regions. A 95 percent confidence region for γ, obtained by collecting all the values of the parameter that cannot be rejected by a subset KJ test of nominal level 5 percent, ranges from -.3 percent to -121 percent. Similarly, a 90 percent confidence region ranges from -5.4 percent to -95 percent. Hence, using the projection in equation (2), we obtain that a 95 percent confidence region for the veteran effect (net of schooling) ranges from a wage reduction of .3 percent to 70 percent, whereas the wage reduction in a 90 percent confidence region ranges from 5.3 percent to 61.3 percent.[17] Of course, this is very imprecise; the test is conservative in the presence of weak instruments. However, it is also interesting to note that, even with such degree of imprecision, we can strongly reject a zero or positive net effect of veteran status.

The confidence regions obtained by the weak-instrument-robust methods are very broad and contain returns that are hard to justify using economic theory. Nevertheless, the TSLS estimates (that are not supposed to be robust to weak instruments) are also included inside these robust confidence regions. Moreover, the TSLS-based conclusion that there is no veteran premium – and that there is a veteran penalty – at the age of 29 is also validated by these confidence regions.[18]

6. Conclusion

Estimating the effect of military service on civilian earnings poses several methodological challenges. The most important is the endogeneity of military service which renders OLS estimates

[17] The not-so-impressive performance of the subset CLR test in this particular sample was disappointing to us. Although the only theoretical result on the optimality of the CLR test (Moreira, 2003 and Andrews, et. al, 2006) is known for the joint test of all the structural coefficients (and not applicable to the subset test done here), from various Monte-Carlo simulations we expected a better performance of this test.

[18] The subset K and CLR statistics are, by definition, zero at the LIML estimates. Hence LIML estimates are always included in the confidence regions obtained by inverting these two tests. The LIML estimates are also extremely likely to be contained in the confidence regions obtained by inverting the subset KJ test unless the regions are heavily influenced by the J statistic (i.e. the difference between the subset AR and the subset K statistics). However, the same does not apply to the TSLS estimates. Hence, the fact that the TSLS estimate is included in the weak-instrument-robust confidence region is reassuring.

of the veteran effect inconsistent and asymptotically biased. IV methods, using lottery parameters as instruments for military service, have enabled researchers to correct for this.

A second issue is the endogeneity of schooling. When schooling is correlated with military service, treating schooling as exogenous or excluded can affect the consistency of the estimate of the veteran effect.

In this paper we use data from the NLS-YM to estimate the Vietnam-era veteran effect in a model that allows for both military service and schooling to be endogenous. We compare the estimated veteran effect to estimates from models in which schooling is treated as exogenous and models in which schooling is omitted entirely. Our point estimates suggest a veteran penalty of about 31.2 percent when both schooling and veteran status are treated as endogenous and 20.9 percent when only veteran status is treated as endogenous. Rosenzweig and Wolpin's (2000) point that schooling is endogenous is validated although it is less clear empirically that the endogeneity significantly biases the estimate of the veteran effect.

In our sample, the key instrumental variables turned out to be weakly correlated with military service and schooling. To avoid the weak instrument problem we, therefore, applied the recently proposed weak-instrument robust methods of inference (see, Kleibergen, 2004, 2008; Moreira, 2003; Kleibergen and Mavroeidis, 2009a). The conclusions remain unchanged.

There is one key substantive implication of this paper. Despite the fact that we have subjected the data to a series of tests for sensitivity to alternative treatments of schooling and applied conservative approaches for drawing inferences in the presence of weak instruments, we still confirm findings in the literature that approximately 9-10 years after Vietnam-era service, veterans suffered a significant penalty.

References

Andrews, D. W. K., M. J. Moreira and J. H. Stock. 2006. Optimal Two-sided Similar Invariant Tests for Instrumental Variables Regression. Econometrica 74: 715 – 752.

Angrist, J. 1989. Using the Draft Lottery to Measure the Effects of Military Service on Civilian Earning. *Research in Labor Economics* 10, Ronald Ehrenberg (ed.) Greenwich, CT, JAI.

Angrist, J. 1990. Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records. *American Economic Review* 80: 313 – 336.

Angrist, J. 1991. The Draft Lottery and Voluntary Enlistment in the Vietnam Era. *Journal of the American Statistical Association* 86: 584 – 595.

Angrist, J. 2004. Treatment Effect Heterogeneity in Theory and Practice. *The Economic Journal* 114: 52 – 83.

Angrist, J. and S. Chen. 2008. Long-Term Economic Consequences of Vietnam-Era Conscription: Schooling, Experience and Earnings. IZA DP No. 3628.

