

# On-the-job-training as a signal\*

*Why low skilled employees invest less in further training*

Olga Meshcheriakova,<sup>\*\*</sup> Stan Vermeulen,<sup>\*\*\*</sup>

## Abstract

Studies of on-the-job training have found that low skilled workers participate less in further training. In this paper, we develop a signalling model of training where training can increase productivity when workers' prior ability matches the level of the training course. Consequently, employers can use observed training participation as information about the quality of their employees when job performance is unobservable. As a result, programs aimed at low ability workers will be underutilized because of the negative signal participation conveys about their initial ability. We show that offering training has spillover effects: introducing high-level courses to a choice set of only low-level courses increases participation in low-level courses, and improves sorting efficiency. We discuss the implications of these results for optimal training provision, and make several testable predictions for assessing the validity of our model.

Keywords: Training, Human Capital, Signaling

JEL: J24, J45, M53

---

\* The authors thank Timothy Bond, Lex Borghans, Tamás Dávid-Barrett, Mirko Draca, Bart Golsteyn, Susanna Loeb, Arjan Non, Trudie Schils, and seminar participants at Maastricht University, Universität Tübingen, Kiel Institute for the World Economy, as well as participants at EALE 2016, and the Oslo Workshop on Education, Skills, and Labor Market Outcomes 2017 for their valuable comments.

\*\* Department of Economics, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, the Netherlands. Phone nr: +31 43 388 3841. [o.meshcheriakova@maastrichtuniversity.nl](mailto:o.meshcheriakova@maastrichtuniversity.nl)

\*\*\* Department of Economics, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, the Netherlands. [c.vermeulen@maastrichtuniversity.nl](mailto:c.vermeulen@maastrichtuniversity.nl)

## 1. Introduction

With the increased importance of a well-trained workforce due to skill-biased technological change (Acemoglu, 2002; Acemoglu and Autor, 2011; Autor, Levy, and Murnane, 2003), it has become essential for low ability workers to upgrade their skills. However, empirical studies on training participation consistently show that low skilled workers participate less in further training than their high skilled colleagues (Carneiro and Heckman, 2003; Heckman, LaLonde, and Smith, 1999; Bassanini et al., 2007). A recent report from the OECD and the European Commission based on PIAAC data shows that people with high skills are three to five times more likely to participate in adult learning activities than low skilled individuals (OECD and European Commission, 2013). Currently, research on on-the-job training explains low training participation of low skilled workers through individual differences in preferences and personality (Borghans, Duckworth, Heckman, and Ter Weel, 2008; Colquitt, LePine, and Noe, 2000; Fouarge, Schils, and de Grip, 2013). Since preferences and personality are relatively invariable (Meier and Sprenger, 2015; Borghans et al., 2008), this explanation implies that the potential for effectively stimulating the low skilled is limited.

In this paper, we develop a signalling model of on-the-job training as an alternative explanation for why low skilled workers participate less in training. We argue that training can increase workers' productivity when workers' prior ability matches the level of the training course. Consequently, employers can use observed training participation as information about the quality of their employees when job performance is (partially) unobservable. Employees will therefore incorporate this signalling value of signing up for courses of different levels into their participation decision. As a result, low ability employees do not sign up for training that is effective at improving their skills because of the negative signal training participation conveys about their initial ability.

This idea offers potential for increasing the training participation of low skilled workers through interventions, in contrast with the preference based explanation. We show that the availability of different types of training attenuates the signal of each particular training program. The intuition behind this result is that in a situation where employees can choose between a basic, low-level, course and not participating in any training, introducing an advanced training program stimulates employees at the high end of the ability distribution to sign up for advanced training. This decreases the average

ability of those who do not take any training, making the option of not taking training less attractive. Basic training then becomes relatively more attractive for those at the low end of the skills distribution. Therefore, offering training has spillover effects: introducing advanced courses increases participation in basic courses. This implies that low skilled workers can be effectively targeted once employers are aware of the signalling function of training.

Our model leads to several empirically testable predictions. First, employees for whom the signal is more important will participate less in low level training, and more in high level training, than those for whom the signal is less important. This implies, for example, that low skilled employees in industries where performance is harder to observe will participate less in training that is geared towards low skilled workers than in industries where performance is more easily observable. It also implies that employees that are on a fixed term contract are predicted to sort into advanced training more than their colleagues on permanent contracts. Second, in companies which offer a higher variety of training programs, participation rates among employees will be higher. Third, employers that offer and stimulate more advanced training opportunities will have a higher average level of productivity, independent of the effectiveness of the advanced training programs.

Our study contributes to the literature on the determinants of on-the-job training and to the literature on signalling. While there have been many papers on on-the-job training from a human capital or screening perspective (e.g. Becker, 1962; Hashimoto, 1981; Acemoglu, 1997; Acemoglu and Pischke, 1998, 1999a, 1999b; Autor, 2001; Leuven, 2005), few studies acknowledge the potential signalling function of on-the-job-training from the worker's perspective. In doing so, we provide a novel explanation for the low participation rate of low skilled employees. This insight can be used by policy makers as well as companies in the private sector to more effectively tailor training to their low skilled employees.

The remainder of this paper is organized as follows. Section 2 gives an overview of the relevant literature and motivation for our model. Section 3 introduces the model. Section 4 discusses the empirical predictions and Section 5 concludes.

## **2. Background Literature and Motivation**

In this section we discuss the theories that have been put forward explaining why low skilled employees sign up for on-the-job training less, and evaluate their empirical support. Firstly, we look at whether employers are less willing to invest in training for their low skilled workers. Secondly, we consider the explanation that the returns to training in terms of wages or future career outcomes are lower for low skilled workers. Thirdly, we discuss the explanation that low skilled workers participate less because of differences in preferences, personality traits, and motivation that are related to the decision to participate in training.

