Alcamentos

Departamento de Economía

1401

ON THE SIZE OF SHEEPSKIN EFFECTS: A META-ANALYSIS

Jhon James Mora[1] y Juan Muro[2]

[1]Universidad ICESI y Alcamétrica
[2]Universidad de Alcalá y Alcamétrica

DEPARTAMENTO DE ECONOMÍA
Plaza de la Victoria, 2
28802 Alcalá de Henares (Madrid)
Teléfono: 91 885 42 01
http://www.uah.es/centros_departamentos/departamentos

Universidad de Alcalá

Alcamentos

Nº: 1401

ON THE SIZE OF SHEEPSKIN EFFECTS: A META-ANALYSIS

Jhon James Mora[1] y Juan Muro[2]

[1]Universidad ICESI y Alcamétrica

[2]Universidad de Alcalá y Alcamétrica

DEPARTAMENTO DE ECONOMÍA
Plaza de la Victoria, 2
28802 Alcalá de Henares (Madrid)
Teléfonos: 91 885 42 01
http://www.uah.es/centros_departamentos/departamentos

ON THE SIZE OF SHEEPSKIN EFFECTS: A META-ANALYSIS

Jhon James Mora[1] and Juan Muro[2]*

[1]Universidad ICESI and Alcamétrica
[2]Universidad de Alcalá and Alcamétrica

January, 2014

Abstract

We use information gathered from 122 studies on the effects of high school degrees on wages in different countries worldwide to carry out a meta-analysis that shows high school degrees have a statistically significant effect on wages of nearly 8%. This effect varies either when the review is made in countries away from the tropics or when factors such as sex, race, and continent are taken into account. Our results also reveal the existence of a publication bias that tends to increase the magnitude of the sheepskin effect. Nevertheless, when the former is included into the analysis the later remains statistically significant.

Keywords: Sheepskin effects, Meta-analysis, Publication bias.
JEL Classification: C80, I21, J24

***Corresponding author.**
Facultad de Ciencias Económicas y Empresariales
Universidad de Alcalá
Plaza de la Victoria, 2
28802 Alcalá de Henares, Madrid (SPAIN)
e-mail: juan.muro@uah.es

ON THE SIZE OF SHEEPSKIN EFFECTS: A META-ANALYSIS

Jhon James Mora and Juan Muro

1. Introduction

The degree equation was first developed by Hungerford and Solon in 1987 and is usually known as the "sheepskin effect equation". The equation, which is estimated from a regression of wages of individuals in a given country, is aimed at determining the effects of schooling degrees on wages in a specific setting. Using cross-section data, Hungerford and Solon (1987) found that there is an additional significant return in the years during which a diploma is earned. Since then many studies have been carried out to test and measure the sheepskin effect hypothesis. From our review most of this research has been completed in Brazil (29.51%), United States (24.59%) and Colombia (10.66%).

Although it is evident that there might be measurement errors in educational attainment when empirical research is based on data of self-reported educational levels, Card (1999), Kane *et. al*. (1999), and that OLS estimates overstate the effects of a schooling degree, we also must observe that if sheepskin effects persist across different countries, their importance should not be neglected. The existence of schooling degree effects is obviously important when it comes to establishing educational policies in any country because of the high social costs involved, particularly in developing countries.

One of the possible ways to determine the magnitude of sheepskin effects is by examining various publications and working papers on this subject. In this paper we conduct a meta-analysis of the schooling degree equation that centers specifically in the effect of high school degrees. We have reviewed a total of 122 articles and working papers published that covers 15 different countries, including, among others, Libya, The Philippines, and Egypt. Our findings show that the effect of a schooling degree is not only statistically significant but depends on factors such as closeness to the tropics, sex, race, and continent. The paper provides an important contribution in that it shows that the effect of a high school degree on wages is true in a statistical sense. This means in other words that said effect is not statistically equal to zero. Additionally, we find that the size of the sheepskin effect is around 8 percent in the case of high school degrees.

The paper is organized as follows: section 2 presents the schooling degree equation and the meta-analysis technique; section 3 discusses relevant data; section 4 reviews the results; and lastly, section 5 provides the conclusions.

