

IZA DP No. 6037

**When the Early Bird Catches the Worm:
The Impact of Training in Retail**

Christiane Hinerasky
René Fahr

October 2011

When the Early Bird Catches the Worm: The Impact of Training in Retail

Christiane Hinerasky

University of Paderborn

René Fahr

*University of Paderborn
and IZA*

Discussion Paper No. 6037
October 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

When the Early Bird Catches the Worm: The Impact of Training in Retail*

We econometrically evaluate the performance effects of a six month e-learning programme in a large retail chain with monthly data on sales revenue, for four years using panel regressions. Participants in early cohorts show positive performance effects during training periods that depreciate afterwards. We conclude that offering training on a voluntary basis leads participants with the highest expected idiosyncratic gains and the highest talent to self-select into early participation. As performance effects already unfold during training, our findings put forward the importance of continuous training with close coaching unlike single training incidences.

JEL Classification: C31, C33, J24, M53

Keywords: evaluation, company training, e-learning, average treatment effect, average treatment effect on the treated, selection effect, continuous learning, continuous vocational training

Corresponding author:

René Fahr
University of Paderborn
Warburger Straße 100
33098 Paderborn
Germany
E-mail: rene.fahr@wiwi.uni-paderborn.de

* We thank Patrick Kampkötter, Ulrich Kaiser, seminar participants at the University of Paderborn, and conference participants at the Colloquium on Personnel 2009 in Wien, AKempor 2010 in Bochum and EEA 2011 in Oslo for very helpful comments on earlier drafts of this paper. The data used in this study are proprietary.

Introduction

For various reasons, investments in skills are becoming more varied in terms of formalization, job-relatedness, as well as the time and place of training. Continuous vocational training - or lifelong learning - is becoming more important for individual careers, firm competitiveness, and macroeconomic performance alike. This calls for continuous informal learning of employees on the job and, at the same time, for more formal training to acquire the knowledge that is required for the diffusion and efficient utilization of new technologies. Despite the importance and high expenditures on company training, the literature on effects of work related training is still rare and inconclusive.

Two issues are of particular importance when studying the performance effects of work related training: Consider at first selection effects, as some employees benefit more from training than others and they or the employer may take this knowledge into account when deciding on training enrollment. A large empirical literature suggested sophisticated econometric methods to account for these selection effects (Angrist and Krueger 1999). A second issue is related to the most fundamental question of how training basically works and is concerned with the timing of productivity effects of trainings. A stylized fact in this respect is that productivity is commonly lower during training periods. Reasons for this can be twofold and attributed to the fact that, firstly, working hours are reduced when workers spend time in the classroom rather than the workplace and, secondly, training costs simply decrease returns (Aragón-Sánchez, Barba-Aragón and Sanz-Valle 2003, Dearden, Reed and Van Reenen 2006). Based on the same reasons, an increase in returns can only be expected only with a timely lag (Aragón-Sánchez et al 2003, Dearden, Reed, Van Reenen 2006). Our data set has several key features that render our data particularly suitable for a detailed econometric analysis of the question on who benefits from company training and how. We have data from one company with more than 500 stores offering a nearly identical product portfolio to all customers. Therefore, training effects of the branch manager as a key employee are measured on the organizational level rather than the firm or the individual level. By using information on one company with detailed information on the training programme, our study is close to the concept of an “econometric case study” (e.g. Bartel 2000). We have a precise objective performance measure (sales revenue) and managers are observed over four consecutive years with monthly performance data. In addition, the panel data set has a quasi-experimental structure as not all managers take part in the training programme at the same point in time. In the present study, we will be able to provide evidence on the following questions, particularly because training is offered ostensibly on a

voluntary basis, weekly time investments in the 6 month training programme are low (about 10 hours during leisure time) and the performance measure is independent from the direct costs of training: 1. we evaluate the performance effects of a training programme by accounting for selection into training in various ways. In particular we will be able to assess the direct effect of training on the individual by comparing training before and after taking part in the programme (the average treatment effect on the treated (ATT) in the terminology of the programme evaluation literature) and the training effect of a randomly selected individual (average treatment effects (ATE) relying on a simple comparison of participants with non-participants). 2. We will study those enrolling early into training in more detail, assessing whether training on a voluntary basis can reveal heterogeneity in the performance effects. 3. As the e-learning training lasts 6 months, we include indicators for the periods during and after participation to learn more about the transfer and sustainability of training. This case study, in particular, allows assessing the importance of continuous training compared to training en bloc.

In the next section we describe the company with its training program in more detail and provide a further discussion of the potential effects of training on firm performance. Section 3 describes several features of our data. The empirical strategy is explained in section 4 and results are presented in section 5. Section 6 concludes.

Training in Retail

The puzzling question in the empirical firm-training literature is how to reconcile the contradicting positions of finding either no or positive training effects. In line with Bartel (1989) and Black and Lynch (1996), Barret and O'Connell (2001) find no productivity effect for firm-“specific” training content. Many training incidences serve the purpose of adjusting to a new job or retraining which suggests that training is rather aiming at preserving productivity than augmenting it (Fahr, Hinerasky and Simons 2010). This might be particularly true if workers' skills have to keep up with the machinery's technological progress as is typically found in manufacturing. With this in mind, we consider retail a good environment to study the returns to training, as adaption to new technologies are, except for electronics retailers, of minor importance. Still, several studies in different industries show a positive and significant impact of training. These are, among others, Bishop (1990), Holzer et al (1993), Bartel (1995), Delaney and Huselid (1996), Barron, Berger, Black (1999), Dearden, Reed, Van Reenen (2006), Liu and Batt (2007), and Morrow, Jarret, Rupinsky (1997). The latter, evaluating a pharmaceutical company's training programmes with a

case study framework similar to ours, find technical and sales training to have an higher impact than managerial training.¹ Many of these studies use cross-sectional data which do not control for unobservable time-invariant characteristics of trainees or allow for an application of models that control for non-random selection. Several studies have shown that econometric methods may provide false estimates compared to experimental data, especially when selection effects are present (LaLonde 1986, Ashenfelter and Card 1985, Dehejia and Wahba 1999). We therefore estimate and account for a pre-training mark up among the training group, if present.

As mentioned in the introduction, the time structure of training effects is rather unexplored. Lower productivity of employees during training relies in parts on the way how training is offered. Training forms usually studied are temporary off-the-job training spells (Bartel 1995, Dearden, Reed, Van Reenen 2000, and Zwick 2006). Naturally, productivity increases, due to higher human capital acquired in off-the-job training, can only unfold after finishing the training. In our case, however, e-learning training is supplemental to regular working hours for a period of 6 months. Accordingly, we might be able to observe the impact of training even during the training phase. Furthermore, participants receive regular online support from varying supervisors, when discussing different topics that the programme has covered in previous weeks. Support is given in chat rooms where an almost natural communication can take place. The psychological literature defines this form of coaching as a process to increase performance levels by encouragement and motivation (Burdett 1998, Evered and Selman 1989, Ellinger and Bostrom 1999). Coaching behaviour is found to be positively associated with individual performance by Ellinger, Ellinger and Keller (2003) and Agarwal, Angst and Mangi (2009). Coaching as a form of managerial support can, hence, improve participant's work performance even in the short-term by its direct stimulation.

Only very few studies investigated the sustainability of training, despite its relevance in keeping the productivity of the workforce. The reason is again that only very few studies can measure productivity effects months or even years after the investment has occurred. One example is Zwick (2005) who reported productivity effects lasting several years after the training incidence. Zwick (2005) finds a long-lasting increase in productivity for formal external training forms even two years after training. Formal internal training, participation at seminars and talks, however, have no lasting positive

¹ There is also a large amount of literature which looks at wage returns to company training. Loewenstein and Spletzer (1998) give an overview for the U.S. and Pischke (2001) for Germany, an excellent survey of the literature is provided by Bassanini et al (2006).

impact, some training forms even exhibit a negative long-term effect on productivity (Zwick 2005). We too will provide a detailed analysis of the long-lasting productivity effects after training using a monthly decomposition.

