ECONOMIC RESEARCH FORUM
منتدى البحوث الاقتصادية
2010

working paper series

EXTERNAL RETURNS TO HIGHER EDUCATION IN TURKEY

Ozan Bakis, Nurhan Davutyan,
Haluk Levent, and Sezgin Polat

Working Paper No. 517

EXTERNAL RETURNS TO HIGHER EDUCATION IN TURKEY

Ozan Bakis, Nurhan Davutyan,
Haluk Levent, and Sezgin Polat

Working Paper 517

April 2010

This paper was presented at the 16th Economic Research Forum Conference (7–9 November 2009, Cairo-Egypt). We thank, without implicating in any way, Cem Baslevent and Ritva Reinikka for useful comments.

Send correspondence to:
Ozan Bakis
Department of Economics, Galatasaray University
Turkey
Email: ozanbakis@gmail.com

First published in 2010 by
The Economic Research Forum (ERF)
7 Boulos Hanna Street
Dokki, Cairo
Egypt
www.erf.org.eg

Abstract

This paper studies local human capital externalities and returns to education in Turkey. Data comes from 2006 Household Labor Survey. Instrumental Variables-OLS estimation indicates internal (external) returns amounting to 4.9% (2.4%), while IV estimates using quantile regression range from 3% to 6.9% (1.3% to 3.5%). We discuss further characteristics of the Turkish labor market segmented by gender and show that external returns are uniformly higher for women. Our results also indicate both internal and external returns increase or equivalently the wage distribution spreads out as education increases.

ملخص

تتناول هذه الورقة تأثير العوامل الخارجية و العائدات الرأسمالية البشرية المحلية علي التعليم في تركيا. و قد أخذت هذه البيانات من خلال المسح الذي تم إجراءه علي العمالة المنزلية لعام 2006. تشير تقديرات المتغيرات المساعدة – باستخدام طريقة المربعات الصغرى العادية إلي أن العائدات الداخلية (المتأثرة بعوامل خارجية) تقدر ب% 4.9 (2.4%)، بينما تشير تقديرات المتغيرات المساعدة- باستخدام الانحدار الكمي إلي أن العائدات الداخلية تتراوح بين 3%و6.9% (1.3% إلي3.5%.) نتناول الورقة أيضا مزايا أخري لسوق العمل التركي الذي يقسم حسب الجنس و توضح أن العائدات الخارجية أعلي بشكل عام لدي السيدات. كما تشير النتائج إلي زيادة في كل من العائدات الداخلية و الخارجية أو توزيع الأجور بالتساوي مع ارتفاع مستوي التعليم.

1. Introduction

In contrast to traditional development theory where the source of growth — technical change— is exogenous, new growth theory endogenizes the source(s) involved. It posits knowledge augmented accumulation of productive factors obviates diminishing returns thereby enabling sustained output increases. This allows room for policy intervention, since increased investment in human capital would accelerate growth. Acemoglu (2009) provides a thorough treatment. Endogenous growth theorists differ in the precise mechanisms connecting knowledge to growth, but external benefits to education — knowledge spillovers— is the common theme. For Lucas (1988) such externalities constitute the foundation of economic development.

Early empirical work on external returns has relied on cross-country growth regressions and macroeconomic time series data, Barro (1997). These studies have been criticized on grounds of data and identification problems, e.g. Krueger and Lindahl (2001). A more recent body of literature investigates these issues at a more disaggregated level. The common approach is to estimate education's external benefits by augmenting an individual earnings model with aggregate measures of schooling. One strand pioneered by Winter-Ebmer (1994) focuses on spillovers occurring within industry sectors (Sakellariou, 2001; Sakellariou and Maysami, 2004; Kirby and Riley, 2009). Starting with Rauch (1993), the other strand assesses knowledge spillovers occurring within sub-national geographic jurisdictions, e.g. Acemoglu and Angrist (2000); Moretti (2004); Shields and Shields (2009), all concerning the United States.

We contribute to this growing literature by presenting evidence on external returns to regional level schooling in Turkey for 2006. Since wage data tends to be skewed and fat-tailed, we use (LAD) least absolute deviations-based methods rather than least squares. We estimate the education-earnings nexus at different points of the wage distribution rather than only at the mean, which allows for a richer interpretation.

2. Modeling

The regional approach pioneered by Rauch (1993) assumes competitive markets where factors earn their marginal product. Here labor's marginal product includes an external return element. Moretti (2004b) provides a thorough exposition.

The model assumes perfectly mobile capital, labor and a fixed amount of land in each region. Following Lucas (1988), a region's aggregate human capital is also a productive factor. Capital is traded nationally, land labor only locally. The rental price of capital is common to all regions while land and labor prices are region specific. Land includes "free" amenities like clean beaches or good public schools. Among identical workers, higher wages may be due to a high capital to labor ratio at the workplace — i.e. private return— or a high regional level of education – i.e. external return. When workers move to a region with high wages or greater amenities, housing becomes expensive. Hence equalization of real wages across regions entails different nominal wages and land prices. Thus a worker's (nominal) wage will depend on her/his characteristics as well as the workplace involved and those of the region. For instance all else being equal, a worker may end up accepting a lower wage for the sake of living in a desirable region.

