IFAU – INSTITUTE FOR LABOUR MARKET POLICY EVALUATION

External effects of education on earnings:

Swedish evidence using matched employee-establishment data

Gunnar Isacsson

WORKING PAPER 2005:10

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Industry, Employment and Communications, situated in Uppsala. IFAU's objective is to promote, support and carry out: evaluations of the effects of labour market policies, studies of the functioning of the labour market and evaluations of the labour market effects of measures within the educational system. Besides research, IFAU also works on: spreading knowledge about the activities of the institute through publications, seminars, courses, workshops and conferences; creating a library of Swedish evaluational studies; influencing the collection of data and making data easily available to researchers all over the country.

IFAU also provides funding for research projects within its areas of interest. There are two fixed dates for applications every year: April 1 and November 1. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The authority has a traditional board, consisting of a chairman, the Director-General and eight other members. The tasks of the board are, among other things, to make decisions about external grants and give its views on the activities at IFAU. A reference group including representatives for employers and employees as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

External effects of education on earnings: Swedish evidence using matched employee-establishment data♠

by

Gunnar Isacsson♣

15 April, 2005

Abstract

This paper provides an empirical investigation of externalities from education in Sweden in an earnings equation framework. The empirical models are estimated on a large sample of matched employees and establishments. External effects of education are identified from the average educational attainment of workers outside the individual's establishment. The paper also investigates the coherence of the evidence with respect to the idea that educational externalities arise through face-to-face interaction between individuals. A set of different specifications and fixed effects models is used to investigate the robustness of the basic cross-sectional model. The cross-sectional models suggest, in general, that externalities are positive and significantly different from zero. The cross-sectional evidence is also broadly coherent with the idea that externalities are declining in spatial transaction costs, such as the Euclidean distance between establishments. However, after accounting for individual fixed effects and dummy variables for the county in which the individual works the results indicate no statistically significant external effects of education on earnings in Sweden.

♠ Funding from the Institute for Labor Market Policy Evaluation and VINNOVA (the Swedish Agency for Innovation Systems), is gratefully acknowledged. Anders Björklund, Mikael Lindahl, Erik Mellander, Jan-Eric Nilsson, seminar participants at the Institute for Labor Market Policy Evaluation, the Swedish Institute for Social Research and Dalarna University provided valuable comments on previous drafts of this manuscript. Helge Bennmarker and Fridtjof Thomas provided valuable programming assistance. Any remaining shortcomings are, of course, my own responsibility.

♣ National Road and Transportations Research Institute, Box 760, 781 27 Borlänge, Sweden

Table of contents

1. Introduction

The idea that localized human capital externalities may have positive effects on economic growth seems to have gained widespread attention among economists in recent years. It has also been hypothesized that such beneficial effects of human capital are likely to operate more efficiently in dense economic environments such as cities (see for example Lucas, 1988, and Jacobs, 1969). The motivation is that face-to-face interaction between individuals is likely to be more frequent in geographical areas with high population or employment density and that it is easier to find the "right" individuals to interact with in such areas. Such interactions may, in turn, be an important channel through which new ideas and knowledge are transmitted in the economy. There are different theoretical formalizations of such ideas (see Duranton & Puga, 2004, for a recent survey of theoretical work in this field). The idea seems, however, to date back at least to Marshall (1920).

As with all kinds of externalities, the issue of human capital externalities is obviously policy relevant. It relates to the question if, and by how much, education should be subsidized. In light of this relevance, it is somewhat surprising to note that empirical work on this issue has been relatively scant, at least relative to the large literature on private economic returns to human capital.[1] This paucity of empirical work might partly be due to the fact that "[k]nowledge flows ... are invisible; they leave no paper trail by which they may be measured and tracked...", as Krugman (1991b, p 53) puts it.[2] In other words, human capital externalities are hard to observe in data so it may be difficult to model them empirically.

One feasible empirical approach is based on wage equations. In a pioneering paper Rauch (1993) demonstrates that conditional on own educational attainment, the wage of an individual is also positively related to the average educational level in the metropolitan area in which he/she lives.[3] Thus, even though human capital externalities "leave no paper trail", Rauch's framework may be useful to measure, at least indirectly, human capital externalities.[4]

---

[1] See Moretti (2004a) for a survey of empirical research on human capital externalities.

[2] However, Jaffe et al (1993) argue convincingly that knowledge flows sometimes do leave a paper trail in the form of patent citations.

[3] Rauch (1993) also demonstrates that housing costs are increasing in the average educational level of a metropolitan area's workforce.

[4] A related approach is based on so-called Macro-Mincer equations, which are employed for cross-country comparisons of gross national incomes rather than data on individuals living in different geographical regions (see, for example, Krueger & Lindahl, 2001).

There are some potential drawbacks of the model used by Rauch, however. The first is that he treats the average educational attainment as an exogenous variable in the wage equation. There is, however, a risk that causation also runs from high wages to high educational attainment, and that relevant variables have been omitted from the regression equation. Thus, it is not surprising that many of the papers following Rauch have focused on the possible endogeneity of the average level of schooling. Acemoglu & Angrist (2000), Moretti (2004b), and Ciccone & Peri (2004) all use some kind of instrumental variable approach to account for this potential problem of the analysis. Chosen instruments are, for example, the time-lagged age-structure in the geographical unit of observation and time-lagged differences in schooling legislation between different geographical units of observations. The results of these papers are mixed. Some report statistically significant external returns whereas others report no statistically significant results.

A second drawback of Rauch's model, noted by Ciccone & Peri (2004) and Moretti (2004b) is that a positive association between the wage of an individual and the average educational attainment in his/her labor market may also be due to imperfect substitutability between workers with different educational levels. In a conventional demand-supply framework, an increase in the share of highly educated workers in a labor market will lower the wage of such workers (in the absence of positive human capital externalities) but raise the wage of lower educated workers when the two types of workers are complements in production. Thus, the net effect on the average wage may be positive even in the absence of human capital externalities.[5] After controlling for imperfect substitut-

[5] Ciccone & Peri (2004) propose an empirical approach that allows the identification of human capital externalities even when workers with different levels of educational attainment are imperfect substitutes in production. The approach is based on the observation that an increase in average human capital of the work-force has no first-order effects on the average wage in a geographical area when wages reflect marginal social products, and the skill-composition of the work-force in the geographical area is held constant. But the increase will have a first-order effect whenever wages of high-skilled workers are below the corresponding marginal social product, and the skill-composition is held constant. Hence, Cicconce & Peri suggest that the external effects of education on wages is identified through the change in average educational attainment between different points in time in a geographical area, holding the weights of the skill-composition in the geographical area constant. Based on this constant-composition approach they find no significant effects of aggregate human capital on wages. Moretti (2004b), on the other hand, argues that if an increase in the share of college educated workers in a city increases the wages of such workers, the effect of human capital externalities may be strong enough to dominate the negative supply-effect on wages. His results suggest that the positive external effects to education are indeed significantly different from zero. However, this evidence is also in line with an explanation relying on skill-biased technological change, where technical change may have a positive effect on wages of high-skilled workers which is large enough to dominate negative ef-

ability in the estimation framework Ciccone & Peri find no statistically significant results of external returns to education.

One restriction of most previous work on external returns to education is that the geographical scope of externalities is dictated by the chosen geographical unit in the analysis. In general, it is assumed that externalities come to a halt at the administrative borders of the geographical unit, and that they are homogeneous within the geographical unit. In other words, distance and transport costs between individuals or production plants are not explicitly modeled. Such costs may obviously be relevant if face-to-face interactions are thought to be important for the transmission of new ideas, skills and knowledge. An exception is a recent investigation by Conley et al (2003) who let externalities depend on the geographical distance between individuals. They find that the external effects are indeed larger for individuals that are located close to each other.[6] The explicit modeling of transport costs may, in addition, also be relevant from a policy perspective, since one major cost regarding the potential face-to-face interaction between individuals is the cost of transporting people (cf Glaeser & Kohlhase, 2004).[7]

The purpose of the present paper is to estimate human capital externalities in an earnings equation framework using data from Sweden. The analyses are based on a large longitudinal sample of matched employees and establishments with detailed information on the educational attainment of the work-force at the establishment level. The analyses also use information on all establishments in Sweden, so that the educational attainment of the workforce in the local labor markets may be described in a relevant way. To investigate whether individuals' face-to-face interaction are important in this framework, the analyses also add a spatial structure to the empirical models. Geographical coordinates (lon-

---

fects on wages of an increased supply of high-skilled workers (see Berman, et al, 1994, for an investigation on US data, or Mellander, 1999, for an investigation on Swedish data).

[6] Using another metric for distance, Moretti (2004c) tests the coherence of the evidence in a similar spirit by investigating whether the magnitude of estimated externalities depends on the economic distance between plants. By using input-output tables to measure the degree of interaction between plants, he finds that educational externalities, or spillovers, are larger for plants that often interact and smaller for plants with relatively little interaction with each other.

[7] The increased use of information technology in the last decade has lowered the cost of contacting people; e g through electronic mail and the internet. This might suggest that the cost of transporting people would tend to lose some of the policy relevance related to facilitating face-to-face interaction. However, Gaspar & Glaeser (1998) note that this reasoning hinges on the assumption that an electronic contact is a substitute for a face-to-face contact. If it instead is a complement, the increased use of information technology might actually increase the policy relevance of the cost of transporting people. Gaspar & Glaeser also provide some evidence from a previous revolution in information technology, namely the introduction of telephones, which suggest that electronic contact actually may be a complement to face-to-face interaction.

gitudes and latitudes) of the establishments' locations are available in the data, which implies that the geographical structure of the local economy can be represented in a flexible way.