The retail company under investigation sells products to private customers in the lower and middle price segment in an industry with tough price competition. After merging with a similar sized competitor, the company reassessed its market positioning away from a former focus on low prices to a stronger orientation towards quality and service. The training programme was introduced in 2005, as part of this change in the marketing strategy, and was especially designed for store managers, who are directly responsible for the customer contact.² The major aim of the programme was to teach store managers business and marketing-related knowledge, which they in turn should communicate to their employees and thereby improve the company's public image. While offering a wide range of different products, the product portfolio is very similar across all 500 stores. Built as an e-learning tool with text, films, and audio tutorials, the programme covers topics such as time management, purchasing, sales, organization, leadership, marketing, law, and managerial-economics. For a period of 6 months, participants spend an average of 10 hours per week on studying and preparing courses supplemental to their regular work load. Initial training participation was voluntary. Admission to training was based on a first come, first served strategy. However, there has been an understanding that all store managers are eventually expected to participate in the training course. Within the observation period from October 2005 until December 2007, we observe 10 training courses starting in different intervals. Each course officially ends with a final exam consisting of a written and an oral test, held by the Chamber of Commerce and Industry (CCI), and is rewarded upon passing with a CCI-certificate. The written test covers all learning material quizzing 120 multiple choice questions. In the oral test, participants are faced with different day-to-day issues, e.g. complaining customers or a low sales volume, and have to deliver problem solving suggestions and their recommended course of action in front of an examination board. 99 percent of all participants have passed the exam obtaining a grade of 4(D) or better, though few needed a second try. Finishing the course is rewarded with a monthly pay raise of 100 Euro, those who passed with distinction (grades 1(A) and 2(B)), were on top of that refunded the entire training costs of about 240 Euro. Within the first three years (2005 to 2007), 165 store managers took part in

² The information in this section originates from personal conversation between the authors and the head of human resource development as well as personal conversation between one of the authors and various company representatives during a 2-week stay in the headquarters of the company.

groups between 20 and 32 participants (\bar{x} 26.1). Within the observation period, a total of 10 groups started the 6-months programme in October 2005, March, June, September and November 2006, and February, March, April, September and October 2007.

Data

The data for this study is taken from the company's personnel records and financial data for the time between January 2004 and December 2007. 340 out of 500 subsidiaries were managed by one supervisor during the entire 4-year period and were therefore selected for observation. Among these, 165 store managers took part in the training programme while 175 store managers did not. The average training participant (non-participant) is 41 (44) years old, has 3.3 (3.2) full-time employees, approximately 10.6 (10.9) years of company experience, and is leading a store that has existed for 11.5 (11.5) years. Even though participation is not being determined randomly, both groups are strikingly similar and only significantly differ in their age and tenure, which we will examine further when discussing the determinants of training. Testing the a-priori differences in their pre-training sales revenues also shows no significant difference between non-participants and participants-to-be. Each store can be categorized into one of three different sales level categories, which depend on size, location, and regional purchasing power. As an objective productivity measure relative monthly sales revenues are observed for each store in all 48 months. This results in a total sample of 16,320 observations. Additional information for each store is available on the number of employees and their respective hours of work, the manager's age and tenure, the age of the store, a dummy indicating whether the store is a franchise store, as well as numerous time-invariant characteristics that further describe the store environment.³ An overview of our control variables as well as descriptive statistics can be seen in Table 1.

[Insert Table 1 about here]

We measure training participation with two dummy variables: in Training and after Training. In Training takes the value 1 if the respective store manager is taking part in training in the month under

³ Certain information (e.g. age, tenure, number of full-time employees) was not available for franchise stores.

consideration and after Training takes the value 1 for all months after the respective store manager has finished the programme.

A store manager's productivity is measured by his respective store's sales revenue $y_{i,t}$ in month t for store i . Due to privacy restrictions, monthly sales are normalized by dividing a store's current sales revenue by its sales revenue in January 2004. The normalized sales y for outlet i in January 2004, hence, equals $y_{i,Jan2004} = 100$. Normalized sales y for outlet i in month t therefore equals:

$$y_{i,t} = \left(Sales_{i,t} / Sales_{i,Jan2004} \right) * 100 \quad (1)$$

As distinct stores usually operate on different sales revenue levels according to their size or location there might be peculiarities in training participation due to better pre-programme outcomes. As each store's performance measure in the base month is automatically set to 100, we can no longer identify variation between stores as stemming from their operational result. Subsequently, selection due to pre-programme performance is eliminated and cannot be estimated from our data using $y_{i,t}$.

[Insert Figure 1 about here]

Figure 1 provides an overview of the company's overall normalized sales revenue pattern and shows that seasonal effects are very strong. Spring and autumn are peak seasons with average sales revenues of around 170 percent of the January 2004 sales levels. Sales during summer and winter regularly decrease following the same pattern each year. While the overall sales level is stable at around 150 percent for the years 2004, 2005 and 2006, it has increased substantially in 2007. To account for these patterns and to eliminate any seasonal effects, we model the seasonal structure by including dummies for each month and year as suggested by Hanssens, Parsons, Schultz (2003). Due to the low variation in sales during the years 2005 and 2006, we include a dummy for those two years and the year 2007 that accounts for higher revenue in the retail market (see Figure 1). Still, as usually a concern in time series, autocorrelation of the error term may be present and a reason for inefficient estimates. We test for first order serial correlation (AR(1)) in the idiosyncratic errors using Wooldridge's (2002) test for panel data and account for autocorrelation by calculating panel-corrected standard error estimates (PCSE) for linear

cross-sectional time-series models.⁴

Another typical measure of firm performance is given by the growth rate of sales. The store managers in the present case sell a very specific range of products in the lower and middle price segment. Due to the nature of the product, they attract only customers in a comparatively narrow local area. By the use of training, store managers might be able to increase the efficiency of doing business in the local store or attract a wider customer group with more widespread marketing campaigns. Thus, the optimal impact of training, measured by the variation in sales at store level, lies in higher sales levels up to market saturation.⁵ Yet due to natural limits in product demand in the local area, it is unlikely to expect a long-term growth in sales as an impact of training. With these arguments it should be obvious that the growth rate of sales is not the favoured variable when investigating individual performance effects. Yet, an advantage of analyzing the growth rate is to overcome the reference dependence of the performance measure ($y_{i,t}$), allowing for a better comparison of the impact of training between store managers. Given (1) we can take logs and write

$$\ln (y_{i,t}) = \ln (Sales_{i,t}) - \ln (Sales_{i,Jan2004}) + \ln(100). \quad (2)$$

As sales revenue in the base month (January 2004) is normalized and set to 100, we deduct the pre-year values and receive the growth rate of sales as the difference between the logarithm of sales in the present month t and the same month in the previous year $t - 12$, which is more conveniently expressed in equation (3). As we can see, the growth rate is independent of the base month.

$$\ln \left(\frac{y_{i,t}}{y_{i,t-12}} \right) = \ln \left(\frac{Sales_{i,t}}{Sales_{i,t-12}} \right) \quad (3)$$

When using the growth rate as dependent variable, we are able to estimate cross-sectional variation between stores and not only within stores.⁶ In addition, this special case allows the estimation of selection effects stemming from different pre-programme performance, which could not be estimated from $y_{i,t}$, but is required when learning more about the mechanisms of the selection process in place and the possible distortion of any training effects estimated from the data.

⁴ Drukker (2003) using Monte Carlo simulations showed good size and power properties of Wooldridge's test in samples of reasonable size for all common estimation procedures. Testing for autocorrelation in the idiosyncratic errors of order AR(p) is, however, only available in time-series, not yet for panel data.

