



HAL
open science

Within-Groups Wage Inequality and Schooling: Further Evidence for Portugal

Corrado Andini

► **To cite this version:**

Corrado Andini. Within-Groups Wage Inequality and Schooling: Further Evidence for Portugal. Applied Economics, 2008, 42 (28), pp.3685. 10.1080/00036840802314564 . hal-00582253

HAL Id: hal-00582253

<https://hal.science/hal-00582253>

Submitted on 1 Apr 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Within-Groups Wage Inequality and Schooling: Further Evidence for Portugal

| | |
|-------------------------------|--|
| Journal: | <i>Applied Economics</i> |
| Manuscript ID: | APE-07-0469.R1 |
| Journal Selection: | Applied Economics |
| Date Submitted by the Author: | 16-Jun-2008 |
| Complete List of Authors: | Andini, Corrado; University of Madeira, Department of Management and Economics |
| JEL Code: | I21 - Analysis of Education < I2 - Education < I - Health, Education, and Welfare, J31 - Wage Level, Structure; Differentials by Skill, Occupation, etc. < J3 - Wages, Compensation, and Labor Costs < J - Labor and Demographic Economics, C29 - Other < C2 - Econometric Methods: Single Equation Models < C - Mathematical and Quantitative Methods |
| Keywords: | Endogeneity, Quantile Regression, Schooling, Wage Inequality |
| | |



Within-Groups Wage Inequality and Schooling: Further Evidence for Portugal

Corrado Andini *

University of Madeira, CEEAplA and IZA

ABSTRACT

This paper provides further evidence on the positive impact of schooling on within-groups wage dispersion in Portugal, using data on male workers from the 2001 wave of the European Community Household Panel. The issue of schooling endogeneity is taken into account by using the newest available instrumental-variable technique for quantile regression, i.e. the control-function estimator due to Lee (2007). The findings are compared with earlier results based on different techniques, i.e. the instrumental-variable estimator due to Arias, Hallock and Sosa-Escudero (2001) and the standard exogeneity-based estimator due to Koenker and Bassett (1978).

Keywords: Endogeneity, Quantile Regression, Schooling, Wage Inequality.
JEL Classification: I21, J31, C29.

* Please send correspondence to: Corrado Andini, Universidade da Madeira, Campus da Penteada, 9000-390 Funchal, Portugal. Tel.: +351 291 70 50 53, Fax: +351 291 70 50 40, Email: andini@uma.pt, and Website: <http://www.uma.pt/andini>.

I gratefully acknowledge financial support from the European Commission (EDWIN Project, HPSE-CT-2002-00108). For valuable comments, I would like to thank Santiago Budría, Joop Hartog, Pedro Telhado Pereira, Robert Wright and the participants at the 22nd National Conference of the Italian Association of Labour Economists. The usual disclaimer applies.

I. INTRODUCTION

Many authors agree on the argument that schooling has a positive impact on within-groups wage inequality in Portugal¹. The evidence is generally based on the standard quantile-regression techniques due to Koenker and Bassett (1978)², applied to several types of wage equations. As a matter of example, Hartog, Pereira and Vieira (2001), Machado and Mata (2001) as well as Martins and Pereira (2004) use different specifications of a Mincerian model and find that the return to schooling is increasing along the conditional earnings distribution. Particularly, the return at the ninth decile seems to be significantly higher than the return at the first decile.

Apart from a recent contribution by Andini (2008), a common feature of the quantile-regression studies on schooling and within-groups wage inequality in Portugal is that no one deals with the endogeneity of schooling which is, typically, a relevant issue in studies focusing on the impact of schooling on the mean of the conditional earnings distribution.

Likewise the ordinary-least-squares estimator, the estimator of Koenker and Bassett (1978) assumes residuals' orthogonality³. Therefore, disregarding the endogeneity of

¹ This section partly borrows from Andini (2008).

² Min and Kim (2004) provide a Monte Carlo comparison between the Koenker-Bassett parametric estimator and the nonparametric approach due to Yu and Jones (1998). The results indicate that the nonparametric quantile-regression method is more appropriate, particularly when the underlying model is nonlinear or the error term follows a non-normal distribution.

