

A New Dataset on Mincerian Returns

Francesco Caselli, Jacopo Ponticelli, Federico Rossi

November 2014

1 Introduction

What is the economic value of an additional year of schooling? How and why does it vary across countries? These questions are at the core of the field of labor economics, and have received enormous attention in the last few decades. The implications of the answers are obviously far reaching, from the design of educational policy to the evaluation of the importance of human capital as a source of differences in standards of living across countries.

The workhorse empirical model to estimate the returns to education is the human capital earning function introduced by Mincer (1974), where the logarithm of earnings is regressed on years of schooling and a quadratic function of years of experience. This specification has strong theoretical foundations, being the outcome of a standard Ben-Porath (1967) model of human capital accumulation, and, given its simplicity, has been shown to fit the data remarkably well¹.

In the last few decades, George Psacharopoulos and his co-authors have provided a great service to the profession by compiling extensive collections of estimates of the returns to education for a wide range of countries (Psacharopoulos, 1981, 1985, 1994; Psacharopoulos et al., 2004). These estimates have been extensively used to analyze cross country patterns and evaluate the contribution of human capital to economic growth.

¹See Card (1999), Heckman et al. (2003), Lemieux (2006) and Polachek (2008) for extensive reviews.

The latest available estimates in the aforementioned collections (Psacharopoulos et al., 2004) are, for most countries, relative to the 1980s. In the last twenty years, however, there has been a burgeoning of new studies estimating the returns to education in different countries, thanks to a wealth of new data and econometric techniques which have become available. In this paper we present a new collection of Mincerian coefficients estimated with data relative to more recent years. In particular, the dataset includes up to two estimates for each country, one for the 1989-1999 period and one for the 2000s; these estimates come from a large number of academic papers and technical reports (see the Online Appendix for a list of sources).

The paper is structured as follows. Section 2 describes the data collection process and the coverage of the dataset. Section 3 offers an overview of the main patterns emerging from the data, and Section 4 concludes.

2 Sources and Criteria

In the latest review, Psacharopoulos et al. (2004) emphasize the importance of a selective approach in selecting estimates of returns to education reasonably comparable across countries. In this section we describe the criteria we adopted for the inclusion of an estimate in our dataset.

Ideally, we would want to limit ourselves to estimates coming from nationally representative samples, specifications with exactly the same controls and variables perfectly comparable across countries. Since this would limit our collection to a handful of observations, some compromise is in order to be able to perform meaningful cross country comparisons.

The estimates included in our dataset come from a large number of academic papers and technical reports (see the Online Appendix for a list of sources). Most of these studies are published in peer reviewed journals; however, to broaden the coverage we included also unpublished works as long as they met adequate standards in terms of sample size, data quality and econometric implementation.

In order to ensure comparability, when selecting the estimates we tried to stick as close as possible to the standard Mincerian specification, which includes years of schooling, experience and experience squared as controls. Many papers we have surveyed estimate richer models, controlling for other individual characteristics; luckily for our purposes, results from the baseline specification are often included as well. A particularly common (and misguided) practice is the inclusion of occupational or sectorial dummies: given the occupation is itself an outcome influenced by education, the regression does not have a causal interpretation². We therefore do not include estimates affected by this problem.

Another obstacle for a direct comparison across studies is that exact definition of the dependent variable depends on the context. Whenever possible, we give preference to measures of hourly gross earnings, which are not directly affected by differences in labor supply (part time versus full time workers) across individuals and in taxation across countries.

As noted by Psacharopoulos et al. (2004), estimates coming from samples of workers employed in the public sector pose additional problems, since their wages are likely not to reflect the market ones. We therefore limit ourselves to studies relative to the private sector.

Finally, as an alternative to the log-linear specification, many papers in the literature estimate models where the returns to schooling are allowed to vary depending on the stage of education. In particular, a common specification consists in regressing the logarithm of earnings on dummies corresponding to the highest level of completed schooling (primary, secondary and higher). on top of the usual experience controls. As shown in Caselli (2016), under some assumptions we can establish a one-to-one mapping between these coefficients and the Mincerian return corresponding to the classic log-linear specification. We therefore follow this method to compute the implied returns and include them whenever an alternative estimate coming from a log-linear specification is not available.