⁵ Of course there are other intended training outcomes like lower turnover rates of store managers and store employees or increasing customer satisfaction which all can lead to higher sales rates, which manifest in the long run and therefore are outside our observation period.

Empirical Strategy

The particular structure of the data allows us to identify the average treatment effect (ATE), the average treatment effect on the treated (ATT), as well as the selection effect due to better pre-programme performance (SE).⁷ Figure 2 provides an illustration of the treatment effects estimated in the present study. Note that the figure serves the purpose of illustration and that sales levels are chosen arbitrarily.

[Insert Figure 2 about here]

To further formalize our strategy and explain in greater detail, we introduce some notation.⁸ Let us denote participation in the training programme by $D = 1$ and non-participation by $D = 0$. The time period before start of the training programme is denoted by t' and that after participation in the training programme by t . Because the length of training spans over a considerable time (6 months), we denote the beginning and end by t'_0 and t_0 , respectively. When denoting the sales of a training participant by y_1 and the sales without training participation by y_0 , the ATE, i.e. the effect of training for a randomly chosen store manager into participation, is given by the difference in expected sales of participants and non-participants and can be interpreted as causal treatment effect given random assignment into both groups:

$$ATE = E(y_1 | D = 1) - E(y_0 | D = 0). \quad (4)$$

As training is not only extended to every sales manager - but managers are even asked to attend - the ATE is seen as an important figure to the company in assessing the effect of its training, offered to the average store manager. However, the ATE may be biased by potential selection effects. Selection effects (SE) are found to be positive or negative, e.g. positive, when only motivated managers register for training and their estimated 'training effect' contains both a non-distinguishable training and motivation effect, which then upwardly biases estimates of the ATE. On the contrary, SE may be negative when under-performing managers register for training, but can only achieve a 'regression to the mean', meaning that their positive treatment effect cannot be detected from comparison of post-training

⁶ When transforming and de-personalizing the data, any variation between stores was eliminated, as $y_{i,t}$ does not correspond to the actual sales level but to a store's sales related to its level in the base month. Translating sales ($y_{i,t}$) into the growth rate of sales we re-introduce the variation between stores as the growth rates are independent of the base month.

⁷ Selection effects can be detected from estimating sales growth. As the growth rate in sales reveals cross-sectional variation that is not given when estimating the performance measure $y_{i,t}$.

⁸ Overviews on the microeconomic evaluation of treatment effects are provided by Caliendo and Hujer (2006) and Blundell and Costa Dias (2009).

performances in both groups. An ATE estimate will then be downwardly biased. The probit regression in Table 2 Specification 1 shows the variables which lead managers to select into the training programme.

[Insert Table 2 about here]

We find tenure to have a positive impact on training participation with a turning point after 14.17 years with the company. Various *Store Location* and *Store Environment* variables also prompt managers to rather sign up for training. More importantly, we find store managers with positive sales development prior training to have a 20% higher chance in taking part in the programme. We can therefore rule out negative selection effects and infer that the ATE estimate can provide an upper bound of the total training effect, in our data, as it may be biased by potential positive selection into training. Besides the different directions selection effects may take, they vary in their origin. A first kind of bias is selection into treatment, based on expected idiosyncratic gains. Such a bias arises when a participant before training does not differ with respect to sales levels from a non-participant, but knows from past experience, that he is a fast learner who can quickly transfer training content. A second selection effect is due to better pre-programme outcomes of the training participants, leading more able managers, who are already performing on a high sales revenue level, to rather select into the training group than less able ones. Both biases are often compound which results in difficulties in discrimination. But with the observation of participants' outcomes before, after and during training, we are able to identify a pre-programme mark-up of future training participants by comparing (2) and (1) in Figure 2 or formally by estimating on the difference in present and the preceding year's sales:

$$SE = E(y_{0,t} | D = 1) - E(y_0 | D = 0) \quad (5)$$

In addition, we will estimate the ATT which denotes the training effect for a randomly chosen participant and which is usually seen as more compelling than the ATE with respect to an estimation of the training's productive efficiency. With the data at hand, we will estimate the returns to training in terms of sales revenue with the before-after estimator as:

$$ATT^{BA} = E(y_{1,t} | D = 1) - E(y_{0,t} | D = 1) \quad (6)$$

Because training lasts 6 months with moderate study requirements of 10 hours per week, we will estimate all effects for the periods during training and after training. Estimating equation (6) corresponds to a comparison of (3) with (2) in Figure 2 ((3a) with (2) for an ATT during treatment) and

the ATE is identified as the comparison of (3) and (1) for the post-training effect and the comparison of (3a) and (1) for the during-training period effect. Our regression estimation is straightforward and is estimated with fixed effects and a generalized least-squares method (Prais-Winsten) whereby $X'_{i,t}$ contains a vector of time-varying and Z'_i a vector of time-invariant control variables:

$$y_{i,t} = \alpha_i + \lambda_t + \rho D_{i,t} + \beta_1 X'_{i,t} + \beta_2 Z'_i + \varepsilon_{i,t} \quad (7)$$

α_i is the individual-level effect and $\varepsilon_{i,t}$ is the idiosyncratic error. When there is serial correlation in the idiosyncratic error, i.e. $\varepsilon_{i,t} = \rho_i \varepsilon_{i,t-1} + u_{i,t}$ with $u_{i,t}$ being uncorrelated across observations and ρ_i being an estimator of autocorrelation between residuals, serial correlation is caused by $\rho_i \varepsilon_{i,t-1}$ and efficiency of estimates is reduced. Prais-Winsten corrects for autocorrelation between panels by computing ρ_i and estimating parameters conditional on the estimates of the autocorrelation parameters (see e.g. Greene 2008 Chapter 13).

When evaluating training courses with simultaneous start dates by comparing sales revenue before and after with OLS, or in the case of serial correlation with Prais-Winsten estimates, one is able to control for (pre-programme) characteristics which are on average different for participants compared to non-participants. A disadvantage of this method is however, that other exogenous effects might well have an impact on individual sales levels right at the moment when training starts, and then bias the estimate of the training coefficient. As we investigate 10 different training cohorts that start and end training at different points in time, the effect of a possible exogenous influence evens out by estimating the training effect for groups before, after, or right at that exogenous occurrence. Since effects are calculated for all training cohorts on average, the influence of a possible exogenous effect cancels out. Formally, the before-after estimator allows a valid estimation of the treatment effect as long as the approximation error defined by $E(y_{0t'} - y_{0t} | D=1)$ is zero (Heckman, LaLonde, Smith 1999), in the sense that the difference in ante and post-training outcomes is not affected by any cause other than training.

Another issue, which has to be considered, is strategic decision making of when managers decide to take part in the training programme. A store manager, for example, who foresees a lower sales period, e.g. due to construction close to or in front of his shop, might choose to take part in the programme during this period on account of lower daily work volume that comes with a reduced amount of occasional customers. Lower sales revenue during the training and construction period would

then be mistaken for a negative training effect. On the other hand, managers who foresee the busy seasons of spring and autumn may choose their timing of training accordingly, to be best prepared for these additional sales opportunities. Then our model could pick up on the increased sales that the manager foresaw, but which is not due to the training received. Given strategic timing decisions of managers, we should a) find higher participation rates before busy seasons and b) compressed sales during training and upwardly biased coefficients after training, if training phases coincide with busy seasons. Consider, however, that trainings are set up, once enough managers have registered to complement a course, so that individual strategic decisions may be destroyed by an insufficient amount of enrolled participants. Yet, if enough managers follow the same strategic calculus and form a cohort, the programme may be set up at the date planned. Looking at the timely allocation of trainings, we find no systematic pattern of starting months of training, which is confirmed by χ^2 tests⁹. To rule out the fact that any boost or slump effect may be misinterpreted for a positive or negative training effect, we provide robustness checks with the difference of sales and pre-year sales, i.e. the growth rate of sales, as dependent variable.