³ Consider the regression model $Y = \beta X + \varepsilon$. The estimator of Koenker and Bassett is consistent at each quantile θ of the conditional distribution of Y if the orthogonality condition holds at each quantile, i.e. $\text{Quant}_{\theta}(\varepsilon_{\theta}|X) = 0 \quad \forall \theta$, while the consistency of the ordinary-least-squares estimator requires that the conditional mean of regression residuals is null, i.e. $E(\varepsilon|X) = 0$.

1
2
3
4 schooling may imply inconsistent estimates of its coefficient along the conditional wage
5
6 distribution, i.e. the estimated impact of schooling on within-groups earnings dispersion
7
8 may be misleading.
9

10
11 A further common feature of existing studies is that no one deals with the total impact
12
13 of schooling on within-groups wage inequality in Portugal, meaning that previously-
14
15 estimated wage equations do not exclude schooling-dependent covariates, such as
16
17 industry dummies⁴ and labour-market experience⁵, from the list of regressors.
18

19
20 To be more precise about the meaning and the relevance of the concept of total return to
21
22 schooling, let us present the following simple example. First, let us label the logarithm
23
24 of hourly earnings as $\ln w$, the number of schooling years as s , the so-called potential
25
26 labour-market experience as $\text{exp} = \text{age} - s - 6$. Second, let us suppose that a wage
27
28 equation includes potential labour-market experience as control-variable, i.e.
29
30 $\ln w = \beta_0 + \beta_1 s + \beta_2 \text{exp} + \varepsilon$.
31
32
33

34
35 In the latter case, the coefficient of schooling years β_1 does not capture the total effect
36
37 of schooling on earnings for the simple reason that the total effect is given by
38
39

$$\frac{\partial \ln w}{\partial s} = \beta_1 - \beta_2.$$

40
41
42 Although in our simple example one can recover the total effect of schooling by
43
44 subtracting the estimate of β_2 from the estimate of β_1 , it is worth stressing that, when
45
46
47
48

49
50
51 ⁴ Jobs in some industries may require more years of schooling than jobs in other industries.

52
53 ⁵ Pereira and Martins (2004) properly argue that in order “to obtain the full impact of education on wages,
54
55 one should be careful not to include in the wage equation covariates whose value can depend on
56
57 education. In the extreme case one should only regress the $\ln(\text{wage})$ in education.” (p. 526). See also
58
59 Andini (2007).
60

1
2
3
4 the mathematical law of the schooling-dependent covariate is unknown (i.e. the general
5
6 case), the recovering exercise becomes impossible. Hence, if the objective of the
7
8 empirical analysis is the total return to schooling, it is important to exclude schooling-
9
10 dependent covariates from the wage equation.
11

12
13 Following Andini (2008), this paper tackles a controversial issue. On the one hand, if
14
15 the potential correlation between residuals and schooling is limited by the insertion in
16
17 the wage equation of a large set of control-variables, standard quantile-regression
18
19 techniques are likely to consistently estimate the coefficient of schooling along the
20
21 conditional wage distribution but unlikely to recover the total returns to schooling due
22
23 to the likely presence of schooling-dependent covariates among the control-variables.
24
25 On the other hand, if the set of control-variables excludes schooling-dependent
26
27 covariates and only includes strictly exogenous regressors, then standard techniques
28
29 may poorly estimate the total impact of schooling on within-groups wage inequality due
30
31 to the likely correlation between residuals and schooling. This paper aims at presenting
32
33 *consistent* estimates the *total* return to schooling along the conditional distribution of
34
35 Portuguese wages and compares the results with previous findings by Andini (2008).
36
37
38
39
40
41
42
43

44 **II. EMPIRICAL MODEL**

45
46 Andini (2008) finds that the difference between the total return at the ninth decile and
47
48 the total return at the first decile of the conditional earnings distribution in Portugal is
49
50 likely to range between 4.2% estimated using the standard quantile-regression
51
52 techniques due to Koenker and Bassett (1978) and 26.2% estimated using the
53
54 instrumental-variable approach due to Arias, Hallock and Sosa-Escudero (2001). In this
55
56 paper, we provide further evidence on the total impact of schooling on within-groups
57
58
59
60