The paper extends and complements the previous literature in some respects. Firstly, it demonstrates that matched employee-establishment data may be useful for dealing with imperfect substitutability when estimating the external returns to education in a wage or earnings equation framework. The use of establishment data resembles the approach of Moretti (2004c) who uses plant level data to estimate the external return to education. However, Moretti (2004c) estimates the external return to education in a production framework rather than in an earnings equation framework. Secondly, the present paper adds spatial structure to the empirical models to see if the evidence on externalities is 'coherent' with the idea that face-to-face interactions are relevant for the transmission of externalities.[8] In this sense it resembles the investigation in Conley et al (2003). But unlike Conley et al, the data used here make it possible to address the issue of imperfect substitutability between workers with different levels of education directly. Finally, in light of the mixed results in previous work, an investigation of external effects of education on Swedish data might be informative regarding such effects in other countries too. The reason is that the private economic returns to education are relatively low in Sweden compared to other developed countries. This might suggest a relatively large difference between the private and the social economic returns to education. Thus, if the Swedish evidence would indicate no external effects of education on earnings, this might be suggestive also for other developed countries.

The paper is organized as follows. The theoretical motivations of the empirical models are provided in section 2. The empirical models are presented in section 3 together with some econometric issues. Section 4 contains a description of the data and section 5 presents the results. Section 6 concludes.

[8] See Rosenbaum (2002) pp 5-6 and chapter 9 for a discussion of 'coherence' in observational studies.

2. Theoretical framework

This section first outlines a basic theoretical framework to measure educational externalities (section 2.1). It then turns to discuss geographical issues regarding the transmission of externalities and how they may be modeled (section 2.2). Finally, it discusses other explanations of the spatial variation in earnings that may be relevant to consider when measuring educational externalities (section 2.3).

2.1 The basic model[9]

Assume that there are no other factors of production but K different kinds of workers characterized by their educational attainment, and that the production technology is characterized by constant returns to scale at the level of the firm. Assume that in a given geographical area g ($g = 1, 2, ..., G$) all establishments e ($e = 1, 2, ..., E_g$) have access to the following production function at time t ($t = 1, 2, ..., T$).

$$Y_{egt} = A_{gt} * F\left(N_{e1t}, N_{e2t}, ..., N_{eKt}\right)$$

where Y_{egt} is the output which is traded in a perfectly competitive market at a price normalized to one; A_{gt} denotes the total factor productivity in area g, assumed to be taken as exogenous by the establishment; and N_{ekt} ($k = 1, 2, ..., K$) is the number of workers employed in the establishment having attained educational level k.

Educational externalities are assumed to enter the model through total factor productivity in area g. Specifically, assume that

$$A_{gt} = \exp\left(\theta_{0t} + \theta_1 S_{-egt}\right), \qquad \text{(TFP1)}$$

where S_{-egt} is a measure of the average educational attainment of the workforce in area g outside establishment e, and θ_1 measures the external effect of education. Note that the externality is assumed to affect total factor productiv-

[9] The theoretical framework presented in this section is basically an adaptation of the theoretical models presented by, for example, Rauch (1993) or Moretti (2004a), to the available data of the present paper.

ity in a Hicks neutral way. The parameter θ_{0t} captures the effect of other time varying variables at the national level on total factor productivity.

Profit maximization implies that $A_{gt}F_k\left(N_{e1t}, N_{e2t}, ..., N_{eKt}\right) = w_{gkt}$ for all *e* and *k* where F_k denotes the first-order partial derivative of $F(\bullet)$ with respect to the *k*th argument and w_{gkt} is the wage of workers with educational level *k* in area *g* at point in time *t*. Taking the natural logarithm of the equilibrium wage for each type of worker *k* in area *g* gives

$$\ln w_{gkt} = \ln A_{gt} + \ln F_k\left(N_{e1t}, N_{e2t}, ..., N_{eKt}\right)$$

A first-order Taylor approximation of this relation around the national symmetric equilibrium gives for each type of worker *k* employed in establishment *e* and area *g*

$$\ln w_{gkt} \approx \ln A_{gt} + F_k\left(\overline{N}_{1t}, \overline{N}_{2t}, ..., \overline{N}_{Kt}\right) + \frac{F_{1k}}{F_k}\left(N_{e1t} - \overline{N}_{1t}\right) + \frac{F_{2k}}{F_k}\left(N_{e2t} - \overline{N}_{2t}\right)$$
$$+ ... + \frac{F_{Kk}}{F_k}\left(N_{eKt} - \overline{N}_{Kt}\right)$$

where $\overline{N}_{kt}$ (k = 1, 2, ..., K) is the national average number of workers having attained educational level *k* with the average taken over all establishments; F_{lk} (l = 1, 2, ..., K) denotes the partial derivative of F_k with respect to the *l*th argument. Diminishing marginal productivity implies that $F_{kk} \leq 0$, but the second-order cross partial derivatives may be either positive or negative depending on whether two types of workers are complements or substitutes in production.

Thus, the equilibrium wage within the geographical area suggests that the marginal product of each type of worker is equalized across establishments located in the same geographical area. However, if workplace related amenities affect the utility of working in a specific establishment, the equilibrium wage may also vary between establishments in the same area to compensate for such amenities according to the standard theory of compensating wage differentials (see for example Rosen, 1986).

Assuming that workers are mobile, the spatial equilibrium distribution of wages and workers is characterized by equalizing utilities across workers in different locations (cf Roback, 1982 and 1988, Rauch, 1993, and Moretti, 2004b). Hence, nominal wages vary across space, since other variables affecting the individual's utility, such as costs of living and local amenities, also vary across space. The spatial equilibrium distribution of establishments is characterized by a zero profit condition for establishments producing the traded good. Therefore, if nominal wages tend to be higher in areas with high levels of human capital, there has to be some advantages to the firm of locating production in such a place to motivate the higher nominal wage costs in these labor markets. Otherwise, it would be profitable to relocate production to other geographical areas. This is also the main motivation for using the nominal wage rather than a wage that is adjusted for geographical differences in the cost of living as the dependent variable in the empirical analysis.

In sum, this basic model suggests that the spatial variation in the equilibrium wage of a worker with educational attainment *k* is due to variation in total factor productivity and to variation in the educational composition of the workforce at the individual's establishment; that is, variation in the vector $\left(N_{e1t}, N_{e2t}, \ldots, N_{eKt}\right)$. Obviously, there are other explanations of this variation, some of which are discussed in section 2.3.

2.2 The geographical scope of externalities

The framework outlined in the previous section does not explicitly model the mechanism through which externalities work. If face-to-face interactions are relevant for the transmission of externalities it seems reasonable to provide some spatial structure to model the costs of these interactions (as in Conley et al, 2003). This section outlines two alternative ways of doing this.

Alternative 1

If individuals cross administrative borders to interact with people outside the geographical area in which the own establishment is located, it seems reasonable to assume that educational externalities cross the same borders. A variable that may be described as a measure of the accessibility to human capital in other geographical areas is used to capture such cross-border effects of educational externalities. This variable will be called the transport-cost-weighted educational attainment ($TCWS_{gt}$) for short in the following. It is given by

$$TCWS_{gt} = \sum_{j \neq g}^{G} d_{gjt}^{-1} S_{jt}$$

where d_{gjt} is a measure of the costs of transporting people from region *g* to region *j* including time costs; and S_{jt} is the average educational attainment of individuals employed in municipality *j*.

Thus, spatial structure is added to equation 1 by modeling the total factor productivity of establishment *e* in area *g* as

$$\ln A_{egt} = \theta_{0t} + \theta_1 S_{-egt} + \theta_2 TCWS_{gt} \qquad \text{(TFP2)}$$

where the parameter θ_1 is used to capture the externalities arising within the own geographical area, and θ_2 is used to assess the effect on total factor productivity of accessibility to human capital in the surrounding geographical areas. If this accessibility is relevant we should expect that the estimate of θ_2 is positive.

Alternative 2

If the data set contains information on the geographical coordinates; i e, the longitude and the latitude of the establishments, another way of adding spatial structure to the model in a relatively flexible manner is to create a set of circles each with a different radius around each establishment. It is then possible to obtain the average educational attainment of the work-force within the borders of each pair of circles; i e, to measure the educational attainment of the work-force at different distances from the establishment. This implies that the empirical analysis does not have to rely on some predetermined geographical unit of observation.

More formally, it is possible to model total factor productivity at each geographical location as follows

$$A_{et} = \exp\left(\theta_{0t} + \theta_1 S_{-et}^{(1)} + \theta_2 S_{-et}^{(2)} + \theta_3 S_{-et}^{(3)} + \theta_4 S_{-et}^{(4)}\right). \qquad \text{(TFP3)}$$

where $S_{-et}^{(d)}$ (d = 1, 2, ..., 4) is the average educational attainment of the work-force outside establishment *e* at a maximum distance of, say, 5 kilometers (d = 1), between 5 and 10 kilometers (d = 2), between 10 and 15 kilometers (d = 3), and between 15 and 20 kilometers (d = 4). Hence, if distance matters to the spread of educational externalities we would expect to see that the effect decays with distance, i e, $\theta_1 > \theta_2 > \theta_3 > \theta_4$. Note that here it is irrelevant in which geographical area the establishment is located, only the coordinates of the establishment are relevant. For this reason the subscript relating to geographical area (*g*) is dropped in TFP3.