The before-after-estimator for the ATT is estimated with Prais-Winsten models as well as a fixed effects panel model allowing for serial correlation of order one. While the Prais-Winsten regression offers a better comparison to our estimates of the ATE, the fixed effects model constitutes our preferred specification since we are able to account for time invariant store manager specific effects in the decision to participate in training.¹⁰ We will provide results for the growth rate of sales as an additional check of robustness.

Empirical Results

In the following we will only discuss the training coefficients. Given the fact that the training programme lasts 6 months and participation does not prevent those enrolled in the programme to follow their regular job, we control with *in Training* for time periods during training participation and *after Training* for time periods after training. Controlling for a detailed list of individual and store specific variables we identify an average treatment effect (ATE) during training of 8.4 percentage points higher sales

⁹ χ^2 tests show, that the observed distribution of starting months does not differ from a uniform distribution across all months.

compared to the individual sales level in January 2004 and of 5.3 after training (Table 3). The difference between these effects is highly significant. For a randomly chosen store manager, training, hence, provides both a positive effect during and after training. Yet, these positive effects attenuate once estimating the average treatment effect on the treated (ATT) with simple Prais-Winsten estimates. During the training period, sales revenue is increased by 3.4 percentage points with no significant variation after participation. Estimating ATT and controlling for unobserved individual heterogeneity by fixed effects estimations, we find no effect during training and even a significant negative effect of 3.5 percent lower sales after training (Specification 3).¹¹ The difference in estimates in (2) and (3) suggests that much of the effect estimated in (2) can be explained by individual heterogeneity for which the fixed-effects model controls.

[Insert Table 3 about here]

To study into this surprising negative effect of participation we provide estimates on the growth rate of sales in Specifications (4) – (7).¹² Note that, as the growth rate of sales is independent of the base month, it offers an additional estimation of selection effects that, if present, can bias estimates on the ATE.

When inspecting average treatment effects, we observe that sales growth increases at a rate of 2.5 percent during training but remains identical to the preceding year after participation. The insignificant estimate on the selection effect in column 5 indicates that training effects estimated with the ATE are not overestimated because of selection due to higher pre-programme outcomes. However, we do not find any training effects during training when considering only the training participants and a significant negative growth rate of sales of 1.6 % in column 6 and of highly significant 4.8 % in the fixed effects specification in column 7. Taking part in training, hence, has a negative effect on sales revenue and sales growth for training participants in the after-training period. Positive effects of training estimated as average treatment effect, which is commonly done in the literature, may thus not only be overestimated

¹⁰ An alternative approach would be a cohort-wise estimation of difference in difference models. However, being left with a maximum of 30 independent observations per cohort, this approach inflates standard errors and therefore does not allow the identification of treatment effects.

¹¹ The Prais-Winsten model compares averages across participants and non-participants whereby only *average* time-invariant (selection) effects are controlled for. Individuals are not directly compared before and after. Only the fixed-effects model can control for *individual* time-invariant effects and therefore contrasts individuals before and after training.

¹² Estimates with the typical definition of sales growth $y_{i,t} = (\text{sales}_{i,t} - \text{sales}_{i,t-12}) / \text{sales}_{i,t-12}$ instead of $\text{sales}_{i,t} / \text{sales}_{i,t-12}$ as dependent variable yield similar results.

from selection of participants with higher pre-programme outcomes -which is however not the case here- but often due to the impossibility of observing the participant's progress in future periods.

These first results can be summarized with two observations. First, at least some estimations suggest that any effects of training, if observed at all, are in line with what is commonly known as a Hawthorne effect (Roethlisberger and Dickson 1949, Levitt and List 2009), i.e. we observe a productivity increase during exposure to treatment followed by a slump to even below the pre-training revenue level. Second, any positive effects of training which are found with simple estimates of ATE and ATT vanish once controlling for manager fixed effects. As we find no hints that higher pre-programme outcomes of participants entail selection effects, the results indicate that training participants do not differ from non-participants in their growth rate of sales before participation.

The data at hand offer a more profound possibility evaluating training effects, when selection may be present. In particular, we figure those who expect the highest gains from training to participate first. This assumption is supported by the variation in final grades between course cohorts as can be seen from Table 4.¹³ The average overall grades seem to be slightly better in the first half of the cohorts compared to the second half. Using Mann Whitney U-Test statistics, we tested various breakdowns of the cohorts with respect to grades which support splitting the sample after group 5. We therefore split the first half (groups 1-5) from the second half (groups 6-10) to see if their respective training results differ. A probit regression on early versus late participation also shows that both groups differ in numerous variables, especially in the value of their store and locational characteristics (Specification 2 in Table 2).¹⁴ Note that we cannot compare the productivity of different clusters of training cohorts, since sales are reported in individually normalized values. Estimates for the preferred specification with interactions for early and late cohorts are reported in Table 5.

[Insert Table 4 about here]

[Insert Table 5 about here]

Estimates for the ATE during training in early and late cohorts amount to 6.4 and 8.3 percentage points significantly higher sales revenue.¹⁵ However, the difference between both effects is not

¹³ A detailed analysis of final grades and their impact on productivity is given in Hinerasky, Fahr and Sliwka (2010).

¹⁴ For an even more detailed analysis of the timing decision we also conducted an Ordered Logit Regression with the inversed cohort number as the variable of interest to predict cohort selection and received qualitatively similar results.

¹⁵ For a full specification one would include dummies for early and late periods, as well as separate dummies for in Training and after Training periods. However, as early and late participation is only defined for training participants, it cannot be specified

significant. We find a slightly higher after training effect for early participants than in the full sample, yet late participants show no after training effect that differs from zero. When inspecting the ATT, we now find a significant increase of sales during training of 4.7 percentage point higher sales levels, compared to the reference month and of 3.8 percentage points after participation among early cohorts. Looking at late cohorts, we find no effect of training and even a significant negative effect after participation. This negative effect for the late cohort, however, has to be interpreted carefully, because of the availability of only few after training periods. Early participants differ from the full sample in a consistent significant positive training effect while trained and do not show a decrease in after-training sales. The driving forces of early training participation are therefore the participants' expected idiosyncratic gains of the training programme. Fixed effects estimates account for the latter as long as they are time invariant. It indicates that coefficients for the ATE during and after training are positive, yet measured imprecisely, as early-takers are a selected, more talented subgroup among which the training has a significantly sales-boosting effect, while being trained. While we report evidence above that our observed distribution of starting months does not suggest anything other than a uniform distribution, we provide a further robustness check with individual sales growth as dependent variable (Specifications 4-7 in Table 5) to rule out any seasonal influences on our outcome variable which come along with strategic assessment on when to take part in training. The latter would lead us to misinterpret seasonal fluctuations for training effects. We find the within-training effect to be robust and not influenced by strategic timing decisions. As the estimated selection effect is insignificant, early participants are not selected based on higher pre-programme outcomes, but in their ability to transfer and use training content in their everyday work.

To catch any non-linear time trends, which might stem from the change in company strategy by the time the first trainings took place, we follow, in a final step, a synthetic cohort approach in which later cohorts are simulated to have taken part in training at an earlier point in time. Assuming significant differences between participants and non-participants and given an approximate similarity among training-takers, this will capture and reveal any structural time effects that might be mistaken for effects of training. The synthetic cohorts show no significant ATT effects and even tend to have a weak significant negative ATE, which strengthens the early bird hypothesis. If all early cohorts are simulated to have participated at this point in time, the ATE remains positive and highly significant.

separate of training participation. We omit the dummies *in Training* and *after Training* as this would cover the training's time horizon twice, which inflates standard errors and entails a more inefficient estimation.

