Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2015-31 | April 23, 2015 | http://www.economics-ejournal.org/economics/discussionpapers/2015-31

Please cite the corresponding Journal Article at
http://www.economics-ejournal.org/economics/journalarticles/2015-37

On the Size of Sheepskin Effects: A Meta-Analysis

Jhon James Mora Rodríguez and Juan Muro

Abstract

The authors use information gathered from 122 studies on the effects of high school diplomas on wages in different countries worldwide to carry out a meta-analysis that shows high school diplomas have a statistically significant effect on wages of nearly 8 percent. This effect varies whether the country is away from the tropics or whether factors such as sex, race, and continent are taken into account. The authors' results also reveal the existence of a publication bias that tends to increase the magnitude of the sheepskin effect. Nevertheless, when the former is factored into the analysis the latter remains statistically significant.

(Published in Special Issue Meta-Analysis in Theory and Practice)

JEL C8 I21 J24
Keywords Sheepskin effects; meta-analysis; publication bias

Authors
Jhon James Mora Rodríguez, ✉ Universidad Icesi, Colombia, jjmora@icesi.edu.co
Juan Muro, Universidad de Alcalá, Spain

Citation Jhon James Mora Rodríguez and Juan Muro (2015). On the Size of Sheepskin Effects: A Meta-Analysis. Economics Discussion Papers, No 2015-31, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2015-31

Received January 15, 2015 **Accepted as Economics Discussion Paper** March 6, 2015 **Published** April 23, 2015

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 individuals’ wages in a given country, is aimed at determining the effects of school diplomas and degrees on wages in a specific setting. Using cross-sectional data, Hungerford and Solon (1987) found that there is a return for each year of education and an additional significant return on the years during which a diploma or degree is earned. Since then many studies have been carried out to test the hypothesis and measure the sheepskin effect. For our review most of this research was completed in Brazil (29.51%), the United States (24.59%), and Colombia (10.66%).

While it is evident that there may be measurement errors in educational attainment when empirical research is based on data of self-reported education levels (see Card, 1999; Kane et al., 1999) and that ordinary least squares (OLS) estimates overstate the effects of a diploma/degree, we also must observe that if sheepskin effects persist across different countries, their importance should not be neglected. The existence of diploma/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 diploma/degree equation that centers specifically on the effect of high school diplomas. We have reviewed a total of 122 published articles and working papers that cover 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 diploma on wages is real in a statistical sense. In other words, the said effect is not statistically equal to zero. Additionally, we find that the size of the sheepskin effect is around 8% in the case of high school diplomas.

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

2. Sheepskin effects and meta-analysis

In general, additional earnings from the complete range from school diplomas and certificates to Ph.D. degrees can be estimated from the following wage regression:

$$\ln(Wh_i) = \alpha_1 S_i + \alpha_2 \exp_i + \alpha_3 \exp_i^2 + \sum_{t=1}^{T} \beta DS_{t,i} + \sum_{t=1}^{T} \beta S_{t,i} (S - S_t)_{t,i} + \mu_i , \tag{1}$$

where $\ln(Wh)$ is the logarithm of hourly wages; S is the number of years of schooling; exp and $\exp^2$ represent an individual's years of 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 credential itself, which can be as advanced as a doctoral degree (PhD). In this article we only consider the high school diploma, and not the number of years to obtain the diploma. For example in Colombia the diploma is obtained after 11 years of schooling while in the USA it takes 12 years and in the UK 13 years of the schooling. The regression in (1) allows us to estimate a β value for each schooling diploma and its standard error (Hungerford and Solon, 1987; Mora and Muro, 2008).

Meta-analysis has been used in medical and psychological studies on a regular basis (see, e.g., Sterling, 1959; Rosenthal, 1979; Begg and Berlin. 1988; Borenstein et al., 2009). It has also been utilized in economics by a number of authors, including 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 diploma. Each publication also supplies information about the estimated standard error. Thus,

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

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 diploma, 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 and is determined using the method of moments (DerSimonian and Laird, 1986), from the following equation:

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

where W_i is the weight of each article or working paper and n is the number of articles and working papers.

Equation (2) shows that the effect of a diploma 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 determine the variability between studies a vector of covariates X_i is incorporated, as shown in equation (4) below.

$$HS - Sheepskin_i \left| \theta_i \right. \sim N(\theta_i, \sigma_i^2) \quad ; \quad \theta_i \sim N(\beta' X_i, \tau^2)$$
$$HS - Sheepskin_i = \beta' X_i + u_i + \varepsilon_i \; ; \quad u_i \sim N(0, \tau^2) \, ; \varepsilon_i \sim N(0, \sigma_i^2) \qquad (4)$$

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, 310–11)

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

$$\text{effect}_i = \beta_1 + \beta_0 Sd_i + e_i \qquad (5)$$

where effect_i is the effect of a school diploma on wages and Sd_i is its standard error. In the absence of publication bias, the estimate of the true effect will have a value close to β_1, regardless of the standard error. The distribution in equation (5), however, is heteroscedastic. A heteroscedasticity-corrected regression is obtained by transforming the standard error:

$$t_i = \beta_0 + \beta_1 (1 / Sd_i) + e_i \qquad (6)$$

Egger et al. (1997) posit that a test of significance of 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 the number of observations tends to infinity and Sd tends to zero. Therefore, a test of β_1 is a test for a true effect of a school diploma that goes beyond the systematic "contamination" that arises from publication biases. Hence, β_1 is the "true" value of the effect of a school diploma.