We use the 2006 Household Labor Survey. We focus on workers in the private sector where productivity considerations can be expected to dominate remuneration decisions. We also discuss the implications for each gender separately. Our model is:

$$w_{ij} = \beta X_{ij} + \lambda R_j + \alpha H_j + \mu_j + \theta_i + \varepsilon_{ij} \qquad (1)$$

Subscripts i and j index over individuals and the 26 regions comprising Turkey. We have logged wages (w_{ij}) for 54,728 workers. X holds all individual and workplace attributes including age, gender, tenure, workplace size, 9 sectoral dummies (e.g. agriculture, mining etc), social security status[1] and years of schooling as individual human capital. H is regional human capital proxied by college share in total employment, R stands for regional indicators. We have data on three regional variables: unemployment rate for each educational stratum, an amenities index, and a per capita openness to foreign trade index.

The error terms comprise θ_i (including ability) possibly correlated with individual human capital; a regional factor μ_j possibly correlated with H_j; finally ε_{ij} is an i.i.d random shock.

3. Identification

The parameters of interest are the coefficients of individual schooling and regional human capital. Identification poses a problem due to omitted variables impacting both wages and educational status. We need but cannot measure individual ability. The impact of ability may be mistaken for that of education. A good instrument, like quarter of birth a la Acemoglu-Angrist (2000), which approximates random schooling assignments irrespective of ability, is unavailable.

Fortunately, in the case of individual schooling, substantial evidence indicates that downward bias caused by measurement error in educational attainment cancels "upward ability bias". Krueger and Lindahl (2001, p1101) reach this conclusion after an extensive literature survey. If, as argued by Moretti (2004a,b), returns to traits like ability which are correlated with individual schooling vary across regions, such that particular individuals self select to particular regions, then regional human capital too will suffer from omitted variables bias. To counteract this possibility and following Vandenbussche et al (2006), we instrument regional knowledge by its own three years lagged value. This "predetermined variables" approach is valid if serial correlation is absent, as is likely to be in our case since we are using cross-sectional data.

4. Estimation

Ordinary least squares regression focuses on the mean. Thus it only provides a partial view of the relationship between the regressors and the dependent variable. Quantile regression (QR) gives a more complete picture since it provides information about the linkage between the outcome variable (in our case wages) and the regressors at different points of the conditional wage distribution.

Other advantages of quantile or least absolute deviations regression include robustness to outliers and avoidance of distributional assumptions regarding regression errors. As pointed out by Cameron and Trivedi (2009), these features make QR especially suitable for heteroskedastic data. We performed the available tests (Breusch-Pagan / Cook-Weisberg) to check for heteroskedasticity. The null hypothesis of constant variance was decisively rejected in every case. In addition both a Kolmogorov-Smirnov test as well as a skewness and kurtosis test soundly rejected the null hypothesis of normality. Consequently and noting that our median (QR) estimates of returns to education are consistently below the mean (OLS) estimates – see Table 1-, we believe adopting this approach is appropriate as well as fruitful.

The quantile regression model posits a linear relationship between the conditional quantile and the set of explanatory variables X[2]:

[1] The extent of informality in Turkey is considerable. Workers in that sector lack social protection which is reflected in their wages. See Davutyan (2008).

[2] For expository convenience here we subsume regional indicators R and regional human capital H under X.

$$w = X'ß + \mu \quad (2)$$

Letting F denote the (unknown) cumulative density function of w, the quantile q Є (0, 1) is defined as that value of w that splits the data into the proportions q below and (1-q) above. In other words F (w_q) = q and $F^{-1}(q) = w_q$. For instance since the median (q=0.5 or 50'Th percentile) of our logged wages equals 0.734, Pr (w <= 0.734) is 0.5. These concepts extend to our predictor, namely the conditional quantile regression function Q_q (w/X). Thus

$$Q_q (w/X = x) = x' \beta_q \text{ for } 0 < q < 1 \quad (3)$$

Unlike the OLS and maximum likelihood estimators, the QR estimator does not have a closed form solution. Its computational implementation requires numerical optimization via linear programming methods. The q'th regression quantile estimate βhatq is the solution to the following minimization problem:

$$Q(\beta_q) = \sum_{w>x'\beta}^{N} q \mid w - x'\beta q \mid + \sum_{w\leq x'\beta}^{N} (1-q) \mid w - x'\beta q \mid \quad (4)$$

When q is 0.5 equal weights are placed on both types of predictions and one gets the median estimator whose ancestry goes back to Laplace's work in 1789, Koenker (2005, p4). If, say, q = 0.75 greater weight is put on predictions where w>x'ß than for those where w<=x'ß, and we get the QR estimator for the 75'Th percentile of the wage distribution.

As Buchinsky (1998) demonstrates, the QR estimator that minimizes (5) is asymptotically normal under quiet general conditions. Thus

$$\beta hat_q \sim N (\beta_q, A^{-1} B A^{-1}) \quad (5)$$

where A = ∑i xixi' B = ∑i f uq (0/xi) xixi' and f uq is the conditional density of the error term μq = w - x' β_q evaluated at μ_q = 0. Since this term is awkward to evaluate bootstrap methods are more commonly used to generate variance-covariance estimates. This adds to the computational burden. Indeed, given that the advantages of median regression over OLS are known since the early 19'Th century, the latter's preponderance is essentially due to its having a closed form solution obviating the need for numerical optimization. From this perspective the recent resurgence of quantile regression is due to the microchip revolution and the concomitant cheapening of computing as well as storage capacity and the resulting emergence of large human resources data sets.

For useful overviews of the quantile method as well as additional insights, the reader is referred to Cameron and Trivedi (2005, pp89-90) and Angrist and Prischke (2009, pp269-291) and the works cited therein.

