

What Makes Firm-based Vocational Training Schemes Successful? The Role of Commitment

Christian Dustmann, University College London

Uta Schönberg, University College London and Institute for Employment
Research

This Draft: September 2011

Abstract

This paper studies a possible market failure in the firm-based vocational training market: training may be too complex to be specified in a contract in a way that is legally enforceable, resulting in the inability of firms to commit to training provision. Our hypothesis is that firm-based vocational training schemes are more successful, as evidenced by higher enrolment rates and lower dropout rates, in countries like Germany, than in Anglo-Saxon countries like the U.K., because commitment to training provision is more widespread in Germany. To support our hypothesis, we present a model of firm-provided training and use it to quantify how much lower training is in the no commitment than in the commitment case, by calibrating our model to match key data moments obtained from survey and administrative data. We find that training in the no-commitment case is, according to our baseline estimate, only 28% of that in the commitment case. We then provide evidence that commitment to training provision is likely to be more problematic in Anglo-Saxon countries than in Germany, and link this to well-structured regulatory and monitoring institutions that exist in Germany but are largely absent in Anglo-Saxon countries.

Keywords: educational policies, training, wage compression

JEL: I2, M53, J24, J31

Correspondence: Christian Dustmann, Department of Economics, University College London, Gower Street, London WC1E 6BT, England. E-mail: c.dustmann@ucl.ac.uk; Uta Schönberg, Department of Economics, University College London, Gower Street, London WC1E 6BT, England. E-mail: u.schoenberg@ucl.ac.uk. We are grateful to V Bhaskar, David Card, Steve Pischke, Hilary Steedman, and especially Paul Ryan for insightful discussions and comments. We thank the editor and two referees for their constructive comments. We gratefully acknowledge financial support from the German Research Foundation (DFG) and the Anglo-German Foundation (AGF).

1 Introduction

Among industrialized countries, workplace related and vocational skill development is (once again) recognized as a key factor in strengthening competitiveness and growth, as reflected in e.g. President Obama's "manufacturing skill speech", or the renewed emphasis on firm-based apprenticeship programmes of the British Government.¹ The key question is how to best develop these skills. Some countries, like the United States, Sweden, France, and Italy, provide vocational training in the form of community colleges or full-time school-based vocational education. Other countries like Germany, Austria and Switzerland run large-scale firm-based apprenticeship schemes that combine firm-provided on-the-job training with state-provided school-based education, lasting between two and three years, and with two in three individuals of each cohort receiving post-secondary education within these schemes.² Firm-based apprenticeship training schemes have a number of advantages over vocational schools: craft techniques and customer interaction may be taught more effectively in a work environment than in the classroom, and firms may know better than schools which skills are needed at the workplace. Firm-based training may also allow for smoother transitions of firm-trained apprentices into employment (see Ryan, 2001a and Parey, 2009 for evidence). Several countries, including the U.S., the U.K., France, and Australia, have attempted in the recent past to implement new, or to expand existing, firm-based apprenticeship schemes (see e.g. Bowers, Sonnet and Bardone, 1999, for an overview), often with little success, as evidenced by low enrolment rates and widespread concerns about the quality of training (see e.g. Wolf, 2011 and Adult Learning Inspectorate, 2006, for the U.K. and Schofield, 2000, for Australia).³

¹See <http://www.whitehouse.gov/the-press-office/2011/06/08/president-obama-and-skills-americas-future-partners-announce-initiatives> and the latest UK 2011 budget (http://cdn.hm-treasury.gov.uk/2011budget_complete.pdf), where the government announces an additional £180 million for up to 50,000 additional apprenticeship places.

²See e.g. Winkelmann (1997) and Harhoff and Kane (1997) for comparisons in post-school training between Germany and the U.S.

³See e.g. Muench (1992), Soskice (1994), and Franz and Soskice (1995) for more details on the German apprenticeship system, Tritscher-Archan (2009) on the Austrian apprenticeship system, Hoeckel et al. (2009) and Hanhart and Bossio (1998) for the Swiss apprenticeship system, Ryan (2000, 2001b), Ryan and Unwin (2001), Gospel (1998), and Steedman et al.(1998) on the U.K. apprenticeship system, and Sharpe and Gibson (2005) and Knight and Mlotkowski (2009) on

In this paper, we address the question why firm based vocational training schemes work in some countries, but not in others. We focus on a possible market failure in the training market which has been discussed in the case of specific training by, for instance, MacLeod and Malcomson (1993) and Prendergast (1993).⁴ Since training takes place inside firms, it is not easily verifiable by a third party and may simply be too complex to be specified in a contract in a way that is legally enforceable. Suppose that firms promise workers a certain level of training in exchange for an apprenticeship wage that is well below the wage offered to unskilled workers. Clearly, firms have a strong incentive to renege on this promise, as they can increase profits by employing the trainee on tasks typically performed by unskilled workers (such as preparing coffee and cleaning machines), rather than teaching him new skills. If training contracts are not enforceable, such a breach of contract will go unpunished. In equilibrium, workers anticipate this and are therefore only willing to accept training contracts which they consider as credible. We refer to the ability of the firm to credibly assure workers that they will not renege on their training promises and deliver the promised training intensity as *commitment to training provision*. Our hypothesis is that apprenticeship training schemes are more successful— as evidenced by higher enrolment rates and lower dropout rates— in countries like Germany than in Anglo-Saxon countries like the U.K. because commitment to training provision is more widespread. We further hypothesize that this may be due to a well-structured regulatory framework and monitoring institutions that exist in Germany but are absent in Anglo-Saxon countries.

To provide support for our hypothesis, we first present a model of firm-provided training. Becker (1962) shows that if training is fully general and labor markets are perfectly competitive, firms are not able to recoup the return of that investment and are therefore not willing to finance general training. Consequently, the training market will break down completely and there will be no training in equilibrium if firms are unable to commit to training provision—any other training level will not be

the Canadian and Australian apprenticeship system.

⁴Acemoglu and Pischke (1999a, 1999b) discuss this problem for the case of general training, but do not develop its implications for training policies.

seen as credible by workers. In contrast, Acemoglu and Pischke 1998, 1999a, and 1999b point out that firms will provide general training even if they cannot commit to training provision if wages are compressed. With wage compression, the profit of firms increases with training, and firms therefore capture some of the return to general training (see also Stevens, 1994a). Based on the work by Acemoglu and Pischke, a common explanation for why firm based training schemes work better in some countries than in others is that countries with successful training schemes exhibit a larger degree of wage compression than countries with less successful training schemes. Our model does not exclude this as an explanation as to why training intensities differ across countries. However, our key point is that two countries with the same degree of wage compression may differ greatly in their training intensities because of differences in firms' ability to commit to training provision.

To demonstrate this, we present a model where wages are compressed due to firm-specific training and asymmetric information between training and outside firms, although the key conclusions from our model do not depend on the exact sources of wage compression.⁵ We show that if firms are able to commit to training provision, they provide a training intensity that is socially optimal and that equates the marginal cost of training with the *worker's and firm's joint* return to training. If, in contrast, firms are not able to commit to training provision, they provide a training intensity that is below the socially optimal level and that equates the marginal cost of training with the *firm's private* return to training. Thus, the training intensity under no commitment is closer to the socially optimal training intensity as provided under commitment, the higher the level of wage compression in the economy.

Figure 1 illustrates this.⁶ In Panel A, we compare the training intensities under commitment and no commitment predicted by our model for increasing levels of wage compression due to firm-

⁵Other reasons for wage compression include asymmetric information between potential trainees and training firms (Autor, 2001), as well as minimum wages and union agreements (e.g. Acemoglu and Pischke, 1999b, 2003; Dustmann and Schönberg, 2007).

⁶See Section 3.1 and footnote 11 for details on the figure.

specific human capital accumulation, assuming that there are no informational asymmetries between training and outside firms. In Panel B, we compare the training intensities under the two scenarios for increasing levels of wage compression due to asymmetric information, now assuming that training is completely general. If training is completely general, and if there are no informational asymmetries between training and outside firms, then the training intensity under no commitment is zero, whereas under commitment, it is equal to the socially optimal level. If, in contrast, training is completely firm-specific, or if the informational advantage of training firms is substantial, the training intensities under no commitment approaches, for the chosen parameter values, 85% and 65% of those under commitment.

For the case of the German apprenticeship system, we then quantify how much lower the training intensities are in the no commitment case than in the commitment case, by calibrating our model to match levels of wage compression due to firm-specificity and asymmetric information that are consistent with those observed in the data. We find that according to our baseline estimate, training under no commitment is only 28% of that under commitment to training provision. This will result in lower enrolment rates into training if there are fixed costs of training and workers are ex ante heterogenous. We finally provide several pieces of evidence that commitment to training provision is more problematic in Anglo-Saxon countries than in Germany, and argue that this may be linked to stricter apprenticeship regulation in Germany.

Our findings suggest that attempts to revitalize apprenticeship training in Anglo-Saxon countries have been less successful than expected, as these countries may not have paid enough attention to the enforceability of apprenticeship contracts and to ensure that firms are able to commit to training provision. Instead, policy makers often provide incentives for workers or firms to participate in apprenticeship programs through subsidies. However, subsidy policies do not change the firm's optimization problem and hence do not help to implement the optimal training intensity under commitment.