Each of these two alternative ways of adding spatial structure to the model has some distinct advantages. There are two advantages of TFP2 compared to TFP3: (1) If the cost of transporting people is relevant to the transmission of educational externalities, it might be relevant from a policy perspective to directly address such costs when adding spatial structure to the model; and (2) it is likely that transport costs provide a more relevant metric for measuring he interaction between individuals in different geographical areas than the Euclidean distance used in TFP3. A potential advantage of TFP3 as compared to TFP2 is, however, that it does not rely on administratively set borders between the geographical units of observation

2.3 Other explanations of the spatial variation in wages

There are, of course, other explanations of the spatial variation in wages and earnings which are relevant to address in an empirical assessment of educational externalities. Beside knowledge spillovers, Marshall (1920) suggested that large markets provide demand and supply linkages that may benefit production. This idea is also the point of departure of the so-called "new economic geography" where agglomeration of economic activity is explained as an interaction between increasing returns to scale at the level of the firm, transportation costs and labor mobility.[10]

Demand linkages may be relevant both within a geographical area and between geographical areas. Linkages within an area may be summarized by a measure of the "economic size" of the area itself, such as the total employment in the area as suggested by Ciccone & Peri (2004), for example. Demand linkages between geographical areas may be summarized by the so-called "market potential", which has a long tradition in empirical regional economics (see Harris, 1954, for an early example). The recent work in the "new economic geography" has also provided the market potential with a solid theoretical underpinning. The market potential of a geographical area is basically a weighted sum of the purchasing power of the surrounding geographical areas, where the weights are related to the transport costs to each one of these areas. Hanson (2004) presents the first empirical investigation of the market potential relying on "structural" estimation equations, and the results are supportive for the new theoretical work on economic geography. The present paper does not provide a structural estimation framework for the market potential since the market po-

[10] See Krugman (1991a and 1991b) for early work in this field or the textbook by Fujita et al (1999) or the surveys by Fujita & Thisse (1996) and Ottaviano & Puga (1998).

tential itself is not in the focus of the paper. Instead it relies on reduced form versions of the market potential.

Measures of demand linkages tend to be strongly correlated with the educational attainment of the labor force. This underscores the relevance of controlling for the market potential in an empirical investigation of educational externalities. When adding measures of demand linkages within and between geographical areas to the empirical model, it seems reasonable to introduce them in a way which is consistent with the model of the educational externality, however. The motivation is that this avoids a possible misspecification of the empirical model that may bias the results in favor of one of the explanations for the spatial variation in wages and earnings. Hence, when TFP1 in section 2.1 is used to model the externality it seems reasonable to include a control variable for the size of the geographical area's internal market and not include a control variable that reflects linkages between geographical units of observation. The reason is that TFP1 constrains the externality to zero between different geographical areas. Ciccone & Peri (2004) suggest that the size of the internal market may be modeled by total employment in the area. In the present paper I basically follow their approach but exclude those employed in the establishment where the individual works.[11]

Furthermore, when modeling the spatial scope of the externality as in TFP2 in section 2.2, the market potential is modeled in a similar spirit as

$$MP_{egt} = \lambda_1 Z_{-egt} + \lambda_2 \left(\sum_{j \neq g}^{G} d_{gjt}{}^{-1} Z_{jt} \right),$$

where Z_{-egt} is total employment in the geographical area *g* outside the own establishment; Z_{jt} is total employment in geographical area *j*; d_{gjt} is a measure of transport costs between geographical areas *g* and *j* at point in time *t*.

Finally, when modeling the geographical scope of educational externalities as in TFP3 in section 2.2, it seems reasonable to model the market potential in the same vein. For this reason the following reduced form market potential function is used together with TFP3 in the empirical model

[11] Usually, it is total income that matters to aggregate demand. Using total income in an earnings equation framework probably creates an endogeneity problem, however, since individual earnings are part of total income. Hence, I choose to use employment instead of income when modeling demand linkages within and between geographical areas.

$$MP_{et} = \lambda_1 Z_{-et}^{(1)} + \lambda_2 Z_{-et}^{(2)} + \lambda_3 Z_{-et}^{(3)} + \lambda_4 Z_{-et}^{(4)}$$

where $Z_{-et}^{(d)}$ (d = 1, 2, ..., 4) is total employment outside establishment e at a maximum distance of, say, 5 kilometers (d = 1), between 5 and 10 kilometers (d = 2), between 10 and 15 kilometers (d = 3), and between 15 and 20 kilometers (d = 4). Hence, if transport costs are relevant to demand linkages we would expect that $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$.

Another explanation of the spatial variation in wages and earnings is spatial variation in the industry structure. The industry structure may in turn vary because of differences in natural advantages of production (see Kim, 1995 and 1999, and Ellison & Glaeser, 1999). One obvious example is mining. In addition, spatial variation in wages may result from spatial variation in individual-specific characteristics related to human capital accumulation and labor market attachment, for example: age, age-squared and the sex of the individual. Hence, it seems relevant to control for the industry classification of the establishments and observable individual-specific characteristics in the empirical models.

3. Empirical models

This section starts with a presentation of the basic empirical model in section 3.1, and then turns to a presentation of the models where more spatial structure is added to the model in section 3.2. Endogeneity problems of these models are subsequently discussed in section 3.3.

3.1 The basic empirical model

The basic empirical model of the present paper is a variant of the model presented by Rauch (1993), but it circumvents the problem of imperfect substitutability of workers with different levels of educational attainment by presuming disaggregate data on individuals and the establishments in which they work. With this type of data it is possible to let the marginal product of each type of worker depend both on the average educational attainment of the work-force outside the individual's establishment, and on the educational attainment of the workforce in the establishment (as suggested in section 2.1). The external effect of education is thus identified with information on the educational attainment of the workforce outside the individual's establishment.[12]

The explanatory variables of the model have been motivated in section 2. The basic theoretical model of section 2.1 also suggested that the parameters related to the educational attainment of the individual's co-workers depended on the educational attainment of the individual. Hence, some parameters in the benchmark model will be educational-specific. The benchmark empirical model of this paper is given by

$$y_{iegt} = \theta_{0t} + \theta_1 S_{-egt} + \alpha_k + \mathbf{N}_{egt}\mathrm{B}_k + Z_{-egt}\lambda + \mathbf{W}_{egt}\Lambda + \mathbf{X}_{it}\Phi + \varepsilon_{iegt} \quad (1)$$

where y_{iegt} is the natural logarithm of annual earnings of individual i (i = 1, 2,…, I_e) working in establishment e (e = 1, 2, ..., E_g), in geographical area g (g = 1, 2, ..., G) in year t (t = 1, 2, ..., T) with educational attainment k; S_{-egt} is the average educational attainment of the work-force in establishments outside the own establishment; $\mathbf{N}_{egt}$ is the vector of the educational attainment of the workforce in the individual's establishment, i e, the number of workers in the individual's establishment that has attained each educational level (k = 1, 2, ..., K); Z_{-egt} is total employment in the geographical area outside the own establishment; $\mathbf{W}_{egt}$ contains a set of industry dummies; and $\mathbf{X}_{it}$ contains the individual-specific characteristics: age, age-squared and a dummy variable that equals one if the individual is male.

Note that the educational-level-specific intercept α_k in equation 1 may be interpreted as the national average marginal product of k level workers. This estimate is, however, obtained by including dummy variables for own educational attainment using the lowest educational level as the reference category. Therefore the estimates of the private return to educational levels are obtained as the relative difference between the lowest educational level and the own educational level. In addition, the vector $\mathbf{B}_k$ captures how the educational attainment of the work-force in the individual's establishment affects the marginal product of k level workers, and hence the demand for such workers. Since these parameters were derived as the quotients between the partial derivatives of the marginal product of k level workers and the marginal product itself, they

[12] Cf Moretti (2004c), who estimates the external effects of education on productivity using plant level data, and identifies the external effect from variation in educational attainment outside the plant.

naturally vary between different educational categories. For a well behaved production function we would also expect to see: (1) $\mathbf{B}_k^k < 0$ for all *k*, where $\mathbf{B}_k^k$ is the *k*th element of the vector $\mathbf{B}_k$, representing the quotient between F_{kk}, the partial derivative of the marginal product of level *k* workers and the corresponding marginal product itself F_k, and (2), due to Young's theorem, sign($\mathbf{B}_k^l$) = sign($\mathbf{B}_l^k$) where $\mathbf{B}_k^l$ is the quotient between F_{lk}; i e, the partial derivative of F_k with respect to the *l*th argument, and F_k, and $\mathbf{B}_l^k$ is similarly defined.

There have been two suggestions in the literature on how to define the variable for the average educational attainment of the work-force outside the individual's establishment. Rauch (1993), Acemoglu & Angrist (2000) and Ciccone & Peri (2004) use the average years of schooling in the geographical unit of observation (city or state) . This choice implies that the external effect of increasing the average educational attainment in an area is assumed to be homogeneous across the educational distribution. Moretti (2004b) uses instead college educated. This choice constrains, instead, the potentially positive external effects of lower educations to be zero. I have chosen to use average years of schooling but I have also tested all empirical models using the share of workers having attained a longer university education or a post graduate education. Qualitatively, the results were not sensitive to whether years of schooling or the share of university educated were used. In the following, I therefore only present the results pertaining to average years of schooling in the following.