1
2
3
4 wage inequality using the latest available instrumental-variable technique for quantile
5 regression, i.e. the control-function estimator due to Lee (2007).
6
7

8
9 To explain the rationale behind our estimation procedure, let us suppose that a variable
10 Y is explained by an exogenous variable X and an endogenous variable S. Further, let
11 us assume that there is an interest in estimating the impact of S (not only on the mean
12 but also) on the shape of the conditional distribution of Y, controlling for X and using
13 quantile-regression techniques. The control-function approach due to Lee (2007) is a
14 two-stage approach similar to the one adopted by Arias, Hallock and Sosa-Escudero
15 (2001, AHS henceforth). The first-stage regression consists of an ordinary-least-squares
16 regression of S on X and instrumental variables Z1, Z2, and so on. The difference
17 between the estimator of AHS and the estimator of Lee is related to the second stage.
18 The method of AHS replaces the actual values of S with first-stage predicted values of
19 S. In contrast, Lee adds a polynomial function of first-stage residuals to the set of the
20 second-stage regressors, formed by the actual values of S and X.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 Both approaches aim at providing consistent estimates of the coefficient of the
37 endogenous explanatory variable at several quantiles of the conditional distribution of
38 the dependent variable, using the estimator of Koenker and Bassett (1978, KB
39 henceforth) in the second stage. And, both approaches suffer from the typical loss of
40 efficiency associated with the implementation of instrumental-variable techniques in
41 finite samples, but the method of Lee turns out to be *more efficient* than the method of
42 AHS, in our specific application, at both the first and the ninth decile of the conditional
43 wage distribution. Note that the latter point has special relevance for the issue of
44 recovering a reliable measure of the impact of schooling on within-groups wage
45 inequality.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Based on the empirical specification proposed by Andini (2008), the estimation procedure of this paper is as follows:

$$(1) \quad \ln w_i = \nu_\theta + \beta_\theta s_i + \sum_j \delta_{\theta j} yb_{ji} + \sum_c \lambda_{\theta c} \hat{\eta}_i^c + \mu_{\theta i}$$

$$s_i = \nu + \sum_j \alpha_j yb_{ji} + \sum_k \phi_k qb_{ki} + \eta_i$$

$$\hat{s}_i = E(s_i | yb_{ji}, qb_{ki})$$

$$\hat{\eta}_i = s_i - \hat{s}_i$$

$$\text{with } \text{Quant}_\theta(\mu_{\theta i} | s_i, yb_{ji}, \hat{\eta}_i) = 0 \text{ for each } \theta.$$

Our notation is as follows: letter i refers to the i -th individual in the sample, $\ln w$ stands for the logarithm of gross hourly earnings, s measures years of schooling, yb is an indicator-variable for year of birth with j going from 1937 to 1984 (year 1936 is the excluded category), qb is an indicator-variable for quarter of birth with k going from 1 to 3 (quarter 4 is the excluded category), θ is a quantile-indicator going from 5 to 95, c represents the order of the control-function.

Data are extracted from the latest available wave of the European Community Household Panel (ECHP), the 2001 wave, and are related to Portuguese male workers.

Summary sample statistics are reported in Table 1.

III. RESULTS

First-stage regression results are presented in the Appendix. Specifically, as already discussed by Andini (2008), the use of a full set of quarters of birth as instrumental variables is not entirely satisfactory because the F-test of excluded instruments does not