One of the reasons that low skilled workers receive less on-the-job training could be that employers are unwilling to offer them training. Hashimoto (1981), following Becker (1962), developed a model in which employers and employees share the costs and gains of training, and training only occurs when the predicted gains outweigh the costs for both parties. In practice, it could be that the predicted gains from training for the employer are too low when training low skilled employees, and this may result in less training opportunities for low skilled workers. However, most empirical studies investigating the supply side of training do not find a difference in the willingness of employers to provide training for workers of different levels of skill. Using Dutch data, Maximiano (2012) finds that different training rates between workers of different education levels reflect differences in the workers' demand for training, rather than differences in employers' willingness to supply training. Leuven and Oosterbeek (2000) also conclude that firms do not have strong preferences with regard to education levels when supplying training in The Netherlands, Canada, United States and Switzerland using data from the IALS. For Italy, Croce and Tancioni (2007) also find the educational level unrelated to willingness of employers to provide training. All in all, the available empirical evidence does not favour employers' unwillingness to provide training opportunities as an explanation for the low training incidence of lower skilled workers.

Another explanation put forward is that low skilled employees train less because the expected returns for this group are lower than for high skilled employees. In this case low skilled employees will demand less training even though employers would be happy to supply it. The empirical evidence for this explanation is mixed. In an early study, Lynch (1992) finds no difference in the returns to

training in terms of wages for workers of different levels of education in the United States. Fouarge et al. (2013) reach the same conclusion using Dutch data. Other studies do find that highly educated workers benefit more from taking training in terms of wage. Arulampalam and Booth (2001) find that the returns to training in Britain are largest for highly educated young men. For Germany, Kuckulenz and Zwick (2004) show that high skilled workers benefit more from training than low skilled workers. However there are also studies that find the opposite relationship. Budría and Pereira (2007) show that in Portugal returns to training are larger for low educated workers. Similarly, the OECD (OECD, 1999) and Brunello (2004) both report higher returns for low educated workers using international data from several European countries. In sum, the explanation that low skilled workers sign up less for training because of lower returns is not convincingly supported by the empirical literature.

The last explanation put forth in the literature is that low skilled workers differ from high skilled workers in terms of personality, motivation, and economic preferences. These factors in turn are related to people's willingness to engage in further training. The empirical evidence seems to provide support for this explanation. Fouarge et al. (2013) find that low educated workers are less future oriented, have a higher preference for leisure, have a more external locus of control, and report a higher level of exam anxiety. They show that these traits predict a lower willingness to participate in training. These findings are in line with Colquitt et al.'s (2000) meta-analysis of the psychological literature on training motivation. They report that among others, locus of control, anxiety, and self-efficacy are strongly related to training motivation. In turn, these traits have been shown to vary systematically with cognitive ability. Coleman and DeLeire (2003) show that internal locus of control is related to high cognitive ability. A high level of self-efficacy also correlates with high cognitive skills (Paunonen and Hong, 2010), while high levels of anxiety are related to lower cognitive performance (Ackerman and Heggstad, 1997). Thus, the explanation that high skilled and low skilled workers differ on traits related to willingness to participate in training is generally supported.

The preference based explanation for differences in the incidence of training leaves little room for interventions aimed at improving low skilled employees. This is because personality and economic preferences have been found to be relatively stable over time (Meier and Sprenger, 2015; Borghans et al., 2008), and any meaningful change that does occur is usually related to significant life

events or internal maturation (Specht, Egloff, and Schmukle, 2011). It is unlikely that any intervention can have a lasting impact on workers' underlying preferences. Indeed, the empirical evidence on the effects of policies intended to stimulate workers to take up training shows that low skilled workers are the least likely to make use of vouchers that significantly reduce the costs of training (Schwerdt, Messer, Woessmann, and Wolter, 2012). Even more discouragingly, Hidalgo, Oosterbeek, and Webbink (2014) find that those low skilled workers that do make use of vouchers are typically the type of low skilled worker that would have participated without the voucher, leading to large deadweight losses. This evidence suggests that cost-based interventions are not optimally designed to induce non-participating low-skilled workers to sign up for training.

While preferences likely are part of the story of workers' willingness to train, it is unclear how many of the current low skilled non-participants of training are never takers. We argue that part of the reason why cost-based interventions may be ineffective at increasing the participation rate of low skilled employees, is that there are additional signalling costs associated with training tailored towards low skilled workers. Participating in training that is effective at increasing productivity of low skilled workers, identifies those workers as having low initial skills. In jobs where productivity is (partly) unobservable, workers may be better off not taking this type of training even when it is effective and free of costs. The negative effect of divulging their initial skill level to the employer may outweigh the gains in productivity from training. In the next section, we formalize our argument by developing a signalling based model on workers' decisions to enter on-the-job training programs.

### **3. On the job training model**

In this section, we describe our model. First, we introduce a single program model separately for basic and advanced training, and state our assumptions. Later, we describe a situation where employees can choose between different types of available training programs.

The framework we adopt is similar to the standard signalling model (Spence, 1973; Stiglitz, 1975), where employers have imperfect information about their workers' productivity, and workers can sign up for education (in this case on-the-job training) as a signal of their ability. Employees are paid according to their expected productivity, conditional on their training decision. Our model differs

from the standard signalling setup in the sense that on-the-job training can increase productivity if the level of training matches the level of the worker's initial ability, and returns to training vary with initial ability. Thus, where in the standard job market signalling model differences in costs lead to (semi)separating equilibria, in our model differences in the productivity returns to training lead employees of different initial levels of ability to select into different types of training.

### *3.1 Single training program model*

We assume that ability ( $a$ ) is uniformly and continuously distributed on  $[0,1]$  (similar to Spence (1974); Lang and Kropp (1986)). Each program has added value ( $v(a)$ ), which is assumed to be linear in ability and consists of the fixed ( $v_0$ ) and marginal returns ( $v_1$ ). Those values are determined by the program characteristics:

$$(1) \quad v(a) = v_0 + v_1 a.$$

We assume that all of the available training programs are advanced or basic by design. The added value of basic training is decreasing with ability ( $v_1 < 0$ ), so low ability employees benefit more from this training. Conversely, for advanced training programs the added value is increasing with ability ( $v_1 > 0$ ), as employees with higher ability can learn more from these programs.