Bound, J., D. A. Jaeger, and R. M. Baker. 1995. Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak. *Journal of the American Statistical Association* 90: 443 – 450.

Bound, J. and S. Turner. 2002. Going to War and Going to College: Did World War II and the G.I. Bill Increase Educational Attainment for Returning Veterans? *Journal of Labor Economics* 20: 784-815.

Bryant, R. and A. Wilhite. 1990. Military Experience and Training Effects on Civilian Wages. *Applied Economics* 22: 69 – 81.

Bryant, R., V. Samaranayake and A. Wilhite. 1993. The Effect of Military Service on the Subsequent Civilian Earnings of the post-Vietnam Veteran. *Quarterly Review of Economics and Finance* 33: 15 – 31.

Card, D. 1995. Using Geographical Variation in College Proximity to Estimate the Return to Schooling. *Aspects of Labour Market Behavior: Essays in Honor of John Vanderkamp*, Christofides, L. N., E. K. Grant and R. Swidinsky (eds). Toronto: University of Toronto Press.

Card, D. 1999. The Causal Effect of Education on Earnings. *Handbook of Labor Economics* 3, Orley Ashenfelter and David Card (eds). Amsterdam: Elsevier.

Card, D. 2001. Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica* 69: 1127 – 1160.

Card, D. and T. Lemieux. 2001. Draft Avoidance and College Attendance: The Unintended Legacy of the Vietnam War. *American Economic Review Papers and Proceedings* 91: 97 – 107.

Chaudhuri, S. and E. Zivot. 2010. A new method of projection-based inference in GMM with weakly identified parameters. Technical Report. University of North Carolina (Chapel Hill) and University of Washington.

De Tray, D. 1982. Veteran Status as a Screening Device. *American Economic Review* 72: 133 – 142.

Dufour, J. M. 1997. Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models. *Econometrica* 65: 1365 – 1388.

Goldberg, M. and J. Warner. 1987. Military Experience, Civilian Experience and the Earnings of Veterans. *Journal of Human Resources* 22: 62 – 81.

Guggenberger, P. 2010. The Impact of a Hausman Pretest on the Asymptotic Size of a Hypothesis Test. *Econometric Theory* 26, 369 – 382. doi:10.1017/S0266466609100026.

Hahn, J., J. Ham and H. R. Moon. 2008. The Hausman Test and Weak Instruments. Technical Report, Department of Economics, University of California and University of Southern California.

Halvorsen, R. and R. Palmquist. 1980. The Interpretation of Dummy Variables in Semilogarithmic Equations. *American Economic Review* 70: 474 – 475.

Kleibergen, F. 2004. Testing Subsets of Parameters in the Instrumental Variables Regression Model. *The Review of Economics and Statistics* 86: 418 – 423.

Kleibergen, F. 2005. Testing Parameters In GMM Without Assuming That They Are Identified. *Econometrica* 73: 1103 – 1123.

Kleibergen, F. 2008. Subset Statistic in the Linear IV Regression Model. Technical Report, Department of Economics, Brown University.

Kleibergen, F. and S. Mavroeidis. 2009a. Inference on subsets of parameters in GMM without assuming identification. Technical report, Department of Economics, Brown University.

Kleibergen, F. and S. Mavroeidis. 2009b. Weak Instrument Robust Tests in GMM and the New Keynesian Phillips Curve. *Journal of Business and Economic Statistics* 27: 293 – 311.

Kling, J. R. 2001. Interpreting Instrumental Variables Estimates of the Returns to Schooling. *Journal of Business and Economics Statistics* 19: 358 – 364.

Joeffe, M. M., D. Small, T. T. Have, S. Brunelli, H. I. Feldman. 2008. Extended Instrumental Variables Estimation for Overall Effects. *The International Journal of Biostatistics* 4, Article 4.

MacLean, A. and G. Elder. 2007. Military Service in the Life Course. *Annual Review of Sociology* 33: 175 – 196.

MacLean, A. 2008. The Privileges of Rank: The Peacetime Drat and Later-life Attainment. *Armed Forces and Society* 34: 682 – 713.

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

Moreira, M. J. 2003. A Conditional Likelihood Ratio Test for Structural Models. *Econometrica* 71: 1027 – 1048.

Oi, Walter. 1967. The Economic Cost of the Draft, *American Economic Review* 67: 39 – 62.

Rosen, S. and P. Taubman. 1982. Changes in Life-Cycle Earnings: What Do Social Security Data Show? *The Journal of Human Resources* 17: 321 – 338.