The remainder of the paper is organized as follows. The next section develops a model of firm-provided training, analyzes it under both commitment and no commitment to training provision, and discusses the firm's incentives to commit to training provision. In Section 3, we describe how we calibrate the model in order to quantify how much lower the training intensity is under no commitment than under commitment. We report results in Section 3.2. In Section 4, we provide some tentative support for the hypothesis that commitment to training provision is more problematic in Anglo-Saxon countries than in Germany, and that this might be due to the stricter apprenticeship regulation in Germany. We conclude in Section 5.

2 A Model of Firm-Provided Training

In this section, we present a model of firm-provided training where firms may finance training because of wage compression due to firm-specific training and asymmetric information between training and outside firms.

2.1 Set-Up

There are two periods, the first period is the training period. There are many workers and firms, both are risk-neutral. Firms maximize expected profits, and workers maximize expected utility. The worker's productivity y in training and outside firms in period 2 depends on her ability, η , the amount (or quality) of on-the-job training received in period 1, τ , and the degree of firm-specificity of training,

α :

$$y = \begin{cases} h(\tau)\eta & \text{(training firm), and} \\ \alpha h(\tau)\eta & \text{(outside firm).} \end{cases}$$

We assume that $h(\tau)$ is concave in τ and that $0 \leq \alpha \leq 1$. If $\alpha = 0$, training is purely firm-specific; if in contrast $\alpha = 1$, training is purely general. The parameter α is therefore a measure for the

transferability of training. Throughout the paper, we treat α as exogenous; hence, firms and workers cannot choose between investing in general or specific skills.⁷ In this production function, ability and training are complements, i.e. training increases the productivity of high-ability workers more than that of low-ability workers. This is necessary for asymmetric information to provide an incentive for firms to finance training. Note that in this model, workers receive training for one period (e.g. one year). Hence, τ measures *how much* training workers receive during this period and is therefore best understood as a measure for the *quality* of training.

Workers are either of low or high ability. The share of low-ability workers in the economy is common knowledge, and denoted by p . We assume that the expected productivity of untrained workers is the same as that of workers in training, and given by y_0 . Training costs are denoted by $c(\tau)$, with $c'(\tau) > 0, c''(\tau) \geq 0$. In the first period, neither firms nor workers observe the worker's ability. In the second period, only training firms get to know the worker's ability whereas outside firms receive no new information about it. However, outside firms do observe the training intensity which the worker has received.

After training is completed, workers may either stay with or leave the training firm. We endogenize mobility in a simple way and assume that during the training period workers experience a utility shock θ . This shock captures the worker's ex post evaluation of his or her work environment. For instance, it could reflect how workers like their co-workers. Only workers, but not firms, observe θ . The worker's utility in period 2 at the training firm is the sum of her wage, w , and her utility from non-pecuniary job characteristics, θ . A worker's utility at outside firms is equal to the wage offer of outside firms. The utility shock is drawn from a distribution with the cumulative distribution function G , with associated pdf g , and support $[\underline{\theta}, \bar{\theta}]$, $\bar{\theta} > 0$, where G belongs to the family of log-concave distribution functions.

⁷A justification for this assumption is that the production technology is likely to be fixed in the short- and medium run. The partial transferability of training can also be motivated by the skill-weight approach of Lazear (2011). According to this approach, skills and hence training are fully general. However, since firms differ with respect to the weights they place on each skill, training is more valuable in the training firm than in outside firms.

We assume that wages are determined in spot-markets and rule out long-term wage contracts. In the second period, outside firms simultaneously make wage offers to workers by maximizing expected profits. Training firms observe the outside offer, and make a counter offer. We further impose the standard free entry condition on firms: No firm earns positive profits in the long-run.

To summarize, the exact timing of events is as follows.

1. At the beginning of period 1, firms decide how much training to offer to workers and offer workers an apprenticeship wage. At this point, they do not know the ability of workers.
2. Training takes place.
3. At the end of period 1, the training firm finds out the ability of each worker.
4. At the beginning of period 2, outside firms make a wage offer to workers. They observe the training intensity which the worker has received, but not her ability.
5. Training firms observe the worker's (best) outside offer, and make a counteroffer.
6. Workers discover their utility shock and decide whether to stay or quit.

We analyze the model under two assumptions. Under the first assumption, training is verifiable and training contracts are enforceable. Hence, firms can commit to training provision. Under the second assumption, training is not verifiable and training contracts are not enforceable. Hence, firms are unable to commit to training provision. Note that, as it is standard in the literature, we assume in both cases that training is observable by outside firms.⁸ We investigate the firm's decision whether or not to commit to training provision in Section 2.3.

⁸Acemoglu and Pischke (1998, 1999b), Autor (2001), and Malcomson (2003) all assume that training is observable, but non-verifiable.

2.2 Analysis

2.2.1 Wage Determination in the Second Period

We begin with wage determination in the second period. While the amount of training offered in the first period depends on whether firms can or cannot commit to training provision, the rules for wage determination do not. Let v denote the wage offer of outside firms, and w the wage offer of training firms. A worker stays with her training firm if the wage offer of the training firm w plus the utility shock θ exceeds her outside wage offer v . Hence, the probability that the worker stays with the training firm is given by $\Pr(\text{stay}) = \Pr(w + \theta > v) = 1 - G(v - w)$.

Training firms observe the worker's ability and therefore offer different wages to low- and high-ability workers. From now onwards, we therefore index wage offers of training firms by ability ($i = L, H$). Taking workers' outside offers and the first period training choice as given, training firms maximize

$$\max_{w_i} (1 - G(v - w_i))(h(\tau)\eta_i - w_i), \quad i = L, H.$$

>From the first-order condition, wage offers of training firms are implicitly defined as

$$w_i = h(\tau)\eta_i - \frac{1 - G(v - w_i)}{g(v - w_i)}, \quad i = L, H. \quad (1)$$

Outside firms, in contrast, do not observe the worker's ability and therefore offer the same wage to low- and high-ability workers. Due to perfect competition, outside wages are equal to the expected productivity of those workers who leave the training firm. Low-ability workers are more likely to

switch firms, and wage offers of outside firms reflect this adverse selection:⁹

$$v = E[\alpha h(\tau)\eta_i|\text{move}] = \frac{pG(v - w_L)(\alpha h(\tau))\eta_L + (1 - p)G(v - w_H)(\alpha h(\tau))\eta_H}{pG(v - w_L) + (1 - p)G(v - w_H)}. \quad (2)$$

The log-concavity of G ensures that the first order conditions given by equation (1) are sufficient for a maximum, and that for given levels of τ , w_L , w_H , and v are uniquely determined (see Schönberg 2007 for a formal proof).¹⁰

2.2.2 Training Provision

Next, we turn to the firm's decision to train in the first period. Suppose first that firms can commit to training provision. Under this assumption, firms choose training by maximizing expected profits subject to the constraint that workers are offered a utility at least as high as that received by outside firms; otherwise, firms will not be able to attract any worker. This maximization problem corresponds to the maximization of the worker's utility subject to the zero profit constraint (see also Acemoglu and Pischke, 1999b):

$$\max_{\tau} W + U(\tau) \quad \text{s.t.} \quad \Pi(\tau) - c(\tau) - W = 0.$$

Here, W denotes the training wage, U denotes the worker's expected utility in the second period, and Π denotes the firm's expected profit in the second period. The training wage W is determined by the free-entry condition and thus bid up to the point where firms make zero expected profits: $W = \Pi(\tau) - c(\tau)$. Substituting this into the maximization problem, the chosen training intensity

⁹Note that because outside wages are equal to expected productivity, 'poaching externalities' as discussed in Stevens (1994b) are absent in our model.

¹⁰Acemoglu and Pischke (1998) set up a similar model of asymmetric information between training and outside firms and show that their model may exhibit multiple equilibria. In an extension to their basic model, Acemoglu and Pischke (1998) point out that if the timing of events is the same as in our model (i.e. outside firms make a wage offer to trained workers and training firms make a counter offer), the equilibrium is unique.

under commitment, τ^C , satisfies:

$$\frac{\partial \Pi(\tau^C)}{\partial \tau} + \frac{\partial U(\tau^C)}{\partial \tau} = \frac{\partial c(\tau^C)}{\partial \tau}. \quad (3)$$

Under commitment to training provision, the training intensity chosen by the firm equates the firm's and worker's joint return to training (left hand side) with the marginal cost of training (right hand side). Notice that this is also the socially optimal training intensity that maximizes total welfare in the economy.

Contrast this with the case of no commitment to training provision, where training contracts are not enforceable. In this case, training firms may initially promise trainees a certain training intensity, but later renege on this promise and provide a different training intensity. Now suppose at the beginning of the training period firms offer workers the socially optimal training intensity τ^C . However, workers do not perceive this training intensity as credible, as firms can increase their profits by providing a different training intensity. The only training intensity workers consider credible is the one that maximizes the firm's future profit. The training intensity chosen by the firm under no commitment, τ^{NC} , therefore equates the marginal cost of training with the firm's private return to training:

$$\frac{\partial \Pi(\tau^{NC})}{\partial \tau} = \frac{\partial c(\tau^{NC})}{\partial \tau}. \quad (4)$$

Equation (4) highlights that under no commitment to training provision, firms provide training only if the firm's expected profit is increasing in training, i.e. $\frac{\partial \Pi}{\partial \tau} > 0$. This is what Acemoglu and Pischke (1999a, 1999b) refer to as wage compression. In Appendix A, we show that in our model wage compression arises due to firm-specific training and asymmetric information between training and outside firms.

