DIW Berlin
German Institute
for Economic Research

Discussion Papers

417

Hans J. Baumgartner

Are There Any Class Size Effects On Early Career Earnings In West Germany?

Berlin, April 2004

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

DIW Berlin

German Institute
for Economic Research

Königin-Luise-Str. 5
14195 Berlin,
Germany

Phone +49-30-897 89-0
Fax +49-30-897 89-200

www.diw.de

ISSN 1619-4535

ARE THERE ANY CLASS SIZE EFFECTS ON EARLY CAREER EARNINGS IN WEST GERMANY?

Hans J. Baumgartner)*

DIW Berlin
31 March 2004

Abstract:

The paper analyses the effect of class size, which stands proxy for school quality, on early career earnings. Using confidential district level information from the German Socio-Economic Panel (GSOEP), it is demonstrated that class size has no discernible effect on early career earnings. This finding is robust to changes in specification and the choice of sub-samples. The economic literature focused so far mainly on the US and the UK. This paper confirms this literature and adds new evidence to the growing empirical literature on this issue from Germany.

JEL Classification: I21, J31, C52
Keywords: school quality, class size, labour market outcomes,
Corresponding address: Hans J. Baumgartner
DIW Berlin
Königin-Luise-Str. 5
14195 Berlin
Germany
phone: +49/(0)30/89789-307
fax: +49/(0)30/89789-114
Email: hbaumgartner@diw.de

---

*) I am grateful to Viktor Steiner for supervising my Ph.D. research. I owe also thanks to Wiji Arulampalam, Thomas K. Bauer, Tilman Brück, Felix Büchel, Boriss Siliverstovs, Jeffrey A. Smith and participants at the 15th EALE Annual Conference 2003, the 14th EEA/IFS Euro Summer School on Microeconometrics, the 6th IZA European Summer School in Labour Economics, and the 1st DIW-Ph.D.-Seminar for helpful discussions and comments. I also appreciate the useful comments of an anonymous referee. The usual disclaimer applies.

Note on data availability: The data used in this paper are available to any researcher. The data on class size are collected from the state offices for statistics (*Landesämter für Statistik*). The German data protection law restricts the regional GSOEP micro data to be analysed on-site at the DIW Berlin. For more information on how to obtain the regionally disaggregated GSOEP see www.diw.de/english/sop/faq/index.html (item 2.13).

1. Introduction

German public spending on education sums up to some 70.5 billion Euro in 2001, which accounts for 13.8 percent of the combined federal, state, and district budget and 3.4 percent of GDP.[1] This massive public investment in education raises the question whether it is spent effectively and what its impact is on the targeted pupils. This question was already publicly discussed before PISA unearthed a relatively poor performance of German pupils and accelerated ever since. The issue is on the top of the political agenda and also vitally debated in academia.

This paper adds evidence to this debate in examining the effects of the quality of schooling on early career earnings. Whether an increased public spending to finance smaller classes is a good investment is evaluated by an education production function approach. Input is measured by average class size in the residential district where schooling was completed. The outcome measure is earnings in the second year in the job. This approach differs from another research branch on the effects of school quality, which is based on in-school performance as an outcome variable.[2]

From an economist's point of view, however, it is the labour market that matters in gauging the effectiveness of schooling.[3] An "*objective 'market test'*", as Card and Krueger (1996b) name it, is met by schools that increase their students subsequent earnings. Applying such a test, Card and Krueger find in a series of papers (1992a, 1992b, 1996c, 1996b, 1996a) supporting evidence for a positive quality-earning relation. Applying different measures for

[1] Statistisches Bundesamt Deutschland 2002: Indikatoren für Bildung.

[2] Card (1998), Dolton and Vignoles (2000), Sander (1993), Barro and Lee (2000), Wilson (2000), Winter-Ebmer and Wirz (2002), and West, et al. (2001) find that available financial resources have a positive impact on in-school performance. But it is important how the financial resources are utilised. Wößmann (2000, 2001) finds that in-school performance is not associated with available resources. The institutional setting of schooling is of much more importance. This is confirmed by Angrist and Lavy (2002), who find that computer-aided instructions are not associated with higher test scores. This literature is surveyed by Hanushek (2002).

[3] I find the literature on the quality-earning link to be much less voluminous than the literature on the quantity-earning relations. The quantity-earning relation is reviews by Ashenfelter, et al. (1999), Blundell, et al. (1999) Card (1995, 1999), and Psacharopoulos (1994). In a recent publication Harmon, et al. (2001) collect evidence on

the quality of schooling – such as pupil-teacher-ratio, average term length, relative teacher pay – they find that the return to an additional year of schooling is higher for those men who are educated in states that provide a higher quality of schooling.

