

Cui, Ying; Martins, Pedro S.

Working Paper

What drives social returns to education? A meta-analysis

GLO Discussion Paper, No. 706

Provided in Cooperation with:
Global Labor Organization (GLO)

Suggested Citation: Cui, Ying; Martins, Pedro S. (2020) : What drives social returns to education? A meta-analysis, GLO Discussion Paper, No. 706, Global Labor Organization (GLO), Essen

This Version is available at:
<https://hdl.handle.net/10419/225637>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

What drives social returns to education? A meta-analysis *

Ying Cui[†]

Pedro S. Martins[‡]

Queen Mary University of London

Queen Mary University of London

& IZA & GLO

October 23, 2020

Abstract

Education can generate important externalities that contribute towards economic growth and convergence. In this paper, we study the drivers of such externalities by conducting the first meta-analysis of the social returns to education literature. We analyse over 1,000 estimates from 31 articles published since 1993 that cover 15 countries. Our results indicate that: 1) spillovers slow down with economic development; 2) tertiary schooling and schooling dispersion increase spillovers; 3) spillovers are smaller under fixed-effects and IV estimators but larger when measured at the firm level; and 4) there is publication bias (but not citation bias).

Keywords: returns to education; education externalities.

JEL Codes: I26, I28, J24, J31, C36.

*We thank comments from Maria Koumenta and Yong Yang. Cui thanks financial support from the China Scholarship Council. All errors are our own.

[†]Email: ying.cui@qmul.ac.uk. Address: School of Business and Management, Queen Mary University of London, 327 Mile End Road, London E1 4NS, United Kingdom. Phone: +44 (0)20 7882 8570

[‡]Corresponding author. Email: p.martins@qmul.ac.uk. Address: School of Business and Management, Queen Mary University of London, 327 Mile End Road, London E1 4NS, United Kingdom. Web: <https://sites.google.com/site/pmrsmartins/>.

1 Introduction

The effects of education may go much beyond the individuals that acquire it. A person's schooling may affect different outcomes of their colleagues at work, neighbours, and possibly even other people in the same country (10). To the extent that education shapes an individual's own thinking, actions, and outcomes - as indicated by a large literature on the private returns of education -, one's learning at school can also influence different economic and non-economic variables regarding other individuals. Specific examples include better informed political participation, increased tax revenues, lower public expenditure, lower crime, and slower spread of diseases.

Given their significance and breadth, such spillovers can promote economic development and convergence. This will apply if the spillovers are higher when economic development is lower. Indeed, the specific examples above may suggest that the marginal social return to education would be higher at lower levels of economic development. For instance, if crime or the spread of disease tends to be higher at lower levels of economic development and if education tends to reduce crime or the spread of disease, then the social effects of education may be greater when countries are at earlier stages in their development. Many of these social effects would be translated into pecuniary dimensions, including productivity and wages.

Note also that the potential for such spillovers or externalities - and the underlying inefficiency from education provided exclusively by markets - has also contributed to large public investments in education. For instance, according to the World Bank, over 15% of governments' total expenditure is devoted to education, corresponding to an average of 4% of GDP. Moreover, education spillovers may follow from non-pecuniary external returns (technological spillovers or knowledge diffusion) or, alternatively, pecuniary external returns (market interactions and prices) (12, 4). In the latter case, more schooling in the general workforce may incentivise firms to invest in capital which may make even the less schooled more productive. However, note that schooling could theoretically also have negative external effects, namely in the context of signaling models.

In this paper, we seek to better understand the drivers of social returns to education, including the role of economic development. Our contribution is to conduct what we believe is the first meta-analysis of this literature. According to our review, that we describe in more detail below, there are 31 journal articles in the (micro-)economics literature that research the magnitude of different types of education spillovers. These studies cover 15 countries, of which five are emerging or developing economies (China, Indonesia, Kenya, South Africa, and Tunisia).

To be able to better compare the studies, we focus on the microeconomic studies that consider pecuniary outcomes. We then study the role of a number of contextual and methodological variables.

Moreover, we also analyse the extent to which the literature suffers from publication and citation biases. The former concerns the more likely publication of particular results, namely those with positive effects. The second type of bias, which we borrow from the medical literature, concerns the extent to which particular results, namely positive effect, are more likely to be cited by other papers.

Our findings are supportive of the hypothesis above, namely that spillovers slow down with economic development. Moreover, we also find that tertiary schooling and schooling dispersion increase spillovers; spillovers are smaller under fixed-effects and IV estimators but larger when measured at the firm level; and that there is publication bias (but not citation bias). These results can also be helpful in allowing researchers to better compare their findings with other studies which adopt different methodological approaches.

The remaining of the paper is organised as follows. Our data are described in Section 2. The research design and results are presented in Section 3. This includes both the analysis of the drivers of the social and external returns along multiple dimensions of the studies as the analysis of publication and citation biases. The final section concludes.

2 Data and variables

2.1 Criteria for selecting studies

Our selection of studies was based on a comprehensive Google Scholar search. This was based on the following keywords entered in our search: education externalities, human capital externalities, education spillovers, social returns to education and external returns to education (using the 'OR' operator). These keywords capture the different phrases that authors have used to capture the concept of externalities in education. Our search was conducted in May 2020 and considered the first 30 pages of results delivered by Google Scholar.

Following the initial stage above, we then considered the studies that met the following subject and methodological criteria. On the subject side, we considered only studies that focused on at least one of three key economic outcomes: productivity (of firms), wages, and rents (of land). These have been identified before as three main avenues for education spillovers (12): more educated workers can drive upwards the productivity of the firms in which they are employed, which can then increase the wages of all employees (through some combination of rent sharing and human capital improvements). Finally, increases in wages in locations where schooling spillovers are large can lead to increases in rental prices.

On the methodological side, our inclusion criteria were the following. First, we consider only

studies published in academic journals list in the Academy of Business Schools (ABS) Journal Guide 2018. This widely used journal list includes over 300 journals in economics alone. Second, we only include estimates that are reported in the main text (excluding those in the appendices of the paper) and in which the number of observations, the time period considered in the estimation (and other variables discussed below) are clearly reported. Third, we considered only studies focused on single countries.¹

In summary, we collected a total of 1,008 estimates from 31 empirical studies on education spillovers. These estimates cover 15 countries: Denmark, Germany, Hungary, Italy, Netherlands, UK, Portugal, US, Switzerland, China, Russia, South Africa, Indonesia, Kenya and Tunisia, which amount to 33.07% of the world population.

2.2 Explanatory variables

We characterise the studies and their estimates along nine dimensions. These dimensions lead to 28 variables that we then consider as potential predictors of the size of the spillover that we analyse, following our discussion above. These dimensions and variables are as follows:

1. *Spillover measure*: We consider the cases of the *average years of schooling* and the *share of college-educated workers*. The former is more comprehensive while the latter is more focused on the individual profiles that may have greater potential to generate spillovers. Of course, lower levels of schooling, such as secondary school completion, may also be relevant in less developed countries where average education levels are lower. Note that we could not find any study that focused on spillovers from qualitative types of education, such as vocational education or particular subjects or course in higher education (e.g. engineering or humanities).

2. *Spillover scope*: These have been classified as *regional*, *industry*, or *firm levels*. These three levels correspond to the ways in which the measures of the potential spillovers (such as the average years of schooling and share of college-educated workers) are calculated. As discussed before, the schooling externalities can arise and be measured within different dimensions, related to the location of the individuals, the sector in which they work, or within their firm (a particular combination of sector and location in most cases, namely in firms located in a single area). For instance, if the average years of schooling (or college-educated share) adopted in the study are calculated within a city/region, then we define the spillover scope of that study to be the regional level. Alternatively, if the measures are calculated within each firm (industry), by considering the relevant schooling measure of those workers, then we define the scope as the firm (industry) level.

¹We also considered Acemoglu & Angrist (2) and Rudd (14), which were not published in journals but are widely cited in the literature, and excluded one very short study which reported only three estimates.

3. *Spillover outcome:* The effects of spillovers may arise in multiple variables. In this meta-analysis, we consider the cases of the productivity of firms, the wages of workers, and rental prices. As discussed above, spillovers are likely to arise in terms of productivity in a first instance, as co-workers (in the same firm, industry and or region) benefit from the skills of more educated individuals. This will increase the productivity of those co-workers and their firms. Such productivity increases will then lead to wage increases through some combination of market competition for more productive workers and rent sharing from firms to their workers. Higher wages and profits in regions driven by the spillovers can also translate into higher prices for goods in limited supply, such as land and housing, driving increases in rents. The latter increases may lead to some equalisation of real wages across locations that are endowed with workforces of different schooling levels.

4. *Spillover type:* The literature studies social returns to education as such (when considering the returns to individuals throughout the domain considered), and not only the private returns that accrue to each person, but also the external returns to education. These can be regarded as the additional returns on top of the private returns. In other words, $SR = PR + ER$, in which SR, PR and ER are, respectively, the social returns, the private returns and the external returns. If the model considered in the paper under analysis considers separately the individual's schooling (as an additional control variable) and then focuses on 'social' schooling, we regard the coefficient of the later as a measure of the external return, over and above the private return. If there is no such control for individual's schooling (and their private returns), then the coefficient (regarding total schooling in a given region, for instance) is regarded as a social return.

5. *Data set characteristics:* Studies draw on the main types of data sets used in (micro-)econometric analyses, namely single cross-sections, pooled cross-sections, and panel data. We compare our estimate along these three dimensions. The case of panel data sets are based on repeated observations of individuals or firms over time. We also considered the time period examined in each study as well as the size of the sample (number of observations).

6. *Estimation method:* Several models are estimated using OLS. However, many studies seek to address the potential endogeneity of their schooling variable through fixed effects (drawing on the repeated availability of the same individual observations over time), instrumental variables or other methods. In some cases, the studies draw on more than one such method, namely when combining instrumental variables and fixed effects. These data points (estimates) are classified correspondingly - unlike in the previous dimensions, the multiple potential outcomes are not mutually exclusive in this case, regarding the estimation method.