By construction, fixed added value in (1) defines the gains of employees at the bottom of the ability distribution, since it outweighs marginal added value for low values of  $a$ . Analogously, marginal added value determines the gains of employees with high ability, and we assume that fixed and marginal added values are negatively correlated. An increase in marginal added value makes the training more beneficial for higher ability employees, and makes it less useful for low ability employees simultaneously. The same logic holds for an increase in fixed added value: by adjusting to serve the needs of employees with low ability, the training becomes less useful for high ability employees.

The total costs of training include monetary costs and time investment. Following Spence (1973) we assume that costs decrease with ability, as the smarter individuals require less time to master the same material. However, the decreasing nature of the costs ( $c$ ) can be incorporated in net added value, subtracting them from (1), i.e. if  $c = c_0 - c_1 a$ :

$$v_{net}(a) = (v_0 - c_0) + (v_1 + c_1)a = v'_0 + v'_1 a.$$

Therefore, without loss of generality we can make a simplifying assumption that the costs are constant and equal  $c$ .

We assume that worker productivity is not perfectly observed, and therefore individual ability and added value of training are private information. The information the market receives about an individual employee is whether or not he went into training, and general information about the program he took. Based on the information about the program, market agents form beliefs about the level of training. For example, programs that lead to a master's degree equivalent will likely be inferred to be an advanced form of training, while a program labeled a 'knowledge refreshment course' is likely to be judged as basic. Using the available information, market agents make assumptions about the average level of ability of workers who did and did not participate in training. Here we analyse the extreme case when training participation is the only available source of information regarding ability.<sup>1</sup> Market agents infer an employee's ability from the type of training he takes. For example, an employee who signs up for advanced training, is expected to be of higher initial ability than a basic program participant. Therefore, employees take into account market beliefs when deciding to take the training.

The decision to participate in training is a single period voluntary decision. Employees are rational to the extent that they choose an option with the highest benefit, i.e. the employee makes a decision to participate in training if gains from participating are higher than those from not participating.<sup>2</sup>

To distinguish between ability before and after training we refer to ability after the participation decision as knowledge:

$$(2) \quad K(a) = \begin{cases} v_0 + (1 + v_1)a & | \text{ participating in training} \\ a & | \text{ not participating} \end{cases}.$$

---

<sup>1</sup> In practice, employers form increasingly accurate beliefs about the productivity of their employees over time (Altonji and Pierret, 2001; Lange, 2007). However, for the results of our model to hold, productivity needs only be partially unobserved, as we show in Section 3.3.

<sup>2</sup> Without loss of generality we assume that if the employee is indifferent between taking and not taking the training, he will participate.



This means that for employees who do not take training, knowledge is equal to their initial ability.

We assume that wages  $w(a)$  are set at the level of observed employee's productivity ( $p(a)$ ):

$$(3) \quad w(a) = p(a).$$

Since the equilibrium solution differs for different types of training, we look separately at both situations, starting with basic and continuing with advanced training. First, we predict participation in the ideal situation when productivity is perfectly observable, and training is only taken by those who benefit from it in terms of ability. Later we show how the results change once productivity becomes unobservable.

### 3.1.1 Basic training

We first consider a situation where productivity is perfectly observable ( $p(a) = a$ ). In this case, a worker's wage is equal to his ability. Training cannot be used for signaling and is only valued for the skills increase it provides, since the increase in ability translates into an employee's wage raise. Therefore the training is only taken by employees, for whom the net added value of training is positive ( $v(a) - c \geq 0$ ). Equating net added value to zero, we find the participation threshold ability ( $a^*$ ):

$$v(a^*) - c = 0,$$

$$(4) \quad a^* = \frac{c - v_0}{v_1},$$

where  $a^*$  is defined as the ability level for which the gains from training are equal to the gains from not taking the training (zero in this case).

As in this case added value only decreases with ability ( $v_1 < 0$ ), all employees with ability lower than the threshold sign up for the training. Thus, the participation rate equals the threshold.

Now let's look at the situation where ability is not observable.

Since individual ability is private information of the employee and training participation is the only available source of information, the best guess the market can make about a person's knowledge is the average knowledge of people taking (or not taking) training ( $EK(a^*)_p$  and  $EK(a^*)_{np}$  respectively). Therefore, unlike in the previous case where employee's wage was equal to his own

knowledge, his wage is equal to the average knowledge of the group taking training in case he signs up, and the average knowledge of the group not taking training otherwise:

$$(5) \quad w(a) = \begin{cases} EK(a^*)_p & | \text{ participating} \\ EK(a^*)_{np} & | \text{ not participating} \end{cases}.$$

Everyone for whom the benefits from training exceed the benefits from not participating will take the training in equilibrium. The employee with marginal ability ( $a^*$ ), for whom the net benefits from going and not going into training are equal, is indifferent and, as mentioned before, is assumed to participate.

*Proposition 1. In equilibrium, all employees are indifferent between taking and not taking the training.*

Since in equilibrium the marginal employee ( $a^*$ ) is indifferent between two options, we find the participation threshold by equating the difference between wages and costs of training of the group taking training with wages of the group not taking training, and using (5):

$$w(a^*)_p - c = w(a^*)_{np},$$

$$(6) \quad EK(a^*)_p - c = EK(a^*)_{np}.$$

Since (6) does not depend on  $a$  directly, the benefits from going (not-going) into training are the same for everyone and boil down to the expected knowledge of each group (minus costs). This means that, in equilibrium, net benefits (wage minus costs of training) of training are the same for the entire employee pool and are equal to benefits (wages) of the group not taking the training. The equilibrium is determined by the number of employees in each group and their individual abilities. This situation is similar to the perfect competition theory in microeconomics, where the profits of an individual firm depend on the number of firms in the market. This mechanism regulates the number of operating firms. Similarly to our case, in the equilibrium all active firms receive zero profits (equal to the profits they would receive by not operating).