This paper uses Swedish municipalities as the basic geographical unit of measurement.[13] The municipality is the lowest level administrative area in Sweden. However, the municipalities differ to a large extent with respect to the density of economic activity. This implies that the geographical unit of measurement used here will be more heterogeneous than, for example, metropolitan areas, which underscores the need to provide more spatial structure to the basic empirical model. One of the models presented below is therefore based on the coordinates of the establishment's location. Hence two types of geographical unit of observation are used here.

[13] Previous work in this field has used such diverse geographical areas as US Standard Metropolitan Statistical Areas, SMSAs (Rauch, 1993, Ciccone & Peri, 2004 and Moretti, 2004b) and the states of the USA (Angrist & Acemoglu, 2000 and Ciccone & Peri, 2004) when constructing the average educational attainment in the individual's labor market.

3.2 Empirical model with additional spatial structure

The second model adds spatial information to equation 1 in the way outlined in section 2.2. The first alternative estimation equation thus becomes

$$\begin{aligned} y_{iegt} &= \theta_{0t} + \theta_1 S_{-egt} + \theta_2 TCWS_{gt} + \\ &+ \alpha_k + \mathbf{N}_{egt}\mathrm{B}_k + \mathbf{Z}_{egt}\boldsymbol{\lambda} + \mathbf{W}_{egt}\Lambda + \mathbf{X}_{it}\Phi + \varepsilon_{iegt} \end{aligned} \tag{2}$$

where the transport-cost-weighted educational attainment ($TCWS_{gt}$) is defined in accordance with S_{-egt} ; that is, it is based on the average years of schooling. Generalized commuting costs between the municipalities are used as weights in the construction of $TCWS_{gt}$.[14] Furthermore, $\mathbf{Z}_{egt}$ in equation 2 is a vector containing total employment in the own municipality outside establishment *e*, and the transport-cost-weighted sum of the surrounding municipalities' total employment. The remaining variables in equation 2 are the same as those in equation 1.

The second alternative of adding spatial structure presented in section 2.2 leads to the following estimation equation

$$\begin{aligned} y_{iet} &= \theta_{0t} + \theta_1 S_{-et}^{(1)} + \theta_2 S_{-et}^{(2)} + \theta_3 S_{-et}^{(3)} + \theta_4 S_{-et}^{(4)} + \\ &+ \alpha_k + \mathbf{N}_{et}\mathrm{B}_k + \mathbf{Z}_{-et}^{(d)}\boldsymbol{\lambda} + \mathbf{W}_{et}\Lambda + \mathbf{X}_{it}\Phi + \varepsilon_{iet} \end{aligned} \tag{3}$$

where $S_{-et}^{(d)}$ (d = 1, 2, ..., 4) is the average years of schooling of the work-force outside establishment *e* at a maximum distance of 5 kilometers (d = 1) from the establishment, between 5 and 10 kilometers (d = 2), between 10 and 15 kilometers (d = 3), and between 15 and 20 kilometers (d = 4); $\mathbf{Z}_{-et}^{(d)}$ is a vector containing the total employment at a maximum distance of 5 kilometers from the establishment (not including the employment in establishment *e*), between 5 and 10 kilometers away, between 10 and 15 kilometers away, and between 15 and 20 kilometers away; the remaining variables are the same as in equation 1.

[14] Generalized commuting costs include both the monetary cost of travel and the value of time. It is, furthermore, a weighted average of different modes of transport with the weights determined by the relevant market shares. The same costs are used when constructing the transport-cost-weighted sum of the surrounding municipalities' total employment; that is, the market potential. See section 4 for further information regarding these costs.

3.3 Endogeneity problems

The basic empirical model might be plagued by different kinds of endogeneity problems; i e, there are different reasons to believe that there exist a correlation between the error term and the average educational attainment of the workforce outside the individual's establishment. This is also true of the models in section 3.2 that add spatial structure to the analysis. Consequently, the ordinary least squares (OLS) estimator applied to equations 1, 2 and 3 might be biased and inconsistent. The primary reason for believing that there exist a correlation between the error term and the different measures of average educational attainment outside the own establishment is that relevant variables have been omitted from the equations.

First of all, an extensive literature has tried to control for unobserved individual specific "ability" in empirical models of the private economic return to education (see Card, 1999, for a survey). Obviously, omitted individual-specific ability could also bias cross-sectional estimators of the external effect of education. The motivation is that high-ability (or "career-oriented") individuals might tend to be located in areas with relatively much human capital; that is, in areas where the average educational attainment of the labor force is high. For this reason I choose to include individual-specific fixed effects in the model, denoted by f_i in the following. This suggests the following structure of the error term in equations 1, 2 and 3

$$\varepsilon_{iegt} = f_i + v_{iegt} ,$$

where the first term on the right hand side is potentially correlated with the average educational attainment of the work-force outside the own establishment and the second term is assumed to be uncorrelated with the explanatory variables in the equations. A conventional fixed effects estimator is then applied to equations 1, 2 and 3; e g the fixed effects version of equation 1 is given by

$$\begin{aligned} \Delta y_{ieg} &= \Delta\theta_0 + \theta_1 \Delta S_{-eg} + \Delta\alpha_k + \Delta \mathbf{N}_{eg} \mathrm{B}_k + \\ &+ \Delta Z_g \lambda + \Delta \mathbf{W}_{eg} \Lambda + \Delta \mathbf{X}_i \Phi + \Delta\varepsilon_{ie} \end{aligned} , \qquad (4)$$

where $\Delta y_{ieg} = y_{iegt} - y_{ie'g's}$; y_{iegt} is defined in equation 1 and $y_{ie'g's}$ is the natural logarithm of the individual's annual earnings in time period *s* in which he/she worked in establishment *e'* (*e'* = 1, 2, ..., $E_{g'}$), in municipality *g'* (*g'* = 1, 2, ..., *G*). Note that *e'* may be different from *e* and *g'* may differ from *g*, since some people change establishment and municipality. All right-hand-side variables are defined accordingly.

Since some individuals change the municipality in which they work, unobserved natural advantages of a municipality that may affect the productivity of establishments located there are not completely accounted for when using equation 4.[15] Hence, if such unobserved variables both raise the marginal productivity of workers and attract high educated workers, the error term in equation 4 may still be positively correlated with the average educational attainment. I choose to deal with this potential correlation by including a set of dummy variables for the county in which the establishment is located. [16] (The county is a larger geographical unit than the municipality.)

The fixed effects estimator together with the dummy variables discussed above remove the effect of unobserved time-invariant variables in the analysis; in other words, unobservable variables that might explain both high (low) earnings and a high (low) average educational attainment. Still, there may also exist time-varying omitted variables, the change in which may be correlated with changes in the average educational attainment of the workforce outside the own establishment. In the empirical analyses I assume, however, that such time varying variables are accounted for by the inclusion of the various measures of demand linkages in the analysis; i e, total employment in the municipality outside the own establishment and the market potential. This assumption is supported by the results in Moretti (2004b). He uses both an instrumental variables approach and a control variable approach to deal with these problems. More specifically, to control for exogenous demand shifts that may induce a spurious correlation between a change in the average educational attainment and the change in wages in a city, he includes a Katz and Murphy index (Katz & Mur-

---

[15] See Kim (1995 and 1999) and Elison & Glaeser (1999) for empirical evidence and discussions of natural advantage vs. the role of spillovers in locational decisions of firms in the US. There may obviously also exist important unobserved aspects of the establishment's production technology that may motivate an establishment-specific fixed effect as well. I investigated such fixed-effects in equations 1, 2 and 3 but all of the results were quite similar to those obtained when accounting for the individuals-specific fixed effects. To save space I only consider unobserved individual fixed effects and spatial heterogeneity in the text. It could be noted, however, that 83 percent of the individuals in the total sample used below are employed in the same municipality in the two time periods, and 60 percent are employed in the same establishment.

[16] I also tested specifications with dummy variables for the municipality in which the establishment is located. However, when using these dummy variables in equation 1, the estimated external effect of education became unreasonably sensitive to the inclusion or exclusion of the variables pertaining to the educational attainment of the individual's co-workers. The estimated external effect of education with municipality dummy variables in equation 1 tended to be positive without these variables in the equation, and tended to be negative when the municipality dummies were included. For this reason I choose to model unobserved spatial heterogeneity through dummy variables defined at the level of the county instead. This leaves more variation in the average educational attainment outside the individual's establishment, and leads to more stable results regarding the estimated external effect of education.

phy, 1992) in the estimation equation. His instrumental variable estimates of the external effects of education in this equation are all within one standard error of the corresponding ordinary least squares estimate. This indicates that with a suitable control variable for exogenous demand shifts, there should be little to be gained from using an instrumental variable approach.

4. Data

The data set is derived from Statistics Sweden's (SCB) registers. It consists basically of two parts: one is a sample of employed individuals and the other is data on all establishments in Sweden. The individuals are sampled from approximately 10 percent of all establishments in 1998 and linked to their establishment in 1998 and 1993.[17] The sample is stratified with respect to the size and county-distribution of the establishments, so as to assure representativity.

The analyses use information on individuals' gross annual earnings, educational attainment, age, sex and identification number for his or her establishment in 1998 and 1993. Information on gross annual earnings is collected from tax records.[18] Earnings in 1998 have been deflated to the 1993 price level with the consumer price index. The register information on educational attainment is collected from administrative records on completed degrees within the regular educational system, and is expressed in terms of the individual's highest attained educational level. There are seven mutually exclusive educational categories and an additional category for missing information on educational level. The sample of individuals is restricted to include individuals with non-missing information on own educational attainment, being at least 25 years old in 1993 and at most 65 years in 1998. The motivation for the lower age limit is to exclude individuals that may still be in the educational system. The motivation for the upper age limit is that this is the mandatory age of retirement in Sweden. The included individuals must also have had non-zero annual earnings in both 1993 and 1998.