7. *Country income level:* As discussed in the Introduction, social returns to education may vary

depending on the income level of the country under analysis. We examine this by classifying countries in three income levels: high, upper-middle, lower-middle and low. Using the World Bank Atlas method, these correspond to GNI per capita in 2018 of \$12,376 or more, between \$3,996 and \$12,375, between \$1,026 and \$3,995, and \$1,025 or less, respectively.

8. *Schooling quantity and quality*: The underlying levels of schooling in each country may also influence its social returns. We consider both quantity and quality dimensions, measured in terms of the percentage of university graduates and the average schooling years (quantity) (3) and the average PISA scores across its three dimensions of maths, science and reading (quality). We used the measurement for the same year as that of the data when the social return is estimated or the closest possible (the latter in the case of the PISA scores). Across the countries and time periods considered, we find an average percentage of graduates of 14% and 10.2 years of schooling. These relatively large means reflect the large percentage of estimates from developed countries and recent years, following significant expansions of their education systems.

9. *Publication characteristics*: We are also interested in uncovering potential relationships between the publication characteristics of the article and the underlying social returns. We consider three different variables in this case, namely the year of publication, the number of (Google Scholar) citations, and the journal rank. The latter is measured using (1), a widely used journal ranking including over 300 journals in economics alone, considering five categories of journals (from 'recognised world-wide as exemplars of excellence', the top level, to 'recognised, but more modest standard in their field', the bottom level).

2.3 Descriptive statistics

The descriptive statistics of the variables described above are presented in **Table 1**. These statistics are computed from the 1,008 estimates from 31 empirical studies, treating each estimate equally. We find that most estimates are computed at the regional level (86%), while the industry and firm levels represent only 6% and 9% of the total, respectively. As to the measurement of the potential spillovers, 58% of the estimates use the average years of schooling and 42% use the share of college-educated workers.

Wages prove to be the key outcome of education spillovers (82% of the estimates), followed by (firm) productivity (15%), while rents play a residual role (only 2% of the cases). Most studies focus on the external returns to education as they consider specifications that control for individual schooling (64% of the observations). The remaining 36% do not control for individual schooling and are thus interpreted as generating social returns to education (including both direct and indirect effects).

In contrast to the previous dimensions, there is considerable dispersion in data types across our sample of studies: 25% of the estimates come from single cross-sections, 49% from pooled cross-sections, and 26% from panel data. Similarly, 35% of the estimates use OLS only, 21% use fixed effects (involving repeated observations of the units upon which the externalities may arise), 26% use instrumental variables, and the remaining 21% use different methods (GLS, Heckman’s selection and multilevel modelling). These statistics also indicate that in only 3% of the cases are there overlaps between the different methods, which correspond to the joint use of fixed effects and instrumental variables.

When considering the economic development of the countries studied, we find that the large majority (77%) are high-income. 19% are upper-middle income and only 4% are lower-middle income. These statistics highlight the geographical dispersion of the existing estimates, but also the limited evidence available from lower-income, developing countries. As discussed above, the latter may be the ones where social returns to education are the highest, given their relatively lower average schooling levels and the greater impact upon social variables.

Interestingly, the average number of observations used in each estimate is very large, at 292,000. Moreover, on average, the estimates concern data regarding the year of 1990, even if the full range of years is very wide, starting at 1950 and ending in 2010. In contrast, the year of (journal) publication of the articles is, on average, 2007, covering the period 1993 to 2020.

Finally, the journals where the estimates are published are typically well reputed, with an average ranking of 3.19, where 5 would be the maximum (namely the top five journals in economics) and 1 the minimum (typically national journals, with limited peer review). The (log) number of Google scholar citations is 4.63 (about 100), with a maximum of 7.49 (nearly 1,800).

3 Results

3.1 Methodology

To investigate the drivers of social returns, we perform a meta-analysis by estimating the following equation:

$$\hat{\beta}_{ij} = \alpha_0 + \sum_{k=1}^k \alpha_k X_{jk} + e_{ij}, \quad (1)$$

where $\hat{\beta}_{ij}$ is the i^{th} estimate effect from j^{th} study and X_{jk} are the meta-independent variables that follow from the study design and are described above.

However, note that in our analysis the estimated effects are not necessarily directly comparable as they are diverse along different dimensions. For instance, we consider in this study both the social returns to education and the external effects of education, depending on the approach adopted in each paper. In order to analyse all estimates jointly, we transform the estimated effects into partial correlation coefficients (PCC). The PCCs, which measure the association of a dependent variable and the independent variable, are widely used in economic meta-analyses to standardise effect sizes.

The formula we use is as follows:

$$PCC\hat{\beta}_{ij} = \frac{t_{ij}}{\sqrt{t_{ij}^2 + df_{ij}}}$$

, where t_{ij} is the t-statistic of the effect under study and df_{ij} are the degrees of freedom for the i^{th} estimation in the j^{th} paper. Note that several selected studies do not report the number of regressors used, which prevent us from considering the exact value of df . In these cases, we used an approximation of the number of regressors.² Finally, the standard errors of the PCCs, SE, are calculated as:

$$SE(PCC_{ij}) = \sqrt{\frac{1 - PCC_{ij}^2}{df_{ij}}}$$

We present the resulting descriptive statistics of the t-statistics and PCCs in **Table 2**. We find that 62.4% of the estimates indicate a significantly positive effect (at the 5% level), 33.9% are insignificant, and 3.7% are significantly negative.³ The mean and median t-statistics for the full sample are 3.693 and 2.462, respectively, while the PCC values have a mean and median values of 0.041 and 0.015. The distributions of t-statistics and PCCs are presented in **Figures B.2** and **B.3**.

PCC values by study characteristics are presented in **Table B.1**, including information on the number of estimates and their range. This analysis indicates some relevant patterns including that: (1) PCCs are higher at the firm level (0.076); (2) PCCs tend to be lower under fixed-effects (0.016) and IV (0.034) estimators, compared to the OLS estimator (0.043); (3) PCCs are higher when measuring social returns (0.043) than when measuring exclusively spillovers (0.039); and (4) PCCs are higher in low-income economies (0.070, compared to 0.037 in high-income economies). Next we test the robustness of these patterns by considering the multiple dimensions simultaneously under the meta-regression below.

²One challenge concerns studies based on panel data sets with a large number of individuals and a small number of time periods, which may create a significant difference between N and df.

³See also **Figure B.1** which illustrates the range of estimates over the years in which they were published.

3.2 Drivers of social returns

Our meta-regression considers the following six estimators: OLS, not assigning any weight to each estimate; meta-regression fixed-effects estimator, which weights each estimate by the within-study variance; meta-regression random-effects estimator, which weights each estimate by the within-study variance plus the between-studies variance; and the weighted least square (WLS) estimator, which uses either the inverse of the number of estimates reported in the study as a sampling weight, or the inverse of the standard error of PCCs and the imputed quality level of the journal where the study was published as an analytical weight. Note that we do not consider the variables that are generated after the study (journal rank, year of publication, and number of citations) as potential explanatory variables of the PCC. OLS and WLS procedures calculate standard errors clustering at the study level.⁴

Our main results, based on the estimation of equation 1, are presented in **Table 3**. These prove to be generally very robust across all the specifications. In particular, we find that: First, compared to the regional level (i.e. the benchmark category), spillovers are higher at the firm level, with an increase of between 0.025 and 0.065. According to our discussion above, this positive effect may arise because people engage much more directly and intensively within firms, including through face-to-face communication. This facilitates deeper interactions, through which the human capital obtained from schooling can more easily spillover to other individuals.

Second, we do not find systematic differences in the estimates of spillovers when using the share of college-educated workers (as opposed to the average years of schooling). This result may be driven from the fact that most variation in average years of schooling is driven by variation in college-educated shares, at least in more developed countries (the large share of our sample).

Third, as expected, social returns are stronger than spillovers, with an increase ranging between 0.006 and 0.068. Social returns include both the individual own effects of schooling and the spillovers or externalities that schooling generates. To the extent that spillovers are positive, social returns would be expected to be greater than spillover estimates, as indeed we find in our analysis.

Fourth, compared to the benchmark OLS estimator, fixed-effects and IV estimators tend to generate a weaker spillover effect. This difference may be explained by upward biases potentially present in OLS analyses. For instance, more educated workers may tend to flow to regions, industries or firms that are more productive and that pay higher wages, regardless of the spillovers. Such positive correlation between co-workers schooling and wages would generate a positive bias in OLS analyses.

⁴We also standardized the two continuous variables, the number of observations and time period considered, in order to make their coefficients more easily interpretable. See Martins & Yang (11) for another illustration of this approach in a different literature context.

This interpretation is consistent with the fact that IV estimators tend to lead to the lower PCCs, even if not in all specifications.

Fifth, compared to high-income economies (the benchmark categories), upper-middle and lower-middle income economies tend to exhibit larger spillover effects, ranging between 0.027 and 0.079. Moreover, the effect is stronger in the case of lower-middle income economies than upper-middle income economies. This finding may be driven by the association between income and schooling levels and the diminishing scope for spillovers as schooling levels increase, as discussed above. If a large share of the workforce has higher levels of schooling, the scope for low-educated workers to learn from their high-educated coworkers is weaker, leading to lower spillovers as we find in our analysis.

Sixth, studies with more observations lead to smaller effects. Moreover, spillover effects tend to diminish with time, although the coefficient is not always significant. The latter result may be consistent with the finding above regarding country income levels: as a country increases its level of development and its average education, the scope for spillovers may also be reduced.