4.1 Instrumental Variables Quantile Regression

The standard quantile method, just like OLS, assumes that all the regressors are exogenous. But if as is the case here, regional human capital H is endogenous, then using conventional quantile regression to infer about its impact over the entire wage distribution will yield biased results. Chernozhukov and Hansen (2006, 2008) develop an instrumental variables quantile regression model (IVQR) that eliminates this source of bias. Given the following structural model:

$$w = \beta X + \alpha H + U \quad (6)$$

Where w is logged wages, H is the endogenous human capital of the relevant region and X is the vector of control variables including regional indicators and U is an error term. The corresponding quantile regression function becomes:

$$Q_q (w/X, H) = \beta_q X + \alpha_q H \quad \text{for } 0 < q < 1 \quad (7)$$

Chernozhukov and Hansen (2008, pp381-383) derive an estimation equation of the following form:

$$P (w <= \beta_q X + \alpha_q H / X, Z) = q \quad (8)$$

using the following assumptions:

(a) $\beta_q X + \alpha_q H$ is strictly increasing in q
(b) U is uniformly distributed with mean 0 and SD equaling 1
(c) The instrument(s) Z is independent of U
(d) Z and H are dependent.

Essentially (8) provides a moment restriction which can be used to obtain IVQR estimates for β_q and α_q. Specifically, for a given value

of α_q. , we regress $(w - \alpha_q H)$ against X and Z to estimate $\beta hat (\alpha_q)$ and $\gamma hat (\alpha_q)$ where γhat represents the estimated coefficient(s) of the instrument(s). The moment equation (8) implies that zero is the quantile solution to $(w - \beta_q X - \alpha_q H)$ conditional on X and Z. Therefore that particular α_q value which makes the coefficients of the instrumental variables $\gamma hat (\alpha_q)$ as close to zero as possible, is our IVQR estimate of the impact of H on w. Formally:

$$\alpha hat_q = \text{arginf } Wn[(\alpha_q)] = n[\gamma hat(\alpha_q)' A(\alpha) \gamma hat(\alpha_q)]$$
$$\alpha \in \Omega \quad (9)$$

Ω is the parameter space for α. As shown in Chernozhukov

and Hansen (2008, p383), $A(\alpha)$ is the inverse of the asymptotic covariance matrix of $sqrt(n)[\gamma hat(\alpha_q) - \gamma(\alpha_q)]$ which means that

$Wn[(\alpha_q)]$ is the Wald statistic for testing $\gamma(\alpha_q) = 0$.

In summary for a given probability level q, the estimation procedure works as follows:

a) Define a set of suitable values $\{\alpha_j , j= 1 \text{ to } J\}$ and run ordinary quantile regressions of $(w – \alpha jH)$ against X and Z to estimate $\beta hat(\alpha j)$ and $\gamma hat(\alpha j)$
b) Use the inverse of the covariance matrix of $\gamma hat (\alpha_j)$ to obtain the Wald statistics – namely z or F values - Wn $[(\alpha_j)]$. Take the α_j value that minimizes the $Wn[(\alpha_j)]$ as the estimate of α for that quantile level. The estimates of β_q vector are the corresponding coefficients on X.

Chernozhukov and Hansen (2008) provide two detailed examples of how their method works in practice. The present study was implemented on STATA 10. Convergence was achieved without much difficulty in every case. However the estimation of the variance-covariance matrix required the use of bootstrap methods. Other studies using this same approach include Eren (2009), Galvao and Rojas (2009).

5. Regression Results

Our data set pertains to 2006 and consists of 54,728 observations on individuals working for private sector firms, in the 26 statistical regions comprising Turkey[3]. We restricted our sample to private sector employees on grounds that – as compared to the public sector- productivity considerations play a more decisive role in remuneration decisions.

We regressed logged hourly wages (WAGE) against a set of control variables. They are: worker age and its squared value (AGE, AGE2), length of service in current job and its square (TENURE, TENURE2), a dummy indicating social security status (SOSEC) and a

[3]Source: TURKSTAT. Turkey is divided into 26 NUTS2 (Nomenclature of Territorial Units for Statistics) regions.

second dummy WORKPLACE to denote large enterprises with more than 50 employees. We also have data on three regional variables: Unemployment rate (UN_RATE) for each educational stratum, an amenities index (INDEX) for each region, and openness to foreign trade index (OPEN). OPEN is calculated as per capita exports plus imports at the regional level. We also have 9 sectoral dummies. They are: AGRICULTURE –subsumed under the constant-, MINING,

MANUFACTURE, UTILITIES, CONSTRUCTION, TRADE, TRANSPORT, FINANCE and SOCIAL SERVICES[4].

We quantify each worker's individual capital stock by her/his years of schooling (EDU). Thus in each regression EDU's coefficient is our measure of private or internal return to education. The share of college graduates in each region's workforce (COLLEGE) is our proxy for the regional human capital stock. Therefore, COLLEGE's coefficient measures the external return to education[5].

Appendices 2-5 display the full set of results. It can be seen that abstracting from the amenity INDEX – discussed subsequently –,all variables have the expected signs. AGE and TENURE are strongly positively, and their squares negatively, significant. The SOSEC and WORKPLACE dummies indicate that workers with social security coverage working in large enterprises (> 50 employees) receive higher wages. The coefficient of the FEMALE dummy - for the full sample of 54,728 observations- is always significantly negative. Wages in non-AGRICULTURAL sectors are in general higher as indicated by their positive coefficients. The negative and significant UN_RATE (unemployment rate) coefficients point out to Phillips curve effects, particularly at the upper echelons of the wage distribution. Since UN_RATE varies on a regional as well as educational stratum basis, we believe it is successful in capturing the impact of unemployment on wages. Thus higher wage flexibility at the upper echelons of the wage curve indicates individual bargaining plays a greater role in wage setting at higher wage levels.