Comparing the training intensity under no commitment (equation (4)) and commitment (equation

(3)), it is apparent that firms provide more training under commitment ($\tau^C > \tau^{NC}$). Hence, under no commitment to training provision, training is below the socially optimal level. The reason is that under no commitment, firms choose training by maximizing their *private* return to training, while under commitment, training is chosen by maximizing the firm’s and worker’s *joint* return to training. Equations (4) and (3) further highlight that training intensity under no commitment is closer to the socially optimal training intensity, as provided under commitment to training provision, the larger wage compression is. Figure 1 in the Introduction illustrates this. In Panel A, we compare the training intensities under commitment and no commitment predicted by our model for increasing levels of wage compression due to firm-specific human capital accumulation, assuming that there are no informational asymmetries between training and outside firms (i.e. we assume that $\eta_L = \eta_H$). If training is completely general (i.e. $\alpha = 1$), the training intensity under no commitment is zero, whilst that under commitment equates the marginal productivity and the marginal cost of training (i.e. $\partial h(\tau)\eta/\partial\tau = \partial c(\tau)/\partial\tau$). In this case, the difference between the two training intensities is largest. If, in contrast, training is completely firm-specific (i.e. $\alpha = 0$), the training intensity under no commitment approaches, for the chosen parameter values, 85% of that under commitment.

In Panel B, we compare the training intensities in the two scenarios for increasing levels of wage compression due to asymmetric information, now assuming that training is completely general. If the productivity of low-ability workers is the same as that of high-ability workers (i.e. $\eta_L = \eta_H$), the training intensity under no commitment to training provision is zero and the difference between the training intensities under commitment and no commitment to training provision is largest. If, in contrast, the productivity of high-ability workers is five times larger than that of low-ability workers, the ratio between the two training intensities is, for the chosen parameter values, 65%.¹¹

¹¹

In the figure, we assume that the share of low-ability workers is 0.5, and that their productivity is 1. The dispersion of the utility shock is 0.4 (which is close to our baseline estimate; see Table 3). In Panel A, we set the productivity of high-ability workers to 2.75 (which is again close to our baseline estimate, see Table 3). See Section 3.1 for the functional forms of the production and cost functions, as well as the chosen form of the distribution of utility shocks.

2.2.3 Who Pays for Training: Workers or Firms?

Who bears the training costs under no commitment and under commitment, firms or workers? Under both commitment and no commitment to training provision, apprenticeship wages are bid up until firms' expected profits are equal to zero: $W^j = y_0 + \Pi(\tau^j) - c(\tau^j)$. Untrained workers receive a first period wage equal to their productivity plus the firms' expected second period profit in the absence of training: $y_0 + \Pi(0)$.

Under no commitment, firms choose training by maximizing expected profits (see equation (4) for the first order condition). In the absence of wage compression, no training takes place. With wage compression, profits with training are higher than profits without training: $\Pi(\tau^{NC}) - c(\tau^{NC}) > \Pi(0) > 0$. Consequently, as apprenticeship wages and wages of untrained workers are bid up until firms' expected profits are equal to zero, apprenticeship wages exceed wages of untrained workers as well as the productivity of apprentices, y_0 , and firms bear all the training costs.¹² This reflects that under no commitment workers are not willing to accept a wage cut to finance training.

In contrast, under commitment, firms choose training by maximizing the workers' utility subject to the zero profit constraint (see equation (3) for the first order condition). In the absence of wage compression, the firm's profit in the second period with training equals that without training ($\Pi(\tau^C) = \Pi(0)$). Consequently, apprenticeship wages are lower than wages of untrained workers by the amount of the training costs $c(\tau^C)$, and apprentices bear all the training costs. With wage compression, $\Pi(\tau^C) > \Pi(0)$. Therefore, the difference between apprenticeship wages and wages of untrained workers is less than the training costs, so that firms and workers share the training costs.

This suggests a possible way to distinguish between commitment and no commitment. Under no commitment, apprenticeship wages exceed wages of untrained workers and the productivity of ap-

¹²Note that if the productivity of workers in training is below that of untrained workers (implying that there is a fixed cost of training), then apprenticeship wages may be below those of untrained workers (but above the productivity of apprentices) and firms and workers may share the total costs of training.

prentices, while under commitment, apprenticeship wages may be below wages of untrained workers and the productivity of apprentices. However, this stark conclusion depends on our assumption that apprentices are as productive as untrained workers. If, as it is likely, apprentices are less productive than untrained workers, then apprenticeship wages may be lower than wages of untrained workers even under no commitment. In this case, the ratio between the apprenticeship wage and the wage of untrained workers will *ceteris paribus* be higher under no commitment than under commitment to training provision. A further distinguishing feature between the two scenarios is that under no commitment apprenticeship wages exceed, whereas under commitment they may be below, the productivity of apprentices. In Section 4.1, we will compare apprenticeship wages, wages of untrained workers, and the productivity of apprentices to infer the ability of firms to commit to training provision in the U.K. and Germany.

2.3 Commitment to Training Provision and Time Inconsistency

So far, we have treated the firm's ability to commit to training provision as exogenous. We now discuss the determinants of that decision. Assume first that firms expect to be in the training market for two periods only. Recall that the training intensity and apprenticeship wage combination that maximizes total welfare, i.e. the sum of workers' utilities and firms' profits, are those under commitment, τ^C and W^C . Consequently, if the firm offers any training-apprenticeship wage combination other than τ^C and W^C that does not maximize the worker's and firm's joint welfare, firms and workers will be able to come up with a private agreement that makes both parties better off. However, once workers have accepted the training offer, firms have an incentive to deviate from the promised training intensity τ^C , and offer the training intensity that maximizes their own, as opposed to the joint, return to training. From equation (4), this is the one under no commitment, τ^{NC} . This is a typical time inconsistency problem.

If training is verifiable in court, and training contracts are enforceable, then deviating firms will be detected and punished. This eliminates the time-inconsistency problem, provided that the expected punishment (left hand side) exceeds the gain from renegeing on the promised training intensity (right hand side):

$$q \cdot P = (\Pi(\tau^{NC}) - c(\tau^{NC})) - (\Pi(\tau^C) - c(\tau^C))$$

Here q denotes the probability that a firm that does not deliver the promised training intensity will be detected, and P is the punishment in case of detection. As training takes place within firms, it may simply be too complex to be specified in a contract in a way that is legally enforceable, and thus not be verifiable. One solution to this problem is external regulation of apprenticeship training, which may help to detect firms that do not deliver the promised training intensity and impose fines on these firms, leading to enforceable training contracts and enabling firms to commit to training provision.

Acemoglu and Pischke (2000) also point out that a successful apprenticeship system may require external regulation. Their explanation is, however, different from ours. Acemoglu and Pischke (2000) analyze their model under no commitment to training provision and focus on the external certification of skills. In their model, certification allows outside firms to observe whether the worker has received some training which in turn forces outside and training firms to pay higher wages to trained workers than to untrained workers. This then provides an incentive for workers to exert effort during training— which, in Acemoglu and Pische’s model, is necessary for training to be productive. In our model, in contrast, outside firms are able to observe training regardless of whether they are able to commit to training provision. Observability of training is, however, not sufficient to ensure that training contracts are legally enforceable. We argue that external apprenticeship regulation makes training contracts legally enforceable, and thereby enables firms to commit to training provision.

If firms live for multiple periods (and workers for only two periods), firms have in principle an incentive to build up a reputation for providing high-quality training above that under no commitment

even in the absence of external regulation which ensures the verifiability of training. Market forces may therefore be strong enough to ensure that firms commit to training provision. However, whether firms are indeed able to develop a reputation for providing high-quality training crucially depends, among other things, on the information potential trainees have about the training intensities which firms have provided in the past. External regulation of apprenticeship training may transmit this type of information to potential trainees, and hence facilitate commitment to training provision. We discuss apprenticeship regulation in detail in Section 4.2.

3 Model Calibration

In our model, the training intensity under no commitment is unambiguously lower than the socially optimal training intensity under commitment to training provision, and the difference between the two training intensities is the larger the lower wage compression. In this section, we quantify *how much* lower training intensities are in the no-commitment than in the commitment case. To this end, we calibrate our model for parameter values that are consistent with the levels of wage compression observed in the data, which we estimate from administrative and survey data for Germany. We then compare the training intensity under the commitment and the no-commitment case.

3.1 Choice of Parameter Values

For our calibration, we first need to impose functional forms on the production- and cost technology, as well as on the distribution of utility shocks. For the production- and cost technology, we assume that $h(\tau) = (1 + \tau)$ and $c(\tau) = 0.5\tau^2$.¹³ We further assume that utility shocks θ are drawn from a logistic distribution with mean 0 and scale parameter b , i.e. $G(\theta) = 1/(1 + e^{-\theta/b})$. We normalize

¹³We have repeated the simulation exercise for two alternative specifications of the production and cost technology: $h(\tau) = (1 + \tau^{1/2})$ and $c(\tau) = 1/3\tau^3$. Our results are robust to these alternative functional forms.

the share of low-ability workers to $p = 0.5$ and the productivity of low-ability workers to $\eta_L = 1$.¹⁴ This leaves three parameters to be determined: α (the degree of transferability of training), η_H (the productivity of high-ability workers), and b (the scale parameter of the logistic distribution of utility shocks).