Finally, we find a potentially surprising result in that spillovers on productivity tend to be lower than those on wages. Presumably spillovers would first arise in productivity and then part of them would be accrued to workers in the form of higher wages. While we leave a more definite explanation for this result for future research, we speculate that at least part of the answer may involve measurement error in productivity. Measurement error is likely to be more significant in productivity than in wages, given that wages are directly observable whereas productivity needs to be estimated, using a range of variables that are not always present in the available data sets.⁵

3.3 Education levels

In an extensions of our main results, we also examine the role of schooling levels in education spillovers. As discussed before, we consider both quantity and quality measures and, with respect to the former, both tertiary schooling and overall schooling years. We test the hypothesis that tertiary schooling and or higher quality of schooling may generate stronger spillover effects.

We also consider the hypothesis that greater dispersion in the distribution of schooling may increase education spillovers. In the limit, if every worker has the same level of schooling, a higher level of schooling for the entire workforce may not generate spillovers in the sense that there are no schooling gaps that would facilitate learning from some workers with respect to their more-educated colleagues.

More specifically, we re-estimate our main specification from the previous equation 1 extended to include the three new variables above: share of tertiary graduates, average schooling years, and

⁵We also replicated the analysis above separately for estimates based on spillovers and social returns. The results are presented in Tables B.2 and B.3, respectively, and are very similar to the main findings discussed in this section.

average schooling quality (as measured from PISA tests). When controlling for (holding constant) average schooling, increases in the share of tertiary graduates must imply a relative increase in low-educated workers and a relative decrease in medium-educated workers and consequently an increase in the dispersion of schooling.⁶

Table 4 presents the results of this new specification. First, we find that the coefficients of the remaining variables than schooling levels remain unchanged from our main results, at least qualitatively when not also quantitatively. Second, the new results indicate that tertiary schooling has a statistically significant positive effect on education spillovers, while average schooling has a statistically significant negative effect. Finally, we do not find significant relationships between PISA scores and education spillovers.

As discussed above, we interpret the opposite signs of the coefficients of the two education quantity variables as supporting the relevance of dispersion in schooling as a driver of education spillovers. In this respect, increases in tertiary education may be a particularly relevant source of such externalities. This result is in line with our finding from the previous subsection regarding the role of economic development in that it is associated with large increases in tertiary education.

Our finding about the lack of a statistically significant positive association between schooling quality and education spillovers is potentially related to the same mechanism regarding education dispersion described above. Higher levels of schooling quality - as measured by PISA - may be particularly relevant for workers with compulsory schooling (when the PISA measurement takes place). This may boost their labour market perspectives in a way that reduces the potential learning from more schooling colleagues, thus diminishing the resulting spillovers. Any statistical noise in the PISA scores regarding schooling quality will also attenuate the effects that can be measured in our approach.

3.4 Publication bias

An important question in the context of meta-analyses is the extent to which particular types of estimates, namely those that are significant, are more likely to be published. Studies that find that there are spillovers to schooling may be regarded as more interesting and relevant in contrast to those that do not find significant effects, which would contribute to a skewed understanding of education externalities.

We obtain a first indication of the extent of publication bias in this literature by using the funnel plot methodology. The funnel plot (9) is a scatter plot of the reported study effect (i.e. PCCs in our case) against measures of study precision (i.e. the inverse standard errors of PCCs). In the absence of

⁶Note that in the new specification we control for schooling variables but we drop the economic development measures that we used before, given the potential significant multicollinearity between the two.

publication bias, the shape of the scatter plot should resemble a symmetric inverted funnel because the sampling error is random.

Figure B.4 presents a funnel plot based on the data that we collected and that was described above. We find that the plot is skewed, given that the right tail is much more prominent compared to the left tail. This suggests that a large share of negative estimates may be missing from the funnel plot, which would indicate publication bias as selection for a positive sign. However, this method of publication bias detection is based on visual inspection only, which may lack objectivity and accuracy.

In order to provide a more rigorous analysis, next we perform a trim-and-fill analysis in order to estimate if studies are potentially missing from our meta data set. This trim and fill method (6) involves dropping (trimming) the smaller studies causing funnel plot asymmetry and introducing (filling) those studies potentially missing from the meta-analysis because of publication bias. This method can generate the overall effect-size estimate using both observed and imputed studies. It can also deliver a funnel plot, in which omitted studies are imputed.

The result of trim and fill is presented in **Table 6**. We find that: (1) the mean spillover effect based on the 1,008 observed estimates is 0.005; (2) there are 327 estimates potentially missing and subsequently imputed; and (3) after including such estimates, we obtain a new estimate (based on the observed plus imputed 1,335 estimates) of the mean spillover effect of 0.004. **Figure B.5** also reveals that a substantial portion of negative estimates is missing from the funnel plot, due to publication bias in the form of selection for a positive sign. As indicated before, if these missing estimates were included in the meta-analysis, the resulting funnel plot would be more symmetrical.

In addition to funnel plot, we also make use of the meta-regression to further detect publication bias. We follow the FAT-PET-PEESE approach proposed by (17) to test whether the effect is genuine or influenced by publication bias. We start out by regressing the t-statistic of the k^{th} estimate on the inverse of the standard error ($1/SE$) using the following equation:

$$t_k = \beta_0 + \beta_1(1/SE_k) + u_k. \quad (2)$$

We then test the null hypothesis that the intercept term β_0 is equal to zero in equation 2 - this corresponds to the funnel-asymmetry test (FAT). If β_0 is statistically significantly different from zero, then the distribution of the effect sizes is regarded as asymmetric. However, regardless of publication selection, we are also able to identify an empirical effect by testing the null hypothesis that the coefficient β_1 is equal to zero in equation 2 - this corresponds to the precision-effect test (PET). If β_1 is statistically significantly different from zero, this indicates the presence of a genuine effect.

Moreover, we can also estimate the magnitude of the empirical effect corrected for publication

selection by estimating the following equation which has no intercept:

$$t_k = \beta_0 SE_k + \beta_1 (1/SE_k) + u_k \quad (3)$$

We test the null hypothesis that the intercept term β_1 is equal to zero in equation 3. If β_1 is statistically significantly different from zero, this indicates that a non-zero effect does actually exist in the literature, and β_1 can be regarded as its estimate. This corresponds to the precision-effect estimate with standard error (PEESE) test. To test the robustness of the β coefficients in equations 2 and 3 above, we use four estimators: OLS estimator, which does not assign any weight to each estimate; weighted least square (WLS) estimator, which uses either the inverse of the number of estimates reported in the study as a sampling weight or the inverse of the imputed quality level of the journal as an analytical weight; and a least square dummy variable (LSDV) estimator which controls for study-level fixed effects. All estimation procedures calculate standard errors at the study level.

The regression results based on the FAT-PET-PEESE test for publication bias are presented in **Table 5**. We find the following: First, in the FAT test, the null hypothesis that $\beta_0 = 0$ of equation 2 is rejected, which indicates the presence of publication bias. Second, in the PET test, the null hypothesis that $\beta_0 = 0$ of equation 2 is rejected as well, which implies that a true empirical effect does exist in the literature even if there is a publication selection bias. Finally, in the PEESE test, the coefficient β_1 in equation 3 is statistically significantly different from zero, which indicates the the magnitude of the empirical effect corrected for publication selection is significantly positive, ranging between 0.017 and 0.02. In conclusion, we find evidence of publication bias, although the FAT-PET-PEESE test also indicates that there is a genuinely positive spillover effect.

3.5 Citation bias

We are also interested in the novel concept of 'citation bias', which we define as the extent to which particular types of results tend to be cited more. Specifically, we want to know if significantly positive estimates tend to receive more citations than other types of results. This analysis is possible given our collection of citation counts from Google Scholar, as mentioned above.

The screening of studies in the literature can in some cases be more relevant at the citation stage than in terms of publication. Indeed, given the large number of journals around the world, most studies may eventually be published, even if only after several rejections from journals of possibly increasingly lower average standing. In this case, the extent to which a paper is cited may be a more relevant dimension of bias in its impact on the literature.

We propose a simple method to shed light on the question above, regarding potentially higher

citations for studies of particular characteristics, namely significantly positive spillovers. Specifically, we regress the log number of citations on the respective PCCs in the context of the following equation:

$$\log C_j = \alpha_0 + \alpha_1 PCC_{ij} + \alpha_2 Z_j + u_{ij}, \quad (4)$$

where $\log C_j$ is log citations of the j^{th} study, and Z_j includes two characteristics of the study that may also affect its citations record, namely the publication age and the imputed journal quality. Studies published in journals of higher standing and less recently will typically have more citations - it therefore may be useful to control for these variables, even if they may not be directly related to the magnitude of the PCC, especially if publication bias is not relevant, as indicated by our earlier findings. More importantly, if the coefficient α_1 is statically significant and positive, that would indicate that studies with larger spillovers are more likely to be cited. In other words, researchers would tend to pay more attention to significant, large results, and the literature would be developing a biased view of social returns to education.

First, when we run a scatter plot of the log number of citations against the average PCC value per study (**Figure B.6**), we find a negative but negligible correlation between the two variables, a preliminary result which is not consistent with citation bias. We then present the OLS estimates of different versions of equation 4 in **Table 7**. Column 1 does not include any controls, column 2 controls only for journal rank, column 3 controls for both journal rank and time since publication. We find in all cases that the PCCs are not positively correlated with the log number of citations. This result is consistent with our preliminary evidence of no citation bias in this literature. When considering all PCCs from the multiple estimates per study (and clustering standard errors accordingly), we again find similar results, indicating no positive relationship between more positive social returns and higher citation counts.

4 Conclusions

Education can generate important externalities and motivate the considerable involvement of governments in this sector around the world. Such externalities may also be stronger at lower levels of education. In this paper, we study the drivers of education externalities by conducting the first meta-analysis of the social returns to education literature. We analyse over 1,000 estimates from 31 journal articles published between 1993 and 2020, covering 15 countries in total, of which five are emerging or developing economies.

In a nutshell, our results indicate that: 1) spillovers fall with economic development; 2) spillovers

tend to be smaller under fixed-effects and IV estimators but larger when measured at the firm level; 3) tertiary schooling and schooling dispersion tends to increase spillovers; and 4) there is evidence of publication bias but not of citation bias.