1
2
3
4 reject the null (p-value 0.1972; this result seems driven by the third quarter). It is also
5
6 known, however, that passing the F-test should not be intended as a strict requirement
7
8 due to the limitations of the test itself, i.e. low power (see Cruz and Moreira, 2005).
9
10 This is particularly true when the Sargan test of over-identification is passed (p-value
11
12 0.1351), i.e. the model specification is not rejected. In addition, our results are
13
14 consistent with the findings of Angrist and Krueger (1991) who argue that people born
15
16 later in the year have slightly more schooling than people born earlier. Note, indeed,
17
18 that the estimated coefficients for the first, the second and the third quarter of birth are
19
20 negative and that the excluded category is the fourth quarter of birth. Further, robustness
21
22 checks highlight that, if the model is just-identified using the fourth quarter of birth
23
24 only, the estimated coefficient for the last quarter of the year has the expected positive
25
26 sign and is statistically significant (p-value 0.055). Finally, if only the last two quarters
27
28 of the year are used as instruments, the estimated coefficients are both positive and the
29
30 F-test of excluded instruments is passed (p-value 0.096).
31
32
33
34
35

36
37 It is worth stressing that our estimation results for the wage equation are robust to the
38
39 identification strategy. As a matter of example, the AHS method provides roughly the
40
41 same results when one uses the fourth quarter of birth as single instrument for schooling
42
43 years rather than a full set of indicator-variables for quarters of birth as in the Appendix.
44
45 Hence, for sake of comparison with our previous estimates of the total return to
46
47 schooling along the conditional wage distribution, we adopt the same identification
48
49 strategy as implemented by Andini (2008).
50
51
52

53
54 Figure 1 plots the estimates of the main parameter of interest, i.e. the estimates of β in
55
56 model (1) which are, in turn, reported in Table 2. Note that the last two columns of
57
58 Table 2 also report the AHS results and the standard KB results presented by Andini
59
60 (2008).

1
2
3
4 Regarding the choice of the order of the control-function, Lee (2004) does not suggest
5
6 the existence of an optimal order because Monte-Carlo experiments in finite samples
7
8 show that “the estimator is not very sensitive to the choice of the order of the
9
10 polynomial approximations” (p. 15). Indeed, our empirical findings based on a number
11
12 of different orders of the control-function are consistent with the simulation results. As
13
14 a matter of example, we report estimates of β based on c going from 1 to 4. The
15
16 corresponding estimates of the total return to schooling along the conditional wage
17
18 distribution are labelled as L1, L2, L3, L4.
19
20
21

22
23 Depending on the order of the control-function, the total impact of schooling on within-
24
25 groups wage inequality varies from 5.1% to 7.4% and fits the interval provided by
26
27 Andini (2008), who refers to the difference between the return at the ninth decile and
28
29 the return at the first decile of the conditional earnings distribution. Therefore, one
30
31 contribution of this paper consists of providing evidence on a smaller interval of impact
32
33 than previously estimated, i.e. 5.1%-7.4% vs. 4.2%-26.2%.
34
35

36
37 Further, note that the standard KB techniques disregarding the endogeneity issue
38
39 suggest a 4.1% gap and therefore underestimate the total impact of schooling on within-
40
41 groups wage inequality in Portugal, as previously suggested by Andini (2008), although
42
43 the magnitude of the downward bias is smaller than one predicted by the AHS method.
44
45

46
47 Remarkably, the estimator of Lee predicts a pattern of the schooling coefficient, along
48
49 the conditional wage distribution, that looks very much like the one predicted by the
50
51 method of AHS, with a peak around the eight decile and a drop around the second
52
53 decile. Hence, the impact of schooling on within-groups wage inequality is higher when
54
55 measured as difference between the 25th quantile and the 75th quantile, ranging from
56
57 7.6% to 12.2%. Again, this result is not captured by the standard quantile-regression
58
59 estimator that only highlights a 3.0% gap.
60

1
2
3
4 In contrast, the standard exogeneity-based techniques seem to perform relatively fine
5
6 when considering the conditional average return to schooling. The ordinary-least-
7
8 squares estimator prospects a 5.7% coefficient, which lies within the interval of 4.7%-
9
10 7.5% obtained using the control-function approach and the AHS method. The
11
12 conclusion is that schooling has a positive impact on *between-groups* wage inequality in
13
14 Portugal. The latter, however, is a well-known result and goes beyond the objective of
15
16 this paper.
17
18
19
20
21
22