Rosenzweig, M. R. and K. L. Wolpin. 2000. Natural “Natural Experiments” in Economics. *Journal of Economic Literature* 38: 827 – 874.

Rostker, B. 2006. I Want You: The Evolution of the All-Volunteer Force. *RAND Corporation*.

Schwartz, S. 1986. The Relative Earnings of Vietnam and Korean-Era Veterans. *Industrial and Labor Relations Review* 39: 564 – 572.

Shea, J. 1997. Instrument Relevance in Multivariate Linear Models: A Simple Measure. *The Review of Economics and Statistics* 79: 348 – 352.

Staiger, D. and J. H. Stock. 1997. Instrumental Variables Regression with Weak Instruments. *Econometrica* 65: 557 – 586.

Stanley, M. 2003. College Education and the Midcentury GI Bills. *The Quarterly Journal of Economics* 118 671 –708.

Stiglitz, J. and L. Bilmes. 2008. *The Three Trillion Dollars War,* W.W. Norton & Co.

Stock, J. H. and J. H. Wright. 2000. GMM with Weak Identification. *Econometrica* 68: 1055 – 1096.

Stock, J. H. and M. Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, D. W. K. Andrews and J. H. Stock (eds), Cambridge: Cambridge University Press: 80 – 108.

Teachman, J and V. Call 1996. The Effect of Military Service on Educational, Occupational and Income Attainment. *Social Science Research* 25: 1 – 31.

Teachman, J. and L.M. Tedrow. 2004. Wages, Earnings, and Occupational Status: Did World War II Veterans Receive a Premium? *Social Science Research* 33: 581 – 605.

Teachman, J. 2004. Military Service during the Vietnam Era: Were There Consequences for Subsequent Civilian Earnings? *Social Forces* 83: 709 – 730.

Teachman, J. 2005. Military Service in the Vietnam Era and Educational Attainment. *Sociology of Education* 78: 50 – 68.

Tables and Figures:

Table 1: Descriptive Statistics

Variables	Mean (std.dev.)		
	Overall	Veterans	Not Veterans
log(real wage: 1981$)	6.734 (.502)	6.761 (.476)	6.717 (.517)
Veteran (proportion)	.392 (.488)	-	-
Schooling: Highest year completed	13.49 (2.67)	13.562 (2.150)	13.439 (2.959)
Black	.252 (.434)	.224 (.417)	.270 (.444)
Proportion of men whose wage is from the year:			
1975	.020 (.139)	.021 (.144)	.019 (.135)
1976	.027 (.162)	.023 (.150)	.029 (.169)
1978	.048 (.214)	.053 (.224)	.045 (.208)
1980	.088 (.283)	.090 (.286)	.087 (.281)
1981	.811 (.392)	.809 (.393)	.812 (.391)
Age at which wage is earned	32.354 (2.289)	32.506 (2.186)	32.257 (2.349)
Residence at the age of 14 (South-Atlantic is omitted category)			
Northeast	.040 (.196)	.040 (.196)	.040 (.196)
Mid-Atlantic	.161 (.367)	.150 (.357)	.168 (.374)
East North Central	.186 (.389)	.191 (.393)	.182 (.386)
West North Central	.095 (.294)	.124 (.330)	.077 (.268)
East South Central	.098 (.297)	.089 (.285)	.104 (.305)
West South Central	.115 (.319)	.101 (.301)	.124 (.330)
Pacific	.089 (.285)	.088 (.283)	.090 (.286)

<u>Table 1:</u> <u>Descriptive Statistics (continued)</u>

Variables	Mean (std.dev.)		
	Overall	Veterans	Not Veterans
Type of area in 1966 (Rural is the omitted category)			
Urbanized	.434 (.496)	.452 (.498)	.422 (.494)
Urban place	.165 (.371)	.169 (.375)	.162 (.368)
Instrumental Variables			
Lottery Number	181.566 (103.689)	173.426 (104.446)	186.817 (102.888)
Lottery Ceiling	180.697 (30.134)	184.389 (26.463)	178.315 (32.065)
Proportion with at least one 4 year accredited college in the neighborhood			
Private College	.580 (.494)	.596 (.491)	.569 (.495)
Public College	.481 (.500)	.494 (.500)	.473 (.499)
Total Number of Observations	2754	1080	1674

Table 2(a): Regression Results from Equation (1)[19]