*Proposition 2. Participation rate of a basic training increases with an increase in added value and decreases with an increase in costs.*

As  $EK(a^*)_p$  ( $EK(a^*)_{np}$ ) is the expected knowledge of the group taking (not taking) the training,  $K(0)$  is the knowledge level of the lowest ability employee, and 1 is the maximum possible ability, we can calculate the average knowledge of each group:

$$(7) \quad EK(a^*)_p = \frac{K(0)+K(a^*)}{2},$$

and

$$(8) \quad EK(a^*)_{np} = \frac{a^*+1}{2}.$$

Taking (2) into account we solve (6) for  $a^*$ :

$$(9) \quad a^* = \frac{2c+1-2v_0}{v_1}.$$

Taking into account that  $v_1 < 0$ , we see from (9) that the participation rate depends positively on added value parameters. When  $v_0$  is increasing, participants gain relatively more from the training. The same is true for an increase in  $v_1$  (decrease in absolute value). As the added value curve becomes flatter, on average total gains increase. Therefore both changes lead to increased participation in the program.

*Proposition 3. Basic training should be extremely effective or appeal to a wide audience in order to attract participants.*

We also notice that in order for a basic training program to attract participants ( $a^*$  to be positive), its net fixed added value should be higher than  $\frac{1}{2}$ . Signing up for a basic training of this kind will signal low initial ability, for which the training should compensate by being effective enough to bring the lowest ability employees to an above average level. This extreme result is explained by our assumption that training participation is the only available source of information about productivity. Once we allow for more sources of information, the magnitude of the result becomes smaller (see Section 3.3).

The negative signal associated with basic training implies that this type of training will attract few (if any) participants. This poses a problem for training programs aimed at low skilled employees.

For example, let's assume that the aim is to improve the employees at the bottom 10% of the ability distribution and to bring them to the level just above this 10<sup>th</sup> percentile. Our model shows that nobody will participate in the training due to the negative signal even if the training program is proven to be effective. Employers that aim to target the bottom of the distribution will therefore necessarily have to design their program so that it is appealing to a broader audience in order to induce their employees to participate. However, generalizing the program will decrease its efficiency in raising the quality at the low end of the distribution. While somewhat counterintuitive, decreasing the knowledge gained from training for those at the bottom of the distribution increases their willingness to participate. The other, less realistic, option would be to make the training so effective that low ability employees are better off after the training than not participating at all and being assumed to be average.

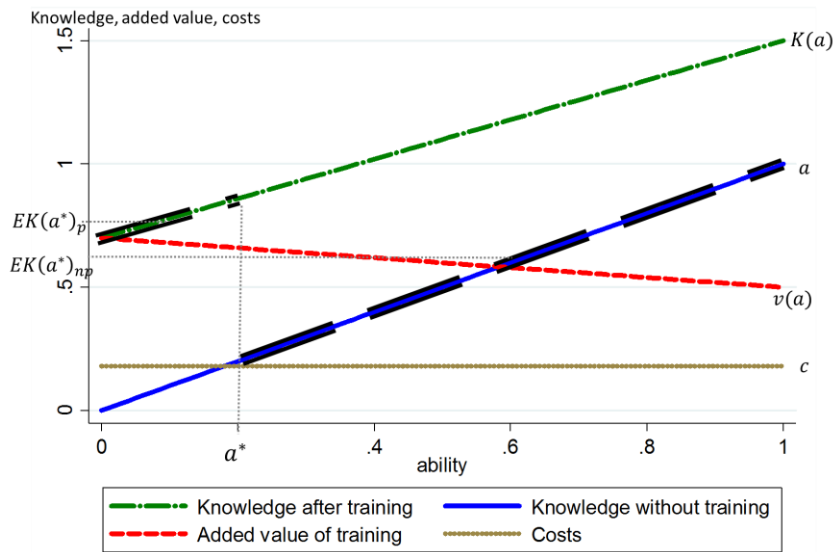
*Proposition 4. Imperfect information decreases the participation rate of the basic program.*

Now we compare this result with the hypothetical situation (4) without signaling discussed earlier. If we rewrite (9) we can treat the second term as a signaling stigma, reducing participation:

$$a^* = \frac{2c-2v_0+1}{v_1} = \frac{c-v_0}{v_1} + \frac{c-v_0+1}{v_1}.$$

A graphical example of a basic training program is provided in Figure 1. The added value and costs of the training are:  $v_0 = 0.7$ ,  $v_1 = -0.2$ ,  $c = 0.18$ . We see that the line of gross added value ( $v(a)$ ) is always above the costs ( $c$ ), and therefore in the situation where productivity is observed and the training does not have any signaling value, all employees sign up for training, and the participation rate is 100%. Everyone improves their initial ability and shifts from  $a$  (solid) to  $K(a)$  (dash-dot) line. The knowledge after training curve ( $K(a)$ ) is the sum of knowledge without training ( $a$ ) and added value of training ( $v(a)$ ). If productivity is not observable, using the formulas derived earlier, we calculate that the stigma of the basic training reduces participation to 20% ( $a^*$ ). Employees with initial ability below  $a^*$  take the training and their knowledge shifts to the  $K(a)$ , while employees with initial ability higher than  $a^*$  stay on  $a$ . The thick dashed black line is used to show the knowledge distribution in the market in equilibrium.

Figure #1. Basic course equilibrium



### 3.1.2 Advanced training

Similar to the previous section we first analyze a situation where signaling is not applicable. Since we use the same formulas, the participation threshold is the same as for the basic program:

$$a^* = \frac{c - v_0}{v_1}.$$

In this case employees with initial ability higher than  $a^*$  take the training, therefore the participation rate equals  $1 - a^*$ .

Now we consider the situation where ability is not observable and employees can sign up for training to signal their ability. Since the program is assumed to be advanced and highly able employees benefit from it the most, smart individuals are expected to sign up. Therefore, if for a particular employee the net added value of the training is negative (added value is lower than the costs), but at the same time his knowledge after training is lower than the average knowledge of employees attending that training, he can signal high knowledge just by subscribing to that training. Therefore, because of the positive signal, more employees are expected to take the training program than just those that benefit from it in terms of increased ability.