2. Sheepskin effects and Meta-analysis

In general, additional earnings from holding a schooling degree can be estimated from the following wage regression

$$LnWh_i = \alpha_1 S_i + \alpha_2 exp_i + \alpha_3 exp^2{}_i + \sum_t^T \beta DS_{t,i} + \sum_t^T \beta S_{t,i} (S - S_t)_{t,i} + \mu_i \quad (1)$$

Where LnWh is the logarithm of hourly wages; S is the number of years of schooling; exp, exp^2 represent an individual's labor experience and its square; DS_t is a dummy variable for the year in which a given degree is earned; S_t is the year in which a degree is earned, and t is the schooling degree itself, which can be as advanced as a doctoral degree (PhD). In analyzing the effect of high school degrees, we observe that in some countries the diploma is obtained after eight years of schooling, while in others it can take either 10 or 11 years. The regression in (1) allows to estimate a β value for each schooling degree and its standard error, Hungerford and Solon (1987), Mora and Muro (2008).

The instrument of meta-analysis has been used in medical and psychological studies on a regular basis, Sterling (1959), Rosenthal (1979), Begg and Berlin (1988), Borenstein et al. (2009). It has also been utilized in economics by a number of authors. Among others, Card and Krueger (1995a, 1995b) to study the effects of minimum wages; Dalhuisen et al. (2003) to analyze income elasticity of water demand; Jarrell and Stanley (2004) to review wage discrimination; Abreu et al. (2005) to quantify beta-type convergence; and Colegrave and Giles (2008) to study school cost functions.

Let us assume there is an article or working paper that provides information about the size of the effect of a high school degree. Each publication also supplies information about the estimated standard error. Thus,

$$\begin{aligned} HS-Sheepskin_i \,|\theta_i &\sim N(\theta_i, \sigma_i^2) \quad ; \quad \theta_i \sim N(\theta, \tau^2) \\ HS-Sheepskin_i &\sim N(\theta\ , \sigma_i^2 + \tau^2) \\ HS-Sheepskin_i &= \theta + u_i + \varepsilon_i \ ; u_i \sim N(0, \tau^2); \quad \varepsilon_i \sim N(0, \sigma_i^2) \end{aligned} \tag{2}$$

In equation (2), HS-sheepskin is the estimated effect derived from equation (1). It is worth noting that θ_i is the effect of a high school degree, which varies from one study to another. It is assumed that it has a normal distribution around the mean effect θ. The between-studies variance, τ^2, is estimated from relevant data.

Between-studies variance τ^2 is determined using the method of moments, DerSimonian and Laird (1986), from the following equation

$$\tau^2 = \frac{\left[\sum_{i=1}^{k} W_i Sheepskin_i^2 - \frac{\left(\sum_{i=1}^{k} W_i Sheepskin_i \right)^2}{\sum_{i=1}^{k} W_i} \right] - (k-1)}{\sum W_i - \frac{\sum W_i^2}{\sum W_i}} \qquad (3)$$

Where W_i is the weight of each article or working paper and K is the number of articles and/or working papers.

Equation (2) shows that the effect of a schooling degree could be explained with both a fixed-effect model and a random effect model. However it does not provide any explanation as to the determinants of variability between studies. To take into account factors that determines the variability between studies a vector of covariates X_i is incorporated, as shown in equation (4) below.

$$
\begin{aligned}
HS - Sheepskin_i \,|\theta_i \quad &\sim N(\theta_i, \sigma_i^2) \quad ; \quad \theta_i \sim N(\beta' X_i, \tau^2) \\
HS - Sheepskin_i \quad &\sim N(\beta' X_i + \tau^2) \\
HS - Sheepskin_i \quad &= \beta' X_i + u_i + \varepsilon_i \; ; \quad u_i \sim N(0, \tau^2) \, ; \varepsilon_i \sim N(0, \sigma_i^2)
\end{aligned} \qquad (4)
$$

In (4) θ_i is again the estimated effect which varies from one study to another. It is assumed that it has a normal distribution around the linear predictor of θ. τ^2, on the other hand, is the between-studies variance which is estimated from relevant data and cannot be accounted for by covariates.