Their result, however, is challenged by Betts (1995) who finds that traditional measures of school quality fail to explain subsequent earnings. He speculates that structural changes may have weakened a quality-earning relation in the US. Card and Krueger (1992a, 1992b) analyse a cohort that enjoyed schooling before the 1960s, while Betts' (1995) cohort went to school there after. The variation in U.S. school quality, however, converged in recent years, which may explain why studies focusing on a younger cohort are not able to find a significant relations between quality and earnings. If school quality has diminishing effects, and if school quality improved over time, then may current studies be largely on the flatter part of the production function, where the variation in quality induces less variation in earnings then compared to the steeper part. Note that Card and Krueger observe rages from an earlier period where quality had stronger effects.[4]

Another explanation for the different findings is shown by Betts (1996, figure 6-1). Surveying 24 articles he finds that studies who measure quality at the school level find insignificant effects of quality on earnings, while those studies that measure schooling on a more aggregate level are rather likely to find a significant relation. A reason for this aggregation bias could be the omission of important state differences in school policy (Hanushek, Rivkin and Taylor, 1996).

Heckman, et al. (1995, 1996) examine directly Card and Krueger's conclusions and find that it is sensitive to the crucial assumption that migration is random and not based on

---

the quantity-earning relation for a series of European countries, which also includes a section on Germany (Lauer and Steiner, 2001).

[4] This argument is based on the assumption that on higher levels of quality the marginal productivity of an improved quality may be significantly less than at lower levels of quality. This hypothesis, however, is challenged since developing countries, which may be rather on the steeper part of the production function, lack significantly stronger quality effects (Hanushek, 1995).

differential earning opportunities. Card and Krueger's strategy in estimating the effect of a particular school input is to compare the earnings of men who received their schooling in a different state than they currently live in. If migration is self-selecting, the results might be biased. It is unclear, notes Burtless (1996), whether non-random migration biases Card and Krueger's quality effects by a large amount.

The focus of the quality-earning literature is clearly on the US (Hanushek, 2002) but recent studies draw conclusions also from UK data. These studies use all the same data source, the NCDS, which is a longitudinal survey of all British citizens born in one certain week in 1958. One of the advantages of this data is that it provides also standardised test scores at the age of seven and eleven that can stand proxy for unobserved ability. Dolton and Vignoles (2000) find that although the quality of schooling has a small positive impact on student attainment, there is no measurable relation between the quality of schooling and subsequent earnings. This finding is generally confirmed by Dearden, et al. (2000), who find that the quality of schooling has only an impact on women's wages with low ability and by Harmon and Walker (2000), who cannot find a significant quality-earning relation either.

This paper addresses the quality-earning issue for the German example, where this matter has not been addressed yet in the economic literature. Although most studies using foreign data cannot confirm a quality-earning relation, the working hypothesis is that there could still be such a link in Germany, since the US-American and the British schooling system differs in many aspects from the schooling system in Germany. A further novelty of this paper is that different specifications commonly used in the economic literature are applied on one unique data set.

The structure of this paper is the following: the next section discusses the econometric specifications that are used to analyse the class size effects on early career earnings in West

Germany. Section 3 describes the data. The main results are presented in section 4 and discussed in section 5. Section 6 concludes with brief review of further research.

2. Econometric specification

The standard earning function regresses the log of hourly earnings (*ly*) on years of schooling (*S*) and a vector of other variables (*X*) that may have an impact on compensation.

$$ly = \alpha + \beta_1 S + \beta_2 X + \varepsilon \tag{1}$$

The coefficient on schooling estimates the percent increase in earnings resulting from one additional year of schooling and is typically interpreted as the rate of return to schooling (Polachek and Siebert, 1993, Borjas, 2000).

This specification, however, does not account for differences in school quality. A given level of schooling yields the same return in this model regardless of which quality of schooling was enjoyed. To allow quality of schooling to have an impact on earnings, I assume β_1 – the return to education – to be a function of school quality (Q).

$$ly = \alpha + \beta_1(Q)S + \beta_2 X + \varepsilon$$

Since I do not know the functional form of $\beta_1(Q)$, I assume a quadratic approximation as proposed for instance by Card and Krueger (1992a):

$$\beta_1 = \gamma_1 + \gamma_2 Q + \gamma_3 Q^2$$

Lazear (2001) provides a theoretical model for optimal class size. He emphasises the public good aspect of class room education. If one pupil disrupts the class, learning is reduced for all other pupils. The probability that one pupil disrupts rises with class size. On the other hand, do pupils learn from their peers. This suggests that class size is strictly concave ($\gamma_2>0$; $\gamma_3<0$), i.e. there are diminishing returns to quality for a given level of schooling. Substitution yields

$$ly = \alpha + \left(\gamma_1 + \gamma_2 Q + \gamma_3 Q^2\right)S + \beta_2 X + \varepsilon\,. \tag{2}$$