23 **IV. FINAL REMARKS**

24
25 This paper uses the control-function estimator for quantile regression due to Lee (2007)
26
27 and provides further evidence of the total impact of schooling of within-groups wage
28
29 inequality in Portugal. In our specific application, the estimator of Lee turns out to be
30
31 more efficient than the estimator of Arias, Hallock and Sosa-Escudero (2001) at two
32
33 key-deciles of the conditional wage distribution, thus providing more reliable measures
34
35 of the corresponding total returns to schooling. Specifically, we find that standard
36
37 exogeneity-based estimator of Koenker and Bassett (1978) underestimates the impact of
38
39 schooling on earnings dispersion through its within-groups dimension, although the
40
41 magnitude of the bias is likely to be less than previously suggested.
42
43
44
45

46
47 The empirical research on the positive association between schooling and within-groups
48
49 wage dispersion in Portugal is relatively rich in contributions. The finding of a positive
50
51 association is also supported by several studies using the KB estimator with education
52
53 levels rather than schooling years. The striking evidence that schooling/education is a
54
55 strong source of the so-called residual earnings inequality should inspire an effort of
56
57 *understanding the reasons* behind this stylized fact.
58
59
60

1
2
3
4 Some authors suggest that one possible explanation of the above-referred fact has to do
5
6 with educational mismatches in the labour market. However, the existing empirical
7
8 evidence is not necessarily consistent with this argument (see Budría, 2006). Our
9
10 perception of the problem is that future research should pay more attention to the issue
11
12 of school quality. The existing differences in the quality of Portuguese universities and
13
14 secondary schools, which are well-known to everybody living in Portugal (although the
15
16 academic research in this field is relatively scarce⁶), are likely to play an important role
17
18 in explaining the stylized fact that the wage returns for individuals with the same
19
20 number of schooling years or the same level of education (and the same observed
21
22 characteristics) are quite heterogeneous. However, an empirical evaluation of the latter
23
24 hypothesis needs individual-level data on school quality, which are not currently
25
26 available.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

56
57 ⁶ Oliveira and Santos (2005) find evidence of heterogeneity in efficiency among secondary schools in
58
59 Portugal. To the best of our knowledge, a more comprehensive study on school quality in Portugal is not
60
available.

REFERENCES

- Andini, C. (2008) The Total Impact of Schooling on Within-Groups Wage Inequality in Portugal, Applied Economics Letters, 15(2), 85-90.
- Andini, C. (2007) Returns to Education and Wage Equations: a Dynamic Approach, Applied Economics Letters, 14(8), 577-579.
- Angrist, J.D. and Krueger, A.B. (1991) Does Compulsory Schooling Attendance Affect Schooling and Earnings?, Quarterly Journal of Economics, 106(4), 979-1014.
- Arias, O., Hallock, K.F. and Sosa-Escudero, W. (2001) Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data, Empirical Economics, 26(1), 7-40.
- Budría, S. (2006) Can Over-Education Account for the Positive Association between Education and Within-Groups Wage Inequality? A Note, MPRA Working Papers, n° 92, Munich Personal RePEc Archive, October.
- Hartog, J., Pereira, P.T. and Vieira, J.A.C. (2001) Changing Returns to Education in Portugal during the 1980s and Early 1990s: OLS and Quantile Regression Estimators, Applied Economics, 33(8), 1021-1037.
- Koenker, R. and Bassett, G. (1978) Regression Quantiles, Econometrica, 46(1), 33-50.
- Lee, S. (2004) Endogeneity in Quantile Regression Models: A Control Function Approach, CEMMAP Working Papers, n° CWP08/04, Centre for Microdata Methods and Practice, December.
- Lee, S. (2007) Endogeneity in Quantile Regression Models: A Control Function Approach, Journal of Econometrics, 141(2), 1131-1158.
- Machado, J.A.F. and Mata, L. (2001) Earnings Functions in Portugal 1982-1994: Evidence from Quantile Regressions, Empirical Economics, 26(1), 115-134.
- Martins, P.S. and Pereira, P.T. (2004) Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries, Labour Economics, 11(3), 355-371.
- Min, I. and Kim, I. (2004) A Monte Carlo Comparison of Parametric and Nonparametric Quantile Regressions, Applied Economics Letters, 11(2), 71-74.
- Oliveira, M.A. and Santos, C. (2005) Assessing School Efficiency in Portugal using FDH and Bootstrapping, Applied Economics, 37(8), 957-968.
- Pereira, P.T. and Martins, P.S. (2004) Returns to Education and Wage Equations, Applied Economics, 36(6), 525-531.
- Yu, K. and Jones, M. (1998) Local Linear Quantile Regression, Journal of the American Statistical Association, 93(441), 228-237.