Estimating (3) or (4) provides an initial estimate of θ. Available literature on the topic of meta-analysis provides discussions of whether the aforementioned value could be biased due to the current publication policies of scientific journals. As an illustrative example, Card and Krueger (1995) and Stanley (2005) contend that there are at least three different sources of publication bias in economics.

> "1.-Reviewers and editors may be predisposed to accept papers consistent with the conventional view. 2.- Researchers may use the presence of a conventionally expected result as a model selection test. 3.- Everyone may possess a predisposition to treat `statistically significant` results more favorably" Stanley (2005, 310-211)

To tackle with this problem a test to identify the potential existence of the aforementioned publication bias has been proposed. The test is based in running the following regression

$$
effect_i = \beta_1 + \beta_0 Sd_i + e_i \qquad (5)
$$

Where effect variable is the effect of a schooling degree on wages, and Sd is its standard error. In the absence of publication bias, the true effect will have a random variance close to the value of β_1, regardless of the standard error. Equation in (5), however, is heteroscedastic. A heteroscedasticity-corrected regression, by making a standard-error adjustment, is

$$t_i = \beta_0 + \beta_1 (1 / Sd_i) + e_i \tag{6}$$

Egger et al. (1997) posit that a test of significance on β_0 is a test of publication bias that indicates the direction of the bias. Stanley (2008), on the other hand, argues that the observed effect comes close to θ when n tends to infinity and Sd tends to zero. Therefore, a test on β_1 is a test for a true effect of informality that goes beyond the systematic "contamination" that arises from publication biases. Hence, β_1 is the “true value of the effect of a schooling degree”.

3. Data

A search on JSTOR, SCOPUS, ISI-Web, EBSCO, and GOOGLE yielded a list of 122 articles and/or papers published between 1987 and 2011. Table 1 contains summary statistics of our sample.

[Insert Table 1]

On average, publications on the topic of sheepskin effects of a high school degree show an additional return on a schooling degree of 19.8% with a standard deviation. Brazil, where most studies have been carried out, is the country that evidences the greatest additional return on a schooling degree.

Canada and Sweden, on the contrary, are the countries with the lowest additional return on a schooling degree. A 44% of the studies consider gender differences (male vs. female), while 31% of the studies incorporate race differences (white vs. black, mestizo, and indigenous populations). Lastly, 71% of all studies were performed in countries on the American continent. When we compute effect/Sd the results show a minimum value of 0.01 and a maximum value of 33.75. With a 5% level of significance 24 (19%) values reject the sheepskin effect hypothesis and 32 (26%) values reject the hypothesis of the sheepskin effects with a 1% level of significance.

4. Results

We carry out a meta-analysis in order to examine whether the studies of the effect of high school degrees share an effect in common, in which case the fixed-effect method should be used, or otherwise, there is a remarkable study heterogeneity, in which case the random effect method should be employed.

[Insert Table 2]

Table 2 shows an estimated value of the schooling degree effect of 7.9% when the fixed effect method is used, while the estimated value is 14.5% with the use of random effects, and the between-studies variance is close to 0.03.

Although both estimates of the effect of a high school degree are statistically significant, various studies in different places around the world and the results of estimates for men and women or people of different races show that there is a large heterogeneity from one study to another.

Therefore, the random effect method should be used for the analysis. In order to explore the issue of heterogeneity, a Q test of heterogeneity, Borenstein et al. (2009) was carried out yielding a value of 1307.384. Under the null that the studies share an effect in common the test follows a chi-squared distribution with k-1 degrees of liberty. The rejection of the null hypothesis reinforces the hypothesis of heterogeneity and the convenience of using the random effect method.

Higgins et al. (2003) use an index that aims to identify to what extent the variance is spurious and to what extent it is real. Higgins et al. index, I^2, is a relative scale ranging from 0 to 100 which is not dependent upon the number of studies. If I^2 is close to zero, the observed variance is largely spurious, but if I^2 is close to 100, it makes sense to draw conjectures about the variance and about what could explain it. In other words, it is reasonable to carry out meta regressions or subset-based analyses. Hence, according to our results in Table 2, it would make sense to incorporate covariates into our analysis.

The set of covariates included in our model are the distance to the equator, a dummy variable for men (sex), a dummy variable for race (race), and a dummy variable for the American continent. The estimation results are listed in Table 3 below.