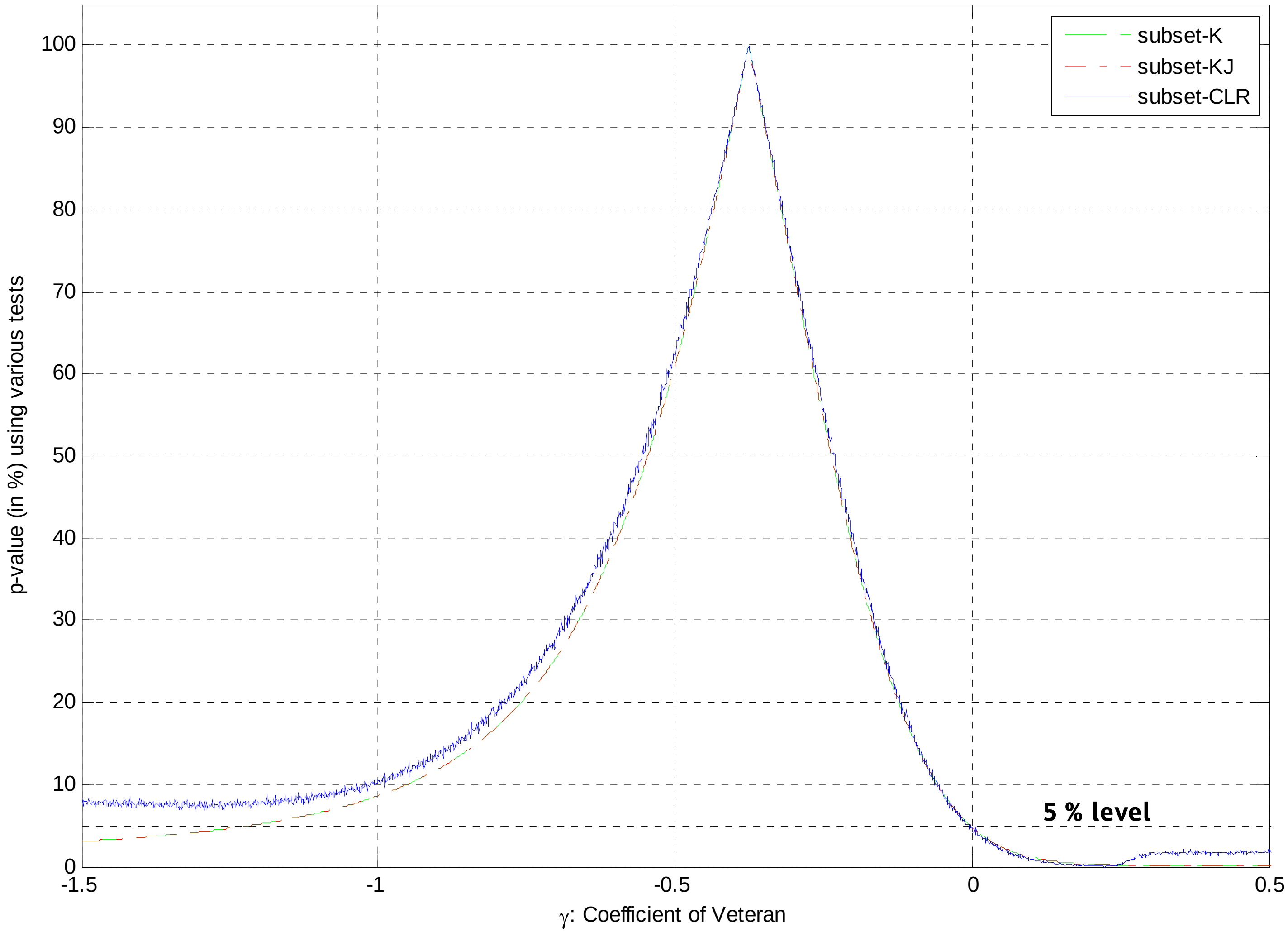
		Specification				
		(A)	(B)	(C)	(D)	(E)
Veteran is treated as		endogenous	endogenous	endogenous	exogenous	exogenous
Schooling is treated as		endogenous	exogenous	excluded	exogenous	excluded
Veteran	Coefficient: γ	-.374*	-.234	-.172	.019	.019
	(SE)	(.222)	(.165)	(.168)	(.018)	(.018)
	Effect: $\delta = e^{\gamma} - 1$	-.312**	-.209	-.158	.019	.019
	(SE)	(.153)	(.131)	(.141)	(.018)	(.018)
Schooling	Coefficient: β	.161**	.049***	-	.049***	-
	(SE)	(.078)	(.003)		(.003)	
p-value from Sargan's test of over-identification		.978	.379	.108	-	-
p-value from the test of endogeneity (low p-value is evidence against exogeneity)	For Veteran	.031	.109	.244	-	-
	For Schooling	.082	-	-	-	-
	For Veteran and Schooling	.055	-	-	-	-
p-value from the generalized Hausman test of whether γ differs between specifications	Compare A	-	.120	.117	.067	.068
	Compare B	-	-	.081	.084	.085
	Compare C	-	-	-	.180	.181
	Compare D	-	-	-	-	.965
Partial R^2 (Shea)	Veteran	.012	.012	.012	-	-
	Schooling	.004	-	-	-	-
F-stat for instruments	Veteran	8.46	8.46	8.46	-	-
	Schooling	2.53	-	-	-	-
Test of Weak Identification: Stock and Yogo (2005)	Cragg-Donald Statistics	1.894	8.47	8.464	-	-
	Bias of IV relative to OLS	more than 30%	between 10% - 20%	between 10% - 20%	-	-
	Size of 5% Wald-test	more than 25%	between 20% - 25%	between 20% - 25%	-	-

