Online Appendix

A. A review of studies in servitization that have considered the implications of employee training

[TABLE A ABOUT HERE]

B. Determinants of training intensity

Table B presents the results of different regression models with the number of training courses undertaken by the service units being the dependent variable.

[TABLE B ABOUT HERE]

A Poisson regression model (Models 1 and 2) is an appropriate estimation strategy here, as the number of courses is a count variable. However, we also present the estimation results from OLS Fixed Effects models (3 and 4), which treat number of courses as a continuous variable, since this model is used in a robustness check that uses an Instrumental Variables (IV) estimation strategy (see section 4.3 of the paper and D of this Appendix). This IV method employs lagged № of service hours as an instrument for training, which is mainly why this variable enters the training regressions, but not the performance equation. Model (5) presents the first-stage of our Two-Stage Least Squares (2SLS) estimation, which is done only for outsourced units (more details are discussed there).

From Models (1) and (3), which include the lagged performance dummies as explanatory variables, we observe that units that achieve high performance (a score of 3 or 4) in a period, tend to undertake more training in the following period. This suggests the presence of reverse causality, which needs to be dealt with in the models that explain service performance. In addition, we note that several other variables affect training (more notably ownership, distance and indicators of unit size). Thus, it is important that these variables are controlled for in the regression framework for service performance, as they are also potential determinants of performance. Models (2) and (4) show that service units that received the monetary bonus in the previous quarter are more likely to undertake more training in the current period (if past performance is not controlled for). Thus, it is important that this variable is included in any non-dynamic model explaining service performance. Such a non-dynamic model is run as a robustness check in our IV approach, which uses lagged № of service hours claimed by units as an instrument.

C. Analysis and results after splitting the sample by ownership status

In the first robustness check we re-run the analysis but separately for the two groups of service units. Splitting the sample into two allows for the possibility that company-owned and outsourced units are qualitatively different in nature. For example, as we have mentioned in the theoretical background, company-owned units are virtually the same firm as TruckBus, hence need to adhere to certain policies, while decision-making for strategic or tactical issues may lie with the HQs. Conversely, outsourced units are, generally small independent businesses over which the manufacturer has considerably less influence.

The results are presented in Table C. It should be noted that due to the smaller sample sizes, particularly for company-owned units, the estimated coefficients are less trustworthy than our base-line models, since the *consistency* of the estimators used here is based on the assumption of an infinite sample size. Of course, obtaining an infinite sample is impossible, but the corresponding asymptotic theory suggests that a 'large enough' sample is required to eliminate small-sample biases. A sample of more than 1000 observations can safely be considered large enough, but it is not straightforward to argue that around 300 observations are enough in these relatively complex econometric frameworks. For company-owned units, the clustered standard errors also need to be treated cautiously, as 20 clusters (i.e. the number of units) may be too 'few' to allow valid inference (Cameron and Miller 2015).

[TABLE B ABOUT HERE]

The main insights derived from the first approach are equally pronounced here: For company-owned units, there is only a statistically weak effect of training on service performance. This effect appears to be linear, since if the quadratic terms are removed, the coefficients of training become significant at the 5% level of significance. For outsourced units, the results are in total agreement with those of the base-line approach: That there is a strong, inverted U-shaped relationship between training and performance. The turning points are also in agreement with our analysis in section 5.2.

D. Additional robustness checks

Table D1 presents some additional robustness checks. Note that all results of Models (4) through (8) are obtained from the sample of outsourced units only.

[TABLE D1 ABOUT HERE]

Model (1) is the base-line dynamic CRE-OP specification, but with time dummies (instead of a time trend) to account for the effect of time-specific events. Crucially, with this specification we explicitly control for the progressively upward adjustment of the thresholds of the 1st KPI (MOT 1st time pass rate) by TruckBus. This happened three times during our timeframe, and it may have induced higher levels of effort by the service units, as well as prompted them to take on more training to be able to cope with the increased TruckBus quality standards. The resulting coefficients of training, for both outsourced and company-owned units, are virtually identical with those of the base-line model, and the main qualitative insights do not change, which indicates that the way we capture the effect of time does not make any material difference.