²See Angrist et al. (2009) for a detailed discussion of the "bad control" problem.

This leaves us with a total of 87 observations for the 1990s and 91 for the 2000s, which are displayed in the Appendix (rounded to the first decimal place). Many of the countries included in this collection were not part of the ones previously available, allowing us to provide a more complete picture on the international patterns.

3 The Main Patterns

The average returns to education by region are shown in Table 1. Overall, the average for all observations included in the dataset is 8.70% for the 1990s and 8.22% for the 2000s; these are approximately 1 percentage point lower compared to Psacharopoulos et al. (2004). For what concerns regional differences, countries in Latin America and the Caribbean stand out for having the highest returns in 1995, on average just below 11%, while countries in South-East Asia and the Pacific have the highest returns in 2005; countries in the Advanced Economies group (as classified by the World Bank) have instead returns below the world average.

Region	Year	
	1995	2005
Advanced Economies	7.79	7.36
Europe and Central Asia	7.37	7.03
Latin America and the Caribbean	10.85	8.17
Africa and Middle East	8.41	8.47
South - East Asia and the Pacific	8.62	10.58
World	8.70	8.22

Table 1: Regional Averages of Mincerian Coefficients

We now move to consider the correlation between estimated returns and the level of economic development. On a theoretical ground, the relationship is ambiguous: on one hand richer countries are endowed with a larger share of educated workers, and if skilled labor is subject to decreasing returns we should expect lower returns there; on the other hand, the availability of more educated workers could encourage firms to adopt more

skill-intensive technologies, widening the productivity gap between skilled and unskilled labor³. According to the estimates we have collected, there does not appear to be a systematic relationship between returns to schooling and real GDP per capita, either in the 1990s or in the 2000s (Figure 1). Even excluding the two outliers (Jamaica for the 1990s and Malta for the 2000s), the correlation remains slightly negative and not significantly different from zero at standard confidence levels. Similar conclusions hold with respect to the relationship with average years of schooling^{4 5} (Figure 2).

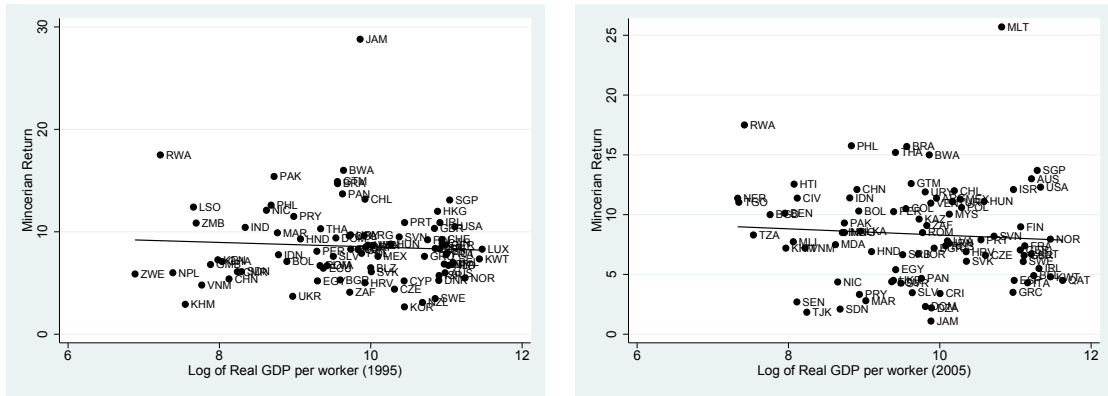


Figure 1: Mincerian Coefficients and Real GDP across countries

Table 2 shows the average returns by gender. In both decades women experience substantially higher returns than men; this is consistent with the pattern documented in previous collections. In recent work, Pitt et al. (2012) document that this gap in returns to schooling can not simply be ascribed to differences in the quantity of education across genders, since in most countries women have higher educational attainment than men. They instead propose an explanation based on comparative advantage due to biological differences in the endowment of skill and brawn.

³Moreover, countries' demographic structures and TFP levels might affect the estimated returns, leading to cross country differences; see Seshadri et al. (2014) for a version of this argument

⁴An exception is represented by the downward relationship between returns to education and average years of schooling in 1995. Excluding the outlier Jamaica, a regression of Mincerian coefficients on years of schooling (and a constant) yields an estimated slope of -0.34, significant at the 5% confidence level.