Table 1. Summary sample statistics

| Variable | Obs. | Mean | S.E. | Min | Max |
|--------------------------------|------|------|------|------|------|
| Logarithm of gross hourly wage | 1782 | 6.55 | 0.47 | 3.32 | 8.49 |
| Schooling years | 1782 | 8.80 | 3.91 | 3.00 | 27.0 |
| Year of birth | 1782 | 1965 | 11.5 | 1936 | 1984 |
| Quarter of birth | 1782 | 2.45 | 1.11 | 1.00 | 4.00 |

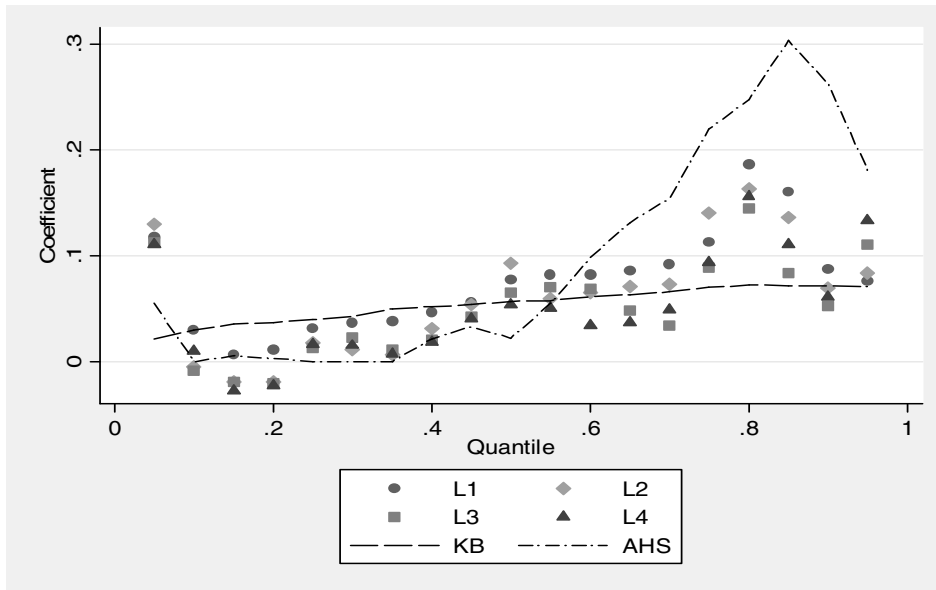
For Peer Review

Table 2. Conditional Returns to Schooling

| Quantile | L1 | L2 | L3 | L4 | KB | AHS |
|----------|--------------------|---------------------|---------------------|---------------------|--------------------|--------------------|
| 5 | 0.1175 (0.0869) | 0.1296 (0.0933) | 0.1124 (0.0859) | 0.1103 (0.0909) | 0.0211 (0.0052) | 0.0550 (0.0828) |
| 10 | 0.0300 (0.0631) | -0.0050 (0.0674) | -0.0084 (0.0611) | 0.0096 (0.0728) | 0.0298 (0.0033) | 0.0000 (0.0782) |
| 15 | 0.0067 (0.0452) | -0.0195 (0.0476) | -0.0195 (0.0455) | -0.0279 (0.0551) | 0.0352 (0.0025) | 0.0054 (0.0536) |
| 20 | 0.0111 (0.0444) | -0.0194 (0.0516) | -0.0207 (0.0633) | -0.0228 (0.0533) | 0.0368 (0.0019) | 0.0028 (0.0376) |
| 25 | 0.0313 (0.0335) | 0.0177 (0.0369) | 0.0128 (0.0324) | 0.0164 (0.0494) | 0.0396 (0.0017) | 0.0000 (0.0155) |
| 30 | 0.0367 (0.0378) | 0.0115 (0.0373) | 0.0225 (0.0300) | 0.0152 (0.0421) | 0.0427 (0.0023) | 0.0000 (0.0151) |
| 35 | 0.0381 (0.0311) | 0.0066 (0.0357) | 0.0116 (0.0462) | 0.0069 (0.0400) | 0.0497 (0.0018) | 0.0000 (0.0108) |
| 40 | 0.0467 (0.0304) | 0.0314 (0.0371) | 0.0205 (0.0302) | 0.0184 (0.0310) | 0.0518 (0.0020) | 0.0214 (0.0526) |