Overall, we believe our results highlight the relevance of the literature on the social returns to education and the importance of its findings for national and international policy as well. In particular, these findings support the continuing investment in schooling - including tertiary education - in developing countries as they highlight the stronger social role of education at lower levels of economic development. Education may promote world development both from an individual private perspective but also in terms of the higher social returns that it generates for poorer countries. From an academic perspective, our results allow researchers to better compare their findings with respect to the existing literature, in particular studies developed under different methodologies.

References

- [1] ABS (2018), Academic journal guide, Technical report, Chartered Association of Business Schools, London.
- [2] Acemoglu, D. & Angrist, J. (2000), ‘How large are human-capital externalities? Evidence from compulsory schooling laws’, *NBER Macroeconomics Annual* **15**, 9–59.
- [3] Barro, R. J. & Lee, J. W. (2013), ‘A new data set of educational attainment in the world, 1950–2010’, *Journal of development economics* **104**, 184–198.
- [4] Cardoso, A. R., Guimaraes, P., Portugal, P. & Reis, H. (2018), The Returns to Schooling Unveiled, IZA Discussion Paper 11419, Institute of Labor Economics.
- [5] Demena, B. A. & van Bergeijk, P. A. (2017), ‘A meta-analysis of FDI and productivity spillovers in developing countries’, *Journal of Economic Surveys* **31**(2), 546–571.
- [6] Duval, S. & Tweedie, R. (2000), ‘Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis’, *Biometrics* **56**(2), 455–463.
- [7] Furnée, C. A., Groot, W. & van Den Brink, H. M. (2008), ‘The health effects of education: a meta-analysis’, *European Journal of Public Health* **18**(4), 417–421.
- [8] Huang, J., Van den Brink, H. M. & Groot, W. (2009), ‘A meta-analysis of the effect of education on social capital’, *Economics of Education Review* **28**(4), 454–464.

- [9] Light, R. J. & Pillemer, D. B. (1984), *Summing up: the science of reviewing research*, Harvard University Press, Cambridge, MA.
- [10] Marshall, A. (1980), *Principles of economics*, Macmillan, London: Macmillan.
- [11] Martins, P. & Yang, Y. (2009), ‘The impact of exporting on firm productivity: a meta-analysis of the learning-by-exporting hypothesis’, *Review of World Economics* **145**(3), 431–445.
- [12] Moretti, E. (2004), Human capital externalities in cities, in ‘Handbook of Regional and Urban Economics’, Vol. 4, Elsevier, pp. 2243–2291.
- [13] Rud, I., Van Klaveren, C., Groot, W. & Maassen van den Brink, H. (2013), Education and Youth Crime: a Review of the Empirical Literature, Working Papers 48, Top Institute for Evidence Based Education Research.
- [14] Rudd, J. B. (2000), Empirical evidence on human capital spillovers, Finance and Economics Discussion Series 2000-46, Board of Governors of the Federal Reserve System (U.S.).
- [15] Sianesi, B. & Reenen, J. V. (2003), ‘The returns to education: Macroeconomics’, *Journal of Economic Surveys* **17**(2), 157–200.
- [16] Stanley, T. D. (2005), ‘Beyond publication bias’, *Journal of Economic Surveys* **19**(3), 309–345.
- [17] Stanley, T. D. & Doucouliagos, H. (2012), *Meta-regression analysis in economics and business*, Vol. 5, Routledge.

Table 1: Definitions and summary statistics of explanatory variables

Variable	Description	Mean	SD	Min	Max
Spillover scope					
Regional level#	=1, if spillovers are estimated at regional level	0.86			
Industry level	=1, if spillovers are estimated at industry level	0.06			
Firm level	=1, if spillovers are estimated at firm level	0.09			
Spillover measure					
Average years of schooling#	=1, if aggregate human capital is measured by average years of schooling	0.58			
Share of college-educated workers	=1, if aggregate human capital is measured by share of college-educated workers	0.42			
Spillover outcome					
Wages#	=1, if estimate is based on wages	0.82			
Rental prices	=1, if estimate is based on rents	0.02			
Productivity (firms)	=1, if estimate is based on productivity	0.15			
Spillover type					
Social returns to education	=1, if estimated model does not control for individual schooling	0.36			
External returns to education#	=1, if estimated model controls for individual schooling	0.64			
Data set type					
Cross-section	=1, if data are cross-section data	0.25			
Pooled cross-sections#	=1, if data are pooled cross-sections	0.49			
Panel data	=1, if data are panel data	0.26			
Estimation method					
OLS#	=1, if estimation method is OLS	0.35			
FE	=1, if estimation method uses fixed effects	0.21			
IV	=1, if estimation method uses instrumental variables	0.26			
Other methods	=1, if estimation method is different from above	0.21			
Country type					
High-income #	=1, if data come from high-income country	0.77			
Upper-middle income	=1, if data come from upper-middle-income country	0.19			
Lower-middle income	=1, if data come from lower-middle-income country	0.04			
National schooling quantity and quality					
Tertiary education completed	% of tertiary education graduates	0.14	0.08	0.01	0.31
Avg. years of total schooling	Average years of total schooling	10.15	2.31	4.54	13.42

Continued on next page

Table 1 – continued from previous page

Variable	Description	Mean	SD	Min	Max
Avg. PISA score	Average PISA score of mathematics, science and reading	495.24	22.63	373	528
Other data characteristics					
Sample size	No. of observations (in thousands)	292.23	804.11	0.05	8034.7
Time period considered	Average year of time period considered	1990.59	10.38	1950	2010
Year of publication	Year of journal publication	2007.12	5.47	1993	2020
Journal rank	Based on ABS (2018) list (from 5 (max) to 1 (min))	3.19	1.10	1	5
Log citations	Log of number of Google Scholar citations (May 2020)	4.63	1.79	0	7.49

Note: When the grouped variables include all possible categories, the benchmark categories omitted in the subsequent analysis are indicated by #.

Low-income economies are defined as those with a GNI per capita, calculated using the World Bank Atlas method, of \$1,025 or less in 2018; lower middle-income economies are those with a GNI per capita between \$1,026 and \$3,995; upper middle-income economies are those with a GNI per capita between \$3,996 and \$12,375; high-income economies are those with a GNI per capita of \$12,376 or more. The data on 'Tertiary education completed' and 'Avg. years of total schooling' is drawn from Barro and Lee (3). The data on 'Avg. PISA score' is drawn from OECD database. The PISA score is unavailable for South Africa and Kenya. Also, we excluded the PISA score for China, which is only available from 2009 (and refers to a specific region of the country), while the period in the two Chinese studies considered in the meta-analysis is 1995–1998. The total number of observations is 1,008.

Table 2: Descriptive statistics for effect size variables

	t-statistics	PCCs
Mean	3.693	0.041
Median	2.462	0.015
Maximum	89.000	0.507
Minimum	-6.000	-0.223
Std.Dev.	5.567	0.084
5%	-1.125	-0.038
10%	-0.227	-0.001
90%	9.250	0.123
95%	12.160	0.215
Observations	1,004	1,004

Note: When calculating the PCC, we assume that the degrees of freedom are equal to the number of observations when the studies do not report degrees of freedom nor the number of regressors used.

Table 3: Meta-regression analysis of heterogeneity among studies of education spillovers

Dependent variable: PCCs		OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Scopes of spillovers (Regional level#)							
Industry level		-0.0233* (0.012)	0.0024*** (0.001)	-0.0236*** (0.008)	-0.0109 (0.011)	-0.0047 (0.007)	-0.0187* (0.011)
Firm level		0.0369** (0.018)	0.0257*** (0.000)	0.0291*** (0.008)	0.0659*** (0.024)	0.0385*** (0.009)	0.0272** (0.012)
Measures of spillovers (Average years of schooling#)							
Share of college-educated workers		0.0069 (0.013)	0.0023*** (0.000)	0.0012 (0.005)	-0.0042 (0.014)	-0.0016 (0.005)	0.0033 (0.010)
Outcomes effect of Spillovers (Wages of workers#)							
Rental prices of land		-0.0865*** (0.028)	-0.0005 (0.001)	-0.0805*** (0.014)	-0.0859*** (0.028)	-0.0216* (0.012)	-0.0827*** (0.026)
Productivity of firms		-0.0783*** (0.022)	-0.0105*** (0.001)	-0.0726*** (0.010)	-0.0687*** (0.024)	-0.0290** (0.011)	-0.0716*** (0.022)
Types of education spillovers (External returns#)							
Social returns		0.0680*** (0.020)	0.0064*** (0.000)	0.0628*** (0.007)	0.0584*** (0.021)	0.0207** (0.009)	0.0665*** (0.019)
Estimation methods (OLS#)							
FE		-0.0212* (0.011)	-0.0064*** (0.000)	-0.0201** (0.008)	-0.0186* (0.011)	-0.0102*** (0.003)	-0.0192* (0.009)
IV		-0.0390*** (0.014)	-0.0051*** (0.000)	-0.0289*** (0.005)	-0.0412** (0.016)	-0.0090*** (0.003)	-0.0267** (0.013)
Other estimation		0.0286* (0.015)	-0.0043*** (0.000)	0.0282*** (0.006)	-0.0039 (0.016)	0.0016 (0.012)	0.0310** (0.015)
Types of dataset (Pooled cross-sections#)							
Panel data		-0.0000 (0.012)	0.0072*** (0.000)	0.0050 (0.008)	-0.0032 (0.016)	0.0122** (0.005)	-0.0005 (0.012)
Cross-section		-0.0046 (0.014)	-0.0022*** (0.000)	-0.0008 (0.006)	-0.0089 (0.020)	-0.0025 (0.005)	0.0013 (0.012)