In summary, selection into treatment is not random, either because it depends on pre-programme potential outcomes or because it depends on programme gains. As we do not find positive selection effects in Tables 3 and 5, participations do not seem to be selected by higher pre-programme outcomes.¹⁶ Early-takers, however, seem to make the most out of training by exhibiting higher sales levels and sales growth rates during training (Table 5).

Conclusion

The institutional environment of the company and the training programme provided a neat framework to investigate performance effects of training and to study into the sources of a potential selection bias. While it is beyond the scope of the paper to analyse the transmission of training, we were able to study productivity effects of training during and after participation which are not confounded by work time reductions or direct expenses for training. Any positive effects of training, which are found with simple estimates of ATE and ATT, vanish once controlling for manager fixed effects. Importantly, we find a significant negative impact on individually normalized sales levels once controlling for individual time-invariant heterogeneity in estimates of the ATT and a negative impact on sales growth rates after the end of the six month e-learning training.

We further explored the puzzling negative impact of training by investigating whether this result might be due to the heterogeneity of training participants. Surprisingly, once restricting the sample to participants who volunteered as early-participants we observe a statistically and economically considerable impact of training on productivity during training periods. We conclude that early-takers form a subgroup among which training has its most prevalent effects. This implies for the provision of training that with a first-come, first-served strategy those with the highest talent and the highest prospects of success will self-select into early-participation. Asking all managers to participate, however, can have the opposing effect as even those managers have to invest time and energy, for whom the training or its form might not be suited. Eventually, the type of training studied in the present paper is a 6 month e-learning programme where participants study in the evening hours while working full-time. We found that any performance effects of training unfold rather during than after training. These

¹⁶ This result might be due to the fact that each manager is supposed to participate in the long run.

however, cannot be interpreted as human capital effects, since no persistent human capital has been built up. In case it had, knowledge was not used and transferred into regular work days. Training increases sales levels through mechanisms that are not covered by human capital theory. As human capital would not vanish as quickly, we figure the given effects rather to be a coaching effect, by which participants feel motivated to use the offered programme up to the highest personal and economic benefit. By that our findings indicate the importance of providing training in the form of continuous training rather than crash courses. Training content and presented strategies are transferred because participants are confronted with addressed topics every day and reminded by regular communication with their chat group mentors. This calls for finding new ways to make company training part of the everyday work life rather than condensing the learning contents in few training incidences. By that our study complements the literature which study how training content is transferred into the workplace.

References

- [1] Agarwal, Ritu, Corey M. Angst, and Massimo Magni (2009) 'The performance effects of coaching: A multilevel analysis using hierarchical linear modeling.' *The International Journal of Human Resource Management* 20(10), 2110–2134
- [2] Angrist, Joshua D., and Alan B. Krueger (1999) *Empirical Strategies in Labor Economics*
- [3] Aragón-Sánchez, Antonio, Isabel Barba-Aragón, and Raquel Sanz-Valle (2003) 'Effects of training on business results.' *International Journal of Human Resource Management* 14(6), 956–980
- [4] Ashenfelter, Orley, and David Card (1985) 'Using the longitudinal structure of earnings to estimate the effect of training programs.' *The Review of Economics and Statistics* 67(4), 648–660
- [5] Barrett, Alan, and Philip J. O'Connell (2001) 'Does training generally work? The returns to in-company training.' *Industrial and Labor Relations Review* 54(3), 647–662
- [6] Barron, John M., Mark C. Berger, and Dan A. Black (1998) 'Do workers pay for on-the-job-training.' *The Journal of Human Resources* XXXIV 2, 235–252
- [7] Bartel, Ann (1989) 'Formal employee training programs and their impact on labor productivity: Evidence from a human resources survey.' NBER Working Paper Series No. 3026
- [8] Bartel, Ann (1995) 'Training, wage growth, and job performance: Evidence form a company database.' *Journal of Labor Economics* 13(3), 401–425
- [9] Bartel, Ann (2000) 'Measuring the employer's return on investments in training: Evidence from the literature.' *Industrial Relations* 39(3), 502–524
- [10] Bassanini, Andrea, Alison Booth, Giorgio Brunello, Maria De Paola, and Edwin Leuven (2006) 'Workplace training in europe.' *Brunello, G., Garibaldi, P. and Wasmer, E (eds.): Education and Training in Europe (Ch. 8-13). Oxford University Press.*
- [11] Bishop, John (1990) 'Job performance, turnover, and wage growth.' *Journal of Labor Economics* 8(3), 363–386

- [12] Black, Sandra E., and Lisa M. Lynch (1996) 'Human-capital investments and productivity.' *The American Economic Review* 86(2), 263–267
- [13] Blundell, Richard, and Monica Costa Dias (2009) 'Alternative approaches to evaluation in empirical microeconomics.' *Journal of Human Resources* 44(3), 565–640
- [14] Burdett, John O. (1998) 'Forty things every manager should know about coaching.' *Journal of Management Development* 17(2), 142–152
- [15] Caliendo, Marco, and Reinhard Hujer (2006) 'The microeconometric estimation of treatment effects - an overview.' *Journal of the German Statistical Society / Allgemeines Statistisches Archiv* 90(1), 197–212
- [16] Dearden, Lorraine, Howard Reed, and John VanReenen (2000) 'Who gains when Workers Train? Training and Corporate Productivity in a Panel of British Industries.' *The Institute for Fiscal Studies (IFS) Working Paper 00/04*
- [17] Dearden, Lorraine, Howard Reed, and John VanReenen (2006) 'The impact of training on productivity and wages: Evidence from British panel data.' *Oxford Bulletin of Economics and Statistics* 68(4), 397–421
- [18] Dehejia, Rajeev H., and Sadek Wahba (1999) 'Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs.' *Journal of the American Statistical Association* 94(448), 1053–1062
- [19] Delaney, John T., and Mark A. Huselid (1996) 'The impact of human resource management practices on perceptions of organizational performance.' *Academy of Management Journal* 39(4), 949–969
- [20] Drukker, David M. (2003) 'Testing for serial correlation in linear panel-data models.' *The Stata Journal* 3, 168–177
- [21] Ellinger, Andrea D., Alexander E. Ellinger, and Scott B. Keller (2003) 'Supervisory coaching behavior, employee satisfaction, and warehouse employee performance: A dyadic perspective in the distribution industry.' *Human Resource Development Quarterly* 18(9), 752–771

- [22] Ellinger, Andrea D., and Robert P. Bostrom (1999) 'Managerial coaching behaviors in learning organizations.' *Journal of Management Development* 18(9), 752–771
- [23] Evered, Roger D., and James C. Selman (1989) 'Coaching and the art of management.' *Organizational Dynamics* 18, 16–32
- [24] Fahr, Rene, Christiane Hinerasky, and Sabine Simons (2010) 'Returns to company training - evidence from a new approach using quasi experimental data.' Working Paper, University of Paderborn
- [25] Greene, William (2008) *Econometric Analysis*, 6th edition ed. (Pearson Prentice Hall, New Jersey)
- [26] Hanssens, Dominique M., Leonard J. Parsons, and Randall L. Schultz (2003) *Market Response Models - Econometric and Time Series Analysis*, 2nd ed. (Kluwer Academic Publishers)
- [27] Heckman, James J., R. J. LaLonde, and J. A. Smith (1999) 'The economics and econometrics of active labor market programs.' *O. Ashenfelter and D. Card, Chapter 31, Handbook of Labor Economics* IV, 1865–2073
- [28] Hinerasky, Christiane, Rene Fahr, and Dirk Sliwka (2010) 'Learning outcomes, Feedback, and the Performance Effects of a Training Program.' *Working Paper, University of Paderborn*
- [29] Holzer, Harry J., Richard N. Block, Marcus Cheatham, and Jack H. Knott (1993) 'Are training subsidies for firms effective? The Michigan experience.' *Industrial and Labor Relations Review* 46(4), 625–636
- [30] LaLonde, Robert J. (1986) 'Evaluating the econometric evaluations of training programs with experimental data.' *The American Economic Review* 76(4), 604–620
- [31] Levitt, Steven D., and John A. List (2009) 'Was there really a Hawthorne Effect at the Hawthorne plant? An analysis of the original illumination experiments.' *NBER Working Paper No. w15016*
- [32] Liu, Xiangmin, and Rosemary Batt (2007) 'Economic pay-offs to informal training: Evidence from routine service work.' *Industrial and Labor Relations Review* 61(1), 75–89
- [33] Loewenstein, Mark A., and James R. Spletzer (1998) 'Dividing the costs and returns to general training.' *Journal of Labor Economics* 16(1), 142–172