| 45 | 0.0559 (0.0335) | 0.0540 (0.0289) | 0.0423 (0.0385) | 0.0401 (0.0391) | 0.0537 (0.0018) | 0.0329 (0.0329) |
| 50 | 0.0770 (0.0668) | 0.0925 (0.0389) | 0.0651 (0.0423) | 0.0535 (0.0501) | 0.0570 (0.0032) | 0.0221 (0.0398) |
| 55 | 0.0819 (0.0445) | 0.0594 (0.0472) | 0.0699 (0.0469) | 0.0501 (0.0563) | 0.0577 (0.0026) | 0.0549 (0.0613) |
| 60 | 0.0822 (0.0509) | 0.0651 (0.0547) | 0.0685 (0.0436) | 0.0340 (0.0517) | 0.0614 (0.0031) | 0.0977 (0.0473) |
| 65 | 0.0855 (0.0440) | 0.0711 (0.0523) | 0.0479 (0.0471) | 0.0371 (0.0522) | 0.0631 (0.0030) | 0.1312 (0.0489) |
| 70 | 0.0918 (0.0542) | 0.0727 (0.0650) | 0.0338 (0.0643) | 0.0488 (0.0647) | 0.0658 (0.0029) | 0.1541 (0.0567) |
| 75 | 0.1129 (0.0675) | 0.1401 (0.0414) | 0.0887 (0.0963) | 0.0934 (0.0632) | 0.0698 (0.0036) | 0.2195 (0.0397) |
| 80 | 0.1861 (0.0551) | 0.1626 (0.0657) | 0.1445 (0.0702) | 0.1556 (0.0589) | 0.0723 (0.0038) | 0.2473 (0.0729) |
| 85 | 0.1600 (0.0642) | 0.1360 (0.0576) | 0.0836 (0.0795) | 0.1103 (0.0797) | 0.0716 (0.0031) | 0.3029 (0.0795) |
| 90 | 0.0872 (0.0813) | 0.0695 (0.0735) | 0.0521 (0.0719) | 0.0606 (0.0680) | 0.0717 (0.0042) | 0.2624 (0.1123) |
| 95 | 0.0756 (0.0990) | 0.0834 (0.1046) | 0.1103 (0.1060) | 0.1329 (0.1105) | 0.0706 (0.0062) | 0.1808 (0.1073) |
| Mean | 0.0751 (0.0447) | 0.0643 (0.0447) | 0.0472 (0.0446) | 0.0479 (0.0445) | 0.0577 (0.0029) | 0.0751 (0.0534) |
| 90-10 | 0.0572 | 0.0745 | 0.0605 | 0.0510 | 0.0419 | 0.2624 |
| 75-25 | 0.0816 | 0.1225 | 0.0759 | 0.0770 | 0.0302 | 0.2195 |