From this perspective the reduced flexibility or stickiness observed at the lower ends of the wage curve would be explained by "fair wage" considerations or minimum wage laws stressed by Akerlof and Yellen (1990) The coefficients of OPEN - measuring openness to foreign trade for each NUTS region- are strongly positive. We will argue this may partially reflect an externality where the knowledge spillover is from the foreign to the domestic sector.

In what follows we highlight our salient findings and present the relevant private (EDU) and external (COLLEGE) returns to education. As mentioned previously our Kolmogorov-Smirnov as well as skewness and kurtosis tests strongly rejected the null of normality in favor of a skewed and fat-tailed wage distribution. The relevant p values were less than 0.0001 and 0.00001 respectively. It is well known that for such data median (rather than the mean) is the better measure of central tendency. Table 1 compares the OLS estimates with the corresponding QR or median (Q=0.5) estimates. In each case, the "more typical" median estimates of educational returns are less than the corresponding mean estimates. We believe this pattern is consistent with the leptokurtic nature of our wage distribution and thus justifies the use of quantile methods.

In view of these results we conclude that the "typical" private return to education does not vary by gender and is about 3.8%, as indicated by the EDU iv-qr row of Table 1. On the other hand, the "typical" external return to education for females is much larger than for males: 3.4% versus 1.4%. See the corresponding entries in Table 1 for the COLLEGE iv-qr row.

[4] Appendix 1 defines their coverage.

[5] In terms of eq (9) of Section 4.1, COLLEGE for 2006 is H, whereas COLLEGE for 2003 is Z.

We now turn to anayze the returns to education information provided by the quantile estimates at the 10'th 25'th 50'th 75'th and 90'th percentiles of the wage distribution.

Both conventional and instrumental variables QR estimates displayed in Table 2 show that, for a given educational level (say EDU = 5 yrs), individual returns get larger as one moves up the wage distribution. This pattern holds for the whole data set as well as female and male workers separately. For ordinary quantile estimates, F-tests for the equality of these coefficients were soundly rejected. In each case the relevant p value was below 0.00001. This finding implies that as education levels increase the corresponding wage curve gets wider[6]. For instance the estimates for ALL indicate an additional year of schooling raises the lower decile of wages by 3%, the median by 3.8% but the upper decile by 6.5%. In other words wage inequality increases with education. The same finding is reported to hold for US wages starting with the 1990'es, Angrist and Pischke (2008, p28). Martins and Pereira (2004) present a similar finding most EU countries.

The underlying reason is still debated, however the weight of the evidence points to skill biased technical change associated with the computer revolution, Lemieux (2008, pp21-22). Given Turkey's position in and integration with the world economy (its foreign trade grew from 31% of GNP in 2000 to 45% in 2008[7]), it is tempting to argue the same phenomenon underlies our finding as well. The policy implication that follows involves promoting skills that complement and facilitate computer usage broadly conceived.

Table 3 lists our estimates for external returns to education. First we note our instrumental variable estimates are always higher than the corresponding conventional QR estimates. Second, as in the case of internal returns, coefficients get larger when moving up the wage curve. For ordinary quantile estimates, F tests of equality of these coefficients were decisively rejected with a p value < 0.0001 for all workers and women but not for men. Finally external returns for women are substantially above those for men. This pattern holds for each quantile level. For instance the median IVQR external return for women is twice larger than the corresponding return for men – 3.1% versus 1.5%.

We can think of two complementary explanations for this phenomenon. Turkish women being relative newcomers to the labor market, have a greater zeal to learn, perhaps as a compensatory mechanism. Second, men display a greater willingness to engage in informationally meaningful interactions with female –as compared with male- colleagues. Enterprise level data could shed more light on this point. However the policy implication of this finding is clear. Social i.e. private plus external returns to education are higher for women. This finding points out to an additional reason towards encouraging female labor force participation. It can also have implications for promotion and other organizational practices within firms.

Two additional findings displayed in Table 4 deserve comment. These numbers are the 2SLS and instrumental variable QR estimates for OPEN and the amenities INDEX.

As mentioned previously, Turkey is well integrated into the world economy. The GNP share of foreign trade in 2008 was about 45%. Given the nature of the production process and of the world supply chain under globalization, a good deal of this trade is of the intra-firm or intra-industry variety. A careful study by Cakmak (2006) demonstrates that such trade involving Turkey and four major EU[8] countries increased by almost 300% from 1991 to

[6] Note: In each case the quantile coefficients successively increase. Thus the "simple regression" line of WAGE against EDU for q=0.90 is steeper than the corresponding q=0.10 regression line. This means a more spread out wage distribution for higher EDU values.

[7] Undersecretariat for Foreign Trade http://www.dtm.gov.tr/dtmweb/

[8] France, Germany, Italy and the UK.

2004. For instance in 2004, 38% of all manufacturing trade between Turkey and France involved intra-industry trade, in the automotive subsector the ratio reached 47.2%. Similar figures hold for Germany and Italy as well. Since such trade requires exchange of personnel during the production process, we believe the significantly positive coefficients for OPEN - shown in the top portion of Table 4 - reflect knowledge spillover effects as well. Use of finer enterprise level data may further clarify this issue.

Finally we turn to the 2SLS and IVQR coefficients of our amenities INDEX. According to the theory expounded in our Section 2, "compensating wage reductions" would occur to equalize wages across different NUTS regions. Basically, the greater the "free" amenities of a region – such as good public schools, cultural facilities, clean beaches etc- valued by workers, the greater would be the wage cut involved, ceteris paribus. Thus we expected a significantly negative coefficient for our amenities INDEX. Unfortunately our evidence does not support this prediction. We conclude our "compensating wage reductions" mechanism is either too weak to reveal itself and/or our amenities index is too crude.