The degree of transferability of training, α , captures wage compression due to firm-specific human capital in our model. From Figure 1, Panel A, the lower α , the larger the wage compression due to firm-specific human capital and the smaller the difference between the training intensity under no commitment and commitment to training provision. We choose α using direct information about the applicability of apprenticeship skills from the *German Qualification and Career Survey*, using the 1979, 1985 and 1992 waves. In these waves, workers are asked about the proportion of skills acquired during apprenticeship training that are applicable at their current job. Survey responses are categorical, distinguishing between very much, a lot, some, not too much, and very little. We quantify the extent to which human capital is specific rather than general by assigning to the five categories the values 90, 75, 50, 25, and 10.¹⁵ Our analysis is based on men in West Germany who completed an apprenticeship and are between 18 and 55 years old.

In Panel A of Table 1, we report estimates for the degree of transferability of apprenticeship training separately by firm size. We compare the applicability of skills for four groups of workers: workers who are still employed in the apprenticeship firm and occupation, workers who left the apprenticeship firm but are still employed in the apprenticeship occupation, workers who switched occupations, but are still employed in the apprenticeship firm, and workers who left both the apprenticeship firm and occupation.¹⁶ Our results pooled for all firms (column 1) indicate that workers who left the

¹⁴We have repeated the simulation exercise for two alternative values: $p = 0.25$ and $p = 0.75$. Our overall conclusions are unchanged.

¹⁵Our results are similar if we assign the values 95, 80, 50, 20, and 5 or 85, 70, 50, 30, and 15.

¹⁶Occupational mobility is defined at the 2-digit level which distinguishes between 93 occupations. Our regressions control for the year the apprenticeship ended, experience and its square, firm size, apprenticeship occupation (20 occupations), high school degree (*Abitur*), and age at end of apprenticeship.

apprenticeship firm but not the apprenticeship occupation can employ 4.40% less of their skills than a worker who is still employed at the apprenticeship firm and occupation. The corresponding number is 8.8% for workers who are still employed at the training firm, but left the training occupation, and 34.6% for workers who have left both the training occupation and the training firm.¹⁷ Hence, the firm-specific component of apprenticeship training is small, around 5%. In Panel B of Table 1, we display our estimates for the degree of the transferability of skills, α . Our baseline estimate, pooled for all firms, for α is 0.956 ($1 - 0.044$).

The productivity of high-ability workers, η_H , captures wage compression due to asymmetric information in our model. From Figure 1, Panel B, the larger η_H , the larger the wage compression due to asymmetric information and the smaller the difference between the training intensity under no commitment and commitment to training provision. The larger the scale parameter of the logistic distribution of utility shocks, b , the more workers quit the firm after training. This in turn reduces the firm's return to training, $\partial\Pi/\partial\tau$, and consequently the difference between the training intensities under commitment and no commitment to training provision.¹⁸

We obtain estimates for η_H and b by matching two data moments, the quit rate after training, and the wage differential between workers who (initially) stay with or move away from the training firm, to respective moments generated by our model.

The quit rate after training in our model is given by

$$q^M = p(1 - G(v - w_L)) + (1 - p)(1 - G(v - w_H)).$$

¹⁷In related research, Fitzenberger and Spitz-Öner (2004) and Fitzenberger and Kunze (2005) conclude that the specificity of training does not prevent apprentices from switching occupations: They find that occupational switches are frequent, and are likely to occur in order to realize better wage and career prospects. Winkelmann (1996) notes the low retention rates of apprentices after training, and concludes that apprentices acquire predominantly portable skills.

¹⁸A figure that plots the training intensities under commitment and no commitment to training provision predicted by our model against b , for constant values of α and η_H is available on request.

The stayer-mover wage differential in our model is $SM^M = E[w|\text{stay}] - v/E[w|\text{stay}]$, where

$$E[w|\text{stay}] = \frac{p(1 - G(v - w_L))w_L + (1 - p)(1 - G(v - w_H))w_H}{p(1 - G(v - w_L)) + (1 - p)(1 - G(v - w_H))},$$

and w_L , w_H , and v are given by equations (1) and (2). Both q^M and SM^M are complicated functions of b and η_H . We compute these model moments for a fine grid of values for b and η_H , using our estimate for the degree of transferability of apprenticeship skills, α . Let q^D and SM^D denote the corresponding data moments. We then pick those values for b and η_H that minimize the sum of the squared distance between the model and the data moments:

$$\min_{b, \eta_H} (q^M(b, \eta_H) - q^D)^2 + (SM^M(b, \eta_H) - SM^D)^2. \quad (5)$$

We estimate the quit rate (q^D) and the stayer-mover wage differential (SM^D) using data from the IABS 1975-2001, a two percent sample of administrative social security records. We restrict the sample to men in West Germany who finished the apprenticeship between 1980 (this restriction ensures that we observe the size of the apprenticeship firm) and 1998 (this restriction ensures that every apprentice is observed for at least three years).

The first row in Table 2 shows that overall, about 36% of apprentices leave the apprenticeship firm after apprenticeship completion. The quit rate declines substantially with firm size: it is 44.76% in training firms with less than 10 employees, but only 25.54% in firms with more than 500 employees. In the second row of Table 2, we display the wage differential between workers who stay with the apprenticeship firm ('stayers') and those who leave the apprenticeship firm ('movers'). Our regressions control for the worker's age at the beginning of the apprenticeship, apprenticeship duration, high school degree (*Abitur*), foreign status, as well as the industry and the size of the apprenticeship firm. Overall, stayers earn 3.6% higher wages than movers. There is again substantial heterogeneity by the size of the

apprenticeship firm: In firms with less than 10 employees, movers earn 2.2% *higher* wages than stayers. In firms with at least 500 employees, in contrast, movers earn 12.7% *lower* wages than stayers.¹⁹

Note that the model moments can be computed either under the assumption of commitment to training provision, or under the opposite assumption of no commitment to training provision. Since training firms in Germany are likely to be able to commit to training provision (see Section 4.2 for empirical evidence), we calculate in our baseline results the model moment under the assumption of commitment to training provision. For robustness, we also compute the model moments under the opposite assumption of no commitment to training provision. We then compare the training intensity under the two scenarios, and report the ratio between the two, τ^{NC}/τ^C .

3.2 Results

In Table 3, we report the values for the productivity of high-ability workers (η_H), and the scale parameter of the distribution of non-pecuniary job characteristics (b), that minimize the squared distance between the data and the model moments (see equation (5)). The model moments are first calculated under the assumption that all firms commit to training provision (Panel A), and then under the opposite assumption that no firm commits to training provision (Panel B). In both cases, we are able to fit the data moments almost perfectly.

Notice that the productivity of high-ability workers, η_H , increases with firm size, regardless of how the model moments are obtained. This is because the stayer-mover wage differential is larger in larger training firms, while the degree of transferability is roughly similar across firms of different sizes. Hence, our model attributes the larger wage disadvantage of movers in large training firms to a higher degree of asymmetric information in these firms. This is reasonable, as the production technology in large firms is likely to be more sensitive to ability than that in small firms. Moreover, the scale parameter of the distribution of utility shocks, b , decreases with firm size. This is because the quit

¹⁹See e.g. Euwals and Winkelmann (2004) and Wachter and Bender (2006) for similar results.

rate after apprenticeship training is higher in smaller training firms.

How much lower is training under no commitment than under commitment? In the final column of Table 3, we report the ratio between the training intensities, τ^{NC}/τ^C . According to our baseline estimate which computes the model moments under the assumption of commitment to training provision (Table 3, Panel A), the training intensity under no commitment is 28.3% of that under commitment. This compares to 43.6% when model moments are calculated under the opposite assumption of no commitment to training provision. There is considerable heterogeneity by firm size. In training firms with less than 10 employees, the ratios between the training intensities under no commitment and commitment are 13.5% and 34.8%, respectively. For large firms with more than 500 employees, in contrast, training under no commitment is 56.1% and 66.8% of that under commitment.

Our model allows for only three reasons why wages of movers differ from wages of stayers: firm-specific training, asymmetric information, and non-pecuniary job characteristics. Differences in wages of stayers and movers may additionally reflect differences in occupation-specific training, differences in search capital, or differences in the productivity of the training firms. In Appendix B, we argue that our key conclusion—i.e. that the lack of commitment results in substantially lower training intensities—is robust to these extensions.

So far, we have inferred wage compression indirectly, by exploiting information on the applicability of apprenticeship skills, and by matching key model moments—the stayer-mover wage differential and the quit rate after apprenticeship training—to their model counterparts. In principle, we could measure wage compression directly by comparing the effect of training on wages with that on productivity. While there are no studies that conduct such an analysis for apprenticeship training, several studies do so for on-the-job training. A typical estimate suggests that the increase in wages due to training is only about 50% of the increase in productivity (see e.g. Dearden et al., 2006, Conti, 2005, Konings and Vanormelingen, 2009), although the hypotheses that the increase is the same typically cannot be

rejected. According to our baseline estimates ($\alpha = 0.956, \eta_H = 2.46, b = 0.45$, see Table 3, Panel A), our model predicts that the wage increase due to training is roughly 68% of the productivity increase due to training. This is not too far off the estimates of wage compression presented in these papers, and implies that the training intensity under no commitment is only 28% of that under commitment. For the parameters calibrated to fit the stayer-mover wage differential and the quit rate in large firms, our model predicts that the wage increase due to training is only 34% of the productivity increase due to training—which corresponds to a larger level of wage compression (and thus to a higher ratio between the training intensity under commitment and no commitment to training provision) than that reported in the literature we cite above. Yet, the training intensity under no commitment to training provision is only 60% of that under commitment to training provision.