Models (2) and (3) operationalize performance in a way that takes into account the increased importance of the 1st KPI to TruckBus. Specifically, in addition to the MOT pass/fail threshold, TruckBus applies another, much lower threshold, which is basically the national average of the MOT 1st time pass rate. Throughout the entire timeframe, if a service unit failed to reach that lower threshold, TruckBus considered its overall service performance to be zero, irrespective of how many other KPIs it may have achieved. As such, this is the context-specific performance measure that TruckBus and the service units live by. Evidently, the main qualitative difference to our preferred performance measure is that the zero level of performance occurs more often, especially in the earlier periods (when the network, as a whole, was performing worse – see Table D2). We believe that an analysis that uses this as an outcome measure, which in effect penalizes service units that obtained another two or three KPIs but not the 1st one, 'misses' some of the effect of training on the other KPIs and does not reflect the true service performance. However, as this is closer to the real-world context, it may affect the decisions of the service units in terms of training, and therefore subsequence performance. Thus, it is important to check if using this performance variable qualitatively changes our results. As expected, the results are

slightly less pronounced compared to the base-line analysis, with estimates of the training terms being a bit smaller and slightly less significant.

In Models (4) and (5) we provide results of a Fixed Effects Instrumental Variables (IV) estimator, which is an alternative strategy to deal with potential endogeneity of training due to reverse causality or other omitted time-varying factors (Baltagi 2013). Specifically, we use a standard Two-Stage Least Squares model (2SLS), that also takes into account unit-specific unobserved heterogeneity. As this is a linear estimator, we need to assume that the values assigned to the performance variable have a cardinal interpretation, which is, however, a standard assumption in the economics literature (see Ferrer-i-Carbonell and Frijters 2004). We use the 'period $t-1$ № of service hours sold to TruckBus' and its squared term as instruments for training at period t. The idea is that units that experience considerable amounts of TruckBus related workloads are more likely to undertake more training in the next period, to be able to deal with the increased needs. Indeed, this is found to be a *relevant* instrument for both 'number of courses' and 'training days', as it has a strong impact on training (see Model (5) in Table B, and F-stats in Table D1). We also assume that this is an *exogenous* instrument; i.e., controlling for unit-specific heterogeneity (via the Fixed-Effects approach and for other control variables), it only affects performance through its impact on training. Although *exogeneity* is not a testable assumption, as an informal test we find that, in our base-line dynamic specification (where we believe that training is not endogenous), the instrument has a totally insignificant effect on performance (p-value>0.9). The 2SLS results indicate that training has a positive and significant effect on performance, strengthening our conclusions from the base-line models. Note that in these models, we only use the linear term of training, as we were unable to instrument for both training and its square term.

Finally, in Models (6) and (7) we control for the reverse causality problem by using the Arellano-Bover/Blundell-Bond (AB-BB) linear estimator for dynamic models in panel data (Arellano and Bover 1995, Blundell and Bond 1998), where we again need to assume cardinality of performance. In these models, reverse causality of performance on training is controlled for by including the lag of performance as a covariate. However, inserting lagged values of the dependent variable generates another endogeneity problem, as the lagged values are now correlated with the 'unobserved heterogeneity' error term. The aforementioned estimator is one of the standard econometric approaches

to control for the endogeneity of the lagged performance, by using a system-Generalised Methods of Moments estimator, where a series of lagged differences in performance and performance before $t-1$ are instruments for performance at $t-1$. The results of this method suggest that, again, number of courses have a significant inverted U-shaped relationship with performance (the same for training days, although the evidence is weaker). However, these results should be treated cautiously, as there is mild evidence of second order serial correlation in the differenced errors, which implies some form of misspecification in the model. Finally, note that a simpler dynamic Fixed Effects estimator that does not account for endogeneity of lagged performance, provides estimates that are qualitatively similar to the AB-BB estimator in terms of the effect of training, suggesting that the endogeneity of lagged performance causes little bias on the estimated effect of training.