I explore two further approaches to model the quality-earning relation. Both alternatives assume that the quality of schooling has an impact on effective schooling (S^*). That is human capital depends on both quality and quantity of schooling:

$$ly = \alpha + \beta_1 S^*(S;Q) + \beta_2 X .$$

Again, I do not know the functional form of $S^*(S;Q)$ and assume therefore first an additive relations between schooling and quality.

$$S^*(S;Q) = \varphi_1 + \varphi_2 S + \varphi_3 Q ,$$

which yields by substitution

$$ly = (\alpha + \phi_1) + \phi_2 S + \phi_3 Q + \beta_2 X + \varepsilon$$

$$\text{with } \phi_i = \beta_1 \varphi_i . \qquad (3)$$

Or, allowing S and Q to interact

$$ly = (\alpha + \phi_1) + \phi_2 S + \phi_3 Q + \phi_4 SQ + \beta_2 X + \varepsilon$$

$$\text{with } \phi_i = \beta_1 \varphi_i . \qquad (4)$$

Equation (3) is in fact the most often used specification if school quality is allowed for in recent British studies (Dearden, Ferri and Meghir, 2000, Dolton and Vignoles, 2000, Harmon and Walker, 2000). In this setting, however, the quality of schooling leads to a parallel shift of the earning function. In other words, the quality of schooling has the same effect at every year of schooling. Model (4), on the other hand, allows the quality of schooling to have a potentially larger effect on early career income if the pupil stays in school longer. This is the setting used, for instance, by Betts (1995).

3. Data and variable definition

The estimates are based on the German Socio-Economic Panel Study (GSOEP). This is a longitudinal survey of individuals living in private households in Germany. The GSOEP covers each year since 1984. Although the GSOEP is a rich data source on individuals and the

households they live in, it provides no information whatsoever on school quality.[5] The strategy is thus to merge district (*Landkreis*) average school quality data with the GSOEP. Using a confidential version of GSOEP[6] provides information on the residential district when secondary schooling is completed within the survey period. Early career earnings are observed at the second year in the job. The focus is thus on those who left school between 1984 and 1997.

Card and Krueger (1996c, 1998) suggest that pupils who are educated at better schools should benefit more per year of schooling than those students attending lower quality schools. There are, however, problems on how to measure the quality of schooling properly, which are discussed, for instance, by Hanushek (2002). This paper defines the quality of schooling through class size[7]. But at which level of aggregation should class size be measured? It might be that the optimal class size for more able pupils is somewhat larger than for less able students. More able pupils might better learn with and from their peers, whereas less able pupils might require more interactions with their teachers and do not perform well in small working groups. Thus, pupils might be sorted into classes conditional on their ability, which may raise problems if class size is measured at the individual-level.

Measuring the quality of schooling at the school-level raises a related problem. It might well be that parents, who care more about their offspring's education, tend to move into catchment areas of higher quality schools (Leech and Campos, 2001). Thus, schools that provide a good quality may attract pupils who have a supportive learning environment at home. These identification problems are avoided if class size is averaged by districts or states. Although aggregation renders these problems it causes further econometric difficulties if

[5] Haisken-DeNew and Frick (2001) provide further information on the GSOEP data.
[6] The German data protection law considers micro data at this level of aggregation as very sensitive and restricts thus this kind of analysis to be conducted on-site at the DIW Berlin. More information on how to obtain this data can be found on http://www.diw.de/english/soep/faq/.
[7] Class size may be just a rough measure for school quality. However, it is a readily available and a easy to understand concept of school quality to parents and policy makers.

important state differences in school policy are omitted (Hanushek, Rivkin and Taylor, 1996). The latter problems is addressed in section 5.1.

The GSOEP provides 693 cohort members who (i) completed schooling between 1984-1997 and (ii) participate in the survey at the second year in the job after having completed either schooling, vocational training, or higher education, which is the time when the early career income is measured. Since the GSOEP does not distinguish between those who gained their highest school degree in a direct way or indirectly (*2. Bildungsweg*), I drop those cohort members that I assume to be too old to have gained their degrees directly; that are, all who were 20 years or older when they completed lower or intermediate secondary schooling and all who were 23 or older when completing upper secondary schooling. Class size at the district level is not available for six cohort members[8]. The analysis is thus based on 445 observations.

The natural logarithm of hourly earnings is the dependent variable. The GSOEP provides information on actual and contractual hours worked per week. I follow Bauer and Haisken-DeNew (2001) and use the maximum of these two. This avoids undercounting the nominal 40 hours of salaried jobs or if a full time employed person actually only worked say 10 hours that interview week due to sickness but would normally work 40 hours. The OECD-MEI consumer price index deflates earnings, since data span more than a decade. The GSOEP provides only information on the awarded degree of secondary education. It is unknown, however, if the cohort member attended schooling straight through or if s/he had to repeat grades. Schooling is thus defined as the minimum years to reach the awarded degree. That is nine years for lower secondary schooling, ten years for intermediate secondary schooling and 13 years for having completed gymnasium. The quality of schooling is approached by class

[8] Four cohort members have missing district information when they completed secondary schooling but state information is available. One district (the city Kassel) does not report class size for every year, thus one cohort member drops; and finally one observation is dropped since the reported class size is unreasonably high (the district Main-Taunus reports a class size that exceeds 200 (sic!)).

size at the district level and at the state level. Since pupils in higher types of schooling (e.g. gymnasium) used to be taught in larger classes then their peers for instance in lower secondary schooling, class size is corrected by the mean of the respective type of schooling. Descriptive statistics of the variables are shown in table 1.