Continued on next page

Table 3 – continued from previous page

	OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Types of country analysed (High-income economies#)						
Lower-middle income economies	0.0581** (0.025)	0.0319*** (0.003)	0.0422*** (0.012)	0.0797** (0.035)	0.0382*** (0.012)	0.0460** (0.020)
Upper-middle income economies	0.0364*** (0.013)	0.0277*** (0.001)	0.0311*** (0.006)	0.0379** (0.016)	0.0316** (0.012)	0.0302*** (0.011)
Other data characteristic						
Time period considered	-0.0133 (0.008)	-0.0003*** (0.000)	-0.0094*** (0.002)	-0.0147* (0.008)	-0.0012 (0.001)	-0.0140 (0.008)
Data size	-0.0184*** (0.005)	-0.0044*** (0.000)	-0.0172*** (0.002)	-0.0131*** (0.003)	-0.0079*** (0.002)	-0.0175*** (0.005)
Constant	0.0260*** (0.009)	0.0096*** (0.000)	0.0247*** (0.004)	0.0346*** (0.011)	0.0135*** (0.003)	0.0232*** (0.008)
Observations	1004	1004	1004	1004	1004	1004
R-squared	0.175		0.281	0.263	0.226	0.167

Note: When the grouped variables include all possible categories, the categories omitted in the subsequent analysis (the benchmark categories) are indicated by #. OLS denotes ordinary least square, which does not assign any weight to each estimate. FE denotes meta-regression fixed-effects, which weight the each estimate by the within-study variance (column 2). RE denotes meta-regression random-effects, which weight the each estimate by the within-study variance plus the between-studies variance (column 3). WLS denotes weighted least square, which uses either the inverse of the number of estimates reported in the study as a sampling weight (column 4), or the inverse of the standard error of PCCs (column 5) and the quality level of the study as an analytical weight (column 6). (1) (4) (5) and (6) procedures calculate standard errors clustering at the study level. We standardized the two continuous variables, i.e. the number of observations and time period considered, in order to make them more interpretable. The significance levels of t-test are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: The effects of schooling quantity and quality on education spillovers in developed countries

Dependent variable: PCCs		OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Scopes of spillovers (Regional level#)							
Industry level		-0.0458* (0.022)	0.0006 (0.001)	-0.0358** (0.014)	-0.0366** (0.014)	-0.0035 (0.008)	-0.0407* (0.021)
Firm level		0.0518*** (0.014)	0.0248*** (0.000)	0.0497*** (0.010)	0.0510*** (0.015)	0.0443*** (0.007)	0.0438*** (0.011)
Measures of spillovers (Average years of schooling#)							
Share of college-educated workers		0.0134 (0.021)	0.0044*** (0.000)	0.0056 (0.008)	-0.0011 (0.020)	0.0017 (0.007)	0.0099 (0.017)
National schooling quantity and quality							
Tertiary education completed (%)		0.7977** (0.312)	0.0838*** (0.004)	0.6448*** (0.137)	0.5532* (0.275)	0.2759** (0.105)	0.7298** (0.298)
Avg. years of total schooling		-0.0307* (0.016)	-0.0048*** (0.000)	-0.0223*** (0.006)	-0.0191 (0.014)	-0.0087** (0.004)	-0.0278* (0.016)
Avg. PISA score		0.0000 (0.001)	-0.0002*** (0.000)	-0.0002 (0.000)	0.0001 (0.001)	-0.0006 (0.000)	-0.0000 (0.001)
Outcomes effect of spillovers (Wages of workers#)							
Rental prices of land		-0.0672*** (0.017)	-0.0093*** (0.002)	-0.0591*** (0.017)	-0.0703*** (0.018)	-0.0302** (0.012)	-0.0661*** (0.016)
Productivity of firms		-0.0759*** (0.023)	-0.0078*** (0.001)	-0.0689*** (0.012)	-0.0898*** (0.027)	-0.0278*** (0.010)	-0.0815*** (0.023)
Types of education spillovers (External returns#)							
Social returns		0.0665*** (0.018)	0.0072*** (0.000)	0.0591*** (0.008)	0.0688*** (0.020)	0.0236** (0.009)	0.0662*** (0.018)
Estimation methods (OLS#)							
FE		-0.0132 (0.015)	-0.0067*** (0.000)	-0.0099 (0.009)	-0.0130 (0.018)	-0.0030 (0.004)	-0.0112 (0.011)
IV		-0.0458** (0.019)	-0.0054*** (0.000)	-0.0336*** (0.006)	-0.0585** (0.023)	-0.0091*** (0.003)	-0.0413** (0.019)

Continued on next page

Table 4 – continued from previous page

	OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Other estimation	0.0286** (0.013)	0.0014*** (0.001)	0.0322*** (0.007)	0.0072 (0.023)	0.0178 (0.015)	0.0305** (0.012)
Types of dataset (Pooled cross-sections#)						
Panel data	-0.0221 (0.017)	0.0027*** (0.000)	-0.0157 (0.011)	-0.0045 (0.026)	-0.0062 (0.010)	-0.0142 (0.018)
Cross-section	-0.0335** (0.012)	-0.0058*** (0.000)	-0.0309*** (0.009)	-0.0167 (0.018)	-0.0223* (0.011)	-0.0255** (0.011)
Other data characteristic						
Time period considered	-0.0215* (0.012)	-0.0004*** (0.000)	-0.0159*** (0.003)	-0.0209* (0.012)	-0.0036 (0.002)	-0.0224* (0.012)
Data size	-0.0172*** (0.005)	-0.0045*** (0.000)	-0.0159*** (0.003)	-0.0138*** (0.003)	-0.0078*** (0.002)	-0.0175*** (0.005)
Constant	0.2140 (0.215)	0.1520*** (0.008)	0.2608** (0.121)	0.0861 (0.269)	0.3726* (0.189)	0.2205 (0.197)
Observations	777	777	777	777	777	777
R-squared	0.211		0.332	0.266	0.245	0.213

Note: When the grouped variables include all possible categories, the categories omitted in the subsequent analysis (the benchmark categories) are indicated by #. OLS denotes ordinary least square, which does not assign any weight to each estimate. FE denotes meta-regression fixed-effects, which weight the each estimate by the within-study variance (column 2). RE denotes meta-regression random-effects, which weight the each estimate by the within-study variance plus the between-studies variance (column 3). WLS denotes weighted least square, which uses either the inverse of the number of estimates reported in the study as a sampling weight (column 4), or the inverse of the standard error of POCs (column 5) and the quality level of the study as an analytical weight (column 6). (1) (4) (5) and (6) procedures calculate standard errors clustering at the study level. We standardized the two continuous variables, i.e. the number of observations and time period considered, in order to make them more interpretable. The significance levels of t-test are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Testing for publication bias
FAT-PET test: $t = \beta_0 + \beta_1(1/SE) + u$

	OLS (1)	WLS (2)	WLS (3)	LSDV (4)
Intercept($H_0:\beta_0=0$)	2.650*** (0.429)	2.958*** (0.488)	2.678*** (0.479)	2.920*** (0.022)
1/SE($H_0:\beta_1=0$)	0.014** (0.005)	0.014** (0.006)	0.014*** (0.005)	0.017*** (0.005)
Observations	1002	1002	1002	1002
R-squared	0.186	0.178	0.169	0.413

PEESE test: $t = \beta_0 SE + \beta_1(1/SE) + u$

	OLS (1)	WLS (2)	WLS (3)	LSDV (4)
SE	1.856*** (0.573)	1.687*** (0.469)	2.132*** (0.765)	-0.395 (1.263)
1/SE($H_0:\beta_1=0$)	0.019*** (0.006)	0.020*** (0.006)	0.020*** (0.007)	0.017*** (0.005)
Observations	1002	1002	1002	1002
R-squared	0.323	0.324	0.303	0.576

Note: OLS denotes ordinary least square, which does not assign any weight to each estimate. WLS denotes weighted least square, which uses either the inverse of the number of estimates reported in the study as a sampling weight (column 2), or the inverse of the quality level of the study as an analytical weight (column 3). LSDV denotes the least square dummy variable, which controls for study-level fixed effects. All estimation procedures calculate standard errors at the study level. The significance levels of t-test are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Trim-and-fill analysis of publication bias

Estimates	No. of estimates	Effect size	[95% Conf.Interval]
Observed	1008	0.005	[0.005, 0.005]
Observed+imputed	1335	0.004	[0.004, 0.004]

Note: The analysis used PCCs as effect sizes based on common-effect model. "Imputed" denotes potential estimates which are missed because of publication bias.

Table 7: Testing for citation bias, OLS models
Dependent variable: Log no. of citations

	(1)	(2)	(3)
PCCs	-4.170 (6.205)	-5.761 (5.599)	-9.329*** (3.097)
Journal rank		0.955*** (0.235)	0.548*** (0.162)
Publication age			0.228*** (0.035)
Constant	4.050*** (0.527)	1.505** (0.656)	0.047 (0.608)
Observations	31	31	31
R-squared	0.013	0.319	0.778

Note: All estimation procedures calculate standard errors at the study level. The significance levels of t-test are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendices

A List of selected papers in the meta-analysis

References

- Acemoglu, D., & Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER macroeconomics annual*, 15, 9-59.
- Battu, H., Belfield, C. R., & Sloane, P. J. (2003). Human capital spillovers within the workplace: evidence for Great Britain. *oxford Bulletin of Economics and Statistics*, 65(5), 575-594.
- Bentsen, K. H., Munch, J. R., & Schaur, G. (2019). Education spillovers within the workplace. *Economics Letters*, 175, 57-59.
- Bratti, M., & Leombruni, R. (2014). Local human capital externalities and wages at the firm level: Evidence from Italian manufacturing. *Economics of Education Review*, 41, 161-175.
- Broersma, L., Edzes, A. J., & Van Dijk, J. (2016). Human capital externalities: effects for low-educated workers and low-skilled jobs. *Regional Studies*, 50(10), 1675-1687.
- Canton, E. (2009). Human capital externalities and proximity: evidence from repeated cross-sectional data. *De Economist*, 157(1), 79-105.
- Ciccone, A., & Peri, G. (2006). Identifying human-capital externalities: Theory with applications. *The Review of Economic Studies*, 73(2), 381-412.
- Czaller, L. (2017). Increasing social returns to human capital: evidence from Hungarian regions. *Regional Studies*, 51(3), 467-477.
- Dalmazzo, A., & De Blasio, G. (2007). Production and consumption externalities of human capital: An empirical study for Italy. *Journal of Population Economics*, 20(2), 359-382.
- Dalmazzo, A., & De Blasio, G. (2007). Social returns to education in Italian local labor markets. *The Annals of Regional Science*, 41(1), 51-69.
- Fan, W., Ma, Y., & Wang, L. (2015). Do we need more public investment in higher education? Estimating the external returns to higher education in China. *Asian Economic Papers*, 14(3), 88-104.