⁵Using an extended version of the dataset constructed by Psacharopoulos et al. (2004), Banerjee et al. (2005) find a small but significant negative relationship between Mincerian returns and both GDP per capita and average years of schooling.

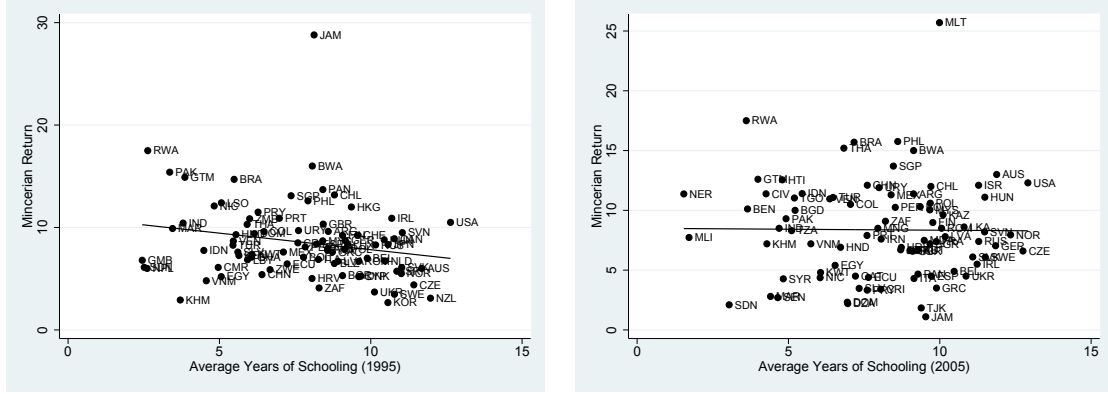


Figure 2: Mincerian Coefficients and Years of Schooling across countries

Gender	Year	
	1995	2005
Men	9.12	7.82
Women	10.01	9.52

Table 2: Average Mincerian Coefficients across Genders

References

- Angrist, J. D. and J. S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- Banerjee, A. V. and E. Duflo (2005): "Growth Theory through the Lens of Development Economics," in Aghion, P. and S. Durlauf, *Handbook of Economic Growth*, Vol. 1, Elsevier, chapter 7, pp. 473-552.
- Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75, 352.
- Card, D. (1999): "The causal effect of education on earnings," in Ashenfelter, O. and D. Card, *Handbook of Labor Economics*, Vol. 3, Elsevier, chapter 30, pp. 1801-1863.
- Caselli, F. (2016): *Technology Differences over Space and Time*, Princeton University Press.
- Heckman, J. J., L. J. Lochner and P. E. Todd (2003): "Fifty Years of Mincer Earnings Regressions," NBER Working Papers 9732, National Bureau of Economic

Research, May 2003.

Lemieux, T. (2006): "The "Mincer Equation" Thirty Years After *Schooling, Experience, and Earnings*," in S. Grossbard, *Jacob Mincer: A Pioneer of Modern Labor Economics*, Springer US, pp. 127-145.

Mincer, J. A. (1974): *Schooling, Experience, and Earnings*, Columbia University Press, New York.

Pitt, M., M. Rosenzweig and M. N. Hassan (2012): "Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy," *American Economic Review*, 102 (7), 3531-60.

Polachek, S. W. (2008): "Earnings Over the Life Cycle: The Mincer Earnings Function and Its Applications," *Foundations and Trends(R) in Microeconomics*, 4 (3), 165-272.

Psacharopoulos, G. (1981): "Returns to Education: An Updated International Comparison," *Comparative Education*, 17 (3), pp. 321-341.

Psacharopoulos, G. (1985): "Returns to Education: A Further International Update and Implications," *Journal of Human Resources*, 20 (4), 583-604.

Psacharopoulos, G. (1994): "Returns to investment in education: A global update," *World Development*, 22 (9), 1325-1343.

Psacharopoulos, G. and H. A. Patrinos (2004): "Returns to investment in education: a further update," *Education Economics*, 12 (2), 111-134.

Seshadri, A. and R. Manuelli (2014): "Human Capital and the Wealth of Nations," *American Economic Review*, 104(9), 2736-62.