The regional variables are constructed from the establishment data. This data set includes the relevant information on the educational attainment of the workforce, industry classification and the location of the establishment. The

[17] I have excluded sailors, farmers and self-employed individuals from the sample.

[18] The earnings definition used here includes wage earnings, work related compensations and earnings from running an own enterprise. To see if the results were sensitive to the inclusion of the latter earnings component, I also estimated all versions of the basic model with a more narrow definition of earnings excluding this component. All results regarding the external effects of education were virtually the same with the more narrow earnings definition, however.

average years of schooling of the workforce outside the individual's establishment is obtained by imputing the average years of schooling. The Swedish Level of Living Survey (SLLS) in 1991 is used to this end.[19] The SLLS contains both register information on the highest attained educational level and self-reported years of schooling. This dual information is used to estimate a model that relates self-reported years of schooling to highest attained educational level. The results are used to impute years of schooling for each educational level, to construct the average years of schooling in each municipality.

Information on travel costs between municipalities is obtained from a model used by the Swedish National Road Administration and the Swedish National Rail Administration. This source provides data on "generalized commuting costs" between municipalities in 1997 which are used to model travel costs both in 1993 and 1998. These costs include both monetary and time costs, and are obtained as a weighted average of costs associated with different modes of travel.[20] I also use the generalized commuting costs when constructing the municipalities' market potential. Municipalities are defined according to the 1995 classification when there were 288 municipalities.

Unfortunately, several establishments lack information on the longitude and latitude of the establishment's geographical location. In terms of employees, information is missing for 22 percent of Sweden's total employment in 1998. The TFP3 and the associated market potential presented in section 2 are based on establishments with non-missing information, and the empirical models are estimated on individuals that work in such establishments both in 1993 and 1998.

Table 1 contains descriptive statistics for the observable individual specific-characteristics in the sample. Since model 3 is estimated on a smaller sample, due to the missing information on longitudes and latitudes of establishments, the table also presents descriptive statistics separately for this smaller sample. The first row shows that average earnings in 1998 in the total sample is approximately equal to 226 000 Swedish kronor (SEK), which is almost 21 percent higher than the corresponding figure in 1993. Columns 3 and 4 display that average earnings in 1998 and 1993 is basically the same in the smaller sample.

The dummy variables *E1 – E7* indicate the highest attained educational level of the individual. The table suggests only minor changes over time in the educational level of the individuals in the sample. It also suggests no substantial difference between the total sample and the smaller one.

---

[19] See Fritzell & Lundberg (1994).

[20] Since there were no major changes in the transportation infrastructure between 1993 and 1998, commuting costs should not be changing dramatically over this period of time.

Table 1. Descriptive statistics: Means and (standard deviations): Individual characteristics

	All		*Smaller sample*	
	1998	1993	1998	1993
Earnings	225949	187406	225389	187961
	(122509)	(96717)	(128001)	(102745)
Educational level 1	0.097	0.098	0.094	0.095
	(0.297)	(0.297)	(0.292)	(0.293)
Educational level 2	0.104	0.104	0.106	0.107
	(0.305)	(0.306)	(0.308)	(0.309)
Educational level 3	0.337	0.343	0.341	0.346
	(0.473)	(0.475)	(0.474)	(0.476)
Educational level 4	0.121	0.124	0.124	0.127
	(0.326)	(0.330)	(0.330)	(0.333)
Educational level 5	0.168	0.167	0.164	0.162
	(0.374)	(0.373)	(0.370)	(0.369)
Educational level 6	0.156	0.151	0.155	0.150
	(0.363)	(0.358)	(0.362)	(0.357)
Educational level 7	0.017	0.013	0.016	0.013
	(0.128)	(0.113)	(0.125)	(0.112)
AGE, in 1998	45.783	-	45.961	-
	(9.075)		(9.092)	
MALE	0.505	-	0.488	-
	(0.500)		(0.500)	
Sample Size	177 591		112 463	

Notes: Earnings are measured in the prices of 1993. Educational levels: 1=Compulsory school (<9 years), 2=Compulsory school (9 or 10 years), 3=Upper secondary school (≤ 2 years), 4=Upper secondary school (> 2 years), 5=Post secondary school (< 3 years), 6=Post secondary school (≥ 3 years), 7=Post graduate education. The smaller sample refers to the sample of individuals working in establishments with information on geographical coordinates.

Table 2 contains descriptive statistics pertaining to municipalities and establishments. Note that the entries in this table are weighted averages with the weights given by the sample share of individuals in each municipality or establishment. The average years of schooling in the municipality (outside the own establishment) is almost 12 years in 1998 and was closer to 11.5 years in 1993. Furthermore, the average individual works in a municipality where some 104000 individuals are employed in 1998 which is almost 16 percent higher than the corresponding figure in 1993. Average establishment employment is increasing between 1993 and 1998. The average individual of the smaller sample seems to be working in a smaller establishment that is located in a larger municipality than the average individual of the total sample.

Table 2. Descriptive statistics: Means and (standard deviations) – Municipality characteristics and establishment characteristics

	All		*Smaller sample*	
	1998	1993	1998	1993
Average years of schooling[a]	11.922	11.611	11.973	11.656
	(0.576)	(0.581)	(0.597)	(0.596)
Employment[b]	103718	89680	120533	102026
	(158418)	(133764)	(172700)	(144078)
Transport-cost-weighted years of schooling[c]	5.789	5.607	5.962	5.772
	(2.232)	(2.194)	(2.279)	(2.236)
Transport-cost-weighted employment[d]	7997	7287	8077	7426
	(5200)	(4664)	(4986)	(4568)
Establishment employment	655	625	336	340
	(1305)	(1282)	(626)	(742)
Sample Size	177 591		112463	

Note: The information is weighted by the sample share of individuals in each municipality or establishment.
(a) Average years of schooling in the municipality outside the own establishment.
(b) Employment in the municipality outside the own establishment.
(c) Weighted sum of average years of schooling in the surrounding municipalities.
(d) Weighted sum of employment in the surrounding municipalities.

Table 3 reports further establishment characteristics for establishments with information on geographical coordinates. Note that the entries are weighted averages with the weights given by the sample share of individuals in each establishment. The table first presents the average years of schooling at different distances from the establishment. These entries suggest that individuals tend to have their establishment located in an area where the average educational attainment is relatively high. For example, average years of schooling among workers outside the own establishment at a maximum distance of 5 kilometers is 12 years in 1998, and the average years of schooling among workers further away is decreasing monotonically with the distance. This decline in the average educational attainment is also seen for the entries referring to 1993. Total employment is also largest within a circle with radius 5 kilometers from the establishment and then falls off with distance from the establishment.

Table 3. Descriptive statistics: Means and (standard deviations) – Establishment characteristics for establishments with information on longitudes and latitudes.

	1998	*1993*
Years of schooling, Distance≤5 km[a]	12.019	11.696
	(0.646)	(0.644)
Years of schooling, 5<distance ≤ 10km[b]	11.715	11.356
	(0.755)	(0.993)
Years of schooling, 10<distance≤15 km[c]	11.634	11.335
	(0.686)	(0.948)
Years of schooling, 15<distance≤20 km[d]	11.595	11.294
	(0.651)	(0.860)
Employment, distance≤5 km[e]	73146	61024
	(110002)	(91245)
Employment, 5<distance≤10 km[f]	57467	45845
	(93510)	(75859)
Employment, 10<distance≤15 km[g]	38338	31494
	(68333)	(57953)
Employment, 15<distance≤20 km[h]	26101	21218
	(47092)	(39072)
Sample Size	112 463	

Note: The information is weighted by the sample share of individuals in each establishment.
(a) Average years of schooling of employed individuals outside the individual's establishment within a circle of radius 5 kilometers from the own establishment.
(b-d) Average years of schooling of employed individuals within two circles with radii 5 and 10; 10 and 15; and 15 and 20 kilometers, respectively from the establishment.
(e-h) Employment outside the individual's establishment. Distances defined in accordance with years of schooling distances.

5. Results

This section begins with a presentation of the cross-sectional results in section 5.1, before turning to the fixed effects results in section 5.2. Within both of these sub-sections the results are first presented for the basic empirical model in equation 1 whereupon the results obtained with equations 2 and 3 are presented; that is, the equations that add spatial structure to the model of the externality.

5.1 Cross-sectional results

5.1.1 The basic empirical model

Table 4 presents the results from a series of different specifications of the basic model; that is, equation 1. The first specification only includes variables for the individual's educational attainment, age and a male dummy variable. The intention of this specification is simply to present a set of benchmark estimates for the private economic return to education, and see how these returns are affected when including the variable used to measure the externality.

The second specification in Table 4 includes the average years of schooling in the municipality outside the individual's establishment. The estimated external effect of education obtained with this specification suggests that a one-year increase in the average years of schooling would raise earnings by approximately 9 percent. The second column suggests, furthermore, that the private returns are negatively, albeit very modestly, affected by the inclusion of the variable for the externality.