- Groot, S. P., & de Groot, H. L. (2020). Estimating the skill bias in agglomeration externalities and social returns to education: Evidence from dutch matched worker-firm micro-data. *De Economist*, 168(1), 53-78.
- Heuermann, D. (2011). Human capital externalities in Western Germany. *Spatial Economic Analysis*, 6(2), 139-165.
- Iranzo, S., & Peri, G. (2009). Schooling externalities, technology, and productivity: Theory and evidence from US states. *The Review of Economics and Statistics*, 91(2), 420-431.
- Kimenyi, M. S., Mwabu, G., & Manda, D. K. (2006). Human capital externalities and private returns to education in Kenya. *Eastern Economic Journal*, 32(3), 493-513.
- Kirby, S., & Riley, R. (2008). The external returns to education: UK evidence using repeated cross-sections. *Labour economics*, 15(4), 619-630.
- Liu, Z. (2007). The external returns to education: Evidence from Chinese cities. *Journal of Urban Economics*, 61(3), 542-564.
- Liu, Z. (2014). Human capital externalities in cities: evidence from Chinese manufacturing firms. *Journal of economic geography*, 14(3), 621-649.
- Michaud, P. C., & Vencatachellum, D. (2003). Human capital externalities in South Africa. *Economic Development and Cultural Change*, 51(3), 603-628.
- Martins, P. S., & Jin, J. Y. (2010). Firm-level social returns to education. *Journal of Population Economics*, 23(2), 539-558.
- Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review*, 94(3), 656-690.
- Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of econometrics*, 121(1-2), 175-212.
- Muller, C., & Nordman, C. J. (2011). Within-firm human capital externalities in Tunisia. *The Journal of Development Studies*, 47(4), 657-675.
- Muravyev, A. (2008). Human capital externalities Evidence from the transition economy of Russia
1. *Economics of Transition*, 16(3), 415-443.
- Purnastuti, L., & Salim, R. (2015). Externalities and the social return to education in Indonesia. *Australian Journal of Labour Economics*, 18(1), 53.

- Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: evidence from the cities. *Journal of urban economics*, 34(3), 380-400.
- Rudd, J. B. (2000). Empirical evidence on human capital spillovers.
- Sand, B. M. (2013). A re-examination of the social returns to education: Evidence from US cities. *Labour Economics*, 24, 97-106.
- Wheeler, C. H. (2007). Do localization economies derive from human capital externalities?. *The Annals of Regional Science*, 41(1), 31-50.
- Winters, J. V. (2014). STEM graduates, human capital externalities, and wages in the US. *Regional Science and Urban Economics*, 48, 190-198.
- Wirz, A. M. (2008). Private returns to education versus education spill-over effects. *Empirical Economics*, 34(2), 315-342.

B Additional tables and figures

Table B.1: Education spillover: PCCs by study characteristics

Study Characteristics	Observations	Mean	Median	Minimum	Maximum
Scope					
Regional level#	862	0.039	0.014	-0.223	0.507
Industry level	48	0.007	0.007	0.006	0.009
Firm level	86	0.076	0.048	-0.008	0.477
Regional-level and Industry-level	8	0.050	0.039	0.023	0.112
Measure					
Average years of schooling#	580	0.047	0.018	-0.223	0.506
Share of college-educated workers	424	0.032	0.011	-0.204	0.507
Outcome					
Wages of workers#	825	0.045	0.015	-0.223	0.507
Rental prices of land	23	0.022	0.019	0.013	0.041
Productivity of firms	156	0.018	0.012	-0.053	0.136
Type					
External returns to education#	637	0.039	0.016	-0.223	0.507
Social returns to education	367	0.043	0.013	-0.204	0.477
Data Set Type					
Pooled cross-sections#	490	0.044	0.008	-0.223	0.507
Cross-section	249	0.057	0.041	-0.060	0.477
Panel data	265	0.019	0.012	-0.053	0.470
Estimation method					
OLS#	354	0.043	0.017	-0.137	0.477
FE	181	0.016	0.014	-0.053	0.136
IV	234	0.034	0.003	-0.204	0.311
FE and IV	30	0.016	0.010	-0.006	0.056
Other estimation	205	0.069	0.042	-0.223	0.507
Country Type					
High-income economies#	777	0.037	0.010	-0.223	0.507
Upper-middle income economies	189	0.049	0.037	-0.053	0.263
Lower-middle income economies	38	0.070	0.056	-0.060	0.306

Note: When the grouped variables include all possible categories, the categories omitted in the subsequent analysis (the benchmark categories) are indicated by #.

Table B.2: Meta-regression analysis of heterogeneity among studies of spillovers

Dependent variable: PCCs		OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Scopes of spillovers (Regional level#)							
Industry level		-0.0239*** (0.006)	0.0041*** (0.001)	-0.0217*** (0.008)	-0.0008 (0.005)	-0.0030 (0.003)	-0.0195*** (0.006)
Firm level		0.0191** (0.010)	0.0289*** (0.001)	0.0149* (0.009)	0.0674*** (0.016)	0.0262*** (0.005)	0.0174** (0.007)
Measures of spillovers (Average years of schooling#)							
Share of college-educated workers		0.0051 (0.008)	0.0016*** (0.000)	0.0056 (0.006)	-0.0028 (0.005)	0.0024 (0.003)	0.0053 (0.006)
Estimation methods (OLS#)							
FE		-0.0127 (0.008)	0.0011 (0.001)	-0.0138 (0.015)	-0.0284*** (0.008)	-0.0092** (0.004)	-0.0110* (0.002)
IV		-0.0120*** (0.004)	-0.0077*** (0.000)	-0.0110* (0.006)	-0.0002 (0.008)	-0.0086*** (0.001)	-0.0020 (0.005)
Other estimation		0.0381*** (0.008)	-0.0055*** (0.000)	0.0360*** (0.006)	-0.0021 (0.008)	0.0076** (0.004)	0.0368*** (0.008)
Types of country analysed (High-income economies#)							
Lower-middle income economies		0.0461*** (0.014)	0.0357*** (0.003)	0.0336*** (0.012)	0.0609*** (0.019)	0.0385*** (0.008)	0.0312** (0.013)
Upper-middle income economies		0.0302*** (0.006)	0.0370*** (0.001)	0.0284*** (0.008)	0.0269** (0.009)	0.0395*** (0.006)	0.0207*** (0.006)
Types of dataset (Pooled cross-sections#)							
Panel data		-0.0066 (0.010)	0.0038*** (0.001)	-0.0038 (0.015)	0.0075 (0.008)	0.0094** (0.004)	-0.0082 (0.010)
Cross-section		0.0064 (0.006)	-0.0038*** (0.000)	0.0083 (0.007)	0.0029 (0.008)	0.0015 (0.003)	0.011* (0.006)
Other data characteristic							
Time period considered		-0.0096*** (0.002)	-0.0006*** (0.000)	-0.0087*** (0.003)	-0.0084*** (0.003)	-0.0024*** (0.001)	-0.0110*** (0.002)

Continued on next page

Table B.2 – continued from previous page

	OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Data size	-0.0106*** (0.003)	-0.0011*** (0.000)	-0.0100*** (0.003)	-0.0098*** (0.002)	-0.0028*** (0.001)	-0.0131*** (0.003)
Constant	0.0203*** (0.004)	0.0093*** (0.000)	0.0189*** (0.005)	0.0213*** (0.003)	0.0105*** (0.001)	0.0189*** (0.003)
Observations	637	637	637	637	637	637
R-squared	0.173		0.267	0.334	0.235	0.165

Note: When the grouped variables include all possible categories, the categories omitted in the subsequent analysis (the benchmark categories) are indicated by #. OLS denotes ordinary least square, which does not assign any weight to each estimate. FE denotes meta-regression fixed-effects, which weight the each estimate by the within-study variance (column 2). RE denotes meta-regression random-effects, which weight the each estimate by the within-study variance plus the between-studies variance (column 3). WLS denotes weighted least square, which uses either the inverse of the number of estimates reported in the study as a sampling weight (column 4), or the inverse of the standard error of PCCs (column 5) and the quality level of the study as an analytical weight (column 6). (1) (4) (5) and (6) procedures calculate robust standard errors. We standardized the two continuous variables, i.e. the number of observations and time period considered, in order to make them more interpretable. The significance levels of t-test are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Meta-regression analysis of heterogeneity among studies of social returns