[TABLE D2 ABOUT HERE]

E. Further insights from matching estimators

We conduct further robustness checks by using a 'matching' estimation approach, following the work by Abadie and Imbens (2006, 2011) and Abadie at al. (2004).[1] According to this matching approach, every company-owned service unit is matched with one outsourced service unit based on their similarity in terms of their training profiles. For each matched-pair we calculate the difference in their performance and then calculate the average difference in performance across all matched-pairs. Based on our hypotheses and the results of this paper, we expect that on average, company-owned service units underperform relative to similar, in terms of training profiles, outsourced ones.

To perform the matching, we use the Mahalanobis distance metric (see Mahalanobis 1936), which is widely used in empirical applications and preferred to other scores, such as the Euclidean distance, as it also accounts for the covariance among covariates (see Rubin 1973, Abadie and Imbens 2006). As we are interested in the similarity of training profiles, our Mahalanobis distance score is computed based on the number of: 'technical courses'; 'non-technical courses'; 'first-time courses'; and 'repeated courses'. To also guarantee that each company-owned observation is matched with an

[1] The latter discusses implementation of the theoretical results by Abadie and Imbens (2006) in Stata.

outsourced observation in the same, or at least tangential, time period, given that we have a panel data structure, the matching process is also based on our time indicator.[2] Once the distance score is created, each company-owned service unit is matched with the outsourced service unit with the lowest distance. A matching 'with replacement' is performed, as this allows for better quality matches. Then, the average difference in performance is calculated using the matched pairs. An issue discussed in Abadie and Imbens (2006, 2011) is that this estimator is inconsistent when the matching process is based on more than one continuous covariates. They suggest a "large-sample bias" reduction approach, which uses linear functions of the covariates. In our case, as all covariates used are treated as continuous variables, we apply this bias reduction approach. Finally, to allow for valid inference, we obtain 'robust' standard errors, as described in Abadie and Imbens (2006).

The results are presented in Table E. Model (1) suggests that the average performance of company-owned service units is 0.15 points lower than the average performance of similar, in terms of training profiles, outsourced service units, a difference that is statistically significant at the 5% level. This result supports and strengthens the conclusions drawn from our baseline models. In model (4), to avoid matching very different service units in terms of other observed characteristics, we include the additional variable "hours claimed" which is a proxy for service units' size and workload. By doing so, the estimated impact increases in magnitude and is still significant at the 5% level.

As in any matching process, there may be some matches that have been of low quality. In our case, for model (1), 38.8% of the matches have a 0 distance, indicating a perfect match, while 98.6% have a Mahalanobis distance smaller than 2. For model (4), although there are no exact matches (this is because 'hours claimed' is a truly continuous variable), 93.6% have a distance smaller than 2. These indicate that quality of matching is overall satisfactory, but slightly better for model 1. To check if 'less good' matches affect our estimates, in models (2) and (4) we remove all company-owned observations with Mahalanobis distance higher than 2, and in models (3) and (6) all company-owned observations with distances higher than 3. For models (2) and (3), we see that the estimated differences remain similar

[2] We find that including this index works almost perfectly well, with most matches happening exactly in the same time period.

in value and significant at 5%. For models (5) and (6), the estimated differences become even stronger, both in magnitude and in statistical significance.

[TABLE E ABOUT HERE]

F. Further support for H2a: Measuring changes in training profiles after switching ownership

We have hypothesized (Hypothesis 2a) that one reason for the differential effect of training on performance between outsourced and company-owned units is that outsourced service units will choose types of manufacturer-led formal training that are more effective. We provided some evidence of this in section 5.3, where we showed that outsourced units take significantly more technical courses (in proportion) than company-owned units, while we also showed that technical training courses have a larger impact on the performance measure than non-technical courses.