The third specification in Table 4 intends to deal with the issue of imperfect substitutability between workers with different educational attainment.[21] This specification thus includes a set of control variables for the number of workers at the individual's establishment having attained each one of eight educational categories, interacted with the dummy variables for own educational attainment.[22] It also includes a set of industry dummies based on two-digit industry codes. The results suggest that the external effect is approximately 25 percent lower than in the specification that excludes these control variables. Hence, this result indicates a relatively large positive bias in estimates of the external effect of education obtained in models that fail to account for imperfect substitutability, which is basically in line with the conclusions of Ciccone & Peri (2004). The private economic returns to education obtained with this specification also tend to be higher than in the second specification.

[21] See the discussions in Ciccone & Peri (2004) and Moretti (2004b).

[22] I have included the category of workers with missing information on educational attainment so that these control variables sum to the total number of workers at the establishment. The results on the external effect of education are, however, not sensitive to this choice.

Table 4. Model 1 – Stacked OLS

Specification	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Municipality years of schooling[a]	-	0.091	0.068	0.061	0.043
		(0.014)	(0.013)	(0.017)	(0.014)
Own education= Level 2[b]	0.087	0.073	0.084	0.074	0.084
	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
Own education= Level 3[c]	0.106	0.093	0.121	0.094	0.121
	(0.007)	(0.005)	(0.004)	(0.005)	(0.004)
Own education= Level 4[d]	0.235	0.207	0.214	0.207	0.214
	(0.012)	(0.006)	(0.005)	(0.006)	(0.005)
Own education= Level 5[e]	0.266	0.239	0.291	0.241	0.292
	(0.010)	(0.007)	(0.007)	(0.007)	(0.007)
Own education= Level 6[f]	0.481	0.446	0.515	0.446	0.515
	(0.013)	(0.008)	(0.009)	(0.008)	(0.009)
Own education= Level 7[g]	0.719	0.651	0.754	0.649	0.759
	(0.043)	(0.056)	(0.026)	(0.059)	(0.026)
Age	0.048	0.049	0.048	0.049	0.048
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age-squared/100	-0.048	-0.049	-0.047	-0.049	-0.047
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Male	0.365	0.365	0.286	0.365	0.286
	(0.007)	(0.006)	(0.005)	(0.006)	(0.005)
Municipality Employment/1000 000[h]	-	-	-	0.171	0.145
				(0.056)	(0.042)
Co-workers education[i]	No	No	Yes	No	Yes
Industry dummies[j]	No	No	Yes	No	Yes
R^2 *– Adjusted*	0.352	0.365	0.420	0.367	0.421
Number of individuals			177 591		

Notes: Dependent variable: Natural logarithm of annual earnings in 1998 and 1993. Standard errors robust to clustering at the level of the municipality in parentheses. All specifications are estimated with separate intercepts for 1993 and 1998.
(a) See note a Table 2.
(b-g) See notes to Table 1.
(h) See note b Table 2.
(i) The number of individuals having attained each of 7 educational levels and a separate category for those with missing information on educational level interacted with dummy variables for the individual's educational level.
(j) Based on two-digit-industry codes.

Specification 4 in Table 4 ignores the issue of imperfect substitutability and instead includes a control variable for the size of the market in the municipality, namely total employment in the municipality outside the individual's establishment. This specification suggests that increasing the average years of schooling by one year will lead to a 6 percent increase in individual earnings. This also implies a positive bias in the estimated external effect of education in the specification that ignores the role of large markets and demand linkages; i e specification 2 in the table. Furthermore, working in a municipality with a high employment level has a positive effect on earnings as suggested by the significantly positive parameter for municipality employment outside the individual's establishment.

The fifth specification reported in Table 4 controls both for imperfect substitutability between workers with different educational attainment, demand linkages (employment outside the own establishment) and industry dummies. The estimated effect of increasing the average years of schooling outside the individual's establishment by one year is now associated with a 4 percent increase in individual earnings. This effect is significantly different from zero at conventional levels of significance. The point estimate is substantially lower than in the second specification that ignored both the issue of imperfect substitutability and demand linkages. The two 95 percent confidence intervals of the parameter in the two specifications are overlapping, however, suggesting that the difference is not significant.

5.1.2 Empirical models with additional spatial structure

Table 5 presents results from the use of two variables to measure the externality: one that measures the average educational attainment in the municipality outside the individual's establishment, and one that consists of a weighted sum of the average educational attainment in the surrounding municipalities (TFP2 in section 2). Specification 1 in Table 5 includes these two variables together with control variables for the individual's educational attainment, age, age-square and a male dummy variable. Both of the estimated parameters for the educational externalities are positive and different from zero at conventional levels of significance. This suggests externalities of education across municipality borders. It can be noted that the external effect of the average years of schooling in the municipality outside the own establishment is substantially lower than the corresponding estimate in Table 4, column 2. This reflects a positive relationship between the average educational attainment in a municipality and the accessibility to human capital in the surrounding municipalities.

Table 5. Model 2 – Stacked OLS

Specification	*(1)*	*(2)*	*(3)*	*(4)*
Municipality years of schooling[a]	0.067 (0.012)	0.048 (0.009)	0.036 (0.016)	0.016 (0.010)
Transport-cost-weighted years of schooling/100[b]	1.161 (0.235)	1.034 (0.207)	-0.205 (0.383)	-0.464 (0.265)
Municipality Employment/1000 000[c]	-	-	0.210 (0.062)	0.210 (0.039)
Transport-cost-weighted employment/100 000[d]	-	-	0.541 (0.193)	0.662 (0.117)
Co-workers education[e]	No	Yes	No	Yes
Industry dummies[f]	No	Yes	No	Yes
R^2 – Adjusted	0.368	0.421	0.369	0.423
Number of individuals	177 591			

Notes: Dependent variable: Natural logarithm of annual earnings in 1998 and 1993. Specifications include control variables for the individual's age, age-squared, educational attainment, a dummy variable that equals one if the individual is male and separate intercepts for 1993 and 1998. Standard errors robust to clustering at the level of the municipality in parentheses.
(a) See note a in Table 2.
(b) See note c in Table 2.
(c) See note b Table 2.
(d) See note d in Table 2.
(e-f) See notes i and j Table 4.

Specification 2 in Table 5 includes control variables to account for industry structure and imperfect substitutability between workers with different educational attainment. This lowers both of the estimated external effects of education. The relative effect of including the control variables is larger on the parameter measuring the external effects occurring within the municipality. This result seems reasonable if most of the effect from failing to account for imperfect substitutability is operating relatively close to the establishment.

In specification 3 the inclusion of the control variables for the market potential has, instead, a more marked effect on the estimated external effect of education in the surrounding municipalities, which becomes negative and insignificantly different from zero. The external effect of education in the own municipality is relatively low but still significantly different from zero at the five percent level of significance. Note also that both of the variables that intend to measure the market potential are positive and different from zero at conventional levels of significance. The results from specification 3 lend support to the idea that externalities are important at shorter distances whereas demand linkages are important over larger geographical areas.[23]

Specification 4 in Table 5 includes both control variables for imperfect substitutability between workers with different educational attainment, industry dummies and market potential. This specification suggest even lower external effects to education in the own municipality. In fact, the estimated external effect of average years of schooling in the municipality outside the individual's establishment is not significantly different from zero at the 10 percent level of significance in this specification. Moreover, the results obtained with this specification actually suggest a negative effect from the average years of schooling in the surrounding municipalities, which is different from zero at the 10 percent level of significance. This might reflect that gainful interactions in the own municipality are relocated to the surrounding municipalities when these municipalities have large stocks of human capital.[24]

Table 6 presents the cross-sectional results pertaining to the second way of adding spatial structure to the empirical model (TFP3 in section 2), by measuring the average educational attainment at four different distances from the establishment. Note that the sample used to estimate the external effects of education with this model is smaller than that used to estimate models 1 and 2. This is due to missing values on the establishments' geographical coordinates. Specification 1 in Table 6 includes the four variables for the average years of schooling at different maximum distances from the individual's establishment (excluding years of schooling in the own establishment). The magnitude of the estimated parameters suggests indeed that the external effect of education is declining in distance, although it does not seem to be declining monotonically. The estimated external effect is largest for average years of schooling close to the establishment, suggesting that an increase in the average years of schooling by one year would increase the individual's earnings by approximately 5 percent. The effect is substantially smaller when the distance is between 5 and 10 kilometers where a corresponding increase in average years of schooling would lead to an increase in the individual's earnings by some 1.5 percent. The corresponding effect is approximately 2 percent at the distance 10 to 15 kilometers. Finally, at a distance between 15 and 20 kilometers from the establishment, an increase in the average years of schooling would lead to an increase in the individual's earnings by approximately 1 percent. All of the estimated parameters are different from zero at the 10 percent level of significance in specification 1.

[23] See for example Ottaviano & Puga (1998).

[24] Boarnet (1998) reports a similar finding for infrastructural capital which seems to have a negative spillover effect on neighbouring areas.

Specification 2 adds the control variables for imperfect substitutability between workers with different educational attainment. This reduces the estimated external effect of average years of schooling at all distances. In fact, it is only the parameters associated with average years of schooling at a maximum distance of 5 kilometers and at the distance between 10 and 15 kilometers that are significantly different from zero at conventional levels of significance.

Specification 3 is used to investigate how the inclusion of employment at different distances from the establishment affects the estimated external effects of education, where the distances are the same as those used to measure the external effects of education. This specification suggests that the inclusion of this measure of the market potential reduces the estimates of the external effects of education at all distances but the one furthest away from the municipality. This is somewhat surprising since the third specification of model 2 suggested that controlling for the market potential in the analysis lowered the external effects from education located far away from the establishment; cf Table 5, column 3. It could also be noted that it is only the parameters related to years of schooling at a maximum distance of 5 kilometers from the establishment and at a distance between 15 and 20 kilometers that are different from zero at the 10 percent level of significance.