6. Summary

Endogenous growth theory emphasizes the role of knowledge spillovers in fostering economic growth. Given the econometric difficulties involved in quantifying such external effects using cross country data, recent work has focused on sub-national regions, e.g. Moretti (2004a, b).

Using 2006 data on the 26 Turkish NUTS regions and 54,728 observations on private sector workers we provide strong evidence on the existence of external returns to education. Our IV median estimate indicates a one percentage point increase in the ratio of college educated workers in a region, raises the "typical" worker's wages by 1.9% (3.1% females, 1.5% males). The comparable mean, namely IV-OLS, estimate is larger, 2.3%. However given the skewed and leptokurtic nature of our data, we believe ordinary least squares estimates will tend to exaggerate the typical impact. LAD or least absolute deviations estimation is more appropriate for such cases.

In addition LAD allows estimating the impact of the regressors on various quantiles of the wage distribution. This allows for a richer analysis. In fact our quantile estimates show that both internal and external returns to schooling increase with higher education levels. This implies rising wage inequality. Similar findings are reported for the US, European as well as Latin American and East Asian wage distributions since the 1990'es, Angrist and Pischke (2009), Lemieux (2008), Martins and Pereira (2004), Patrinos et al (2009). This phenomenon is ascribed to skill biased technical change associated with the computer revolution. The policy implications include training and education programs consistent with such changes in production technology.

Finally our quantile estimates for external effects indicate knowledge spillovers in Turkey, involving women are substantially larger than for men. Further research clarifying the validity of this pattern for comparable countries e.g. Balkan and Middle Eastern ones, would have great practical significance.

References

A. O. Cakmak. 2006. "Pattern of Intra-industry Trade in Manufacturing between Turkey and Germany, Italy, France and the UK: 1991-2004". *Ekonomik ve Sosyal Arastirmalar Dergisi*, 3(1):30–47.

Acemoglu, D. and J. Angrist. 2001. "How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws". In *NBER Macroeconomics Annual* 2000, ed. B. Bernanke and K. Rogoff. Cambridge: MIT Press, 9-59.

Akerlof, G and J. Yellen. 1990. "The Fair-Wage Effort Hypothesis and Unemployment". *Quarterly Journal of Economics* 105: 255–283.

Angrist, J. and J.S. Pischke. 2008. "*Mostly Harmless Econometrics: An Empiricist's Companion*". Princeton, NJ: Princeton University Press.

C. Sakellariou. 2001. "Identifying the External Effects of Human Capital: A Two-Stage Approach". *Applied Economics Letters* 8: 191–194.

Cameron, A.C. and P.K. Trivedi. 2005. "*Microeconometrics: Methods and Applications".* Cambridge, UK: Cambridge University Press.

Cameron, A.C. and P.K. Trivedi. 2009. "*Microeconometrics using Stata*". Texas: STATA Press.

Chernozhukov, V. and C. Hansen. 2006. "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models". *Journal of Econometrics* 132 (2):491–525.

Chernozhukov, V. and C. Hansen. 2008. "Instrumental Variable Quantile Regression: A Robust Inference Approach". *Journal of Econometrics* 142(1): 379–398.

D. Acemoglu. 2009. "*An Introduction to the Economics of Growth*". Princeton, NJ: Princeton University Press.

E. Moretti. 2004a. "Human Capital Externalities in Cities". Chapter 51, Vol. 4. In "*Handbook of Regional and Urban Economics*", ed. J.V. Henderson and J.F. Thisse. Amsterdam: Elsevier.

E. Moretti. 2004b. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data". *Journal of Econometrics* 121:175–212.

Galvao, A.F and G.M. Rojas. 2009. "Instrumental Variables Quantile Regression for Panel Data with Measurement Errors". *Discussion Paper* 09/06. London: Dept. of Economics, City University.

J. E. Rauch. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities". *Journal of Urban Economics* 34:380–400.

Kirby, S. and R. Riley. 2008. "The External Return to Education: UK Evidence using Repeated Cross Sections". *Labour Economics* 15:619–630.

Krueger, A. and M. Lindahl. 2001. "Education for Growth: Why and for Whom?". *Journal of Economic Literature* 39:1101–1136.

M. Buchinsky. 1998. "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research". *Journal of Human Resources* 33:88–126.

Martins, P.S. and P.T. Pereira. 2004. "Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 European Countries". *Labour Economics* 11:355–371

N. Davutyan. 2008. "Estimating the Size of Turkey's Informal Sector: An Expenditure Based Approach". *Journal of Economic Policy Reform* 11(4):261–271.

O. Eren. 2009. "Ability, Schooling Inputs and Earnings: Evidence from the NELS". Discussion Paper, Dept. of Economics, University of Nevada, Las Vegas.

Patrinos H. A, C. R Cano and C. Sakellariou. 2009. "A Note on Schooling and Wage Inequality in the Public and Private Sector". *Empirical Economics* 37(2):383–392.

R. E. Lucas. 1988. "On the Mechanics of Economic Development". *Journal of Monetary Economics* 22:3–42.

R. J. Barro. 1997. "*Determinants of Economic Growth: A Cross-Country Empirical Study*". Cambridge, MA: MIT Press.

R. Koenker. 2005. "*Quantile Regression*". Cambridge, UK: Cambridge University Press.

R. Winter-Ebmer. 1994. "Endogenous Growth, Human Capital and Industry Wages". *Oxford Bulletin of Economic Research* 46: 289–314.