Standard errors in parentheses

Figure 1. Conditional Returns to Schooling



Peer Review

Appendix. First-stage regression of schooling years

| | Coeff. | Robust S.E. | t | P-value |
|------------------|---------|----------------|-------|---------|
| Quarter of birth | | | | |
| 1 | -0.5303 | 0.2733 | -1.94 | 0.053 |
| 2 | -0.5162 | 0.2780 | -1.86 | 0.064 |
| 3 | -0.3004 | 0.2675 | -1.12 | 0.262 |
| Year of birth | | | | |
| 1937 | -0.1968 | 0.4193 | -0.47 | 0.639 |
| 1938 | 0.4486 | 0.8273 | 0.54 | 0.588 |
| 1939 | 0.1668 | 1.1782 | 0.14 | 0.887 |
| 1940 | -0.2851 | 0.3714 | -0.77 | 0.443 |
| 1941 | 1.2821 | 1.2007 | 1.07 | 0.286 |
| 1942 | 0.0151 | 0.7509 | 0.02 | 0.984 |
| 1943 | 2.6175 | 1.3599 | 1.92 | 0.054 |
| 1944 | 0.4641 | 1.3862 | 0.33 | 0.738 |
| 1945 | 3.5350 | 1.1231 | 3.15 | 0.002 |
| 1946 | 1.7825 | 0.9022 | 1.98 | 0.048 |
| 1947 | 0.1080 | 0.5157 | 0.21 | 0.834 |
| 1948 | 2.4706 | 1.2484 | 1.98 | 0.048 |
| 1949 | 1.3233 | 0.7644 | 1.73 | 0.084 |
| 1950 | 2.5171 | 0.9047 | 2.78 | 0.005 |
| 1951 | 1.4911 | 0.6433 | 2.32 | 0.021 |
| 1952 | 1.2574 | 0.6183 | 2.03 | 0.042 |
| 1953 | 1.2702 | 0.4896 | 2.59 | 0.010 |
| 1954 | 1.5278 | 0.8376 | 1.82 | 0.068 |
| 1955 | 0.5447 | 0.4065 | 1.34 | 0.180 |
| 1956 | 1.6778 | 0.6129 | 2.74 | 0.006 |
| 1957 | 1.6236 | 0.6541 | 2.48 | 0.013 |
| 1958 | 1.4677 | 0.5944 | 2.47 | 0.014 |
| 1959 | 0.6664 | 0.4798 | 1.39 | 0.165 |
| 1960 | 1.1161 | 0.4322 | 2.58 | 0.010 |
| 1961 | 1.6299 | 0.6509 | 2.50 | 0.012 |
| 1962 | 2.3032 | 0.5585 | 4.12 | 0.000 |
| 1963 | 2.2418 | 0.5484 | 4.09 | 0.000 |
| 1964 | 1.4711 | 0.5304 | 2.77 | 0.006 |
| 1965 | 3.2459 | 0.7877 | 4.12 | 0.000 |
| 1966 | 3.1548 | 0.8257 | 3.82 | 0.000 |
| 1967 | 3.8398 | 0.9662 | 3.97 | 0.000 |
| 1968 | 3.0704 | 0.6384 | 4.81 | 0.000 |
| 1969 | 3.1249 | 0.5733 | 5.45 | 0.000 |
| 1970 | 3.7298 | 0.6285 | 5.93 | 0.000 |
| 1971 | 3.8072 | 0.4555 | 8.36 | 0.000 |
| 1972 | 3.9569 | 0.5066 | 7.81 | 0.000 |
| 1973 | 3.1319 | 0.4842 | 6.47 | 0.000 |
| 1974 | 4.3609 | 0.5063 | 8.61 | 0.000 |
| 1975 | 4.4149 | 0.5147 | 8.58 | 0.000 |
| 1976 | 3.6441 | 0.4565 | 7.98 | 0.000 |
| 1977 | 3.9069 | 0.4371 | 8.94 | 0.000 |
| 1978 | 3.5783 | 0.4024 | 8.89 | 0.000 |
| 1979 | 3.8873 | 0.4110 | 9.46 | 0.000 |
| 1980 | 3.9421 | 0.3377 | 11.67 | 0.000 |
| 1981 | 4.0952 | 0.2941 | 13.92 | 0.000 |
| 1982 | 2.7029 | 0.4154 | 6.51 | 0.000 |
| 1983 | 2.6506 | 0.3425 | 7.74 | 0.000 |
| 1984 | 2.6250 | 0.4438 | 5.91 | 0.000 |
| Constant | 6.4083 | 0.2554 | 25.09 | 0.000 |

Sargan test of over-identifying restrictions

0.1351

F-test of excluded instruments

0.1972