We therefore conclude that, for reasonable levels of wage compression consistent with those observed in the data, the training intensity under no commitment is substantially lower than that under commitment to training provision. Note that this will result in lower enrollment rates into training schemes if there are fixed costs of training and workers are ex ante heterogeneous. While we have not explicitly modeled this, a particularly simple way to do so is to allow for heterogeneous fixed costs of training, reflecting for instance workers' motivation to undergo training.

To see this, let F_i denote these fixed costs, and assume that F_i is drawn from a distribution with cdf $H(\cdot)$. Workers prefer training over no training if their utility with training, $W^j + U(\tau^j) - F_i$, $j = C, NC$, exceeds their utility without training, $W^0 + U(\tau^0)$:

$$[W^j + U(\tau^j)] - [W^0 + U(0)] > F_i, \quad j = C, NC.$$

There is a threshold $F^{*j}, j = C, NC$, such that all workers with a fixed cost below F^{*j} choose to be trained while all workers with a fixed cost above F^{*j} prefer not to be trained. Since workers' utility with training is higher in the commitment than in the no commitment case (i.e. $W^C + U(\tau^C) >$

$W^{NC} + U(\tau^{NC})$), the threshold is higher in the commitment than in the no commitment case (i.e. $F^{*NC} < F^{*C}$). Consequently, enrolment into training schemes is higher under commitment than under no commitment to training provision: $H(F^{*NC}) < H(F^{*C})$.

4 Discussion and Policy Implications

4.1 Is Commitment to Training Provision Lower in the U.K. than in Germany?

Could commitment therefore be the key element to understand why firm-based apprenticeship training schemes are less successful in countries other than Germany, Austria and Switzerland? In this section, we provide evidence that is in line with this hypothesis, focusing on the U.K. The first piece of evidence that commitment to training provision is likely to be more problematic in the U.K. than in Germany is the low quality of apprenticeship training, which is a prime concern in the U.K. (see e.g. Ryan and Unwin, 2001; Ryan, Gospel and Lewis, 2007).²⁰ The low quality of training could also be responsible for the low enrollment rates in apprenticeship programs: the percentage of 16-18 year olds that participate in work-based learning decreased from 9.3 percent in 2000 to 6.4 percent in 2009.²¹ A further indication of the low quality of apprenticeship training in the U.K. are the high dropout rates (see row (1) of Table 4). In 2002, 76% of apprentices did not complete the apprenticeship program. By 2005, this number declined to 47% (Lewis and Ryan, 2007; Adult Learning Inspectorate, 2006). The corresponding numbers in Germany are 23.4% in 2002 and 19.4% in 2005, respectively (Berufsbildungsbericht 2006, 2007).²² Our interpretation of the low quality of apprenticeship training in the U.K. (and other Anglo-Saxon countries) is that workers obtain the training intensity under no

²⁰The 5th report of the House of Lords Select Committee on Economic Affairs (2007) also identifies the low quality as a main problem of the U.K. apprenticeship program. Similar concerns are voiced in the U.K. government commissioned "Wolf Report" (Wolf 2011), which gives a damning verdict on the quality of vocational education.

²¹See Department for Education <http://www.education.gov.uk/rsgateway/DB/SFR/s000938/index.shtml>

²²The evidence for other Anglo-Saxon countries points in a similar direction. Dropout rates range from 63% in the U.S., 52% in Australia, and 39% in Canada (see Table 10 in National Centre for Vocational Education Research (2009) for Australia, Table 5 in Sharpe and Gibson (2005) for Canada; the number for the U.S. refers to the construction sector, for apprentices who enrolled in the program in 1998 (United States Government Accountability Office, 2005)). Sharpe (1999) provides additional evidence for low completion rates in Canada.

commitment, τ^{NC} , instead of the higher one under commitment, τ^C .

Recall from Section 2.2.3 that a further indication of the lack of commitment to training provision is a high ratio between apprenticeship wages and wages of untrained workers. Table 4, row (2) shows that this ratio is almost twice as high in the U.K. than in Germany (0.83 versus 0.43).²³ A further difference between commitment and no commitment to training provision is that under no commitment apprenticeship wages exceed, whereas under commitment they may be below, the productivity of apprentices. We investigate this in row (3) of Table 4. Here, we first report the mean difference between the productivity and the wage of apprentices in 3-year apprenticeship programmes in Germany in the private sector. We base our analysis on a survey data set, the *German Apprenticeship Training Survey*, and construct an annual productivity measure.²⁴ Overall, the productivity of apprentices somewhat exceeds the wage of apprentices. Comparable data for the U.K. exists for the construction sector, which is also the sector that trains the most apprentices in the U.K. (38% in 2002 to 2004, based on the British Labor Force Survey). In Germany, the productivity of apprentices in this sector exceeds their wage by 550 Euros—which is similar to the difference for all sectors. In the U.K., in contrast, the annual wage of apprentices exceeds their productivity by 2,800 GBP.

To sum up, the evidence in Table 4 suggests that apprentices in Germany are more willing to accept a wage cut to finance training than apprentices in the U.K.—which they are only willing to do if firms are able to commit to training provision.

4.2 Apprenticeship Regulation and Commitment to Training Provision

A key question is which policies encourage commitment to training provision? And why is commitment to training provision likely to be lower in the U.K. and other Anglo-Saxon countries than in Germany,

²³Steedman (2008) provides additional evidence that England has far higher apprenticeship wages than continental countries with well-functioning firm-based training systems, like Germany, Switzerland and Austria. In Australia, the ratio between the apprenticeship wage and the wage of unexperienced untrained workers is around 87% (derived from Bittman et al. 2007). Lynch (1993) also reports that the ratio between apprenticeship wages and adult untrained wages is much higher in the U.S. than in Germany.

²⁴See Beicht et al. (2004) for more details.

Austria, and Switzerland? We offer one possible explanation: The apprenticeship system in the German-speaking countries is highly institutionalized and externally regulated, whereas the systems in Anglo-Saxon countries are more market driven. We now describe the regulation of apprenticeship training in Germany and contrast it with that in other Anglo-Saxon countries, focusing on the U.K.²⁵

Most importantly, in Germany the *Vocational Training Act* provides apprentices with the right to take legal action if the apprenticeship firm consistently violates its obligations (e.g. if apprentices consistently perform activities during training with no or little learning content), and are advised by external bodies throughout this process. A similar detailed legislation does not exist in the Anglo-Saxon countries.

There are several other mechanisms in place in the German apprenticeship system which make the system more transparent and thus help firms to build a reputation for providing high-quality training and which are absent in the Anglo-Saxon system. First, chambers of crafts, industry and trade regularly monitor training firms, and have the power to withdraw the firm's permission to train apprentices if firms do not meet the required standard. In the U.K., in contrast, the Training and Enterprise Councils are voluntary, operate at the local level, have purely advisory status, and lack statutory power. Second, in Germany apprentices are obliged to attend vocational schools once or twice per week, where they are taught general subjects, such as Math and English, as well as subjects specific to their occupation. This provides an opportunity to find out about training in other firms. In the U.K., Canada, Australia, and the U.S., in contrast, vocational school attendance is not compulsory.²⁶ Finally, and perhaps most importantly, in Germany apprentices sit final exams at the

²⁵See e.g. Muench (1992), Soskice (1994), and Franz and Soskice (1995) for more details on the German apprenticeship system, Tritscher-Archan (2009) on the Austrian apprenticeship system, Hoeckel et al. (2009) and Hanhart and Bossio (1998) for the Swiss apprenticeship system, Ryan (2000, 2001b), Ryan and Unwin (2001), Gospel (1998), and Steedman et al.(1998) on the U.K. apprenticeship system, and Sharpe and Gibson (2005) and Knight and Mlotkowski (2009) on the Canadian and Australian apprenticeship system. Lynch (1994) provides an insightful comparison between the U.S. training system and the German training system, emphasizing that training institutions are far less developed in the U.S., with a high degree of firm-individual autonomy in training investment, and low assessment standards.

²⁶For instance, while in the U.K. many apprentices (57% according to a 2007 survey) receive some off-the-job instruction, it is not required and the educational content varies (Department for Innovation, Universities, and Skills, 2008; see also Ryan, 2000, 2001b). In the U.S., federal regulation recommends 144 hours of related technical instructions

end of the apprenticeship, which are centralized and organized by the chambers. This may help to detect firms that do not deliver on training promises, as the following example shows. Suppose that the exam is designed such that a (high-ability) worker who received the training intensity under no commitment, τ^{NC} , fails the exam while a (low-ability) worker who received the training intensity under commitment, τ^C , passes the exam. In this case, the firm's pass rate will signal whether the firm has offered the training intensity under commitment or no commitment. In Australia, Canada, U.S. and the U.K, apprentices also receive a certificate at the end of the apprenticeship program. However, the testing of apprentices is often carried out by internal assessors, or the employer itself.²⁷ If firms let apprentices pass the exam even if they received the training intensity under no commitment, τ^{NC} , then the firm's pass rate will no longer signal commitment to training provision.