[Insert Table 3]

Table 3 shows that the effect of a high school degree decreases as the distance to the equator increases, is larger for men than for women and, when the race variable is included in the model, the effect of the degree is greater for white people than for black, indigenous, and other populations. With respect to the geographic variable, the studies conducted on the American continent reveal that a degree is recognized to a lesser extent than in other countries.

4.1. Publication biases and the true effect of a high school degree

Until now we have obtained estimates that as mentioned above are likely affected by publication bias. In order to test this hypothesis, we estimate equations (5) and (6) above. Our results are in Table 4.

[Insert Table 4]

Equation (5) estimates are shown in the first column of Table 5. They suggest an effect of a schooling degree around 100%, which would mean a schooling degree would increase wages by 100%. The bias direction is positive (constant), which would imply most studies tend to report a larger effect than actually observed. Due to the presence of heteroscedasticity in equation (5) we use in our analysis equation (6) estimates in the second column of Table 5. They show a much more moderate effect of a schooling degree on wages of 5.8%.

In order to discuss if the so far estimated effect is true, we run a regression between the t values of each study and the sample size of the study (n). As thoroughly shown by Stanley (2005), in the case there is indeed a true effect of schooling degrees, and given that $t = \beta/Sd$ when $\beta \neq 0$, in the regression $Ln(t) = \alpha_0 + \alpha_1 Ln(n)$ the value of α_1 will be statistically equal to ½. Our estimated value was 0.487 (third column in Table 4), and F for hypothesis $\alpha_1 = ½$ was 0.06. This means that the observed effect of schooling degrees is far from being different from zero, which shows that the effect is true.

Since the true effect variable is estimated with an error where 1/SE incorporates sampling errors, which are also estimated we must conclude there are measurement errors in equation (5), Sterne et al. (2000) and Macaskill et al. (2001). To solve this problem we estimate equation (6) using as instrument the inverse of the square root of the number of observations, Stanley (2005).

[Insert Table 5]

The results in table 5 show again that the bias direction is positive and that the "true" effect is close to 9%.[1] In column (2) we incorporate a dummy variable for ISI or Scopus journal to take into account differences arising from the quality of the publication. The result shows that if the paper was published in an ISI or Scopus journal the estimated sheepskin effect diminishes and is only 2 percent (0.07 minus 0.05).

Finally, we incorporate a variable to capture the likely obsolescence of the "sheepskin effect" paradigm and its impact on the size of estimated high school degree effect. To do that we construct a time to origin variable calculated as the time gap between the publication year of each study and the publication year of the seminal paper by Hungerford and Solon (Year of publication - 1987). The augmented regression with the time to origin variable gives an estimated reduction of the sheepskin effect of 0.002 by year.[2]

[1] A regression was also carried out with the 30, 60, and 90 percentiles of the distribution. The IV-quantile regression does not yield statistically different results between percentiles [F for the difference between percentiles 30 and 90 was 1.5 with a probability of 0.223, F for the difference between percentiles 30 and 90 was 0.04 with a probability of 0.83, and F for the difference between percentiles 60 and 90 was 0 with a probability of 0.979].

[2] The total effect over the 24 years since the first publication of the sheepskin equation is around -0.05 (-0.002*24).

5. Conclusions

There is no doubt that return is an important aspect of education. Following this track of thoughts, not only the amount of education (understood as the number of years of education received by a student) is important, but also the ability of education to signal productivity of individuals in the labor market, Spence (2002), Mora and Muro (2008).

One of the used instruments to estimate the ability schooling degrees have to signal is the degree equation. A review of the literature on the topic shows the relevance of the study of sheepskin effects worldwide. Concerning the size of the effect we find a high heterogeneity in the published results. We utilize a meta-analysis framework to offer a robust estimate of the size of high-school degree on wages. First of all our research undoubtedly shows that there is an additional and statistically significant wage increase for individuals who have earned a high school degree. The size of the effect, however, is not identical for all individuals but varies with their sex, race or continent they live in. In addition, interesting geographic differences can be appreciated when the published studies refer to countries away from the equator line.