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.

Table 1. Data of Studies

Variable	Percentage (%)	*n*	
Sex	44	122	
Race	32	122	
The Americas	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
Weighted Average or Total	0.20	0.07	122

Source: Authors' computation.

On average, publications on the topic of sheepskin effects of a high school diploma show an additional return on a schooling degree of 19.8% with a standard deviation of 0.07. Brazil, where most studies have been carried out, is the country that evidences the greatest additional return on a school diploma. Canada and Sweden, on the other hand, are the countries with the lowest additional return. 44% of the studies consider gender differences (male vs. female), while 31% of the studies incorporate race differences (white vs. black, mestizo, and/or indigenous populations). Lastly, 72% of all studies were performed in countries on the American continents. When we compute effect/*Sd* the results show a minimum value of 0.01 and a maximum value of 33.75. At the 5% level of significance, 24 studies (19%) reject the sheepskin effect hypothesis, while 32 studies (26%) reject the hypothesis of the sheepskin effect at the 1% level of significance.

4. Results

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

Table 2. Random and Fixed Meta-Analysis

Method	B	τ^2	95% Confidence Interval Lower	Upper	Z (value)	I^2	Number of Studies
Fixed	0.08		0.08	0.09	54.9	90.7%	122
Random	0.15	0.003	0.13	0.16	22.3	90.7%	122

Source: Authors' computation.

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

Although both estimates of the effect of a high school diploma are statistically significant, various studies in different places around the world and the 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 hypothesis that the studies share an effect in common the test follows a chi-squared distribution with $k - 1$ degrees of freedom. The rejection of the null hypothesis reinforces the appropriateness 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. Their index, I^2, is on a relative scale ranging from 0 to 100 that is independent of 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 factors that 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 Americas. The estimation results are listed in Table 3 below.

Table 3. Meta-Regressions

	Meta-Reg[1]	Meta-Reg[2]	Meta-Reg[3]	Meta-Reg[4]
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)
The Americas				−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 3 shows that the effect of a high school diploma 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, is greater for white people than for black, indigenous, and other populations. With respect to the geographic variable, the studies conducted on the American continents reveal that a diploma is recognized to a lesser extent than in other countries.

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

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

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^2	0.195	0.538	0.399
Number of Cases	122	122	122

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

Equation (5) estimates are shown in the first column of Table 4. They suggest an effect of a high school diploma around 100%, which would mean a high school diploma 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 high school diploma on wages of 5.8%.

In order to discuss whether the effect as so far estimated is true, we run a regression between the *t*-values of each study and the sample size of the study (*n*). As shown by Stanley (2005), if there is indeed a true effect of high school diplomas, 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 the hypothesis $\alpha_1 = ½$ was 0.06. This means that the observed effect of diplomas is statistically far from zero, which shows that the effect is true.

To solve the problem of measurement error in equation (5) (see Sterne et al., 2000; Macaskill et al., 2001) we estimate equation (6) with IV regression using as instrumental variable the inverse of the square root of the number of observations (Stanley, 2005).

Table 5. Publication Bias–Corrected Estimates

	Bias-corrected (1)	Bias-corrected (2)
Sheepskin Effect	0.088***	
	(0.017)	
Sheepskin Effect		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^2	0.397	0.551
Number of Cases	122	122

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

The results in Table 5 show again that the bias 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 to only 2% (0.07 − 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 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 (i.e., year of publication − 1987). In this case our results show a reduction of 0.2% per year of the sheepskin effects.[2]

---

[1] A regression was also carried out with the 30th, 60th, and 90th percentiles of the distribution. The IV-quantile regression does not yield statistically different results between percentiles [*F* for the difference between percentiles 30 and 60 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 −4.8% (= −0.2% × 24).

5. Conclusions

There is no doubt that return is an important aspect of education. Following this train of thought, 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 instruments used to estimate the capability of school diploma as a signal is the Sheepskin equation. A review of the literature on this topic shows the relevance of the study of sheepskin effects worldwide. Concerning the size of the effect we find a high heterogeneity in published results. We utilize a meta-analysis framework to offer a robust estimate of the effect of a high school diploma 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 diploma. The size of the effect, however, is not identical for all individuals but varies with their sex, race, or the continent they live in. In addition, interesting geographic differences can be appreciated when the published studies refer to countries' distances from the equator.

Our results also corroborate the presence of publication biases and provide evidence for the conclusion that most articles tend to overestimate the diploma effect. Finally, we present a publication bias–corrected meta-analysis regression that allows us to conclude that a high school diploma has an effect on wages of around 8%, with a substantial shrinkage when the article has a quality label (i.e., has been published in a journal with high impact – ISI or SCOPUS). In the latter case the size of the high school diploma effect is only 3%.

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.