4.3 Are Subsidies a Solution?

Rather than implementing policies that help to encourage commitment to training provision, a popular policy to increase apprenticeship training are subsidies, either given to apprentices or to training firms. For instance, in the U.K., apprenticeship funding is available from the National Apprenticeship Service, which covers 100% of the cost of off-job training for apprentices aged 16-18, and 50% of that cost for apprentices older than 19.²⁸ Schemes like the 2010 Apprenticeship Grants for Employees (AGE), which offers up to £2,500 to employers who take on unemployed 16-17 year olds as apprentices, are a further example. In Canada, a maximum \$2,000 taxable cash grant is available to apprentices who complete the first or second year of their apprenticeship program. A further \$2,000 taxable cash grant may be paid upon completion of the program.²⁹ In Australia, training firms typically receive \$1,250 for

per year, which may consist of instruction delivered by the employer at the employer's place of business ("Standards of apprenticeship", *Code of Federal Regulations* Title 29, pt. 29.5).

²⁷For more information, see the 'Training and Employment' tab at <http://www.apprenticeships.org.uk/Employers/Other-Questions.aspx> for the U.K., Knight and Mlotkowski (2009) for Australia, and Lynch (1994) and Blanchflower and Lynch (1994) for the U.S.

²⁸See <http://www.apprenticeships.org.uk/Employers/Training-and-Funding.aspx> .

²⁹See http://www.hrsdc.gc.ca/eng/workplaceskills/trades_apprenticeship/index.shtml

apprenticeship programs leading to basic qualification, and \$4,000 for apprenticeship programs leading to more advanced qualification.³⁰ In Germany, in contrast, subsidies are paid to neither apprentices nor apprenticeship firms (see Brunello et al. 2007).

Consider here the subsidy of apprenticeship wages, which we denote by S . The subsidy does not change the firm's maximization problem (given by equations (4) and (3) under no commitment and commitment, respectively) in any way. Hence, firms that cannot commit to training provision will continue to offer the sub-optimal training intensity τ^{NC} to workers who receive the subsidy. Hence, a subsidy does not solve the fundamental problem of low training quality under no commitment.³¹

5 Conclusion

In this paper, we analyze a possible market failure in the firm-based apprenticeship training market: the non-verifiability of firm-based training. Since training takes place within firms, it may not be easily verifiable by a third party. Moreover, training may simply be too complex to be specified in a contract in a way that is legally enforceable. As a consequence, training firms may be able to renege on their training promises without getting punished. We refer to the ability of firms to credibly assure workers that they will not renege on their training promises and deliver the promised training intensity as *commitment to training provision*. Our hypothesis is that apprenticeship training schemes are more successful in countries like Germany than in Anglo-Saxon countries like the U.K. because more firms are able to commit to training provision in Germany than in Anglo-Saxon countries. We further hypothesize that this may be due to a well-structured regulatory framework and monitoring institutions that exist in Germany but are largely absent in Anglo-Saxon countries.

We first present a model of firm-provided training and wage compression, and analyze it under

³⁰See Knight and Mlotkowski (2009).

³¹An apprenticeship wage subsidy may, however, induce workers to enrol in apprenticeship programs if there are fixed costs of training. With fixed costs, workers accept training offers under no commitment to training provision only if the sum of their utility with training and the subsidy S , $U(\tau^{NC}) + S$, exceeds their utility without training, $U(0)$. If $U(\tau^{NC}) < U(0)$, a subsidy may make workers choose training over no training.

commitment and no commitment to training provision. We show that the training intensity under no commitment to training provision is unambiguously lower than the socially optimal training intensity under commitment to training provision. The training intensity under no commitment is the closer to the socially optimal level the larger wage compression in the economy.

We then quantify how much lower the training intensities are in the no commitment case than in the commitment case, by calibrating our model to match levels of wage compression observed in the data. We find that, according to our baseline estimates, training intensities under no commitment are only about 28% of those under commitment to training provision. Hence, the inability of firms to commit to training provision results in substantially lower training intensities—which will result in lower enrolment rates into apprenticeship training if there are fixed costs of training and workers are ex ante heterogenous. We finally provide several pieces of evidence that are consistent with the hypothesis that commitment to training provision is more problematic in Anglo-Saxon countries, in particular in the U.K., than in Germany, and argue that this may be linked to differences in the regulation of apprenticeship training between these countries.

Which policy implications can be drawn from our findings? We believe that countries that would like to expand firm-based apprenticeship training should pay careful attention that apprenticeship contracts are enforceable and that firms are able to commit to training provision, possibly through stricter regulation of the apprenticeship system, like the monitoring of training firms, and examination of training achievements by external institutions. Subsidizing apprenticeship programs, in contrast, may not be the most effective way of expanding apprenticeship training, as it does not address the commitment problem.

References

- [1] Acemoglu, D. and J.-S. Pischke (1998), "Why do Firms train? Theory and Evidence", *Quarterly Journal of Economics*, 112, 79-119.
- [2] Acemoglu, D. and S. Pischke (1999b), "The Structure of Wages and Investment in General Training", *Journal of Political Economy*, 107, 539-573.
- [3] Acemoglu, D. and J.-S. Pischke (2000), "Certification of Training and Training Outcomes", *European Economic Review*, 44, 917-927.
- [4] Acemoglu, D. and S. Pischke, (2003), "Minimum Wages and On-the-Job Training", *Research in Labor Economics*, 22, 159-202.
- [5] Adult Learning Inspectorate (2006), "The Final Annual Report of the Chief Inspector, 2005-2006", Coventry.
- [6] Autor, D. (2001), "Why Do Temporary Help Firms Help to Provide Free General Skills Training?" *Quarterly Journal of Economics*, 116, 1409-1448.
- [7] Becker, G. (1962), "Investment in Human Capital: A Theoretical Analysis", *Journal of Political Economy*, 70, 9-49.
- [8] Beicht, U., Walden, G., and H. Herget. (2004), "Cost and Benefits of In-Company Vocational Education and Training in Germany", Federal Institute for Vocational Training, Bielefeld.
- [9] BIBB Report (2007), "More training companies - More training places - Less quality?" March, Bonn.
- [10] D. G. Blanchflower and L. M. Lynch, 1994. "Training at Work: A Comparison of U.S. and British Youths," NBER Chapters, in: Training and the Private Sector, pages 233-260 National Bureau of Economic Research, Inc.
- [11] Bowers, N., A.Sommet and L.Bardone (1999), "Giving young people a good start: the experience of OECD countries", in: OECD, "Preparing Youth for the Twenty-first Century: the Transition from Education to the Labour Market", 7-84, Paris.
- [12] Brunello, Garibaldi and Wasmer (2007), Education and Training in Europe, Oxford University Press, Oxford, Chapter. 13
- [13] Conti, G. (2005), "Training, Productivity and Wages in Italy", *Labour Economics*, 12, 557-576.
- [14] Dearden, L., Reed, H. and J. Van Reenen (2006), "The Impact of Training on Productivity and Wages: Evidence from British Panel Data", *Oxford Bulletin of Economics and Statistics*, 68, 397-421
- [15] Department for Innovation, Universities, and Skills (2008), "Apprenticeship Pay: 2007 Survey of Earnings by Sector", DIUS Research Report 08 05.
- [16] Dustmann, C. and U. Schoenberg (2009), "Training and Unions", *The Review of Economics and Statistics*, 91(2): 363-376

- [17] Euwals, R. and R. Winkelmann (2004), "Training Intensity and First Outcomes of Apprenticeship Graduates", *International Journal of Manpower*, 25, 447-462.
- [18] Fitzenberger, B. and A. Kunze (2005), "Vocational Training and Gender: Wages and Occupational Mobility among Young Workers", *Oxford Review of Economic Policy*, 21, 392-415.
- [19] Fitzenberger, B. and A. Spitz-Öner (2004), "Die Anatomie des Berufswechsels: Eine empirische Bestandsaufnahme auf Basis der BiBB/IAB-Daten 1998/99", in: W. Franz, H.J. Ramser und M. Stadler (eds.), *Bildung, Wirtschaftswissenschaftliches Seminar in Ottobeuren*, Bd. 33, Tübingen, 29-54.
- [20] Franz, W. and D. Soskice (1995), "The German Apprenticeship System", in F. Buttler, W. Franz, R. Schettkat, and D. Soskice (eds.), *Institutional frameworks and labour market performance*, London: Routledge.
- [21] Hanhart, S. and S. Bossio (1998), "Costs and benefits of dual apprenticeship: Lessons from the Swiss system", *International Labour Review*, 137, 483-500.
- [22] Harhoff, D. and T.J. Kane (1997), "Is the German apprenticeship system a panacea for the U.S. labour market?" *Journal of Population Economics*, 10, 171-196.
- [23] Hoeckel, S. F. and W. N. Grubb (2009), "Learning for Jobs, OECD Reviews of Vocational Education and Training, Switzerland", *OECD reviews of Vocational Education and Training*, OECD.
- [24] Hogarth, T. and C. Hasluck (2003), "Net Cost of Modern Apprenticeship Training to Employers", Research Report RR418, Department for Education and Skills.
- [25] House of Lords, Select Committee on Economic Affairs, 5th report (2007), "Apprenticeship: A Key Route to Skill", Published by the Authority of the House of Lords, London: The Stationery Office Limited.
- [26] Konings, J. and S. Vanormelingen (2009), "The Impact of Training on Productivity and Wages: Firm Level Evidence", C.E.P.R. Discussion Papers, CEPR Discussion Papers: 7473.
- [27] Lewis, P. and P. Ryan (2007), "The Role of External Inspection in the U.K. Training Market", mimeo, King's College London.
- [28] Lynch, L. M. (1993), "The Economics of Youth Training in the United States", *The Economic Journal*, 103, 1292-1302.
- [29] Lynch, L. M. (1994), "Introduction", in: L.M. Lynch (editor), *Training and the Private Sector*, NBER comparative labor market series, University of Chicago Press, Chicago, 233-260.
- [30] MacLeod, W. B. and J.M. Malcomson (1993), "Investments, Holdup, and the Form of Market Contracts", *American Economic Review*, 83, 811-837.
- [31] Malcomson, J.M., Maw, J.W., and B. McCormick (2003), "General Training by Firms, Apprentice Contracts, and Public Policy", *European Economic Review*, 47, 197-227.
- [32] Muench, J. (1992), "Das Berufsbildungssystem in der Bundesrepublik Deutschland", European Centre for the Development of Vocational Training (CEDEFOP), Berlin.