- [34] Morrow, Charley C., M. Quintin Jarrett, and Melvin T. Rupinski (1997) 'An investigation of the effect and economic utility of corporate-wide training.' *Personnel Psychology* 50, 91–110
- [35] Pischke, Jörn-Steffen (2001) 'Continuous Training in Germany.' *Journal of Population Economics* 14, 523–548
- [36] Roethlisberger, F. J., and William J. Dickson (1949) *Management and the Worker* (Cambridge: Harvard University Press)
- [37] Wooldridge, Jeffrey M. (2002) *Econometric Analysis of Cross Section and Panel Data* (Cambridge, MA. MIT Press)
- [38] Zwick, Thomas (2005) 'Continuous vocational training forms and establishment productivity in Germany.' *German Economic Review* 6(2), 155–184
- [39] Zwick, Thomas (2006) 'The impact of training intensity on establishment productivity.' *Industrial Relations* 45(1), 26–46

Tables

Table 1: Description of Variables

<i>Individual time-variant Variables</i>	Obs.	Mean	Std. Dev.
Age	13056	43.18	8.41
Tenure	13008	11.06	6.45
Existence	16320	13.26	5.72
	Obs.	Percent	Cum.
Training Participation:	16320		
Participation (<i>everT=1</i>)	7920		
While Training (<i>inT=1</i>)	869		
After Training (<i>afterT=1</i>)	1364		
Franchise Branch:	15936		
Yes (<i>partner=1</i>)	3,304	20.73	20.73
No	12,632	79.27	100

<i>Store – specific time-invariant Variables</i>	Obs.	Mean	Std. Dev.
Sales Revenue (<i>sales</i>)	16320	145.84	40.49
Full Time Equivalents (<i>fullemp</i>)	10356	3.26	0.94
	Obs.	Percent	Cum.
Store Environment:	16320		
Shopping Center (<i>standcenter</i>)	6,384	39.12	39.12
Downtown (<i>standinner</i>)	2,976	18.24	57.35
Stand Alone (<i>standalone</i>)	6,960	42.65	100.00
Store Environment Detail:	16320		
1a-Location (<i>stand1a</i>)	1,200	7.35	7.35
1b-Location (<i>stand1b</i>)	1,776	10.88	18.24
< 10 tsqm floor space & discount (<i>standun10tqmdisc</i>)	2,544	15.59	33.82
< 10 tsqm floor space & market (<i>standun10tqmretail</i>)	768	4.71	38.53
> 10 tsqm floor space & discount (<i>standov10tqmdisc</i>)	1,536	9.41	47.94
> 10 tsqm floor space & market (<i>standov10tqmretail</i>)	1,536	9.41	57.35
Weak commerce (<i>standtypp7</i>)	3,840	23.53	80.88
Strong commerce (<i>standtypp8</i>)	3,120	19.12	100.00

Table 1: Description of Variables (continued)

	Obs.	Percent	Cum.
Sales Revenue Category:	13056		
Small (<i>salesun1m</i>)	9,216	70.59	70.59
Medium (<i>salesbet12m</i>)	2,592	19.85	80.15
Large (<i>salesov2m</i>)	1,248	19.85	100
Region:	16320		
West Germany (<i>westger=1</i>)	10032	61.47	61.47
East Germany	6288	38.53	100
Federal State:	16320		
Baden-Württemberg (<i>bw</i>)	3,024	18.53	18.53
Bayern (<i>ba</i>)	2,496	15.29	33.82
Berlin (<i>ber</i>)	528	3.24	37.06
Brandenburg (<i>bra</i>)	1,008	6.18	43.24
Bremen (<i>bre</i>)	144	0.88	44.12
Hamburg (<i>ha</i>)	96	0.59	44.71
Hessen (<i>he</i>)	576	3.53	48.24
Mecklenburg-Vorpommern (<i>mv</i>)	720	4.41	52.65
Niedersachsen (<i>ns</i>)	720	4.41	57.06
Nordrhein-Westfalen (<i>nw</i>)	1,824	11.18	68.24
Rheinland-Pfalz (<i>rp</i>)	768	4.71	72.94
Saarland (<i>sa</i>)	240	1.47	74.41
Sachsen (<i>sac</i>)	1,872	11.47	85.88
Sachsen-Anhalt (<i>saa</i>)	1,152	7.06	92.94
Schleswig-Holstein (<i>sh</i>)	144	0.88	93.82
Thüringen (<i>th</i>)	1,008	6.18	100.00
Size of Town:	16320		
Provincial Town < 20,000 inhabitants (<i>townsmall</i>)	8,544	52.35	52.35
Middle Town 20,000 - 100,000 inhabitants (<i>townmid</i>)	4,512	27.65	80.00
Major City > 100,000 inhabitants (<i>townbig</i>)	3,264	20.00	100.00
Branch Management:	13056		
Line Manager (<i>ffk=1</i>)	12288	94.12	94.12
Supervising Manager (<i>supffk=1</i>)	48	0.37	94.49
Supervising Employee (<i>empffk=1</i>)	720	5.51	100
Regional Executive Manager:	16320		
Dummy for each Executive Manager (<i>mm1-mm3</i>)			
District:	16320		
Dummy for each of 23 districts (<i>regioncode1-23</i>)			

Table 2: Determinants of Training Participation/ Probit Regression

Exogenous Variables	(1)		(2)	
	Probit Regression on Training Participation		Probit Regression on Early Training Participation	
Sales Revenue	0.1921***	(0.0268)	0.2956***	(0.0456)
Full time equivalents	-0.0048	(0.0055)	0.0046	(0.0101)
Age	0.0017	(0.0050)	-0.0317***	(0.0093)
Age sq	-0.0001*	(0.0001)	0.0003***	(0.0001)
Tenure	0.0170***	(0.0028)	0.0296***	(0.0049)
Tenure sq	-0.0006***	(0.0001)	-0.0013***	(0.0002)
Age Store	-0.0076	(0.0052)	-0.0388***	(0.0094)
Age Store sq	0.0006***	(0.0002)	0.0020***	(0.0004)
Franchise store	0.0346	(0.0402)	0.0301	(0.0756)
South Germany	0.0593	(0.0491)	-0.1801***	(0.0696)
West Germany	0.1421***	(0.0548)	-0.0804	(0.0778)
North Germany	-0.0447	(0.0435)	-0.1970***	(0.0552)
Town 20,000-100,00 inhabitants	0.0226*	(0.0119)	0.0021	(0.0204)
Major City >100,000 inhabitants	0.0084	(0.0139)	0.1055***	(0.0289)
Sales Revenue Category between 1 and 2 million Euro	0.0330***	(0.0125)	0.0654***	(0.0220)
Sales Revenue Category over 2 million Euro	-0.0117	(0.0190)	-0.0141	(0.0335)
1a Store Location	-0.0720***	(0.0178)	-0.2100***	(0.0255)
Location Shopping Center	0.0916***	(0.0153)	0.1192***	(0.0263)
Location Downtown	0.1592***	(0.0230)	0.3503***	(0.0352)
Weak commerce	0.0005	(0.0141)	0.0207	(0.0258)
Manager is supervising manager	0.3203***	(0.0934)	0.0704	(0.0980)
Manager is supervising employee	-0.1659***	(0.0101)	0.2689***	(0.1001)
Regional executive manager 1	-0.1138***	(0.0193)	-0.0888**	(0.0361)
Regional executive manager 2	-0.0832**	(0.0421)	-0.2994***	(0.0636)
Observations	7668		3888	
Pseudo R-squared	0.0503		0.0707	

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Coefficients can be interpreted as marginal effects. References are East Germany, Medium Cities, Sales Revenue Category under 1 million Euro, 1b Store Location, Greenfield Location, Strong commerce, Manager, and Regional executive manager 3.