*Proposition 5. Participation rate of an advanced training increases with increases in added value and decreases with an increase in costs.*

Again the threshold is derived from (6), but this time representations for expected knowledge of the groups are different due to the fact that the tails of the distribution for which the training is most and least beneficial change places:

$$EK(a^*)_p = \frac{K(1)+K(a^*)}{2} \text{ and } EK(a^*)_{np} = \frac{a^*+0}{2}.$$

$K(1)$  is the knowledge level of the employee with the highest ability and 0 is the lowest possible ability. Then the threshold is

$$(10) \quad a^* = \frac{2(c-v_0)-v_1-1}{v_1},$$

and the participation rate is

$$(11) \quad 1 - a^* = \frac{2(v_1+v_0-c)+1}{\alpha_1}.$$

Here again the participation rate depends positively on added value parameters and negatively on the costs. The underlying intuition is the same as in the previous case. An increase in added value makes more employees better off taking the training and shifts the threshold to the left, increasing the participation rate.

*Proposition 6. The decision to go into training does not have to be related to program effectiveness and can have a pure signaling effect.*

We notice that for low fixed added value ( $v_0$ ) and marginal returns ( $v_1$ ), in the equilibrium without signaling nobody signs up for training because the net benefits are negative. However, when employees' productivity is unobservable and signaling is possible, employees do go into training, because they benefit in terms of wages. If the program is useless but assumed to be for highly able individuals, it provides an opportunity for employees to signal high ability. Therefore, employees at the margin will sign up for a program from which they gain no knowledge if the program is assumed to be advanced.

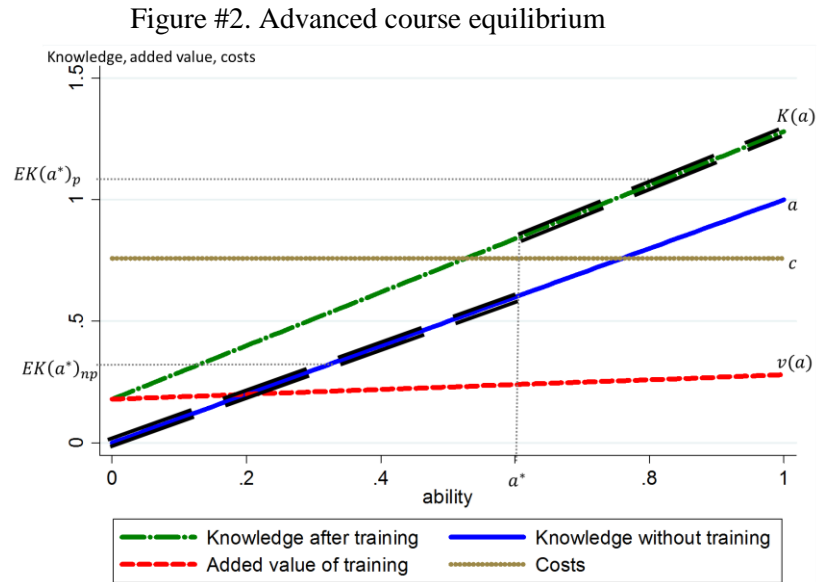
*Proposition 7. Advanced training programs are oversubscribed.*

Now let's rewrite (10) to compare it with (4) and see what happens to the threshold due to signaling:

$$a^* = \frac{c-v_0}{v_1} + \frac{c-v_0-1-v_1}{v_1}.$$

We see that the threshold is lower and more employees are participating, including those for whom training is too costly in addition to those who would take the training without signaling.

Figure 2 provides an example of an advanced program. The added value parameters and costs are:  $v_0 = 0.18$ ,  $v_1 = 0.1$ , and  $c = 0.76$ . We see that for any ability level the costs are higher than total added value. Therefore in the case when worker's productivity is observable, nobody signs up for the training and the participation threshold  $a^* = 1$ . In the case worker's productivity is not observable, participation rate rises to 40% ( $a^* = 0.6$ ), entirely due to the positive signal. Again, the thick dashed black line shows the knowledge distribution in the market.



### 3.2 Multiple training programs model

So far we discussed a situation where employees can only choose one training program. Now we extend the model to a situation where employees can choose between an advanced and a basic program. We assign additional indices  $i = 1$  to the parameters corresponding to a basic training and  $i = 2$  to the parameters of an advanced training. Added value and costs of training are now denoted as:

$$v_i(a) = v_{0i} + v_{1i}a \text{ and } c_i, i = 1,2. i = 1,2.$$

Now employees can choose between three options: basic training, advanced training, and no training. Each option provides a signal, as described in the previous section. As the nature of the programs does not change, we assume that in equilibrium some employees from the bottom of the distribution take basic training (with abilities below some  $a_1^*$ ), and some from the top take an advanced one (with abilities above some  $a_2^*$ ). Then two cases are possible. First, employees in the middle of the distribution (between  $a_1^*$  and  $a_2^*$ ) are better off not taking any training and there are two different thresholds  $a_1^*$  and  $a_2^*$  (Figure 3). Second, there is full training participation and only one threshold  $a^*$ . Employees with ability lower than  $a^*$  take basic training and employees with ability above  $a^*$  take an advanced one (Figure 4).