- [33] National Centre for Vocational Education Research (2009), "Australian vocational education and training statistics: Apprentices and trainees 2008 - Annual", Adelaide, Australia.
- [34] Parey, M. (2009), "Vocational Schooling versus Apprenticeship Training - Evidence from Vacancy Data ", mimeo, University College London.
- [35] Prendergast, C. (1993), "The Role of Promotion in Inducing Specific Human Capital Acquisition", *Quarterly Journal of Economics*, 108, 523-534.
- [36] Ryan, P. (2000), "The institutional requirements of apprenticeship: evidence from smaller EU countries", *International Journal of Training and Development*, 4, 42-65.
- [37] Ryan, P. (2001a), "The School-to-Work Transition: A Cross-National Perspective", *Journal of Economic Literature*, 39, 34-92.
- [38] Ryan, P. (2001b), "Apprenticeship in Britain: Tradition and Innovation", in: Deissinger, T. (ed.), *Berufliche Bildung zwischen nationaler Tradition und globaler Entwicklung*. Nomos, Baden-Baden, 133-157.
- [39] Ryan, P., H. Gospel, and P. Lewis (2007), "Large employers and apprenticeship training in Britain", *British Journal of Industrial Relations*, 45, 127-53.
- [40] Ryan, P. and L. Unwin (2001), "Apprenticeship in the British "training market"", *National Institute Economic Review*, 178, 99-114.
- [41] Schönberg, U. (2007), "Testing for Asymmetric Employer Learning", *Journal of Labor Economics*, 25, 651-692.
- [42] Schofield, K. (2000), "The Good, the Bad, and the Ugly: Perspectives from Three States on the Quality of Australia's Apprenticeship and Traineeship System", Coffs Harbour, 4-7 July.
- [43] Sharpe, A. (1999), "Apprenticeship in Canada: A Training System Under Siege?" Report prepared for the CLFDB National Apprenticeship Committee.
- [44] Sharpe, A. and J. Gibson (2005), "The Apprenticeship System in Canada: Trends and Issues", CLSL Research Report 2005-04, Ottawa, Ontario.
- [45] Stevens, M. (1994a), "A Theoretical Model of On-the-Job Training with Imperfect Competition", *Oxford Economic Papers*, 46, 537-562.
- [46] Stevens, M. (1994b), "An Investment Model for the Supply of Training by Employers", *Economic Journal*, 104, 556-570.
- [47] Steedman, H., H. Gospel, P. Ryan (1998), "Apprenticeship: A Strategy For Growth", CEP discussion paper No.CEPSP11, London.
- [48] Steedman, H. (2008), "Time to look again at Apprentice Pay? Getting Cost-Sharing right", *SSDA Catalyst*, Issue 5.
- [49] Soskice, D. (1994), "Reconciling Markets and Institutions: the German Apprenticeship System", in L.M. Lynch (ed.) *Training and the private sector, International comparisons*. NBER Series in Comparative labour Markets. Chicago: University of Chicago Press, 1-24.

- [50] Tritscher-Archan, S. (2009), "VET in Europa –Country Report Austria", VET policy report: Austria 2008, Vienna : IBW.
- [51] United States Government Accountability Office (2005), "Registered Apprenticeship Programs - Labor Can Better Use Data to Target Oversight", Report to Congressional Requesters GAO-05-886, August 2005, Washington, D.C.
- [52] Wachter, T. and S. Bender (2006), "In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers", *American Economic Review*, 96, 1679-1705.
- [53] Wolf, Alison (2011), "Review of Vocational Education – The Wolf Report" UK Department for Education, London
- [54] Winkelmann, R. (1996), "Employment Prospects and Skill Acquisition of Apprenticeship-Trained Workers in Germany", *Industrial and Labor Relations Review*, 49, 658-672
- [55] Winkelmann, R. (1997), "How young workers get their training: A survey of Germany versus the United States", *Journal of Population Economics*, 10, 159–170.

A Firm-Specific Training, Asymmetric Information, and Wage Compression

In our model, both firm-specific training and asymmetric information lead to wage compression. To see this, note that $\frac{\partial \Pi}{\partial \tau}$ satisfies (under both no commitment and commitment):³²

$$\frac{\partial \Pi}{\partial \tau} = p(1 - G(v - w_L))(h'(\tau)\eta_L - \frac{\partial v}{\partial \tau}) + (1 - p)(1 - G(v - w_H))(h'(\tau)\eta_H - \frac{\partial v}{\partial \tau}).$$

Rearranging this equation, profits are increasing in training ($\frac{\partial \Pi}{\partial \tau} > 0$) if

$$\frac{\partial v}{\partial \tau} < \frac{p(1 - G(v - w_L))h'(\tau)\eta_L + (1 - p)(1 - G(v - w_H))h'(\tau)\eta_H}{p(1 - G(v - w_L)) + (1 - p)(1 - G(v - w_H))} = E[h'(\tau)\eta | \text{stay}].$$

This condition says that wages are compressed if training results in a larger increase in the productivity than in the outside option v , for *those workers who stay with the training firm*. It holds in equilibrium for two reasons. First, due to firm-specific training, an increase in training by one unit raises productivity at incumbent firms by $h'(\tau)\eta$, but at outside firms only by $\alpha h'(\tau)\eta$. Second, due to asymmetric information, more able workers are more likely to stay with the training firm and training raises the productivity of high-ability workers more than that of low-ability workers.

B Results: Robustness Checks

In this section, we argue that our key conclusion that the lack of commitment to training provision results in substantially lower training intensities is robust to alternative ways of computing the transferability of human capital, α , the productivity of high-ability workers, η_H , and the dispersion of utility shocks, b .

Our results in Table 3 are based on an estimate for α that takes into account only the loss of firm-specific skills, but not that of occupation-specific skills. This was motivated by only firm-specificity, and not occupation-specificity leading to wage compression. However, the occupation-specific component of apprenticeship training is quite high (around 35%, Table 2), and ignoring the loss of occupation-specific skills could lead us to overstate the importance of asymmetric information when explaining the wage disadvantage of movers. In Panel A of Table A.1, we therefore report the model parameter values for the productivity of high-ability workers, η_H , the scale parameter of the logistic distribution of utility shocks, b , the ratio between the training intensities under no commitment and commitment when we use an alternative estimate for α that takes into account the loss of occupation- as well as firm-specific skills. This estimate is obtained by comparing the applicability of skills for workers who are no longer employed at their training firm (some of which may have switched occupations) with that of workers who are (see column (1)). This has only a small impact on our findings: for all firms, the ratio between the training intensities under no commitment and commitment is 14.4% (compared to 28.3% in our baseline estimate) when the model moments are

³²Here, we have used the first-order condition for w_L and w_H (equation (1)).

obtained under the assumption of commitment to training provision, and 40.2% (compared to 43.6% in our baseline estimate) when they are obtained under the opposite assumption of no commitment to training provision.

An additional reason why movers earn different wages from stayers that is not captured by our model is job search: movers may be worse or better matched with their new firm on average than stayers. If this were the only difference for the wage differential, we would expect the stayer-mover wage differential to gradually disappear over time, as after some time in the labor market, movers and stayers should have accumulated the same amount of search capital. In Panel B of Table A.1, we report the stayer-mover wage differential 5 years after apprenticeship training, the corresponding model parameter values for η_H , and b , as well as the ratio between the training intensities under no commitment and commitment based on the stayer-mover wage differential 5 years after training. Our overall conclusions are again unchanged.

Yet another reason for why movers earn different wages from stayers is, as stressed by von Wachter and Bender (2006), worker sorting. Suppose that high-ability workers sort into firms that offer apprenticeship programs of higher quality, and that the separation rate after training is lower in these firms. Such a sorting model also predicts, just like asymmetric information, a lower ability of movers, but does not necessarily lead to wage compression. To address this concern, we have compared wages of workers who have been trained in the same firm, and among whom some stay and others leave the training firm, which corresponds to a fixed firm effects wage regression. The results are now based on a sample that heavily select on large firms. We find that once we control for observable firm and worker characteristics, including fixed training firm effects reduces the stayer-mover wage differential after apprenticeship training only slightly, from 0.095 to 0.090. We thus conclude that worker sorting is unlikely to bias our estimates for the degree of wage compression due to asymmetric information, and thus the decline in the training intensity due to no commitment to training provision.