As we mentioned in section 4.2.3, in our sample there are 4 service units switching from outsourced to company-owned and 5 switching from company-owned to outsourced at some point within the study timeframe. To provide further evidence on Hypothesis 2a, here we look at whether the training profile changed for these units after switching ownership status. For this purpose, we calculate the mean 'number of training courses' for these nine units before and after switching ownership. Recognising that training increases over time, for each time period we 'normalize' these means by dividing by the overall mean training in that period. In addition, recognising that the calculation of the overall mean includes units that are very different from the ones that switched, we also calculate another normalised mean for each period, this time only using units that are 'similar' to the ones that switched ownership status in terms of their observed characteristics.

The results are provided in Table F, separately for overall training, for technical training and for non-technical training. We see that both groups increased their overall training, but the difference is relatively higher for those who switched from outsourced to company-owned. This is to some extent as expected, as company-owned units are expected to follow a more 'uniform' training profile dictated by the parent company, while outsourced ones are allowed to choose what training courses to take. Nevertheless, the striking results come when we split training in technical and non-technical courses.

We see that those who switched to outsourced changed their training profile substantially, significantly increasing their technical training and decreasing their non-technical training. On the other hand, those who switched to company-owned, experienced only a small increase in technical training, but a substantial increase in their non-technical training. Of course, as we deal with a sample of 9 service units only, these results need to be treated cautiously, but they do provide further evidence in line with the intuition behind the hypothesis.

[TABLE F ABOUT HERE]

References

Abadie, A., D. Drukker, J.L. Herr, G.W. Imbens. 2004. Implementing matching estimators for average treatment effects in Stata. The stata journal, 4(3), 290-311.

Abadie, A., G.W. Imbens. 2006. Large sample properties of matching estimators for average treatment effects. Econometrica, 74, 235-267.

Abadie, A., G.W. Imbens. 2011. Bias-corrected matching estimators for average treatment effects. Journal of Business and Economic Statistics, 29, 1-11.

Arellano, M., O. Bover. 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics, 68(1), 29–51.

Blundell, R., S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1), 115–143.

Cameron, A.C., D.L. Miller. 2015. A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317-372.

Ferrer-i-Carbonell, A., P. Frijters. 2004. How important is methodology for the estimates of the determinants of happiness?. The Economic Journal, 114(497), 641–659.

Forkmann, S., C. Ramos, S.C. Henneberg, P. Naudé. 2017. Understanding the service infusion process as a business model reconfiguration. Industrial Marketing Management, 60, 151–166.

Johnstone, S., A. Dainty, A. Wilkinson. 2009. Integrating products and services through life: an aerospace experience. International Journal of Operations & Production Management, 29 (5), 520–538.

Kreye, M.E. 2017. Can you put too much on your plate? Uncertainty exposure in servitized triads. International Journal of Operations and Production Management, 37 (12), 1722–1740.

Johnstone, S., A. Wilkinson, A. Dainty. 2014. Reconceptualizing the service paradox in engineering companies: Is HR a missing link? IEEE Transactions on Engineering Management, 61 (2), 275–284.

Mahalanobis, P.C. 1936. On the generalized distance in statistics. National Institute of Science of India.

Paiola, M., H. Gebauer, B. Edvardsson. 2012. Service Business Development in Small- to Medium-

Sized Equipment Manufacturers. Journal of Business-to-Business Marketing, 19 (1), 33–66.

Raja, J.Z., T. Frandsen. 2017. Exploring servitization in China: Challenges of aligning motivation, opportunity and ability in coordinating an external service partner network. International Journal of Operations & Production Management, 37 (11), 1654–1682.

Rubin, D.B. 1973. Matching to remove bias in observational studies. Biometrics, 29, 159–183.

Ulaga, W., J.M. Loveland. 2014. Transitioning from product to service-led growth in manufacturing firms: Emergent challenges in selecting and managing the industrial sales force. Industrial Marketing Management, 43 (1), 113–125.

