Labour Economics 16 (2009) 97–106

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/econbase

The return to firm investments in human capital

Rita Almeida [a,*], Pedro Carneiro [b]

[a] The World Bank, 1818 H Street, NW MC 3-348, Washington, DC, 20433, USA
[b] University College London, Institute for Fiscal Studies and Center for Microdata Methods and Practice, United Kingdom

ARTICLE INFO

Article history:
Received 9 March 2007
Received in revised form 13 June 2008
Accepted 21 June 2008
Available online 2 July 2008

JEL classification codes:
C23
D24
J31

Keywords:
On-the-job training
Panel data
Production function
Rate of return

ABSTRACT

In this paper we estimate the rate of return to firm investments in human capital in the form of formal job training. We use a panel of large firms with detailed information on the duration of training, the direct costs of training, and several firm characteristics. Our estimates of the return to training are substantial (8.6%) for those providing training. Results suggest that formal job training is a good investment for these firms possibly yielding comparable returns to either investments in physical capital or investments in schooling.

1. Introduction

Individuals invest in human capital over the whole life-cycle, and more than one half of life-time human capital is accumulated through post-school investments on the firm (Heckman et al., 1998). This happens either through learning by doing or through formal on-the-job training. In a modern economy, a firm cannot afford to neglect investments in the human capital of its workers. In spite of its importance, economists know surprisingly less about the incentives and returns to firms of investing in training compared with what they know about the individual's returns of investing in schooling[1] Similarly, the study of firm investments in physical capital is much more developed than the study of firm investments in human capital, even though the latter may be at least as important as the former in modern economies. In this paper we estimate the internal rate of return of firm investments in human capital. We use a census of large manufacturing firms in Portugal, observed between 1995 and 1999, with detailed information on investments in training, its costs, and several firm characteristics.[2]

Most of the empirical work to date has focused on the return to training for workers using data on wages (e.g., Bartel, 1995; Arulampalam et al., 1997; Mincer, 1989; Frazis and Lowenstein, 2005). Even though this exercise is very useful, it has important drawbacks (e.g., Pischke, 2005). For example, with imperfect labor markets wages do not fully reflect the marginal product of labor, and therefore the wage return to training tells us little about the effect of training on productivity. Moreover, the effect of training on wages depends on whether training is firm specific or general (e.g., Becker, 1962; Leuven, 2005).[3] More importantly, the literature estimating the effects of training on productivity has little or no mention of the costs of training (e.g. Bartel, 1991, 1994, 2000; Black and Lynch, 1998; Barrett and O'Connell, 2001; Dearden et al., 2006; Ballot et al., 2001; Conti, 2005). This happens most probably due to lack of adequate data. As a result, and as emphasized by Mincer (1989) and Machin and Vignoles (2001), we cannot interpret the estimates in these papers as well defined rates of return.

The data we use is unusually rich for this exercise since it contains information on the duration of training, direct costs of training to the firm as well as productivity data. This allow us to estimate both a production and a cost function and to obtain estimates of the marginal benefits and costs of training to the firm. In order to estimate the total marginal costs of training, we need information on the direct cost of training and on the foregone productivity cost of training. The first is observed in our data while the second is the marginal product of

* Corresponding author. 1818 H Street, NW MC 3-348, Washington, DC, 20433, USA.
E-mail address: ralmeida@worldbank.org (R. Almeida).

[1] An important part of the lifelong learning strategies are the public training programs. There is much more evidence about the effectiveness (or lack of it) of such programs compared with the available evidence on the effectiveness of the private on-the-job training.

[2] We will consider only formal training programs and abstract from the fact that formal and informal training could be very correlated. This is a weakness of most of the literature, since informal training is very hard to measure.

[3] For example, Leuven and Oosterbek (2004, 2005) argue that they may be finding low or no effects of training because they are using individual wages as opposed to firm productivity.

0927-5371/$ – see front matter
doi:10.1016/j.labeco.2008.06.002

worker's time while training, which can be estimated. We do not distinguish whether the costs and benefits of training accrue mainly to workers or to the firm. Instead, we quantify the internal rate of return to training jointly for firms and workers.[4] This implies that, to obtain estimates of the foregone opportunity cost of training we will not take into account whether firms or workers support the costs of training.

The major challenge in this exercise are possible omitted variables and the endogenous choice of inputs in the production and cost functions. Given the panel structure of our data, we address these issues using the estimation methods proposed in Blundell and Bond (2000). In particular, we estimate the cost and production functions using a first difference instrumental variable approach, implemented with a system-GMM estimator. By computing first differences we control for firm unobservable and time invariant characteristics. By using lagged values of inputs to instrument current differences in inputs (together with lagged differences in inputs to instrument current levels) we account for any correlation between input choices and transitory productivity or cost shocks. Our instruments are valid as long as input decisions in period $t-1$ are made without knowledge of the transitory shocks in the production and cost functions from period $t+1$ onwards.[5]

Several interesting facts emerge from our empirical analysis. First, in line with the previous literature (e.g., Pischke, 2005; Bassanini et al., 2005; Frazis and Lowenstein, 2005; Ballot et al., 2001; Conti, 2005) our estimates of the effects of training on productivity are high: an increase in training per employee of 10 h (hours) per year, leads to an increase in current productivity of 0.6%. Increases in future productivity are dampened by the rate of depreciation of human capital but are still substantial. This estimate is below other estimates of the benefits of training in the literature (e.g., Dearden et al., 2006; Blundell et al., 1996). If the marginal productivity of labor were constant (linear technology), an increase in the amount of training per employee by 10 h would translate into foregone productivity costs of at most 0.5% of output (assuming all training occurred during working hours).[6] Given this wedge between the benefits and the foregone output costs of training, ignoring the direct costs of training is likely to yield a rate of return to training that is absurdly high (unless the marginal product of labor function is convex, so that the marginal product exceeds the average product of labor).