Specification 4 controls for imperfect substitutability between workers with different educational attainment, industry dummies and the employment at different distances from the establishment. This specification suggests that the parameter related to years of schooling at a maximum distance of 5 kilometers is positive and different from zero at conventional levels of significance. This estimate suggests that increasing the years of schooling by 1 year will tend to raise the individual's earnings by some 1.5 percent. The parameter related to years of schooling at a distance between 5 and 10 kilometers from the establishment is negative and different from zero at the 10 percent level, hence suggesting negative spillovers between close locations. As in model 2, this negative spillover might be interpreted as gainful interactions being relocated from the own location if the stock of human capital is relatively large in nearby locations. Years of schooling at further distances do not have any significant effects in this specification.

Table 6. Model 3 – Stacked OLS

Specification	*(1)*	*(2)*	*(3)*	*(4)*
Years of schooling, distance ≤ 5 km[a]	0.053 (0.010)	0.045 (0.005)	0.017 (0.009)	0.015 (0.005)
Years of schooling, 5 < distance ≤ 10 km[b]	0.015 (0.008)	0.000 (0.003)	0.003 (0.006)	-0.005 (0.003)
Years of schooling, 10 < distance ≤ 15 km[c]	0.020 (0.006)	0.016 (0.004)	0.008 (0.006)	0.004 (0.003)
Years of schooling, 15 < distance ≤ 20 km[d]	0.009 (0.005)	0.005 (0.003)	0.009 (0.005)	0.003 (0.003)
Employment/1000 000, distance ≤ 5 km[e]	-	-	0.152 (0.094)	0.192 (0.042)
Employment/1000 000, 5 < distance ≤ 10 km[f]	-	-	0.366 (0.158)	0.184 (0.058)
Employment/1000 000, 10 < distance ≤ 15 km[g]	-	-	0.137 (0.123)	0.211 (0.074)
Employment/1000 000, 15 < distance ≤ 20 km[h]	-	-	0.091 (0.124)	0.124 (0.085)
Co-workers education[i]	No	Yes	No	Yes
Industry dummies[j]	No	Yes	No	Yes
R^2 – Adjusted	0.359	0.416	0.364	0.419
Number of individuals	112 463			

Notes: Dependent variable: Natural logarithm of annual earnings in 1998 and 1993. Specifications include control variables for the individual's age, age-squared, educational attainment, a dummy variable that equals one if the individual is male and separate intercepts for 1993 and 1998. Standard errors robust to clustering at the level of the establishment in parentheses.
(a-d) See notes a-d to Table 3.
(e-h) See notes e-h to Table 3.
(i-j) See notes i and j to Table 4.

5.2 Fixed effects estimates

5.2.1 The basic empirical model

Table 7 summarizes the results obtained with the fixed effects estimator applied to model 1. They suggest that the external effect is positive and significantly different from zero at the 10 percent level in specifications 1, 2 and 3; cf Table 7, row 1. But the fourth specification indicates no statistically significant external effects of education.

The estimator may be biased because of omitted unobserved spatial characteristics that the individual fixed effects estimator fails to remove for individuals that change the municipality in which they work. To investigate this, dummy variables for the county of the location was included in the model. Apart from specification 1, these results suggest that the external effect of education is not different from zero at conventional levels of significance; cf Table 7, row 2.

Table 7. Model 1 – Fixed effects
Measure of average educational attainment in the municipality: Municipality years of schooling[a]

Specification	*(1)*	*(2)*	*(3)*	*(4)*
Individual Fixed Effects[b]	0.026	0.020	0.013	0.006
	(0.006)	(0.005)	(0.007)	(0.007)
Individual Fixed Effects	0.011	0.009	0.004	0.000
and County Dummy Variables[b,c]	(0.006)	(0.006)	(0.008)	(0.007)
Municipality employment[d]	-	-	Yes	Yes
Co-workers education[e]	No	Yes	No	Yes
Industry dummies[f]	No	Yes	No	Yes
Number of individuals	177 591			

Notes: Specifications include control variables for the individual's educational attainment, age-squared and an intercept term. Standard errors robust to clustering at the level of the municipality in parentheses.
(a) See note a to Table 2.
(b) The dependent variable is the difference between the natural logarithm of the individual's annual earnings in 1998 and in 1993.
(c) County dummy variables refer to the within-individual difference between 1998 and 1993 in these dummy variables.
(d) See note b to Table 2.
(e-f) See notes i-j to Table 4.

5.2.2 Empirical models with additional spatial structure

The first two rows of Table 8 summarizes the basic fixed effects results for model 2; that is, the model that includes a transport-cost-weighted sum of the average years of schooling in the surrounding municipalities. These results suggest, in general, no statistically significant positive external effects of education within the own municipality. The results do suggest, however, positive external effects of education from the surrounding municipalities in specifications 1 and 2. In specification 3 the effect is negative and significantly different from zero at conventional levels of significance, however. In addition, in the fourth and last specification both of the estimated parameters suggest negative external effects of education, both of which are significantly different from zero at the 5 percent level. This result could be due to relocations of gainful interactions between nearby geographical locations. The negative effect of education in the own municipality might indicate negative effects between locations within the municipality (cf the results in Table 9).

Table 8. Model 2 – Fixed effects

Specification		*(1)*	*(2)*	*(3)*	*(4)*
Individual Fixed Effects[c]	*Municipality years of schooling*[a]	0.009	0.001	-0.008	-0.013
		(0.007)	(0.006)	(0.007)	(0.006)
	Transport-cost-weighted years of schooling[b]	0.009	0.011	-0.012	-0.007
		(0.002)	(0.002)	(0.003)	(0.003)
Individual Fixed Effects and County Dummy Variables[c,d]	*Municipality years of schooling*[a]	0.002	-0.003	-0.003	-0.009
		(0.008)	(0.007)	(0.008)	(0.007)
	Transport-cost-weighted years of schooling[b]	0.008	0.011	0.007	-0.012
		(0.004)	(0.004)	(0.007)	(0.005)
Municipality employment and transport-cost-weighted employment[e]		-	-	Yes	Yes
Co-workers education[f]		No	Yes	No	Yes
Industry dummies[g]		No	Yes	No	Yes
Number of individuals		177 591			

Notes: Specifications include control variables for the individual's educational attainment, age-squared and an intercept term. Standard errors robust to clustering at the level of the municipality in parentheses.
(a) See note a in Table 4.
(b) See note c in Table 2.
(c-d) See notes b-c in Table 7.
(e) See notes b and d in Table 2.
(f-g) See notes i-j in Table 4.

The first set of results obtained with the individual fixed effects estimator applied to equation 2 may, however, still be biased since unobserved spatial heterogeneity is not completely accounted for due to individuals that change the municipality in which they work. Therefore dummy variables for the county of location are included in the individual fixed effects estimation equation. The corresponding results are reported in rows 3 and 4 of Table 8. They indicate no statistically significant external effects of education within the own municipality. Specifications 1 and 2 indicate, however, positive and statistically significant external effects from education in the surrounding municipalities. Specification 3 indicates no statistically significant external effects of education from the surrounding municipalities, whereas specification 4 suggests negative spillovers from education in the surrounding municipalities.

Table 9 presents the fixed effects results for model 3. As shown in the upper part of the table, the individual fixed effects estimator tends to produce results indicating both significantly positive and negative external effects of education. Specifications 1 and 2 suggest positive effects of years of schooling in the circle around the establishment with a radius of 5 kilometers and of years of schooling within the two circles with radii 10 and 15 kilometers. However, these effects are insignificantly different from zero in specifications 3 and 4. The latter specifications indicate instead a negative effect from years of schooling located within the two circles with radii 5 and 10 kilometers around the establishment. Including dummy variables for the county of location does not alter these conclusions in any substantial way, cf the lower part of the table.

Table 9. Model 3 – Fixed effects

Specification		*(1)*	*(2)*	*(3)*	*(4)*
Individual Fixed Effects[e]	*Years of schooling, distance ≤ 5 km*[a]	0.015	0.018	-0.002	0.002
		(0.005)	(0.005)	(0.006)	(0.005)
	Years of schooling, 5 < distance ≤ 10 km[b]	-0.003	-0.005	-0.006	-0.007
		(0.003)	(0.003)	(0.003)	(0.003)
	Years of schooling, 10 < distance ≤ 15 km[c]	0.005	0.006	0.003	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
	Years of schooling, 15 < distance ≤ 20 km[d]	0.003	0.004	0.003	0.003
		(0.003)	(0.003)	(0.003)	(0.003)
Individual Fixed Effects and County Dummy Variables[e,f]	*Years of schooling, distance ≤ 5 km*[a]	0.006	0.012	0.000	0.004
		(0.005)	(0.005)	(0.006)	(0.005)
	Years of schooling, 5 < distance ≤ 10 km[b]	-0.003	-0.006	-0.006	-0.007
		(0.003)	(0.003)	(0.003)	(0.003)
	Years of schooling, 10 < distance ≤ 15 km[c]	0.003	0.004	0.003	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
	Years of schooling, 15 < distance ≤ 20 km[d]	0.001	0.002	0.002	0.003
		(0.003)	(0.003)	(0.003)	(0.003)
Employment at different distances[g]		-	-	Yes	Yes
Co-workers education[h]		No	Yes	No	Yes
Industry dummies[i]		No	Yes	No	Yes
Number of individuals		112 463			

Notes: Specifications include control variables for the individual's age-squared, educational attainment, a dummy variable that equals one if the individual is male and an intercept term. Standard errors robust to clustering at the level of the establishment in parentheses.
(a-d) See notes a-d to Table 6.
(e-f) See note b-c to Table 7.
(g) See notes e-h to Table 6.
(h-i) See notes i-j to Table 4.