Dependent variable: PCCs						
	OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Scopes of spillovers (Regional level#)						
Firm level	0.190*** (0.030)	0.0094*** (0.001)	0.147*** (0.021)	0.126*** (0.030)	0.0671*** (0.016)	0.189*** (0.028)
Measures of spillovers (Average years of schooling#)						
Share of college-educated workers	0.0415** (0.020)	-0.0916*** (0.002)	0.0092 (0.011)	-0.0448 (0.034)	-0.0322*** (0.012)	0.0369* (0.022)
Outcomes effect of Spillovers (Wages of workers#)						
Rental prices of land	-0.0581** (0.026)	-0.0866*** (0.002)	-0.0804*** (0.017)	-0.159*** (0.033)	-0.0652*** (0.016)	-0.0545** (0.025)
Productivity of firms	-0.0031 (0.021)	-0.0011 (0.001)	-0.0046 (0.019)	-0.0103 (0.033)	-0.0143 (0.011)	-0.0016 (0.024)
Estimation methods (OLS#)						
FE	-0.0591*** (0.017)	-0.0026*** (0.000)	-0.0499*** (0.012)	-0.0606*** (0.022)	-0.0144*** (0.005)	-0.0297* (0.016)
IV	-0.0872*** (0.020)	-0.0036*** (0.000)	-0.0695*** (0.009)	-0.137*** (0.026)	-0.0111*** (0.003)	-0.0710*** (0.019)
Other estimation	-0.0295 (0.040)	0.0788*** (0.003)	0.0056 (0.029)	0.0157 (0.024)	0.0453** (0.023)	-0.0218 (0.042)
Types of country analysed (High-income economies#)						
Upper-middle income economies	0.0911*** (0.025)	-0.0543*** (0.002)	0.0543*** (0.015)	-0.0186 (0.035)	-0.0006 (0.013)	0.0896*** (0.026)
Types of dataset (Pooled cross-sections#)						
Panel data	-0.0892*** (0.025)	0.0054*** (0.001)	-0.0756*** (0.019)	-0.0635** (0.031)	-0.0185 (0.013)	-0.111*** (0.027)
Cross-section	-0.114*** (0.024)	0.0033*** (0.001)	-0.0950*** (0.017)	-0.0155 (0.041)	-0.0317*** (0.011)	-0.124*** (0.026)
Other data characteristic						
Time period considered	-0.0394*** (0.005)	0.0005* (0.001)	-0.0291*** (0.001)	-0.0107 (0.001)	-0.0086** (0.001)	-0.0436*** (0.001)

Continued on next page

Table B.3 – continued from previous page

	OLS (1)	FE (2)	RE (3)	WLS (4)	WLS (5)	WLS (6)
Data size	(0.009) -0.0145** (0.006)	(0.000) -0.0032*** (0.000)	(0.006) -0.0138*** (0.004)	(0.010) -0.0200*** (0.006)	(0.004) -0.0114*** (0.004)	(0.009) -0.0103 (0.007)
Constant	0.105*** (0.021)	0.100*** (0.002)	0.119*** (0.011)	0.189*** (0.031)	0.0834*** (0.015)	0.105*** (0.020)
Observations	367	367	367	367	367	367
R-squared	0.278		0.480	0.515	0.327	0.290

Note: When the grouped variables include all possible categories, the categories omitted in the subsequent analysis (the benchmark categories) are indicated by #. OLS denotes ordinary least square, which does not assign any weight to each estimate. FE denotes meta-regression fixed-effects, which weight the each estimate by the within-study variance (column 2). RE denotes meta-regression random-effects, which weight the each estimate by the within-study variance plus the between-studies variance (column 3). WLS denotes weighted least square, which uses either the inverse of the number of estimates reported in the study as a sampling weight (column 4), or the inverse of the standard error of PCCs (column 5) and the quality level of the study as an analytical weight (column 6). (1) (4) (5) and (6) procedures calculate robust standard errors. We standardized the two continuous variables, i.e. the number of observations and time period considered, in order to make them more interpretable. The significance levels of t-test are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

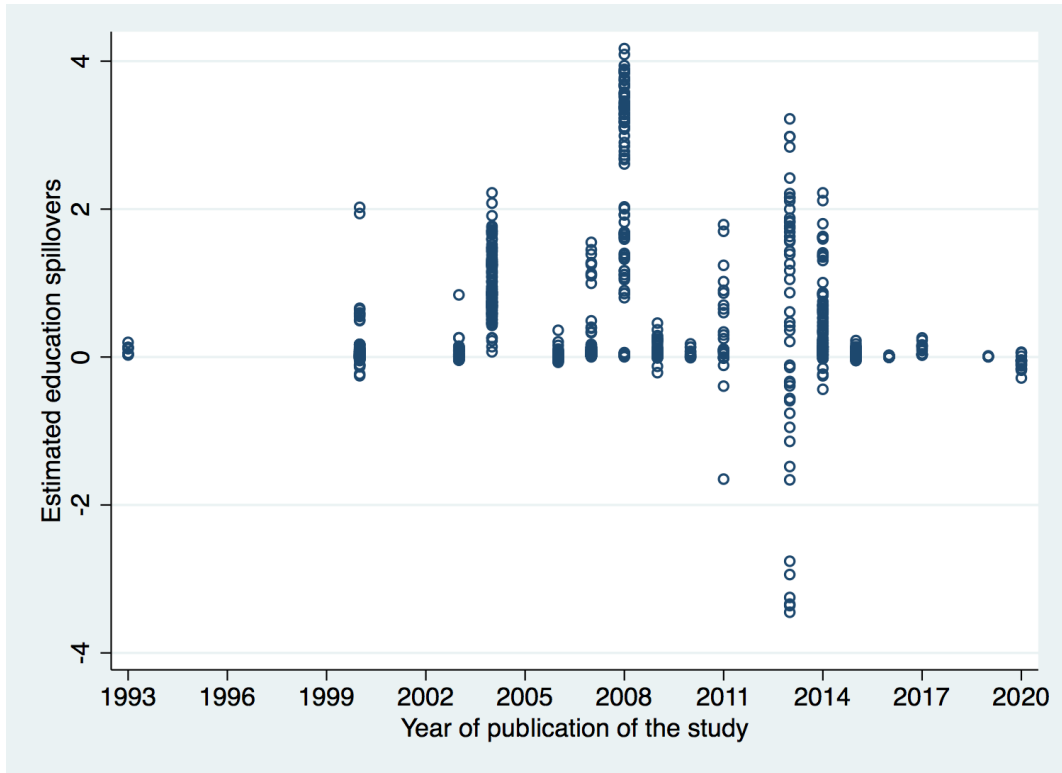


Figure B.1: Education spillover effects by the year of the publication
 Note: In order to get a clearer picture, we dropped 2 observations which have extreme high/low values. The number of estimates is 1,006.