4. Main results

Table 2 summarises the regression results using White (1980) heteroscedasticity corrected standard errors. I am interested in two issues: (i) does class size affect early career earnings; and (ii) if it does, which specification fits best to the German data. Column one of table 2 shows regression results for the base model, which does not allow for class size effects. The other columns allow for various class size effects. The second column represents specification (2) set out above and comes very close to Card and Krueger's (1992a) approach. The third column allows quality to have an effect on the intercept of the earning equation but assumes the slope to be independent of the quality of schooling. This model is the closest representation of the recent British studies.[9] And the last column applies Betts' (1995) modelling strategy and allows the slope to be quality depended.

All regressions have a gender dummy and an indicator for being employed by a public firm when the early career income is measured. The focus is on secondary schooling but time elapsed between having completed secondary schooling and the entry into the labour market. Any human capital acquired during this time is captured by the binary variables for vocational training and higher education. The socio-economic background is captured by the marital status, whether the mother or the father has achieved a degree from gymnasium or higher education, whether the parents are catholic, and by an indicator for having spent childhood in a large city.

These six observations do not drop if the analysis is on state averages.

[9] Dearden, et al. (2000), Dolton and Vignoles (2000), and Harmon and Walker (2000).

Starting with the base model, the working hypothesis is that an adequately specified earning function should allowed for class size effects. Hence, Ramsey's (1969) RESET test should indicate missing variables in the base model. I can reject the null, that the base model has no omitted variables, at the ten percent level but not at the five percent level of significance[10] which is a first indication that quality effects might omitted.

Introducing quality, however, does not add more explanatory power to the models with the R-square measure remaining at 37 percent. The base model is nested in the specifications that allow for quality. Hence it is possible to apply a Wald test to see if the quality inclusive specifications can be restricted to the base model. I reject the null, however, only for model (2). For model (3) and (4) I am not able to reject the null, which suggests that the latter two models can be restricted to the base model and that there are no class size effects on early career earnings in Germany.[11] Although the Wald test indicates quality effects in specification (2), only coefficient γ_3 is individually significant. Moreover, quality has no sizeable effect in this setting and the implied β_1 is 0.043, which is enclosed by the confidence interval of β_1 of model (1) and thus not significantly different. These results suggest, hence, that class size effects has no effect on early career earnings.

5. Discussion

5.1. Aggregation issues

The approach in this paper is to match micro data from the GSOEP with aggregated data at the district level. That is, I assign the district average class size to school leavers who lived in the district in the school leaving year. Moulton (1990) raises the issue that with such an approach OLS standard errors will be downward biased and that inferences will be spurious if the random disturbance terms are correlated within groups. Cohort members who share an

[10] F-test (3, 430)= 2.22 (p-value = 0.085).

[11] As it was suggested by an anonymous referee, I also tested for the joint significance of all quality measures, since the functional relationship is unknown. I can not reject the null hypothesis with an F-test (3, 421)=1.99.

observable characteristic – residence – also may share unobservable characteristics that could lead the disturbance terms to be correlated across individuals within the same distinct or state.

Roger (1993) provides an estimation method that controls for correlated disturbance terms. Applying this method I cluster the observations by the district where schooling was completed. That is, I specify the observations to be independent across clusters (districts) but not necessarily independent within clusters. This changed specification has virtually no effect on the conclusions presented in section 4 and shown in table 2.[12]

It might be, however, that the unobserved characteristics are not district specific but state specific. A vital example for such an unobserved state specific characteristic is school policy which is independently determined by state governments. The observations are therefore also clustered by the state in which schooling was completed.[13] The conclusions suggested by statistical tests do not alter at all. Thus, I conclude that the disturbance terms are not likely to be correlated within clusters.

A related, topic is discussed by Hanushek et al. (1996). They postulate that if important state differences in school policy are omitted, aggregation of the school input variable implies a clear upward biased effect on the outcome measure of schooling. Moreover, Betts (1996) finds that aggregating quality to the state level leads almost always to a positive impact on earnings. Card and Krueger (1996b, 1996a), on the other hand, show econometrically that neither school-level nor more aggregated data is automatically biased. Aggregating the data, they note, would reduce the bias that may arise from potential unobserved correlation between earnings and quality. Hence, it is unclear if my results are influenced by an aggregation bias.