Sakellariou, C. and R. Maysami. 2004. "Lucas Type External Effects of Human Capital: Strong Evidence Using Micro Data". *Applied Economics Letters* 11: 343–346.

T. Lemieux. 2008. "The Changing Nature of Wage Inequality". *Journal of Population Economics* 21:21–48.

Vandenbussche, J, P. Aghion and C. Meghir. 2006. "Growth, Distance to Frontier and Composition of Human Capital". *Journal of Economic Growth* 11: 97–127.

Table 1: Mean vs Median (Q=.05) Estimates for Social Returns. (p value <0.001 in every case**)**

	ALL	FEMALE	MALE
EDU ols	0.0479	0.0433	0.0488
EDU qr	0.0376	0.0343	0.0381
COLLEGE ols	0.0119	0.0195	0.0090
COLLEGE qr	0.0088	0.0174	0.0064
EDU iv-ols	0.0482	0.0433	0.0492
EDU iv-qr	0.0378	0.0342	0.0384
COLLEGE iv-ols	0.0230	0.0303	0.0193
COLLEGE iv-qr	0.0186	0.0310	0.0150
Mean logged wages (WAGE)	0.7534	0.6816	0.7730
Median logged wages (WAGE)	0.7340	0.6827	0.7340
# of workers	54,728	11,776	42,952

Table 2: Private returns to education (p value <0.001 in every case**)**

	Q=0.10	Q=0.25	Q=0.50	Q=0.75	Q=0.90
ALL, QR	0.0295	0.0319	0.0380	0.0511	0.0655
ALL,IVQR	0.0300	0.0323	0.0384	0.0513	0.0651
FEMALE,QR	0.0293	0.0314	0.0343	0.0466	0.0568
FEMALE,IVQR	0.0292	0.0314	0.0342	0.0457	0.0569
MALE,QR	0.0274	0.0304	0.0381	0.0520	0.0693
MALE,IVQR	0.0280	0.0306	0.0384	0.0528	0.0689

Table 3: External Returns to education (p value <0.001 in every case**)**

	Q=0.10	Q=0.25	Q=0.50	Q=0.75	Q=0.90
ALL, QR	0.0063	0.0075	0.0085	0.0112	0.0127
ALL,IVQR	0.0130	0.0156	0.0186	0.0250	0.0336
FEMALE,QR	0.0057	0.0117	0.0174	0.0249	0.0303
FEMALE,IVQR	0.0090	0.0170	0.0310	0.0370	0.0500
MALE,QR	0.0054	0.0059	0.0064	0.0081	0.0085
MALE,IVQR	0.0120	0.0135	0.0150	0.0205	0.0295

Table 4: The Impact of Trade OPENness and Amenities INDEX

	2SLS	Q=0.10	Q=0.25	Q=0.50	Q=0.75	Q=0.90
OPEN	0.000007	0.000008	0.000007	0.000007	0.000006	0.0000055
N= 54,728						
INDEX	***-0.0051***	0.0333	0.0782	0.0648*	***0.0364***	***-0.0093***
N=54,728						

bold italics indicate insignificance, * p<=0.01, No * indicates p value <= 0.000

Appendix 1: Sectoral Dummies

AGRICULTURE: Agriculture, forestry, hunting and fishing
MINING: Mining and quarrying
MANUFACTURING: Manufacturing
UTILITIES: Electricity, gas and water
CONSTRUCTION: Construction
TRADE: Wholesale and retail trade, restaurants and hotels
TRANSPORT: Transportation, communication and storage
FINANCE: Finance, insurance, real estate and business services
SERVICES: Community, social and personal services

Appendix 2

OLS: Dependent Variable: WAGE (logged wages)

	ALL	FEMALE	MALE
N	54728	11776	42952
AGE	.0340	.0302	.0349
AGE2	.0003	-.0003	-.0003
SOSEC	.2016	.2672	.1827
TENURE	.0205	.0236	.0194
TENURE2	-.0003	-.0006	-.0002
INDEX	.1427	.2022	.1497
UN_RATE	.0147	-.0110	-.0161
W_PLACE	.1532	.1439	.1574
OPEN	8.5e-06	9.6e-06	8.2e-06
EDU	.0484	.0433	.0488
COLLEGE	.0116	.0195	.0090
MINING	.3279	.5487	.3049
MNFCTR	.1577	.0635*	.1679
UTLTES	.3618	***-.0230***	.3654
CNSTRCTN	.2903	.1060*	.2758
TRADE	.0737	.0485**	.0642
TRNSPRT	.2310	.2916	.2095
FINANCE	.2275	.2641	.1911
SERVICES	.1791	.2499	.1017
FEMALE	**-**.0825	N/A	N/A
CONS	.8833	-1.0333	-.8387
Adj.R2	0.3756	0.4279	0.3627

2SLS:	ALL	FEMALE	MALE
AGE	.0335	.0297	.0345
AGE2	-.0003	-.0003	-.0003
SOSEC	.2028	.2661	.1845
TENURE	.0204	.0235	.0193
TENURE2	.0003	-.0006	-.0002
INDEX	***-.0051***	***.0602***	***.0089***
UN_RATE	.0179	-.0141	-.0192
W_PLACE	.1539	.1430	.1583
OPEN	7.1e-06	8.1e-06	6.9e-06
EDU	.0487	.0433	.0492
COLLEGE	.0226	.0302	.0193
MINING	.3332	.5503	.3104
MNFCTR	.1578	.0660*	.1681
UTLTES	.3599	***-.0295***	.3635
CNSTRCT	.2866	.0960*	.2733
TRADE	.0720	.0486**	.0629
TRNSPRT	.2255	.2875	.2049
FINANCE	.2207	.2602	.1846
SERVICE	.1729	.2448	.0967
FEMALE	-.0812	N/A	N/A
CONS	.9005	-1.0444	-.8570
Adj.R2	0.3735	0.4263	0.3608