Tables and Figures

Table A – Implications of training in servitization: Empirical evidence from relevant studies

Source	**Empirical Context; Research Design**	**Relevant findings**
Forkmann et al. (2017)	European manufacturer and its network; single in-depth case-study	The focal firm invested extensive resources into developing training (sales, installation, service) for its distributors. As the product-service (PS) offering was continuously under development, training needed to be constantly updated: the focal firm had to synchronize distributor training schedules according to the newly introduced features/upgrades.
Johnstone et al. (2009)	International aerospace OEM; single in-depth case-study	In order to change behaviors, and increase the manufacturer's customer responsiveness, training and skill development of existing employees were required. Managers *"recognised that 'service' (as a functional area) needed a strong development pathway for employees"* (p.531).
Johnstone et al. (2014)	Two global engineering organizations; multiple case-studies	Training and development were required to change the culture of employees. But they emphasized product innovation and new product development, while the skills (both technical and service) required for PS (hybrid) provision were different. Hence, there were concerns that development opportunities were not sufficiently reconfigured to support the evolving strategy.
Kreye (2017)	One European-North African triad; single in-depth case-study	To reduce the uncertainty arising from the local availability of qualified staff, the European buyer and Service provider initiated a cooperation with an engineering university and created an institute to train engineers in the last year of their bachelor degree.
Paiola et al. (2012)	Nine European SMEs; multiple case-studies	Continuous training (on-the-job in installation, repair, and maintenance of equipment, as well as coaching and sales seminars) increases customer-focus and helps the employees of the SMEs assume responsibility and provide flexibility for the broad range of issues that arise in customized service provision. SMEs delivering offerings through distributors stressed the importance of ongoing technical training (and tailored educational programs) for the distributors' staff.
Raja and Frandsen (2017)	European OEM and its Chinese network of services partners; single in-depth case-study	External service partners lacked the necessary skills to provide advanced services, since technicians had not received sufficient formal training (training manuals were not enough), tailored to the needs of high-end customers and the intricacies of the hybrid offerings. Partners felt that the training by the HQs should have been more extensive, and that their ability to deliver quality work was impaired as a result.
Ulaga and Loveland (2014)	Multiple firms from multiple industries and countries; focus groups and interviews	Managers explained how their firms had improved the efficiency and effectiveness of their sales processes over time through training programs. But despite extensive training efforts, salespeople often could not handle selling advanced hybrid offerings, creating a bottleneck for successful service transition. Retraining of 'traditional' salespeople did not guarantee results, as exemplified by an interviewee: *"You are going to have problems with the guy who comes out of a training session and says, 'Yes, that's fine but I knew all that all ready.' It's a rule, and I've seen it with more than 40 (terminated) salespeople."* (p.120).