6. Conclusions

This paper provided an empirical investigation of externalities from education in Sweden in an earnings equation framework. It used matched employee-establishment data to deal with the issue of imperfect substitutability between workers with different educational attainment. It also investigated the idea that educational externalities arise through some form of face-to-face interaction between individuals. Furthermore, a set of different specifications and fixed effects models was used to test the robustness of the basic cross-sectional model. Two different geographical units of observation were also used: Swedish municipalities with administratively determined borders and a set of circles with different radii around establishments with information on geographical latitudes and longitudes.

The basic cross-sectional models suggested that externalities are positive and significantly different from zero. The cross-sectional results were also, in general, coherent with the idea that externalities are declining in the Euclidean distance and that externalities might cross administratively determined borders. This might suggest the relevance of transport costs for the transmission of human capital externalities. All of these results were, however, rather sensitive to the inclusion of variables that intended to control for the effect of imperfect substitutability between workers with different educational attainment in the analysis, industry dummies and measures of the geographical unit's so-called market potential.

Furthermore, few of the cross-sectional results survive the inclusion of individual fixed effects and county dummy variables in the estimation equation. In other words, the cross-sectional results may to a large extent be due to unobserved individual characteristics and unobserved spatial heterogeneity. Thus, even though some geographical areas have relatively high average educational attainment, this does not necessarily cause higher earnings in these areas.

One qualification of this conclusion seems warranted, however. The models used here do not include an explicit dynamic adjustment process of wages to changes in the average educational attainment in the geographical unit of observation. Using a dynamic framework for estimating the external effects of education on wages or earnings might be a fruitful area for future research.

References

Acemoglu, Daron & Joshua Angrist (2000), "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws," *NBER Macroeconomic manual* 2000, pp 9-59.

Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers

Berman, Eli, John Bound & Zvi Griliches (1994), "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures," *Quarterly Journal of Economics*, Vol. 109, pp 367-397.

Boarnet, Marlon G. (1998), "Spillovers and locational effects of public infrastructure," *Journal of Regional Science*, Vol. 38 (3), pp 381-400.

Card, David (1999), "The Causal Effect of Education on Earnings," in Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics, Volume 3*, Amsterdam: Elsevier Science Publishers.

Ciccone, Antonio & Giovanni Peri (2004), "Identifying Human Capital Externalities: Theory with Applications," Mimeo August 2004.

Conley, Timothy G., Fredrick Flyer & Grace Tsiang (2003) "Spillovers from Local Market Human Capital and the Spatial Distribution of Productivity in Malaysia," *Advances in Economic Analysis & Policy*, Vol. 3 No. 1 Article 5.

Duranton, Gilles & Diego Puga (2004), "Micro-foundations of urban agglomeration economies," *Handbook of Regional and Urban Economics*, Vol. IV, editors J. Vernon Henderson and Jacques-François Thisse, Elsevier: North-Holland

Ellison, Glenn & Edward L. Glaeser (1999), "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?," *The American Economic Review: Papers and Proceedings*, May 1999, pp 311-316.

Fritzell, J. & O. Lundberg editors (1994), *Vardagens villkor. Levnadsförhållanden i Sverige under tre decennier*, Stockholm: Brombergs Förlag.

Fujita, Masahisa & Jacques-François Thisse (1996), “Economics of Agglomeration,” *Journal of the Japanese and International Economies* 10, pp 339-378.

Fujita, Masahisa, Paul Krugman & Anthony J. Venables (1999), *The Spatial Economy - Cities, Regions and International Trade*, Cambridge MA: The MIT Press.

Gaspar, Jess & Edward L. Glaeser (1998), “Information Technology and the Future of Cities,” *Journal of Urban Economics* 43, pp 136-156.

Glaeser, Edward L. & Janet E. Kohlhase (2004), “Cities, regions and the decline of transport costs,” Papers in Regional Science 83, pp 197-228.

Hanson, Gordon H. (2004), “Market Potential, Increasing Returns, and Geographic Concentration,” Forthcoming in *the Journal of International Economics*.

Harris, C.D. (1954), “The Market as a Factor in the Localization of Industry in the United States,” *Annals of the Association of American Geographers* 44, pp 315-348.

Jacobs, Jane (1969), *The Economy of Cities*, New York: Vintage Books.

Jaffe, Adam B., Manuel Trajtenberg & Rebecca Henderson (1993), “Geographical Localization of Knowledge Spillovers as Evidenced by Patent Citations,” *Quarterly Journal of Economics*, August 1993, pp 577-598.

Katz, Lawrence F. & Kevin M. Murphy (1992), “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *Quarterly Journal of Economics* 107, pp 35-78.

Kim, Sukkoo (1995), “Expansion of Markets and The Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860-1987,” *Quarterly Journal of Economics* November (1995), pp 881-907.

Kim, Sukkoo (1999), “Regions, resources, and economic geography: Sources of U.S. regional comparative advantage, 1880-1987,” *Regional Science and Urban Economics* 29, pp 1-32.

Krueger, Alan B. & Mikael Lindahl (2001), "Education for Growth: Why and for Whom?," *Journal of Economic Literature* Vol. XXXIX Number 4, December 2001, pp 1101-1136.

Krugman, Paul (1991a), "Increasing Returns and Economic Geography," *Journal of Political Economy* 99(3), pp 483-499.

Krugman, Paul (1991b), *Gegraphy and Trade*, Cambridge MA: The MIT Press.

Lucas, Robert E. (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22, pp 3-42.

Marshall, Alfred (1920), *Principles of Economics: An introductory volume*, Eighth Edition, London: MacMillan.

Mellander, Erik (1999), "The multi-dimensional nature of labor demand and skill-biased technical change," Working Paper 1999:9, Institute for Labor Market Policy Evaluation.

Moretti, Enrico (2004a), "Human Capital Externalities in Cities," *Handbook of Regional and Urban Economics*, Vol. IV, editors J. Vernon Henderson and Jacques-François Thisse, Elsevier: North-Holland.

Moretti, Enrico (2004b), "Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data", *Journal of Econometrics* 121 (2004), pp 175-212.

Moretti, Enrico (2004c), "Workers' Education, Spillovers and Productivity: Evidence From Plant-Level Production Functions," *American Economic Review* 94(3), pp 656-690.

Ottaviano, Gianmarco I.P. & Diego Puga (1998), "Agglomeration in the Global Economy: A Survey of the 'New Economic Geography'," *The World Economy*, 21(6), August 1998, pp 707-731.

Rauch, James E. (1993), "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities," *Journal of Urban Economics*, 34, pp 380-400.

Roback, Jennifer (1982), "Wages, rents and the quality of life," *Journal of Political Economy* 90(6), pp 1257-1278.

Roback, Jennifer (1988), “Wages, rents and amenities: Differences among workers and regions, ” *Economic Inquiry* 26(1), pp 1257-1278.

Rosen, Sherwin (1986), “The Theory of Equalizing Differences,” in Orley C. Ashenfelter & Richard Layard, editors, *Handbook for Labor Economics, Volume 1*, Amsterdam: North Holland, pp 641-692.

Rosenbaum, Paul R. (2002), *Observational Studies*, New York: Springer Verlag.

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

2005:1 Ahlin Åsa & Eva Mörk ”Vad hände med resurserna när den svenska skolan decentraliserades?”

2005:2 Söderström Martin & Roope Uusitalo ”Vad innebar införandet av fritt skolval i Stockholm för segregeringen i skolan?”

2005:3 Fredriksson Peter & Olof Åslund ”Påverkas socialbidragsberoende av omgivningen?”

2005:4 Ulander-Wänman Carin ”Varslad, uppsagd, återanställd. Företrädesrätt till återanställning enligt 25 § LAS i praktisk tillämpning”

2005:5 Isacsson Gunnar ”Finns det en skillnad mellan samhällets och individens avkastning på utbildning?”

Working Papers

2005:1 Ericson Thomas "Personnel training: a theoretical and empirical review"

2005:2 Lundin Martin "Does cooperation improve implementation? Central-local government relations in active labour market policy in Sweden"

2005:3 Carneiro Pedro, James J Heckman & Dimitriy V Masterov "Labor market discrimination and racial differences in premarket factors"

2005:4 de Luna Xavier & Ingeborg Waernbaum "Covariate selection for non-parametric estimation of treatment effects"

2005:5 Ahlin Åsa & Eva Mörk "Effects of decentralization on school resources"

2005:6 Cunha Flavio, James J Heckman & Salvador Navarro "Separating uncertainty from heterogeneity in life cycle earnings"

2005:7 Söderström Martin & Roope Uusitalo "School choice and segregation: evidence from an admission reform"

2005:8 Åslund Olof & Peter Fredriksson "Ethnic enclaves and welfare cultures – quasiexperimental evidence"

2005:9 van der Klaauw Bas, Aico van Vuuren & Peter Berkhout "Labor market prospects, search intensity and the transition from college to work"

2005:10 Isacsson Gunnar "External effects of education on earnings: Swedish evidence using matched employee-establishment data"

Dissertation Series

2003:1 Andersson Fredrik "Causes and labor market consequences of producer heterogeneity"

2003:2 Ekström Erika "Essays on inequality and education"