[12] Roger (1993) states that there is no bias as long as the largest cluster is 5 percent or less of the sample size. The largest cluster by district is 3.5 percent in my sample.
[13] N.B. in this setting are the estimates not efficient, i.e. I have ten clusters (states) and 13 parameters. I drop thus the control variables and run the regression again on ten clusters and seven parameters.

Table 3 in the appendix replicates table 2, with quality of schooling measured at the mean corrected average class size aggregated to the state level. It is sufficient to compare the values from the Wald tests. The marginal significance of the quality effects is higher at the state level, as it is suggested by Hanusheck et al. (1996). This difference, however, does not alter the conclusion drawn in section 4.[14]

Hanusheck et al. (1996) point out that studies *"which use less aggregated data are likely to produce more reliable estimates"*. I note that it were desirable to evaluate class size effects at a lower level of aggregation, e.g. at the zip code level or at the school level. Since such data is not available, however, class size at the district level produces the most reliable and feasible results for the German case.

5.2. State effects

Card and Krueger (1996b) note that their earlier findings (Card and Krueger, 1992a) depend heavily on the inclusion of controls for the permanent differences in the rate of return to schooling across states. If they did not control for state effects in their setting, the impact of the quality of schooling on income would not be significant.

Table 4 expands table 2 by including school state dummies that control for differences across the schooling system. Controlling for such difference, the marginal significance of quality effects drops such in model two that a different conclusion as before is suggested. It is indicated that class size effects are not significantly different from zero on the five percent level. Hence, if it is adequately controlled for specific differences in state policy on schooling, class size seems to have virtually no impact on early career income.

[14] A quality measure related to class size is the pupil-teacher-ratio. This measure is also commonly applied in the literature. It is, however, available on the state level only and not on lower levels of aggregation, because teachers stay on the states' payroll in Germany. Approaching the quality of schooling by state averages of the pupil-teacher-ratio suggests that specification (2) can also be restricted to the base model. I.e. quality effects on early career income are not indicated.

6. Conclusion

Research on the relation between school quality and earning has focused mainly on the US and recently also on the UK. The main finding of this literature is that there appears to be no significant effect of class size on early career earnings. It was hypothesised in this paper that there could still exist a quality-earning relation for the German example, since the US-American and the British schooling system differs in many aspects from the schooling system in Germany. The findings presented in this paper, however, reject this hypothesis. I could not find discernible effects of the quality of schooling measured by class size on early career earnings in Germany.

My cohort is too young to observe their complete age-earning profile. The explanatory variable is actually defined as earnings in the early career. But Burtless (1996) postulates that beneficial effects of school quality will not begin to turn up until the cohort has reached its peak earnings years, that is about the age of forty. And indeed the two often cited studies by Card and Krueger (1992a, 1992b) do find a positive and significant quality-earning relation with a cohort that spans an entire working age. All other studies[15], including this one, reject the quality-earning relation but analyse data that is limited by the fact that respondents are fairly young. When data is available that spans an age-earning profile in some years time, further research on this issue will reveal if the quality-earning relation is age dependent.

The policy conclusion suggested by these findings is that reducing class size has virtually no effect on early career earnings. No doubt there exist schools in impoverished areas in which more spending is needed to raise school quality but on average, school quality approached by class size has no significant effect on early career earnings. But earnings are just one indicator of labour market outcomes. An issue that is not addressed in this paper is whether school quality has also other economic gains. I observe earnings only if individuals

[15] C.f. Betts (1995, 1996), Dolton (2000), Dearden, et al. (2000), Harmon and Walker (2000), Hanushek (2002), and Heckman, et al. (1995, 1996).

are employed. But employability itself could be influenced by the quality of schooling. Further research should thus investigate if the risk to become unemployed is related to school quality. Furthermore, for a given educational attainment there appears to be no quality effect on earnings. School quality, however, may encourage pupils to acquire higher levels of schooling. Hence, it might be that their exists an indirect link between quality and earnings, since more schooling returns higher earnings. More research on these issues will yield further interesting insights.

7. References:

Angrist, J. and V. Lavy, 2002, New Evidence on Classroom Computers and Pupil Learning, *The Economic Journal* 112, 735-765.

Artelt, C., W. Schneider and U. Schiefele, 2002, Ländervergleiche zur Lesekompetenz, in: D. Pisa-Konsortium eds., *PISA 2000 - Die Länder der Bundesrepublik Deutschland im Vergleich* (Opladen Leske + Budrich).

Ashenfelter, O., C. Harmon and H. Oosterbeek, 1999, A review of estimates of the schooling/earnings relationship, with tests for publication bias, *Labour Economics* 6, 453-470.

Barro, R. J. and J.-W. Lee, 2000, Schooling Quality in a Cross Section of Countries, *Economica* 68, 465-488.