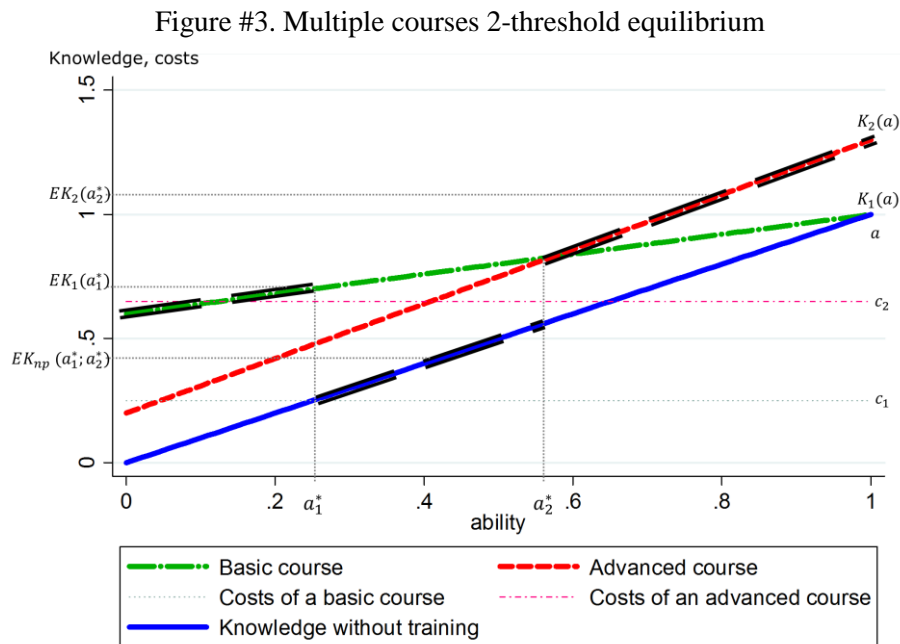
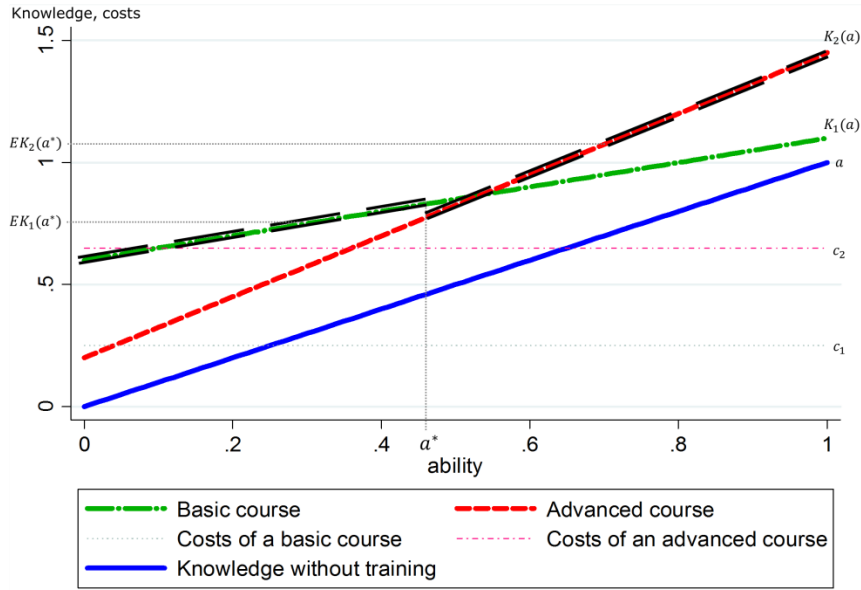




Figure #4. Multiple courses 1-threshold equilibrium



*Proposition 8. Parameters of different programs affect each other's participation.*

We start with the first situation. The threshold of the basic program is strictly lower than the threshold of the advanced program ( $a_1^* < a_2^*$ ). Rewriting (6) for both programs gives a system of equations:

$$(12) \quad \begin{cases} EK_1(a_1^*) - c_1 = EK_{np}(a_1^*, a_2^*) \\ EK_2(a_2^*) - c_2 = EK_{np}(a_1^*, a_2^*) \end{cases}$$

Solving for  $a_1^*$  and  $a_2^*$ :

$$(13) \quad \begin{cases} a_1^* = \frac{2v_{02} + v_{12}[2v_{01} + 1 - 2c_1] + 1 - 2c_2}{1 - v_{11}v_{12}} \\ a_2^* = \frac{2v_{01} + v_{11}[2v_{02} + v_{12} + 1 - 2c_2] - 2c_1}{1 - v_{11}v_{12}} \end{cases}$$

We can see from the parametrical representations of  $a_1^*$  and  $a_2^*$  that each threshold depends on the specifics of both programs.

*Proposition 9. Interventions targeting advanced training have an indirect effect on demand for basic training.*

For example, if the advanced training program gets better (increase in fixed or marginal added value), the share of participants of the basic program increases (together with its threshold). It also

increases with a decrease in costs of the advanced program. Improvements in design of the advanced program ( $\uparrow v_2$ ) as well as availability to a wider audience ( $\downarrow c_2$ ) boost demand for the basic program without any direct interventions.

The mechanism behind this result is the following. A rise in  $v_{02}$  or  $v_{12}$  and a drop in  $c_2$  make the advanced program appealing to more employees and shifts the participation threshold to the left ( $a_2^*$ ). As  $a_2^*$  is also a right margin of the group not taking the training, its average ability decreases ( $\downarrow EK_{np}$ ). As a result, the expected wage of the group not taking any training goes down, while the wage of the group taking basic training does not change. Therefore for employees with abilities close to the threshold  $a_1^*$ , who were better off not taking basic training, the option of taking it is now relatively better. This shifts  $a_1^*$  to the right, which increases both the participation rate and the average knowledge of employees who take the basic training ( $\uparrow EK_1$ ).

*Proposition 10. Increased participation of a basic training reduces the positive signal of an advanced training.*

Conversely, higher added value of basic training ( $\uparrow v_{01}, v_{11}$ ) and lower costs ( $\downarrow c_1$ ) lead to a higher participation threshold ( $\uparrow a_2^*$ ) and a lower participation rate of the advanced program ( $\downarrow 1 - a_2^*$ ). However, as in the previous case, only employees close to margin ( $a_2^*$ ) are affected. Therefore only employees who sign up purely because of the signal choose not to participate in the program, and sorting into training becomes more efficient.

*Proposition 11. Different training programs mutually reduce the absolute value of each other's signal.*

Comparing (9) and (11) with (13) we see that both training programs mutually reduce the absolute value of each other's signal. Moreover, improvements in basic or advanced training alone affect participants of both types of programs, as well as employees not taking training at all.