Figure 1: Wage Compression and Training Intensities under Commitment and No Commitment to Training Provision



Note: In Panel A, we compare the training intensities under commitment and no commitment to training provision predicted by our model for increasing levels of wage compression due to firm-specific training, and assume that there is no informational asymmetry between training and outside firms. In Panel B, we compare the training intensities under commitment and no commitment to training provision for increasing levels of wage compression due to asymmetric information between training and outside firms, and assume that training is completely general. In the figures, we assume that the share of low-ability workers is 0.5, and that the productivity of low-ability workers is 1. The dispersion of the utility shock is set to 0.4 (which is close to our baseline estimate; see Table 3). In Panel A, we set the productivity of high-ability workers to 2.75 (which is again close to our baseline estimate, see Table 3). See Section 3.1 or the functional forms of the production and cost functions, as well as the chosen form of the distribution of utility shocks.

Table 1: The Transferability of Apprenticeship Skills

<u>Panel A: The Specificity of Apprenticeship Skills</u>					
	<i>all</i>	<i>by size of apprenticeship firm</i>			
		<i><10</i>	<i>10-49</i>	<i>50-499</i>	<i>>499</i>
Firm Mover, Occ. Stayer	-0.044 (0.004)**	-0.064 0.006	-0.045 0.006	-0.045 0.007	-0.047 0.01
Firm Stayer, Occ. Mover	-0.088 (0.007)**	-0.078 0.002	-0.051 0.014	-0.06 0.011	-0.115 0.012
Firm Mover, Occ. Mover	-0.346 (0.004)**	-0.408 0.008	-0.343 0.007	-0.303 0.008	-0.316 0.01
<u>Panel B: Share of Training that is General (α)</u>					
	<i>all</i>	<i>by size of apprenticeship firm</i>			
		<i><10</i>	<i>10-49</i>	<i>50-499</i>	<i>>499</i>
Firm Mover, Occ. Stayer	0.956	0.936	0.955	0.955	0.953

Notes: In Panel A, we report estimates for the specificity of apprenticeship training by firm size, where we distinguish between firm movers and occupation stayers, firm stayers and occupation movers, as well as firm and occupation movers. The base category are workers who are still employed at their apprenticeship firm and apprenticeship occupation. The dependent variable is the proportion of skills obtained during apprenticeship training that is applicable at the current job. We control for the year the apprenticeship ended, experience, experience squared, size of the apprenticeship firm (4 categories), apprenticeship occupation (19 categories), age at the end of the apprenticeship, and type of high school. Robust standard errors in parentheses. In Panel B, we report our estimate for the degree of transferability of apprenticeship training, α .

Source: German Qualification and Career Survey, waves 1 to 3, West German men who completed an apprenticeship. N=24,828.

Table 2: Data Moments: Quit Rate and the Mover-Stayer Wage Differential

	all firms	<i>by size of apprenticeship firm</i>			
		<10	10-49	50-499	>499
quit rate	35.86%	44.74%	38.27%	32.30%	25.51%
stayer-mover wage differential	0.036 (0.003)	-0.022 (0.005)	0.027 (0.005)	0.058 (0.006)	0.127 (0.007)
N	43,996				

Notes: We first report the share of workers who leave the training firm after training completion. We then display the wage advantage of stayers after apprenticeship training. Our regressions control for citizenship, age at the beginning of apprenticeship, apprenticeship duration, high school degree (Abitur), 16 apprenticeship industry dummies, and the size of apprenticeship firm (4 dummies). We first show results pooled for all firms, and then separately by the size of the apprenticeship firm. Robust standard errors in parentheses.

Source: IAB Social Security Records, men in the West German labor market who completed an apprenticeship between 1980 and 1998.

Table 3: Data Moments and the Productivity of High-Ability Workers and the Scale Parameter of the Distribution of Utility Shocks (Baseline)

<u>Panel A: Model moments are computed under commitment</u>			
	<i>dispersion</i>	<i>productivity</i>	<i>ratio training</i>
	<i>utility shocks (b)</i>	<i>high-ability (η_H)</i>	<i>NoCom/Com</i>
all firms	0.45	2.46	28.3%
small firms (1-9)	0.56	2.14	13.5%
medium-sized firms (10-49)	0.47	2.39	24.8%
medium-sized firms (50-499)	0.41	2.71	38.5%
large firms (>499)	0.38	5.10	56.1%
<u>Panel B: Model moments are computed under no commitment</u>			
	<i>dispersion</i>	<i>productivity</i>	<i>ratio training</i>
	<i>utility shocks (b)</i>	<i>high-ability (η_H)</i>	<i>NoCom/Com</i>
all firms	0.44	3.01	43.6%
small firms (1-9)	0.55	2.81	34.8%
medium-sized firms (10-49)	0.46	2.96	41.5%
medium-sized firms (50-499)	0.40	3.19	47.8%
large firms (>499)	0.36	5.18	66.8%

Notes: The table first displays the model parameter values for the productivity of high-ability workers (η_H) and the scale parameter of the distribution of utility shocks (b). These parameter values are chosen by minimizing the squared distance between the quit rate and stayer-mover wage differential observed in the data (see Table 2) and their model counterparts. We obtain the model moments first under the assumption that all firms commit to training provision (Com, Panel A), and then under the opposite assumption that no firm commits to training provision (NoCom, Panel B). See Table 1, Panel B for our estimates for the degree of transferability of apprenticeship training, α . For each set of parameter values, we display the ratio of the training intensities under no commitment and commitment to training provision predicted by our model.

Table 4: Apprenticeship Training in the U.K. and (West) Germany

		U.K.	Germany
(1) <u>Dropout Rates</u>			
	2002	76%	23.5%
	2005	47%	19.4%
(2) <u>Ratio Apprenticeship Wages and Wages of Inexperienced Unskilled Workers</u>			
	(Median, 2002-2004)	0.83	0.43
(3) <u>Productivity of Apprentices Minus Apprenticeship Wage:</u>			
	All Sectors (3-year apprenticeship occupations)		€: -324.07
	Construction	£: 2800	€: -558

Notes: The first row compares the shares of apprentices who do not complete the apprenticeship program in West Germany and the U.K. The numbers for the U.K. are taken from Lewis and Ryan (2007), and the numbers for Germany from the Federal Education Report (Bundesbildungsbericht) for the years 2006 and 2007. In the second row, we show the ratio between the median wage of apprentices and that of inexperienced unskilled employees, which we define as employees below the age of 20 who left full-time education at age 16 and are employed full-time. The numbers are own calculations based on the British Labour Force Survey and the IAB Employment Sample. In the third row, we report the difference between the productivity of apprentices and the apprenticeship wage in the construction sector. The numbers for the U.K. (in pounds) are taken from Hogarth and Hasluck (2003). The numbers for Germany (in Euros) are own calculations based on the German Apprenticeship Survey.

Appendix B: Robustness Checks

Table A.1: Training under Commitment and No Commitment to Training Provision: Occupation-Specific Human Capital and Job Search

<i>Panel A: Occupation-Specificity</i>									
	<u>Data Moments</u>			<u>Corresponding Model Parameter Values</u>				<u>Ratio Training</u>	
	α	q^D	SM^D	<i>Com</i>		<i>NoCom</i>		<i>Com</i>	<i>NoCom</i>
all firms	0.808	0.36	0.036	2.13	0.51	2.64	0.47	14.4%	40.2%
small firms (1-9)	0.764	0.45	-0.022	1.94	0.65	2.59	0.62	16.4%	34.1%
medium-sized firms (10-49)	0.807	0.38	0.027	2.11	0.54	2.68	0.51	21.4%	38.6%
medium-sized firms (50-499)	0.827	0.32	0.058	2.27	0.45	2.78	0.43	29.2%	43.2%
large firms (>499)	0.824	0.26	0.127	3.26	0.41	3.53	0.38	53.1%	56.3%

<i>Panel B: Job Search (and Firm-Specificity)</i>									
	<u>Data Moments</u>			<u>Corresponding Model Parameter Values</u>				<u>Ratio Training</u>	
	α	q^D	SM^D	<i>Com</i>		<i>NoCom</i>		<i>Com</i>	<i>NoCom</i>
all firms	0.956	0.36	0.061	2.69	0.46	3.2	0.45	35.6%	46.8%
small firms (1-9)	0.936	0.45	0.032	2.41	0.59	3.05	0.58	19.8%	39.3%
medium-sized firms (10-49)	0.955	0.38	0.061	2.67	0.49	3.2	0.48	34.1%	43.9%
medium-sized firms (50-499)	0.955	0.32	0.063	2.76	0.41	3.24	0.4	39.6%	48.6%
large firms (>499)	0.953	0.26	0.108	4.37	0.37	4.35	0.35	61.7%	61.1%

Notes: In Panel A, we first report alternative estimates for the degree of transferability of skills (α) which take into account the loss of occupation-specific skills. The quit rate (q^D), and the stayer-mover wage differential (SM^D) are the same as in Table 2. In Panel B, we report alternative estimates for the stayer-mover wage differential that take into account search capital. The estimates for q^D and α are the same as in Tables 1 and 2. We then display the corresponding model parameter values for the productivity of high-ability workers (η_H) and the scale parameter of the distribution of utility shocks (b). These parameter values are chosen by minimizing the squared distance between the data and the model moments. We obtain the model counterparts first under the assumption that all firms commit to training provision (*Com*), and then under the opposite assumption that no firm commits to training provision (*NoCom*). Finally, we display the ratio between the training intensities under commitment and no commitment to training provision predicted by our model.