Bauer, T. K. and J. P. Haisken-Denew, 2001, Employer learning and the returns to schooling, *Labour Economics* 8, 161-180.

Betts, J. R., 1995, Does school quality matter? Evidence from the National Longitudinal Survey of Youth, *The Review of Economics and Statistics* 77, 231-250.

Betts, J. R., 1996, Is there a link between school inputs and earnings?, in: G. Burtless eds., *Does Money Matter?* (Washington, D.C. Brookings Institution).

Blundell, R., L. Dearden, C. Meghir and B. Sianesi, 1999, Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy, *Fiscal Studies* 20, 1-23.

Borjas, G. J. eds., 2000, *Labor Economics* (USA, McGraw-Hill).

Burtless, G., 1996, Introduction and Summary, in: G. Burtless eds., *Does money matter?* (Washington D.C. The Brookings Institute).

Card, D., 1999, The causal effect of education on earnings, in: O. Ashenfelter and Card, D. eds., *Handbook of Labor Economics* (Amsterdam Elsevier).

Card, D., 1995, Earningns, Schooling, and Ability Revisited, *Research in Labor Economics* 14, 23-48.

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

Card, D. and A. B. Krueger, 1996a, Labor market effects of school quality: theory and evidence, *NBER Working Paper Series* 5450,

Card, D. and A. B. Krueger, 1996b, Labour Market Effects of School Quality: Theory and Evidence, in: G. Burtless eds., *Does Money Matter? The Link Between Schools, Student Achievement and Adult Success* (Washington D.C. Brookings Institution).

Card, D. and A. B. Krueger, 1992b, School Quality and Black-White Relative Earnings: A Direct Assessment, *Quarterly Journal of Economics* 107, 151-200.

Card, D. and A. B. Krueger, 1998, School Resources and Student Outcomes, *Annals, AAPSS* 39-53.

Card, D. and A. B. Krueger, 1996c, School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina, *Journal of Economic Perspectives* 10, 31-50.

Dearden, L., J. Ferri and C. Meghir, 2000, The Effect of School Quality on Educational Attainment and Wages, *ifs working paper* 00/22,

Dolton, P. and A. Vignoles, 2000, The Impact of School Quality on Labour Market Success in the United Kingdom, *mimeo*

Fertig, M. and C. M. Schmidt, 2002, The Role of Background Factors for Reading Literacy: Straight National Scores in the PISA 2000 Study, *IZA Discussion Paper Series*

Haisken-Denew, J. P. and J. R. Frick eds., 2001, *Desktop Companion to the German Socio-Economic Panel* (Berlin, DIW-SOEP).

Hanushek, E. A., 1995, Interpreting Recent Research on Schooling in Developing Countries, *The World Bank Research Observer* 10, 227-264.

Hanushek, E. A., 2002, Publicly provided education, *NBER Working Paper Series*

Hanushek, E. A., S. G. Rivkin and L. L. Taylor, 1996, Aggregation and the estimated effects of school resources, *The Economic and Social Review* 78, 611-627.

Harmon, C. and I. Walker, 2000, The Returns to the Quality and Quantity of Education: Evidence for Men in England and Wales, *Economica* 67, 19-35.

Harmon, C., I. Walker and N. Westergaard-Nielsen eds., 2001, *Education and Earnings in Europe* (Northampton, Edward Elgar Publishing).

Heckman, J., A. Layne-Farrar and P. Todd, 1996, Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship, in: G. Burtless eds., *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success* (Washington, D.C. Brookings Institution Press).

Heckman, J., A. Layne-Farrar and P. Todd, 1995, The Schooling Quality-Earnings Relationship: Using Economic Theory to Interpret Functional Forms Consistent with the Evidence, *NBER Working Paper Series*

Lauer, C. and V. Steiner, 2001, Germany, in: C. Harmon, Walker, I. and Westergaard-Nielsen, N. eds., *Education and Earnings in Europe* (Northampton Edward Elgar Publishing).

Lazear, E. P., 2001, Educational Production, *Quarterly Journal of Economics* 116, 777-803.

Leech, D. and E. Campos, 2001, Is comprehensive education really free? A case study of the effects of secondary school admissions policies on house prices in one local area, *Warwick Economic Research Papers*

Moulton, B. R., 1990, An illustration of a pitfall in estimating the effects of aggregate variables on micro units, *The Review of Economics and Statistics* 334-338.

Polachek, S. W. and S. Siebert eds., 1993, *The Economics of Earnings* (Cambridge, Cambridge University Press).

Psacharopoulos, G., 1994, Returns to Investment in Education: A global Update, *World Development* 22, 1325-1343.

Ramsey, J. B., 1969, Test for Specification Error in classical Linear Least Squares Regression Analysis, *Journal of the Royal Statistical Society* B31, 250-271.