No * p <= 0.000, * p<=0.01, **p<=0.05, ***p<=0.1, ***bold italics indicate insignificance***

Appendix 3

Full Sample N=54728

Dependent Variable: WAGE (logged wages)

ORD-QR	**Q=0.1**	**Q=0.25**	**Q=0.5**	**Q=0.75**	**Q=0.9**
AGE	.0526	.0429	.0354	.0299	.0185
AGE2	-.0006	-.0005	-.0003	-.0002	-.0000
SOSEC	.3399	.2258	.1656	.1252	.1078
TENURE	.0119	.0135	.0188	.0245	.0277
TENURE2	-.0002	-.0002	-.0003	-.0003	-.0004
INDEX	.1615	.1970	.1813	.1788	.1779
UN_RATE	-.0070	-.0093	-.0117	-.0156	-.0197
W_PLACE	.1326	.1145	.1087	.1245	.1646
OPEN	8.6e-06	8.1e-06	8.0e-06	8.0e-06	8.6e-06
EDU	.0295	.0319	.0380	.0511	.0655
COLLEGE	.0063	.0075	.0085	.0112	.0127
MINING	.5209	.4103	.3516	.2810	.1334*
MNFCTR	.3927	.2708	.1670	.0665	-.0536*
UTLTES	.5476	.3842	.3289	.3093	.2512*
CNSTRCTN	.4440	.3404	.284	.2329	.1817
TRADE	.2890	.1544	.0644	-.0254***	-.1388
TRNSPRT	.3782	.2676	.2066	.1533	.1219
FINANCE	.3910	.2758	.2033	.1620	.1475
SERVICES	.2465	.1809	.1481	.1481	.1706
FEMALE	-.0678	-.0586	-.0713	-.0907	-.0849
CONS	-1.787	-1.210	-.7645	-.4466	***-.0204***
Adj. R2	0.2548	0.2189	0.1924	0.2155	0.2490

IV-QR	**Q=0.1**	**Q=0.25**	**Q=0.5**	**Q=0.75**	**Q=0.9**
AGE	0.0525	0.0427	0.0353	0.0290	0.0182
AGE2	-0.0006	-0.0005	-0.0004	-0.0003	-0.0001**
SOSEC	0.3410	0.2252	0.1666	0.1300	0.1080
TENURE	0.0120	0.0135	0.0187	0.0246	0.0280
TENURE2	-0.0002	-0.0002	-0.0003	-0.0004	-0.0005
INDEX	0.0333	0.0782	0.0648*	***0.0364***	***-0.0093***
UN_RATE	-0.0094	-0.0122	-0.0142	-0.0189	-0.0233
W_PLACE	0.1315	0.1190	0.1117	0.1235	0.1707
OPEN	7.9e-06	7.2e-06	6.9e-06	6.1e-06	5.5e-06
EDU	0.0300	0.0323	0.0384	0.0513	0.0651
COLLEGE	0.0130	0.0156	0.0186	0.0250	0.0336
MINING	0.5295	0.4205	0.3654	0.2870	0.1439
MNFCTR	0.3926	0.2631	0.1695	0.0692	-0.0577**
UTLTES	0.5444	0.3731	0.3318	0.3179	0.2133**
CNSTRCTN	0.4420	0.3326	0.2850	0.2299	0.1758
TRADE	0.2878	0.1483	0.0678	***-0.0278***	-0.1458
TRNSPRT	0.3825	0.2574	0.2030	0.1501	0.0984*
FINANCE	0.3898	0.2626	0.2019	0.1551	0.1179
SERVICES	0.2472	0.1694	0.1454	0.1462	0.1473
FEMALE	-0.0651	-0.0552	-0.0701	-0.0907	-0.0909
CONS	-1.7880	-1.2100	-0.8018	-0.4824	-0.1066*
Adj.R2	0.2381	0.1956	0.1701	0.1907	0.2245

No * p <= 0.000, * p<=0.01, **p<=0.05, ***p<=0.1, ***bold italics indicate insignificance***

Appendix 4

Only Women N=11776

Dependent Variable: WAGE (logged wages)

ORD-QR	**Q=0.1**	**Q=0.25**	**Q=0.5**	**Q=0.75**	**Q=0.9**
AGE	.0248	.0346	.0314	.0272	.0282
AGE2	-.0002	- .0004	- .0003	- .0002	- .0002**
SOSEC	.5449	.3442	.236	.1434	.0707
TENURE	.0165	.0182	.0208	.0279	.0332
TENURE2	-.0008*	-.0006	- .0005	- .0005	-.0005*
INDEX	.3741	.2880	.2280	.1697	.1708**
UN_RATE	***- .0028***	-.0073	- .0091	-.0149	- .0202
W_PLACE	.1357	.1098	.0803	.0872	.0992
OPEN	.0000	8.9e-06	7.6e-06	7.0e-06	7.9e-06
EDU	.0293	.0313	.0342	.0465	.0568
COLLEGE	.0056*	.0117	.0174	.0248	.0302
MINING	.4017***	.5689**	.5193***	.6181*	.4192*
MNFCTR	.1857	.1763	.1227	***.0372***	- .1222
UTLTES	.4386	.2738	***.0390***	- .2225	- .5713
CNSTRCTN	.2828	.2356	.1486*	***.0457***	- .1668**
TRADE	.1625*	.1168*	.0459***	***- .0285***	- .1656
TRNSPRT	.2726	.2640	.2309	.2779	.3679
FINANCE	.3186	.2898	.2410	.2430	.2096
SERVICES	.2520	.2241	.2217	.2373	.2576
CONS	-1.5495	-1.2783	- .8995	-.6073	- .3427
Adj.R2	0.3079	0.2742	0.2263	0.2506	0.2970