Second, we estimate that, on average, foregone productivity accounts for less than 25% of the total costs of training. This finding shows that the simple returns to schooling intuition is inadequate for studying the returns to training, since it assumes negligible direct costs of human capital accumulation. In particular, the coefficient on training in a production function (or in a wage equation) is unlikely to be a good estimate of the return to training. Moreover, without information on direct costs of training, estimates of the return to training will be too high since direct costs account for the majority of training costs (see also the calculations in Frazis and Lowenstein, 2005).

Our estimates indicate that, while investments in human capital have on average zero returns for training for all the firms in the sample, the returns for firms providing training are quite high (8.6%). Such high returns suggest that on-the-job training is a good investment for firms that choose to undergo this investment, possibly yielding comparable returns to either investments in physical capital or investments in schooling.[7]

The paper proceeds as follows. Section 2 describes the data we use. In Section 3, we present our basic framework for estimating the production function and the cost function. In Section 4 we present our empirical estimates of the costs and benefits of training and compute the marginal internal rate of return for investments in training. Section 5 concludes.

2. Data

We use the census of large firms (more than 100 employees) operating in Portugal (Balanco Social). The information is collected with a mandatory annual survey conducted by the Portuguese Ministry of Employment. The data has information on hours of training provided by the employers and on the direct training costs at the firm level. Other variables available at the firm level include the firm's location, ISIC 5-digit sector of activity, value added, number of workers and a measure of the capital, given by the book value of capital depreciation, average age of the workforce and share of males in the workforce. It also collects several measures of the firm's employment practices such as the number of hires and fires within a year (which will be important to determine average worker turnover within the firm). We use information for manufacturing firms between 1995 and 1999. This gives us a panel of 1,500 firms (corresponding to 5,501 firm–year observations). On average, 53% of the firms in the sample provide some training. All the variables used in the analysis are defined in the Appendix A.

Relative to other datasets that are used in the literature, the one we use has several advantages for computing the internal rates of return of investments in training. First, information is reported by the employer. This may be better than having employee reported information about past training if the employee recalls less and more imprecisely the information about on-the-job training. Second, training is reported for all employees in the firm, not just new hires. Third, the survey is mandatory for firms with more than 100 employees (34% of the total workforce in 1995). This is an advantage since a lot of the empirical work in the literature uses small sample sizes and the response rates on employer surveys tend to be low.[8] Fourth, it collects longitudinal information for training hours, firm productivity and direct training costs at the firm level. Approximately 75% of the firms are observed for 3 or more years and more than 60% of

[4] Dearden et al. (2006) and Conti (2005) estimate the differential effect of training on productivity and wages. The former find that training increases productivity by twice as much as it increase wages, while the latter finds only effects of training on productivity (none on wages).

[5] This assumption is valid as long as there does not exist strong serial correlation in the transitory shocks in the data, and firms cannot forecast future shocks. Given the relatively short length of our panel our ability to test this assumption is limited. Dearden et al. (2006) apply an identical methodology (using industry level data for the UK) for a longer panel and cannot reject that second order serial correlation in the first differences of productivity shocks is equal to zero. In their original application, Blundell and Bond (2000) also do not find evidence of second order serial correlation using firm level data for the UK.

[6] For an individual working 2,000 h a year, 10 hours corresponds to 0.5% of annual working hours.

[7] As a consequence, it is puzzling why firms that choose to undergo this investment in training, train on average such a small proportion of the total hours of work (less than 1%). We conjecture that this could happen for different reasons but unfortunately we cannot verify empirically the importance of each of these hypotheses. First, it may be the result of a coordination problem (Pischke, 2005). Given that the benefits of training need to be shared between firms and workers, each party individually only sees part of the total benefit of training. This may be also due to the so called "poaching externality" (Stevens, 1994). See also Acemoglu and Pischke (1998, 1999) for an analysis of the consequences of imperfect labor markets for firm provision of general training. Unless investment decisions are coordinated and decided jointly, inefficient levels of investment may arise. Second, firms can be constrained (e.g., credit constrained) and decide a suboptimal investment. Third, uncertainty in the returns of this investment may lead firms to invest small amounts even though the *ex post* average return is high, although what really matters for determining the risk premium is not uncertainty per se, but its correlation with the rest of the market. However, it is unlikely that uncertainty alone can justify such high rates of return. In our model uncertainty only comes from future productivity shocks, since current costs and productivity shocks are assumed to be known at the time of the training decision. The R-Squared of our production functions (after accounting for firm fixed effects) is about 85%, suggesting that temporary productivity shocks explain 15% of the variation in output. Since productivity shocks are correlated over time this is an overestimate for the uncertainty faced by firms.

[8] Bartel (1991) uses a survey conducted by the Columbia Business School with a 6% response rate. Black and Lynch (1997) use data on the Educational Quality of the Workforce National Employers survey, which is a telephone conducted survey with a 64% "complete" response rate. Barrett and O'Connell (2001) expand an EU survey and obtain a 33% response rate. Ballot, Fakhfakh and Taymaz (2001) use information for 90 firms in France between 1981 and 1993 and 250 firms in Sweden between 1987 and 1993. One exception is Conti (2005). She uses a large panel of Italian firms between 1996 and 1999 but the analysis is done at the more aggregated industry level.