Roger, W., 1993, Regression standard errors in clustered samples, *Stata Technical Bulletin Reprints* 3, 88-94.

Sander, W., 1993, Expenditures and student achievement in Illinois, *Journal of Public Economics* 52, 403-416.

West, A., H. Pennell, T. Travers and R. West, 2001, Financing school-based educatoin in England: poverty, examination results, and expenditure, *Environment and Planning C: Government and Policy* 19, 461-471.

White, H., 1980, A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica* 48, 817-838.

Wilson, K., 2000, Using the PSID to study the effects of school spending, *Public Finance Review* 28, 428-451.

Winter-Ebmer, R. and A. Wirz, 2002, Public Funding and Enrolment into Higher Education in Europe, *IZA Discussion Paper Series*

Wößmann, L., 2000, Schooling Resources, Educational Institutions, and Student Performance: The International Evidence, *Kiel Working Papers*

Wößmann, L., 2001, Schulsystem und Schülerleistung im internationalen Vergleich: Was Institutionen ausmachen, *Die Weltwirtschaft* 283-304.

Table 1:
DESCRIPTIVE STATISTICS

VARIABLES	MEAN	STD. DEV.
Dependend variable:		
Log of hourly wages	2.89	0.41
Explanatory variables:		
Years of schooling	10.55	1.60
Mean corrected class size (district averages)	0.43	4.01
Mean corrected class size (state averages)	-0.10	2.38
Class size (district averages)	27.67	6.86
Class size (state averages)	29.04	8.05
Dummies (yes=1):		
Vocational training	0.72	
University education	0.14	
Female	0.52	
Public sector	0.24	
Married	0.31	
Mother has gymnasium or higher	0.03	
Father has gymnasium or higher	0.10	
Mother is catholic	0.42	
Father is catholic	0.37	
Childhood spent in a large city	0.11	

Note:
1) Numbers of observations: 445.
2) Source: GSOEP 1984-2000.

Table 2:
REGRESSION OF LOG HOURLY EARNINGS ON
CLASS SIZE AT THE DISTRICT LEVEL

	(1)	(2)	(3)	(4)
Female	-0.217	-0.221	-0.217	-0.217
	(6.63)**	(6.76)**	(6.62)**	(6.65)**
Public sector	0.034	0.039	0.034	0.034
	(0.97)	(1.11)	(0.97)	(0.97)
Vocational training	0.358	0.362	0.359	0.359
	(7.51)**	(7.60)**	(7.54)**	(7.33)**
University education	0.475	0.488	0.479	0.478
	(8.62)**	(8.74)**	(8.53)**	(8.47)**
Years of schooling	0.035	0.043	0.033	0.033
	(3.08)**	(3.11)**	(2.68)**	(2.65)**
(Years of schooling)*(Class size)*10^3		0.467		-0.178
		(1.37)		(0.04)
(Years of schooling)*(Class size)^2*10^3		-0.090		
		(2.49)*		
Class size			0.002	0.004
			(0.48)	(0.08)
Married	0.071	0.070	0.071	0.072
	(2.15)*	(2.16)*	(2.18)*	(2.14)*
Mother has gymnasium or higher	0.031	0.073	0.038	0.038
	(0.35)	(0.83)	(0.44)	(0.44)
Father has gymnasium or higher	-0.042	-0.029	-0.043	-0.043
	(0.73)	(0.52)	(0.75)	(0.75)
Mother is catholic	-0.028	-0.035	-0.025	-0.025
	(0.59)	(0.72)	(0.52)	(0.52)
Father is catholic	-0.009	-0.011	-0.010	-0.009
	(0.19)	(0.23)	(0.20)	(0.19)
Childhood spent in a large city	0.035	0.054	0.040	0.040
	(0.76)	(1.10)	(0.77)	(0.77)
Constant	2.297	2.225	2.311	2.311
	(18.04)**	(15.32)**	(17.35)**	(17.19)**
R-squared	0.37	0.37	0.37	0.37
RESET	2.22	1.62	2.12	2.10
Prob>RESET	0.09	0.18	0.10	0.10
Wald[6)]		3.68	0.23	0.12
Prob>Wald		0.03	0.63	0.89

Note:
1) Robust t-statistics in parentheses.
2) Class size is the mean corrected district average.
3) Source: GSOEP 1984-2000.
4) Number of observation is 445.
5) The levels of significance are: ** at 1%, at 5%.
6) The Wald test tests if the corresponding model can be restricted to model (1). The degrees of freedom for this test are 2, 431 for model (2) and (4) and 1, 432 for model (3).