Table 3: Training Effects on Sales Revenue and Sales Growth

Exogenous Variables	Regression on sales revenue			Regression on sales growth			
	(1) Prais-W. ATE	(2) Prais-W. ATT	(3) FE AR(1) ATT	(4) Prais-W. ATE	(5) Prais-W. SE	(6) Prais-W. ATT	(7) FE AR(1) ATT
In Training	8.3681*** (1.55)	3.3846** (1.33)	0.5420 (1.45)	0.0248*** (0.0079)		0.0102 (0.0085)	-0.0097 (0.0103)
After Training	5.2475*** (1.36)	2.2404 (1.48)	-3.5118** (1.73)	-0.0114 (0.0074)		-0.0164* (0.0094)	-0.0476*** (0.0130)
Ever Training					0.0079 (0.0051)		
Months Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years 05 & 06	0.8059 (1.00)	1.5122 (1.02)	-2.1668 (1.85)	-0.0683 (0.0948)	-0.0980*** (0.0050)	-0.0836*** (0.0080)	^a
Year 2007	12.0206*** (1.21)	12.4949*** (1.68)	7.1113** (3.28)	0.0190 (0.0948)	0.0000 (0.0000)	0.0000 (0.0000)	0.0510*** (0.0123)
Age	0.6100 (0.50)	4.3760*** (0.88)	-0.7529 (1.93)	0.0037 (0.0026)	-0.0007 (0.0024)	0.0108** (0.0053)	0.0143 (0.0198)
Age sq	-0.0049 (0.01)	-0.0551*** (0.01)	0.0080 (0.02)	-0.0000 (0.0000)	-0.0000 (0.0000)	-0.0001** (0.0001)	-0.0001 (0.0002)
Tenure	-0.1269 (0.26)	-0.2449 (0.37)	2.6005** (1.32)	0.0013 (0.0015)	-0.0020 (0.0014)	-0.0015 (0.0024)	-0.0040 (0.0118)
Tenure sq	-0.0135 (0.01)	0.0056 (0.02)	-0.0911*** (0.03)	-0.0001 (0.0001)	0.0001 (0.0001)	0.0000 (0.0001)	-0.0000 (0.0003)
Full time equivalents	2.0625*** (0.51)	2.5249*** (0.63)	2.0497** (0.89)	0.0089*** (0.0031)	0.0094*** (0.0027)	0.0061 (0.0043)	-0.0003 (0.0072)
Age Store	-0.7868 (0.54)	-1.0264 (0.64)	6.1266*** (2.09)	-0.0118*** (0.0032)	-0.0060** (0.0027)	-0.0108*** (0.0040)	0.0282 (0.0176)
Age Store sq	-0.0042 (0.02)	0.0114 (0.03)	-0.1602*** (0.04)	0.0003*** (0.0001)	0.0001 (0.0001)	0.0003** (0.0002)	0.0002 (0.0004)
Franchise store	-3.3818 (2.28)	1.8844 (3.10)	-0.5494 (5.69)	0.0007 (0.0180)	-0.0012 (0.0153)	0.0308 (0.0245)	0.0085 (0.0504)
Constant	0.0000 (0.0000)	0.0000 (0.0000)	43.9065 (30.18)	0.0000 (0.0000)	0.1129 (0.0804)	0.0000 (0.0000)	-0.6799* (0.3797)
Observations/ Groups	6597/214	5160/109	5051/109	5361/214	6087/213	3888/109	3779/109
R-sq (Adj.R-sq)	0.85	0.86	(0.62)	0.17	0.13	0.13	(0.06)

Note: Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Specifications (1) & (2) and (4)-(6) assume that, within panels, there is first-order autocorrelation and that the coefficient of the AR(1) process is specific to each panel. The disturbances are assumed to be panel-level heteroskedastic with no contemporaneous correlation across panels.

In Prais-Winsten regressions we explicitly control for the store environment by including 2 dummies for towns from 20,000-100,000 inhabitants or major cities with >100,000 inhabitants (Reference medium cities); 1 dummy for 1a store location (Reference 1b store location); 3 dummies for <10k sqm&retail, >10k sqm&retail, >10k sqm&discount (Reference <10k sqm&discount); 1 dummy for weak commerce (Reference: strong commerce), 2 dummies for shopping center, downtown (Reference: Greenfield location); 2 dummies for Manager is Supervising manager or Supervising employee (Reference: Manager), 2 regional executive manager dummies, 22 district dummies, 15 federal state dummies, West Germany dummy, 2 dummies for sales revenue category.

^a the coefficient could not be estimated in this specification due to multicollinearity.

Table 4: Mean Final Grades by Cohort

Cohort	Mean Grade	Str. Error	[95% Conf. Intervall]	
1	2.17	0.0284	2.11	2.22
2	2.22	0.0267	2.17	2.27
3	2.60	0.0303	2.54	2.66
4	2.46	0.0536	2.36	2.57
5	2.47	0.0249	2.42	2.52
6	2.69	0.0210	2.65	2.73
7	2.44	0.0179	2.40	2.47
8	2.64	0.0336	2.57	2.70
9	2.39	0.0213	2.35	2.43
10	2.81	0.0318	2.75	2.87

Note: Grades are given on a standardized basis: 100-92 Points Grade 1(A); <92-81 Points Grade 2(B); <81-67 Points Grade 3(C); <67-50 Points Grade 4(C); <50-30 Points Grade 5(E); <30-0 Points Grade 6(F).