Our results also corroborate the presence of publication biases and provide evidence on the fact most articles tend to overestimate the degree effect. Finally, we present a publication bias corrected meta-analysis regression that allows us to conclude a high school degree has an effect on wages that is around an 8 percent, size that has a substantial shrinkage when the article has a quality label, has been published in a journal with high impact – ISI or SCOPUS. In the later situation the size of the high school degree effect is only a 2 percent.

References

Abreu, M., H. deGroot and R. Florax (2005). A meta-analysis of beta-convergence: the legendary two-percent. *Journal of Economic Surveys,* 19(3), 389-420.

Begg, C. and J. Berlin (1988). Publication bias: a problem in interpreting medical data. *Journal of the Royal Statistical Society (Series A),* 151, 419-445.

Borenstein, M., L. Hedges, J. Higgins and H. Rothstein (2009). *Computing Effect sizes for Meta-Analysis.* Chichester: John Wiley & Sons, Ltd.

Card, D. and A. Krueger (1995a). Time-series minimum wages studies: a meta-analysis. *American Economic Review,* 85, 238-243.

Card, D. and A. Krueger (1995b). *Myth and Measurement: The New Economics of the Minimum Wage.* Princeton: Princeton University Press.

Colegrave, A.D. and M.J. Giles. (2008). School cost functions: A meta-regression analysis. *Economics of Education Review,* 27, 688-696.

Dalhuisen, J., R. Florax, H. deGroot and P. Nijkamp (2003). Price and income elasticity's of residential water demand: a meta-analysis. *Land Economics,* 79, 292-308.

DerSimonian, R. and N. Laird (1986). Meta-Analysis in clinical trials. *Controlled Clinical Trials,* 7, 177-188.

Egger, M., G. Davey Smith, M. Schneider and C. Minder (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ,* 315, 629-634.

Higgins, J., S. Thompson, J. Deeks and D. Altman (2003). Measuring Inconsistencies in Meta-Analyses. *BMJ,* 327, 557-560.

Hungerford, T. and G. Solon (1987). Sheepskin Effects in the Returns to Education, *The Review of Economics and Statistics,* 69(1), 175-177.

Jarrell, S. and T. Stanley (2004). Declining bias and gender wage discrimination? A meta-regression analysis. *Journal of Human Resources,* 38, 828-838.

Macaskill, P., S. Walter and L. Irwig (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine,* 20, 641-654.

Mora, J. J. and J. Muro (2008). Sheepskin effects by cohorts in Colombia. *International Journal of Manpower*, 29,111 – 121.

Rosenthal, R. (1979). The 'file drawer problem' and tolerance for null results. *Psychological Bulletin,* 86, 638-641.

Spence, M. (2002). Signaling in Retrospect and the Informational Structure of Markets. *American Economic Review,* 92(3), 434-59.

Stanley, T. (2005). Beyond Publication Bias. *Journal of Economic Surveys,* 19(3), 309-345.

Stanley, T. (2008). Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection. *Oxford Bulletin of Economics and Statistics,* 70(1), 103-127.

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance. *Journal of the American Statistical Association,* 54, 30-34.

Sterne, J., D. Gavaghan, and M. Egger (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*, 53, 1119-1129.

Table 1. Summary statistics of data (Sample of Sheepskin effect publications)

Variable	Percentage (%)	N	
Sex	0.44	122	
Race	0.32	122	
America	0.72	122	
By Country	Beta (High School)	Standard Deviation (High School)	N
Brazil	0.34	0.08	36
Canada	0.05	0.005	7
Colombia	0.12	0.02	13
Egypt	0.16	0.15	2
Spain	0.34	0.10	10
United States	0.09	0.05	30
The Philippines	0.13	0.03	2
Japan	0.20	0.06	2
Libya	0.16	0.08	1
Mexico	0.10	0.02	2
New Zealand	0.07	0.08	6
Pakistan	0.28	0.41	3
The Czech Republic	0.22	0.08	4
Czechoslovakia	0.19	0.10	2
Sweden	0.05	0.01	2
Total	0.20	0.07	122

Source: Authors' computation.