Table 3:
REGRESSION OF LOG HOURLY EARNINGS ON
CLASS SIZE AT THE STATE LEVEL

	(1)	(2)	(3)	(4)
Female	-0.216	-0.217	-0.219	-0.219
	(6.68)**	(6.75)**	(6.76)**	(6.75)**
Public sector	0.034	0.030	0.032	0.032
	(0.97)	(0.86)	(0.92)	(0.93)
Vocational training	0.357	0.361	0.362	0.362
	(7.53)**	(7.63)**	(7.61)**	(7.47)**
University education	0.475	0.476	0.487	0.488
	(8.71)**	(8.76)**	(8.91)**	(8.82)**
Years of schooling	0.035	0.044	0.034	0.034
	(3.17)**	(3.53)**	(3.06)**	(3.05)**
(Years of schooling)*(Class size)*10^3		0.485		0.339
		(0.85)		(0.07)
(Years of schooling)*(Class size)^2*10^3		-0.236		
		(2.20)*		
Class size			0.010	0.006
			(1.43)	(0.09)
Married	0.072	0.082	0.079	0.079
	(2.24)*	(2.50)*	(2.42)*	(2.42)*
Mother has gymnasium or higher	0.029	0.065	0.033	0.034
	(0.33)	(0.78)	(0.38)	(0.38)
Father has gymnasium or higher	-0.041	-0.045	-0.040	-0.040
	(0.73)	(0.80)	(0.72)	(0.71)
Mother is catholic	-0.028	-0.035	-0.030	-0.030
	(0.59)	(0.73)	(0.62)	(0.62)
Father is catholic	-0.008	-0.009	-0.009	-0.010
	(0.17)	(0.18)	(0.19)	(0.20)
Childhood spent in a large city	0.033	0.043	0.046	0.046
	(0.74)	(0.96)	(1.03)	(1.01)
Constant	2.291	2.214	2.303	2.304
	(18.25)**	(16.31)**	(18.48)**	(18.50)**
R-squared	0.37	0.37	0.37	0.37
Wald[6)]		4.50	2.06	1.17
Prob>Wald		0.01	0.15	0.31

Note:
1) Robust t-statistics in parentheses.
2) Class size is the mean corrected state average.
3) Source: GSOEP 1984-2000.
4) Number of observation is 445.
5) The levels of significance are: ** at 1%, * at 5%.
6) The Wald test tests if the corresponding model can be restricted to model (1). The degrees of freedom for this test are 2, 431 for model (2) and (4) and 1, 432 for model (3).

Table 4:
REGRESSION OF LOG HOURLY EARNINGS ON CLASS SIZE AT THE DISTRICT LEVEL CONTROLLING FOR STATE EFFECTS

	(1)	(2)	(3)	(4)
Female	-0.217	-0.224	-0.218	-0.218
	(6.66)**	(6.84)**	(6.65)**	(6.67)**
Public sector	0.041	0.043	0.041	0.041
	(1.14)	(1.20)	(1.13)	(1.14)
Vocational training	0.357	0.357	0.356	0.356
	(7.48)**	(7.47)**	(7.46)**	(7.25)**
University education	0.486	0.488	0.483	0.484
	(8.82)**	(8.62)**	(8.54)**	(8.48)**
Years of schooling	0.036	0.047	0.038	0.037
	(3.10)**	(3.23)**	(2.95)**	(2.90)**
(Years of schooling)*(Class size)*10^3		0.050		0.252
		(0.13)		(0.05)
(Years of schooling)*(Class size)^2*10^3		-0.092		
		(2.27)*		
Class size			-0.002	-0.005
			(0.38)	(0.09)
Married	0.069	0.067	0.068	0.068
	(2.08)*	(2.05)*	(2.07)*	(2.01)*
Mother has gymnasium or higher	0.065	0.090	0.061	0.062
	(0.75)	(1.04)	(0.71)	(0.72)
Father has gymnasium or higher	-0.054	-0.041	-0.055	-0.054
	(0.93)	(0.72)	(0.93)	(0.92)
Mother is catholic	-0.030	-0.036	-0.031	-0.031
	(0.60)	(0.71)	(0.62)	(0.62)
Father is catholic	-0.006	-0.006	-0.005	-0.005
	(0.13)	(0.11)	(0.11)	(0.11)
Childhood spent in a large city	0.036	0.046	0.032	0.032
	(0.78)	(0.95)	(0.64)	(0.63)
Constant	2.344	2.243	2.332	2.332
	(17.83)**	(14.59)**	(16.84)**	(16.82)**
School state dummies	Yes	Yes	Yes	Yes
R-squared	0.38	0.38	0.38	0.38
Wald[6)]		2.63	0.15	0.07
Prob>Wald		0.07	0.70	0.93

Note:
1) Robust t-statistics in parentheses.
2) Class size is the mean corrected state average.
3) Source: GSOEP 1984-2000.
4) Number of observation is 445.
5) The levels of significance are: ** at 1%, * at 5%.
6) The Wald test tests if the corresponding model can be restricted to model (1). The degrees of freedom for this test are 2, 422 for model (2) and (4) and 1, 423 for model (3).