IV-QR	**Q=0.1**	**Q=0.25**	**Q=0.5**	**Q=0.75**	**Q=0.9**
AGE	.0261	.0344	.0329	.0278	.0276
AGE2	-.0003	-.0004	- .0009	- .0002	- .0002**
SOSEC	.5481	.3485	.2370	.1434	.0643
TENURE	.0169	.0193	.0191	.0267	.0333
TENURE2	-.0008*	-.0007	- .0004	- .0005	-.0005
INDEX	.2972	.2126	.0824**	***.0613***	***- .0224***
UN_RATE	***- .0041***	-.0093	-.0120	- .0173	-.0256
W_PLACE	.1340	.1069	.0812	.0894	.0976
OPEN	9.8-06	8.0e-06	5.5e-06	5.1e-06	4.9e-06
EDU	.0291	.0313	.0342	.0457	.0569
COLLEGE	.0090	.0170	.0310	.0370	.0500
MINING	.4072***	.5590**	.5054***	.6172**	.4143**
MNFCTR	.1907	.1797	.1254	***.0400***	- .1278
UTLTES	.4543	.2890	***.0385***	- .2898	-.5833
CNSTRCTN	.2909	.2341	.1311**	***.0428***	- .1511**
TRADE	.1687	.1173	.0526**	***- .0227***	-.1643
TRNSPRT	.2751	.2628	.2280	.2787	.3496
FINANCE	.3304	.2938	.2453	.2405	.1754
SERVICES	.2561	.2193	.2128	.2366	.2227
CONS	-1.5656	-1.2835	- .9747	-.6614	- .3918
Adj.R2	0.2947	0.2446	0.1863	0.2016	0.2441

No * p <= 0.000, * p<=0.01, **p<=0.05, ***p<=0.1, ***bold italics indicate insignificance***

Appendix 5

Only Men N=42952

Dependent Variable: WAGE (logged wages)

ORD-QR	**Q=0.1**	**Q=0.25**	**Q=0.5**	**Q=0.75**	**Q=0.9**
AGE	.0621	.0473	.0364	.0296	.0148
AGE2	-.0007	-.0005	-.0004	-.0002	***-.0000***
SOSEC	.2732	.1869	.1468	.1249	.1197
TENURE	.0095	.0120	.0182	.0236	.0263
TENURE2	-.0001**	-.0001	-.0002	-.0003	-.0004
INDEX	.1491	.1954	.1905	.1935	.1948
UN_RATE	-.0094	-.0100	-.0128	-.0165	-.0198
W_PLACE	.1314	.1235	.1209	.1334	.1761
OPEN	8.2e-06	8.0e-06	8.0e-06	7.9e-06	8.2e-06
EDU	.0273	.0304	.0380	.0520	.0693
COLLEGE	.0054	.0059	.0063	.0080	.0085
MINING	.4902	.3942	.3295	.2666	.1160*
MNFCTR	.3825	.2676	.1673	.0840	***-.0171***
UTLTES	.5445	.3834	.3372	.3106	.2366
CNSTRCTN	.3899	.3120	.2714	.2417	.2141
TRADE	.2520	.1394	.0579	***-.0190***	-.1082
TRNSPRT	.3258	.2368	.1897	.1501	.1302
FINANCE	.3490	.2334	.1683	.1199	.1094
SERVICES	.1467	.1055	.0788	.0829	.0933
CONS	-1.803	-1.2057	-.7338	-.4118	***.0383***
Adj.R2	0.2415	0.2054	0.1844	0.2068	0.2407

IV-QR	**Q=0.1**	**Q=0.25**	**Q=0.5**	**Q=0.75**	**Q=0.9**
AGE	.0610	.0467	.0368	.0290	.0145
AGE2	-.0007	-.0005	-.0004	-.0002	***-.0000***
SOSEC	.2723	.1893	.1480	.1270	.1202
TENURE	.0098	.0122	.0178	.0235	.0266
TENURE2	-.0001**	-.0001	-.0002	-.0003	- .0004
INDEX	***.0303***	.0813	.0811	***.0468***	- ***.0062***
UN_RATE	-.0119	-.0128	-.0148	-.0193	- .0236
W_PLACE	.1309	.1228	.1236	.1375	.1795
OPEN	7.5e-06	7.1e-06	6.9e-06	6.3e-06	5.4e-06
EDU	.0280	.0306	.0383	.0528	.0689
COLLEGE	.0120	.0135	.0150	.0205	.0295
MINING	.5027	.4052	.3355	.2718	.1438
MNFCTR	.3949	.2699	.1640	.0763	***-.0227***
UTLTES	.5694	.3628	.3270	.3078	.2592
CNSTRCTN	.3934	.3105	.2696	.2344	.2064
TRADE	.2602	.1408	.0538	***-.0259***	-.1221
TRNSPRT	.3367	.2360	.1788	.1370	.1108
FINANCE	.3500	.2302	.1627	.1105	.0814*
SERVICES	.1572	.0996	.0705	.0675*	***.0630***
CONS	-1.8015	-1.208	-.7673	-.4480	-.0463
Adj.R2	0.2260	0.1860	0.1666	0.1877	0.2221

No * p <= 0.000, * p<=0.01, **p<=0.05, ***p<=0.1, ***bold italics indicate insignificance***