As mentioned above, this holds for the situation where  $a_1^* < a_2^*$ . Plugging representations for  $a_1^*$  and  $a_2^*$  from (13) into this inequality we get

$$(14) \quad 2[v_{01} - v_{02} + c_2 - c_1] - 1 - v_{12}[2v_{01} + 1 - 2c_1] + v_{11}[2v_{02} + v_{12} + 1 - 2c_2] > 0.$$

(14) can be used as a condition to distinguish between full and partial participation. If (14) holds, we are in the first case and  $a_1^* < a_2^*$ . Otherwise, if the left-hand side of (14) is less or equal to zero we get the second case and  $a_1^* = a_2^* = a^*$ . To find out participation threshold, we substitute  $a_1^*$  and  $a_2^*$  with  $a^*$  in (12) and get

$$(15) \quad a^* = \frac{2(v_{01}-v_{02}+c_2-c_1)-1-v_{12}}{v_{12}-v_{11}}.$$

In this situation all employees sign up for training: employees with ability lower than  $a^*$  choose the basic training program, and employees with higher ability sign up for the advanced program. It is clear from (15) that the participation threshold positively (negatively) depends on the added value of the basic (advanced) training and negatively (positively) on the costs of the basic (advanced) one. This is quite intuitive, as due to the full participation, an increase in the added value (as well as decrease in costs) of the basic program attracts employees who were better off signing up for the advanced training. Therefore the threshold just shifts to the left.

### 3.3 Multiple signals

So far in this chapter we introduced two simple cases where an employee's productivity is either perfectly observed or not observed at all, and the training program is the only available source of information about ability. In practice, however, agents usually have additional ways to infer employees' ability (e.g. personal observations, characteristics such as work experience). Therefore it is logical to include more sources of information into the model.

We assume that in the general case, employee's observed productivity ( $p(a)$ ) is a combination of the signal of a training program ( $EK(a^*)$ ) and information about knowledge ( $K(a)$ ) from other sources:

$$p(a) = \lambda EK_i(a_i^*) + (1 - \lambda)K(a), \quad i = 1,2,$$

where  $\lambda \in [0,1]$  characterizes information asymmetry in the market and can be perceived as weight of the signal of the training program in employee's wage. If  $\lambda = 1$ , we are in the case of perfect asymmetry analysed in detail earlier. With a decrease in  $\lambda$  the importance of the signal of the training program decreases, as information from other sources becomes more reliable. Although a decrease in

$\lambda$  attenuates the signal of training compared to (5), it is clear that, while becoming smaller in magnitude, all the results of the model still hold.

#### 4. Empirical predictions

The results from our model lead to several empirically testable predictions. In this section we suggest some ways to validate our model. We also discuss how the predictions of our model differ from predictions following from the human capital, preferences, personality, and motivation based models.

First, in professions where productivity is easier to observe, low skilled employees should sort into training more often than in professions where productivity is harder to observe. This prediction is somewhat related to Riley's (1979) finding that in jobs where productivity is observable, wages are less strongly related to workers' level of education. In other words, employers disregard the signal sent through the education (in our case training) decision and set their wages based on the level of productivity they observe. Therefore, when productivity is observable, signalling has no value, and sorting into training should solely be based on predicted gains in productivity.

Second, we note that the penalty from sending a negative signal is greater at certain points in a worker's career. In our model, we assumed for simplicity that the signal each course sends is valued equally among all employees. However, those for whom the evaluation by their employer is more important in terms of their future employment, benefit more from not signalling low quality. In terms of the model, these employees will put relatively more weight on the value of the signal ( $EK(a^*)_p$ ) than on the gains in actual knowledge ( $K(a)$ ). In practice, this implies that employees who are on fixed-term contracts are less likely to sort into basic courses than employees that are on permanent contracts. When the signal gains importance relative to actual gains in knowledge, it also means that sorting into training becomes less efficient. At the high participation threshold, employees will undertake training which adds nothing in terms of knowledge for the positive signal. At the low participation threshold, employees who would gain knowledge from basic training will not take it because of the negative signal.

A strongly related prediction is that, considering employers form increasingly accurate beliefs of their employees' productivity over time (Altonji and Pierret, 2001; Lange, 2007), workers have more influence over their perceived quality early in the employment relationship. Therefore, our model predicts less participation in basic courses with negative signalling value for employees in companies that have come under new management. Again, we would expect that this effect is stronger for employees on fixed term contracts.

Importantly, the preceding predictions diverge from predictions made by the human capital model and models that considers individual differences as the main factor in explaining differences in training participation of low skilled workers. If individual differences would drive training participation, there should not be a difference in training behaviour related to contract status or length of the employment relationship. Each employee would sort into the training program that increases their productivity most effectively, regardless of the signal (provided that the returns to training would outweigh the costs). However individual differences may still explain differences in training participation between low skilled workers across occupations because of non-random selection into occupations. For example, it could be the case that a preference for work in which performance is observable is related to training motivation.

Furthermore, because the existence of multiple courses targeted at different parts of the ability distribution reduces the signaling value of each individual course, companies that offer a higher variety of training programs will have a higher training participation rate of their low skilled workers. As we have shown before, adding the option of an advanced course to a choice set consisting only of basic courses reduces the average ability of those employees who do not sort into any training. This decreases the negative signal associated with signing up for a basic course, increasing participation in this course. A related prediction is that, because the existence of advanced courses increases participation in basic courses, companies that offer advanced training opportunities will have a higher level of average worker productivity. This holds regardless of the effectiveness of the advanced training programs, provided that the basic course is effective.

Again, these predictions differ from predictions made by a model based on individual preferences for training or the human capital model. If individual differences drive training

participation, or the decision to enter training is purely based on the costs and predicted returns of that particular course, the existence of an advanced course should not influence participation in a basic course when there are no signalling effects. If sorting into training is based only on gains in productivity, the existence of a course that is dominated in terms of gains should not influence the rational employee's sorting decision under the independence of irrelevant alternatives assumption.

## **5. Concluding comments**

Preference based explanations for the observed low training incidence of low skilled workers leave little potential for interventions aimed at increasing their productivity through on-the-job training. We offer a novel, signalling based explanation of training participation, where low skilled workers do not sign up for training because of the negative signal training participation conveys about their initial ability. We argue that training can increase productivity when workers' prior knowledge matches the level of the training course. Consequently, in jobs where performance is (partly) unobservable, employers can use observed training participation as information about the quality of their employees. This implies that training programs aimed at low ability workers will be underused.