Table 2. Random and Fixed Meta-Analysis

		τ^2	95% Confidence Intervals		Z (value)	I^2 (Higgins et.al 2003)	No. of Studies
Method	β		Lower	Upper			
Fixed	0.08		0.08	0.08	54.9	90.7%	122
Random	0.15	0.003	0.13	0.16	22.3	90.7%	122

Source: Authors' computation.

Table 3. Meta-regressions

	Meta-Reg[1]	Meta-Reg[2]	Meta-Reg[3]	Meta-Reg[3]
Diff-Latitude	-0.289***	-0.217***	-0.141**	-0.174***
	(0.065)	(0.046)	(0.042)	(0.042)
Sex		0.149***	0.079***	0.066**
		(0.017)	(0.020)	(0.020)
Race			0.114***	0.132***
			(0.025)	(0.025)
America				-0.060**
				(0.020)
Constant	0.259***	0.170***	0.139***	0.199***
	(0.024)	(0.018)	(0.016)	(0.025)
τ^2	0.009	0.004	0.002	0.002
Q	953.6	780.9	640.2	634.7
I^2	0.874	0.848	0.816	0.816
R^2 adjusted	0.230	0.672	0.823	0.825
N	122	122	122	122

Source: Authors' computation. Note: * p<0.05, ** p<0.01, *** p<0.001.

Table 4. Publication Bias estimates

	Publ. Bias[1]	Publ. Bias[2]	Meta-significance
Sheepskin Effect	0.997*	0.058***	
	(0.392)	(0.014)	
Ln(n)			0.487***
			(0.055)
Constant	0.130***	2.228***	-3.330***
	(0.023)	(0.382)	(0.552)
Log-likelihood	67.21	-295.82	-147.68
Adj. R-square	0.195	0.538	0.399
Number of cases	122	122	122

Source: Authors' computation. Note: * p<0.05, ** p<0.01, *** p<0.001.

Table 5. Publication bias corrected estimates

	Bias corrected (1)	Bias corrected (2)
True Sheepskin Effect	0.088***	
	(0.017)	
True Sheepskin		0.079***
		(0.012)
(ISI or Scopus)/Sd		-0.051***
		(0.011)
(Year of publication - 1987)/Sd	-0.002*	
		(0.001)
Direction-Bias	1.126*	1.300***
	(0.449)	(0.337)
Adj. R-Square	0.397	0.551
Number of Cases	122	122

Source: Authors' computation. Note: * p<0.05, ** p<0.01, *** p<0.001.

Relación de títulos publicados en la colección ALCAMENTOS.

Nº	Autor/es	Título
0801	Juan Muro Cristina Suárez Maria del Mar Zamora	The impact of e-commerce on the tourist purchase decision: An empirical micro analysis
0802	Jhon James Mora Juan Muro	Diploma earning differences by gender in Colombia
0803	José Mª Arranz Ana I. Gil	Alcoholic beverages as determinants of traffic fatalities
0804	Juan Muro Cristina Suárez Maria del Mar Zamora	A note on computing Murphy-Topel corrected variances in a heckprobit model with endogeneity in Stata
0805	Elena López Rigoberto A. López	Demand for differentiated milk products: Implications for price competition
0806	Pablo Alonso González Irene Albarrán Lozano	Long term care in Spain: Extent, costs and challenges
0901	Rigoberto A. López Elena López Carmen Liron-Espana	When is concentration beneficial? Evidence from U.S. manufacturing
0902	Pablo Alonso González Irene Albarrán Lozano	Private long term care insurance: Theoretical approach and results applied to the Spanish case
0903	Cecilia Albert Verdú María A. Davia Rodríguez	Monetary poverty, education exclusion and material deprivation amongst youth in Spain
1001	Jhon James Mora Juan Muro	A cohort-based analysis of the influence of minimum wage levels on labour force participation in the informal sector: Quantitative and substitution effects
1002	Ana Karina Alfaro José Javier Núñez	The Spanish technical change: A regional and a dynamic analysis (1994-2007)
1101	Juan Muro	Dual Farrell measures of efficiency
1201	José Mª Arranz Carlos García-Serrano Virginia Hernanz	How do we pursue labormetrics? An application using the MCVL
1301	Juan Muro Joaquín Vera	Homotheticity, duality and efficiency measures