Table B – The effects of explanatory variables on number of courses

	Poisson		OLS Fixed Effects		Fixed Effects Outsourced
Variables	(1)	(2)	(3)	(4)	(5)
Ownership dummy (=1 if outsourced)	-.3503***	-.3489***	-1.1254	-1.1829*	
	(0.0755)	(0.0749)	(0.6840)	(0.7069)	
Performance at $t-1$ (base dummy is Performance=0)					
Performance = 1	.1776		.3717		
	(0.2097)		(0.5404)		
Performance = 2	.3034		.6649		
	(0.2159)		(0.5069)		
Performance = 3	.4256*		1.0878*		
	(0.2277)		(0.5631)		
Performance = 4	.5308**		1.7715***		
	(0.2447)		(0.6478)		
Lagged bonus achieved	.0538	.2449***	-.1233	.5899**	.2923
	(0.0949)	(0.0585)	(0.2700)	(0.2245)	(0.184)
№ of incidents	-.0009	-.0009	-.0031	-.0030	.0027
	(0.0006)	(0.0006)	(0.0062)	(0.0063)	(0.0064)
Lagged № of service hours	.000063	.000076	.00050	.00058*	0.00086**
	(0.000058)	(0.000056)	(0.00033)	(0.00031)	(0.00035)
Lagged № of service hours squared	-8.90e-09**	-9.62e-09**	-5.34e-08**	-6.14e-08***	-6.33e-08***
	(4.47e-09)	(4.34e-09)	(2.03e-08)	(1.98e-08)	(1.60e-08)
Distance from HQs	-.0011***	-.0011***			
	(0.0002)	(0.0002)			
Total № of employees trained	.0455***	.0454***			
	(0.0040)	(0.0039)			
Log of county density	-.0273	-.0264			
	(0.0240)	(0.0244)			
Dummy for switching	.2148***	.2134***			
	(0.0754)	(0.0767)			
Time dummies	✓	✓	✓	✓	✓
Constant	.1990	.4201**	2.2992**	2.7391***	.8310
	(0.2497)	(0.1833)	(0.9813)	(0.9923)	(0.7208)
№ of observations	1,118	1,118	1,118	1,118	804
№ of clusters	64	64	64	64	53
Log pseudolikelihood	-2,553.16	-2,561.93			
R squared			0.1615	0.1490	0.2102

Robust (clustered by unit) standard errors to account for serial correlation in the idiosyncratic error term within units (and heteroskedasticity for the Fixed Effects model) are presented in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C – Separate results for company-owned and outsourced service units

	(1)		(2)		(3)	
	Training stock (static)		№ of courses (dynamic)		№ of training days (dynamic)	
Variables	C-o.	Out.	C-o.	Out.	C-o.	Out.
Training	.0469*	.0823***	.0301	.0827***	.0187	.0386***
	(0.0273)	(0.0167)	(0.0292)	(0.0241)	(0.0162)	(0.0094)
Training squared	-.00043	-.00083***	-.00011	-.0029**	-.000083	-.00049**
	(.00040)	(0.00018)	(0.00076)	(0.0014)	(.000026)	(0.00025)
№ of observations	314	804	314	802	314	802
№ of clusters	20	53	20	53	20	53
Log pseudolikelihood	-324.71	-844.10	-320.64	-829.14	-320.28	-827.56
Wald stat for joint effect of training	4.20	25.31***	9.73***	13.62***	6.94**	19.40***
Wald stat. (for $\boldsymbol{\delta}_1 = \mathbf{0}, \boldsymbol{\delta}_2 = \mathbf{0}$)	25.73***	82.95***	101.51***	117.15***	81.84***	120.79***

Robust (clustered by unit) standard errors to account for serial correlation in the idiosyncratic error term within units and heteroskedasticity are presented in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Notes: All models include a constant, all explanatory variables and the unit-specific means of the time-varying variables. The dynamic specifications also include performance in period $t-1$ and in period 1 (initial period) as explanatory variables.

Table D1 – More robustness checks

	Sum of KPIs	Sum of KPIs conditional on MOT		Fixed Effects 2SLS		Arellano-Bover / Blundell-Bond		Dynamic Fixed Effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Time dummies	Time trend	Time dummies	№ of courses	№ of training days	№ of courses	№ of training days	№ of courses
C-o. training	.0339	.0193	.0279					
	(0.0306)	(0.0220)	(0.0246)					
C-o. training squared	-.00018	.00016	.000003					
	(0.00080)	(0.00065)	(0.00069)					
Out. training	.0859***	.0767***	.0810***	0.1946*	.0655***	.0631**	.0236**	.0625***
	(0.0252)	(0.0255)	(0.0253)	(0.1116)	(0.0243)	(0.0259)	(0.0108)	(0.0179)
Out. training squared	-.0028*	-.0027*	-.0026*			-.0035**	-.00049	-.0025***
	(0.0015)	(0.0014)	(0.0014)			(0.0016)	(0.00033)	(0.00087)
Lagged Performance	✓	✓	✓			.4452***	.4469***	.1655***
						(0.0665)	(0.0657)	(0.0553)
Time dummies	✓		✓	✓	✓			
Time trend and its square		✓				✓	✓	✓
№ of observations	1,116	1,116	1,116	802	802	802	802	802
№ of groups	64	64	64	53	53	53	53	53
Log pseudolikelihood	-1,148.56	-1,219.51	-1,205.03					
Wald stat c-o.	9.50***	7.71**	9.04**					
Wald stat out.	15.01***	9.95***	13.36***			5.95*	5.31*	6.19***
First-stage F-stat				27.32***	64.99***			
R squared				0.1629	0.1727			0.2681
z-stat of AB test						1.6753*	1.6993*	