[19] Results are based on 2754 observations. Standard errors are reported within parentheses. The symbols *, ** and *** represent significance at the 10 percent, 5 percent and 1 percent level respectively.

Table 3: Testing Exogeneity/Orthogonality restrictions of the instruments

Orthogonality of instruments (tested)	4 year public College	4 year private College	4 year public College	4 year private college	4 year public college 4 year private college
C-statistic (p-value)	.041 (.839)	.043 (.836)	.001 (.979)	.003 (.957)	.044 (.978)
Sargan-statistic (p-value)	.041 (.839)	.043 (.836)	.044 (.978)	.044 (.978)	.044 (.978)
Instruments used in the model	1) 4 year public 2) Lottery number 3) Ceiling in Lottery	1) 4 year private 2) Lottery number 3) Ceiling in Lottery	1) 4 year public 2) 4 year private 3) Lottery number 4) Ceiling in Lottery	1) 4 year public 2) 4 year private 3) Lottery number 4) Ceiling in Lottery	1) 4 year public 2) 4 year private 3) Lottery number 4) Ceiling in Lottery

Figure 1: p-values using weak identification robust tests at different points in the range of the coefficient γ.



Appendix:

Table 2(b): Regression Results for Control Variables from Equation (1)[20]

	Specifications				
	(A)	(B)	(C)	(D)	(E)
Method of estimation	IV	IV	IV	OLS	OLS
Veteran is treated as	endogenous	endogenous	endogenous	exogenous	exogenous
Schooling is treated as	endogenous	exogenous	excluded	exogenous	excluded
Year in which wage is earned[21]	-.027*** (.010)	-.018*** (.007)	-.015** (.007)	-.018*** (.006)	-.014** (.007)
Age at which wage is earned	.029*** (.005)	.030*** (.004)	.030*** (.005)	.027*** (.004)	.028*** (.004)
Black	-.111 (.098)	-.245*** (.025)	-.305*** (.025)	-.231*** (.023)	-.294*** (.023)
Region: Northeast	.059 (.061)	.066 (.050)	.069 (.052)	.079 (.048)	.079 (.050)
Region: Mid-Atlantic	.026 (.061)	.092*** (.033)	.121*** (.034)	.109*** (.030)	.134*** (.031)
Region: East North Central	.074 (.057)	.137*** (.030)	.164*** (.031)	.143*** (.029)	.169*** (.030)
Region: West North Central	-.011 (.070)	.064* (.038)	.098** (.039)	.044 (.034)	.083** (.036)
Region: East South Central	.053 (.042)	.055 (.035)	.056 (.035)	.064* (.033)	.063* (.034)
Region: West South Central	.027 (.043)	.042 (.034)	.048 (.035)	.058* (.031)	.061* (.033)
Region: Pacific	.042 (.071)	.118*** (.038)	.151*** (.039)	.130*** (.036)	.161*** (.038)
Area: Urbanized	.006 (.069)	.097*** (.022)	.138*** (.022)	.085*** (.020)	.128*** (.020)
Area: Urban place	-.030 (.054)	.032 (.026)	.060** (.027)	.026 (.025)	.055** (.026)
Intercept	5.96*** (.769)	6.630*** (.504)	6.929*** (.514)	6.56*** (.485)	6.88*** (.504)

[20] Results are based on 2754 observations. Standard errors are reported within parentheses. The symbols *, ** and *** represent significance at the 10 percent, 5 percent and 1 percent level respectively.
[21] Had this been the variable of interest, once should use dummies to control for the years in which wage is earned to obtain practically meaningful coefficients.