In contrast with the preference based explanation, approaching on-the-job training from a signalling perspective offers potential for interventions. Because courses have spillover effects and mutually affect each other's participation rates, companies and policy makers can increase the training participation of their low skilled employees by offering a variety of courses. The existence of an advanced course increases participation in basic courses, as the average ability of workers that do not take any training decreases. As a result, basic training becomes relatively more attractive for workers at the bottom of the skills distribution.

Our model provides an explanation for why low skilled workers participate less in training, and offers suggestions to increase their participation. However, empirical research should address the question whether our model can explain observed training participation patterns before policy makers take up suggestions based on theory alone. For this reason, we make several empirical predictions that diverge from predictions made by human capital and preference based models. Ultimately, when workers use training as a signal in practice, insights from our model can help companies and policy

makers in designing the most efficient implementation strategy for their on-the-job training programs targeting their low skilled employees.

### Reference list

- Acemoglu, D. (1997). Training and innovation in an imperfect labour market. *Review of Economic Studies*, 64:445–464.
- Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of economic literature*, 40(1), 7-72.
- Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics*, 4, 1043-1171.
- Acemoglu, D., & Pischke, J.-S. (1998). Why do firms train? Theory and evidence. *Quarterly Journal of Economics*, 113(1):79–119.
- Acemoglu, D., & Pischke, J.-S. (1999a). Beyond Becker: Training in imperfect labor markets. *Economic Journal*, 109:F112–F142.
- Acemoglu, D., & Pischke, J.-S. (1999b). The structure of wages and investment in general training. *Journal of Political Economy*, 107(3):539–572.
- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: evidence for overlapping traits. *Psychological bulletin*, 121(2), 219.
- Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1), 313-350.
- Arulampalam, W. N., & Booth, A. L. (2001). Learning and earning: do multiple training events pay? A decade of evidence from a cohort of young British men. *Economica*, 68(271), 379-400.
- Autor, D. H. (2001). Why do temporary help firms provide free general skills training?. *Quarterly Journal of Economics*, 1409-1448.
- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4), 1279-1333.
- Bassanini, A., Booth, A., Brunello, G., De Paola, M. & Leuven, E. (2007) Workplace training in Europe, in *Education and Training in Europe* (Eds) G. Brunello, P. Garibaldi and E. Wasmer, Oxford University Press, Oxford, pp. 143–309.

- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70:9–49.
- Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of human Resources*, 43(4), 972-1059.
- Brunello, G. (2004). Labour market institutions and the complementarity between education and training in Europe. In *Education, Training and Labour Market Outcomes in Europe* (pp. 188-210). Palgrave Macmillan UK.
- Budría, S., & Pereira, P. T. (2007). The wage effects of training in Portugal: differences across skill groups, genders, sectors and training types. *Applied Economics*, 39(6), 787-807.
- Carneiro, P. & Heckman, J. J. (2003). “Human Capital Policy.” In *Inequality in America: What Role for Human Capital Policies?*, edited by James J. Heckman, Alan B. Krueger, and Benjamin M. Friedman. Cambridge, MA: MIT Press.
- Coleman, M., & DeLeire, T. (2003). An economic model of locus of control and the human capital investment decision. *Journal of Human Resources*, 38(3), 701-721.
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an Integrative Theory of Training Motivation: A Meta-Analytic Path Analysis of 20 Years of Research. *Journal of Applied Psychology*, 85(5), 678-707
- Croce, G., & Tancioni, M. (2007). *Disentangling factors behind training participation in Italy* (No. 101). University of Rome La Sapienza, Department of Public Economics.
- Fouarge, D., Schils, T., & Grip, A. de (2013). Why do low-educated workers invest less in further training? *Applied Economics*, 45(18), 2587-2601.
- Hashimoto, M. (1981). Firm-specific human capital as a shared investment. *American Economic Review*, 71(3):475–482.
- Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of labor economics*, 3, 1865-2097.
- Hidalgo, D., Oosterbeek, H., & Webbink, D. (2014). The impact of training vouchers on low-skilled workers. *Labour Economics*, 31, 117-128.



- Kuckulenz, A. & Zwick, T. (2004) *The Impact of Training on Earnings: Differences Between Participant Groups and Training Forms*, Centre for European Economic Research (ZEW), Mannheim.
- Lang, K., & Kropp, D. (1986). Human capital versus sorting: the effects of compulsory attendance laws. *The Quarterly Journal of Economics*, 101(3), 609-624.
- Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, 25(1), 1-35.
- Leuven, E. (2005), The economics of private sector training: A survey of the literature. *Journal of Economic Surveys*, 19: 91–111.
- Leuven, E., & Oosterbeek, H. (2000). The demand and supply of work-related training: Evidence from four countries. In *Research in labor economics* (pp. 303-330). Emerald Group Publishing Limited.
- Lynch, L. M. (1992). Private-sector training and the earnings of young workers. *The American Economic Review*, 82(1), 299-312.
- Meier, S., & Sprenger, C. D. (2015). Temporal stability of time preferences. *Review of Economics and Statistics*, 97(2), 273-286.
- OECD (1999) *Employment Outlook 1999*, Organisation for Economic Co-operation and Development, Paris.
- OECD & European Commission (2013). "The Survey of Adult Skills (PIAAC): Implications for education and training policies in Europe." *Paris and Brussels: OECD Publishing and European Commission, Directorate-General for Education and Culture*.
- Paunonen, S. V., & Hong, R. Y. (2010). Self-Efficacy and the Prediction of Domain-Specific Cognitive Abilities. *Journal of personality*, 78(1), 339-360.
- Riley, J. G. (1979). Testing the Educational Screening Hypothesis. *Journal of Political Economy*, 87(5), S227-52.
- Schwerdt, G., Messer, D., Woessmann, L., & Wolter, S. C. (2012). The impact of an adult education voucher program: Evidence from a randomized field experiment. *Journal of Public Economics*, 96(7), 569-583.

- Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: the impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of personality and social psychology*, *101*(4), 862.
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355-374.
- Spence, M. (1974). Competitive and optimal responses to signals: An analysis of efficiency and distribution. *Journal of Economic theory*, *7*(3), 296-332.
- Stiglitz, J. E. (1975). The theory of " screening," education, and the distribution of income. *The American Economic Review*, *65*(3), 283-300.