Table 5: In Training and After Training Effect of Early and Late Participation

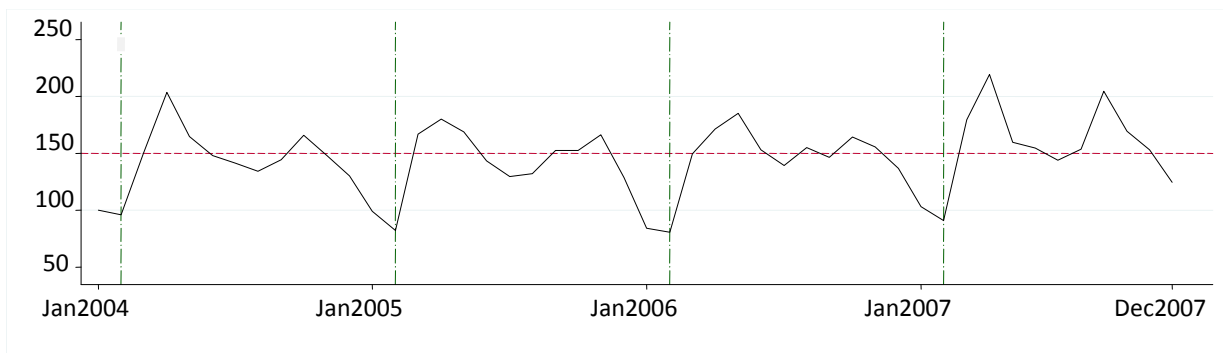
Exogenous Variables	Regression on sales revenue			Regression on sales growth			
	(1) Prais-W. ATE	(2) Prais-W. ATT	(3) FE AR(1) ATT	(4) Prais-W. ATE	(5) Prais-W. SE	(6) Prais-W. ATT	(7) FE AR(1) ATT
Early In Training	6.3824*** (1.8991)	4.7485*** (1.5195)	0.6098 (1.6972)	0.0390*** (0.0102)		0.0469*** (0.0097)	0.0449*** (0.0121)
Late In Training	8.2678*** (2.2205)	0.9298 (1.8718)	1.1202 (1.9135)	0.0094 (0.0100)		-0.0214* (0.0118)	-0.0415*** (0.0134)
Early After Training	5.7617*** (1.4399)	3.7653*** (1.4244)	-2.6364 (1.6781)	-0.0046 (0.0077)		-0.0027 (0.0090)	-0.0085 (0.0124)
Late After Training	0.3036 (2.4247)	-4.1119** (2.0691)	-4.1594* (2.2139)	-0.0345*** (0.0115)		-0.0589*** (0.0125)	-0.0908*** (0.0153)
Ever Training					0.0079 (0.0051)		
Months Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes ^a
Years 2005 & 2006	0.7939 (1.0041)	1.2569 (1.0180)	-2.0217 (1.8490)	-0.0897*** (0.0051)	-0.0980*** (0.0050)	-0.3723*** (0.1254)	
Year 2007	11.9799*** (1.2130)	13.0057*** (1.6889)	7.4174** (3.3497)	0.0000 (0.0000)	0.0000 (0.0000)	-0.2781** (0.1264)	0.0726*** (0.0129)
Age	0.5217 (0.5062)	4.4711*** (0.8872)	-0.7952 (1.9284)	0.0038 (0.0026)	-0.0007 (0.0024)	0.0130** (0.0053)	0.0127 (0.0197)
Age sq	-0.0039 (0.0057)	-0.0561*** (0.0107)	0.0087 (0.0224)	-0.0000 (0.0000)	-0.0000 (0.0000)	-0.0002** (0.0001)	-0.0001 (0.0002)
Tenure	-0.0993 (0.2630)	-0.3563 (0.3668)	2.6143** (1.3248)	0.0011 (0.0015)	-0.0020 (0.0014)	-0.0030 (0.0024)	-0.0088 (0.0117)
Tenure sq	-0.0150 (0.0121)	0.0119 (0.0175)	-0.0910*** (0.0313)	-0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0003)
Full time equivalents	1.9783*** (0.5139)	2.5223*** (0.6317)	2.1310** (0.8891)	0.0089*** (0.0030)	0.0094*** (0.0027)	0.0076* (0.0042)	0.0032 (0.0071)
Age Store	-0.6935 (0.5569)	-0.9188 (0.6335)	5.9971*** (2.0915)	-0.0108*** (0.0032)	-0.0060** (0.0027)	-0.0074* (0.0040)	0.0201 (0.0174)
Age Store sq	-0.0069 (0.0209)	0.0078 (0.0262)	-0.1629*** (0.0374)	0.0003** (0.0001)	0.0001 (0.0001)	0.0002 (0.0002)	0.0002 (0.0004)
Franchise Store	-3.4266 (2.2979)	2.0117 (3.1813)	-0.4910 (5.6944)	0.0006 (0.0182)	-0.0012 (0.0153)	0.0381 (0.0249)	0.0189 (0.0498)
Constant	0.0000 (0.0000)	-14.4577 (20.1411)	45.7004 (30.1676)	0.0504 (0.0896)	0.0480 (0.0661)	0.0000 (0.0000)	-0.5402 (0.3820)
Observations/Groups	6597/214	5160/109	5051/109	5361/214	6087/213	3888/109	3779/109
R-sq (Adj. R-sq)	0.85	0.86	(0.62)	0.17	0.13	0.14	(0.07)

Note: Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Specifications (1) and (2) & (4) – (6) assume that, within panels, there is first-order autocorrelation and that the coefficient of the AR(1) process is specific to each panel. The disturbances are assumed to be panel-level heteroskedastic with no contemporaneous correlation across panels. In Prais-Winsten regressions we explicitly control for the store environment by including 2 dummies for towns from 20,000-100,000 inhabitants or major cities with >100,000 inhabitants (Reference medium cities); 1 dummy for 1a store location (Reference 1b store location); 3 dummies for <10k sqm&retail, >10k sqm&retail, >10k sqm&discount (Reference <10k sqm&discount); 1 dummy for weak commerce (Reference: strong commerce), 2 dummies for shopping center, downtown (Reference: Greenfield location); 2 dummies for Manager is Supervising manager, Supervising employee (Reference: Manager), 2 regional executive manager dummies, 22 district dummies, 15 federal state dummies, West Germany dummy, 2 dummies for sales revenue category.

^a the coefficient could not be estimated in this specification due to multicollinearity.

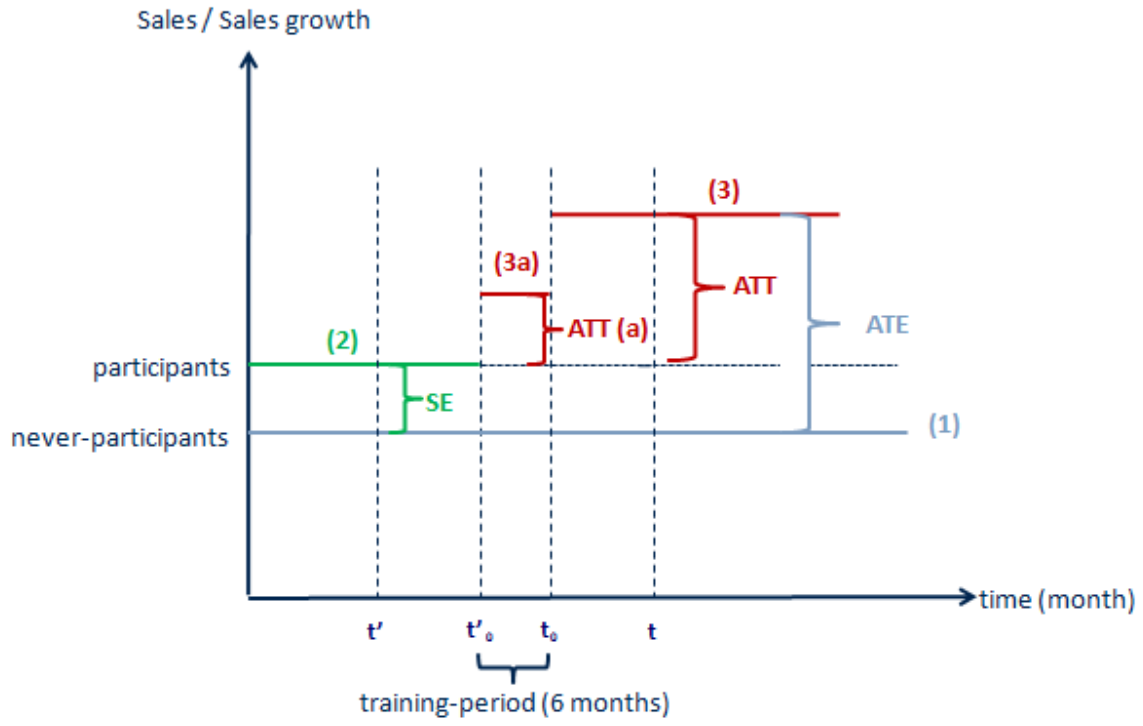
Figures

Figure 1: Normalized Company Sales Revenue for 2004 – 2008



Note: Normalized non-deflated sales revenue is calculated as the average of all shop sales revenues. Sales revenue reaches its low each February as labelled by vertical lines.

Figure 2: Empirical Strategy



Note: ATE: Training Participants afterT vs. Never-Participants
 SE: Training Participants vs. Never-Participants;
 ATT (a): Training Participants inT ($t'_0 - t_0$) vs. Training Participants before (t');
 ATT: Training Participants afterT (t) vs. Training Participants before (t');