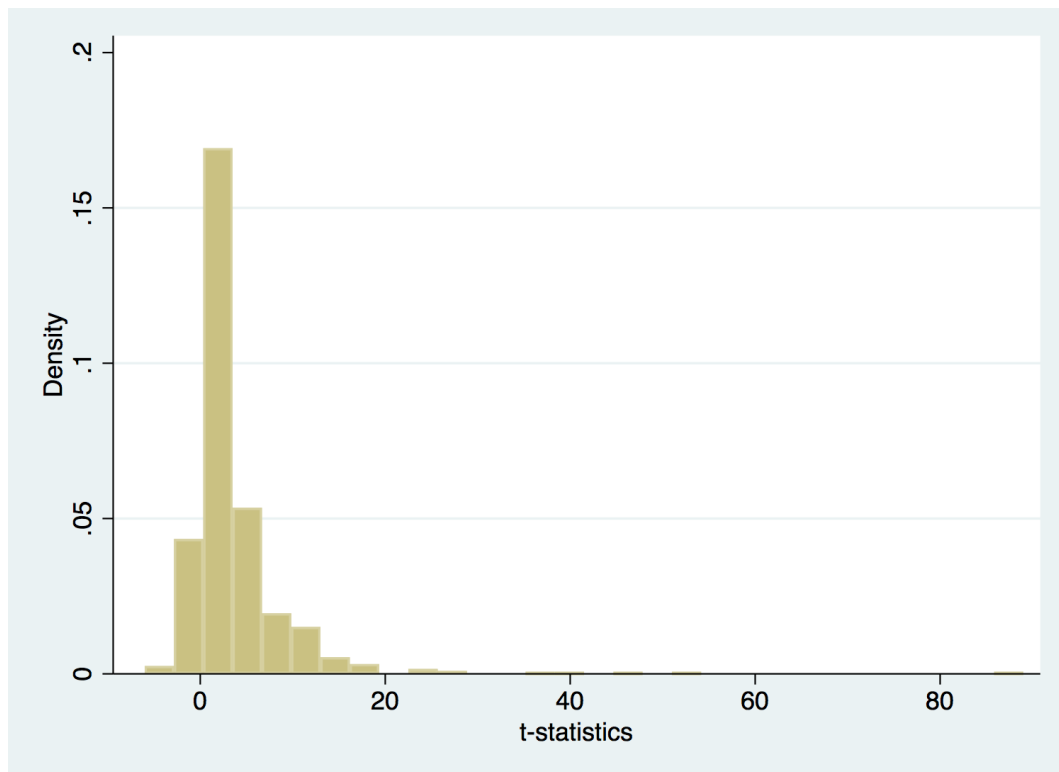


Figure B.2: The distribution of t-statistics

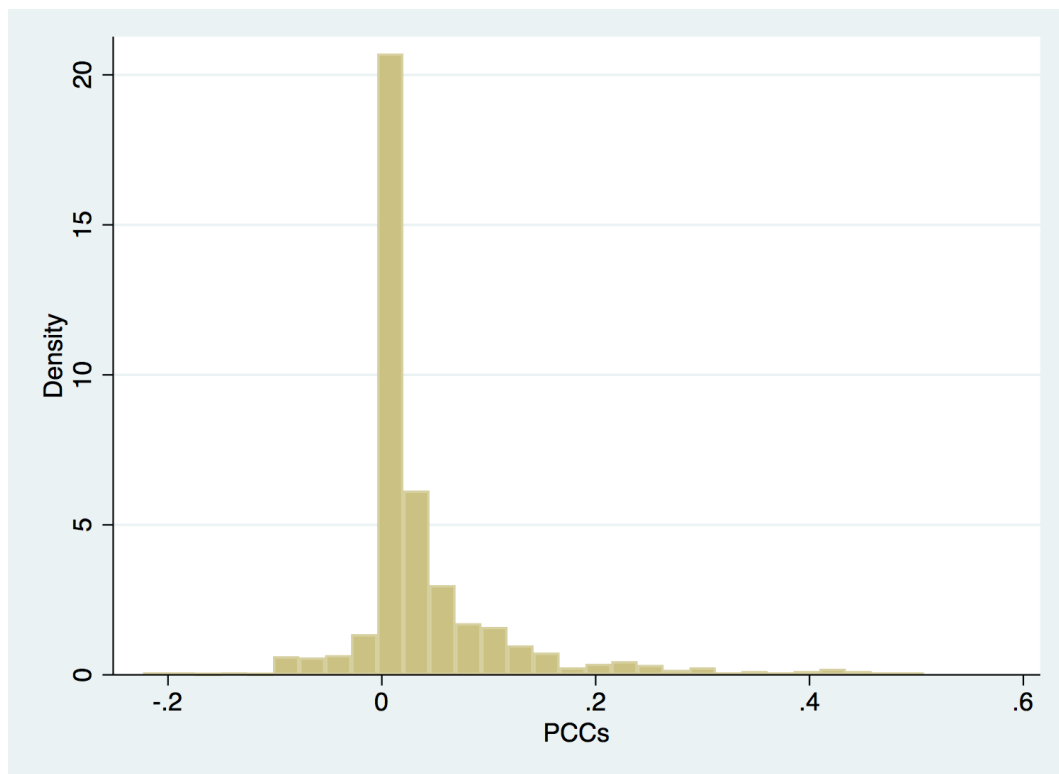


Figure B.3: The distribution of PCCs

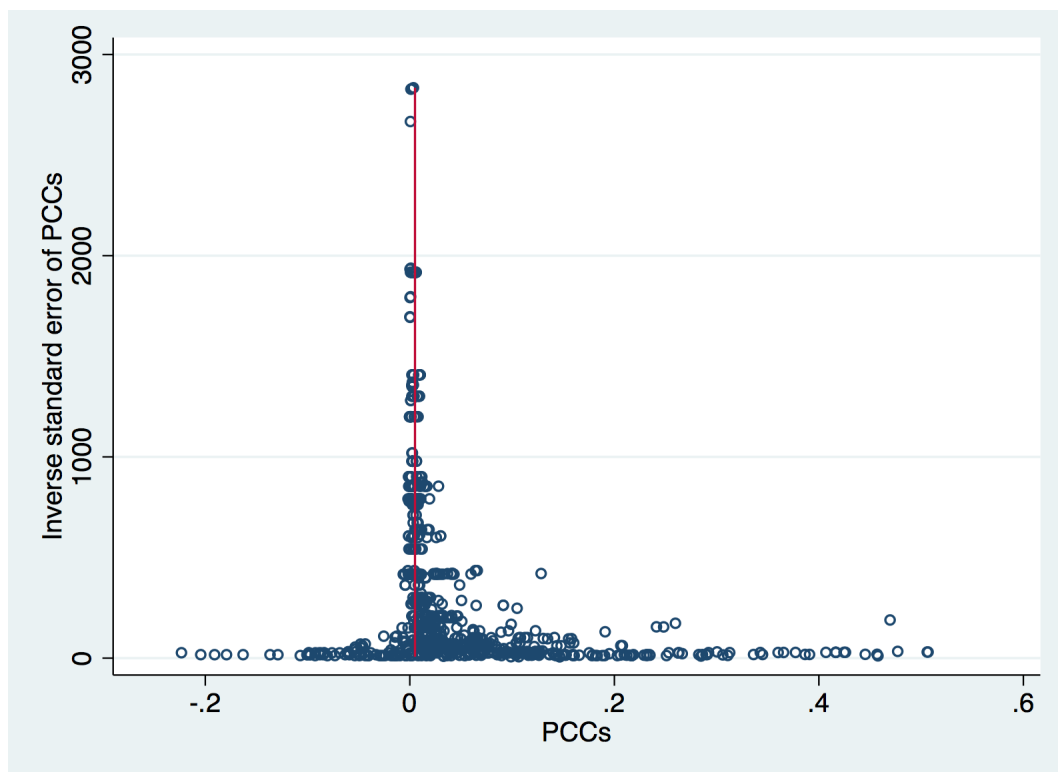


Figure B.4: Funnel-Asymmetry Plot

Note: The values of PCCs obtained are plotted against the inverse standard errors

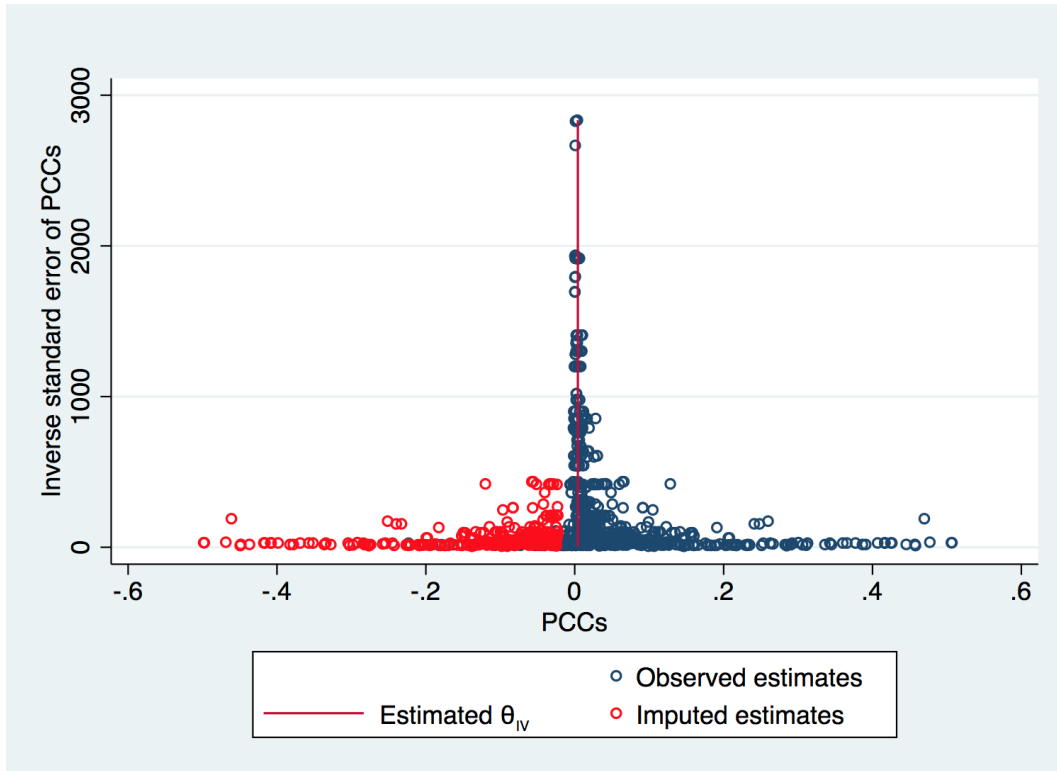


Figure B.5: Trim-and-fill-funnel plot

Note: The values of PCCs obtained are plotted against the inverse standard errors. "Imputed" denotes potential estimates which are missed because of publication bias.

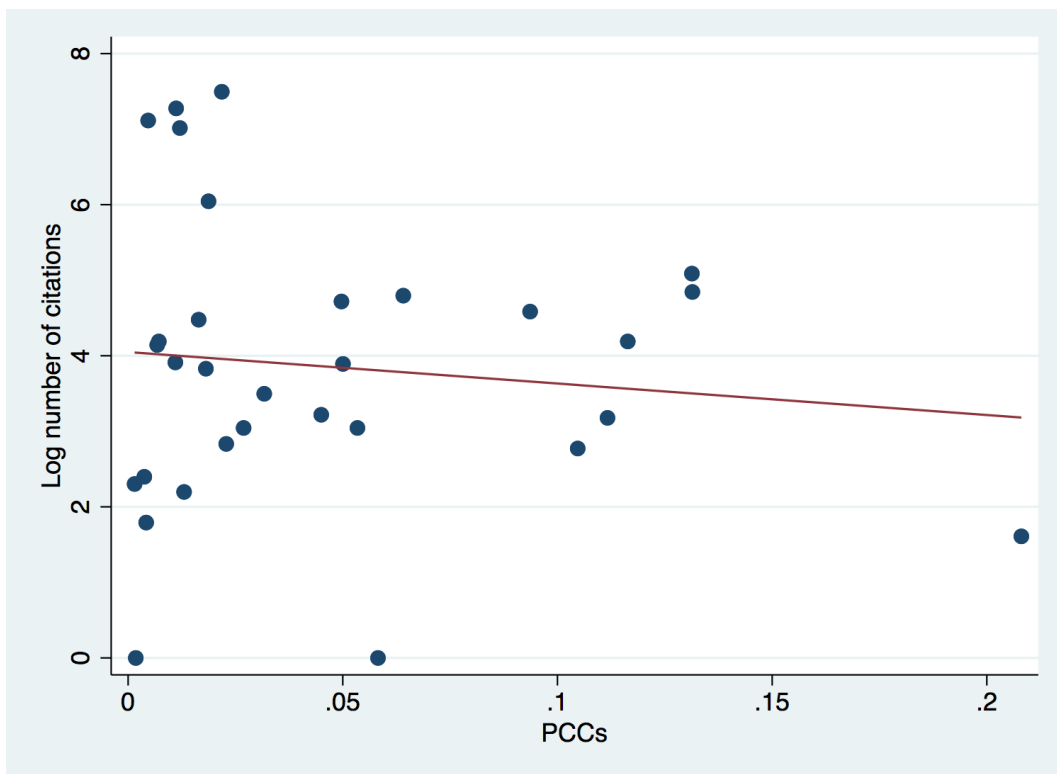


Figure B.6: The relationship between log number of citations and PCCs

Note: The log number of citations are plotted against PCCs