*** p<0.01, ** p<0.05, * p<0.1

Notes: 'Wald stat c-o.' and 'Wald stat out.' are the Wald stats for the null hypothesis that training and training squared have no joint effect on performance for company-owned and outsourced companies respectively.

Models 1, 2, and 3 include a constant, all explanatory variables including past performance dummies, unit-specific means of the time-varying variables, initial conditions and interaction terms between the training variables and the ownership dummy. Robust (clustered by unit) standard errors to account for serial correlation in the idiosyncratic error term within units are presented in parentheses.

Models 4 and 5 include a constant and the time-varying explanatory variables. 'Lagged № of service hours' and its square term are used as instruments. The standard errors are robust to within-unit serial correlation and heteroskedasticity. For both 4 and 5, the instrument pass the *relevance* test, as the F-stats are much higher than 10.

Models 6 and 7 include all explanatory variables and the lagged performance, but not a constant, as the model is in differences. The SEs are obtained using the Arellano-Bond robust VCE estimator. The Arellano-Bond (AB) test is testing for 2nd order zero autocorrelation in 1st differenced errors. Rejection implies some form of misspecification.

Model 8 is a standard FE estimator, which includes the same variables as in 6 and 7, but does not account for endogeneity of lagged performance.

Table D2 – Cross tab of Conditional KPI Sum versus KPI Sum

	Sum of KPIs					
Cond. Sum of KPIs	***0***	***1***	***2***	***3***	***4***	Total
0	22	23	23	12	0	80
1	0	81	0	0	0	81
2	0	0	177	0	0	177
3	0	0	0	363	0	363
4	0	0	0	0	481	481
Total	22	104	200	375	481	1,182

Table E – Matching Estimators

	(1)	(2)	(3)	(4)	(5)	(6)
Difference on Average Performance	-.1492**	-.1560**	-.1464**	-.2104**	-.2335***	-.2403***
	(0.0734)	(0.0712)	(0.0733)	(0.0839)	(0.0815)	(0.0846)
N for C-o.	329	323	315	329	323	308
N for Out.	210	209	206	221	219	214

Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05, * p<0.1
"N for C-o" is the number of company-owned observations used in the matching process
"N for Out" is the number of outsourced observations used in the matching process
Covariates include a time index and number of: 'technical courses'; 'non-technical courses'; 'first-time courses'; and 'repeated courses'.
Models (4)-(6), add the proxy for size, "hours claimed"

Table F – Switch of ownership: Before and after comparison of training profiles

'Normalized' means of 'no of courses'		From C-o. to Out. (N=4)			From Out. To C-o. (N=5)		
		Before	After	% Change	Before	After	% Change
All Training	Using all workshops	0.671	0.888	32.4%	1.265	1.873	48.1%
	Using only similar workshops	0.601	0.636	5.8%	1.257	1.467	16.7%
Technical Training	Using all workshops	0.640	0.970	51.5%	1.515	1.929	27.3%
	Using only similar workshops	0.559	0.716	28.0%	1.470	1.506	2.4%
Non-technical Training	Using all workshops	0.746	0.276	-63.0%	0.692	1.661	140.1%
	Using only similar workshops	0.709	0.164	-76.9%	0.724	1.317	82.0%